SOLENT UNIVERSITY

FACULTY OF BUSINESS, LAW AND DIGITAL TECHNOLOGIES

MSc Applied AI and Data Science

Academic Year 2021-2022

D. Tineghe

Use of computer vision and predictive analysis for council's "pay-and-display" parking

Supervisor: Dr Drishty Sobnath September 2022

This report is submitted in partial fulfilment of the requirements of Solent University for the degree of MSc Applied AI and Data Science.

Acknowledgements

I would like to thank and express my recognition to my supervisor for the guidance, feedback and support offered.

Also, I would like to express my appreciation to my family and loved ones for their endless support.

And lastly, I would like to mention the participants of the study and to thank them for their time and thoughts given.

Abstract

Parking and searching for parking are becoming an increased concern. Understanding parking behaviours and patterns by local authorities could help in providing better parking management that would tackle those issues. This paper aimed to present an artefact and further suggestions that could be used to optimise or improve the council's pay-and-display parking with the use of computer vision and predictive analysis, as well as enhancing other parking space types. Parking video footage and multiple datasets have been used and created, and the artefact has been subject to usability testing. The results suggest that the parking locations are not used to their full capacity, the weather and time/day information influences the behaviours, but no correlation was found for the school holidays presence. Discrepancies in the number of *valid parking tickets* and *CCTV/video count* were found. Popularity and non-popularity times, days and charging points were discovered for electric vehicles, as well as time differences between charging and just being plugged in. Moreover, the usability testing showed that the proposed artefact's best features are *CCTV/video occupancy check or the live count, the up to 2-hour vacancy predictions, the map view* and *the individual page locations*. The obtained results suggest that the proposed artefact could serve the local authorities as a tool that could be used to manage their pay-and-display parking and to use it for future policies, funding, workforce distribution, revenue improvements and to tackle the parking search issue.

Table of Contents

List of Figures

List of Tables

Acronyms

CCTV = Closed-Circuit Television

CNN = Convolutional Neural Networks

DSDM = Dynamic System Development Method

EV = Electric Vehicle

GBP = Great British Pound

ID = Identity Document

IEEE = Institute of Electrical and Electronics Engineers

LSTM = Long-Short Term Memory

MAE = Mean Absolute Error

MoSCoW = Must Have, Should Have, Could Have and Won't Have

MSE = Mean Square Error

N/A = Not Applicable

RMSE = Root Mean Squared Error

SVM = Support Vector Machine

UEQ = User Experience Questionnaire

UK = United Kingdom

vs = versus

1.Introduction & Background

In the United Kingdom, in 2020, there were ~32.7 million passenger licensed cars in operation (Carlier, 2022). This high volume results in the necessity of better utilising the available parking areas. The highest increase in land demand is seen in urban areas (Padmasiri, et al., 2020), (Dogru, et al., 2017) with London only having 16% of its central business district designated for parking areas, compared with 31% which is the big cities' average (Lin, et al., 2017).

Because of this, one of the most repetitive problems an urban-located citizen is having is finding a free parking space (Enríquez, et al., 2017) which results in daily traffic congestion. Different studies suggest different findings regarding the percentage of traffic flow generated from this and the amount of time spent cruising for parking: 40% with an average of 12 minutes (Pflügler, et al., 2016); 35% (Bulan, et al., 2013); 30% with an average of 7.8 minutes (Acharya, et al., 2018), (Valipour, et al., 2016); 30% with an average of more than 20 minutes (Mangiaracina, et al., 2017); 8% to 74% with an average of 3.5 to 14 minutes (Shoup, 2006). Besides the congestion caused by cruising for parking, other factors include increased CO_2 emissions, air pollution, safety issues, driver frustration (Simhon, et al., 2017), (Mangiaracina, et al., 2017), increased vehicle mileage (Shoup, 2006) and fuel consumption (Bibi, et al., 2017). Therefore, good parking management should aim at reducing those factors and help the city to expand and increase its revenue (Lin, et al., 2017) as low parking occupancy is the product of inefficient parking space utilisation (Jakob & Menendez, 2020).

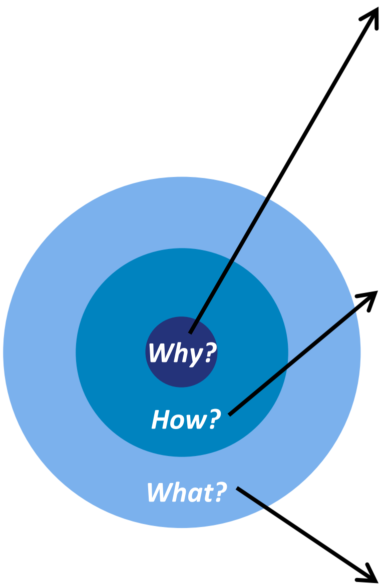
Further, in the United Kingdom, surveys show that people under 50 years old are more likely to switch to electric vehicles (EVs) in the next 10 years (ONS, 2021). This goes in alignment with the government's strategy to become net-zero by stopping the sale of new petrol/diesel cars by 2030 (Gov.uk, 2020). This switch will see an increased need for electric charging stations. Knowing where to install such stations is not an easily answered question which results in a slow rollout of charging infrastructure (Wan, et al., 2015).

Therefore, if the councils are focusing on understanding the parking behaviours and patterns, they could implement solutions that could optimise or improve the parking management leading to tackling those factors created by cruising for parking. The paper focuses only on *"pay-and-display"* parking (on-street or off-street/car park) which is delimited by parking lines.

For that reason, the proposed research question of this project is

> ***How could the council better optimise or improve their "pay-and-display" parking with the use of computer vision and predictive analysis?***

The basis for conducting this research could be explained using Simon Sinek's golden circle theory applied in the research context. Its core principles are to answer the questions of *"why"* (the purpose of doing something), *"how"* is it done and *"what"* is done (Sinek, 2011), *figure 1*.



The <u>motivation</u> for conducting this research has been affected by the wish to help the council have better parking management. This leading to behaviour understanding, improved land utilisation, policies, revenue, reduced air pollution and congestion and driver's frustration.

The <u>approach</u> in doing this research followed the Dynamic System Development Method (DSDM) methodology with the use of multiple datasets and video footage for predictive analysis and computer vision. Everything is wrapped in a usability-tested visual interface.

The <u>outputs</u> of the research include the contribution in this area, a proof of concept artefact and other possible optimisations or improvements that could be applied by the council for their *"pay-and-display"* parking management.

Figure 1: Golden circle

1.2.Aims & Objectives

The **main aim** of the project is to create an artefact and suggest further ways that could be used for optimising or improving the council's *pay-and-display* parking with the use of computer vision and predictive analysis.

A **secondary aim** is to enhance other types of parking within the pay-and-display on-/off-street, such as electric vehicles and *disabled bay spaces*.

To achieve these aims, the following **objectives** need to be met:

- ✓ Usage of multiple datasets and video footage.
- ✓ Apply data cleansing and data analysis to find patterns/insights.
- ✓ Overlap video occupancy with the respective dataset occupancy.
- ✓ Make vacancy predictions based on previous data.
- ✓ Include different parking space types.
- ✓ Create a visual interface that could be used by the council to understand behaviours and patterns of different locations.
- ✓ Propose further ways to optimise or improve the council's pay-and-display parking.

2.Literature Review

Due to the increased growth of urbanisation and car ownership, parking became a serious problem (Dogru, et al., 2017). Thus, **having good parking management and implicit efficient parking policies, the parking behaviours and patterns need to be understood**. In England, local authorities have direct control over their public on-/off-street parking. **General parking policies** include no parking zones, time-limited parking restrictions, space for loading but not parking, paid public parking, residents only parking areas (Rye & Koglin, 2014). **Studies looked at possibly changing the pricing policy into a dynamic one**, aiming at increasing the occupancy rate to a target level for a specific parking area (Simhon, et al., 2017), (Harris, 2014). Simhon, et al. (2017) mentioned that *"public parking has historically been priced ineffectively"* due to its lack of *"demand based smart pricing"*.

An article published by Brooke et al. (2017) looked at **UK's local government officials' views on parking search**. Their findings revealed that the interviewees did not find the parking search a serious problem but mentioned that due to the increase in car ownership, this will become a serious problem in a few years. Further, at the time of the interview, the councils' officials raised the lack of recorded evidence of parking search as an issue, stating that if this was the case, then they would look at implementing policies to address it. Moreover, one of the councils suggested that a thorough audit of current policies is essential in supporting the current parking and traffic needs. Further, when asked about possible solutions for parking search, their answers included improving the information available to drivers, review of current prices, time restrictions and capacity, and lastly, local sensors connected to mobile apps to inform of the vacancy status. But the latter was considered at the time *"financially prohibitive"* and only an option if in the future this becomes a serious problem.

Currently, there have been different solutions that attempt at solving the parking search issue, including sensors, vision-based systems and social crowdsensing and crowdsourcing. Using sensors requires expensive installation and maintenance (Rahman, et al., 2020). The usage of sensors is less expensive and not affected by weather conditions but has difficult installation needs and high maintenance requirements (Bibi, et al., 2017). Moreover, sensors cannot be applied for on-street parking due to the traffic flow (Bulan, et al., 2013). Using cameras is more cost efficient and it can monitor many parking spaces (Rahman, et al., 2020). Also, allowing surveillance and security purposes (Tatulea, et al., 2019) or if the infrastructure is already installed, it can be used for this purpose too. Lastly, regarding the smart parking apps that use crowdsensing/crowdsourcing to guide the user to the best locations of possibly vacant spaces or even allowing parking reservations/payments, their main disadvantages are the *"lack of maturity and the problems inherent in social systems"* (Enríquez, et al., 2017).

Focusing on vision-based systems such as cameras/CCTVs, there are multiple proposed methods and techniques. Yusnita, et al., (2012) proposes the use of image processing to detect space occupancy. The system is made of five modules: (1) detecting unobscured parking spaces using brown rounded patches from inside the parking space, (2) image acquisition, (3) image segmentation using greyscale and threshold techniques, (4) image enhancement (using dilation and erosion) and lastly, (5) image detection module indicating the number of vacant parking spaces (Yusnita, et al., 2012). Their system stands as a cost-efficient solution and their experimental results show correct vacancy counting. However, heavy weather conditions impact its ability of correct detection (Bibi, et al., 2017) and the camera needs to capture a top-down view of the spaces (Padmasiri, et al., 2020). Bibi, et al. (2017) propose a similar approach as Yusnita, et al. (2012) but without having brown rounded patches in each space. Their system gets an image frame from the video, converts it to greyscale, divides spaces into blocks, applies binary and inverse binary, the threshold is calculated and based on it is determined the occupancy status (Bibi, et al., 2017). Their results show that the proposed algorithm has an accuracy of 94% to 100% based on different weather and vehicle appearance conditions. Overall, their study aims to optimising parking detection. Another paper by Kommey, et al. (2018) looks at using image processing on aerial images of parking lots. The system's modules are similar to the ones seen above ((1) initialisation, (2) image acquisition and processing, (3) interpretation and (4) results) (Kommey, et al., 2018). The results showed that it can accurately detect occupancy status and it can be a viable solution. Additionally, the shown used image for testing looks more like an image from a CCTV camera. Furthermore, to help with the possible issues presented in the image/frame feed used for image processing, Tutika, et al. (2018) proposes an algorithm to help with the *"uneven illumination, distorted slot lines and overlapping of cars"* using *"image pre-processing and false contour detection techniques"*.

Other studies look at using more sophisticated solutions. Amato, et al. (2016) look at using convolutional neural network (CNN) for real-time occupancy detection, but it must be run directly on a smart camera and the testing has not happened in unsatisfactory weather situations. Bulan, et al. (2013) use video cameras and the frames are sent to a central processing unit for determining the occupancy status using video processing and computer vision techniques. Acharya, et al. (2018) use CNN with support vector machine (SVM) classifier to determine the occupancy status. The obtained results were highly accurate, but illumination and sun reflections are system limitations with the night-time evaluation remaining in future research (Acharya, et al., 2018).

Alternative studies look at forecasting occupancy rates. Guerrini, et al. (2021) use a Prophet model to forecast a month of occupancy rate for different locations, resulting in a viable solution with a reliable occupancy forecast. Tamrazian, et al. (2015) compared two unsupervised learning approaches in predicting the occupancy rate with the online/real-time data approach proving to be more accurate (than the offline/historical data approach). The results showed an efficient online approach to predicting occupancy rates for different times of the day (Tamrazian, et al., 2015). Moreover, with enough real-time occupancy data collected, the prediction error declined drastically (Tamrazian, et al., 2015). Lastly, Simhon, et al. (2017) proposed the use of least-squares regression to predict the occupancy rate and an optimisation method to adapt the prices to meet an occupancy target level rate.

In addition to those, Dogru, et al. (2017) looked at **understanding the parking behaviours** to help in choosing more suitable parking management policies. Their study suggests that the parking behaviours and the most user-preferred parking policies are particular to each sub-area (Dogru, et al., 2017). Emphasis was put on the fact that understanding the parking behaviour brings great benefit to the city planners in proposing the best parking facilities and policies (Dogru, et al., 2017). Another study by Pflügler, et al. (2016) assessed the **factors that influence parking prediction**. Their results show that the most significant impact on traffic behaviours is given by time (including holidays periods), location and weather (most importantly temperature) (Pflügler, et al., 2016). The limitations of their study include the short period of time when the data was collected (July-September 2015) and the one city evaluation (Pflügler, et al., 2016).

Moving forward, in alignment with the government's strategy, the expectation of having public **electric vehicle charging stations** on-/off-streets across the city is highly expected (Liu, et al., 2016). Moreover, Brandstätter, et al. (2017) look even more to the future, that is when the EVs are shared. Raising the issue that to support EV car-sharing systems, the charging stations need to be strategically placed across an area. By having this system built, other researchers raise the problem of vehicle relocation (Deza, et al., 2022). As of July 2022, the UK has 32,011 public EV charging stations of which 5,974 are rapid chargers (Gov.uk, 2022). The findings of Morrissey, et al. (2016) who looked at the charging behaviours of EV owners suggest that the fast-charging infrastructure located in car parks is the most desired one and the planning of future infrastructure needs to be strategically located to meet the users' overall needs.

Additionally, another study proposes an optimal charging scheme that would save EVs' owners money while helping to balance the supply and demand of the main grid (He, et al., 2016).

To round off, Mangiaracina, et al. (2017) conducted **simulations in the City of Milan regarding a smart parking** app that focused on pay-and-display parking spots in which the driver is informed of their occupancy. The results showed that this could save the driver 77.2 hours a year, ~£74 in fuel costs and the CO_2 emissions in the city will reduce by 44,470 tons per year. In addition, revenue improvements could be seen each year due to reducing the number of cars parked that have not paid for parking. In the UK, in 2019, transport produced 27% of the total greenhouse gas emissions with the biggest contributors being cars and taxis (61%) (Gov.uk, 2021).

3.Methodology

This section presents the methodologies used in collecting the datasets, artefact's evaluation, gathering research papers, professional, legal and ethical issues and the project management applied.

3.1.Datasets & Video Footage Collection

The datasets are either collected as is from the internet, collected from the internet and altered or manually created. Those datasets were chosen to support the aims of the project and to help answer the research question.

Collected as is:

- ***pay-and-display ticket machine logs*** between 01.12.2012 and 30.03.2012 for *River Road 1 & 2 Yarmouth car park (Isle of Wight)* found under a Freedom of Information request to Isle of Wight Council
  (source:
  https://www.whatdotheyknow.com/request/pay_and_display_ticket_machine_l).
- ***EV charging transactions*** from the London Borough of Barnet
  (source: https://data.gov.uk/dataset/16c7326b-57fe-4803-88f8-9286c387f68a/electric-vehicle-charging-transactions).

Partially altered:

- ***weather report*** from Newport, Isle of Wight
  (sources: 2012 - http://www.isleofwightweather.co.uk/2012_data.htm;
  2013 - http://www.isleofwightweather.co.uk/2013_data.htm)*.
  The data was not available in a downloadable format. Therefore, it was manually transferred within an Excel format and not all of the columns available were transferred over.
- ***Bank holidays in England*** in 2012 & 2013
  (sources: 2012 - https://www.ukbankholidays.co.uk/year/2012;
  2013 - https://www.ukbankholidays.co.uk/year/2013)*.
  The data was not available in a downloadable format. Therefore, a new Excel sheet called "*holidays*" was created. Part of the information presented within those links was added to the column *"Event"* and only focused on the period 01.12.2012-31.03.2012.

- **School holiday dates for 2012-2013**

  (source:

  https://moderngov.kingston.gov.uk/documents/s25942/TERM%20DATES%202012%2013.html?CT=2)

  The data was not available in a downloadable format. Therefore, a new Excel sheet called "*holidays*" was created. Part of the information presented within this link was added to the column *"School holiday?"* and only focused on the period 01.12.2012-31.03.2012.
- For the purposes of the project, the dataset copied as is for *"pay-and-display ticket machine logs"* was duplicated and the copy was altered to symbolise a location that has a maximum of 69 *regular* parking bay spaces. Therefore, some rows were deleted, and some duplicate rows were added.

Manually created:

- **Count of *regular/disabled* occupied spaces** based on video* for each location. This has been explained in 4.4.3.Video Dataset Creation.

*The video is only 73 seconds, and the remaining of the data was *almost* randomly produced.

11.1.Appendix A – Datasets shows an in-depth look at all the datasets used within this project, including other datasets created based on the presented datasets.

The **video footage** was collected from the YouTube channel *Tom Berrigan* (source: https://www.youtube.com/watch?v=RY6eu3fZ-Lg). The video was split into two areas **to represent two different locations with diverse amounts of parking spaces available**.

3.2.Usability Testing & User Experience Questionnaire Methodology

The usability testing and the user experience questionnaire have been conducted by the same five participants. The tests have either been conducted face-to-face or online and in both scenarios the meeting has been recorded (including screen recording).

Their demographics include being aged between 27 to 33, 4 males and 1 female, Solent University graduates with job titles such as data analyst, software engineer/developer, sensory panel manager and receptionist.

3.2.1.Usability Test Design

Usability testing is a method which consists in asking users to perform some tasks to assess the product's ease of use and overall perception of the experience (Niranjanamurthy, et al., 2014). The aim is to use the findings to improve the created artefact. It was found that for usability testing, by having 4/5 users that think aloud, the tester will get the maximum insights (Nielsen & Budiu, 2021).

The tasks were designed to evaluate different metrics of the artefact. These include the ***layout*** (ability to detect something the user needs to find), ***terminology*** (understanding the artefact's wording), ***navigation*** (understanding the ways around the artefact), ***feedback*** (receiving a response when an action is made), ***comprehension*** (understanding the instructions given) and ***data entry*** (inserting/changing information within a field).

The **testing techniques** used were *concurrent probing* for the tester as follow-up questions were asked after each task to better understand the participant's thoughts and overall *concurrent thinking aloud* was chosen to be used by participants as they were instructed to either think aloud while doing the task or to verbalise their thoughts after finishing the task.

A **success rate** was calculated based on the time required to complete the task and based on the number of clicks made for each task by each participant. It is important to note that the participants have not been instructed to accomplish the tasks as quick as possible. Therefore, the number of clicks gives a quantitative view.

A **benchmark** for each task was calculated based on a user who is familiar with the artefact. Of course, the time and number of clicks will be different between the familiar user and a first-time user, but it puts in perspective the difference once familiar with the system.

Few **post-test questions** were also asked (at the end of the test) to gain a better understanding of the participant's thoughts and feedback. The results could be found in 5.2.Usability Testing.

3.2.2.User Experience Questionnaire Test Design

The user experience questionnaire (UEQ) is a quick way to assess the users' comprehensive impression of the product's user experience (Schrepp, et al., 2017). It can be used as part of the usability testing or as a stand-alone and it is made of 26 pairs of terms with opposite meanings on a 7-point scale (Schrepp, et al., 2017), *figure 2*. It takes about 3-5 minutes to complete and it is instructed to add some demographic questions (Schrepp, et al., 2017) which were *age, sex* and *job title*.

annoying	o o o o o o o	enjoyable	1
not understandable	o o o o o o o	understandable	2
creative	o o o o o o o	dull	3
easy to learn	o o o o o o o	difficult to learn	4
valuable	o o o o o o o	inferior	5
boring	o o o o o o o	exciting	6
not interesting	o o o o o o o	interesting	7
unpredictable	o o o o o o o	predictable	8
fast	o o o o o o o	slow	9
inventive	o o o o o o o	conventional	10
obstructive	o o o o o o o	supportive	11
good	o o o o o o o	bad	12
complicated	o o o o o o o	easy	13
unlikable	o o o o o o o	pleasing	14
usual	o o o o o o o	leading edge	15
unpleasant	o o o o o o o	pleasant	16
secure	o o o o o o o	not secure	17
motivating	o o o o o o o	demotivating	18
meets expectations	o o o o o o o	does not meet expectations	19
inefficient	o o o o o o o	efficient	20
clear	o o o o o o o	confusing	21
impractical	o o o o o o o	practical	22
organized	o o o o o o o	cluttered	23
attractive	o o o o o o o	unattractive	24
friendly	o o o o o o o	unfriendly	25
conservative	o o o o o o o	innovative	26

Figure 2: English version of the UEQ (Schrepp, et al., 2017)

Answering closest to a negative term means *-3* and answering closest to a positive term means *+3*. They are grouped into 6 scales (with blue, *figure 3*).

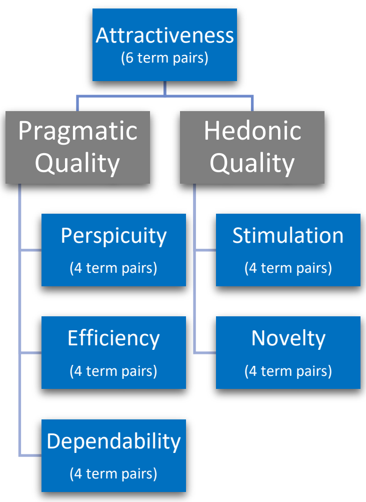


Figure 3: Scale structure

Attractiveness gives the overall impression of the product (Schrepp, et al., 2017).

Perspicuity looks at the product's ease of learn, understandability, familiarity (Schrepp, et al., 2017).

Efficiency looks at the product's time response, interaction efficiency and speed, ease of achieving the task (Schrepp, et al., 2017).

Dependability looks at user's feel of control within the interaction and overall safety and system behaviour prediction (Schrepp, et al., 2017).

Stimulation looks at the user's excitement and motivation to use the product (Schrepp, et al., 2017).

Novelty looks at product's creativity and innovation alongside the user's attention towards it (Schrepp, et al., 2017).

The ***pragmatic qualities*** are goal orientated (achieving the tasks/goals) and ***hedonic qualities*** are not goal orientated but related to enjoyment of using the product (Schrepp, et al., 2017).

3.2.3.Overall Testing Path

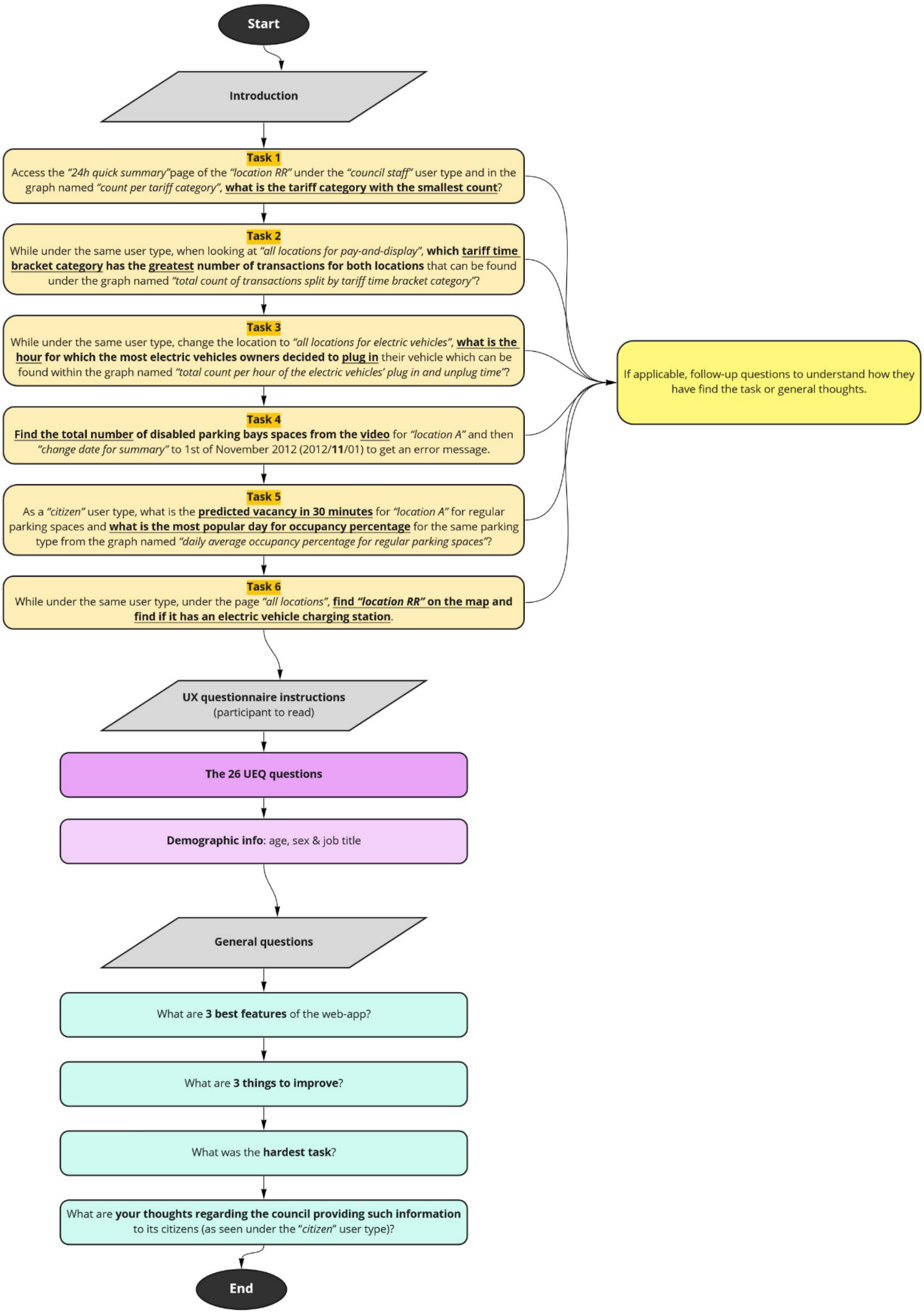


Figure 4: Overall testing path

3.3.Literature Review Methodology

The literature review methodology followed the approach proposed by Solent University which includes defining the *"inclusion and exclusion criteria, identifying databases, conducting searches, reviewing resulting literature and presenting findings"* (Solent University, 2021).

The **inclusion parameters** were to be primary/secondary sources related to the research question, to be written in English, to be peer-reviewed (excluding conferences) and to be published on or after 2006. The **exclusion criteria** are anything that does not meet those parameters.

The **used databases** included (but were not limited to): *Google Search, IEEE, Research Gate, Science Direct, Solent Library's online catalogue, Springer.*

The searches included sentences such as: *"pay and display parking", "parking occupancy detection using computer vision", "parking meter and computer vision", "council parking management", "electric cars parking council", "image segmentation for parking occupancy", "weather and holidays influence on parking".* Some of the articles have been found via the references in other articles.

Lastly, the articles that did not meet the inclusion criteria have been marked with red, *figure 5*.

ID	Title	Year	Comments	Available at:	Search criteria	Aim of the study	Main findings	Peer-reviewed?
1	Statistical Analysis and Prediction of Parking Behavior	2019	downloaded & reviewed	https://link.springer.com/chap	weather and holidays influence on parking	In this paper, we analyze t	Experiments show that t	Yes
2	Car parking occupancy detection using smart camera networks a	2016	downloaded & reviewed	https://ieeexplore.ieee.org/ab	parking occupancy detection and prediction	In this paper, we provide a	Our experiments show th	Conference
3	Video-based real-time on-street parking occupancy detection sy	2013	downloaded & reviewed	https://www.spiedigitallibrary	parking occupancy detection using computer visio	In this article, we present	This article paves the wa	Yes
4	Real-time image-based parking occupancy detection using deep	2018	downloaded & reviewed	https://www.researchgate.net	parking occupancy detection using computer visio	The hypothesis of this rese	We report detection accu	Conference
5	Convolutional Neural Network Customization for Parking Occupa	2020	downloaded & reviewed	https://ieeexplore.ieee.org/ab	parking occupancy detection using computer visio	Parking space classificatio	cmAlexnet has better acc	Conference
6	Automated Vehicle Parking Occupancy Detection in Real-Time	2020	downloaded & reviewed	https://ieeexplore.ieee.org/ab	parking occupancy detection using computer visio	This paper presents an ap	The proposed solution is	Conference
7	Estimation of Free Space on Car Park Using Computer Vision Alg	2019	downloaded & reviewed	https://link.springer.com/chap	parking meter and computer vision	In our work, we proposed	95% accuracy.	Yes
8	Lean smart parking	2014	downloaded & reviewed	https://www.parking-mobility	parking meter and computer vision	It is about how to cost-effe	The best choice of lean s	Cannot find
9	Optimal parking occupancy with and without differentiated park	2020	downloaded & reviewed	https://www.research-collecti	parking meter analysis	Our study proposes a mac	In this paper, we propose	No
10	Parking Management Policies Based on Behavior Analysis at Fati	2017	downloaded & reviewed	https://www.sciencedirect.co	parking meter analysis	Thus, this paper aims to ex	Parking management po	Yes
11	Case analysis of simultaneous concessions of parking meters and	2013	downloaded & reviewed	https://www.sciencedirect.co	parking meter analysis	In this article we assess th	The results corroborate t	Yes
12	Analysis, Design and Implementation of a Forecasting System for	2021	downloaded & reviewed	https://dl.acm.org/doi/pdf/10	parking meter analysis	Study looks at providing th	The preliminary results r	Conference
13	Deep learning for decentralized parking lot occupancy detection	2017	downloaded & reviewed	http://www.nmis.isti.cnr.it/fal	Used the references from article number 6	The paper proposes a dece	Our experiments show th	Yes
14	DiG-Park: A smart parking availability searching method using V2	2017	downloaded & reviewed	https://ieeexplore.ieee.org/do	Used the references from article number 6	In this paper we propose a	We use a smart combina	Conference
15	Parking-stall vacancy indicator system, based on deep convolutio	2016	downloaded & reviewed	https://ieeexplore.ieee.org/do	Used the references from article number 6	In this paper, we present a	In this paper we designe	Conference
16	Fast Classification of Empty and Occupied Parking Spaces Using I	2016	downloaded & reviewed	https://ieeexplore.ieee.org/do	Used the references from article number 6	In this paper we present a	The experimental evalua	Conference
17	Residential Permit Parking: Better Off Without It?	2014	downloaded & reviewed	https://journals.sagepub.com	pay and display parking	The study presented in thi	The study found that this	Yes
18	Intelligent Parking Management System Based on Image Process	2012	downloaded & reviewed	http://ijimt.org/papers/228-G	Used the references from article number 6	This paper aims to present	It makes the process of a	Yes
19	A Survey of Smart Parking Solutions	2017	downloaded & reviewed	https://ieeexplore.ieee.org/do	Used the references from article number 6	We go through the literatu	Our survey gives an exha	Yes
20	The changing face of parking-related data collection and analysis	1991	downloaded & reviewed	https://link.springer.com/artic	Used the references from article number 7	Current trends in requirem	New methods of analysis	Yes
21	Smart parking pricing: A machine learning approach	2017	downloaded & reviewed	https://ieeexplore.ieee.org/ab	Used the references from article number 7	Instead, we propose a ma	We conduct a numerical	Conference
22	Cruising for parking	2006	downloaded & reviewed	https://www.sciencedirect.co	Used the references from article number 11	This paper presents a mod	City governments therefo	Yes
23	On the optimal target curbside parking occupancy rate	2014	downloaded & reviewed	https://www.sciencedirect.co	Used the references from article number 9	This paper develops a simp	This suggests that, in pra	Yes
24	A dynamic macroscopic parking pricing and parking decision mo	2018	downloaded & reviewed	https://www.tandfonline.com	Used the references from article number 9	The proposed responsive p	In summary, the propose	Yes
25	Parking Pricing vs. Congestion Pricing: A Macroscopic Analysis of	2020	downloaded & reviewed	https://www.researchgate.net	Used the references from article number 9	This paper focuses on park	We develop a multimoda	Yes
26	Where is My Parking Spot?: Online and Offline Prediction of Tim	2015	downloaded & reviewed	https://journals.sagepub.com	Used the references from article number 9	This study proposed efficie	The offline procedure sig	Yes
27	Smart parking management in a smart city: Costs and benefits	2017	downloaded & reviewed	https://ieeexplore.ieee.org/ab	pay and display parking	In this article we attempt t	The results obtained sho	Conference
28	Problems and prospects of curbside parking in Lahore: Policy imp	2017	downloaded & reviewed	https://hal.archives-ouvertes.f	pay and display parking	This paper provides a criti	It concludes that develop	Yes
29	A new policy tool: dynamic pricing of on-street parking	2014	downloaded & reviewed	https://digital-library.theiet.or	pay and display parking	The paper presents the cu	Results from the Los Ang	Conference
30	Predicting the Availability of Parking Spaces with Publicly Availab	2016	downloaded & reviewed	https://www.researchgate.net	weather and holidays influence on parking	This article deals with the	The results show that we	Conference
31	Parking guiding system with occupation prediction	2019	downloaded & reviewed	https://repositorio.iscte-iul.pt	weather and holidays influence on parking	MSc Thesis	N/A	N/A
32	The Optimal Distribution of Electric-Vehicle Chargers across a Cit	2016	downloaded & reviewed	https://ieeexplore.ieee.org/ab	electric cars parking and chargers	In this work, we aim to op	The extensive tests verify	Conference
33	Optimal Charging Strategy of Electric Vehicles Customers in a Sm	2016	downloaded & reviewed	https://ieeexplore.ieee.org/do	electric cars parking and chargers	An optimal charging schen	It is demonstrated that t	Conference
34	Future standard and fast charging infrastructure planning: An an	2016	downloaded & reviewed	https://www.sciencedirect.co	electric cars parking and chargers	This study provides an ext	Car park locations were	Yes
35	Existing Approaches to Smart Parking: An Overview	2017	downloaded & reviewed	https://link.springer.com/chap	parking apps uk	In this paper, we give an o	In this paper, the main a	Conference
36	Reducing Parking Space Search Time and Environmental Impacts	2020	downloaded & reviewed	https://ieeexplore.ieee.org/ab	parking apps uk	Presenting a technology d	In this article, we present	Yes
37	Parking futures: The relationship between parking space, everyd	2020	downloaded & reviewed	https://www.sciencedirect.co	parking apps uk	The paper proposes and d	The paper sets out to put	Yes
38	Investigating the use of Machine Learning for Smart Parking App	2019	downloaded & reviewed	https://ieeexplore.ieee.org/ab	parking apps uk	This paper provides an ove	While most algorithms p	Yes
39	An Image Feature-Based Method for Parking Lot Occupancy	2019	downloaded & reviewed	https://www.mdpi.com/1999-	image segmentation for parking occupancy	We are proposing a metho	The system has shown hi	Yes
40	Automatic Parking Space Detection System	2017	downloaded & reviewed	https://ieeexplore.ieee.org/ab	image segmentation for parking occupancy	In this paper, we have des	The performance accura	Conference
41	An Algorithm for Parking Lot Occupation Detection	2008	downloaded & reviewed	https://ieeexplore.ieee.org/ab	image segmentation for parking occupancy	This paper presents unsup	Experimental results fron	Conference
42	A vehicle parking detection method using image segmentation	2001	downloaded & reviewed	https://onlinelibrary.wiley.com	image segmentation for parking occupancy	The present study aims at	The proposed method wa	N/A
43	China's electric car frustrations	2015	downloaded & reviewed	https://www.sciencedirect.co	electric cars parking council	Talks about overcoming lo	It is uncertain which are	Yes
44	Charging station optimization for balanced electric car sharing	2022	downloaded & reviewed	https://www.sciencedirect.co	electric cars parking council	This work focuses on findi	Results of a numerical ex	Yes
45	Determining optimal locations for charging stations of electric ca	2017	downloaded & reviewed	https://www.sciencedirect.co	Used the references from article number 44	In this article, we introduc	Results show that the de	Yes
46	Parking Management	2014	downloaded & reviewed	https://www.emerald.com/ins	council parking management	This chapter explains how	The findings of this chap	Yes
47	The causes and consequences of curb parking management	2021	downloaded & reviewed	https://www.sciencedirect.co	council parking management	Looks at the causes and co	Conclusion: A "Low Paym	Yes
48	On-street parking management and pricing policies: An evaluatio	2021	downloaded & reviewed	https://www.sciencedirect.co	council parking management	This paper introduces an a	A list of findings are as fo	Yes
49	On-street parking search: A UK local authority perspective	2017	downloaded & reviewed	https://www.jstor.org/stable/	council parking management	The aim of this paper is to	City and county council d	Yes
50	Parking Policy	2014	downloaded & reviewed	https://www.emerald.com/ins	council parking management	This chapter provides an o	The reviews reveal that t	Yes
51	A Smart Image Processing-based System for Parking Space Vacan	2018	downloaded & reviewed	https://www.researchgate.net	Parking lot detection using image processing met	This paper presents a sma	The proposed system acc	Yes
52	Parking Space Detection Using Image Processing	2021	downloaded & reviewed	https://www.ijsr.net/archive/v	Parking lot detection using image processing met	In this paper proposes par	The main contribution of	Yes
53	Image Segmentation for Efficient Parking Space Analysis	2018	downloaded & reviewed	https://www.researchgate.net	Parking lot detection using image processing met	Aim: To develop an algorit	The proposed algorithm	Conference

Figure 5: Literature review

The above figure can be found below:

https://docs.google.com/spreadsheets/d/1_H6a6FsyvoCfOmKvy4tfAlX359IHDKQ5/edit?usp=sharing&ouid=101649927425582474373&rtpof=true&sd=true

3.4.Professional, Legal & Ethical Issues

The principles of research ethics presented by Mondal (2020) have been aimed at being followed while completing this project. They included:

- ✓ The research participants have been subject to no harm, respected and their privacy and anonymity ensured.
- ✓ Full written consent.
- ✓ Any misleading information, bias, deception/exaggeration of the aims of the project and discrimination was aimed at being prevented.
- ✓ The communication was done honestly and transparently.
- ✓ The intellectual property was respected.

When conducting the usability tests, the participants have been informed about the purpose of the project and their consent to having the data collected has been obtained in a written format (see 11.3.Appendix C – Consent Forms Example for the consent forms example that has been signed). Their names have been anonymised in this report. The meeting recordings are stored securely and will be deleted once no longer needed. No incentives were offered.

Lastly, the project has received approval of ethical clearance for research and innovation from Solent University (11.4.Appendix D – Ethical Clearance).

3.5.Project Management

3.5.1.Project Management Approach

The project management approach taken to complete this *"one-person"* project fell under the **Dynamic System Development Method** (DSDM), *figure 6*. This approach is beneficial as it allows looping through previous phases due to its dynamic development (Nazir, et al., 2017). Further, DSDM is part of the agile software development methodology (Nazir, et al., 2017).

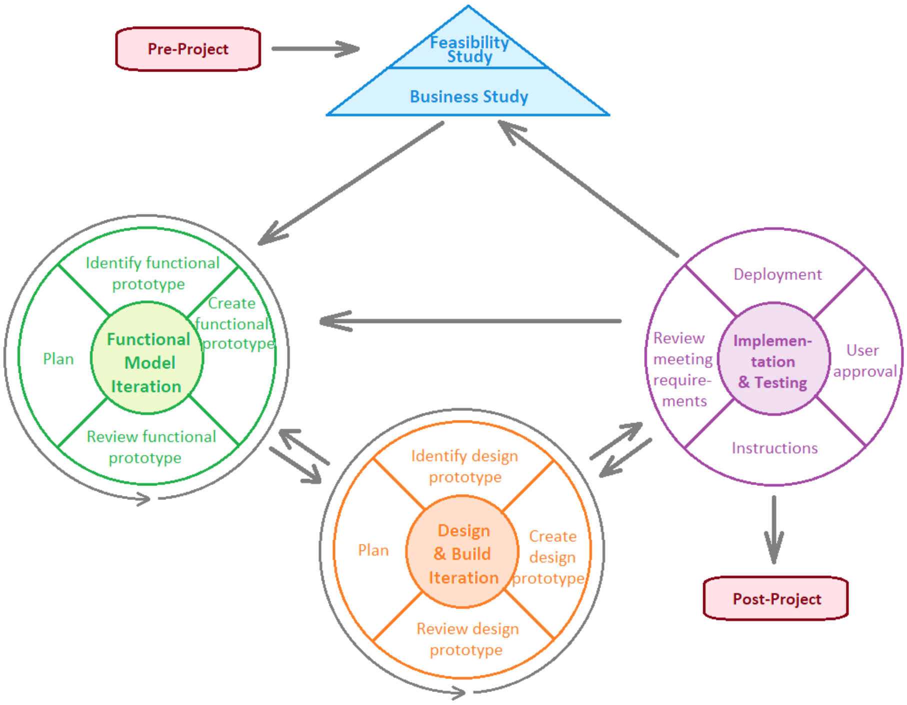


Figure 6: DSDM

The main advantage of the DSDM is the dynamic development which helps in firstly building the most important functionalities and using the iterative and incremental process to carry on with the remaining prioritised functionalities (Anwer, et al., 2017). Furthermore, it is used to provide a rapid application development which blends in best practices and possibly other approaches (Anwer, et al., 2017). **The applicable disadvantage of DSDM in this project** is the lack of guidance regarding the length of each iteration (Anwer, et al., 2017).

Because DSDM allows the integration of other approaches, the **Kanban board** (from the Kanban method) was integrated into this project. It helped in visualising the workload which was split by statuses of *"to do"*, *"doing"* and *"done"* (Klipp, 2014).

3.5.2.Project Timeline & Milestones

The project's timeline is based on the DSDM phases and the unit requirements. Using **Monday.com** which is a workload tracking and management tool (Monday.com, 2021) the project's Gantt chart, timeline and milestone have been defined.

The project has in total 3 milestones representing each submission for the unit. *Figure 7* shows the Gantt chart and milestones of the project.

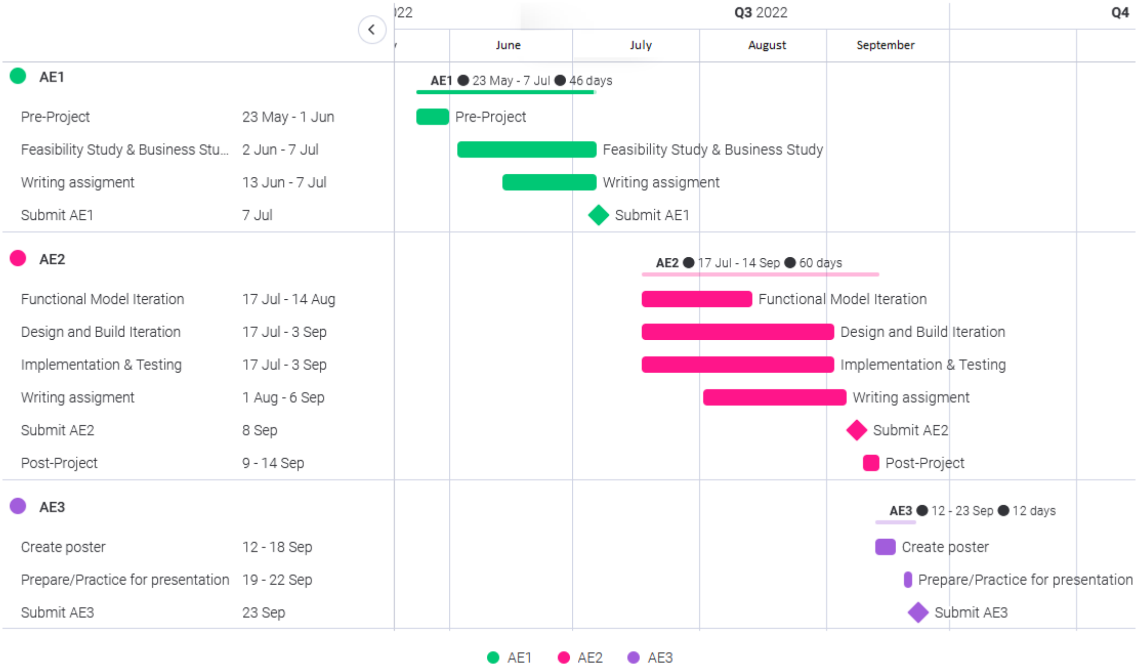


Figure 7: Gantt chart

Figures 8 & 9 show the in-depth breakdown of the 600 hours spent on the project.

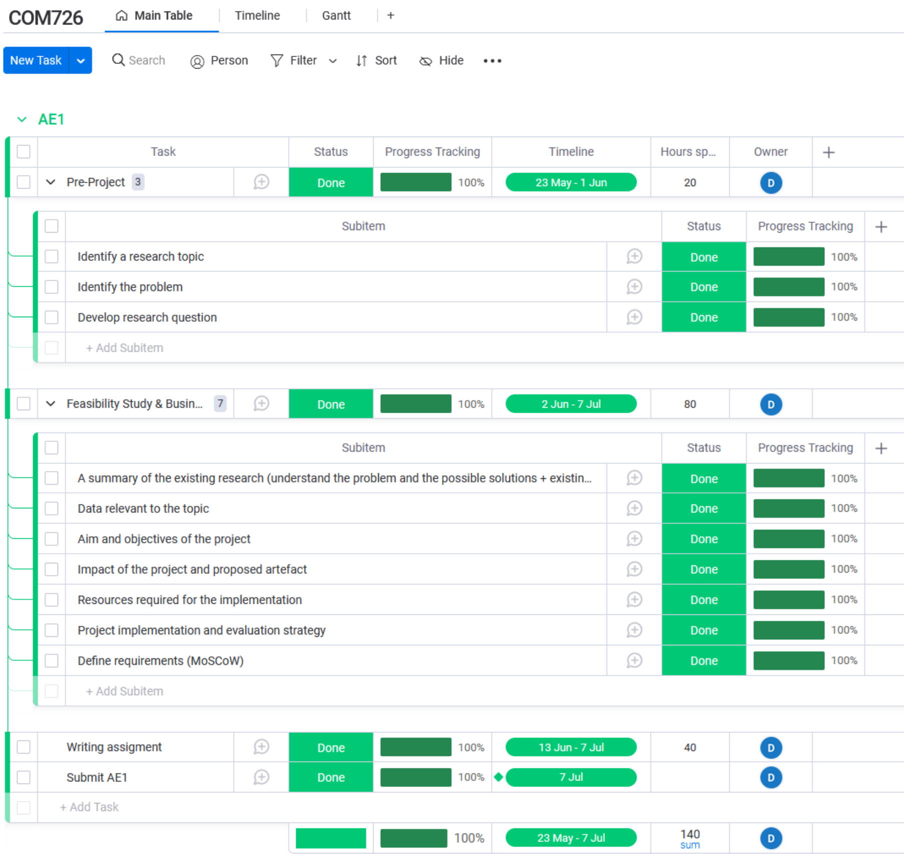


Figure 8: In-depth breakdown (1)

COM726 — Main Table | Timeline | Gantt | +

New Task | Search | Person | Filter | Sort | Hide | ...

AE2

Task	Status	Progress Tracking	Timeline	Hours sp...	Owner
Functional Model Iteration 4	Done	100%	17 Jul - 14 Aug	50	D

Subitem	Status	Progress Tracking
Identify key functionalities to be built	Done	100%
Create the prototype based on the functional requirements (+test & improve)	Done	100%
Review the implemented functionalities (=>further improvements areas)	Done	100%
Create a plan for finalising the functionalities	Done	100%
+ Add Subitem		

Task	Status	Progress Tracking	Timeline	Hours sp...	Owner
Design and Build Iteration 4	Done	100%	17 Jul - 3 Sep	150	D

Subitem	Status	Progress Tracking
Identify design prototype needs and prioritise them	Done	100%
Create the design prototype to achieve satisfactory standards (and apply possible improvements)	Done	100%
Review the design prototype by testing it and apply possible improvements	Done	100%
Create a plan for implementing the design prototype needs	Done	100%
+ Add Subitem		

Task	Status	Progress Tracking	Timeline	Hours sp...	Owner
Implementation & Testing 4	Done	100%	17 Jul - 3 Sep	110	D

Subitem	Status	Progress Tracking
Artefact/Solution deployment to operational environment	Done	100%
User approval/Testing	Done	100%
Review if the solution meets requirements	Done	100%
Artefact instructions	Done	100%
+ Add Subitem		

Task	Status	Progress Tracking	Timeline	Hours sp...	Owner
Writing assigment	Done	100%	1 Aug - 6 Sep	110	D
Submit AE2	Doing	0%	8 Sep		D
Post-Project 2	Doing	0%	9 - 14 Sep	10	D

Subitem	Status	Progress Tracking
Reflect on the project	Doing	0%
Future work	Doing	0%
+ Add Subitem		

Task	Status	Progress Tracking	Timeline	Hours sp...	Owner
+ Add Task					
		67%	17 Jul - 14 Sep	430 sum	D

AE3

Task	Status	Progress Tracking	Timeline	Hours sp...	Owner
Create poster	To do	0%	12 - 18 Sep	25	D
Prepare/Practice for presen...	To do	0%	19 - 22 Sep	15	D
Submit AE3	To do	0%	23 Sep		D
+ Add Task					
		0%	12 - 23 Sep	40 sum	D

Supervisor Meetings

Task	Status	Progress Tracking	Timeline	Hours sp...	Owner
Meeting 1	Done	100%	13 Jun		D
Meeting 2	Done	100%	21 Jun		D
Meeting 3	Done	100%	1 Jul		D
Meeting 4	Done	100%	29 Jul		D
Meeting 5	Done	100%	4 Aug		D
Meeting 6	Done	100%	1 Sep		D
+ Add Task					
		100%	13 Jun - 1 Sep	0 sum	D

Figure 9: In-depth breakdown (2)

3.5.3.Project Contingency Planning

To ensure a quick restoration to normality in the event of an unfortunate scenario, a contingency planning has been put together and split into four categories, *figure 10*.

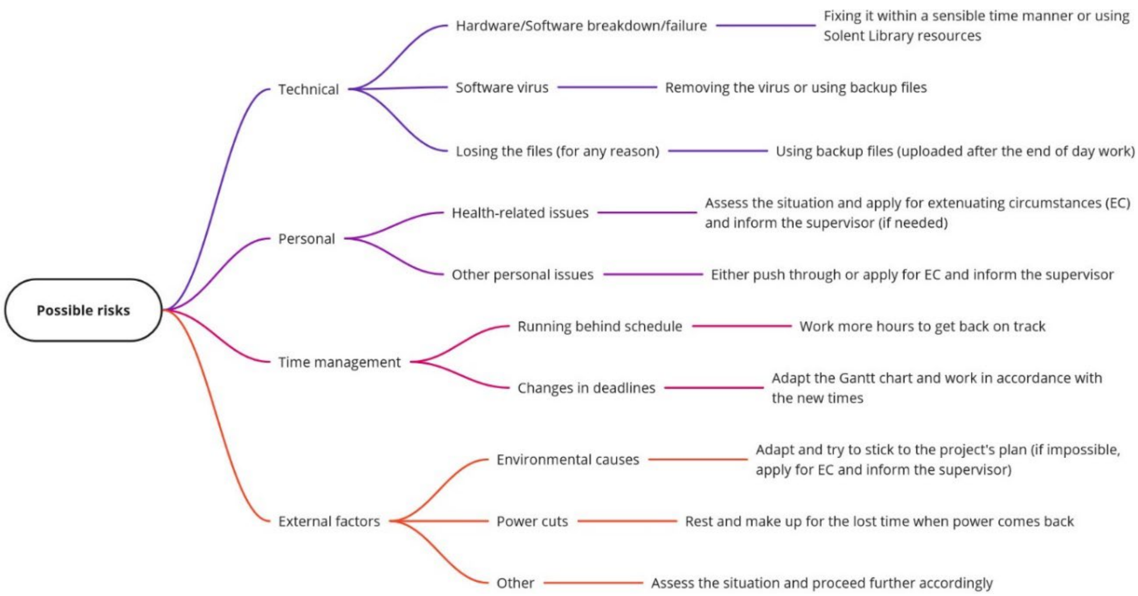


Figure 10: Contingency planning

Fortunately, no unfortunate scenarios happened, but daily work backup was applied.

4.Design & Implementation

4.1.Artefact Requirements

Initially, the general requirements have been defined using MoSCoW method (*Must Have, Should Have, Could Have and Won't Have*), *figure 11*. This method is a way to classify requirements based on their own value and it has been proposed by Clegg and Baker in 1994 (Miranda, 2022).

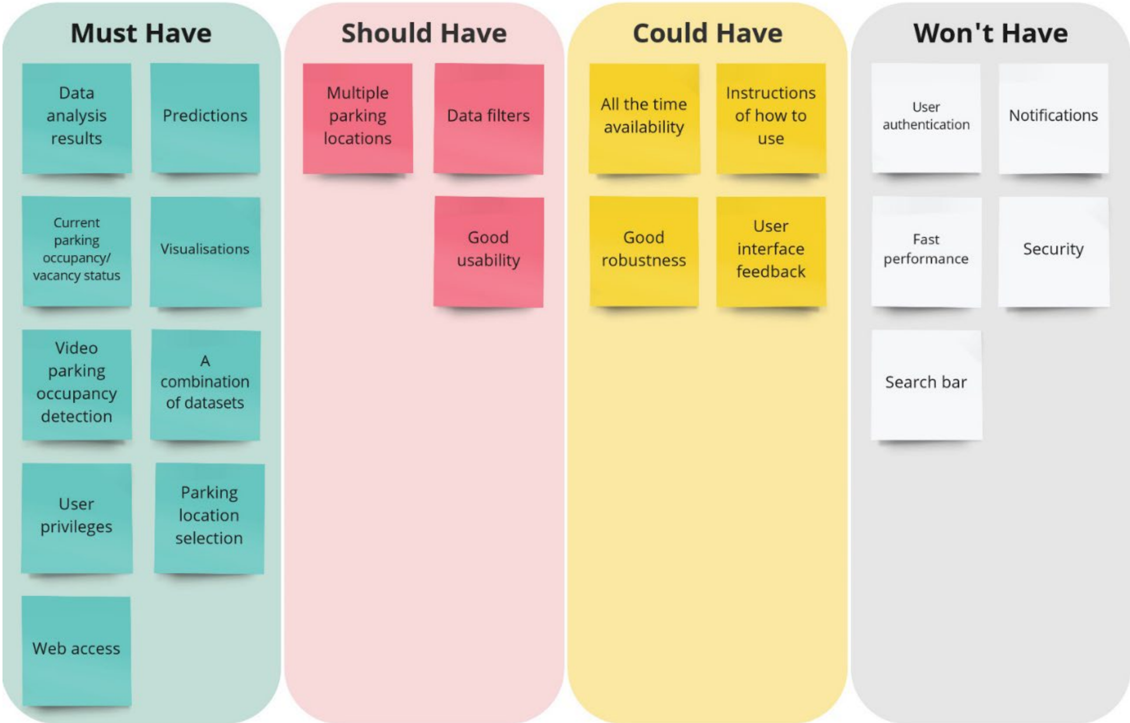


Figure 11: MoSCoW

Further, the functional and non-functional requirements of the artefact have been defined based on the MoSCoW diagram, *figure 12*. The functional requirements consider the functionalities of the system (what it should do) and the non-functional requirements consider describing properties and constraints (how it should do it) (Becker, et al., 2019), (Kurtanović & Maalej, 2017).

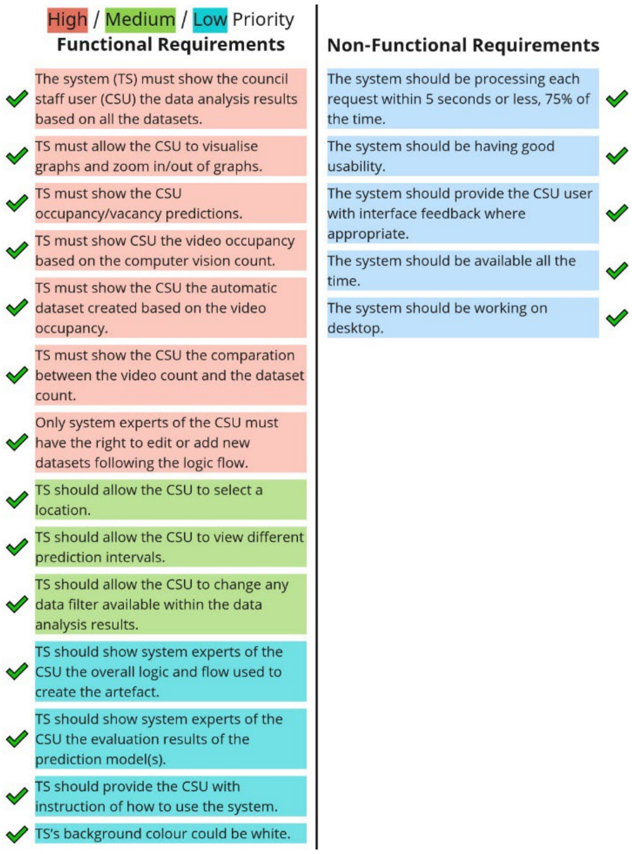


Figure 12: Functional and non-functional requirements

4.1.1.Tools Chosen

Experiments with **PyCharm**, **OpenCV, Anaconda.Navigator** and **Streamlit** were conducted in the feasibility study. Due to their successful experimentation and implementation, they have been chosen for this artefact. Furthermore, **Jupyter Notebook** was used for the purposes of data cleansing, data analysis testing, prediction model evaluation and other general testing.

4.2.Use Case Diagrams

Additionally, two use case diagrams have been created, *figure 13*. The first one shows a high-level council staff perspective overview, displaying the proposed options the system will provide to such users. The extended relationship (marked *<<extend>>*) may only happen sometimes and not all the time.

The second diagram suggests a possible improvement suggestion that could be provided by the council to its citizens. The included relationships (marked *<<included>>*) should always happen every time, opposite to the extended relationship that may only happen sometimes.

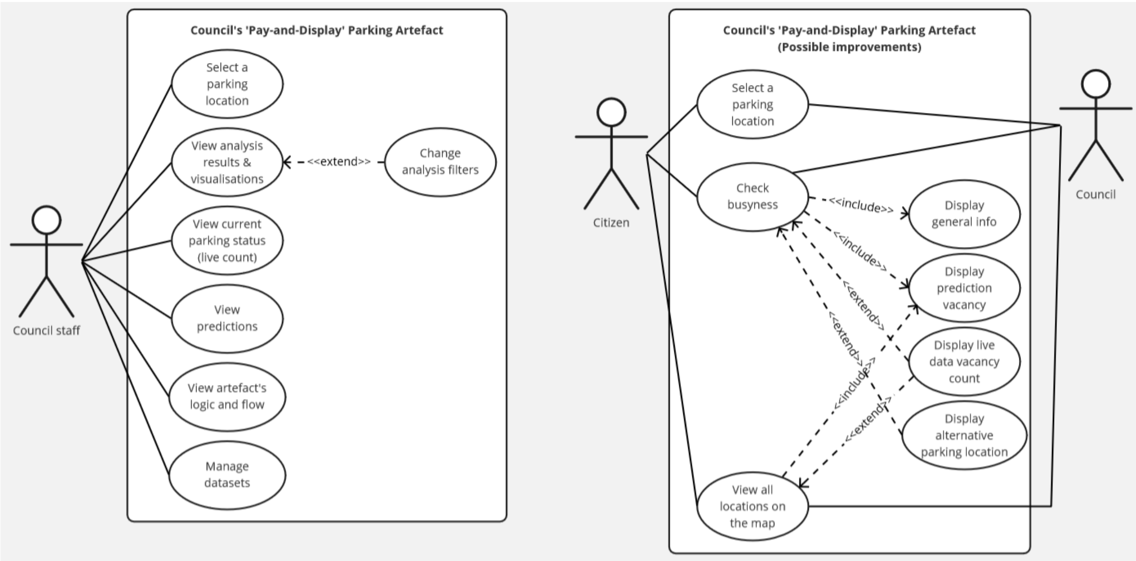


Figure 13: Use case diagrams

4.3.Data Cleansing

The following datasets have had data cleansing applied to them: *weather report, holidays, charging point transactions, pay-and-display ticket machine logs original* and *altered duplicate* (*id_temperature.csv, id_holidays.csv, id_chargingpoint.csv, id_payanddisplay.csv, id_payanddisplay_LOC_A.csv* from folder "*Original_datasets*").

id_temperature.csv:

- Changes were applied to data types.
- Column name changes applied.
- No missing data was found.
- 2 new columns were added to describe the *rainfall* and *wind classification.*

id_holidays.csv:

- Data types changes applied.
- No missing data was found.

id_chargingpoint.csv:

- Data types changes applied.
- Rows with missing data and rows that had the *"total kWh"*=0 were dropped.

id_payanddisplay.csv and *id_payanddisplay_LOC_A.csv:*

- Data types changes applied.
- Sorted values based on "*date*".
- Column name changes applied.
- No missing data was found.
- For the *"Cash paid (GBP)"* column, the *"£"* was removed.
- 3 new columns were added for *"tariff amount (GBP)", "tariff time bracket"* and *"leaving time".*

The file showing all the above changes applied can be found under the name *Data_Preparation_and_Cleansing.ipynb*.

After data preparation and cleansing, the datasets were saved under the names *c_temperature.csv, c_holidays.csv, c_chargingpoint.csv, c_payanddisplay.csv, c_payanddisplay_LOC_A.csv* in the folder *"Datasets_used"*.

Because only one video footage was used for this project, the parking space was split into two areas to represent 2 <u>different</u> locations.

When the split was made, two factors were considered:

(1) Location RR had to have 251 *regular* spaces (mirroring the pay-and-display dataset and location)
(2) The activity level within that area. It was aimed that that area of the video has a medium to high activity.

Therefore, the two locations are defined by the following areas:

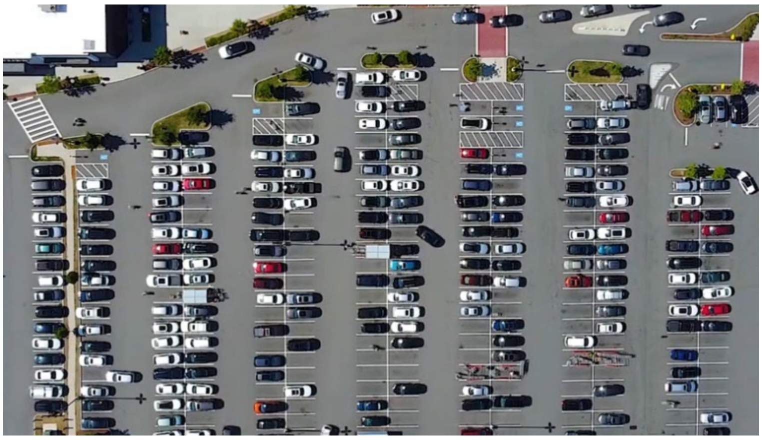


Figure 14: Location RR from the video

Figure 15: Location A from the video

They were treated as two different locations, with different parking counts and they are used in this way to show a proof of concept idea.

4.4.Computer vision

4.4.1.Parking Spaces Positions

The first step in detecting the occupancy is to delimit each parking space for each location.

Taking as example the location RR, which contains vertical parking spaces positions too.

2 files are created to delimit the parking spaces with rectangles:

- one file is delimiting the horizontal spaces (*Loc_RR_Rectangles_regular.py*) and
- the other file is delimiting the vertical spaces (*Loc_RR_Rectangles_regular_v.py*)

Using Paint, it was delimited the width and height needed for the rectangles accordingly to the space type.

For a horizontal rectangle, the *width=35* and *height=16*. These values are opposite for a vertical rectangle.

The code is structured as follows (for *Loc_RR_Rectangles_regular.py*):

- Importing the packages.
- Trying to load an existing *pickle* file that would contain the rectangles' position points or if that is not possible, creating a *positions list*. This was done to enable changes to (some) position points without having to re-add all of them.
- *Width* and *height* of the rectangle are determined.
- A function is created that will check which mouse button (left or right) has been clicked. If the left button is clicked, the position point (*x* and *y*) is added to the position list. If the right button is clicked within *a rectangle* that would be represented by a position point, then the respective position point will be removed from the list. (It uses the *width* and *height* to determine the rectangle and to check if the right click position point is within a possible rectangle.) Finally, it saves the *positions list* in a pickle file, *figure 16*.

```python
# checking which mouse button has been clicked and acting accordingly (add/delete rectangle)
# and dumping the positions in the pickle file
def mouse_click_for_rectangle(events, x, y, flags, params):
    # if the left button is clicked
    if events == cv2.EVENT_LBUTTONDOWN:
        # appending the position of x and y
        position_list.append((x, y))

    # if the right mouse button is clicked
    # check if the position point is in between one of the appended points
    if events == cv2.EVENT_RBUTTONDOWN:
        for i, position in enumerate(position_list):
            x1, y1 = position
            if (x1 <= x <= x1+width and y1 <= y <= y1+height):
                # remove the "i" value from the position list
                position_list.pop(i)

    with open("Loc_RR_Rectangles_regular_pickle", 'wb') as file:
        pickle.dump(position_list, file)
```

Figure 16: Mouse click function

- Lastly, in a *while* loop, the image is imported, the mouse clicks are detected, and the *positions list* is used to place rectangles accordingly to the list's position points. The rectangle's coordinates are given using the position points tuple and adding the *width* and *height* to the position points tuple, *figure 17*.

```python
while True:
    # import the image
    img = cv2.imread("Loc_RR_image.png")
    for position in position_list:
        # put a rectangle on the img
        # rectangle, coordinates () (), colour, thickness
        cv2.rectangle(img, position, (position[0]+width, position[1]+height), (255, 0, 191), 2)
    cv2.imshow("Image", img)
    cv2.setMouseCallback("Image", mouse_click_for_rectangle)
    if cv2.waitKey(1) & 0xFF == ord('b'):
        break
```

Figure 17: While loop

The **result for placing the *horizontal regular spaces* for location RR** looks like this:

Figure 18: Horizontal regular spaces for location RR

The **result for placing the *vertical regular spaces* for location RR** looks like this:

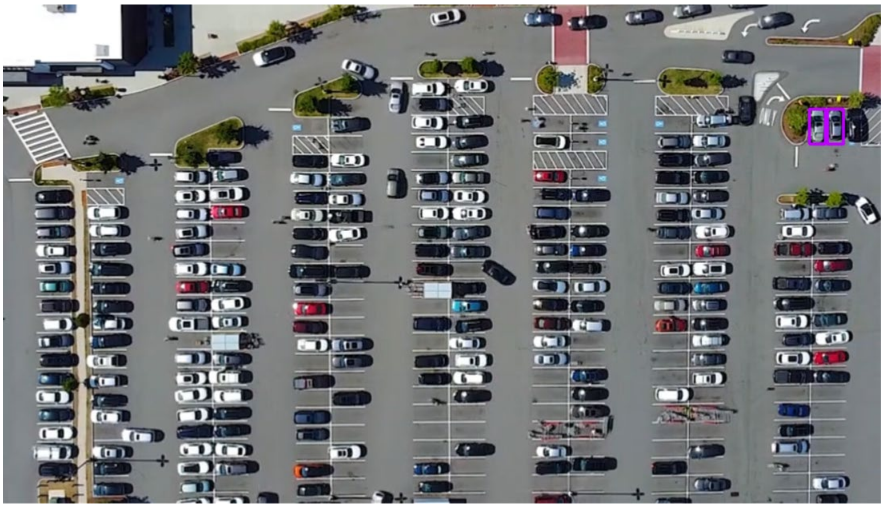


Figure 19: Vertical regular spaces for location RR

The approach is identical for determining the positions of the *disabled bays parking spaces* for location RR: *Loc_RR_Rectangles_disabled.py* (*horizontal spaces*) and *Loc_RR_Rectangles_disabled_v.py* (*vertical spaces*).

The saved positions are stored within:

- *Loc_RR_Rectangles_regular_pickle (horizontal regular spaces)*
- *Loc_RR_Rectangles_regular_v_pickle (vertical regular spaces)*
- *Loc_RR_Rectangles_disabled_pickle (horizontal disabled bays spaces)*
- *Loc_RR_Rectangles_disabled_v_pickle (vertical disabled bays spaces)*

The same approach is taken for location A and the files are named identically, except it will be *"Loc_A_...".*

4.4.2.Occupancy Check

Determining if the parking space is occupied or not.

The overall structure of the *occupancy check* code consists of:

- Importing packages.
- Opening the video file.
- Loading all the *pickle* files that contain the position points.
- Determining the 2 *widths* and 2 *heights* of the rectangles (horizontal and vertical).
- Creating 2 functions (one for *regular* and one for *disabled parking bays spaces*).
- Creating a *while* loop for the video's frames.

Explaining the *while* loop:

- The *while* is happening as long as the video is opened.
- To re-loop the video, an *if statement* is added to check if the current frame is equal to the total number of frames and if true, the current frame is set to zero.
- For each frame of the video, the following is applied:
    - The frame is converted to a ***greyscale*** from BGR (*Blue, Green* and *Red*), *figures 20 & 21.*
    - Then ***cv2.GaussianBlur*** is applied to the image, *figures 20 & 22*.
      This is an **image pre-processing technique** that helps reduce the frame's details (noise) (Rosebrock, 2021). The kernel size was set to *5 x 5* and the standard deviation of the distribution to 1. Those values gave the best results.
    - Next, the frame was converted to a binary image using the ***cv2.adaptiveThreshold***, *figures 20 & 23*.
      To help with possible varying lighting conditions of the frame, adaptive thresholding was applied as it uses a smaller region around it to determine the pixel threshold (OpenCV, 2022). The *adaptiveMethod* used was ***cv2.ADAPTIVE_THRESH_GAUSSIAN_C*** which is *"a gaussian-weighted sum of the neighbourhood values minus the constant C"* (OpenCV, 2022). The *weighted Gaussian mean* is over a 25x25 area with C=18.
    - **Dilation** is applied using ***cv2.dilate*** with a kernel of 3, *figures 20 & 24*. This is applied to increase the white region in the frame (OpenCV, 2016) helping with the non-zero count.

- Lastly, the *dilate* frame was given to the 2 functions.

```
# converting the img to grey scale
grey_img = cv2.cvtColor(frame, cv2.COLOR_BGR2GRAY)
# add some blur
blur_img = cv2.GaussianBlur(grey_img, (5, 5), 1)
# convert to binary img
threshold_img = cv2.adaptiveThreshold(blur_img, 255, cv2.ADAPTIVE_THRESH_GAUSSIAN_C, cv2.THRESH_BINARY_INV, 25, 18)

kernel = np.ones((3, 3), np.uint8)
dilate_img = cv2.dilate(threshold_img, kernel, iterations=1)

checking_parking_space(dilate_img)
checking_parking_space_disabled(dilate_img)

cv2.imshow("Image", frame)
```

Figure 20: Image processing

Figure 21: Greyscale image

Figure 22: Gaussian Blur image

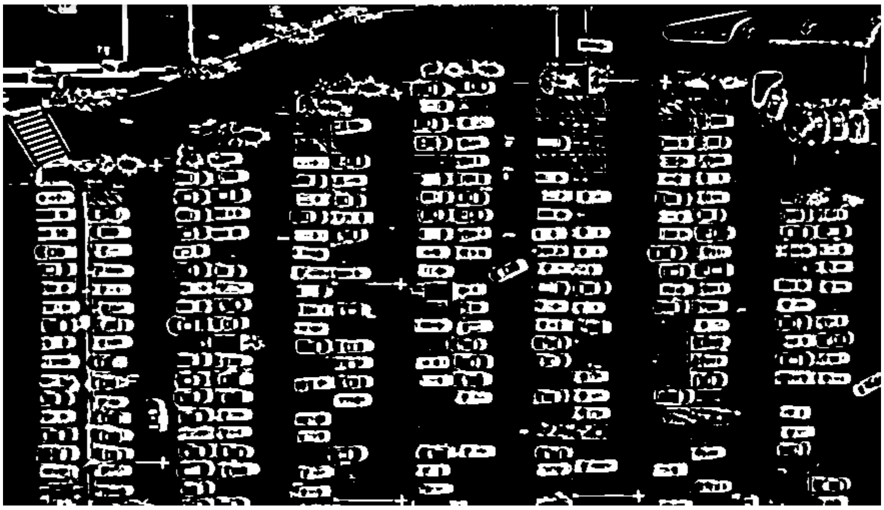


Figure 23: Adaptive Threshold image

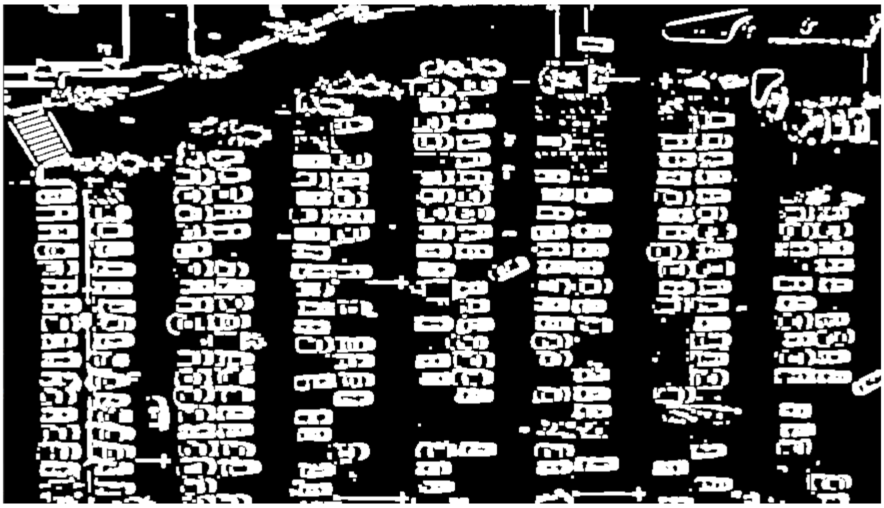


Figure 24: Dilate image

Note: *cv2.medianBlur()* was added to remove some of the noise before dilate, but the overall results were not as good.

Further, taking as an example one of the 2 functions and explaining it (*checking_parking_space(processed_img)*):

- Defining the *space_counter*=0.
- For each position points in the *regular* ***horizontal*** *spaces position list*, the following is applied:
    - Taking the position points in *x, y*.
    - Isolating that parking space using the *x* and *y* and *width* and *height*.
    - Counting the **non-zero pixels** in that isolated/cropped parking space.
    - Using an *if-else statement*, if the count of non-zero pixels is smaller or equal than *165* then the space is vacant and the *space_counter* increments by 1, else the space is occupied. Now a rectangle is placed in that parking position which is coloured and thick accordingly to the space's occupancy/vacancy, *figure 25.*
- Same *for* loop approach is taken for the *regular* ***vertical*** *spaces position list*.
- At the end and outside the loops, a text is added to the frame which will display the *free spaces count* based on the *space_counter*. (Only this function contains the whole legend.)

```
def checking_parking_space(processed_img):
    # for the regular horizontal spaces
    space_counter = 0
    for position in position_list:
        x, y = position
        # cropped rectangles
        cropped_img = processed_img[y:y+height, x:x+width]
        # counting the non-zero pixels
        count = cv2.countNonZero(cropped_img)

        # checking pixels count for confirming the occupancy
        if count<=165:
            # not occupied space
            colour = (30, 250, 30) # green
            thickness = 3
            space_counter = space_counter + 1
        else:
            # occupied space
            colour = (0, 0, 255) # red
            thickness = 2
        cv2.rectangle(frame, position, (position[0] + width, position[1] + height), colour, thickness)
```

Figure 25: Checking the occupancy of regular horizontal spaces

Identical approach is taken for checking the *disabled parking bays spaces (checking_parking_space_disabled(processed_img))*.

Identical approach is taken for location A.

4.4.3.Video Dataset Creation

For each frame, the *space_counter* has been saved within a list. Resulting in 2 lists of *free spaces counts* for *regular and disabled parking bay spaces*.

As example, location RR:

Those 2 lists have been used in *Loc_RR_video_dataset_creation.py* to create 2 *.csv* files:

- *Loc_RR_video_vs_pad_count_regular.csv*
- *Loc_RR_video_count_disabled.csv*

Explaining the *Loc_RR_video_vs_pad_count_regular.csv* file creation:

- To the dataframe containing the pay-and-display *valid tickets count* (*Loc_RR_pad_valid_tickets_no.csv*) the following columns are added *"CCTV count", "Location", "Space type"*.
- To populate the *"CCTV count"* for *regular* spaces, two methods are applied:
  - 1.Instead of randomly populating *"CCTV count"*, the method applied is based on the value of the *"Valid tickets number"*, meaning that **a random value** between *["Valid tickets number"-5, "Valid tickets number"+5]* is generated. That value is added to the *"CCTV count"* only if it is positive (*>=0*) otherwise to the random value +5 is added, *figure 26*.

```
for i in range(0, len(values_tickets), 1):
    p_v_t_n = values_tickets[i] #=pad_valid_tickets_no['Valid tickets number']
    random_spaces = random.randint(p_v_t_n - 5, (p_v_t_n + 5))
    if random_spaces < 0:
        pad_valid_tickets_no["CCTV count"][i] = (random_spaces + 5)
    else:
        pad_valid_tickets_no["CCTV count"][i] = random_spaces
```

Figure 26: Almost randomly populating the "CCTV count" for Location RR, regular bays

  - 2.Starting at the date and time of *9:50 AM, 2013/3/30* and finishing at *10:00 PM, 2013/3/30*, the *CCTV/video count* of occupied spaces is overlapped in the *"CCTV count"* column.

Explaining the *Loc_RR_video_count_disabled.csv* file creation:

- A new dataframe is created with columns of *"Date and time", "Occupied spaces", "Location", "Space type".*

- Starting at the time of *6:30 AM, 2012/12/1* and finishing at *9:40 AM, 2013/3/30* the following is applied:
    - 2 variables are added outside the *while* loop, one to control the randomness (*not_so_random=0*) and one to control the spikes (*var=0*).
    - In the *while* loop, if *var<=3* then the *random_spaces* value does not change and *var=var+1*. **This was done to attenuate the jumps up and down.** Therefore, a jump only happens every 40 minutes.
    - When *var>3*, a new random value is generated based on a $\pm$3 value of the *not_so_random*. If that value is smaller than 0, then the *random_spaces=random_space+3, not_so_random=not_so_random+3* and *var=0*. If that value is not smaller than 0, it goes into another *if-else* statement.
    - That *if-else* statement checks if the *random_spaces* value is higher than 15 (15 being the maximum available spaces), if true, a reset of the *not_so_random* is applied (between 0 and 5) and *var=0*; else, the *random_spaces* value is good and *var=0*.
    - The *figure 27* shows when and what values are added to the *"Occupied spaces"* column.

```
time_plus_10 = datetime(2012, 12, 1, 6, 30, 00)

not_so_random = 0
var = 0 #added to reduce the spikes up and down by limitating the times random is generated (eg every 40 mins)
random_spaces = 0
while time_plus_10 != datetime(2013, 3, 30, 9, 50, 00):
    cctv_RR_data_disabled["Date and time"].append(time_plus_10)
    # this is done to attenuate the high jumps up and down from random
    if var <= 3:
        cctv_RR_data_disabled["Occupied spaces"].append(random_spaces)
        var = var+1
    else:
        random_spaces = random.randint(not_so_random-3, not_so_random+3)
        if random_spaces < 0:
            cctv_RR_data_disabled["Occupied spaces"].append(random_spaces+3)
            random_spaces = random_spaces+3
            not_so_random = random_spaces+3
            var = 0
        else:
            if random_spaces > 15:
                # reset
                not_so_random = random.randint(0, 5)
                cctv_RR_data_disabled["Occupied spaces"].append(not_so_random)
                random_spaces = not_so_random
                var = 0
            else:
                cctv_RR_data_disabled["Occupied spaces"].append(random_spaces)
                not_so_random = random_spaces
                var = 0
    cctv_RR_data_disabled["Location"].append("RR")
    cctv_RR_data_disabled["Space type"].append("Disabled")
    time_plus_10 = time_plus_10 + timedelta(minutes=10)
```

Figure 27: Almost randomly populating the "Occupied spaces" for Location RR, disabled bays

- Lastly, the remaining time until *10:00 PM, 2013/3/30* is populated with the *CCTV/video count* of the total occupied spaces for *disabled parking bays*.

For location A, the same approach is done for creating the *Loc_A_video_vs_pad_count_regular.csv* file as for location RR, but for creating the *disabled bays parking spaces*, a slightly different approach is employed as the total number of spaces is only 2.

There still is a *not_so_random* value that only changes the random value every 40 minutes but the random values vary from 0 to 2 with no reset needed.

4.5.Prediction Model(s)

The chosen model used for making the 4 types of predictions was Long-Short Term Memory (LSTM).

The code structure consists of:

- Importing the packages.
- Reading the dataset.
- Isolating only the needed columns (*date and time* and the *count*).
- *if applicable, altering the dataset to have less frequent data (e.g. every 30 minutes or 1 or 2 hours records).
- Transforming the *count* in array format.
- Getting the size of what would represent 80% of the data.
- Applying normalisation (*MinMaxScaler*).
- Creating a *training dataset* which will contain the first 80% of the data.
- Creating two lists (*x_train* and *y_train*).
- In a *for* loop, starting with the first 60 records, they are added to the *x_train* and the 61st record is added to the *y_train* → this is applied for the whole length of the *train_data*. Basically, the model will get 60 records in *x_train* and the 60+1 record in the *y_train* so it learns how to predict the 61 value.
- Transforming *x_train* and *y_train* to arrays.
- Reshaping *x_train.*
- **Creating and compiling the model** and adding a summary.
- **Fitting the model** and adding *validation_split=0.2* for calculating the loss (80% used for training, 20% used for testing).
- Plotting the training and validation loss on a graph.
- Saving the model.
- Testing the model:
  - Prepare *test_data* (which is the last 60 records of the training data + the remaining 20% of the data).
  - *y_test* list will contain the last 20% of the data.
  - *x_test* list will contain arrays of 60 records until the end of the length of *test_data.*
  - Making *x_test* array and reshaping.
  - Applying prediction to the *x_test* data.
  - Inversing the scaler of the prediction results.

- **Calculating RMSE** (Root Mean Squared Error)**, MAE** (Mean Absolute Error)**, MSE** (Mean Square Error)**, R2-score.**
- Plotting on a graph the training data used, actual vs predicted results.

- **Making the next prediction** using the last 60 records to predict the 61 by following the same steps as presented above.

The same concept as above is applied to each model created and saved. The only difference those models have is the model creation, the optimiser, complier and fitting, *figure 28*.

[Saved]LSTM_Model_Loc_RR_10min_valid_tickets Last Checkpoint: 16/08/2022 (autosaved)

View Insert Cell Kernel Widgets Help

Run Markdown

```
#create the model
model = Sequential()
model.add(LSTM(units=64, input_shape=(x_train.shape[1],1), return_sequences=True))
model.add(LSTM(units=128, return_sequences=True))
model.add(LSTM(units=64, return_sequences=False))
model.add(Dropout(0.1))
model.add(Dense(units=1, activation='linear'))
```

```
optimizer = keras.optimizers.Adam(learning_rate=0.0001)

model.compile(optimizer='adam', loss='mse', metrics=[RootMeanSquaredError()])
model.summary()

history=model.fit(x_train, y_train, batch_size=32, epochs=40, validation_split=0.2, verbose=1)
```

```
Model: "sequential_1"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 lstm_3 (LSTM)               (None, 60, 64)            16896

 lstm_4 (LSTM)               (None, 60, 128)           98816

 lstm_5 (LSTM)               (None, 64)                49408

 dropout_1 (Dropout)         (None, 64)                0

 dense_1 (Dense)             (None, 1)                 65

=================================================================
Total params: 165,185
Trainable params: 165,185
Non-trainable params: 0
```

Figure 28: LSTM model (Location RR, 10 minutes prediction based on valid tickets)

Adam optimiser was chosen as it is a popular and effective algorithm with fast good results (Brownlee, 2017). The default *learning rate* of 0.001 was changed to 0.0001 because it gave better results. The dropout layer was added to prevent overfitting.

In total, 8 LSTM models were created and saved (11.5.Appendix E – LSTM Models). They used the following datasets:

- Every 10 minutes records of *valid tickets* for *location RR*
- Every 10 minutes records of *CCTV/video disabled bay counts* for *location RR*
- Every 30 minutes records of *valid tickets* for *location RR*
- Every 30 minutes records of *CCTV/video disabled bay counts* for *location RR*
- Every 1-hour records of *valid tickets* for *location RR*

- Every 1-hour records of *CCTV/video disabled bay counts* for *location RR*
- Every 2-hour records of *valid tickets* for *location RR*
- Every 10 minutes records of *CCTV/video disabled bay counts* for *location A*

Overall, throughout practice and testing, it was sought to get the best possible results of each LSTM model based on the given dataset. Experimentations can be found in 11.6.Appendix F – LSTM Models Experimentation.

Initially, two other different models were tested (*Prophet* and *linear regression*) but the results were very unsatisfactory (11.7.Appendix G – Other Models Testing).

4.6.Overall Artefact Implementation

The file that contains the Streamlit implementation can be found under the name ***artefact.py***.

The file structure is as follows:

- Importing packages.
- Page configuration and colours set up.
- Creating the sidebar and hiding the first radio-button selection.
- *Instructions* function.
- *Welcome page* info.
- Reading the datasets and combining datasets.
- Multiple functions (with *st.cache* which is a memoisation technique which stores the results locally without repeating the computation unless necessarily (Streamlit, 2022)).
- Council staff → all menu options.
- System expert → all menu options.
- Citizen → all menu options.
- *Artefact instructions* which calls again the *instructions* function.

The file is programmed in Python, but to display something, Streamlit must be used. For example, *"print()"* is replaced by *st.write()* or *st.markdown()*. For *title* or *subtitle* it is used *st.title(), st.subheader()*; to display graphs is *st. plotly_chart(fig), st.pyplot()*; to display an image is *st.image()*; to display a caption is *st.caption()*; to split a page split in columns is *st.columns()*. There are also other available elements such as *st.metric(), st.button(), (with) st.expander, st.empty()*, etc. .

Few snippets of the artefact are presented below.

The whole artefact screenshots (or video demo link) can be found in 11.8.Appendix H – Artefact.

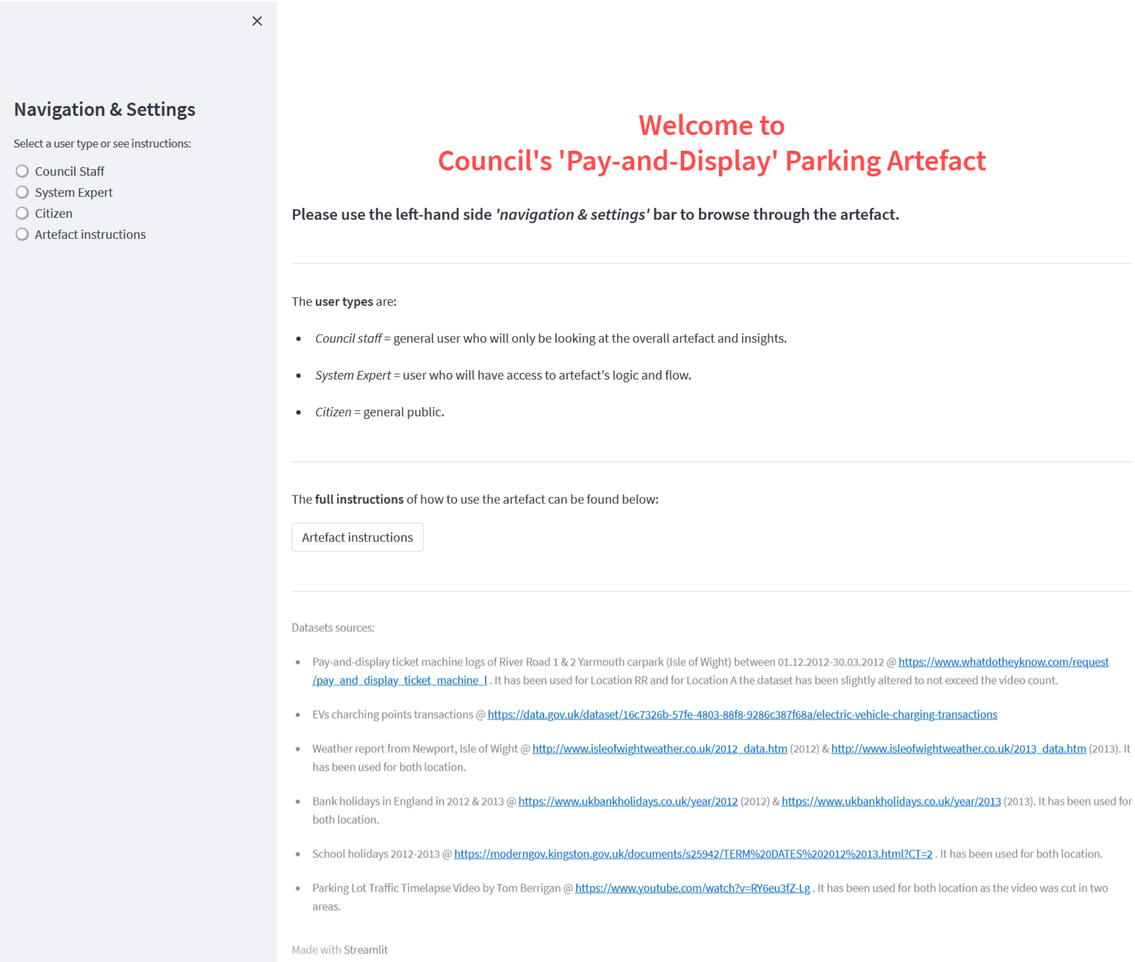


Figure 29: Welcome page

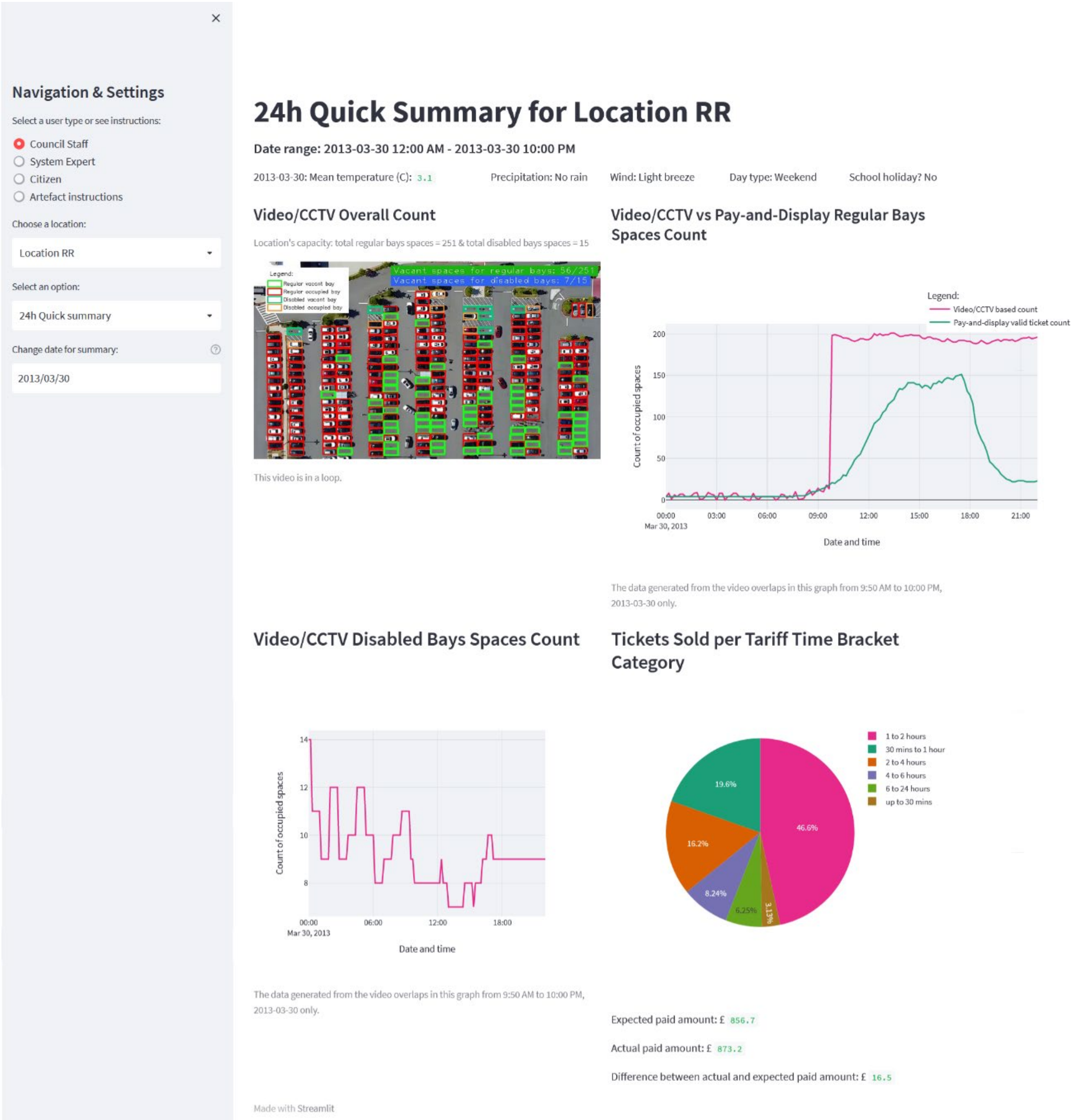


Figure 30: Council Staff → Location RR → 24h Quick summary

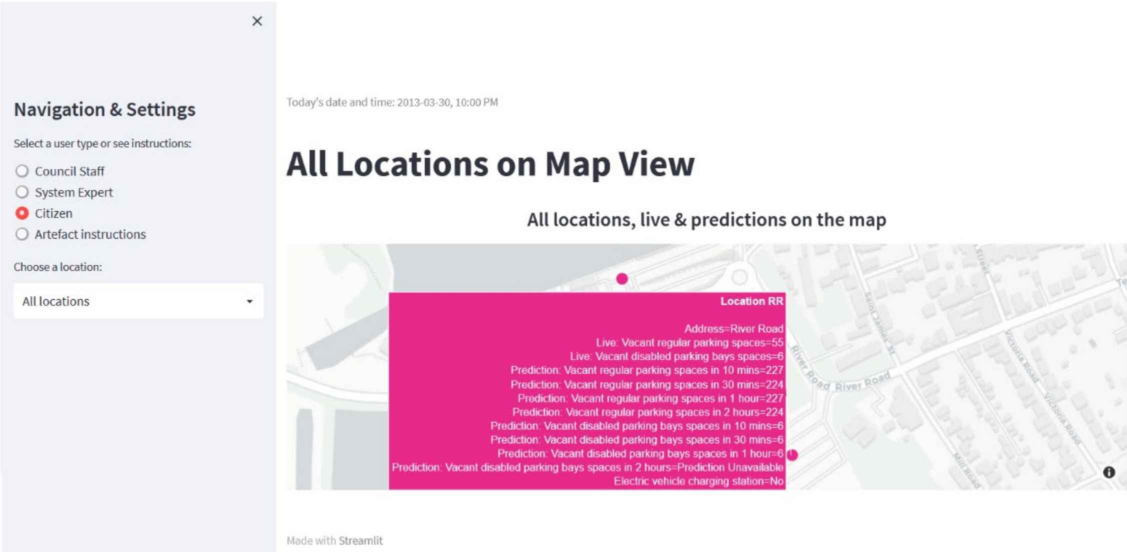


Figure 31: Citizen → All locations

Lastly, multiple resources have been used/accessed to help in creating the overall artefact. They include Streamlit Library (2022), Plotly Graphing Library (2022), OpenCV Documentation (2022), Keras Documentation (2022), Hassan (2021).

4.7. Challenges & Solutions

Throughout this project, there were multiple challenges encountered. Mentioning some of the most prolific ones and their found/applied solution:

- **Training the LSTM models** to get the best results. The solution applied to accomplish this involved a lot of training time, testing, research and practice in the model's creation (11.6.Appendix F – LSTM Models Experimentation).
- Some **programming challenges** involving computer vision and data analysis tasks were encountered. To solve them, research, trial and error and documentation checks were made.
- **Implementing everything using Streamlit** served some challenges (making the *welcome page* disappear when the *"instructions"* button was clicked; having no radio-option selected when displaying the *welcome page*; correctly displaying graphs) but they were fixed with the use of documentation, practice/testing or using a different visualisation library.
- **Gathering the video footage and data**. Initially, only non-fixed view-down footages were found and after a long time of searching, the chosen YouTube video was found. The *pay-and-display machine logs* were found after numerous searches of those terms. Unfortunately, the project was aiming at using more than one such dataset and because it could not be found, a partial duplicate was created to prove as a proof of concept.
- **Finding personnel that works for the council or manages car parks** and is willing to conduct the usability testing for the artefact. *Southampton City Council, Isle of Wight Council* and *Solent University's Estate and (Parking) Facilities* have been contacted and still, no reply was received (11.9.Appendix I – Emails). Unfortunately, this challenge could not of being solved in the given time and future work would involve a more in-person interaction.

4.7.1. Personal Reflection

This project has broadened my understanding and knowledge in the areas of computer vision, predictive analysis and web-app presentation using Streamlit. Overall, this project has been pressuring and demanding and learning when to stop pursuing a task has been crucial in effectively managing the time. The personal growth gained will be useful in the future career encounters.

5.Evaluation & Results

5.1.Data Analysis

The cleansed datasets and the other datasets created based on those datasets were analysed.

An in-depth view and explanation of each graph presented within the artefact can be found in 11.10.Appendix J – Graphs. The overall analysis is based on a summary of those graphs.

The overall insights are:

- **Location RR** – *the true/original dataset*, period 01.12.2012–30.03.2013:
  - Overall, the most preferred *tariff time bracket category* is *"1 to 2 hours"* leading with 2036 transactions, followed by *"30 minutes to 1 hour", "2 to 4 hours"* and in the 4th place having the *"6 to 24 hours"* with 1027 transactions.
  - Moreover, the total collected pay was £18385.8 which is £309.2 more than expected.
  - One of the pay-and-display ticket machines was more used overall than the other, with 238 transactions more. Both machines had an almost identical numbers of transactions for *"1 to 2 hours"*.
  - For the *valid ticket count*, a clear pattern can be seen, during the day the count oscillates but during the night it almost always stays identical.
  - During those 4 months, there are some peaks for occupancy with values equal to or above 72 (maximum 151) and it was noticed that all those days had *no rain*, were *weekends/bank holidays* and had a *low average temperature*. Also, it shows that the parking location was never fully occupied (full capacity is 251).
  - It was noted that for *"no rain"* the *daily average valid ticket count* data was more spread out than for *light, moderate and heavy rain*. The same was noted for *CCTV/video daily average count*.
  - *"Light air"* and *"light breeze"* saw the *daily average valid ticket count* data more spread out with *"gentle breeze"* tighter together. The same was noted for *CCTV/video daily average count*.
  - When analysing the correlation between the whole 4 months' *daily average valid ticket count* and the weather information, a significant correlation could not be seen. But by analysing specific date ranges, such as the month of *January*, there is a moderate positive correlation (0.56, 0.57, 0.55) between *mean, high and low temperature* and the *daily average valid ticket* or *CCTV/video count.* Also, a weak negative correlation can be seen between the *daily average counts* and *rain, high* and *average wind speed (mph)* (under -0.24). For *February*, a moderate negative

correlation is seen between *the daily average counts* and *low temperature*. For *March*, the overall correlation is very weak (under $\pm$0.16). For *December* there is a weak negative correlation between the *daily average counts* and the whole *weather data*.

- The biggest *sum of the daily average valid ticket count* was *Saturday* and *Sunday* followed by *Thursday* and *Wednesday*.
- Even though the *weekdays* are leading in the *sum of daily average count for valid tickets* with 931, the *weekends* are relatively close with 597.
- No significant correlation was found between the *daily average counts* and *school holidays (yes/no)* using point biserial correlation coefficient.
- *[hypothetically]* When comparing the *CCTV/video occupancy count* with the *valid ticket count*, discrepancies can be seen, this could be due to multiple reasons.
- *[hypothetically]* The *disabled bays CCTV/video occupancy count* would show how popular it is and when is the most used.

➢ **Comparing location RR with location A***

*Location A is almost based on an identical dataset with Location RR

- The boxplots of both locations are right-skewed for the cash paid amount. For location RR, *Q1-median* has the most concentrated data. For location A, *min-Q1* has the most concentrated data and both locations have *Q3-max* the most spread-out data. For location RR, 50% of the data (*Q1-Q3*) is between £1.9 and £4.5, and for location A, 50% of the data is between £1 and £4.5.
- Both locations have the most transactions for *"1 to 2 hours"* with the 4^{th} place being *"6 to 24 hours"* and a similar overall transaction number.
- When calculating the mean per weekday for each month for *valid tickets*. Both locations have strong similarities for December, January and March, and only partial similarities for February. Both locations seem to either have the biggest and second biggest mean during the weekend, mid-week (Wednesday-Thursday) or mid-week and Saturday.
- *[hypothetically]* When calculating the mean per weekday for each month for *CCTV/video regular parking spaces count*. Both locations have strong similarities for December-January and only partial similarities for February-March.
- *[hypothetically]* Comparing the *valid ticket* and *CCTV/video count* calculated as a *mean per weekday per each month* for *locations RR* and *A*, the similarities are identical for when the biggest and second biggest mean is.

- *[hypothetically]* For location RR, the highest mean per weekday per month for *CCTV/video disabled parking bays spaces count* varies from 7.5 to 8 across the 4 months and for location A the highest mean per month is 1.1.
- The highest occupancy percentage per month based on *valid tickets* for location RR was in February and the lowest was in January for *regular parking spaces*, for *disabled bays* was highest in January-February and lowest in March. For location A, the patterns were almost identical with location RR, but for *disabled bays*, the lowest occupancy was in December and March.

➢ **Electric vehicles charging points transactions** dataset, period 31.03.2021–01.07.2021:

- The boxplot for the total number of kWh is right-skewed with 50% of the data (*Q1-Q3*) between 7.9 to 23.4. *Q1-median* (or 25% of the data) is the most concentrated and *Q3-max* (or 25% of the data) is the most spread-out.
- Few of the charge points are more preferred over the others. Some charging points have over 30 transactions overall, and others have 1 transaction each.
- The *total time (in minutes)* for charging is spread-out over a smaller data range compared to the *total time (in minutes)* when the vehicle is plugged in. Both boxplots are right-skewed with 50% of the data for plug/unplug time being between 170 to 707.5 and with 50% of the data for charging start/end being between 146 to 466.5. The outliers are more spread-out for plug/unplug times than for charging start/end.
- The difference in minutes between plug/unplug and charging start/end is approximately 100 minutes for April and May and 135 minutes for June.
- The most popular hours for unplugging the vehicles are 4-7AM and 7-11PM and the most popular hours for plugging in the vehicles are 2-6PM. 8AM-1PM have relatively similar plugging/unplugging events. 12-3AM are very unpopular with almost no transactions.
- The most popular days for unplugging the vehicles are the 2nd, 18th, 23rd, 26th, 27th and 30th and the most popular days for plugging in the vehicles are the 11th, 17th, 23rd, 26th, 27th and 30th. Making the 23rd, 26th, 27th and 30th the busiest days. The least popular days for unplugging the vehicles are 3rd, 10th, 24th, 31st and for plugging-in the vehicles are 6th, 10th, 12th, 16th and 31st. Making the 10th and 31st the least popular days. Moreover, overall, for example, if the number of plugged-in vehicles is bigger one day, after one to a few days the number of unplugged vehicles will increase, to balance it out.

- Overall, in the whole dataset, few of the charging points IDs have transactions almost every hour whereas some have a very limited number of transactions.
- Overall, in the whole dataset, in June, there are about 4 charging points IDs that had at least a transaction a day almost every day of the month. (For June there were only 3 such charging points IDs, for May there were only 2 and for April there were only 1.)

5.2. Usability Testing

5.2.1. Overall Tasks Analysis

This section is presenting the finding captured based on the completion time and the number of clicks applied until the task was finished, as well as the participant's subjective perspective of the hardest task. All the tasks have been completed by each participant.

Figure 32 shows each participant's task completion time in seconds. By looking per task*, tasks 2* and *3* were completed by 4 out of 5 participants at an approximate close time to each other. *Tasks 5* and *6* were completed by all participants at an approximate close time to each other. *Tasks 1* and *4* had a wider range of completion. *Participant 1* has completed both tasks in a short time.

For *task 1*, *participants 2* and *4* were the closest to each other and *participants 3* and *5* were the ones that took the longest time.

For *task 4*, *participant 5* was the second to finish the task, then *participants 2* and *4* were almost at the same time (150 seconds) while *participant 3* took the longest.

Overall, *participant 1* completed the tasks very quickly whilst *participant 3* took the longest.

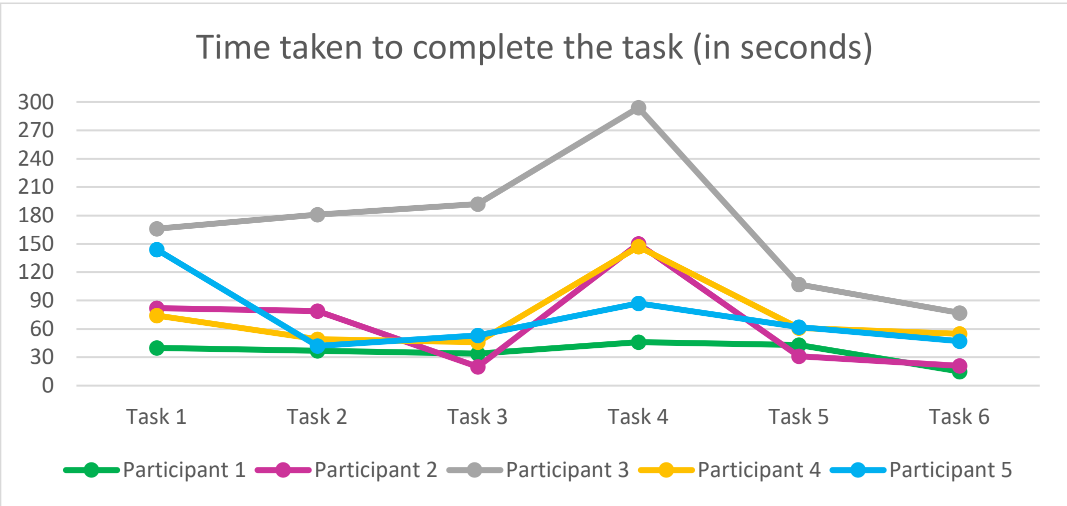


Figure 32: Time taken to complete the task (in seconds)

From *figure 33*, *participant 1* was the closest to the benchmark (with a difference of 13 seconds). 2nd place was *participant 2*, 3rd place were *participants 4* and *5* as they have almost identical times and lastly, *participant 3* took about 7 times more than the benchmark. *Participant 3* explored the web-app while doing the tasks.

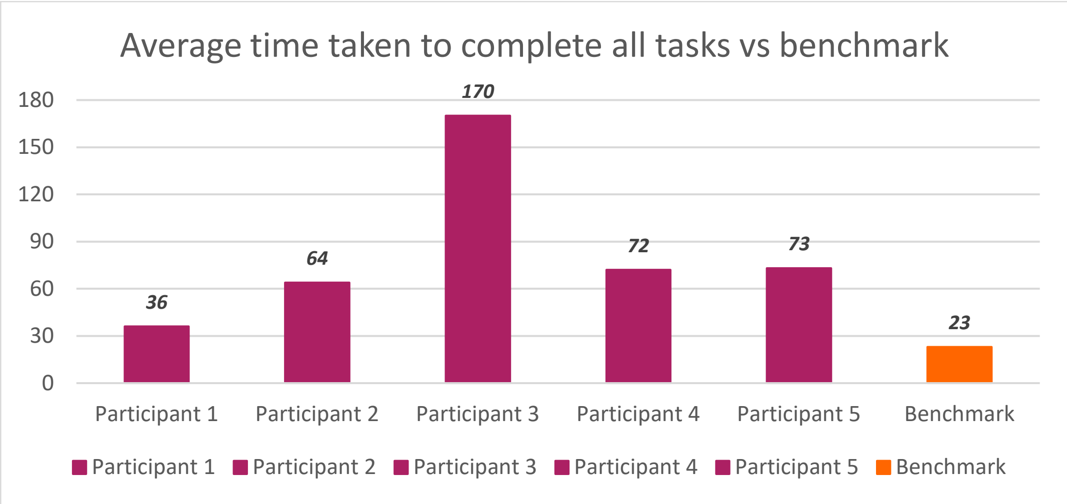


Figure 33: Average time taken to complete all tasks vs benchmark

Figure 34 shows that *task 6* was the closest to the benchmark while *task 4* was the most distant. The other tasks have a significant difference.

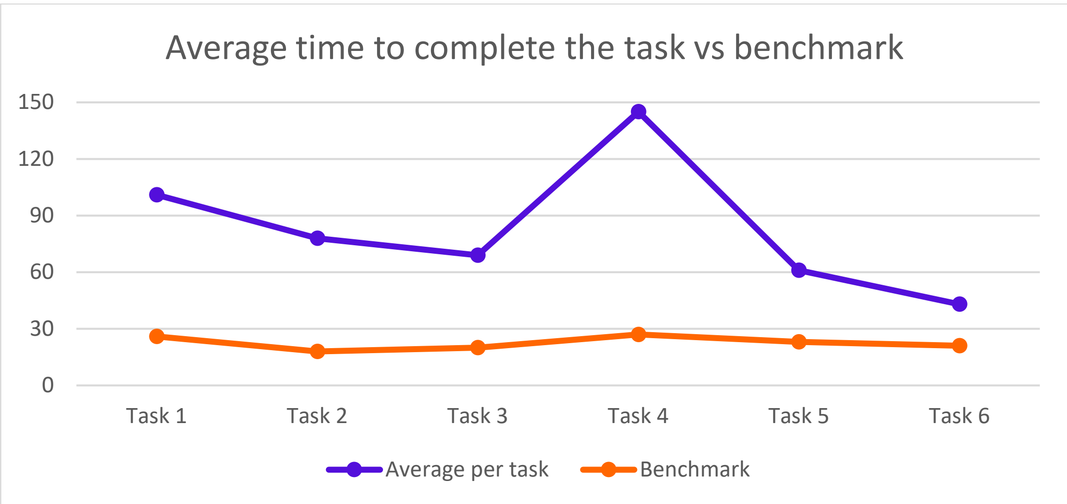


Figure 34: Average time to complete the task vs benchmark

Further, it is important to mention that those results may be not a true representation as Nielsen (2001) is suggesting that having only 5 participants for the time of completion for each task is not enough as it does not give a *“reasonably tight confidence interval on the results”*.

By looking at the number of clicks made until task completion, a more quantitative information can be retrieved, *figure 35*.

As *tasks 5* and *6* took similar times to complete, the number of clicks is identical for *task 5* and only *participant 4* made 4 clicks instead of 2.

Tasks 1, 2 and *3* were set as having a similar – to a higher range of times taken to complete, when looking at the number of clicks, *task 1* was completed between 5-7 clicks, *task 2* between 2-6 clicks and *task 3* was completed by everybody within 2 clicks.

Only *task 4* shows a correlation between the board number of clicks and the highest variety of time taken to complete. Even though *participant 3* took the longest, *participant 2* was the one with the most clicks.

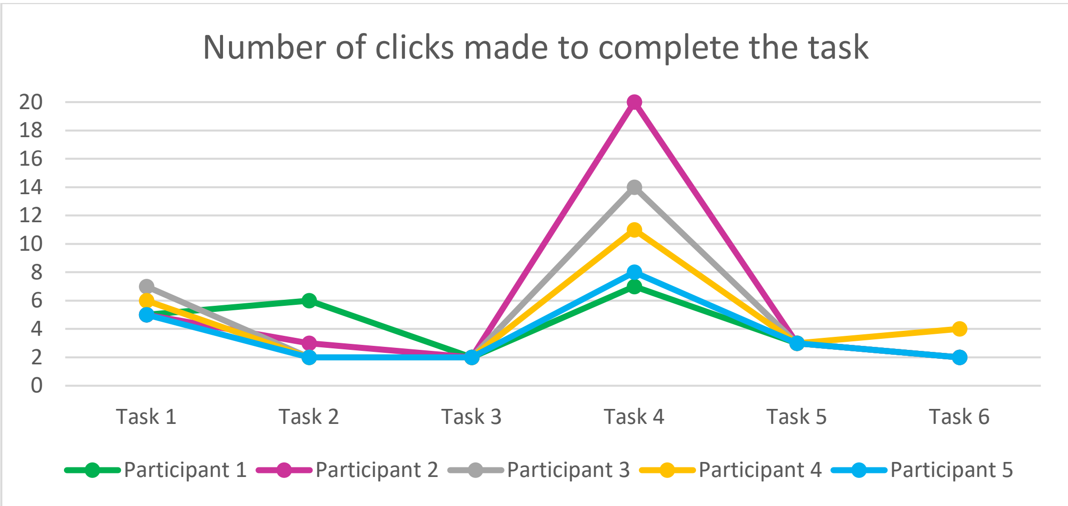


Figure 35: Number of clicks made to complete the task

Participant 5 was the closest to the benchmark for the average number of clicks made to complete a task, *figure 36*. Second place was *participant 1*, then *participant 4,* then *participant 3* and lastly, *participant 2*.

This shows that even though *participant 1* finished the tasks quickest, *participant 5* was the one with the least number of clicks. *Participant 2* who finished the tasks second achieved the highest number of clicks. *Participant 3* with the longest time to complete is the second-to-last for the average number of clicks made.

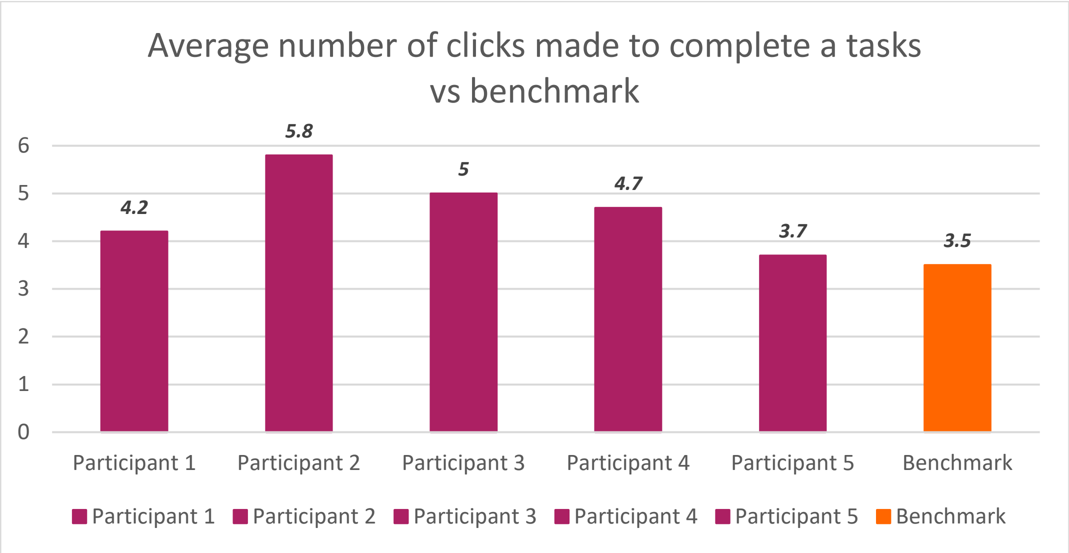


Figure 36: Average number of clicks taken to complete a tasks vs benchmark

Figure 37 shows that the average number of clicks vs the benchmark is identical for *tasks 3, 5* and *6* and with one click difference for *tasks 1* and *2*. *Task 4* has the highest difference (5 clicks). Even though the time required to complete the tasks vs the benchmark is significant overall, when looking at the number of clicks, the only correlation seems to be for *task 4*.

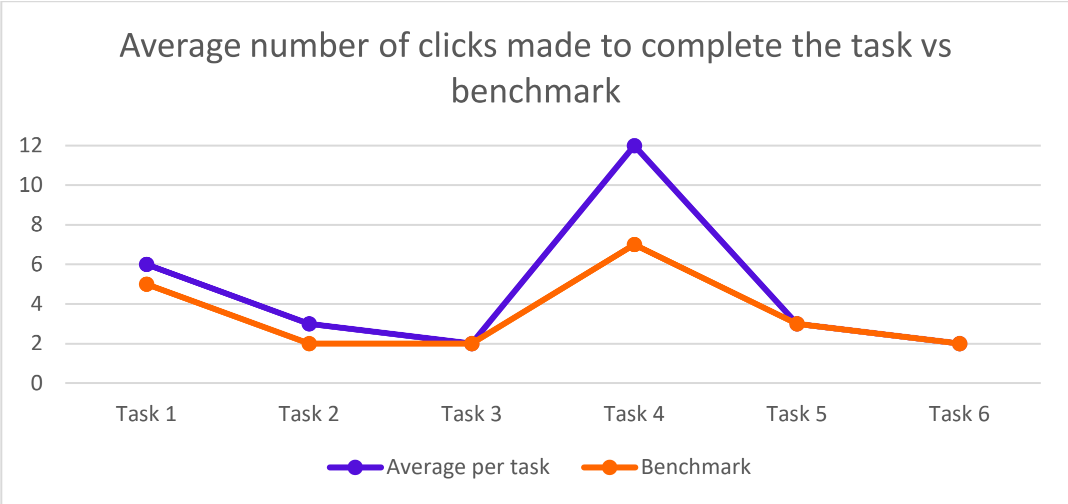


Figure 37: Average number of clicks made to complete the task vs benchmark

The participants have been asked which task they found the hardest. The answers show that *tasks 1* (*participants 3* & *5*) and *2* (*participants 1 & 2*) have been deemed equally the hardest, followed by *task 4* (*participant 4*).

For *participant 3*, *task 1* has been completed 3rd the quickest and with the second most clicks.

For *participant 5, task 1* has been completed the slowest, and with the same number of clicks as the benchmark. They mentioned that they chose this task as the hardest as it is a first-time introduction to the system.

For *participant 1, task 2* has been completed 3rd the quickest with 3 times the benchmark's number of clicks.

For *participant 2, task 2* has been completed 4th the quickest with one click more than the benchmark.

For *participant 4, task 4* took the longest to complete and with the greatest number of clicks.

5.2.2.Participants’ Feedback (per participant)

5.2.2.1.Participant 1

The test was conducted face-to-face.

The participant has found the tasks overall simple and easy to do.

For the **first task**, a bit of confusion was raised over the double headings of the graphs.

The **second task** saw the participant being curious about the graphs’ interaction functionalities which were considered *“really smart”*.

The page of the **third task** was possibly considered having too much information but this would only apply from a citizen perspective and not from a council staff perspective, as it would be a good amount of data and representation. The participant mentioned that they have not seen some of the graph types before.

The error message and the *quick summary* page’s structure of the **fourth task** were considered clear enough. It was mentioned that it is interesting to have the weather and holiday information too.

The information presented on the **fifth task** page was *“really clear and useful”*. The *“popular days”* data would help them better plan their week as they would avoid days that have a high occupancy percentage within a location and having the current day highlighted is good. They asked what is the *popular day* graph based on. It was mentioned that having the predictions of different locations linked with Google Maps *“would be amazing”*.

The **last task**’s page would be the main page the participant would prefer to use, and they would prefer it even more if it were to be integrated with Google Maps. Getting all the essential information needed while hovering over a location *“is perfect”* as there is no information overload due to only displaying the information when a location is selected.

The **best three features** mentioned included:

- The map view.
- The hover-over feature to get all the information needed on the map view.
- The live camera recording of the locations with the occupancy check.

The **improvements suggested** included:

- Swapping the last two titles of the citizens' *individual locations* pages as *tariff and capacity* would be more important than *popular days.*
- Adding a Google Maps link of the address for each location.
- Separating each section of the *live and prediction* title with light faded lines.
- Possibly making the icons bigger in the *live and prediction* section.
- Integrating the *live and prediction vacancy* with Google Maps.

Having such information as presented within the *citizen* user type would be *"very useful and really good"* and a great opportunity to integrate it with other platforms. They would use *all locations on the map* page, and they would *"especially use it if it was integrated with Google Maps".*

Moreover, it was mentioned that the council could better focus funding on more popular locations by having such information analysed.

5.2.2.2.Participant 2

The test was conducted face-to-face.

The participant found some of the tasks confusing and only gave the correct answer after a few guesses.

For the **first task**, they were not sure what they were looking for as it seemed that there is a lot of data. The correct answer was given after the third guess. They have mentioned that the information displayed on the page could be useful for someone familiar with the data, but for a citizen, it will not. They have suggested changing the *count per tariff category* graph to a pie chart.

The **second task** found the participant confused over what a *tariff time bracket category* is. They gave the correct answer after a second guess, as initially, they were only focusing on *the total number of transactions* that are not split by *tariff time bracket category*.

The **third task** was completed easily and quickly, as they were now more focused on finding the graph's name on the page. Regarding the structure of the page, they assumed that the data is displayed in a priority hierarchical way to help the daily council staff user find the most needed data quickly.

The **fourth task** saw the participant very confused about where to find the needed page. They visited the *predictions* page, the *citizens*' options, the *system expert* options. Firstly, they completed the second part of the task (the error message) and then they went back to find the answer to the first part of the task. Initially, they were looking for the total number of parking spaces (giving the answer as 71) but with a further confirmation question, they realised what the correct answer was. This confusion has happened due to misunderstanding the question. They have also mentioned that they were not able to find the needed page easily as they were initially looking for *change date for summary* and not *quick 24h summary*. Further mentioning that they remember seeing the video but not how to get to it.

The **fifth task** was completed easily. They would add the *tariff and capacity* section parallel to the *live and prediction* section. Mentioned that they could possibly gauge the popularity based on the prediction. Moreover, it will be good to have other location suggestions for *regular spaces* only if the prediction is below 5 or 10. Lastly, the predictions should be based on the location's length stay. If it is a long stay location (the type where you would park when you go on holiday), 2-hours in advance may not be enough as it will be the case for the locations presented in this report.

They have found the map of the **last task** displaying *"a lot"* of information, suggesting the removal of predictions for 10 minutes and 1-hour as it feels like *"repeating almost the same data"*.

It was found that the participant understood by *"unavailable"* that there will be no available parking spaces at that time and not that the prediction is unavailable. The suggestion was to change it to *"prediction unavailable"* or to take them out.

The **best features** were:

- Predictions.
- Individual pages for each location as it presents all the data.
- A lot of information that is presented in a readable way for council staff and citizens.
- Easy to understand what the graphs are about.

Improvements suggestions:

- Adding the *citizen's* current location within the map and displaying the nearby available parking locations.
- Different marker colours on the map for locations that are fully occupied vs locations that have vacant spaces.
- Possibly adding another drop-down option for the *citizens' individual location* page with the 3 subheadings options.
- Re-order the subheadings of the *citizens individual location* page.

When asked their thoughts regarding the information presented under the *citizen* user type, they mentioned that the council should properly advertise such system so that the citizens know of its existence. Also, such a system would be useful for anyone who owns and maintains a car park.

5.2.2.3.Participant 3

The test was conducted online. Due to a technical error, the participant's voice has not been recorded. The issue was noticed only after the meeting finished. Therefore, the researcher has re-listened to the recording and tried to write down everything the participant has mentioned/suggested.

The participant found the **first task** quite overwhelming due to the amount of information and animation displayed at once. Concerns were raised over the *section title* vs the *graph title* – suggestions were made in terms of having the graph's title bigger or more visible.

For the **second task**, the participant had a look through the page and to find the graph needed they used *ctrl-F*. This is due to the high amount of information displayed with which they were not familiar. It was mentioned that if they would interact with the system daily, they would probably know better where to locate the needed information. Suggestions for improvement were made towards the headings and the graphs display sections.

The **third task**'s graph was found with the help of *ctrl-F*. It was mentioned that they expected a different graph representation, meaning that they were expecting to see only the *plug-in* time and not the *unplug time* alongside it. The participant seemed to understand the meaning of the graph when asked. This tasked seemed easier to do as they became more familiar with the system.

For the **fourth task**, the participant requested some help to find the needed page, very light instructions were given. The *predictions* page was also accessed. Once the correct page loaded, they knew that the answer is in the video, but a bit of confusion was seen in terms of knowing where to look to find the total number of *disabled bay spaces*. They needed some time to understand the video's legend. A suggestion given was to change the colours for *disabled parking bays* into orange for occupied space. When asked to change the *date for summary*, they could not find the option easily and had to use *ctrl-F*. The error message was considered clear.

Task five was achieved with the use of *ctrl-F*. They have mentioned that the predictions sections are very close to each other and colour improvements should be added to them. They would reorder the page as follows: *tariffs and capacity*, *live and prediction vacancy* and *popular days*.

An interesting indication was that for *disabled parking bays*, other location suggestions (with more free spaces) may not be appropriate as the driver may need to park very closely to the desired location.

Moreover, other suggestions made included:

- When displaying the alternative location also include if the pay is similar.
- A savings deal that would suggest the driver if a different tariff time bracket were more money efficient than another.
- If the driver is about to buy a pay-and-display parking ticket close to the time when the tariffs are not applicable anymore, they could be informed that they would only need to buy for a shorter amount of time instead the time they intend to stay.

Having the current day highlighted was considered good.

The **last task** was completed easily. The main suggestion was to change the display in the hover-over and transform the *predictions* information into a table format. Possibly add icons too.

Lastly, being able to link those parking locations to Google Maps or even other shops to build a route would be interesting.

The **best features** were the live count and the capacity and the map view.

The **improvements** included headings structure and the display of the *prediction* intervals.

When asked their thoughts regarding the council providing such information to its citizens (as seen under the *citizen* user type), it was mentioned that the council should have this information available for its citizens.

5.2.2.4.Participant 4

The test was conducted online.

Overall, after the first task when they understood how to use the navigation bar, they have done the tasks easily (except the 4th one).

For the **first task**, they liked the iteration of actions in the navigation bar, making it easy to use. They liked the structure of this page and that the graphs have hover-over information and can be expanded.

For the **second task**, they find the graphs as *"being labelled correctly"*. They suggested that at the beginning of the page to have an information section that lets the user know that the graphs have interaction functionalities. Overall, they *really liked the colouring* of the page as it made it easy to follow.

Even though they completed the **third task** quickly and found the graph needed *"very clear, I like the colours and everything is well described, very easy to read the data and the legend is there"*, they suggested having a search functionality for graphs to make it easier to find what the user may be looking for.

The **fourth task** saw them unsure what the video meant. An explanation was given suggesting that is similar to the video they have seen in task 1. Therefore, the participant redone the navigation steps for the first task and then changed to location A. To answer the first question of the task, they enlarged the video and manually counted the *disabled bay parking spaces* from the video. Then, they easily completed the second part of the task and found the error message clear. Only in the follow-up questions, they did saw *the total count of the disabled bay parking spaces*. They like the colours chosen for marking the parking spaces from the video and they suggested having the legend on separate lines as -as it is now- it requires a few seconds to understand it.

For the **fifth task**, they had no issues reading the information from the graphs as it is *"super clear and well labelled"*. They mentioned that the prediction section is clearly defined and having the current day highlighted is *"good/nice to have"*. Their suggestions were: moving more up on the page the *charging station* availability, moving the *tariffs and capacity* before the *popular days* section and having the suggested location clickable (taking the user to that location's page). The "*unavailable*" issue was raised here too.

Lastly, they consider that having the prediction for *"2 hours is fine, enough"* as the *popular days* estimates could help the user with the long planning.

For the **last task**, they initially went to the location RR page to find out if there is *charging station* availability. After re-reading the task, they redirected to the correct page. They found the overall page useful as the user may not always know an address by heart and it is easier to see the distances between locations. Overall, after contemplating a bit, they found the amount of hover-over information good ("*it could stay like this*").

The **best features** were the prediction ranges, the map and the live count.

The **improvements** were to have more visible instructions regarding the graphs' functionalities, graph searches (or graphs lists at the top of the page) and "*unavailable*" text changes.

Lastly, they wish the council would provide them with such information (as seen under the *citizen* user type) as it will be good and useful, saving them time knowing if a location is busy or not.

5.2.2.5.Participant 5

The test was conducted online.

Overall, the participant completed the tasks easily.

The **first task** was a bit more challenging due to the introduction to the system, but once they figure out how it worked, they had no other problems navigating. They mentioned that the navigation menu was *"really simple to use"* and they liked the iterative process of displaying information. They took a bit of time to explore the page to find the needed graph. A suggestion given to the task's graph was to change the order of the *tariff category* based on the smallest to biggest tariff category and not displaying them as biggest to smallest count. They would prefer it to stay as a bar chart and not be changed to a pie chart.

The **second task** was considered descriptive and clear. They liked the *tariff time bracket category* order of the graph. They suggested that overall, maybe having specific colours for specific legend types (such as *"location RR"* pink, *"location A"* green, etc.). Therefore, if the user sees those colours, they know the graph is about legend type X without having to check the legend.

For the **third task**, they liked the divisions between bars, hover-over information, and the ability to select/zoom in a graph's part. Suggested a note near the graph which will let the user know of those graph interaction functionalities.

For the **fourth task**, the user remembered where they have seen the video before (task 1) and assumed that each page is consistent. Therefore, they easily accessed the correct page without help. They firstly answered the second part of the task and then, they were asked to answer the first part of the task. They noticed straightaway the total number of *disabled bay parking spaces*. They liked the fact that the *date for summary* can also be changed using the keyboard. Lastly, the suggestions given involved changing the colour of the *occupied disabled bay parking* to orange, adding the legend to the side of the video and splitting it into 4 lines and adding the capacity of the location under the weather information for a quicker glimpse at the parking location.

The second part of the **fifth task** was done first. They liked the page's order, the live info and the prediction. Mentioned that having popular days is *"really useful"* as would give an idea of when to go to those locations. Moreover, having other parking suggestions *"is perfect"* as it will save time searching around. They would like to have this type of page even if it did not contain predictions but only live counts.

Mentioned to change the order of the *tariffs brackets* and to put *"6 to 24 hours"* last. Lastly, when asked if 2 hours of prediction is enough their answer was *"it would be alright"* as they assume more people would look at a location closer to the time of parking.

For the **last task**, they liked the page and that it shows all locations on the map. They would remove the 1- and 2-hours predictions and they would add the location's name under the *marker* (if the name is clicked, to take them to the individual location's page).

The **best features** were the predictions, live parking video (also useful for security measures) and the citizen's individual location page.

To improve was the consistency in labelling and colours, reducing the hover-over information of the map and adding the location's name on the map under the *marker*.

Their thoughts regarding having such information as presented under the *citizen* user type were: *"having something like this would be absolutely amazing for the council to have, not just for local citizens but also for people that come in from outside the town".*

5.3.Improvements of the Artefact based on Feedback (summary)

These suggestions **have been implemented** in the artefact.

"24h Quick summary" page:

- Remove the double headings for the section and graph.
- Changing the *count per tariff category* graph to a pie chart.
- Changing the legend for location A's video (making it on 4 lines).
- Changing the colour of the occupied *disabled bay parking space* to orange.
- Adding the location space capacity information.

Individual location page for *citizen* user type:

- Add the date range for which the *popular days* graph is based on.
- Re-ordering the section titles to *live and prediction vacancy*, *tariffs and capacity* and *popular days.*
- Change the wording for the *unavailable* prediction.
- For the other suggested location, adding a note if the tariffs are the same.
- In the *tariff brackets* table, have *"6 to 24 hours"* last.
- Adding a Google Maps link with the location of the address.

5.3.1.Future Improvements of the Artefact based on Feedback (summary)

These represent **future suggestions** to investigate and implement.

Individual location page for *citizen* user type:

- Making the suggested location clickable which will take the user to the location's individual page.
- Informing the citizen/user if they do not need to buy a parking ticket for a longer time because the charges times are about to expire.

"All locations for pay-and-display" & *"All locations for electric vehicles"* pages:

- Improvements to the display structure (headings, graphs sections).
- Ensuring that the data is presented in a priority hierarchical manner.
- Possibly a *content title* for the graphs at the beginning of the page.
- Making it clear that the graphs are interactive in a non-intrusive way.

"All locations" page for *citizen* user type:

- Adding the citizen's current location on the map.
- Add the location's name under the marker on the map.
- Different marker colours on the map for locations that are fully occupied vs locations that have vacant spaces.
- Transforming the hover-over information of the map into a table format.

Others:

- Vacancy predictions based on location length stay (for a long stay location where you would park the car when you go on holiday, the vacancy prediction to be longer than for a short stay location as presented in this report).
- Colour improvements (e.g. *location RR* is always defined by pink, and *location A* is always defined by green, etc.).
- Investigating linking Google Maps with the vacancy predictions/live counts.

5.4.Strengths of the Artefact based on Feedback (summary)

From the **overall best features**:

- The CCTV/video occupancy check or the live count → mentioned 4 out of 5 times.
- Vacancy predictions → 3/5.
- The map view → 3/5.
- The individual page for each location → 2/5.
- The hover-over information from the map view → 1/5.
- Overall information presented and ease of graph understanding → 1/5.
- Location's capacity information → 1/5.

Some of the participants (in the follow-up questions) have been asked about their **thoughts on the vacancy prediction range**. Their almost overall responses included that the current range is *"fine, enough"*/*"it would be alright"* as some participants mentioned that *popular days* could help in estimating the parking occupancy percentage in the long planning. One participant suggested that the vacancy prediction could be based on the location's length of stay, mentioning that for a long stay location (where you would park when you go on holiday), 2-hours may not be enough as is the case for the presented locations.

Lastly, all participants have been asked about their **thoughts regarding the council providing such information** (as seen under the *citizen* user type) to them. Their responses summarise as having such information would be "*absolutely amazing*", *"very useful and really good"* and *"not just for local citizens but also for people that come in from outside the town"*. Moreover, it was mentioned that the council could better manage the funding and that they should properly advertise the existence of such system to its citizens. Additionally, a participant mentioned that such a system would be useful for anyone who owns and maintains a car park.

5.5.Questionnaire Analysis

To help with the data analysis of the questionnaire responses, the data analysis tool provided by Schrepp (2022) was used, including the creation of the graphs.

Table 1 shows that *participants 4* and *1* found the artefact to be reflecting an overall very positive impression (*average above 2*). *Participants 5 and 3* have an overall positive impression (*average above 1.4*) and *participant 2* being closer to an overall neutral impression (*average of 0.8*).

Table 1: Scale average per participant

	Attractiveness	Perspicuity	Efficiency	Dependability	Stimulation	Novelty
Participant 1	2	2.25	2.5	1.25	1.75	2.5
Participant 2	0.83	0	1.75	1	0.25	1
Participant 3	1.5	1.5	1.25	1.75	2	0.75
Participant 4	3	2	2.5	2.5	3	2.75
Participant 5	1.67	2.25	2	1.5	1.5	2

Figure 38 shows that overall, most of the item pairs have a positive impact, with *impractical/practical* having a 100% score in the category of *"7"* (which is interpreted as *"+3"/extremely good*). There is one neutral score that stands out which is for *not secure/secure*.

Further, 4 item pairs have each a score of 20% for the category of *"3"* (which will be interpreted as a *"-1"*). Their scales in question would be *perspicuity* (2 items), *novelty* (1 item) and *efficiency* (1 item). Additionally, 2 item pairs have each a score of 20% for the category of *"2"* (which will be interpreted as a *"-2"*). Their scales in question would be *novelty* and *stimulation*. No answers were given for category *"1"* (which will be interpreted as *"-3"/horribly bad*).

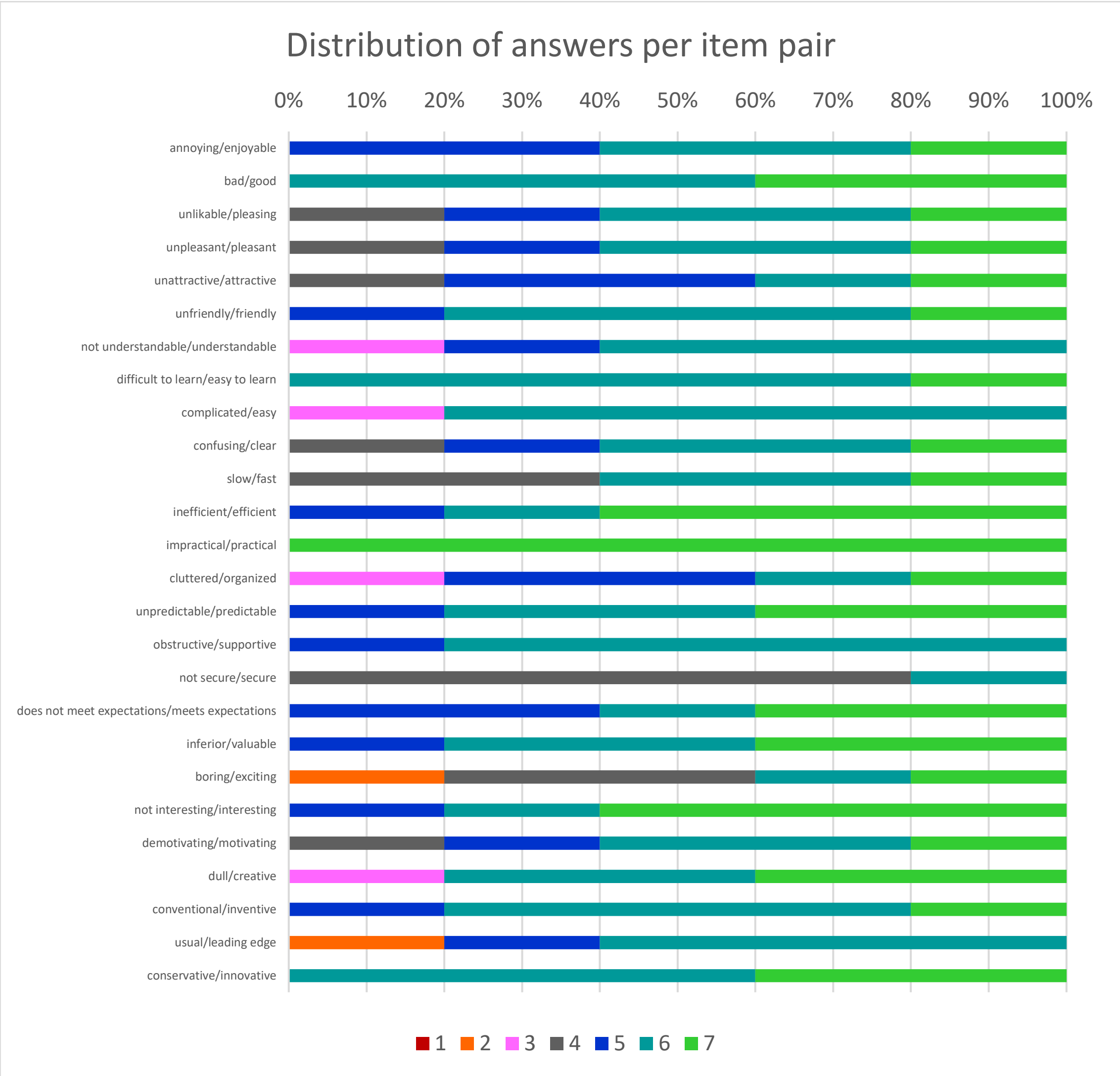


Figure 38: Distribution of answers per item pair

Figure 39 shows that the average value per item pair does not go below 0.4 (*not secure/secure*) which could be overall interpreted as a neutral to (very) positive evaluation per item pair.

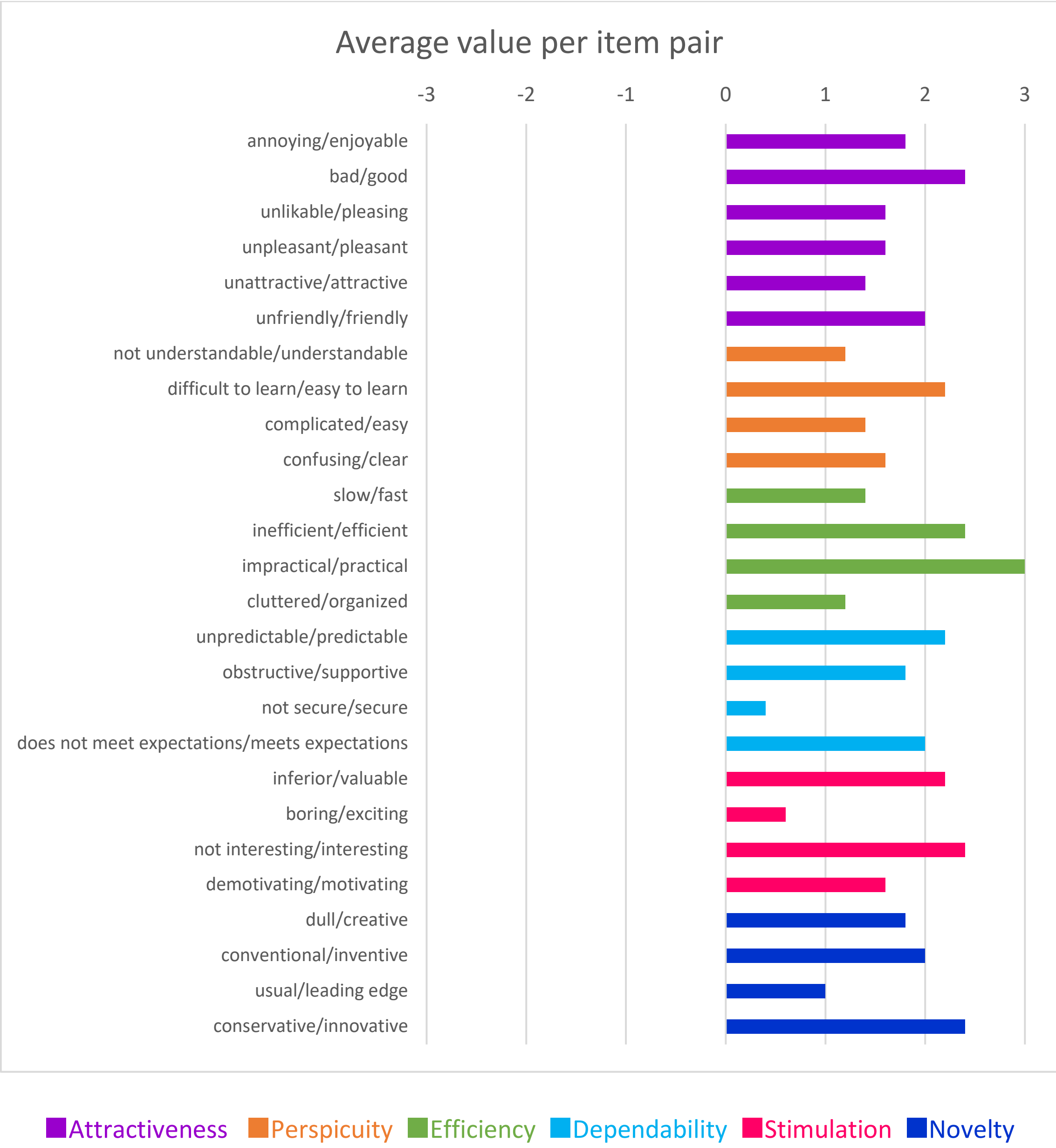


Figure 39: Average value per item pair

Figure 40 shows the average per scale with the variance interpretation. As expected, the mean does not exceed +2 (or -2) due to the participants' different opinions and answer tendencies (Schrepp, 2022).

Overall, the results are trending towards a positive scale evaluation, with the highest variances in *simulation* (0.98), *perspicuity* (0.89) and *novelty* (0.79).

Efficiency has the highest mean (2), showing that the artefact is having a good time response and interaction efficiency, with ease of use in achieving the task.

Attractiveness and *novelty* are second (1.8) showing a good overall impression of the artefact and its proposed creativity and innovation.

Stimulation is third (1.7) showing the user's excitement and motivation in using the artefact.

Perspicuity and *dependability* are last (1.6) showing the ease of learning, understandability, familiarity and the user's interaction control and overall behaviour prediction.

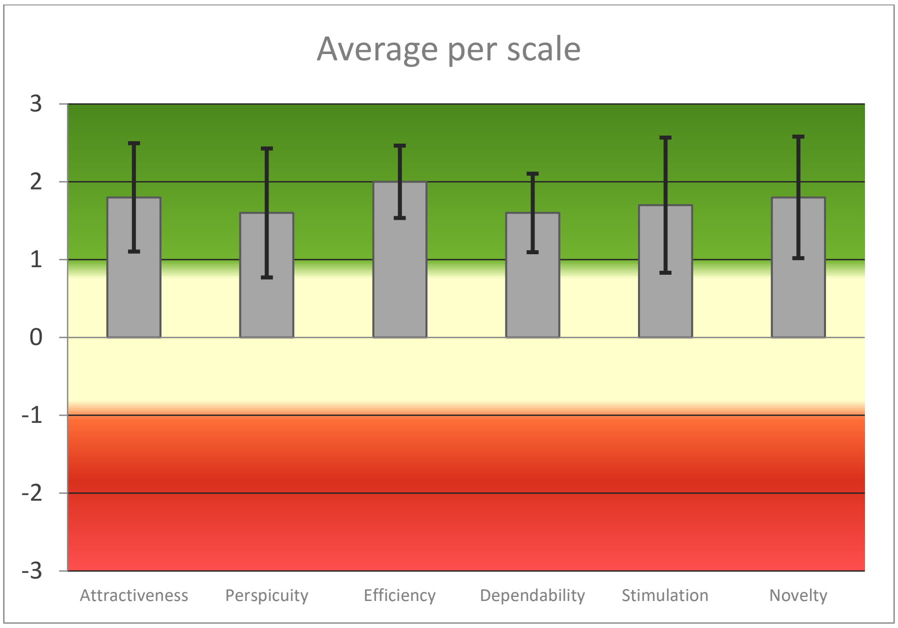


Figure 40: Mean per scale

Moreover, the *pragmatic quality* average (1.73, the goal-related quality aspects) is lower than the *hedonic quality* average (1.75, the non-goal-related quality aspects) by a very small amount, almost insignificant.

Lastly, the artefact is evaluated against a benchmark dataset. It contains the data from 21175 persons from 468 studies evaluating different products including webpages (Schrepp, 2022). Evaluating it against the benchmark allows conclusions concerning the strengths and weaknesses of the artefact (Schrepp, et al., 2017). The only limitation is that the benchmark is not split by different product types, but it is useful in situations when the artefact has not been evaluated before (Schrepp, et al., 2017).

Figure 41 shows that the artefact's evaluation is displaying *"above average"* results compared to the benchmark.

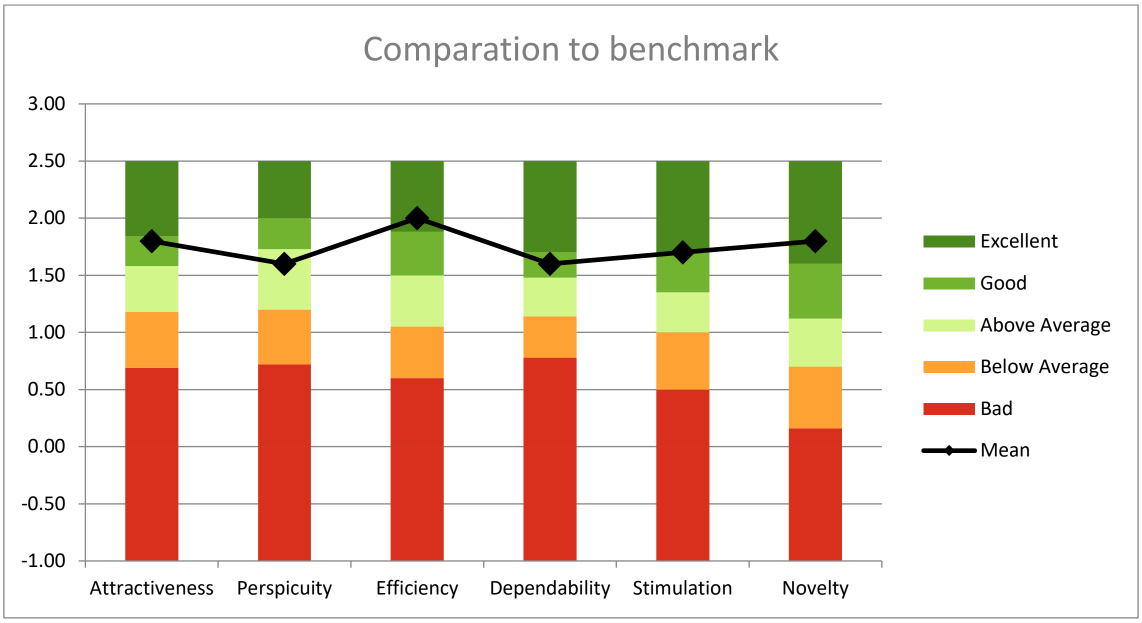


Figure 41: Comparation to benchmark

Finally, no inconsistent answers were detected, meaning that all participants answered seriously and not randomly.

5.6.Prediction Model(s)

8 models have been trained and saved. The validation was assessed in two ways:

- using *validation_split* (80/20) to evaluate the *loss* at the end of each epoch (RMSE was also added as a model metric here). This produced the *training/validation loss* graph.
- splitting the dataset 80/20 to evaluate the model and get the RMSE, MAE, MSE and R2-score.

RMSE measures how spread the residuals (prediction errors) are from the line of best fit. The lower the value the better the model is performing.

MAE measures the absolute average distance between the true data and predicted data but it omits large errors in prediction. Large MAE shows issues in certain areas whilst small MAE shows that the model has a good prediction of the outputs.

MSE measures the squared average distances between the true data and the predicted data. A larger value means that the values are dispersed widely around the mean (meaning larger error) while a small value means the opposite.

R-squared (R^2/R2) shows the goodness of fit measure, the higher the value the better.

For location RR, using the *valid tickets counts* dataset for training and prediction the following results were obtained:

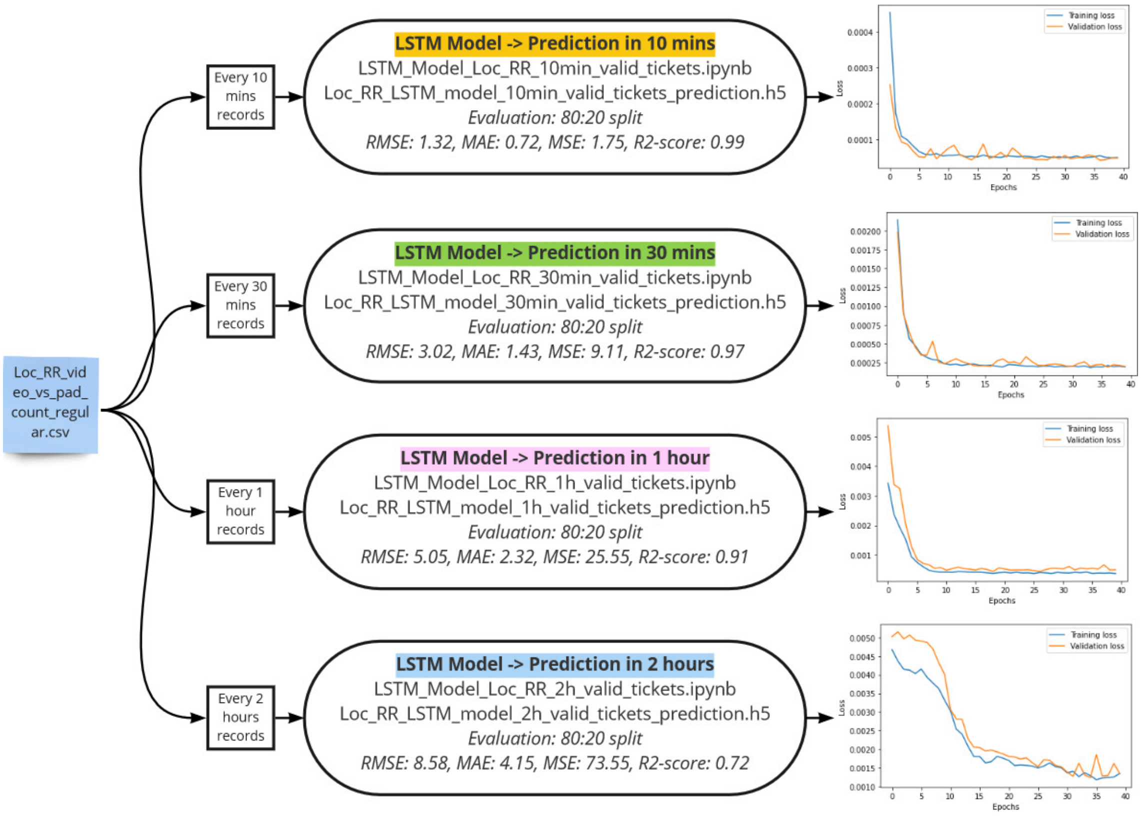


Figure 42: LSTM models for valid tickets counts, Location RR

Note: *Valid tickets counts* were used as the *CCTV/video counts for regular parking spaces* are based on *the valid tickets counts*.

Figure 42 shows that by increasing the time gaps within the *valid tickets* the evaluation results worsen. This is due to decreasing the number of records used for the model.

The prediction in 10 minutes model shows promising results for both the evaluation metrics and *training/validation loss* graph with a good fit learning curve. R^2 score is high (99% of the data fits the model) and the other metrics values are small showing a model that is performing well.

The prediction in 30 minutes model shows an increase in values for RMSE, MAE, MSE and a decrease in the goodness of fit (97% of the data fits the model). As expected, there were fewer values used for this model. The *training/validation loss* graph still shows a good learning curve.

The prediction in 1-hour model has its metrics getting not so good (91% of data fits the model), but the *training/validation loss* graph is still considered a good fit.

Lastly, *the prediction in 2-hours model* has the worse results with a learning curve that shows unrepresentative data used (72% of the data fits the model). Having more values used to train/test the model would greatly help in improving these results and it was chosen to be left in the artefact as a proof of concept.

To continue, for location RR, the *CCTV/video counts for disabled bay spaces* have been used to create and train new models, *figure 43*. The data was created *almost randomly* for these counts.

The prediction in 10 minutes model gives good metrics and a good fit for the learning curve, but an R^2 score of 0.93.

The prediction in 30 minutes model shows an R^2 of 0.78 and a good fit for the learning curve. The other metrics have increased, as expected.

The prediction in 1-hour model has its metrics worsen and an R^2 of 0.6 with an unrepresentative training dataset accordingly to the graph.

Because of the *1-hour model* results, it was concluded that the dataset is not big enough to continue.

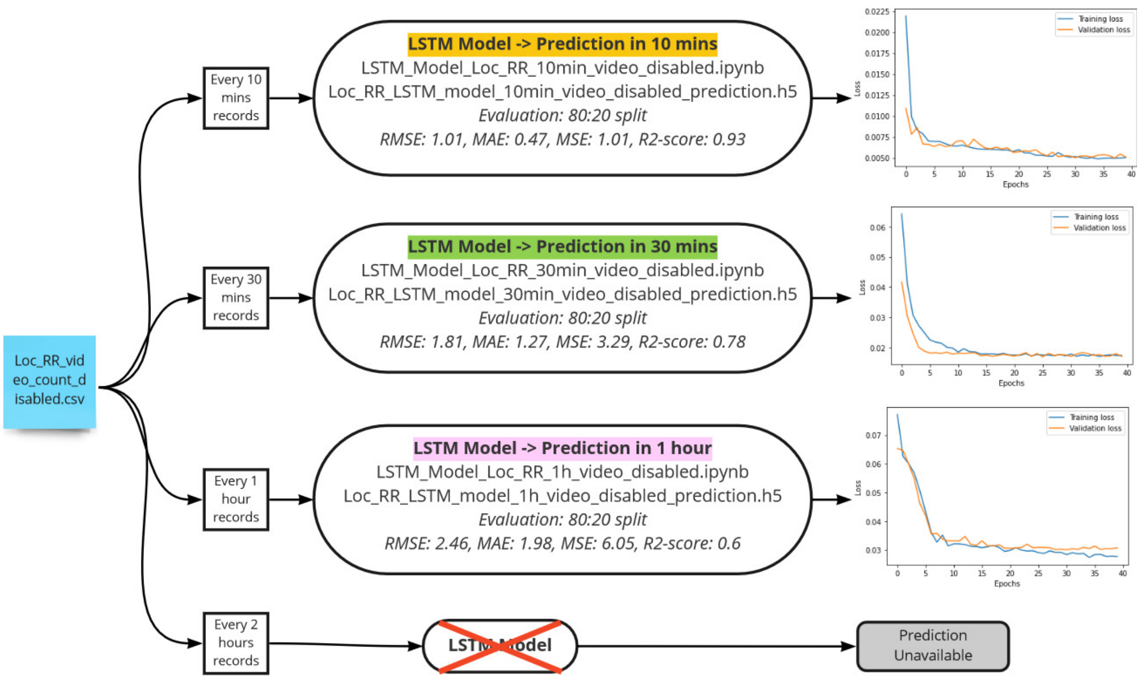


Figure 43: LSTM models for disabled bays CCTV/video counts, Location RR

Additionally, because location A's *valid ticket counts* are mainly based on the same dataset as location RR, the trained models were applied to location A's dataset, *figure 44*. The results are promising for t*he prediction in 10 minutes model* and worsen as you move along. This was expected as the same behaviour has been seen for location RR. *The prediction of 2-hours model* shows that only 68% of the data fits the model.

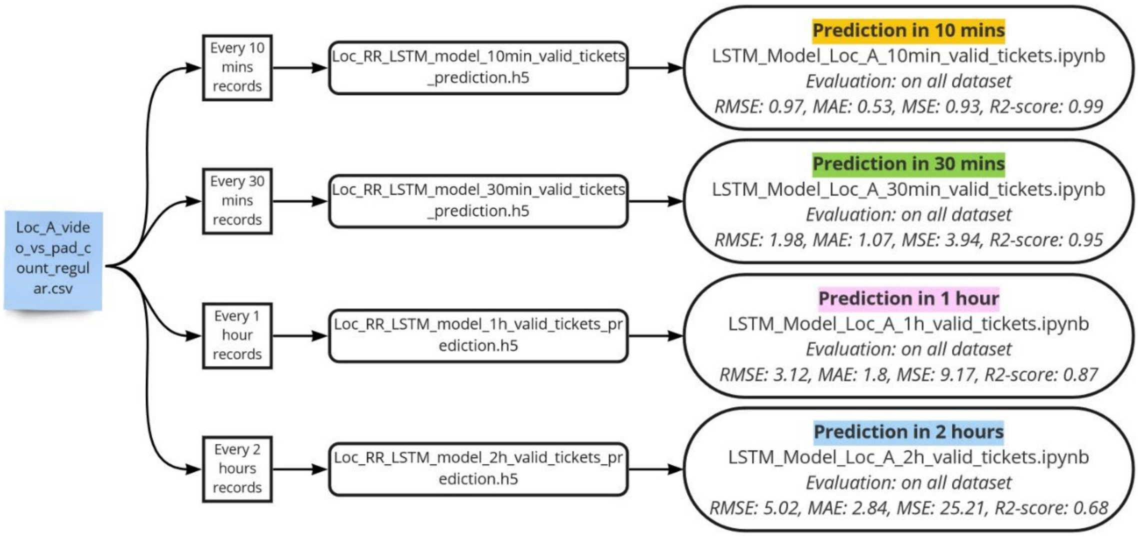


Figure 44: Using saved LSTM models for Location A's valid tickets counts

Lastly, the saved models for location RR's *disabled bay spaces from CCTV/video count* could not be applied to location A's dataset. The results were very bad as those datasets were created *almost randomly* and had a high difference in parking spaces.

A new model for *every 10 minutes records* was created based on location A's values, *figure 45*. The RMSE, MAE & MSE have low values, but the R^2 score is not promising. The *training/ validation loss* graph shows an unrepresentative validation dataset. This prediction was left in the artefact as a proof of concept and the next brackets predictions were not added because the results were very unsatisfactory (11.6.Appendix F – LSTM Models Experimentation).

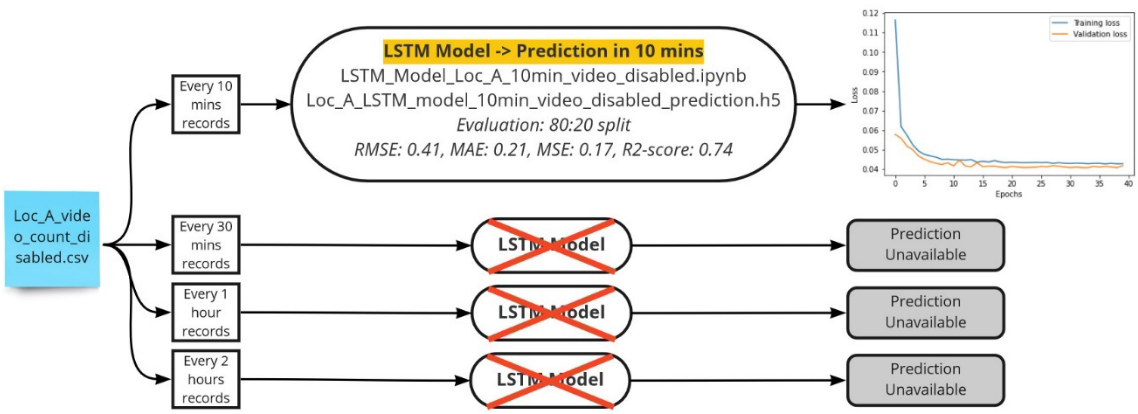


Figure 45: LSTM models for disabled bays CCTV/video counts, Location A

Additionally, to ensure that the predicted value does not exceed the total number of parking spaces or goes below zero. After the prediction, the value is rounded and transformed into an integer and placed through a function that will check if the value is lower than zero or bigger than the total number of parking spaces for that space type. If true, the returned value would be either zero or the total number of parking spaces for that space type.

```
#making sure the prediction is in range
@st.cache
def in_range_prediction(field, max_space):
    if field < 0:
        field=0
    elif field > max_space:
        field = max_space
    return field
```

Figure 46: Prediction in range function

5.7.Occupancy Check Testing

The testing of the occupancy check is restricted by two main factors:

- The video's short time-lapse (and having only daytime recording).
- Manual segmentation of each individual location and the optimal count of non-zero values.

Therefore, to test the accuracy of the occupancy check, the approach taken was to identify how many false detections were made in every 30 frames (meaning 1 second) of the video. Therefore, in total, 73 frames have been extracted from the video and manually checked in terms of false detections of occupied spaces*, *table 2*.

*Caveat: It was considered a false detection only if the space was occupied by something else than a vehicle. If a vehicle was interacting with the space (e.g.: pulling in/out, passing by) and that space occupancy would change or not, that was not considered a false detection.

Table 2: Occupancy check testing

Location	Type	Out of 73, the total number of frames with **correct detections**	Out of 73, the total number of frames with **false detections**	False detections caused by	Frame numbers with false detections	Comments
RR	*Regular space*	63 (86%)	10 (14%)	Shades, people or/with objects	150, 1590, 1920, 2010, 2040, 2070, 2100, 2130, 2160 & 2190	Frame 1920 was the only one with 2 incorrect spaces detections. The rest had only 1 incorrect space detection out of the total number of spaces.
RR	*Disabled bay spaces*	72 (99%)	1 (1%)	Possibly shade	450	1 incorrect space detection out of the total number of spaces.
A	*Regular space*	71 (97%)	2 (3%)	People/ shade	1890 &1920	1 incorrect space detection out of the total number of spaces.
A	*Disabled bay spaces*	73 (100%)	0	N/A	N/A	N/A

6.Discussion & Limitations

This study focuses on helping the councils better optimise or improve their *pay-and-display* on-/off-street parking (delimited by parking lines) with the use of computer vision and predictive analysis. So that they would understand parking behaviours and patterns which could lead to improved parking management that is also tackling factors created by cruising for parking.

The analysis suggests that the created artefact would help in optimising the parking management based on identified patterns and it will improve the citizen's experience regarding pay-and-display parking.

The data analysis for location RR based on the pay-and-display true/original dataset suggests that this location is mostly visited by drivers that intend to stay a relatively short time, with *"6 to 24 hours"* stays being 4th and with a ticket machine more preferred than the other. This location was never fully occupied and only had some peaks of up to 151 out of 251 available *regular bay spaces*. From this, it can be concluded that this location is not used to its maximum capacity. Therefore, in other nearby busy parking locations, this location can be advertised as a less occupied location. Further, when looking at the whole dataset, no significant correlation was seen between the *daily average valid tickets* and the *weather data*. As a standalone, these results contradict the claims made by Pflügler, et al. (2016) who suggested that the weather (most importantly the temperature) has a significant impact on traffic behaviours and it influences parking predictions. But, when analysing individual months, sometimes weak to moderate correlation can be seen (excluding very weak), and not just for the *temperature information* but also for *rain* and *wind speed*. Pflügler, et al.'s (2016) caveat on the findings (July-September 2015 data range and evaluated on one city) could be the reasoning for different results. Moreover, a partial agreement with Pflügler, et al. (2016) could be seen on the time, the findings suggest that during the day the *valid ticket count* oscillates and during the night it is almost the same and the weekends are busier than the weekdays. The disagreement could be seen for holidays. Although Pflügler, et al. (2016) do not define the holidays specifically, it was found that the presence of school holidays has a weak (/non-significant) point biserial correlation coefficient with the *daily average valid ticket* or *CCTV/video counts*. This could be caused by the location and limited data range used. Lastly, hypothetically speaking, discrepancies between the *valid ticket counts* and the *CCTV/video occupancy count*s can be seen, this could be due to multiple reasons such as cars parked without a ticket, parked with an expired ticket, the time difference between parking and buying a ticket and cars leaving before their expiry time on the ticket. Therefore, the council could optimise their workforce tasks (traffic wardens) and have them focus on locations with a higher discrepancy between the two counts. Bringing benefits both to the council (revenue) and the citizens (more spaces as illegally parked cars are identified sooner). Additionally, an idea of how busy the *disabled bay parking spaces* can be used for better parking management. Due to the lack of true data regarding the *CCTV/video count* for *regular* and *disabled parking bay spaces*, the results cannot confirm something objective but show a concept that could be used by the council to optimise their parking management.

When comparing the two locations (RR and A) the data suggest that the locations are similar in terms of preferred stay, the mean per weekday per each month for *CCTV/video regular parking spaces count* and the highest occupancy percentages per month for *valid tickets*. Moreover, partial similarities were seen in the 50% of the data spread for cash paid, the mean per weekday for each month for *valid tickets*, the lowest occupancy for *disabled bays spaces* and overall transaction number. The reliability of those results is impacted by the high similarity of the two datasets. They are added as a proof of concept in terms of what the council could evaluate. They were compared and contrasted, and they could be useful for future strategies that involve funding or policies as Dogru, et al. (2017) suggested, different areas may prefer different policies. Additionally, dynamic pricing as presented by Simhon, et al. (2017) and Harris (2014) could be applied to increase the occupancy rate to a target level.

From analysing the **electric vehicle charging points transactions** dataset it was found that the total time when the vehicle is plugged in and the total time when the vehicle is actually charging has an approximate difference of 100 to 135 minutes per month. The boxplots showed that 50% of the data for plug-in/unplug time was between 170 to 707.5 minutes whereas 50% of the data for charging start/end was between 146 to 466.5 minutes. These results could possibly support the proposal of He, et al. (2016) where the EV owners would help balance the supply and demand of the main grid using an optimal charging scheme as the EV owners have on average the vehicle plugged in for a longer time than the charging time needed. Moreover, it was found that some charging stations IDs were barely used whereas others were frequently used. This could reinforce the suggestions made by Morrissey, et al. (2016) that the future charging station infrastructure needs to be located strategically. Lastly, patterns of popularity and non-popularity were discovered. It was found that the dates of the 23rd, 26th, 27th and 30th are the busiest days for plugging-in and unplugging the vehicles. Whereas the 10th and 31st are the least popular days for plugging-in and unplugging the vehicles. 31st could be influenced by the caveat that the dataset only contains 2 dates of 31st. In terms of the times, as expected, during the night (12-3AM) has almost no transactions, 8AM-1PM has very similar plugging-in/unplugging events and 2-6PM is very popular for plugging-in events, whereas 4-7AM and 7-11PM is the most popular for unplugging events. These results contribute to a clear understanding of the charging patterns and they seem to relate to the working hours' lifestyle. Meaning that the driver unplugs their vehicles early in the morning or late at night to either go to work or spend the evening out, plugging-in is higher in the afternoon, when the driver could come back from work. Similar events from 8AM-1PM could show drivers that are using the charging stations while sorting things out in the town. These time and date patterns could be useful information in efficiently preparing the vehicles that would support the EV car-sharing systems that are proposed by Brandstätter, et al. (2017). Although, the generalisability of the results is limited by the reduced data and data range and the high number of different charging points IDs.

From the overall task analysis, the data suggests that even though the time spent to complete each task varies, the number of clicks is almost identical for all the participants and with the benchmark for all the tasks except task number 4. This could be caused by the fact that this task was lacking every-step instructions as it was aiming to also test the participant's familiarity and ease of learning of the system. Interesting results were seen regarding the participants' subjective opinion of the hardest task. For 3 out of 5 participants, the data shows that they had other tasks that they either took longer to complete or had more clicks than the task they have chosen as the hardest. One of the participants reasoned their choice (task 1) because of the first-time introduction to the system.

The participants' overall feedback suggest that the best features of the system would be the *CCTV/video occupancy check or the live count, the vacancy predictions, the map view* and *the individual page locations* showing that the proposed artefact would improve the citizens' experience as defining this information availability as *"absolutely amazing", "very useful and really good"* and timesaving. This could bring the same outcomes for the council and for the citizens as seen in Mangiaracina, et al. (2017) results. Showing that they would not just save drivers time and money but reduce the city's CO_2 emissions and increase revenue. Moreover, one of the participant's suggestion that the council could better manage funding based on popularity goes in line with the suggestions made in this project.

The questionnaire analysis shows that the artefact's results regarding the 6 scales are *"above average"* to *"excellent"* when compared to the benchmark. Although, when analysing more in-depth the findings, the graphs indicate that the distribution of answers per item pair requires improvements for *perspicuity, novelty, efficiency* and *stimulation*, but the *average per* *scale* shows that *efficiency* scored the highest (2), followed by *novelty*, then *simulations* and then *perspicuity* (1.6). The 1.6 mean is considered *"above average"*. Moreover, the *hedonic quality* scored higher than the *pragmatic quality*. Therefore, even though some participants chose the negative impression, the overall average is good. Lastly, one interesting finding was the *not secure/secure* item pair which scored the lowest average value. This could be due to the participants' confusion regarding this item pair. Although, it is important to note that those results could be subjective (the participants were not strangers to the researcher) and highly influenced by the small sample size as Schrepp, et al. (2017) recommends 20-30 users for providing a *"quite stable measurement".*

Multiple LSTM models were created and used to determine the vacancy predictions. The results show that the datasets with *every 10 minutes records* for the *valid tickets count* perform the best. The more the gap times of the counts are spread, the evaluation results worsen. This is due to the reduced number in dataset size, but they have been left within the artefact as a proof of concept. The results for the *disabled bays spaces* vacancy predictions are more subjective and could not be performed for up to 2 hours predictions. The evaluation results became worse faster and for location A, for *every 10 minutes records*, the *training/validation loss* graph shows an unrepresentative validation dataset. The results for both locations are highly influenced by the nature of the dataset. These behaviours were expected due to the almost randomly created datasets for *CCTV/video counts*. Additionally, the proof of concept regarding the prediction range was evaluated by some of the participants. They have been asked their thoughts regarding the vacancy prediction intervals and the answers show that 2-hours prediction is *"fine, enough"*. Therefore, even though this study is slightly limited in terms of prediction's accuracy, the proof of concept shows that 2-hours is optimal for improving the drivers' experience. Lastly, as Guerrini, et al. (2021) used the Prophet model to determine the monthly occupancy forecast, the experiments of the Prophet model used on the *valid tickets* dataset did not provide good evaluation metrics.

The **created occupancy check** follows similar concepts/steps as Yusnita, et al., (2012) and Bibi, et al. (2017). The results show that the proposed method correctly classifies each parking space with an accuracy that varies from 86% to 100% (based on location and parking space type). The results are partially similar to the two studies. The algorithm's performance was influenced by shadows, people and objects and no more than 2 spaces (out of the total number of spaces) in a checked frame were affected by this. Shows that even though the smallest accuracy is 86%, in reality, almost all of the spaces are classified correctly. Besides the limitations mentioned above, this approach contains other limitations such as induvial location parking spaces segmentation, the optimal count of non-zero and fixed top-view camera. Moreover, the evaluation was not done in unperfect conditions.

Furthermore, the participants that took part in the test only represent the *citizen* user type and not the *council staff*. The results could be highly impacted by this as Brooke et al. (2017) study shows that the local government officials do not see the parking search as a serious problem. Because the study was published in 2017 and the officials suggested that parking search could become a serious problem in few years, it is argued that this artefact would contribute as a possible solution to this problem as now, after 5 years, there are multiple studies that show the devastating effects of parking searching.

7.Demonstration of Achieving the Project's Aims

The main and the secondary aim of the project have been achieved because all the objectives have been achieved.

☑Usage of multiple datasets and video footage. → 3.1.Datasets & Video Footage Collection & 11.1.Appendix A – Datasets

☑Apply data cleansing and data analysis to find patterns/insights. →4.3.Data Cleansing & 5.1.Data Analysis

☑Overlap video occupancy with the respective dataset occupancy. → 11.2.Appendix B – Valid Tickets & 4.4.3.Video Dataset Creation

☑Make vacancy predictions based on previous data. → 4.5.Prediction Model(s)

☑Include different parking space types. → 4.4.2.Occupancy Check

☑Create a visual interface that could be used by council to understand behaviours and patterns of different locations. → 4.6.Overall Artefact Implementation & 11.8.Appendix H – Artefact

☑Propose further ways to optimise or improve council's pay-and-display parking. → 6.Discussion & Limitations

The artefact is a usability tested solution that could benefit the council and implicit its citizens. The artefact contains information that could be used by the council to optimise or improve their pay-and-display parking. The data analysis made could pave the way for future parking management policies and the creation of charging station points.

7.1.Future work

The artefact requires further work based on the usability testing. The suggestions made in 5.3.1.Future Improvements of the Artefact based on Feedback (summary) should be evaluated further and implemented within the system. Moreover, some of the suggestions given to explore the system's implementation with Google Maps could lead to new opportunities for improvement.

Further research is needed to establish that these solutions could be embraced by the local authorities now and could be used as a new insight into optimising or improving the pay-and-display parking. Moreover, future studies should consider broader data (especially the one that involves the pay-and-display ticket machine logs) to establish links between different locations and how could their parking management be improved with the use of pattern understanding.

Lastly, another avenue for future research includes the proposed occupancy check method and how the weather and night-time would affect the efficiency, and how could this be applied to locations that do not have a fixed top-view camera.

8.Conclusion

The study aimed at helping the council better optimise or improve their pay-and-display parking with the use of computer vision and predictive analysis as well as enhancing other parking types.

The proposed artefact (which contains two user type categories: the council staff and the citizens) paves the way for providing a solution that combines computer vision, predictive analysis and data analysis to help the local authorities better manage their parking, find parking patterns and behaviours and implicit improve their customers' experience.

The key findings of the project show that the analysed parking locations are not used to their full capacity. The weather information has a weak to moderate correlation with the parking behaviours but only on some of the months and not on the whole dataset at once. Moreover, it was found that the presence of school holidays has a weak point biserial correlation coefficient with the *daily average of valid tickets* or *CCTV/video counts*. As expected, the time and the day leave a mark on the parking behaviours, meaning that during the day the occupancy oscillates and during the night the occupancy stays almost the same with weekends relatively busy. Future strategies for funding, policies and parking re-routing can be applied to a location based on the location comparison analysis. In addition, the *CCTV/video* occupancy check gives an insight into the actual occupancy vs the ticket machine logs showing that a discrepancy is happening. Therefore, traffic wardens can be sent to locations that have a higher discrepancy. Additionally, the *disabled bay spaces* are monitored and their occupancy could be used for future decisions.

Furthermore, the analysis of the electric vehicles dataset shows that the time a vehicle is plugged-in and unplugged is greater than the time it actually charges. Moreover, patterns of popular days and times were discovered and some charging point station IDs were more frequently used than others.

Overall, the usability testing results showed that the proposed artefact is considered *"above average"* for the six scales when compared to a benchmark and the best features suggested by the participants included the *CCTV/video occupancy check or the live count, the vacancy predictions, the map view* and *the individual page locations* showing that such solution would improve the council's parking management. Also, the participants suggested that up to 2-hours of vacancy prediction is sufficient.

This project's contribution consists in proposing an artefact that could serve the local authorities as a tool that could be used to manage their pay-and-display parking and to use it for future strategical policies, funding, traffic wardens distribution, revenue improvements, to tackle the parking search issue, to reduce the CO_2 emissions and even security. Moreover, the project supports the claims made by Pflügler, et al. (2016) regarding the weather and time influencing parking behaviours and possibly challenges the claims regarding the holiday influence. Additionally, the found insights could be useful in Morrissey, et al. (2016) study.

While the usability testing **limits** the local authority's acceptance of the system, this proposal supports the insights seen regarding the citizen's acceptance of such a system and the growing issue of cruising for parking. Further, while the datasets and video footage used limit the generalisability of the results, this approach provides a tool and insights into areas that could be used for optimising or improving the council's pay-and-display parking.

Lastly, to better comprehend the implications of these results, future studies could address the new local authority perspective on the parking search and the use of broader datasets. More recommendations for future work are made within 7.1.Future work.

9.References

Acharya, D., Yan, W. & Khoshelham, K., 2018. *Real-time image-based parking occupancy detection using deep learning.* Adelaide, Australia, Proceedings of the 5th Annual Research@Locate Conference, Volume 2087, pp. 33–40.

Amato, G. et al., 2016. Car parking occupancy detection using smart camera networks and Deep Learning. *2016 IEEE Symposium on Computers and Communication (ISCC),* pp. 1212-1217, doi: 10.1109/ISCC.2016.7543901.

Anwer, F., Aftab, S., Waheed, U. & Muhammad, S., 2017. Agile Software Development Models TDD, FDD, DSDM, and Crystal Methods: A Survey. *International Journal of Multidisciplinary Sciences and Engineering,* 8(2), pp. 1-10.

Becker, P., Tebes, G., Peppino, D. & Olsina, L., 2019. Applying an Improving Strategy that embeds Functional and Non-Functional Requirements Concepts. *Journal of Computer Science & Technology,* 19(2), pp. 153-174, doi: 10.24215/16666038.19.e15.

Bibi, N., Majid, M., Dawood, H. & Guo, P., 2017. *Automatic Parking Space Detection System.* Wuhan, China, 2017 2nd International Conference on Multimedia and Image Processing (ICMIP), pp. 11-15, doi: 10.1109/ICMIP.2017.4, pp. 11-15, doi: 10.1109/ICMIP.2017.4.

Brandstätter, G., Kahr, M. & Leitner, M., 2017. Determining optimal locations for charging stations of electric car-sharing systems under stochastic demand. *Transportation Research Part B: Methodological,* Volume 104, pp. 17-35, doi: 10.1016/j.trb.2017.06.009.

Brooke, S., Ison, S. & Quddus, M., 2017. On-street parking search: A UK local authority perspective. *Journal of Transport and Land Use,* 10(1), pp. 13-26, doi: 10.5198/jtlu.2015.600.

Brownlee, J., 2017. *Gentle Introduction to the Adam Optimization Algorithm for Deep Learning.* [Online]
Available at: https://machinelearningmastery.com/adam-optimization-algorithm-for-deep-learning/
[Accessed 27 08 2022].

Bulan, O. et al., 2013. Video-based real-time on-street parking occupancy detection system. *Journal of Electronic Imaging,* 22(4), pp. 1-15, 041109, doi: 10.1117/1.JEI.22.4.041109.

Carlier, M., 2022. *Number of licensed cars in the United Kingdom (UK) from 2000 and 2020.* [Online]
Available at: https://www.statista.com/statistics/299972/average-age-of-cars-on-the-road-in-the-united-kingdom/
[Accessed 18 06 2022].

Deza, A., Huang, K. & Metel, M., 2022. Charging station optimization for balanced electric car sharing. *Discrete Applied Mathematics,* Volume 308, pp. 187-197, doi: 10.1016/j.dam.2020.01.042.

Dogru, A. et al., 2017. Parking Management Policies Based on Behavior Analysis at Fatih District in Istanbul, Turkey. *Transportation Research Procedia,* Volume 25, pp. 5205-5219, doi: 10.1016/j.trpro.2018.02.048.

Enríquez, F. et al., 2017. Existing Approaches to Smart Parking: An Overview. In: E. Alba, F. Chicano & G. Luque, eds. *Smart Cities. Smart-CT 2017. Lecture Notes in Computer Science.* Volume 10268: Springer, Cham, pp. 63-74, doi: 10.1007/978-3-319-59513-9_7.

Gov.uk, 2020. *Government takes historic step towards net-zero with end of sale of new petrol and diesel cars by 2030.* [Online]
Available at: https://www.gov.uk/government/news/government-takes-historic-step-towards-net-zero-with-end-of-sale-of-new-petrol-and-diesel-cars-by-2030
[Accessed 20 06 2022].

Gov.uk, 2021. *Transport and environment statistics: Autumn 2021.* [Online]
Available at: https://www.gov.uk/government/statistics/transport-and-environment-statistics-autumn-2021/transport-and-environment-statistics-autumn-2021
[Accessed 19 08 2022].

Gov.uk, 2022. *Electric vehicle charging device statistics: July 2022.* [Online]
Available at: https://www.gov.uk/government/statistics/electric-vehicle-charging-device-statistics-july-2022/electric-vehicle-charging-device-statistics-july-2022
[Accessed 19 08 2022].

Guerrini, G., Romeo, L., Alessandrini, D. & Frontoni, E., 2021. *Analysis, Design and Implementation of a Forecasting System for Parking Lots Occupation.* Nashville, TN, USA, CPS-IoT Week '21: Cyber-Physical Systems and Internet of Things Week 2021.

Harris, R. G., 2014. *A new policy tool: dynamic pricing of on-street parking.* London, UK, Road Transport Information and Control Conference 2014 (RTIC 2014).

Hassan, M., 2021. *Parking Space Counter using OpenCV Python.* [Online]
Available at: https://www.youtube.com/watch?v=caKnQlCMIYI&t=2170s
[Accessed 29 08 2022].

He, T., Bai, Y & Zhu, J., 2016. *Optimal Charging Strategy of Electric Vehicles Customers in a Smart Electrical Car Park.* Glasgow, UK, 8th IET International Conference on Power Electronics, Machines and Drives (PEMD 2016), doi: 10.1049/cp.2016.0298.

Jakob, M. & Menendez, M., 2020. Optimal Parking Occupancy with and without Differentiated Parking: A Macroscopic Analysis. *Preprints, 2020060227, doi: 10.20944/preprints202006.0227.v1.*

Keras Documentation, 2022. *Keras API reference.* [Online]
Available at: https://keras.io/api/
[Accessed 27 08 2022].

Klipp, P., 2014. *Getting Started with Kanban.* California: CreateSpace Independent Publishing Platform, pp. 4, 6.

Kommey, B., Addo, E. & Agbemenu, A., 2018. A Smart Image Processing-based System for Parking Space Vacancy Management. *International Journal of Computer Applications (0975 - 8887),* 182(5), pp. 1-6, doi: 10.5120/ijca2018917540.

Kurtanović, Z. & Maalej, W., 2017. *Automatically Classifying Functional and Non-functional Requirements Using Supervised Machine Learning.* Lisbon, Portugal , 2017 IEEE 25th International Requirements Engineering Conference (RE), doi: 10.1109/RE.2017.82.

Lin, T., Rivano, H. & Le Mouël, F., 2017. A Survey of Smart Parking Solutions. *IEEE Transactions on Intelligent Transportation Systems,* 18(12), pp. 3229-3253, doi: 10.1109/TITS.2017.2685143.

Liu, C. et al., 2016. The Optimal Distribution of Electric-Vehicle Chargers across A City. *2016 IEEE 16th International Conference on Data Mining,* pp. 261-270 , doi: 10.1109/ICDM.2016.142.

Mangiaracina, R. et al., 2017. Smart parking management in a smart city: Costs and benefits. *2017 IEEE International Conference on Service Operations and Logistics, and Informatics (SOLI),* pp. 27-32, doi: 10.1109/SOLI.2017.8120964.

Miranda, E., 2022. *Moscow Rules: A Quantitative Exposé.* Copenhagen, Denmark, 23rd International Conference on Agile Software Development, XP 2022 (LNBIP, Volume 445, pp. 19-34).

Mondal, H., 2020. *Overview of research ethics to make a secure and progressive research.* Barrackpore, India, Conference: National Conference on Research Ethics and Library.

Monday.com, 2021. *Product.* [Online]
Available at: https://monday.com/product
[Accessed 17 02 2021].

Morrissey, P., Weldon, P. & O'Mahony, M., 2016. Future standard and fast charging infrastructure planning: An analysis of electric vehicle charging behaviour. *Energy Policy,* Volume 89, pp. 257-270, doi: 10.1016/j.enpol.2015.12.001.

Nazir, A., Zafar, I. & Abbas, M., 2017. *The Impact of Agile Methodology (DSDM) on Software Project Management.* Kuala Lumpur, Malaysia, International Conference on Engineering, Computing & Information Technology (ICECIT), pp. 1-6.

Nielsen, J., 2001. *Usability Metrics.* [Online]
Available at: https://www.nngroup.com/articles/usability-metrics/
[Accessed 28 08 2022].

Nielsen, J. & Budiu, R., 2021. *Success Rate: The Simplest Usability Metric.* [Online]
Available at: https://www.nngroup.com/articles/success-rate-the-simplest-usability-metric/
[Accessed 28 08 2022].

Niranjanamurthy, M., Archikam, N., Himaja, G. & Puneeth, S., 2014. Research Study on Importance of Usability Testing/ User Experience (UX) Testing. *International Journal of Computer Science and Mobile Computing,* 3(10), pp. 78-85, ISSN 2320-088X.

ONS, 2021. *Over half of younger drivers likely to switch to electric in next decade.* [Online]
Available at:
https://www.ons.gov.uk/economy/environmentalaccounts/articles/overhalfofyoungerdriverslikelytoswitchtoelectricinnextdecade/2021-10-25
[Accessed 20 06 2022].

OpenCV Documentation, 2022. *OpenCV modules.* [Online]
Available at: https://docs.opencv.org/4.x/index.html
[Accessed 29 08 2022].

OpenCV, 2016. *Morphological Transformations.* [Online]
Available at: https://opencv24-python-

tutorials.readthedocs.io/en/latest/py_tutorials/py_imgproc/py_morphological_ops/py_morphological_ops.html
[Accessed 26 08 2022].

OpenCV, 2022. *Image Thresholding.* [Online]
Available at: https://docs.opencv.org/4.x/d7/d4d/tutorial_py_thresholding.html
[Accessed 26 08 2022].

Padmasiri, H., Madurawe, R., Abeysinghe, C. & Meedeniya, D., 2020. Automated Vehicle Parking Occupancy Detection in Real-Time. *2020 Moratuwa Engineering Research Conference (MERCon),* pp. 1-6, doi: 10.1109/MERCon50084.2020.9185199.

Pflügler, C. et al., 2016. *Predicting the Availability of Parking Spaces with Publicly Available Data.* Bonn, INFORMATIK.

Plotly Graphing Library, 2022. *Plotly Open Source Graphing Library for Python.* [Online]
Available at: https://plotly.com/python/
[Accessed 29 08 2022].

Rahman, S. et al., 2020. Convolutional Neural Network Customization for Parking Occupancy Detection. *2020 International Conference on Electrical Engineering and Informatics (ICELTICs),* pp. 1-6, doi: 10.1109/ICELTICs50595.2020.9315509.

Rosebrock, A., 2021. *OpenCV Smoothing and Blurring.* [Online]
Available at: https://pyimagesearch.com/2021/04/28/opencv-smoothing-and-blurring/
[Accessed 26 08 2022].

Rye, T. & Koglin, T., 2014. Parking Management. *Parking Issues and Policies (Transport and Sustainability),* Volume 5, pp. 157-184, doi: 10.1108/S2044-994120140000005027.

Schrepp, M., 2022. *Data Analysis Tool.* [Online]
Available at: https://www.ueq-online.org/
[Accessed 29 08 2022].

Schrepp, M., Hinderks, A. & Thomaschewski, J., 2017. Construction of a benchmark for the User Experience Questionnaire (UEQ). *International Journal of Interactive Multimedia and Artificial Intelligence,* 4(4), pp. 40-44, doi: 10.9781/ijimai.2017.445.

Schrepp, M., Hinderks, A. & Thomaschewski, J., 2017. Design and Evaluation of a Short Version of the User Experience Questionnaire (UEQ-S). *International Journal of Interactive Multimedia and Artificial Intelligence,* 4(6), pp. 103-108, doi: 10.9781/ijimai.2017.09.001.

Shoup, D., 2006. Cruising for parking. *Transport Policy,* 13(6), pp. 479-486, doi: 10.1016/j.tranpol.2006.05.005.

Simhon, E., Liao, C. & Starobinski, D., 2017. Smart parking pricing: A machine learning approach. *2017 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS),* pp. 641-646, doi: 10.1109/INFCOMW.2017.8116452.

Sinek, S., 2011. *Start with Why.* England: Penguin Books Limited.

Solent University, 2021. *How to conduct a literature review.* [Online]
Available at: https://learn.solent.ac.uk/mod/resource/view.php?id=1764929
[Accessed 16 02 2021].

Streamlit Library, 2022. *Streamlit API reference.* [Online]
Available at: https://docs.streamlit.io/library/api-reference
[Accessed 29 08 2022].

Streamlit, 2022. *st.cache.* [Online]
Available at: https://docs.streamlit.io/library/api-reference/performance/st.cache
[Accessed 29 08 2022].

Tamrazian, A., Qian, Z. & Rajagopal, R., 2015. Where is My Parking Spot?: Online and Offline Prediction of Time-Varying Parking Occupancy. *Transportation Research Record: Journal of the Transportation Research Board,* 2489(1), pp. 77–85, doi: 10.3141/2489-09.

Tatulea, P. et al., 2019. An Image Feature-Based Method for Parking Lot Occupancy. *Future Internet,* Volume 11(8), 169, doi: 10.3390/fi11080169.

Tutika, C. et al., 2018. *Image Segmentation for Efficient Parking Space Analysis.* Chennai, India, 2018 IEEE International Conference on Informatics Computing in Engineering Systems (ICICES), doi: 10.48550/arXiv.1803.04620.

Valipour, S., Siam, M., Stroulia, E. & Jagersand, M., 2016. Parking Stall Vacancy Indicator System Based on Deep Convolutional Neural Networks. *2016 IEEE 3rd World Forum on Internet of Things (WF-IoT),* pp. 655-660, doi: 10.1109/WF-IoT.2016.7845408.

Wan, Z., Sperling, D. & Wang, Y., 2015. China's electric car frustrations. *Transportation Research Part D: Transport and Environment ,* Volume 34, pp. 116-121, doi: 10.1016/j.trd.2014.10.014.

Yusnita, R., Norbaya, F. & Basharuddin, N., 2012. Intelligent Parking Management System Based on Image Processing. *International Journal of Innovation, Management and Technology,* 3(3), pp. 232-235.

10.Bibliography

Aftab, K. et al., 2020. Reducing Parking Space Search Time and Environmental Impacts: A Technology Driven Smart Parking Case Study. *IEEE Technology and Society Magazine*, 39(3), pp. 62-75, doi: 10.1109/MTS.2020.3012329.

Ahrnbom, M., Åström, K. & Nilsson, M., 2016. Fast Classification of Empty and Occupied Parking Spaces Using Integral Channel Features. *2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 1609-1615, doi: 10.1109/CVPRW.2016.200.

Amato, G. et al., 2017. Deep Learning for Decentralized Parking Lot Occupancy Detection. *Expert Systems with Applications*, Volume 72, pp. 327-334, doi: 10.1016/j.eswa.2016.10.055.

Arnott, R., 2014. On the optimal target curbside parking occupancy rate. *Economics of Transportation*, 3(2), pp. 133-144, doi: 10.1016/j.ecotra.2014.07.001.

Balzano, W. & Vitale, F., 2017. DiG-Park: A Smart Parking Availability Searching Method Using V2V/V2I and DGP-Class Problem. *2017 31st International Conference on Advanced Information Networking and Applications Workshops (WAINA)*, pp. 698-703, doi: 10.1109/WAINA.2017.104.

Barker, J. & Rehman, S., 2019. Investigating the use of Machine Learning for Smart Parking Applications. *2019 11th International Conference on Knowledge and Systems Engineering (KSE)*, pp. 1-5, doi: 10.1109/KSE.2019.8919291.

Bonsall, P., 1991. The changing face of parking-related data collection and analysis: The role of new technologies. *Transportation*, Volume 18, pp. 83-106, doi: 10.1007/BF00150560.

Bukowski, M., Luckner, M. & Kunicki, R., 2020. Estimation of Free Space on Car Park Using Computer Vision Algorithms. *Advances in Intelligent Systems and Computing*, Volume 920, doi: 10.1007/978-3-030-13273-6_30.

Caicedo, F. & Diaz, A., 2013. Case analysis of simultaneous concessions of parking meters and underground parking facilities. *Transportation Research Part A: Policy and Practice*, Volume 49, pp. 358-378, doi: 10.1016/j.tra.2013.01.038.

Dance, C., 2014. Lean Smart Parking: How to cost-effectively collect high-quality data. *International Parking Institute*, pp. 26-29, available at: https://www.parking.org/wp-content/uploads/2016/01/TPP-2014-06-Lean-Smart-Parking.pdf.

Fabian, T., 2008. An Algorithm for Parking Lot Occupation Detection. *2008 7th Computer Information Systems and Industrial Management Applications*, pp. 165-170, doi: 10.1109/CISIM.2008.53.

Feng, N., Zhang, F., Lin, J., Zhai, J. & Du, X., 2019. Statistical Analysis and Prediction of Parking Behavior. *Lecture Notes in Computer Science*, Volume 11783, doi: 10.1007/978-3-030-30709-7_8.

Hattale, P. et al., 2019. Parking Space Detection Using Image Processing. *International Journal of Science and Research (IJSR)*, 10(3), pp. 1440-1442, doi: 10.21275/SR21321183644.

Jakob M. & Menendez, M., 2020. Parking Pricing vs. Congestion Pricing: A Macroscopic Analysis of their Impact on Traffic, *Preprints*, doi: 10.20944/preprints202006.0257.v1.

Jakob, M. et al., 2018. A dynamic macroscopic parking pricing and decision model. *Transportmetrica B: Transport Dynamics*, Volume 8, pp. 307-331, doi: 10.1080/21680566.2018.1488226.

Manville, M. & Pinski, M., 2021. The causes and consequences of curb parking management. *Transportation Research Part A: Policy and Practice*, Volume 152, pp. 295-307, doi: 10.1016/j.tra.2021.07.007.

Marsden, G., 2014. Parking Policy. *Parking Issues and Policies*, Volume 5, pp. 11-32, doi: 10.1108/S2044-994120140000005016.

Moylan, E. et al., 2019. Residential Permit Parking: Better Off Without It? *Transportation Research Record: Journal of the Transportation Research Board*, *2469*, pp. 23–31, doi:10.3141/2469-03.

Najmi, A. et al., 2021. On-street parking management and pricing policies: An evaluation from a system enhancement perspective. *Transportation Research Part A: Policy and Practice*, Volume 146, pp. 128-151, doi: 10.1016/j.tra.2021.02.009.

Sabir, S. & Anjum, G., 2017. Problems and Prospects of Curbside Parking in Lahore: Policy Implications for Effective Management. *Mehran University Research Journal of Engineering and Technology*, 36(4), pp. 867-880, doi: 10.22581/muet1982.1704.11.

Spurling, N., 2020. Parking futures: The relationship between parking space, everyday life and travel demand in the UK. *Land Use Policy*, Volume 91, 103872, doi: 10.1016/j.landusepol.2019.02.031.

Yamada, K. & Mizuno, M., 2001. A vehicle parking detection method using image segmentation. *Electronics and Communications in Japan (Part III: Fundamental Electronic Science)*, 84(10), pp. 25-34, doi: 10.1002/ecjc.1039.

11.Appendices

11.1.Appendix A – Datasets

Table 3: Datasets descriptions (1)

Dataset description /name	***Used for, alterations, Excel name***	***Original attribute name***	***New name***	***Description***	***Type***	***Was it used?***
pay-and-display ticket machine logs	♦*No alterations made to the original data itself.* ♦*Used for Location RR.* ♦**c_payanddisplay.csv**	*Date*	*[N/A]*	Dates between 01.12.2012-30.03.2012.	Date	Yes
		Machine	*Machine no.*	Machine number.	String	Yes
		Description	*Machine location*	Machine location.	String	No
		Tariff	*Tariff type*	Tariff type acronym.	String	Yes
		Description	*Tariff description*	The tariff description.	String	No
		Cash	*Cash paid (GBP)*	The amount actually paid for the ticket.	Float	Yes
		[inexistent]	*Tariff amount (GBP)*	*[created during data preparation and cleansing]* Expected cash paid (GBP).	Float	Yes
		[inexistent]	*Tariff time bracket*	*[created during data preparation and cleansing]* Tariff time bracket description based on tariff type.	String	Yes
		[inexistent]	*Leaving time*	*[created during data preparation and cleansing]* The ticket expiry date and time based on the paid amount.	Date	Yes
[duplicate] ***pay-and-display ticket machine logs***	♦*Alterations made to the original data (removal/add of rows)* ♦*Used for Location A.* ♦**c_payanddisplay_Loc_A.csv**	*[same as above]*	*[same as above]*	*[same as above]*	*[same as above]*	*[same as above]*

Table 4: Datasets descriptions (2)

Dataset description /name	***Used for, alterations, Excel name***	***Original attribute name***	***New name***	***Description***	***Type***	***Was it used?***
weather report	⬧*Not all the available data was collected from the website.* ⬧*No alterations made to the data itself.* ⬧*Used for both locations.* ⬧***c_temperature.csv***	*Date*	*[N/A]*	Dates between 01.12.2012-31.03.2012.	Date	Yes
		Location	*[N/A]*	The location from which the information applies.	String	No
		Mean temp (C)	*[N/A]*	The mean temperature registered that day.	Float	Yes
		High (C)	*[N/A]*	The highest temperature registered that day.	Float	Yes
		Time	*Date and time high temp*	The time when it was registered.	Date	No
		Low (C)	*[N/A]*	The lowest temperature registered that day.	Float	Yes
		Time	*Date and time low temp*	The time when it was registered.	Date	No
		Rain (mm)	*[N/A]*	Total rainfall depth in millimetres on that day.	Float	Yes
		Avg wind speed (mph)	*[N/A]*	The average wind speed in miles per hour of the day.	Float	Yes
		High (mph)	*[N/A]*	Highest wind speed.	Float	Yes
		Time	*Date and time high wind*	The time when it was registered.	Date	No
		[inexistent]	*Rainfall classification*	*[created during data preparation and cleansing]* Based on the *"rain (mm)"* it was determined what rainfall category will be.	String	Yes
		[inexistent]	*Wind classification*	*[created during data preparation and cleansing]* Based on the *"avg wind speed (mph)"* it was determined what wind category will be.	String	Yes

Table 5: Datasets descriptions (3)

Dataset description /name	Used for, alterations, Excel name	Original attribute name	New name	Description	Type	Was it used?
Bank /school holidays 2012-2013 in England	♦*Not all the available data was collected from the websites.* ♦*No alterations made to the information itself.* ♦*Used for both locations.* ♦**c_holidays.csv**	*[N/A]*	*Date*	*[created manually]* Dates between 01.12.2012-31.03.2012.	Date	Yes
		*Event**	*[N/A]*	*[based on true dataset + the gaps created manually]* What type of day/event is it (*weekend, bank holiday, weekday*). *If it was not *bank holiday* then it will be manually populated with the correct event.	String	Yes
		[N/A]	*Day*	*[created manually]* What day is it.	String	Yes
		School holiday?	*[N/A]*	Was it school holiday or not (*yes* or *no*).	String	Yes
EV charging transactions	♦*No alterations made to the original data.* ♦*Not used in comparison with location RR or A.* ♦*Used for electric vehicles analysis.* ♦**c_chargingpoint.csv**	*Charging event*	*[N/A]*	Charging event ID.	Int	Yes
		Chargepoint ID	*[N/A]*	ID of the charging point.	Int	Yes
		Borough	*[N/A]*	The district where is located.	String	No
		Operator	*[N/A]*	Operator of the charging station.	String	No
		Plug in Date and Time	*[N/A]*	When the charger was plugged.	Date	Yes
		Unplug Date and Time	*[N/A]*	When the charger was unplugged.	Date	Yes
		Charge start Date and Time	*[N/A]*	When the charging started.	Date	Yes
		Charge end Date and Time	*[N/A]*	When the charging ended.	Date	Yes
		Total kWh	*[N/A]*	Total kWh used.	Float	Yes
		[inexistent]	*Total time in min (Plug/Unplug)**	*[*calculated in the artefact]* Total time in minutes between plugging-in and unplugging.	Int	Yes
		[inexistent]	Total time in min (Charge start/end)*	*[*calculated in the artefact]* Total time in minutes between charging start and end.	Int	Yes
		[inexistent]	*Difference between plugging and charging**	*[*calculated in the artefact]* Difference between the above two columns.	Int	Yes

Table 6: Datasets descriptions (4)

Dataset description /name	***Used for, alterations, Excel name***	***Original attribute name***	***New name***	***Description***	***Type***	***Was it used?***
Valid tickets number	⬧*Made based on the original datasets for "pay-and-display ticket machine logs".* ⬧*Location RR will be based on the "pay-and-display ticket machine logs" for location RR and location A will be based on the dataset for location A.* ⬧***Loc_RR_pad_valid_tickets_no.csv*** ⬧***Loc_A_pad_valid_tickets_no.csv*** Creation approach can be found in 11.2.Appendix B – Valid Tickets.	*[N/A]*	*Date and time*	Every 10 minutes records, starting at 01/12/2012 06:30:00 AM and ending at 30/03/2013 10:00:00 PM.	Date	Yes
		[N/A]	*Valid tickets number*	Count of valid tickets number until that time.	Int	Yes
Video/CCTV count for disabled bays spaces	⬧*Made partially based on the video count and the rest based on almost randomly generated data.* ⬧*Location RR will be based on video RR and location A will be based on video A.* ⬧**Loc_RR_video_count_disabled.csv** ⬧**Loc_A_video_count_disabled.csv**	*[N/A]*	*Date and time*	Every 10 minutes records, starting at 01/12/2012 06:30:00 AM and ending at 30/03/2013 10:00:00 PM.	Date	Yes
		[N/A]	*Occupied spaces*	Number of occupied spaces.	Int	Yes
		[N/A]	*Location*	The location of the parking area.	String	No
		[N/A]	*Space type*	The space type (*disabled* or *regular*).	String	No

Table 7: Datasets descriptions (5)

Dataset description /name	***Used for, alterations, Excel name***	***Original attribute name***	***New name***	***Description***	***Type***	***Was it used?***
Video/CCTV count vs valid tickets number	♦*It uses the "valid tickets number" dataset and it combines it with a/the video count of occupancy (which is partially based on the video count and the rest almost randomly generated data).* ♦*Location RR combined to data from location RR and location A to data from location A.* ♦***Loc_RR_video_vs_pad_count_regular.csv*** ♦***Loc_A_video_vs_pad_count_regular.csv***	*[N/A]*	*Date and time*	Every 10 minutes records, starting at 01/12/2012 06:30:00 AM and ending at 30/03/2013 10:00:00 PM.	Date	Yes
		[N/A]	*CCTV count*	The number of occupied spaces based on the video/CCTV.	Int	Yes
		[N/A]	*Location*	The location of the parking area.	String	No
		[N/A]	*Space type*	The space type (*disabled* or *regular*).	String	No
		[N/A]	*Valid tickets number*	The number of valid tickets.	Int	Yes

11.2.Appendix B – Valid Tickets

Every 10 minutes, it was calculated the number of vehicles that could of being parked based on their *"leaving time"*.

Initially, the calculation logic and concepts were practised in Excel (*figure 47*) to gain an understanding of how to create an automatic calculation using Python.

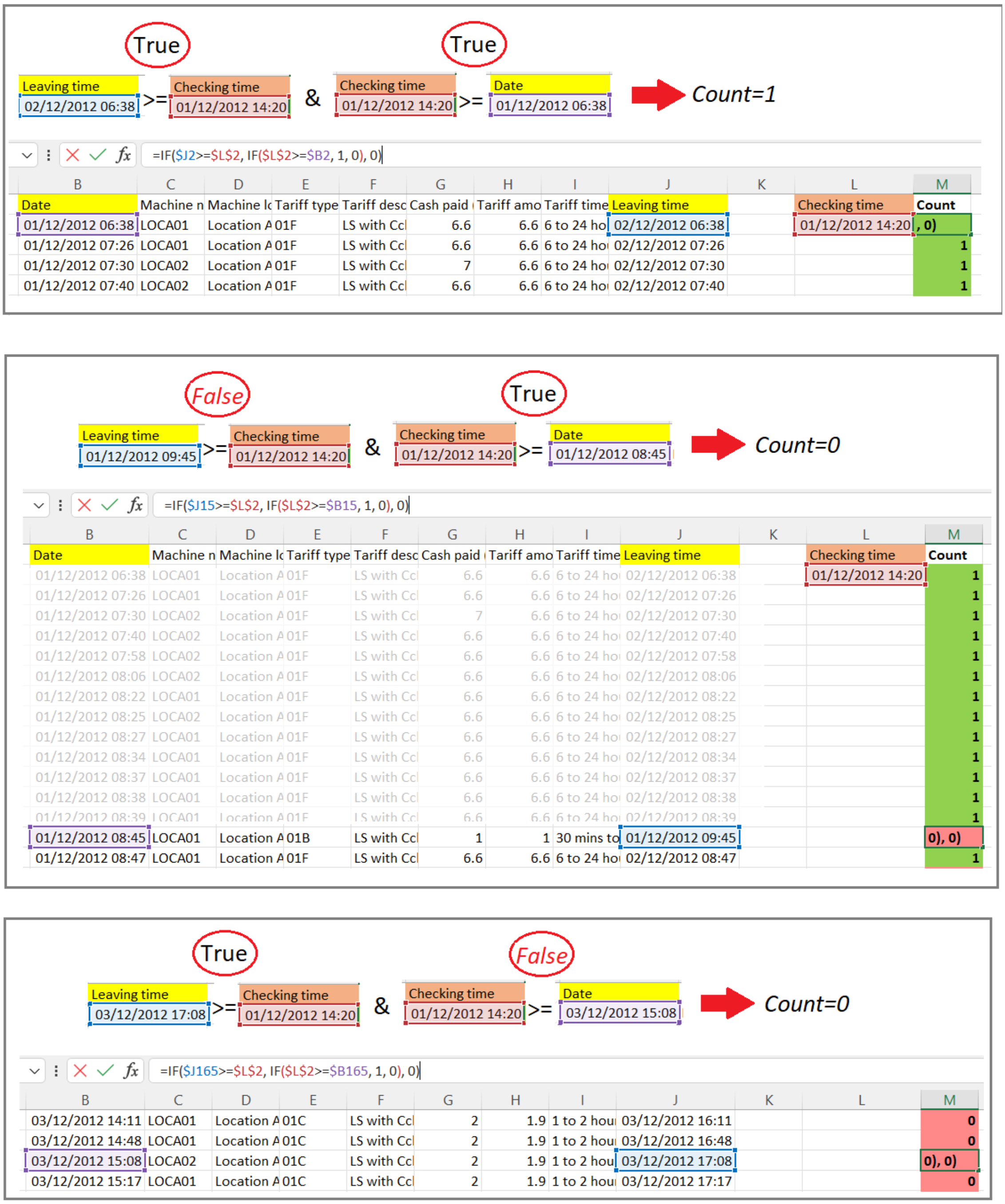


Figure 47: Valid tickets count in Excel

Then, these concepts are applied using Python to automatically calculate for each 10 minutes the number of *valid ticket counts* based on the pay-and-display dataset.

Figure 48, shows the code used to create the 10-minutes count:

```
time_plus_10 = datetime(2012, 12, 1, 6, 30, 0)

pad_valid_tickets = {"Date and time":[], "Valid tickets number":[]}

while time_plus_10 != datetime(2013, 3, 30, 22, 10, 0):
    count = 0
    for index in pad_data.index:
        if (pad_data["Leaving time"][index] >= time_plus_10) and (time_plus_10 >= pad_data["Date"][index]):
            count = count+1
        elif pad_data["Leaving time"][index] >= (time_plus_10 + timedelta(days=1, minutes = 10)):
            break
    pad_valid_tickets["Date and time"].append(time_plus_10)
    pad_valid_tickets["Valid tickets number"].append(count)
    time_plus_10 = time_plus_10 + timedelta(minutes = 10)
    print(time_plus_10)

pad_valid_tickets_no = pd.DataFrame(pad_valid_tickets)
```

```
2012-12-01 06:40:00
2012-12-01 06:50:00
```

Figure 48: Valid tickets count in Python

The *if statement* is based on the previous logic (*figure 47*) and the *elif statement* is added to reduce the computation time and because a ticket cannot be bought for longer than 24 hours.

The files used to create these counts can be found under the names of:

- *Count of legally parked cars in PAD dataset Loc RR.ipynb* → for location RR
- *Count of legally parked cars in PAD dataset Loc A.ipynb* → for location A

11.3.Appendix C – Consent Forms Example

SOLENT
UNIVERSITY

Informed Consent Form

Project Title: *Use of computer vision and predictive analysis for council's "pay-and-display" parking*

I confirm that (please tick as appropriate):

1.	I have been told about the purpose of the project and I understand this.	☑
2.	I have been given the opportunity to ask questions about the project and my participation.	☑
3.	I voluntarily agree to participate in the project.	☑
4.	I understand I can leave the project at any time without giving reasons and that I will not be questioned about why I have left the project.	☑
5.	The procedures regarding anonymity and confidentiality have been clearly explained to me (e.g. not using my real name, so that anything I contributed to this project cannot be recognized unless I give my consent; that only anonymised data will be shared outside the research team).	☑
6.	The procedures regarding data anonymity have been clearly explained to me (e.g. not using my real name, so that anything I contributed to this project cannot be recognized).	☑
7.	I agree to the use of voice recording if telephone, skype or in-person interviews are used.	☑
8.	The use of the data in research, publications, sharing and archiving has been explained to me.	☑
9.	I understand that other researchers will have access to this data only if they agree to preserve the confidentiality of the data and if they agree to the terms I have specified in this form.	☑
10.	I agree to the use of direct quotations in publications provided that my anonymity is preserved.	☑
11.	I understand what I have said or written as part of this project will be used in reports, publications and other research outputs.	☑
12.	I, along with the Researcher, agree to sign and date this informed consent form.	☑

Participant: Name & Signature ____________________

Date: ____________

Researcher: Name & Signature ____________________

Date: ____________

Figure 49: Informed Consent Form

SOLENT
UNIVERSITY

Consent Form

I agree to participate in the usability study conducted by Solent University.

I understand that participation in this usability study is voluntary and I agree to immediately raise any concerns or areas of discomfort I might have with the study administrator.

We are going to record what happens on the screen and our conversations. The recordings will be used to help us to improve the site.

Please sign below to indicate that you have read and you understand the information on this form and that any questions you might have about the session have been answered.

We would appreciate it if the information you see could be kept confidential.

Date: ________________

Name: ____

Please sign your name: ____

Thank you!

We appreciate your participation.

Figure 50: Usability Testing Consent Form

11.4.Appendix D – Ethical Clearance

Ethical clearance for research and innovation projects

Project status

Status

Approved

Actions

Date	Who	Action	Comments
10:46:00 22 August 2022	Drishty Sobnath	Supervisor approved	
12:34:00 16 August 2022	Denise Tineghe	Principal investigator submitted	

Get Help

Ethics release checklist (ERC)

Project details

Project name:	Use of computer vision and predictive analysis for council's "pay-and-display" parking
Principal investigator:	Denise Tineghe
Faculty:	Faculty of Business, Law and Digital Technologies
Level:	Postgraduate
Course:	MSc Applied AI and Data Science
Unit code:	COM726
Supervisor name:	Drishty Sobnath
Other investigators:	N/A

Figure 51: Ethical Clearance (1)

Checklist

Question	Yes	No
Q1. Will the project involve human participants other than the investigator(s)?	◉	○
Q1a. Will the project involve **vulnerable participants** such as children, young people, disabled people, the elderly, people with declared mental health issues, prisoners, people in health or social care settings, addicts, or those with learning difficulties or cognitive impairment either contacted directly or via a **gatekeeper** (for example a professional who runs an organisation through which participants are accessed; a service provider; a care-giver; a relative or a guardian)?	○	◉
Q1b. Will the project involve the use of **control groups** or the **use of deception**?	○	◉
Q1c. Will the project involve any **risk to the participants' health** (e.g. intrusive intervention such as the administration of drugs or other substances, or vigorous physical exercise), or involve psychological stress, anxiety, humiliation, physical pain or discomfort to the investigator(s) and/or the participants?	○	◉
Q1d. Will the project involve **financial inducement** offered to participants other than reasonable expenses and compensation for time?	○	◉
Q1e. Will the project be carried out by individuals unconnected with the University but who wish to use staff and/or students of the University as participants?	○	◉
Q2. Will the project involve sensitive materials or topics that might be considered offensive, distressing, politically or socially sensitive, deeply personal or in breach of the law (for example criminal activities, sexual behaviour, ethnic status, personal appearance, experience of violence, addiction, religion, or financial circumstances)?	○	◉
Q3. Will the project have detrimental impact on the environment, habitat or species?	○	◉
Q4. Will the project involve living animal subjects?	○	◉
Q5. Will the project involve the development for export of 'controlled' goods regulated by the Export Control Organisation (ECO)? (This specifically means military goods, so called dual-use goods (which are civilian goods but with a potential military use or application), products used for torture and repression, radioactive sources.) Further information from the Export Control Organisation	○	◉
Q6. Does your research involve: the storage of records on a computer, electronic transmissions, or visits to websites, which are associated with terrorist or extreme groups or other security sensitive material? Further information from the Information Commissioners Office	○	◉

Declarations

I/we, the investigator(s), confirm that:

☑ The information contained in this checklist is correct.

Figure 52: Ethical Clearance (2)

- [x] I/we have assessed the ethical considerations in relation to the project in line with the University Ethics Policy.

- [x] I/we understand that the ethical considerations of the project will need to be re-assessed if there are any changes to it.

- [x] I/we will endeavor to preserve the reputation of the University and protect the health and safety of all those involved when conducting this research/enterprise project.

- [x] If personal data is to be collected as part of my project, I confirm that my project and I, as Principal Investigator, will adhere to the General Data Protection Regulation (GDPR) and the Data Protection Act 2018. I also confirm that I will seek advice on the DPA, as necessary, by referring to the Information Commissioner's Office further guidance on DPA and/or by contacting information.rights@solent.ac.uk. By Personal data, I understand any data that I will collect as part of my project that can identify an individual, whether in personal or family life, business or profession.

- [x] I/we have read the prevent agenda.

Figure 53: Ethical Clearance (3)

11.5.Appendix E – LSTM Models

The creation of all the LSTM models.

```
[Saved]LSTM_Model_Loc_RR_30min_valid_tickets Last Checkpoint: 16/08/2022 (autosaved)

View   Insert   Cell   Kernel   Widgets   Help
Run   Markdown

#create the model
model = Sequential()
model.add(LSTM(units=64, input_shape=(x_train.shape[1],1), return_sequences=True))
model.add(LSTM(units=128, return_sequences=True))
model.add(LSTM(units=16, return_sequences=False))
model.add(Dropout(0.1))
model.add(Dense(units=1, activation='linear'))

optimizer = keras.optimizers.Adam(learning_rate=0.0001) #, epsilon=0.000001

model.compile(optimizer='adam', loss='mse', metrics=[RootMeanSquaredError()])
model.summary()

history=model.fit(x_train, y_train, batch_size=32, epochs=40, validation_split=0.2, verbose=1)

Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 lstm (LSTM)                 (None, 60, 64)            16896

 lstm_1 (LSTM)               (None, 60, 128)           98816

 lstm_2 (LSTM)               (None, 16)                9280

 dropout (Dropout)           (None, 16)                0

 dense (Dense)               (None, 1)                 17

=================================================================
Total params: 125,009
Trainable params: 125,009
Non-trainable params: 0
```

Figure 54: LSTM model (Location RR, 30 minutes prediction based on valid tickets)

```
[Saved]LSTM_Model_Loc_RR_1h_valid_tickets Last Checkpoint: 08/17/2022 (autosaved)

View   Insert   Cell   Kernel   Widgets   Help
Run   Markdown

#create the model
model = Sequential()
model.add(LSTM(units=64, input_shape=(x_train.shape[1],1), return_sequences=True))
model.add(LSTM(units=258, return_sequences=True))
model.add(LSTM(units=64, return_sequences=False))
model.add(Dropout(0.1))
model.add(Dense(units=1, activation='linear'))

optimizer = keras.optimizers.Adam(learning_rate=0.0001) #, epsilon=0.000001

model.compile(optimizer='adam', loss='mse', metrics=[RootMeanSquaredError()])
model.summary()

history=model.fit(x_train, y_train, batch_size=32, epochs=40, validation_split=0.2, verbose=1)

Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 lstm (LSTM)                 (None, 60, 64)            16896

 lstm_1 (LSTM)               (None, 60, 258)           333336

 lstm_2 (LSTM)               (None, 64)                82688

 dropout (Dropout)           (None, 64)                0

 dense (Dense)               (None, 1)                 65

=================================================================
Total params: 432,985
Trainable params: 432,985
Non-trainable params: 0
```

Figure 55: LSTM model (Location RR, 1 hour prediction based on valid tickets)

[Saved]LSTM_Model_Loc_RR_2h_valid_tickets Last Checkpoint: 08/17/2022 (autosaved)

View Insert Cell Kernel Widgets Help

Run Code

```
#create the model
model = Sequential()
model.add(LSTM(units=64, input_shape=(x_train.shape[1],1), return_sequences=True))
model.add(LSTM(units=258, return_sequences=True))
model.add(LSTM(units=32, return_sequences=False))
model.add(Dropout(0.1))
model.add(Dense(units=1, activation='linear'))
```

```
optimizer = keras.optimizers.Adam(learning_rate=0.0001) #, epsilon=0.000001

model.compile(optimizer='adam', loss='mse', metrics=[RootMeanSquaredError()])
model.summary()

history=model.fit(x_train, y_train, batch_size=32, epochs=40, validation_split=0.2, verbose=1)
```

```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 lstm (LSTM)                 (None, 60, 64)            16896

 lstm_1 (LSTM)               (None, 60, 258)           333336

 lstm_2 (LSTM)               (None, 32)                37248

 dropout (Dropout)           (None, 32)                0

 dense (Dense)               (None, 1)                 33

=================================================================
Total params: 387,513
Trainable params: 387,513
Non-trainable params: 0
```

Figure 56: LSTM model (Location RR, 2 hours prediction based on valid tickets)

[Saved]LSTM_Model_Loc_RR_10min_video_disabled Last Checkpoint: 16/08/2022 (autosaved)

View Insert Cell Kernel Widgets Help

Run Markdown

```
#create the model
model = Sequential()
model.add(LSTM(units=64, input_shape=(x_train.shape[1],1), return_sequences=True))
model.add(LSTM(units=128, return_sequences=True))
model.add(LSTM(units=16, return_sequences=False))
model.add(Dropout(0.1))
model.add(Dense(units=1, activation='linear'))
```

```
optimizer = keras.optimizers.Adam(learning_rate=0.0001) #, epsilon=0.000001

model.compile(optimizer='adam', loss='mse', metrics=[RootMeanSquaredError()])
model.summary()

history=model.fit(x_train, y_train, batch_size=32, epochs=40, validation_split=0.2, verbose=1)
```

```
Model: "sequential_2"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 lstm_6 (LSTM)               (None, 60, 64)            16896

 lstm_7 (LSTM)               (None, 60, 128)           98816

 lstm_8 (LSTM)               (None, 16)                9280

 dropout_2 (Dropout)         (None, 16)                0

 dense_2 (Dense)             (None, 1)                 17

=================================================================
Total params: 125,009
Trainable params: 125,009
Non-trainable params: 0
```

Figure 57: LSTM model (Location RR, 10 minutes prediction based on CCTV/video disabled bay spaces count)

[Saved]LSTM_Model_Loc_RR_30min_video_disabled Last Checkpoint: 16/08/2022 (autosaved)

View Insert Cell Kernel Widgets Help

Run Markdown

```
#create the model
model = Sequential()
model.add(LSTM(units=64, input_shape=(x_train.shape[1],1), return_sequences=True))
model.add(LSTM(units=128, return_sequences=True))
model.add(LSTM(units=2, return_sequences=False))
model.add(Dropout(0.1))
model.add(Dense(units=1, activation='linear'))
```

```
optimizer = keras.optimizers.Adam(learning_rate=0.0001) #, epsilon=0.000001

model.compile(optimizer='adam', loss='mse', metrics=[RootMeanSquaredError()])
model.summary()

history=model.fit(x_train, y_train, batch_size=32, epochs=40, validation_split=0.2, verbose=1)
```

```
Model: "sequential_2"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 lstm_6 (LSTM)               (None, 60, 64)            16896

 lstm_7 (LSTM)               (None, 60, 128)           98816

 lstm_8 (LSTM)               (None, 2)                 1048

 dropout_2 (Dropout)         (None, 2)                 0

 dense_2 (Dense)             (None, 1)                 3

=================================================================
Total params: 116,763
Trainable params: 116,763
Non-trainable params: 0
```

Figure 58: LSTM model (Location RR, 30 minutes prediction based on CCTV/video disabled bay spaces count)

[Saved]LSTM_Model_Loc_RR_1h_video_disabled Last Checkpoint: 17/08/2022 (autosaved)

View Insert Cell Kernel Widgets Help

Run Markdown

```
#create the model
model = Sequential()
model.add(LSTM(units=64, input_shape=(x_train.shape[1],1), return_sequences=True))
model.add(LSTM(units=128, return_sequences=True))
model.add(LSTM(units=8, return_sequences=False))
model.add(Dropout(0.1))
model.add(Dense(units=1, activation='linear'))
```

```
optimizer = keras.optimizers.Adam(learning_rate=0.0001) #, epsilon=0.000001

model.compile(optimizer='adam', loss='mse', metrics=[RootMeanSquaredError()])
model.summary()

history=model.fit(x_train, y_train, batch_size=32, epochs=40, validation_split=0.2, verbose=1)
```

```
Model: "sequential_1"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 lstm_3 (LSTM)               (None, 60, 64)            16896

 lstm_4 (LSTM)               (None, 60, 128)           98816

 lstm_5 (LSTM)               (None, 8)                 4384

 dropout_1 (Dropout)         (None, 8)                 0

 dense_1 (Dense)             (None, 1)                 9

=================================================================
Total params: 120,105
Trainable params: 120,105
Non-trainable params: 0
```

Figure 59: LSTM model (Location RR, 1 hour prediction based on CCTV/video disabled bay spaces count)

[Saved]LSTM_Model_Loc_A_10min_video_disabled Last Checkpoint: 16/08/2022 (autosaved)

View Insert Cell Kernel Widgets Help

Run Markdown

```
#create the model
model = Sequential()
model.add(LSTM(units=64, input_shape=(x_train.shape[1],1), return_sequences=True))
model.add(LSTM(units=258, return_sequences=True))
model.add(LSTM(units=32, return_sequences=False))
model.add(Dropout(0.1))
model.add(Dense(units=1, activation='linear'))
```

```
optimizer = keras.optimizers.Adam(learning_rate=0.0001) #, epsilon=0.000001

model.compile(optimizer='adam', loss='mse', metrics=[RootMeanSquaredError()])
model.summary()

history=model.fit(x_train, y_train, batch_size=32, epochs=40, validation_split=0.2, verbose=1)
```

```
Model: "sequential_5"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 lstm_15 (LSTM)              (None, 60, 64)            16896

 lstm_16 (LSTM)              (None, 60, 258)           333336

 lstm_17 (LSTM)              (None, 32)                37248

 dropout_5 (Dropout)         (None, 32)                0

 dense_5 (Dense)             (None, 1)                 33

=================================================================
Total params: 387,513
Trainable params: 387,513
Non-trainable params: 0
```

Figure 60: LSTM model (Location A, 10 minutes prediction based on CCTV/video disabled bay spaces count)

11.6.Appendix F – LSTM Models Experimentation

Initial experimentation to obtain the best results were made for **Location RR, for 10 minutes prediction based on *valid ticket counts.***

The below figures show the *training/validation loss* results and based on the best representation the values were chosen.

Note: These are not all the tests, as previous experimentations were also done by changing the values of the *epsilon* and *learning rate*.

Figure 61: Experimentation (1)

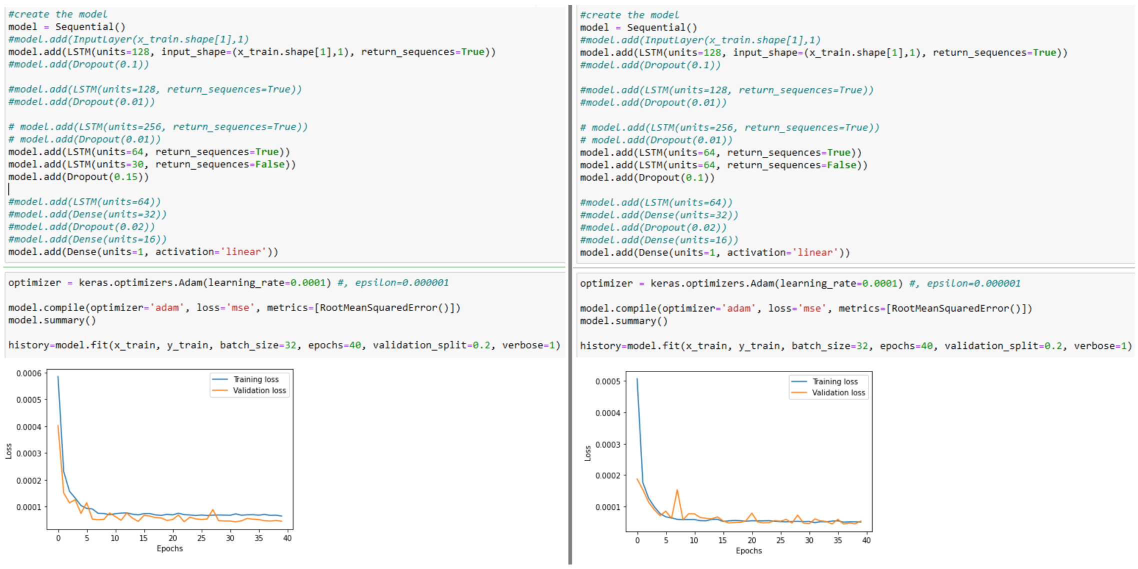


Figure 62: Experimentation (2)

```python
#create the model
model = Sequential()
model.add(LSTM(units=64, input_shape=(x_train.shape[1],1), return_sequences=True))
model.add(LSTM(units=128, return_sequences=True))
model.add(LSTM(units=128, return_sequences=True))
model.add(LSTM(units=64, return_sequences=False))
model.add(Dropout(0.1))
model.add(Dense(units=1, activation='linear'))

optimizer = keras.optimizers.Adam(learning_rate=0.0001) #, epsilon=0.000001

model.compile(optimizer='adam', loss='mse', metrics=[RootMeanSquaredError()])
model.summary()

history=model.fit(x_train, y_train, batch_size=32, epochs=40, validation_split=0.2, verbose=1)
```

```python
#create the model
model = Sequential()
model.add(LSTM(units=64, input_shape=(x_train.shape[1],1), return_sequences=True))
model.add(LSTM(units=128, return_sequences=True))
#model.add(LSTM(units=128, return_sequences=True))
model.add(LSTM(units=64, return_sequences=False))
model.add(Dropout(0.1))
model.add(Dense(units=1, activation='linear'))

#optimizer = keras.optimizers.Adam(learning_rate=0.0001) #, epsilon=0.000001

model.compile(optimizer='adam', loss='mse', metrics=[RootMeanSquaredError()])
model.summary()

history=model.fit(x_train, y_train, batch_size=32, epochs=40, validation_split=0.2, verbose=1)
```

Figure 63: Experimentation (3)

```python
#create the model
model = Sequential()
model.add(LSTM(units=64, input_shape=(x_train.shape[1],1), return_sequences=True))
model.add(LSTM(units=128, return_sequences=True))
#model.add(LSTM(units=128, return_sequences=True))
model.add(LSTM(units=64, return_sequences=False))
model.add(Dropout(0.1))
model.add(Dense(units=1, activation='linear'))

optimizer = keras.optimizers.Adam(learning_rate=0.0001) #, epsilon=0.000001

model.compile(optimizer='adam', loss='mse', metrics=[RootMeanSquaredError()])
model.summary()

history=model.fit(x_train, y_train, batch_size=32, epochs=40, validation_split=0.2, verbose=1)
```

```python
#create the model
model = Sequential()
model.add(LSTM(units=64, input_shape=(x_train.shape[1],1), return_sequences=True))
model.add(LSTM(units=128, return_sequences=True))
model.add(LSTM(units=128, return_sequences=False))
model.add(Dropout(0.1))
model.add(Dense(units=1, activation='linear'))

optimizer = keras.optimizers.Adam(learning_rate=0.0001) #, epsilon=0.000001

model.compile(optimizer='adam', loss='mse', metrics=[RootMeanSquaredError()])
model.summary()

history=model.fit(x_train, y_train, batch_size=32, epochs=40, validation_split=0.2, verbose=1)
```

Figure 64: Experimentation (4)

```python
#create the model
model = Sequential()
model.add(LSTM(units=64, input_shape=(x_train.shape[1],1), return_sequences=True))
model.add(LSTM(units=256, return_sequences=True))
model.add(LSTM(units=64, return_sequences=False))
model.add(Dropout(0.1))
model.add(Dense(units=1, activation='linear'))

optimizer = keras.optimizers.Adam(learning_rate=0.0001) #, epsilon=0.000001

model.compile(optimizer='adam', loss='mse', metrics=[RootMeanSquaredError()])
model.summary()

history=model.fit(x_train, y_train, batch_size=32, epochs=40, validation_split=0.2, verbose=1)
```

```python
#create the model
model = Sequential()
model.add(LSTM(units=64, input_shape=(x_train.shape[1],1), return_sequences=True))
model.add(LSTM(units=135, return_sequences=True))
model.add(LSTM(units=64, return_sequences=False))
model.add(Dropout(0.1))
model.add(Dense(units=1, activation='linear'))

optimizer = keras.optimizers.Adam(learning_rate=0.0001) #, epsilon=0.000001

model.compile(optimizer='adam', loss='mse', metrics=[RootMeanSquaredError()])
model.summary()

history=model.fit(x_train, y_train, batch_size=32, epochs=40, validation_split=0.2, verbose=1)
```

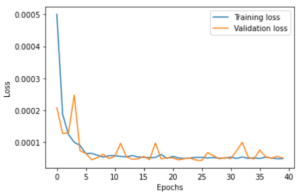
Figure 65: Experimentation (5)

```python
#create the model
model = Sequential()
model.add(LSTM(units=64, input_shape=(x_train.shape[1],1), return_sequences=True))
model.add(LSTM(units=130, return_sequences=True))
model.add(LSTM(units=64, return_sequences=False))
model.add(Dropout(0.1))
model.add(Dense(units=1, activation='linear'))
```

```python
optimizer = keras.optimizers.Adam(learning_rate=0.0001) #, epsilon=0.000001

model.compile(optimizer='adam', loss='mse', metrics=[RootMeanSquaredError()])
model.summary()

history=model.fit(x_train, y_train, batch_size=32, epochs=40, validation_split=0.2, verbose=1)
```




```python
#create the model
model = Sequential()
model.add(LSTM(units=64, input_shape=(x_train.shape[1], 1), return_sequences=True))
model.add(LSTM(units=128, return_sequences=True))
model.add(LSTM(units=64, return_sequences=False))
model.add(Dropout(0.1))
model.add(Dense(units=1, activation='linear'))
```

```python
optimizer = keras.optimizers.Adam(learning_rate=0.0001) #, epsilon=0.000001

model.compile(optimizer='adam', loss='mse', metrics=[RootMeanSquaredError()])
model.summary()

history=model.fit(x_train, y_train, batch_size=32, epochs=60, validation_split=0.2, verbose=1)
```

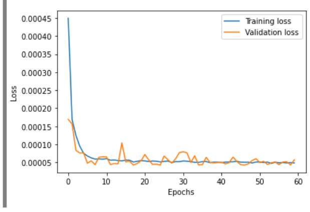


Figure 66: Experimentation (6)

Experimentations were made for Location A, 30 minutes prediction based on *CCTV/video disabled bay spaces count*. The results were unsatisfactory. Therefore, the model was not implemented within the artefact.

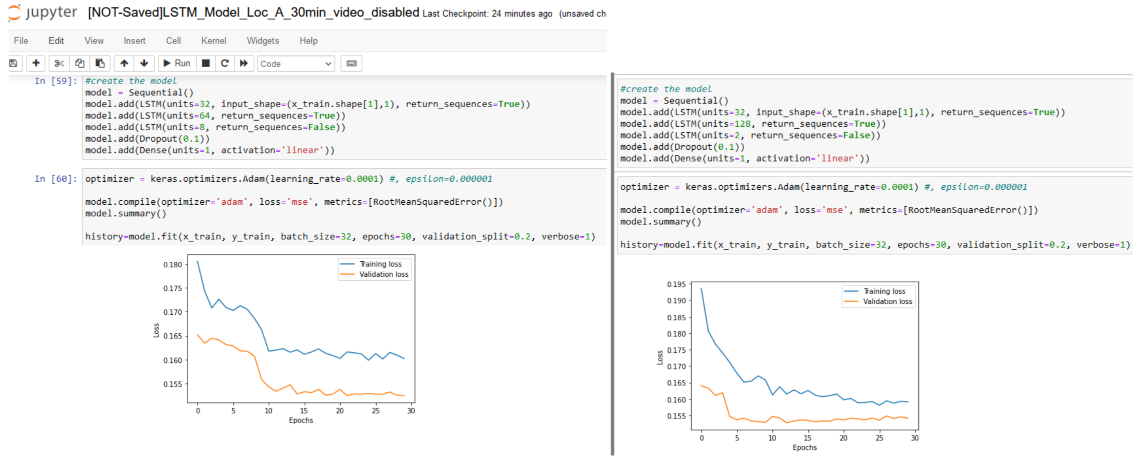


Figure 67: Experimentations, Location A (1)

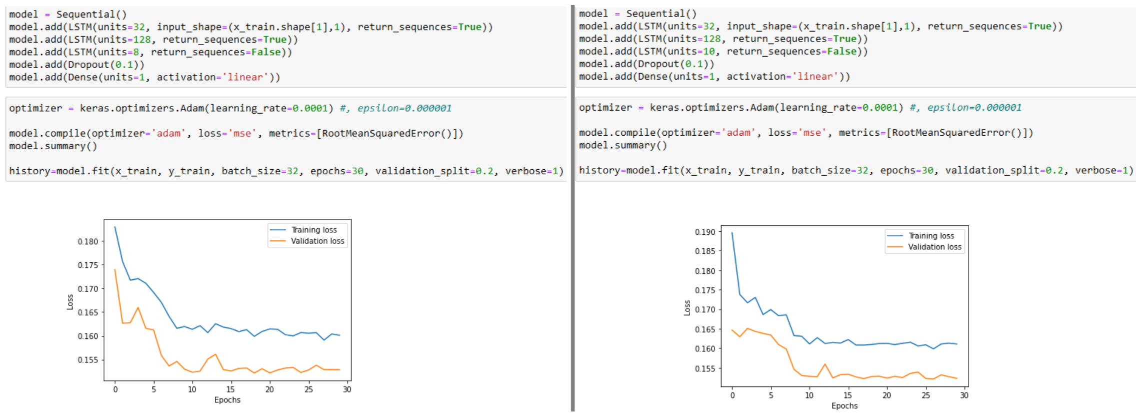


Figure 68: Experimentations, Location A (2)

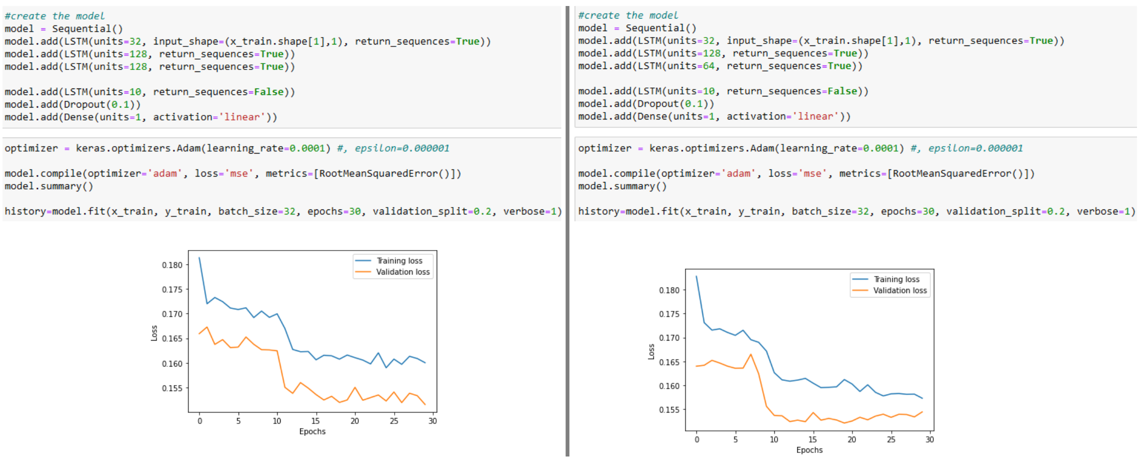


Figure 69: Experimentations, Location A (3)

11.7.Appendix G – Other Models Testing

Folder *"Tests"* contains different tests made to better decide on the chosen models.

The *Prophet model* was tested, but the results were unsatisfactory.

```
from sklearn.metrics import mean_absolute_error, mean_squared_error, r2_score
mse = mean_squared_error(test['y'], forecast[15507:]['yhat'])
rmse = math.sqrt(mse)
mae=mean_absolute_error(test['y'],forecast[15507:]['yhat'])
r_two=r2_score(test['y'],forecast[15507:]['yhat'])

print("Root Mean Squared Error (RMSE): ", round(rmse,2))
print("Mean absolute error (MAE): ", round(mae,2))
print("Mean square error (MSE): ", round(mse,2))
print("R2-score:", round(r_two,2))
```

```
Root Mean Squared Error (RMSE):  22.2
Mean absolute error (MAE):  10.49
Mean square error (MSE):  493.02
R2-score: -0.1
```

Figure 70: Prophet model

A *linear regression* model was created, but no good results were obtained.

```
mse = mean_squared_error(y_test, predictions)
rmse = math.sqrt(mse)
mae=mean_absolute_error(y_test,predictions)
r_two=r2_score(y_test,predictions)

print("Root Mean Squared Error (RMSE): ", round(rmse,2))
print("Mean absolute error (MAE): ", round(mae,2))
print("Mean square error (MSE): ", round(mse,2))
print("R2-score:", round(r_two,2))
```

```
Root Mean Squared Error (RMSE):  16.36
Mean absolute error (MAE):  8.5
Mean square error (MSE):  267.7
R2-score: 0.01
```

Figure 71: Linear regression model

Therefore, those models were not used further.

11.8.Appendix H – Artefact

Video of the artefact:

https://drive.google.com/file/d/1AyzLcBuO44tgXQVXGZ_S1N6NF68nFBbk/view?usp=sharing

Council Staff → Location RR → In-depth analysis *(figures 72 to 75)*

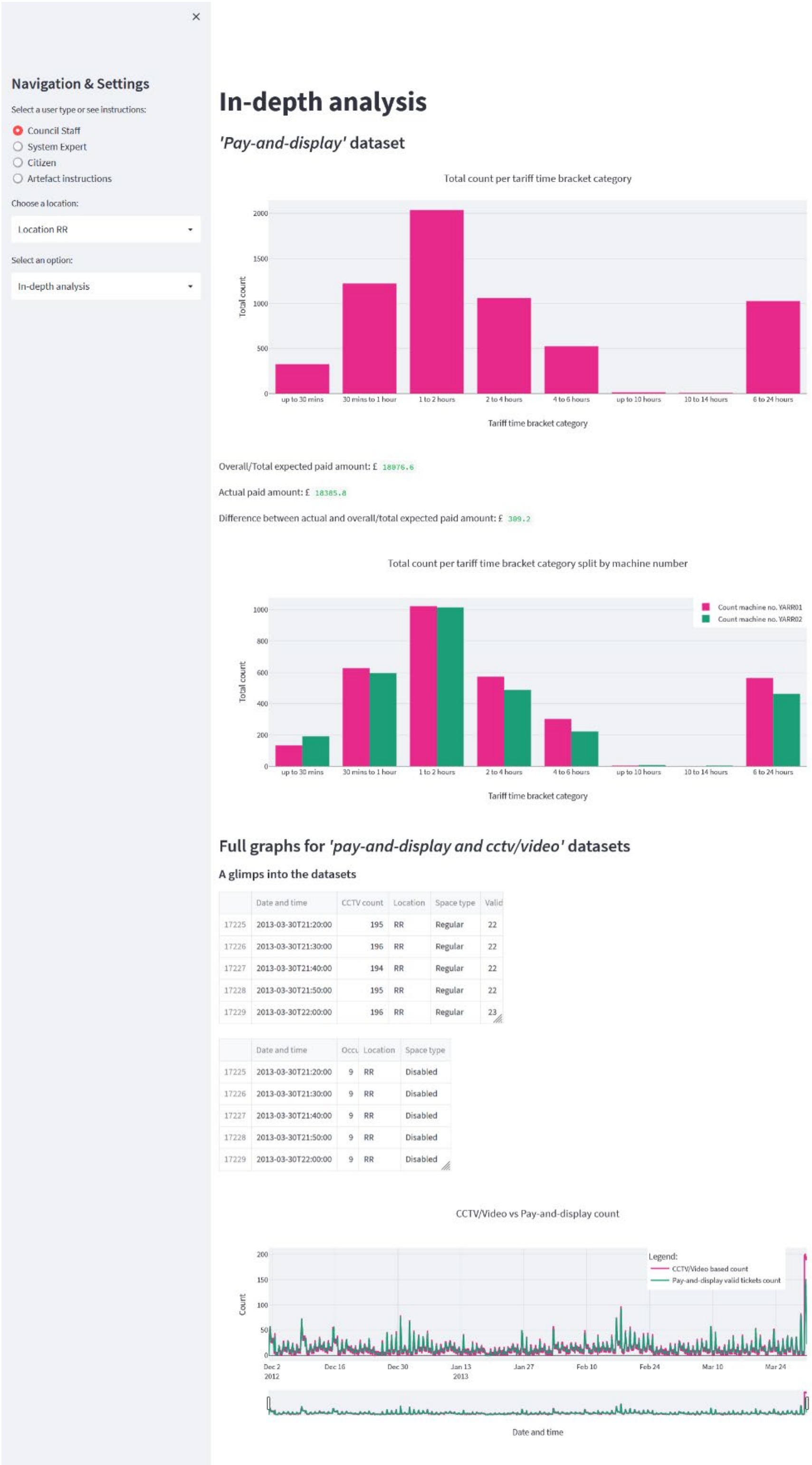
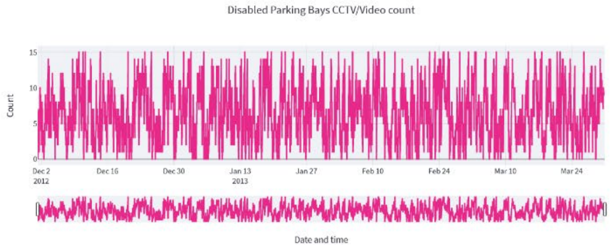
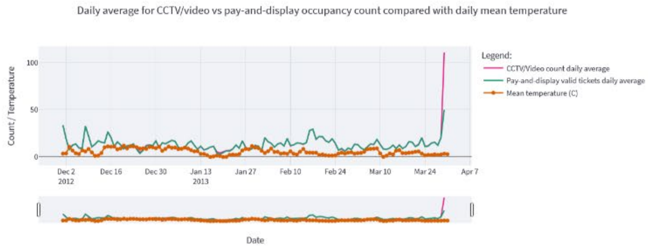
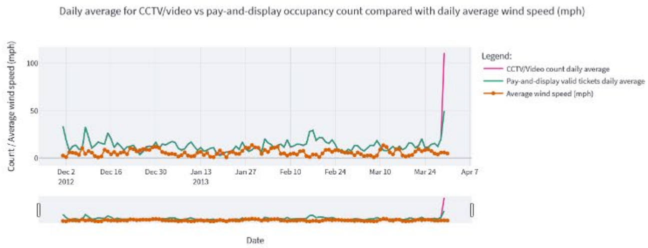


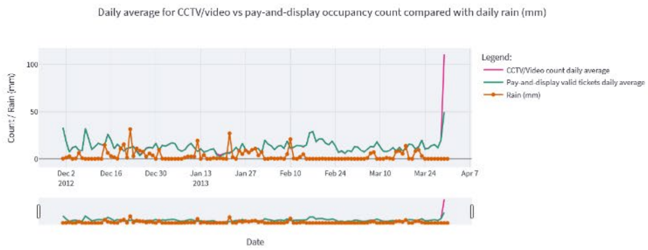
Figure 72: Council Staff → Location RR → In-depth analysis (1)

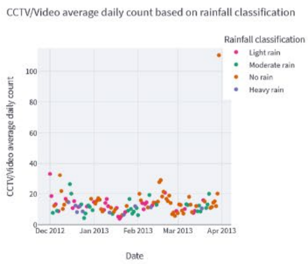


Average count per day for *'pay-and-display and the cctv/video'* data compared with *'temperature'* data









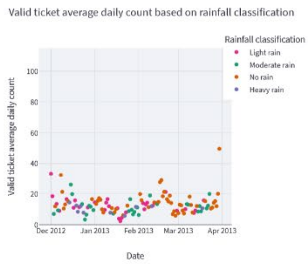


Figure 73: Council Staff → Location RR → In-depth analysis (2)

Figure 74: Council Staff → Location RR → In-depth analysis (3)

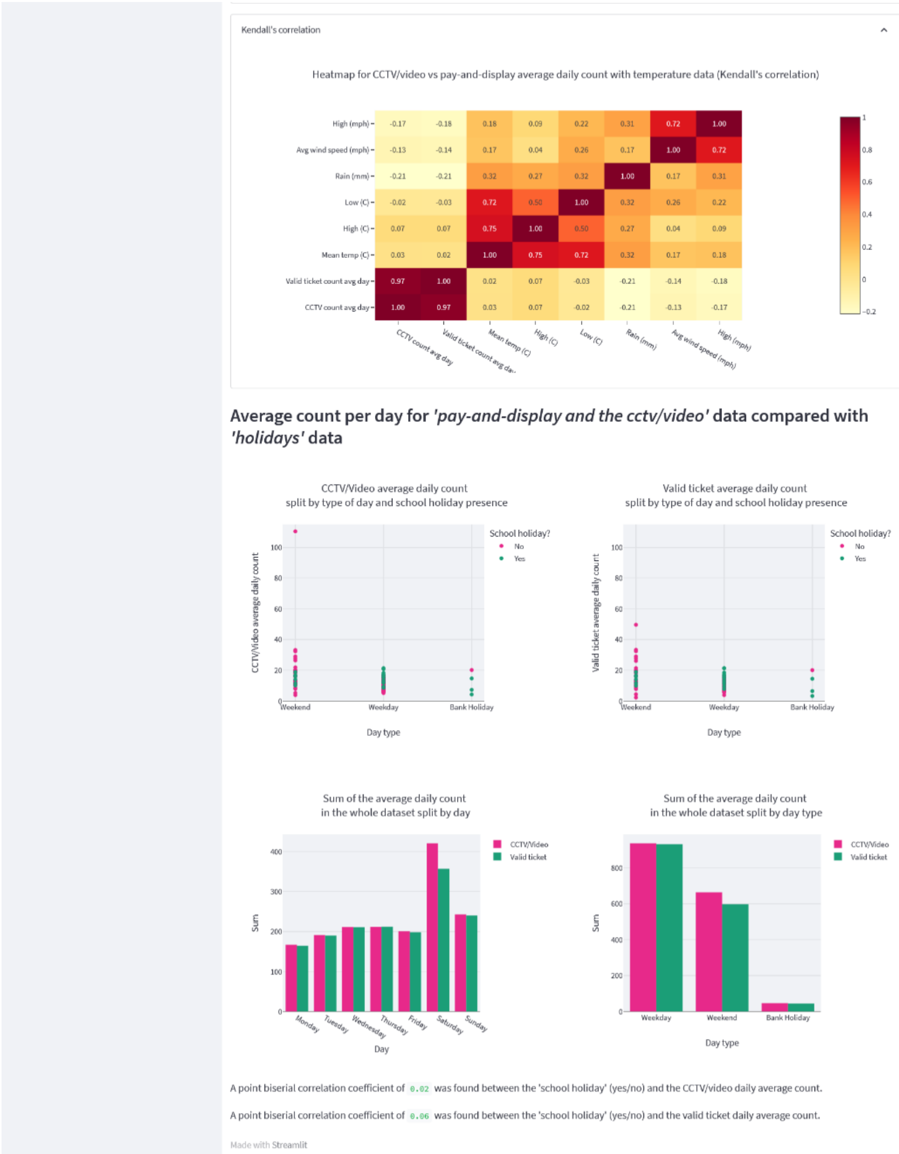


Figure 75: Council Staff → Location RR → In-depth analysis (4)

Council Staff → Location RR → Predictions *(figure 76)*

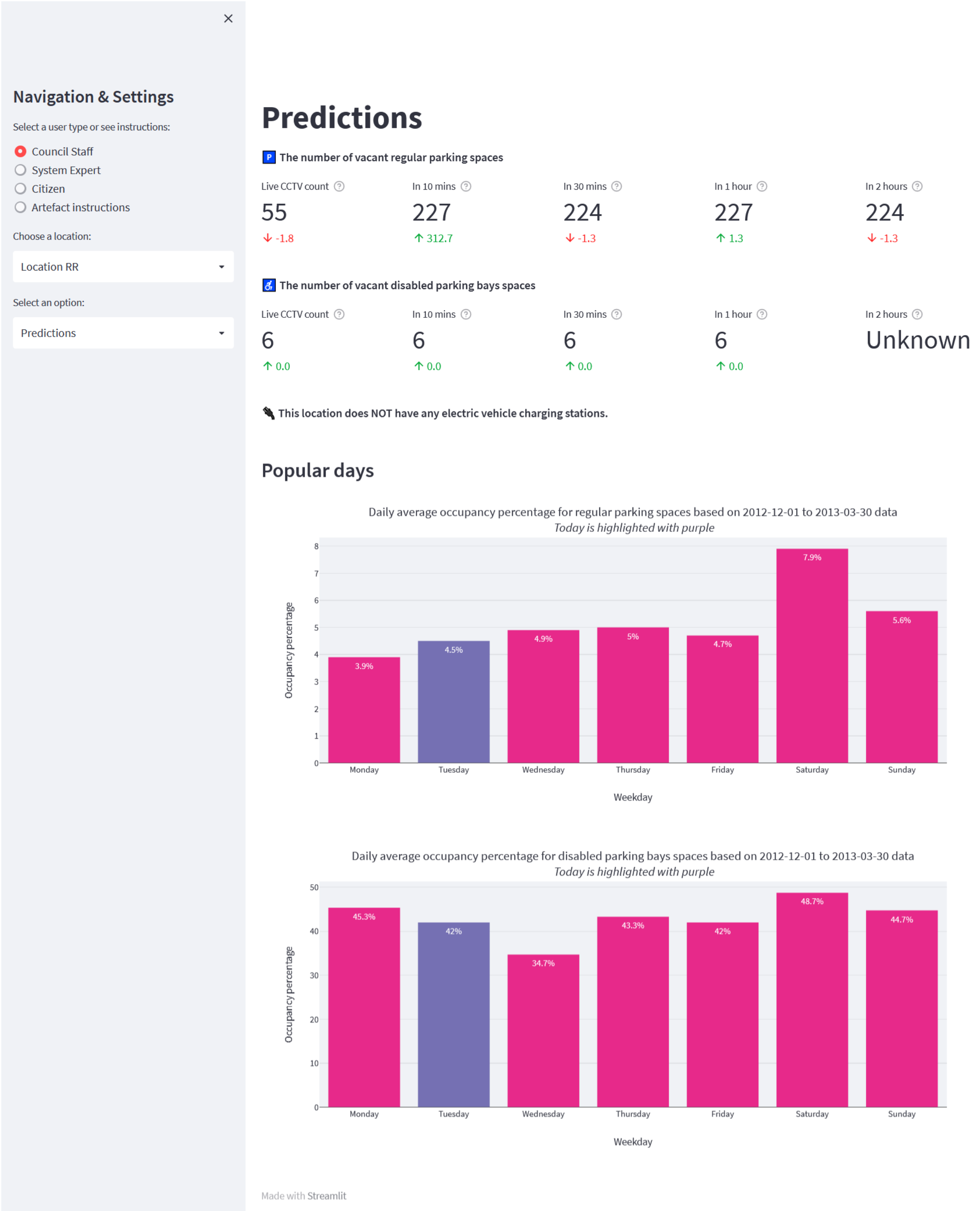


Figure 76: Council Staff → Location RR → Predictions

Council Staff → Location A → 24h Quick summary *(figure 77)*

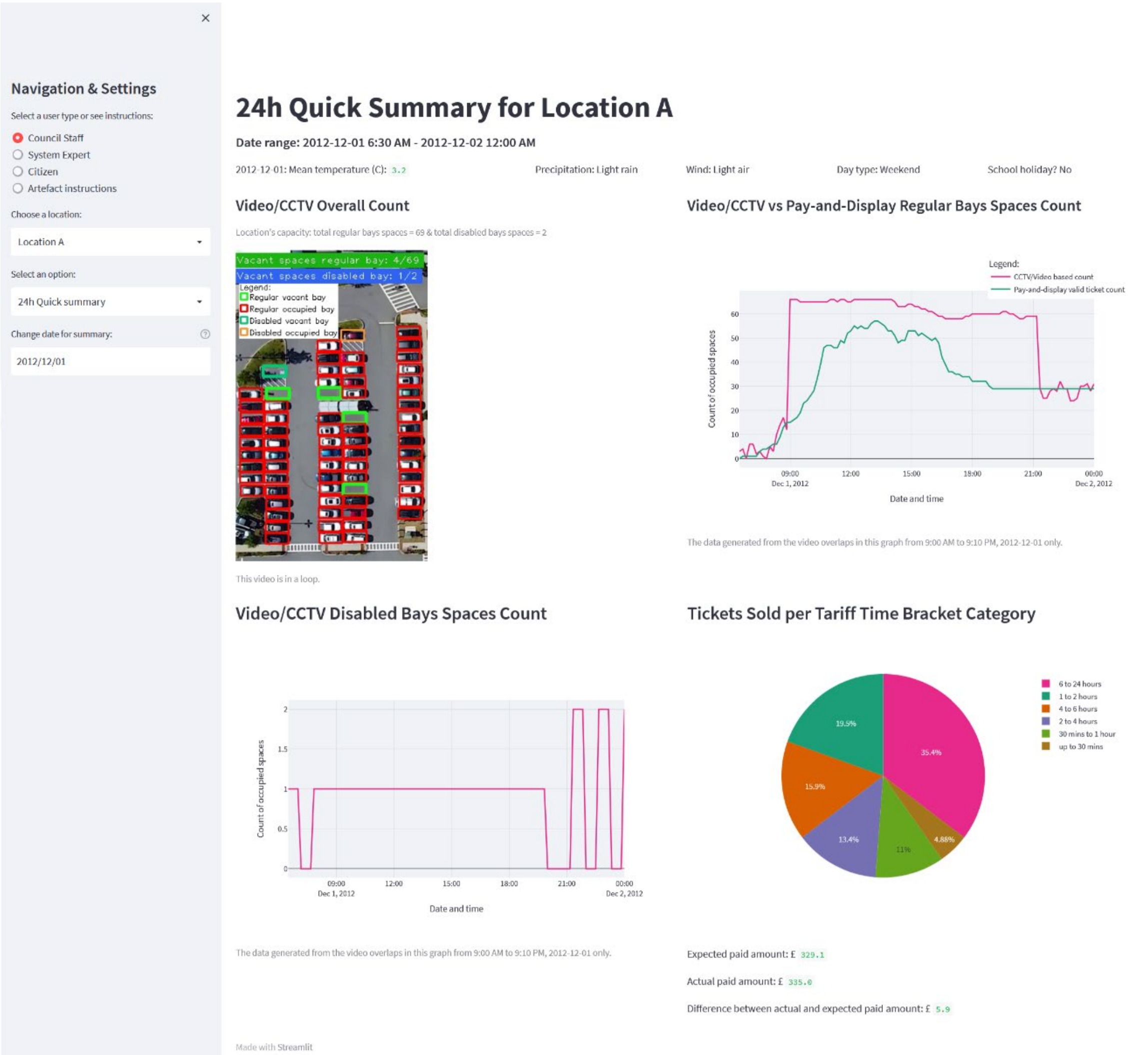


Figure 77: Council Staff → Location A → 24h Quick summary

Council Staff → Location A → Predictions *(figure 78)*

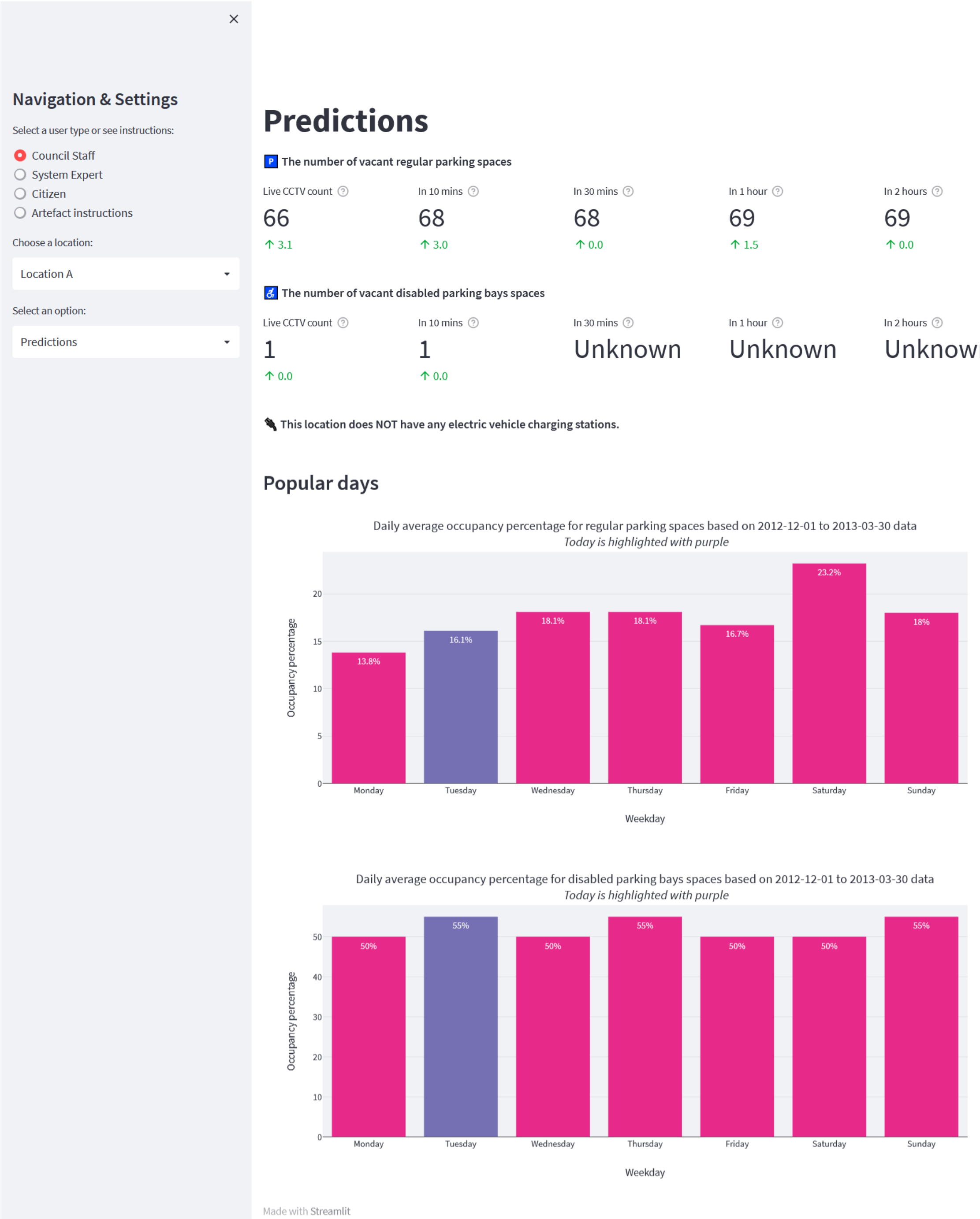


Figure 78: Council Staff → Location A → Predictions

Council Staff → All locations for pay-and-display *(figures 79 & 80)*

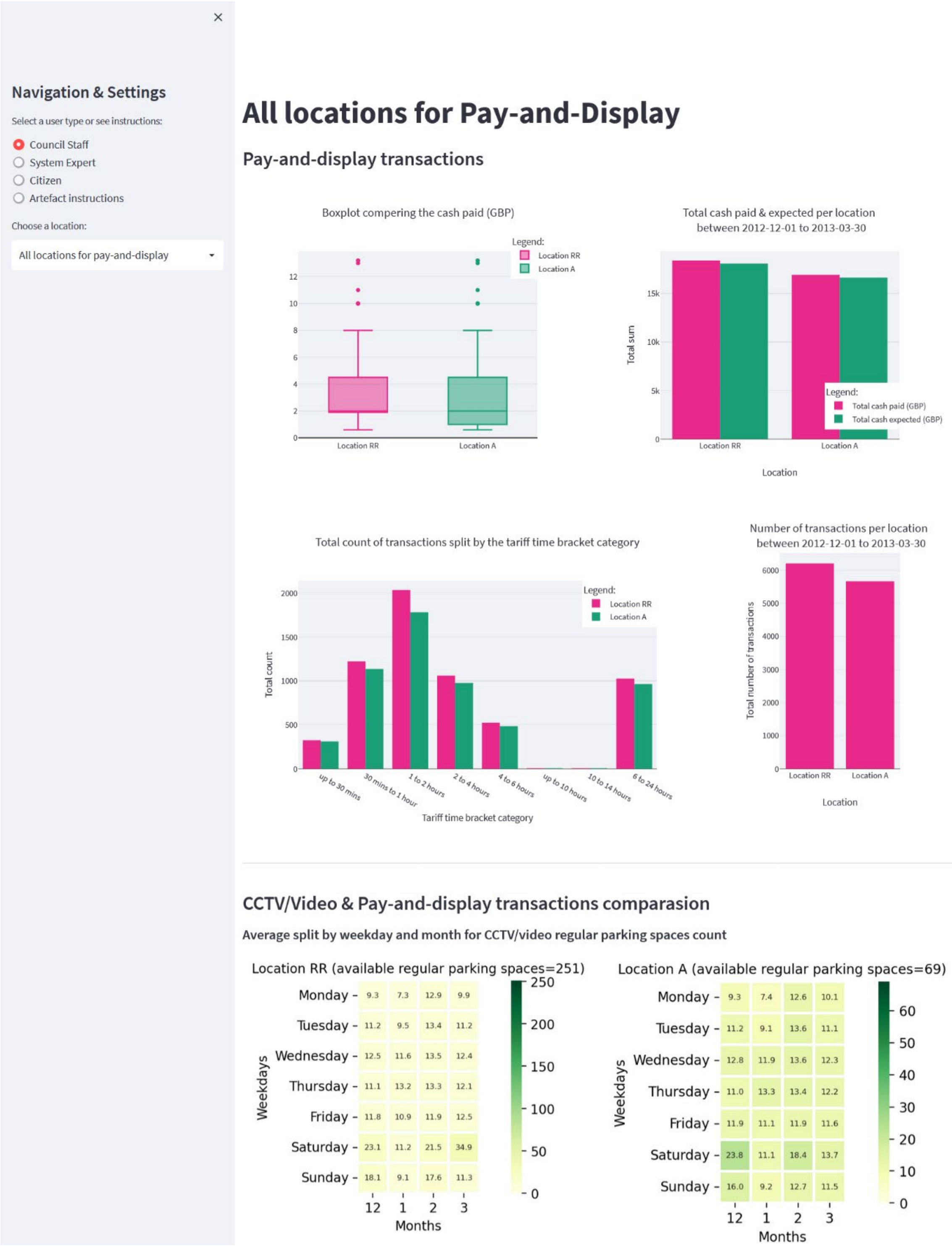


Figure 79: Council Staff → All locations for pay-and-display (1)

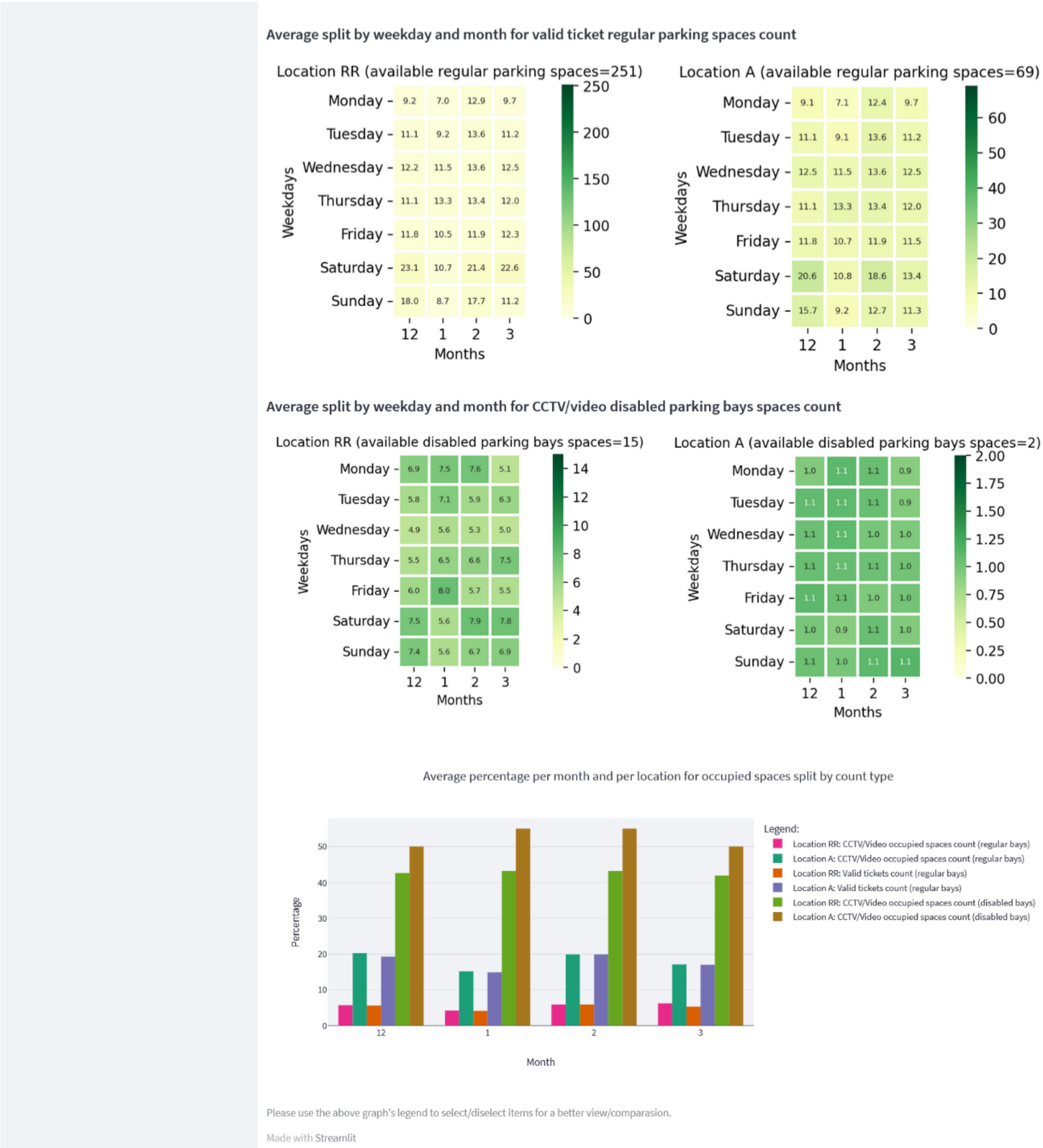


Figure 80: Council Staff → All locations for pay-and-display (2)

Council Staff → All locations for electric vehicles *(figures 81 & 82)*

Figure 81: Council Staff → All locations for electric vehicles (1)

Figure 82: Council Staff → All locations for electric vehicles (2)

System Expert → Information architecture *(figure 83)*

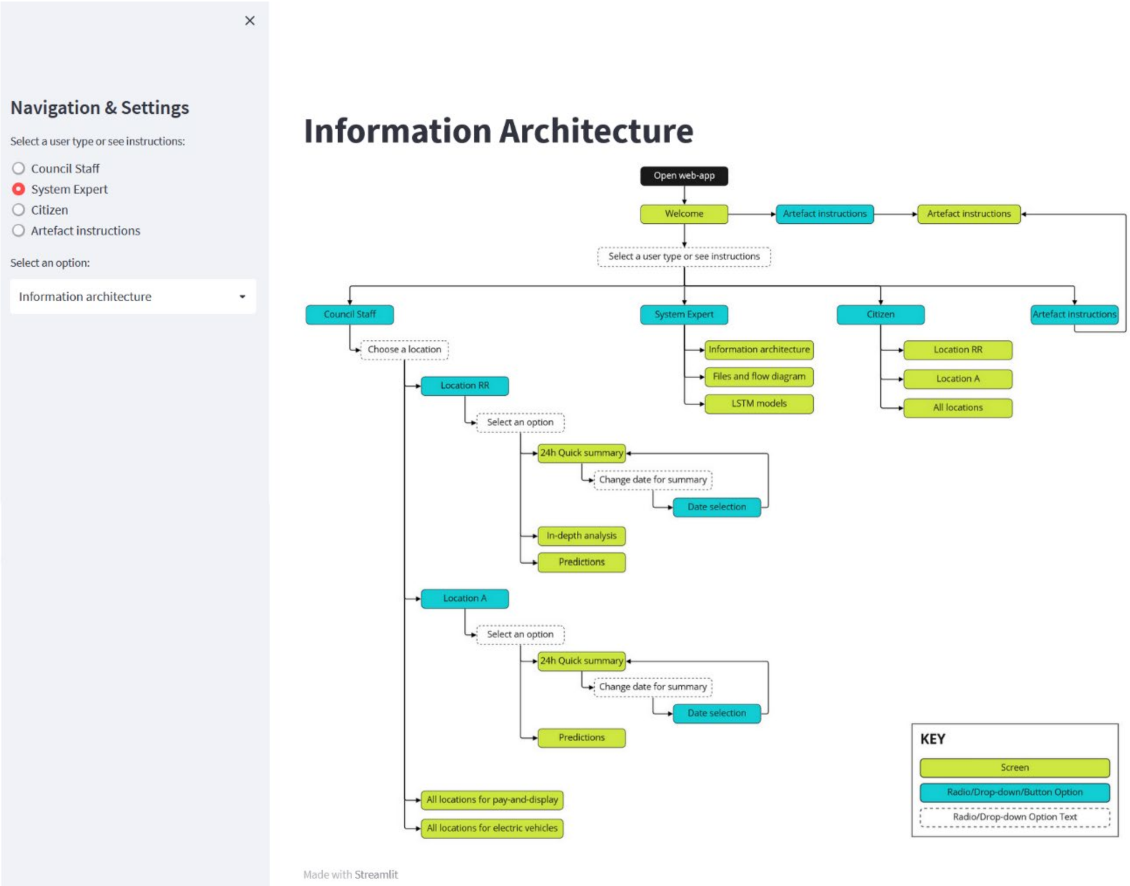


Figure 83: System Expert → Information architecture

System Expert → Files and flow diagram *(figures 84 to 86)*

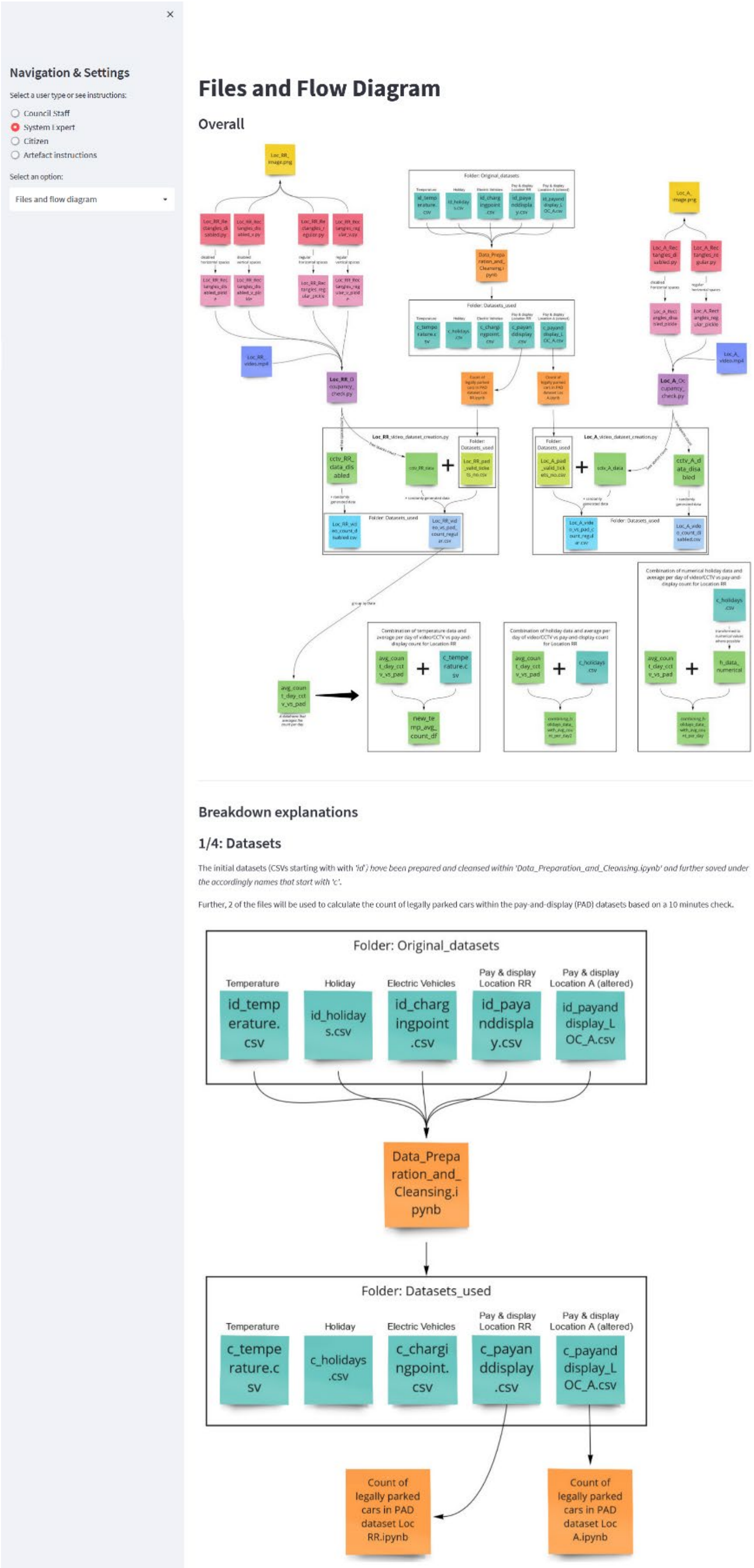
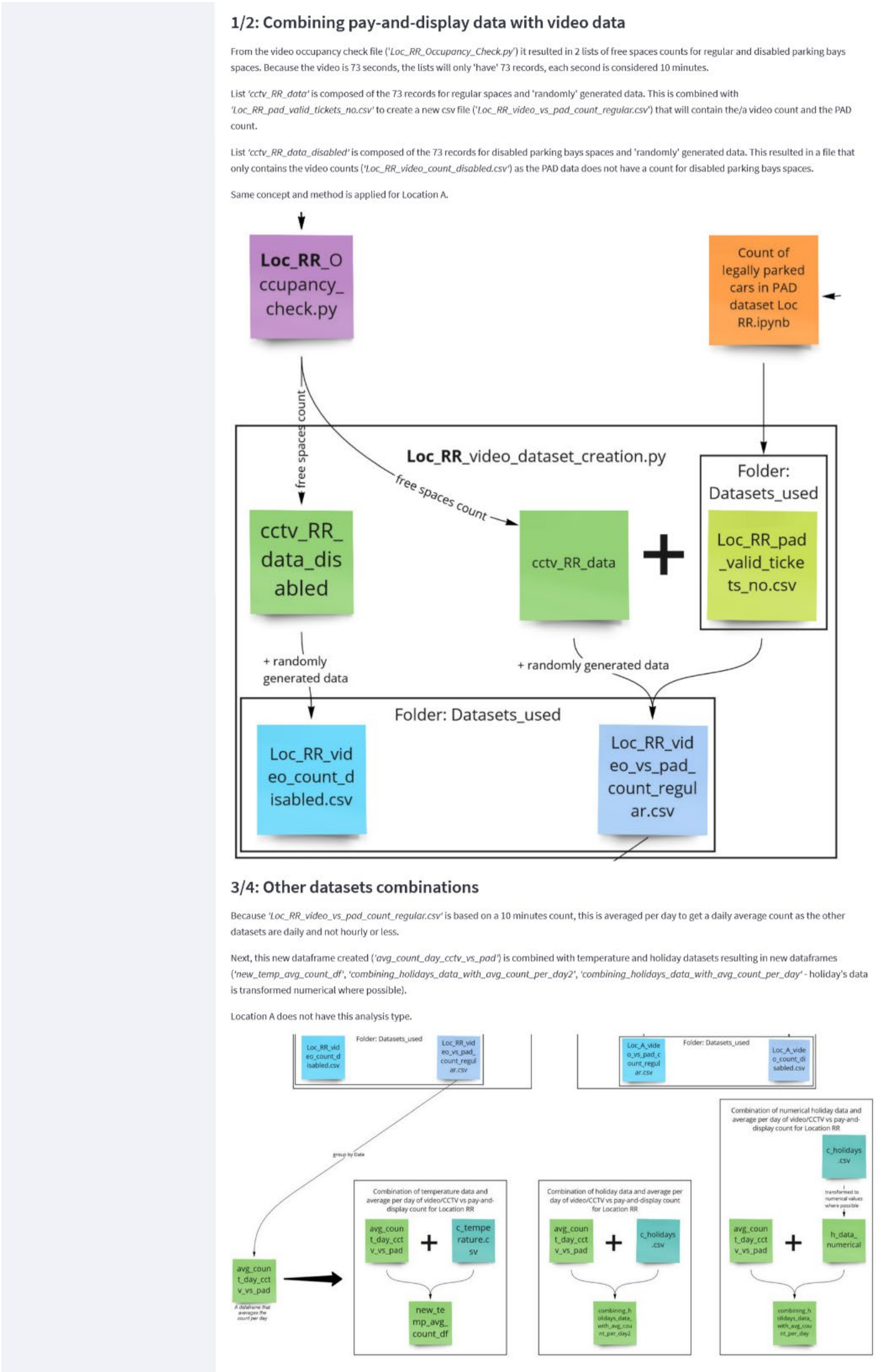


Figure 84: System Expert → Files and flow diagram (1)



1/2: Combining pay-and-display data with video data

From the video occupancy check file ('*Loc_RR_Occupancy_Check.py*') it resulted in 2 lists of free spaces counts for regular and disabled parking bays spaces. Because the video is 73 seconds, the lists will only 'have' 73 records, each second is considered 10 minutes.

List *'cctv_RR_data'* is composed of the 73 records for regular spaces and 'randomly' generated data. This is combined with *'Loc_RR_pad_valid_tickets_no.csv'* to create a new csv file ('*Loc_RR_video_vs_pad_count_regular.csv*') that will contain the/a video count and the PAD count.

List *'cctv_RR_data_disabled'* is composed of the 73 records for disabled parking bays spaces and 'randomly' generated data. This resulted in a file that only contains the video counts ('*Loc_RR_video_count_disabled.csv*') as the PAD data does not have a count for disabled parking bays spaces.

Same concept and method is applied for Location A.

3/4: Other datasets combinations

Because *'Loc_RR_video_vs_pad_count_regular.csv'* is based on a 10 minutes count, this is averaged per day to get a daily average count as the other datasets are daily and not hourly or less.

Next, this new dataframe created ('*avg_count_day_cctv_vs_pad*') is combined with temperature and holiday datasets resulting in new dataframes ('*new_temp_avg_count_df*', '*combining_holidays_data_with_avg_count_per_day2*', '*combining_holidays_data_with_avg_count_per_day*' - holiday's data is transformed numerical where possible).

Location A does not have this analysis type.

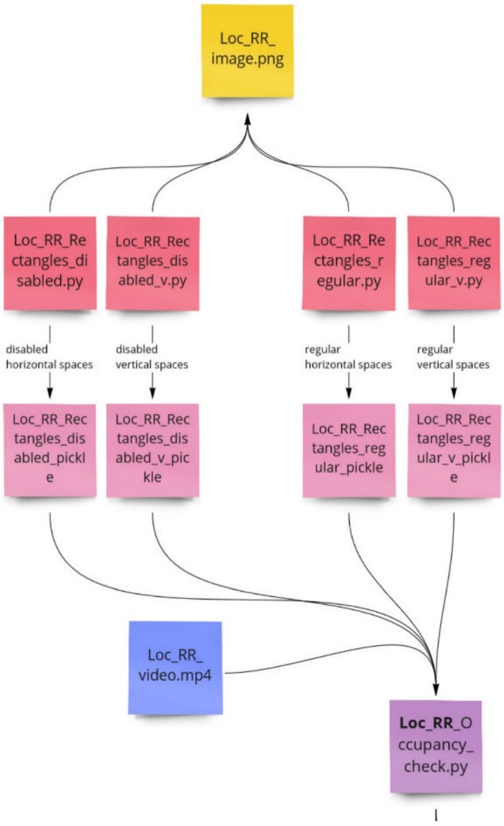
Figure 85: System Expert → Files and flow diagram (2)

4/4: Video creation

Using a snippet from the video (*'Loc_RR_image.png'*) the rectangles are created with the help of *.py* files and the positions are stored within the *pickle* files. Because this location contains 2 types of parking spaces (vertical and horizontal), 2 files per parking space type have been created.

The *'Loc_RR_Occupancy_check.py'* uses the video of the location (*'Loc_RR_video.mp4'*) and the *pickle* files to create the occupancy check.

Same concept and method is applied for Location A (except it does not have a vertical parking space type).



Made with Streamlit

Figure 86: System Expert → Files and flow diagram (3)

System Expert → LSTM models *(figures 87 & 88)*

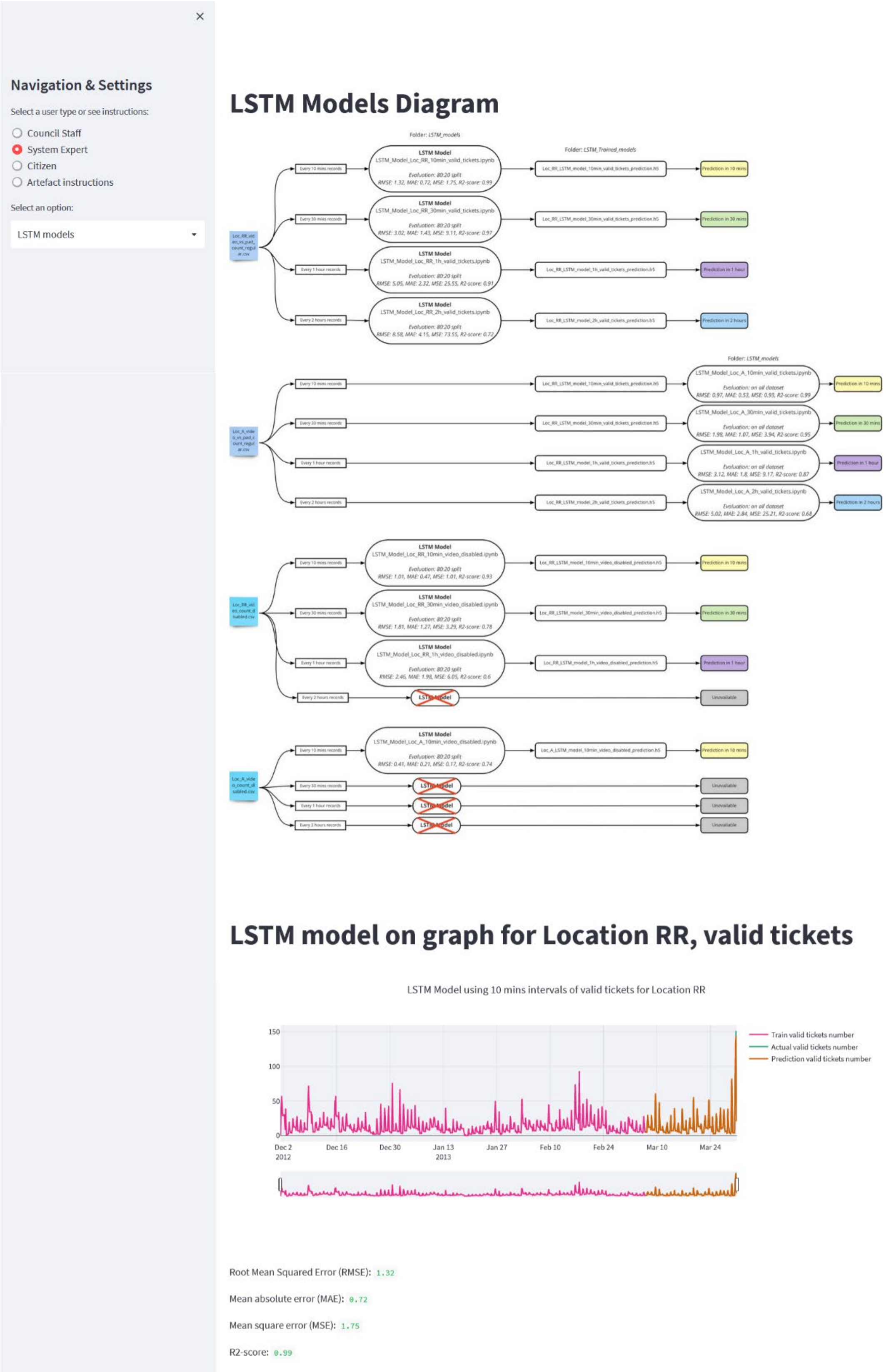


Figure 87: System Expert → LSTM models (1)

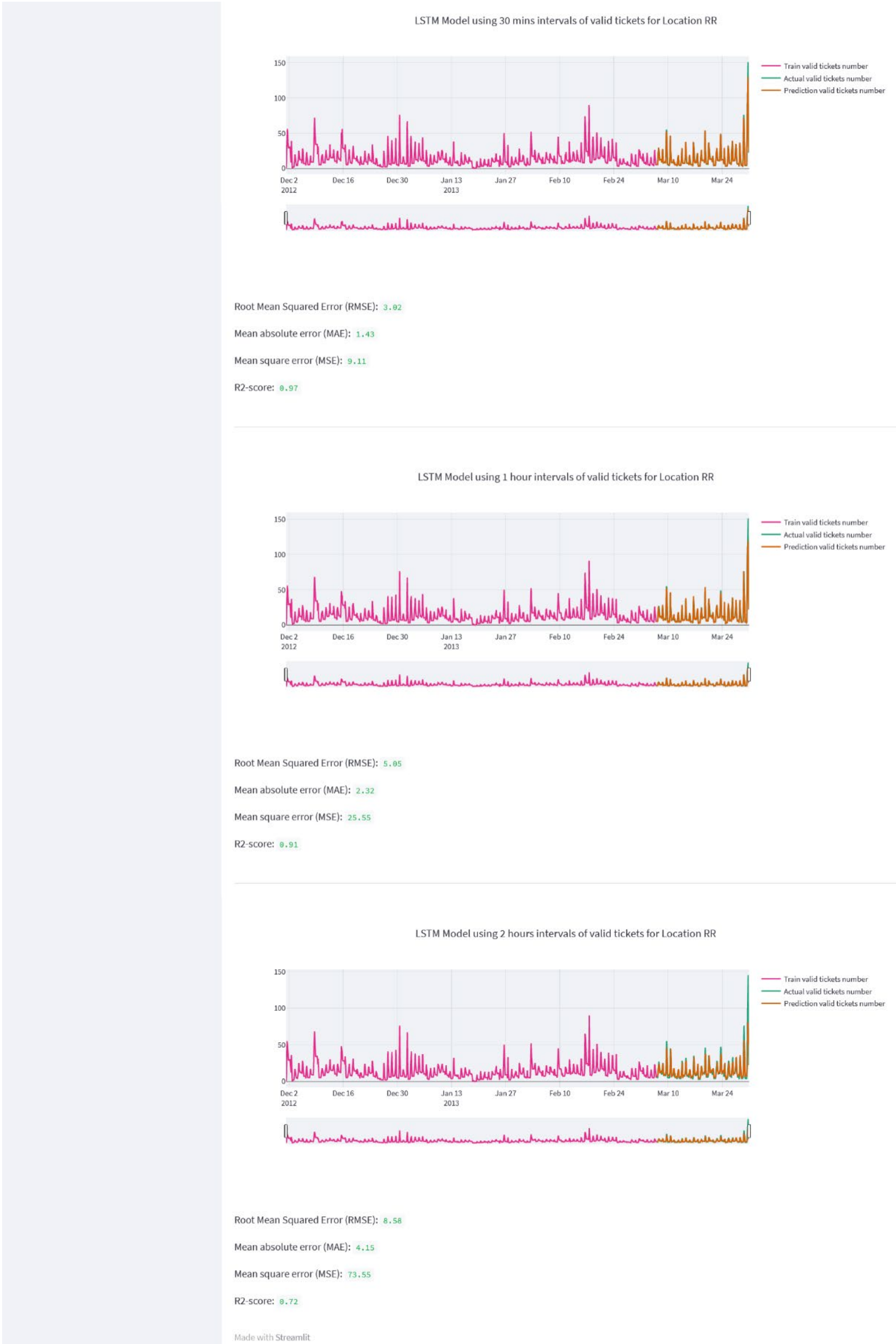


Figure 88: System Expert → LSTM models (2)

Citizen → Location RR *(figure 89)*

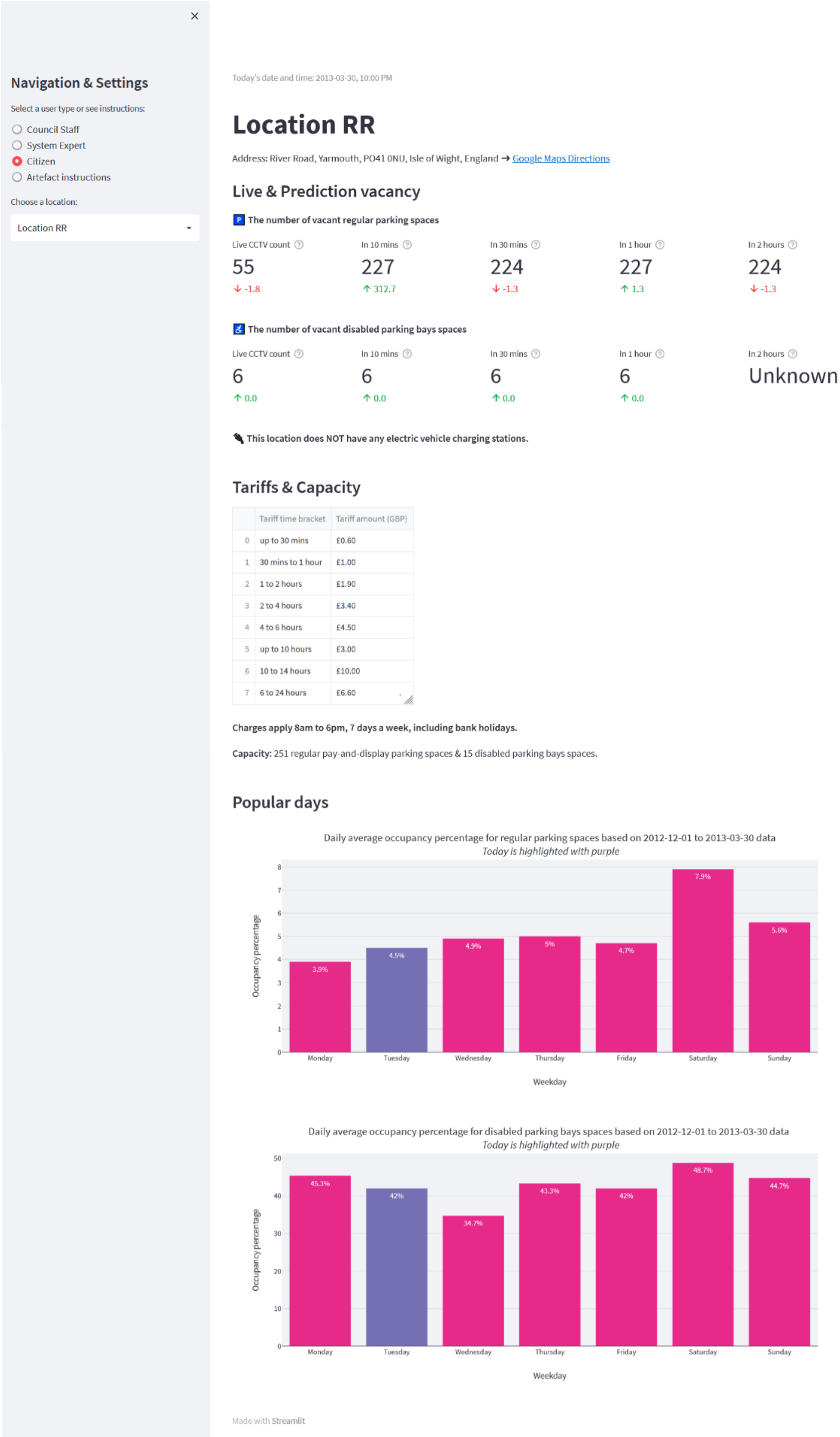


Figure 89: Citizen → Location RR

Citizen → Location A *(figure 90)*

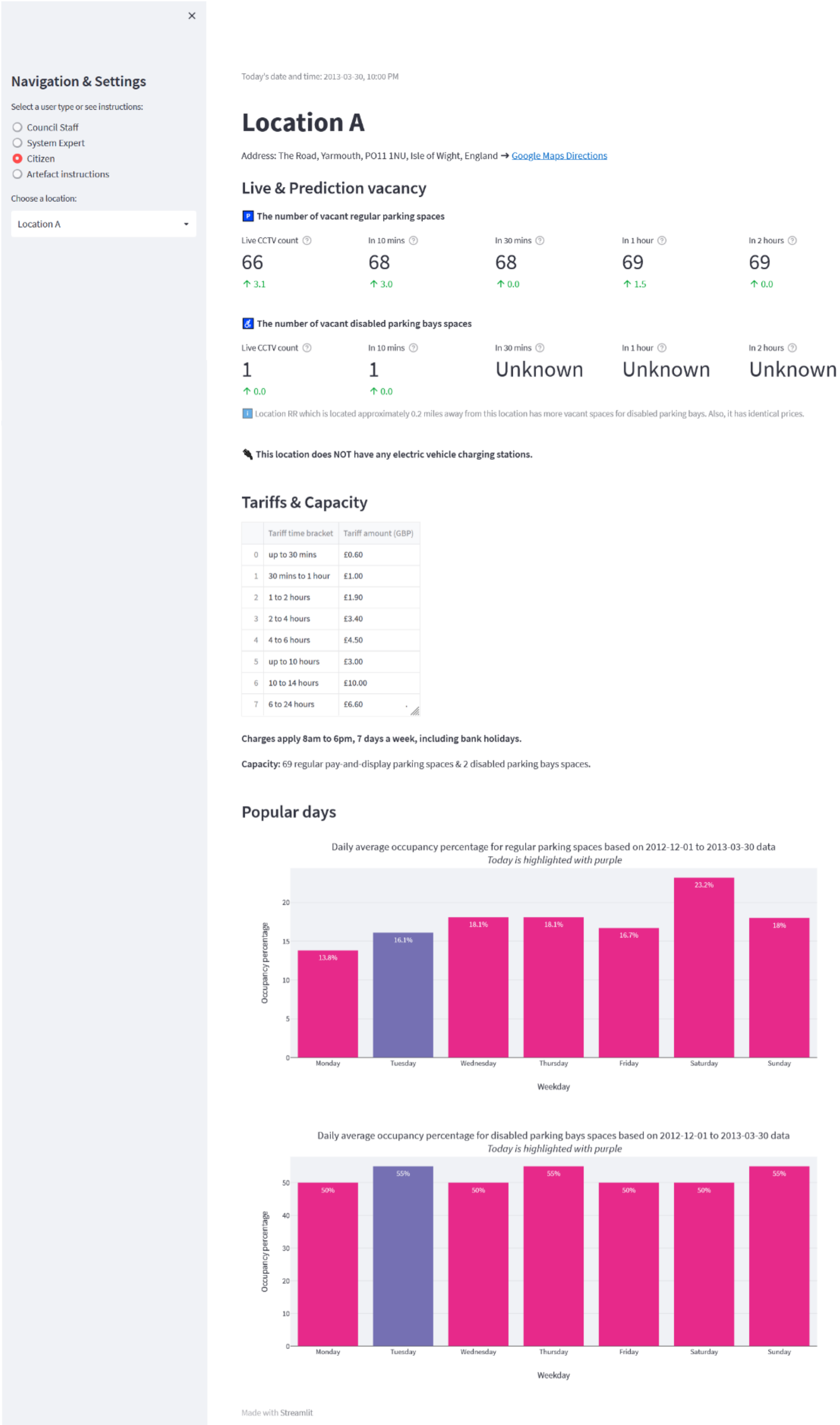


Figure 90: Citizen → Location A

Artefact instructions *(figures 91 & 92)*

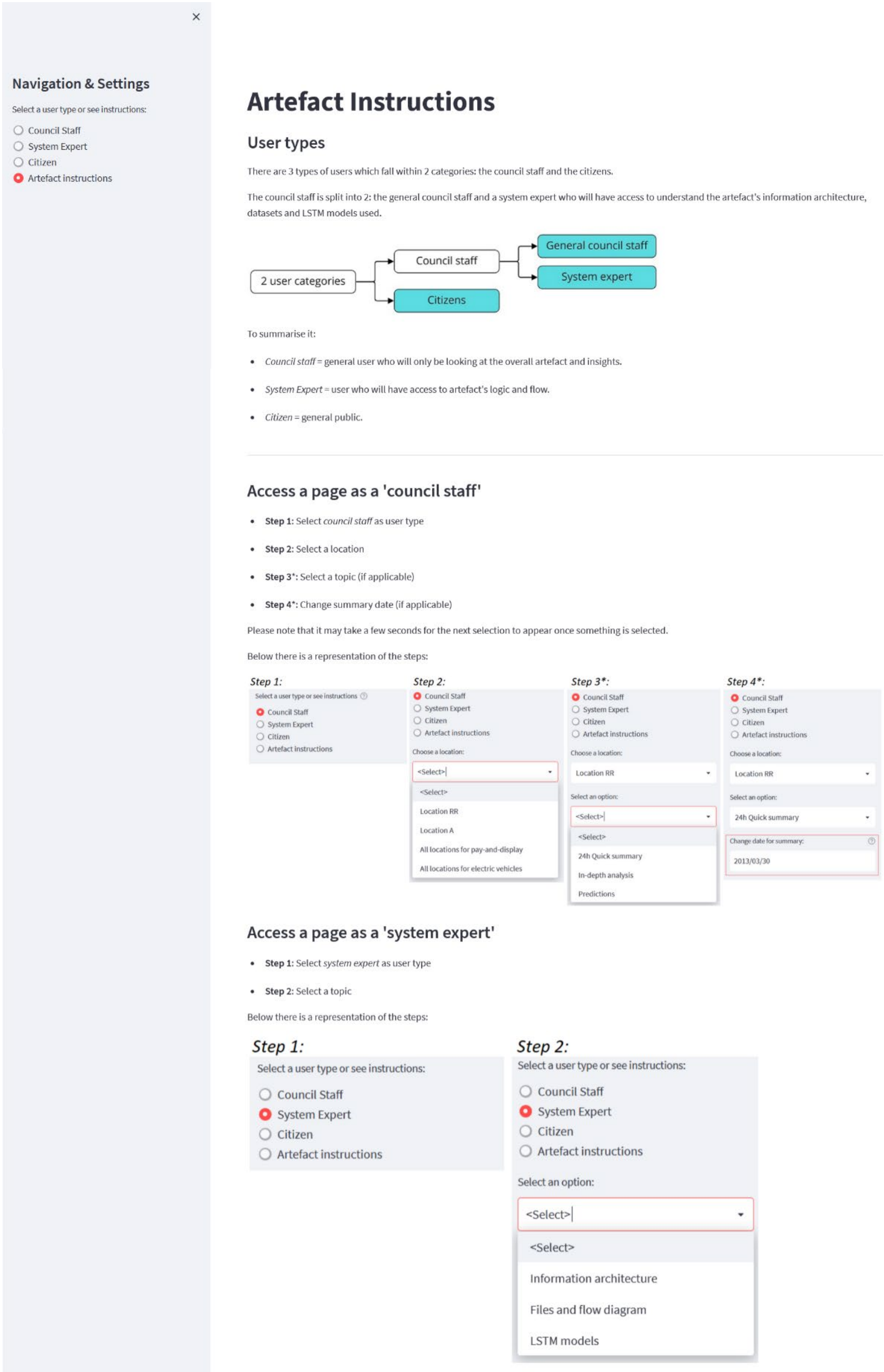


Figure 91: Artefact instructions (1)

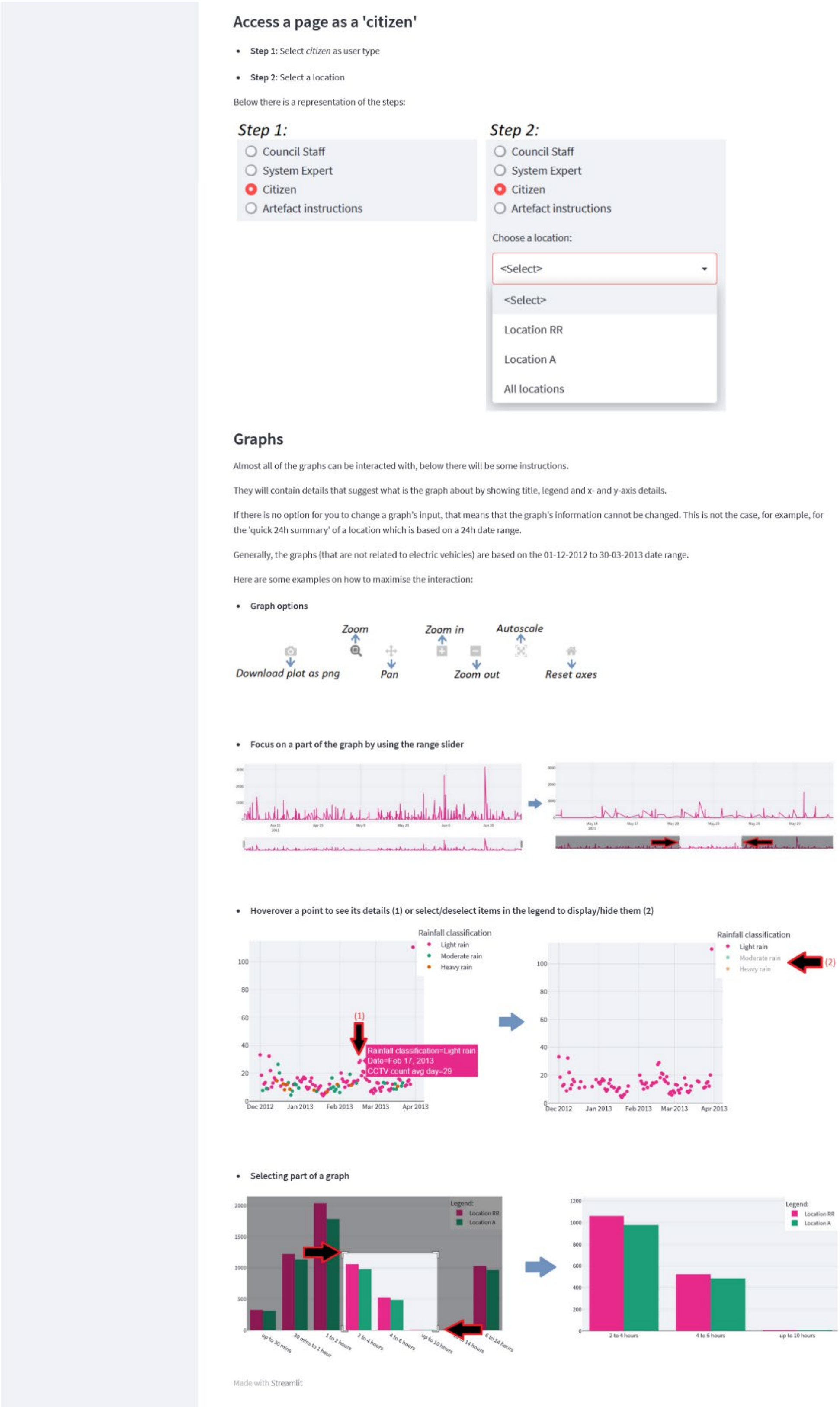


Figure 92: Artefact instructions (2)

11.9.Appendix I – Emails

Southampton City Council *(figures 93 & 94)*

Firefox https://outlook.office.com/mail/sentitems/id/AAQkADg1M2MwNjJm...

Assistance regarding my thesis

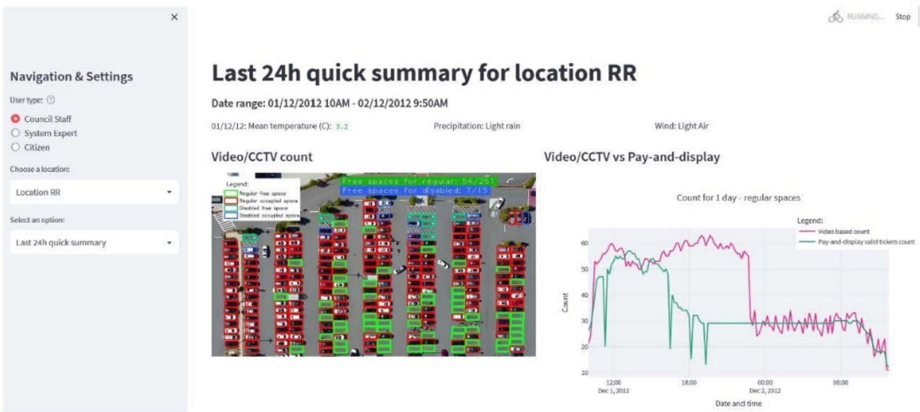
Denise Tineghe <3tinec42@solent.ac.uk>
Sun 07/08/2022 22:38

To: parking.services@southampton.gov.uk <parking.services@southampton.gov.uk>

Hello,

I am currently doing my postgraduate at Solent University and I am contacting you as I would very much appreciate your assistance regarding my thesis.

My work looks at helping the council better optimise or improve their "pay-and-display" parking using artificial intelligence and predictive analysis. Therefore, I have created a product/artefact that would facilitate this (please see a draft snippet below).



One of the requirements of the unit is to have the artefact tested by personnel that works for the council (regardless of the department). The test should not take longer than 20 mins and it will be done via a recorded meeting call. It involves completing up to 5 tasks in the artefact (screen sharing will be needed for this task) followed by a short questionnaire. Overall, the whole meeting would not take more than half an hour.

I appreciate the time taken to consider and read this email and I am looking forward to hearing from you.

Thank you and have a lovely day!

Denise

P.S. 1: At the moment, I am still working on the artefact, the testing is planned to take place after 16/08 (to be further confirmed).
P.S. 2: Please feel free to share this email if you think someone else will be willing to help too, I need to conduct 5 such tests - thank you.

1 of 1 01/09/2022, 09:21 pm

Figure 93: Southampton City Council email

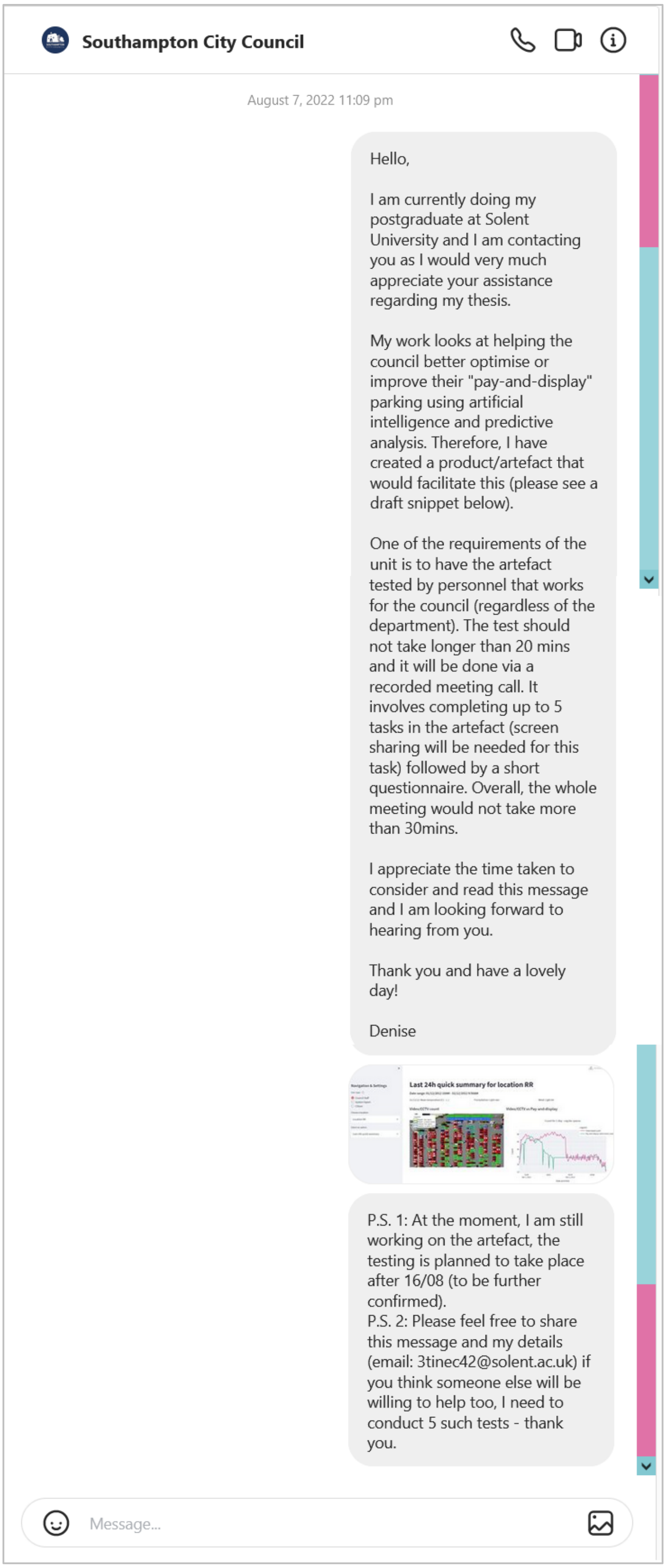


Figure 94: Southampton City Council social media message

Isle of Wight Council *(figures 95 & 96)*

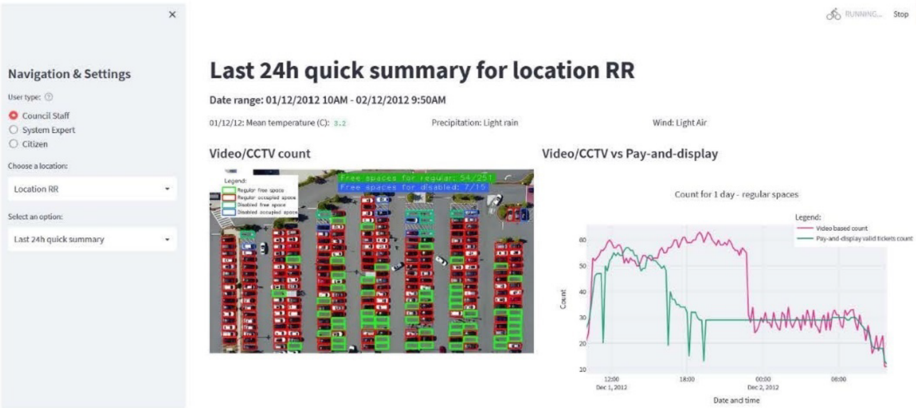
Assistance regarding my thesis

Denise Tineghe <3tinec42@solent.ac.uk>
Sun 07/08/2022 23:03
To: customer.services@iow.gov.uk <customer.services@iow.gov.uk>

Hello,

I am currently doing my postgraduate at Solent University and I am contacting you as I would very much appreciate your assistance regarding my thesis.

My work looks at helping the council better optimise or improve their "pay-and-display" parking using artificial intelligence and predictive analysis. Therefore, I have created a product/artefact that would facilitate this (please see a draft snippet below).



One of the requirements of the unit is to have the artefact tested by personnel that works for the council (regardless of the department). The test should not take longer than 20 mins and it will be done via a recorded meeting call. It involves completing up to 5 tasks in the artefact (screen sharing will be needed for this task) followed by a short questionnaire. Overall, the whole meeting would not take more than half an hour.

I appreciate the time taken to consider and read this email and I am looking forward to hearing from you.

Thank you and have a lovely day!

Denise

P.S. 1: At the moment, I am still working on the artefact, the testing is planned to take place after 16/08 (to be further confirmed).
P.S. 2: Please feel free to share this email if you think someone else will be willing to help too, I need to conduct 5 such tests - thank you.

Figure 95: Isle of Wight Council email

Helen V. is the person who has replied to the Freedom of Information request regarding the *pay-and-display ticket machine logs* (https://www.whatdotheyknow.com/request/pay_and_display_ticket_machine_l) from where the dataset was collected.

Figure 96: Isle of Wight Council - undelivered email

Solent University *(figures 97 & 98)*

Firefox https://outlook.office.com/mail/sentitems/id/AAQkADg1M2MwNjJm...

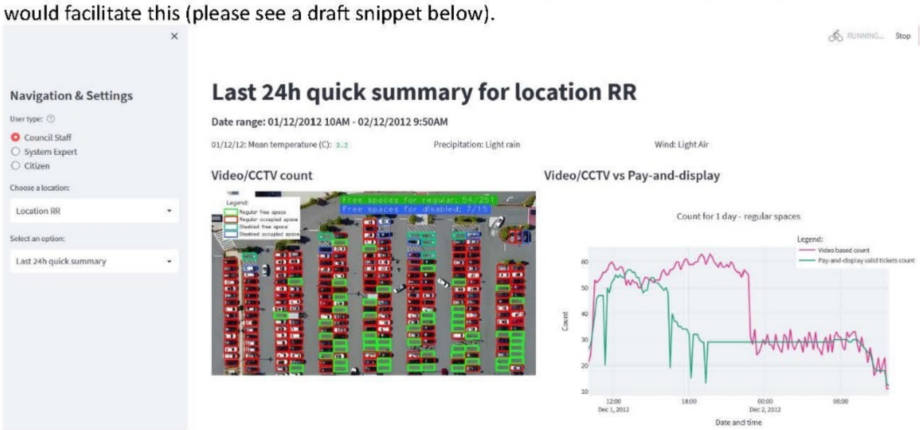
MSc support

Denise Tineghe <3tinec42@solent.ac.uk>
Mon 15/08/2022 14:38
To: **Estates & Facilities <EstatesFacilities@group.solent.ac.uk>**

Hello,

I am currently doing my postgraduate at Solent University and I am contacting you as I would very much appreciate your assistance regarding my thesis.

My work looks at helping the council better optimise or improve their "pay-and-display" parking using artificial intelligence and predictive analysis. Therefore, I have created a product/artefact that would facilitate this (please see a draft snippet below).



I have been advised to email you as I understood that this department is also looking after Solent's parking facilities. One of the requirements of the unit is to have the artefact tested by personnel who works in a department related to parking. The test should not take longer than 20 mins and it will be done via a recorded meeting call. It involves completing up to 5 tasks in the artefact (screen sharing will be needed for this task) followed by a short questionnaire. Overall, the whole meeting would not take more than half an hour.

I appreciate the time taken to consider and read this email and I am looking forward to hearing from you.

Thank you and have a lovely rest of the day!

Denise

1 of 1 01/09/2022, 09:21 pm

Figure 97: Estates and Facilities email

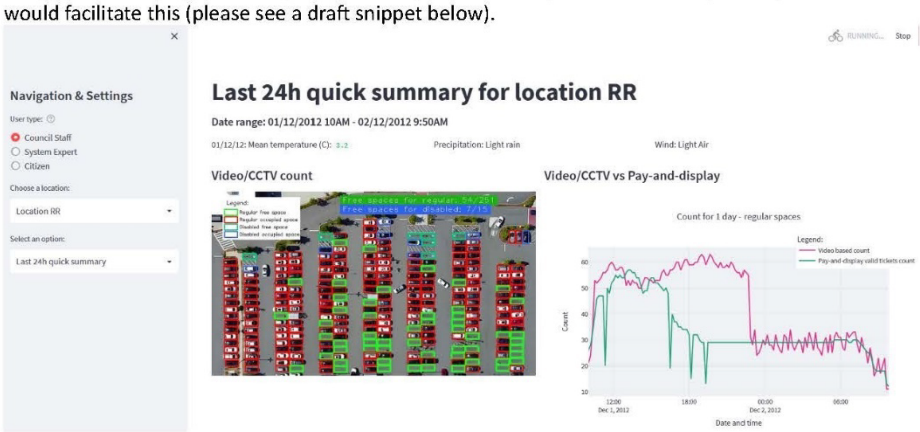
MSc support

Denise Tineghe <3tinec42@solent.ac.uk>
Mon 15/08/2022 12:49
To: Parking Facilities <parking.facilities@solent.ac.uk>

Hello,

I am currently doing my postgraduate at Solent University and I am contacting you as I would very much appreciate your assistance regarding my thesis.

My work looks at helping the council better optimise or improve their "pay-and-display" parking using artificial intelligence and predictive analysis. Therefore, I have created a product/artefact that would facilitate this (please see a draft snippet below).



One of the requirements of the unit is to have the artefact tested by personnel who works in this field. The test should not take longer than 20 mins and it will be done via a recorded meeting call. It involves completing up to 5 tasks in the artefact (screen sharing will be needed for this task) followed by a short questionnaire. Overall, the whole meeting would not take more than half an hour.

I appreciate the time taken to consider and read this email and I am looking forward to hearing from you.

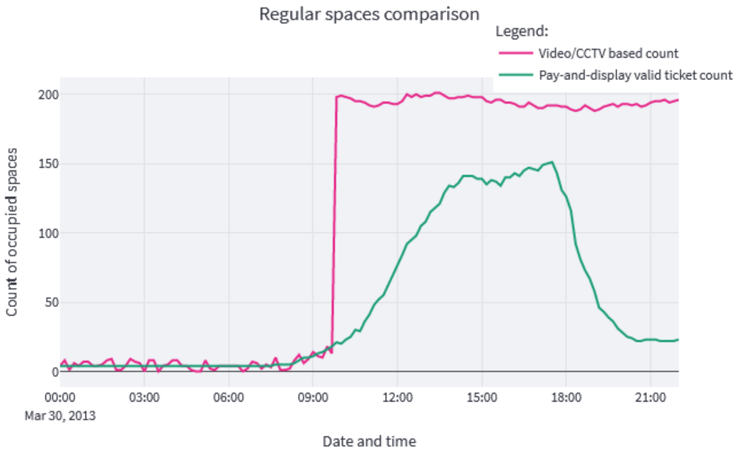
Thank you and have a lovely rest of the day!

Denise

Figure 98: Parking Facilities email

11.10.Appendix J – Graphs

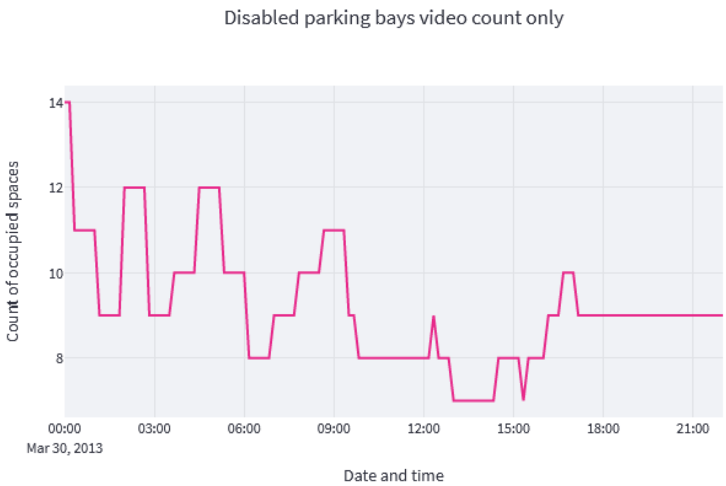
Council staff → Location RR → 24h Quick summary → 2013/03/30



The data generated from the video overlaps in this graph from 9:50 AM to 10:00 PM, 2013-03-30 only.

Figure 99: Video/CCTV vs Pay-and-Display Regular Bays Spaces Count

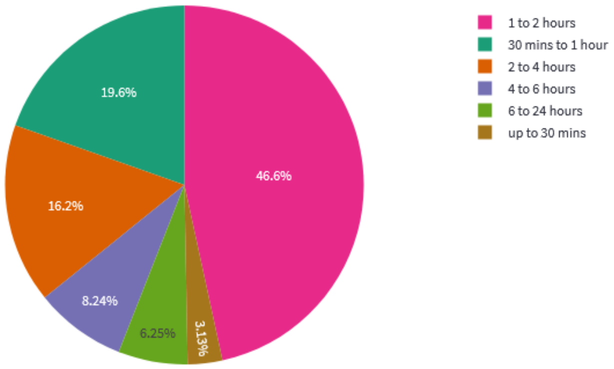
Figure 99 compares the *CCTV/video count* and the *valid pay-and-display tickets count*. It can be seen that after almost half a graph there is a high discrepancy between the two.



The data generated from the video overlaps in this graph from 9:50 AM to 10:00 PM, 2013-03-30 only.

Figure 100: Video/CCTV Disabled Bays Spaces Count

Figure 100: From this graph, the *disabled parking bays* are relatively more occupied during the night to mid-day and less occupied mid-day to night.



Expected paid amount: £ `856.7`

Actual paid amount: £ `873.2`

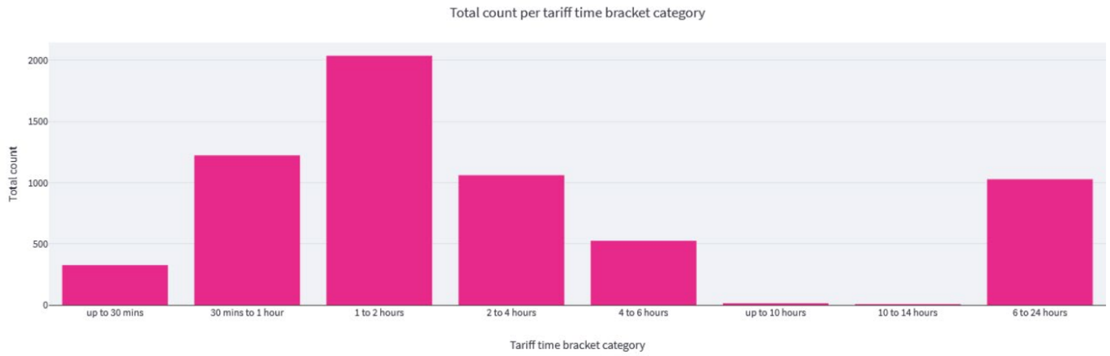
Difference between actual and expected paid amount: £ `16.5`

Figure 101: Tickets sold per tariff time bracket category

Figure 101: For this day, most drivers tended to stay for *"1 to 2 hours"* within the location RR. Also, because the ticket machines do not offer change, the drivers overall paid £16.5 more.

Council staff → Location RR → In-depth analysis

Caveat 1: The *CCTV/video count* for location RR is *almost fully randomly* generated based on the values of the *valid tickets* (for location RR). Only between 9:50 AM to 10:00 PM on 30/03/2013 the *CCTV/video count* is based on the actual video data.



Overall/Total expected paid amount: £ `18076.6`

Actual paid amount: £ `18385.8`

Difference between actual and overall/total expected paid amount: £ `309.2`

Figure 102: Total count per tariff time bracket category

Figure 102: Overall, the dataset shows that the most common frequent ticket was bought for *"1 to 2 hours"*, with *"10 to 14 hours"* being the least. *"2 to 4 hours"* and *"6 to 24 hours"* had relatively similar values. Also, overall, the drivers paid £309.2 more than expected.

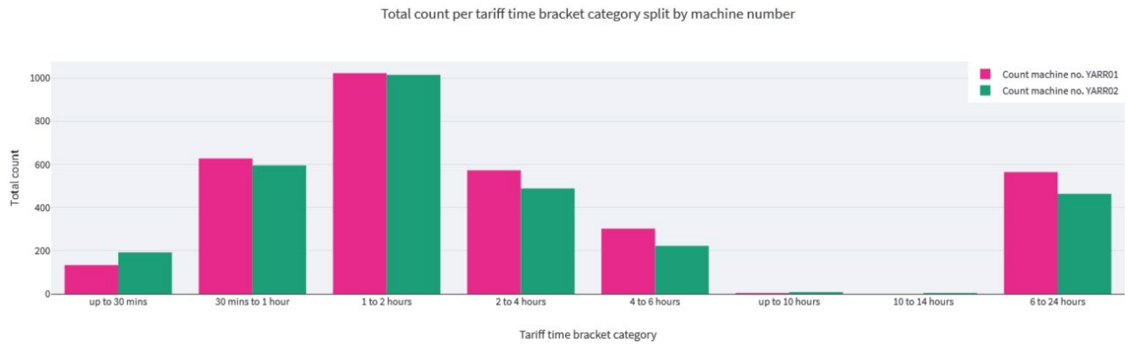


Figure 103: Total count per tariff time bracket category split by machine number

Figure 103: Overall, the ticket machine *YARR01* is more preferred than the *YARR02* excepting for *“up to 30 minutes”, “up to 10 hours”* and *“10 to 14 hours”*. The biggest difference between *YARR01* and *YARR02* is for *“6 to 24 hours”*.

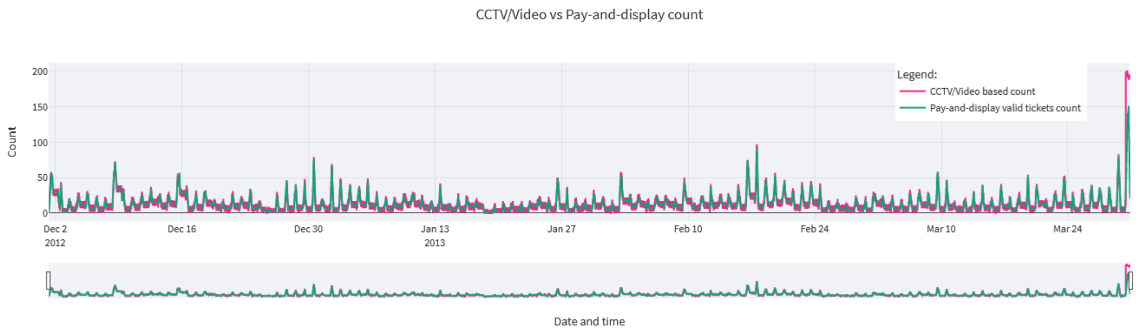


Figure 104: CCTV/Video vs Pay-and-display count

Figure 104: Overall, the *video/CCTV count* is *similar* with the *valid tickets* except towards the end of the graph due to the caveat 1. Moreover, by zooming in it can be seen that during the day the *valid ticket count* oscillates and during the night the *valid ticket count* almost stays the same.

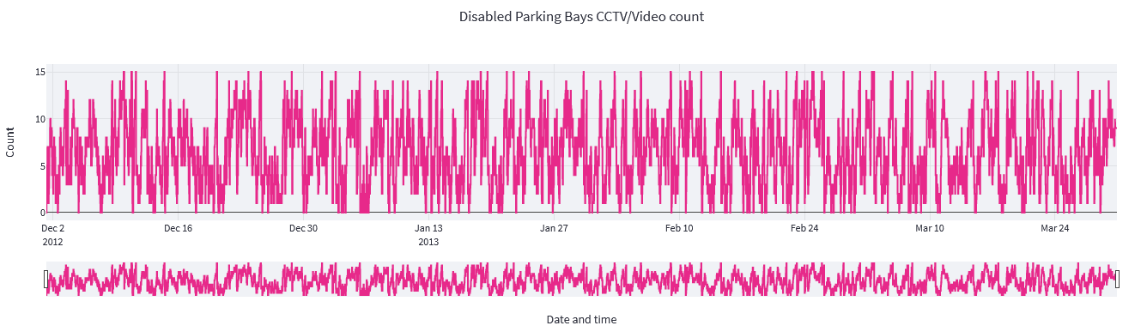


Figure 105: Disabled Parking Bays CCTV/Video count

Figure 105: No clear pattern can be extracted from this graph as this dataset was created based on *almost random* values.

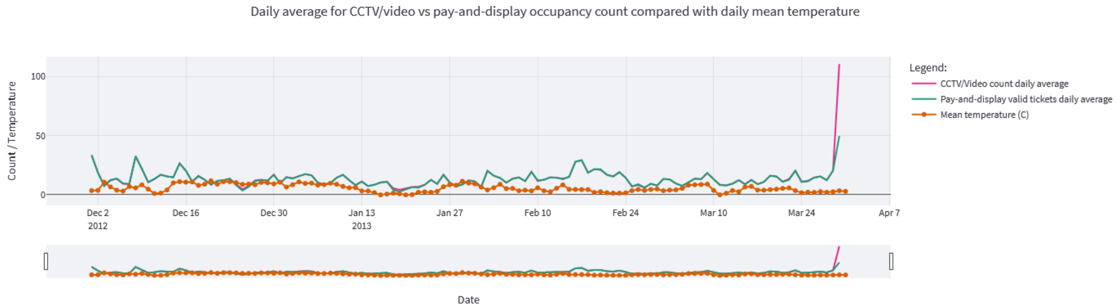


Figure 106: Daily average for CCTV/video vs pay-and-display occupancy count compared with daily mean temperature

Figure 106: Overall, it can be noticed some areas of the *mean temperature* (Jan 4-31; Feb 26-Mar 17) seem to be moving in accordance with the *pay-and-display/CCTV daily average count*.

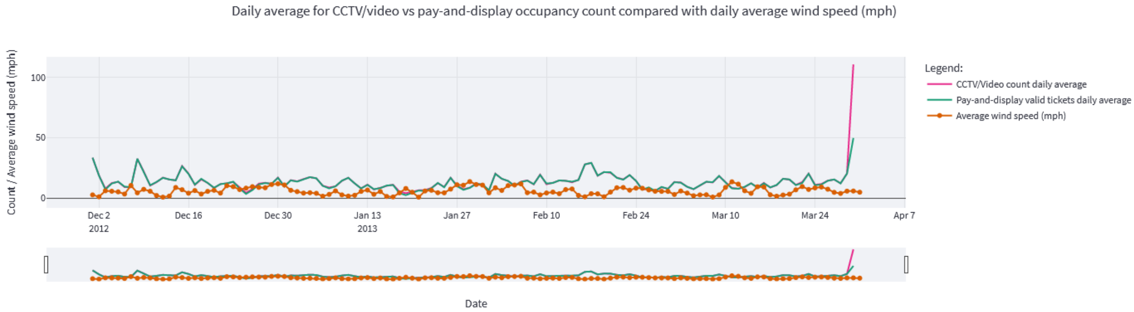


Figure 107: Daily average for CCTV/video vs pay-and-display occupancy count compared with daily average wind speed (mph)

Figure 107: No concrete patterns can be seen in the above graph.

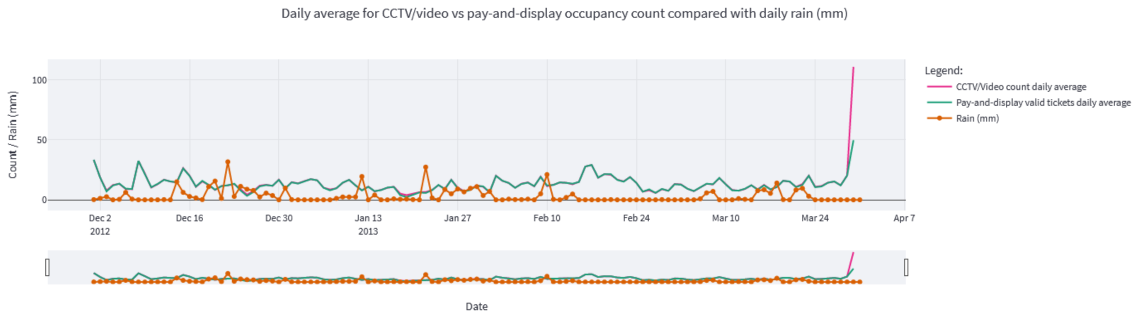


Figure 108: Daily average for CCTV/video vs pay-and-display occupancy count compared with daily rain (mm)

Figure 108: No concrete patterns can be seen in the above graph. There are a few spikes for *"rain (mm)".*

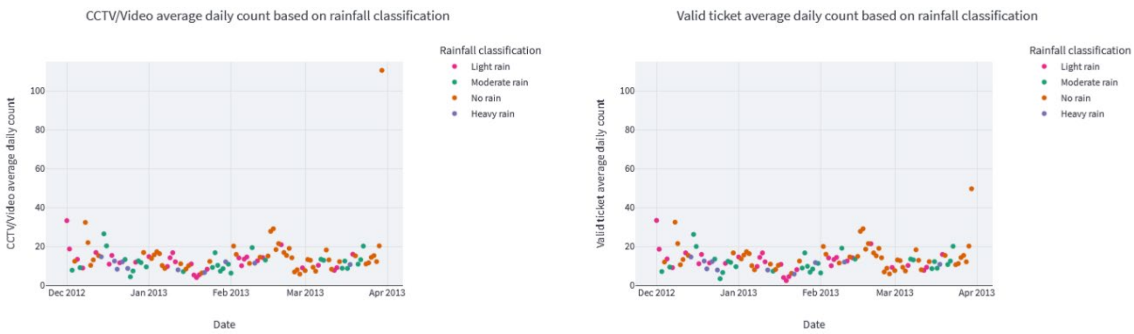


Figure 109: Counts and rainfall classification

Figure 109: The main difference seems to be the values for March 30th, but this is due to *caveat 1*. It is noted that for *"no rain"* the *daily average valid ticket count* data is more spread out than for *light, moderate and heavy rain*, same applies to *CCTV/video daily average count*.

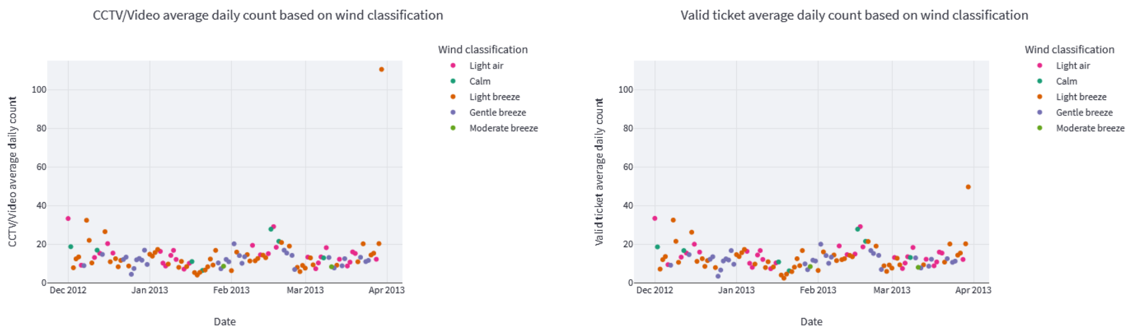


Figure 110: Counts and wind classifications

Figure 110: The main difference seems to be the values for March 30th, but this is due to *caveat 1*. *"Light air"* and *"light breeze"* saw the *daily average valid ticket count data* more spread out with *"gentle breeze"* tighter together. Same applies to *CCTV/video daily average count*.

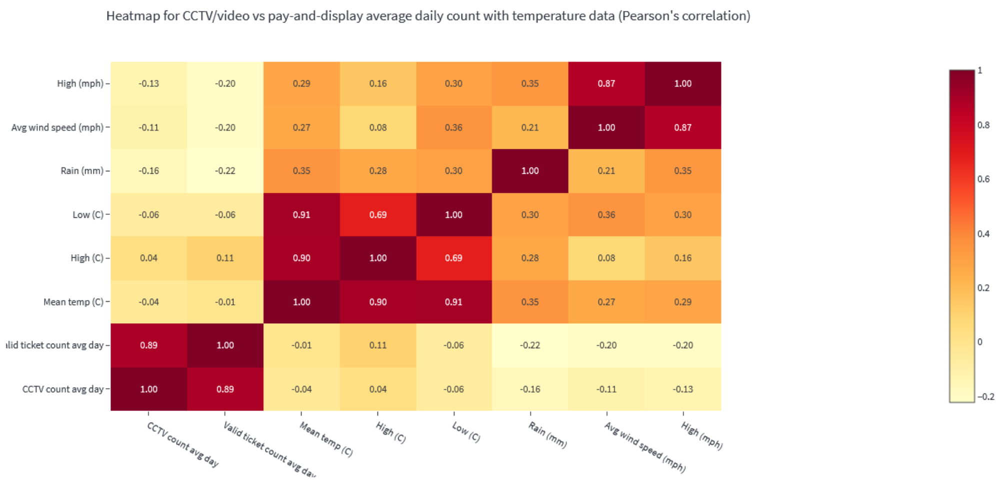


Figure 111: Heatmap for CCTV/video vs pay-and-display average daily count with temperature data (Pearson's correlation)

Figure 111: Interestingly, overall, in this 4-month dataset there is no significant (or very weak) correlation between the *weather* and the *daily average count for valid tickets* or *CCTV/video*.

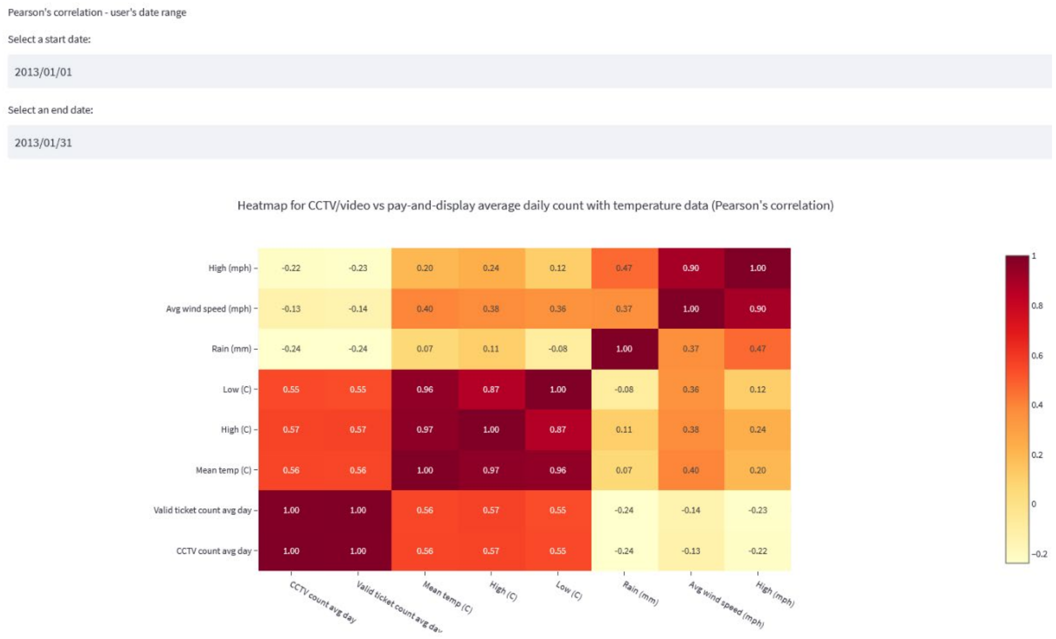


Figure 112: Heatmap for CCTV/video vs pay-and-display average daily count with temperature data (Pearson's correlation) – user input

Figure 112: **But when focusing on more shorter periods of time, such as month January,** a <u>moderate positive correlation</u> can be seen between the *daily average count for valid tickets and CCTV/video* and the *mean, highest* and *lowest temperature*. A <u>weak negative correlation</u> can be seen between the *daily average count for valid tickets and CCTV/video* and *rain, average wind speed* and *high wind speed*.

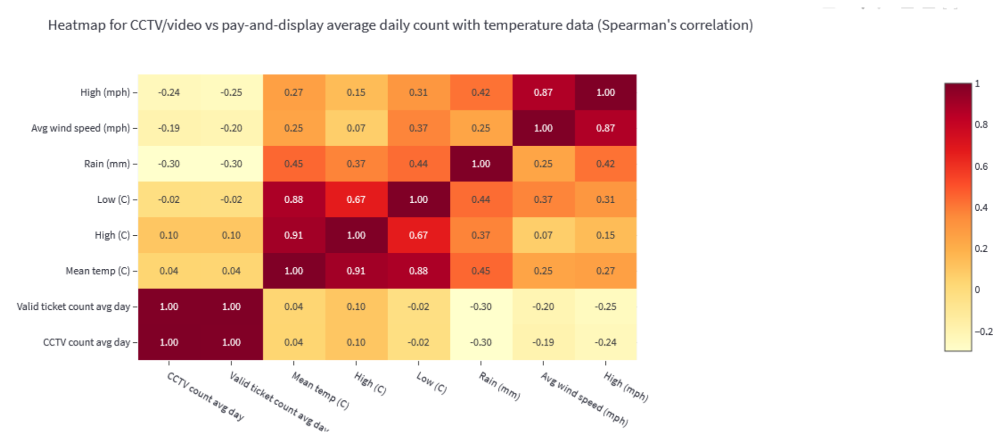


Figure 113: Heatmap for CCTV/video vs pay-and-display average daily count with temperature data (Spearman's correlation)

Figure 113: Spearman's correlation showed a very weak (positive and negative) correlation for *mean, high, low temperature* and weak negative correlation for *rain, average* and *high wind speed*.

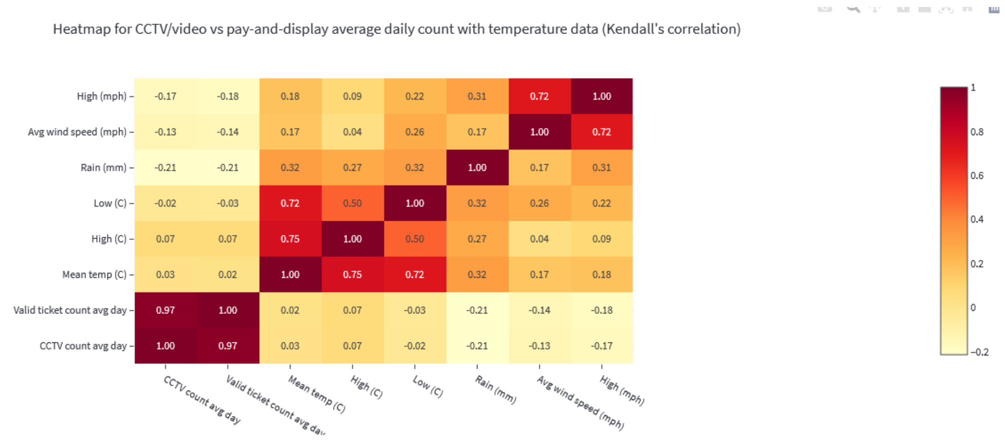


Figure 114: Heatmap for CCTV/video vs pay-and-display average daily count with temperature data (Kendall's correlation)

Figure 114: Kendall's correlation was also applied, no significant correlation can be seen between the *weather* and the *daily average count for valid tickets or CCTV/video*.

Figure 115: Counts, day types and school holidays

Figure 115: Comparing the *daily average count for CCTV/video* and the *valid tickets* based on the *school holiday*, the main difference seems to happen on the *weekend*, but this close similarity between the two graphs is due to *caveat 1*.

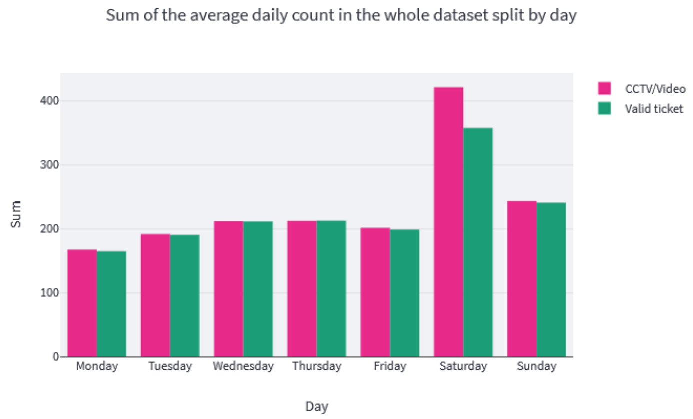


Figure 116: Sum of the average daily count in the whole dataset split by day

Figure 116: Overall, the sum is almost identical or with very little difference except for *Saturday* which is caused by the fact that in that day the *CCTV/video's data* was true to the video and not *almost randomly* generated (*caveat 1*).

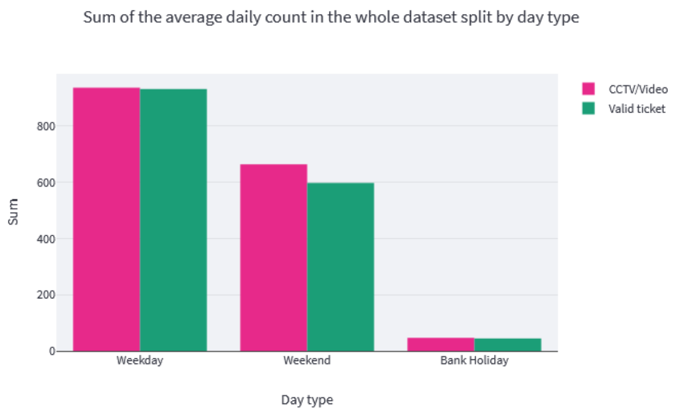


Figure 117: Sum of the average daily count in the whole dataset split by day type

Figure 117: Very close similarities between the two counts. *Caveat 1* can be seen in the *weekend CCTV/video count* bar.

A point biserial correlation coefficient of `0.02` was found between the 'school holiday' (yes/no) and the CCTV/video daily average count.

A point biserial correlation coefficient of `0.06` was found between the 'school holiday' (yes/no) and the valid ticket daily average count.

Figure 118: Point Biserial correlation

Figure 118: The point biserial correlation coefficient was calculated for the holiday's dataset *school holiday* (*yes/no*) against the *daily average count for CCTV/video* and *valid ticket*. This correlation was used as the *school holiday* is a dichotomous variable and the *count* is a continuous variable. The results show no significant correlation.

Council staff → Location A → 24h Quick summary → 2012/12/01

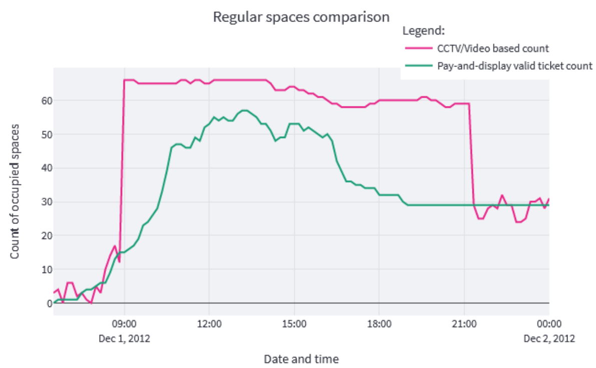


Figure 119: Video/CCTV vs Pay-and-Display Regular Bays Spaces Count, Location A

Figure 119: This graph compares the *CCTV/video count* and the *valid pay-and-display tickets count*. It can be clearly seen that the video count is predominantly higher than the valid ticket count. This overlap happens when the actual video's occupancy is added to the dataset.

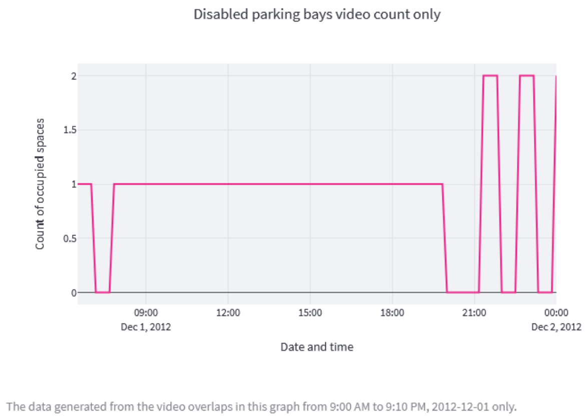
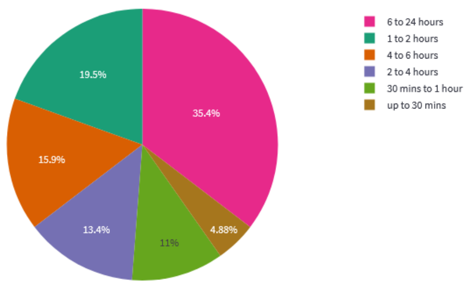


Figure 120: Video/CCTV Disabled Bays Spaces Count, Location A

Figure 120: The disabled parking bays occupancy is almost not changing at all during 6:30AM–7:50PM. The values after 9:10 PM are *almost randomly* generated.



Expected paid amount: £ `329.1`

Actual paid amount: £ `335.0`

Difference between actual and expected paid amount: £ `5.9`

Figure 121: Tickets sold per tariff time bracket category, Location A

Figure 121: For this day, most drivers tended to stay for *"6 to 24 hours"* within the location A, with the lowest value being *"up to 30 mins"*. Also, because the ticket machines do not offer change, the drivers overall paid £5.9 more.

Council staff → All locations for pay-and-display

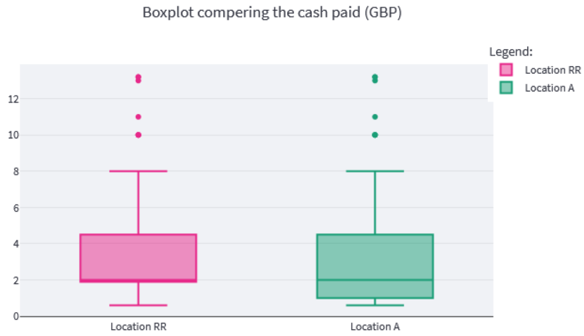


Figure 122: Boxplot compering the cash paid (GBP)

Figure 122: Both datasets have the same outliers. For location RR, *Q1-median* has the most concentrated data, for location A, *min-Q1* has the most concentrated data and both locations have *Q3-max* the most spread-out data. Both boxplots are right-skewed.

Figure 123: Total cash paid & expected per location between 2012-12-01 to 2013-03-30

Figure 123: Total cash paid is higher than *total cash expected* for both locations. Location RR has the most difference.

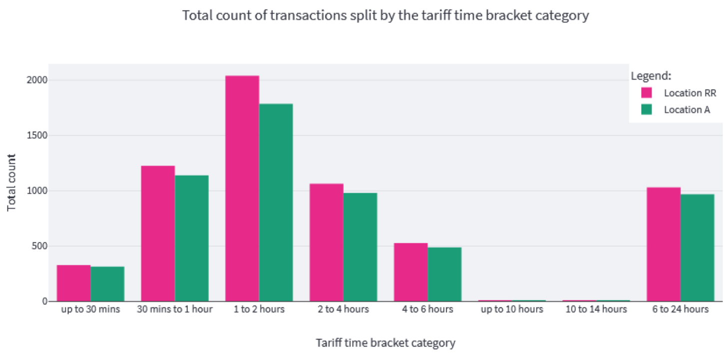


Figure 124: Total count of transactions split by the tariff time bracket category

Figure 124: Most drivers choose the *"1 to 2 hours"* ticket followed by *"30 mins to 1 hour"* for both locations. Shows that these locations have predominantly short-stay drivers, but *"6 to 24 hours"* is not that low either, being the 4th most frequently bought ticket type.

Figure 125:Number of transactions per location between 2012-12-01 to 2013-03-30

Figure 125: Location RR has more ticket machine transactions than location A, the difference being 538.

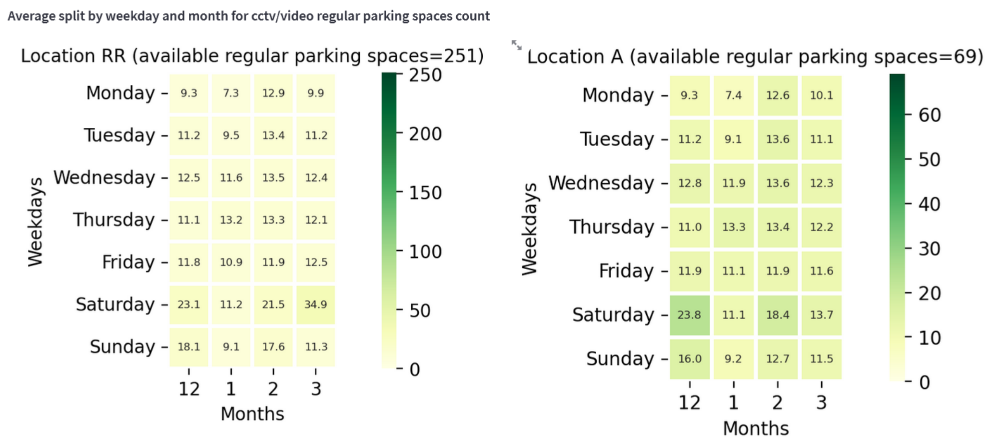


Figure 126: Average split by weekday and month for CCTV/video regular parking spaces count

Figure 126: The mean per weekday is calculated for each month. The first graph for location RR shows that the biggest and second biggest mean for December and February is Saturday and Sunday, for January is Thursday and Wednesday, and for March is Saturday and Friday.

The location A's graph shows the biggest and second biggest mean for December in the weekend, for January is the Thursday, Wednesday, for February is the Saturday with Tuesday/Wednesday and for March it is Saturday with Wednesday.

Strong similarities are seen for December-January for both locations with only partial similarities for February-March.

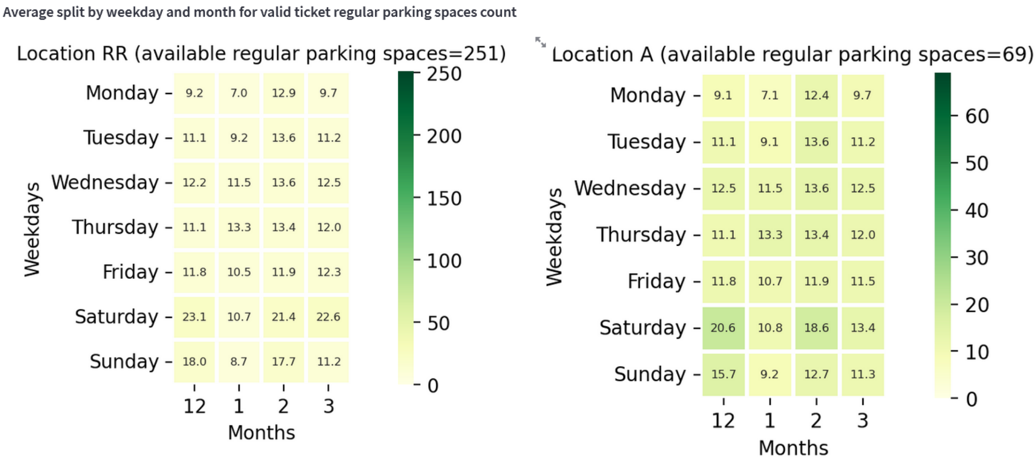


Figure 127: Average split by weekday and month for valid ticket regular parking spaces count

Figure 127: The mean per weekday is calculated for each month. The first graph for location RR shows that the biggest and second biggest mean for December and February is the weekend, with January having the mid-week (Wednesday and Thursday) and March having partially mid-week and partially weekend (Wednesday and Saturday).

The location A's graph shows the biggest and second biggest mean for December in the weekend, for January is Thursday-Wednesday, for February is Saturday and Tuesday/Wednesday and for March is Saturday and Wednesday.

Comparing the last 2 graphs, it can be concluded that for location RR, the months December-February share the same characteristics, and March shares only partial characteristics. For location A, all the months share the same characteristic.

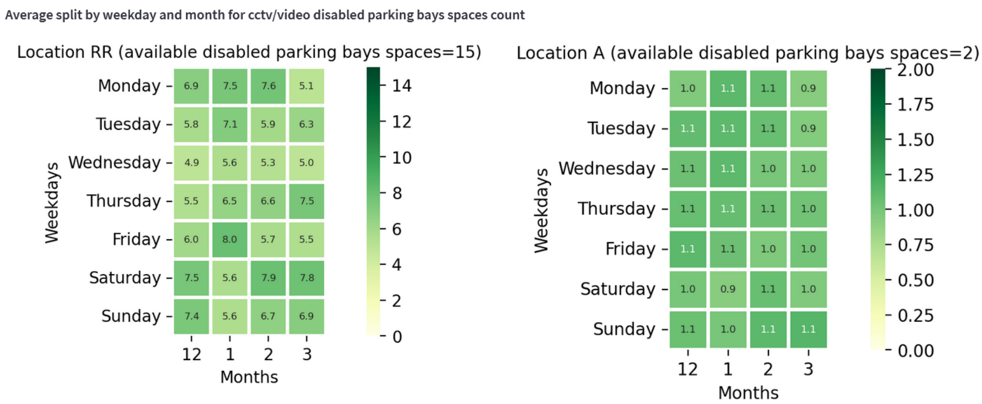


Figure 128: Average split by weekday and month for CCTV/video disabled parking bays spaces count

Figure 128: The mean per weekday is calculated for each month. A true pattern evaluation cannot be done as the data was *almost fully randomly* created for both locations. The highest mean for location RR varies from 7.5 to 8 across the 4 months and for location A the highest mean is 1.1 and the lowest 0.9 or 1.

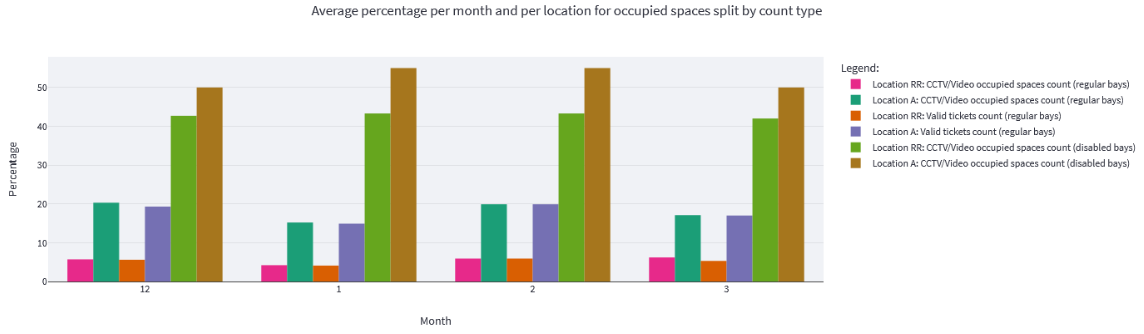


Figure 129: Average percentage per month and per location for occupied spaces split by count type

Figure 129: For location RR, the occupancy percentage is relatively equal for *CCTV/video count* and *valid tickets count*. The smallest percentage is in January (4.2% and 4.1%) and the biggest percentage is in March for *CCTV/video count* (6.2%) and equal in February for both counts (5.9%). The occupancy percentage for *disabled bays* varies from 42% (in March) to 43.3% (in January and February).

For location A, the occupancy percentage is more monthly different for *CCTV/video count* and *valid tickets count.* The smallest percentage is in January for both counts (15.2% and 14.9%) and the biggest percentage is in December for *CCTV/video count* (20.3%) and equal in February for both counts (19.9%). The occupancy percentage for *disabled bays* varies from 50% (December and March) to 55% (January and February).

Council staff → All locations for electric vehicles

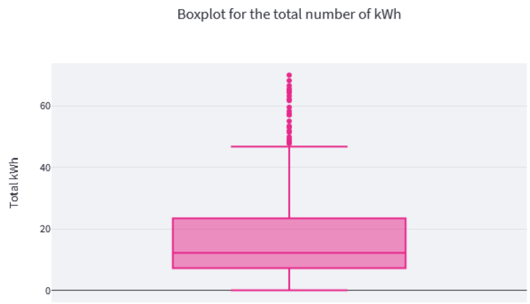


Figure 130: Boxplot for the total number of kWh

Figure 130: Quite a few outliers can be seen after the max value. It seems that the data is more concentrated up to the median and after the median is more spread-out having higher variability, with *Q1-median* (or 25% of the data) the most concentrated and with *Q3-max* (or 25% of the data) being the most spread-out. 50% of the data (*Q1 to Q3*) is between 7.2 to 23.4. The boxplot is right-skewed.

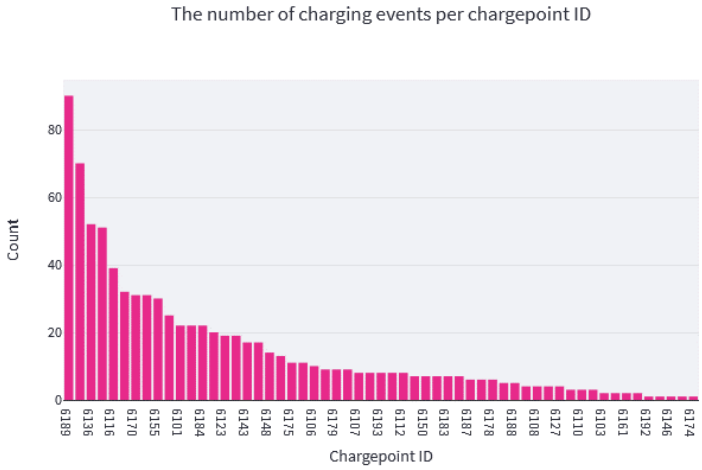


Figure 131: The number of charging events per chargepoint ID

Figure 131: Few of the charge points are more preferred over the others. The first 8 charge points have minimum of 30 transactions and the last 5 have one transaction each.

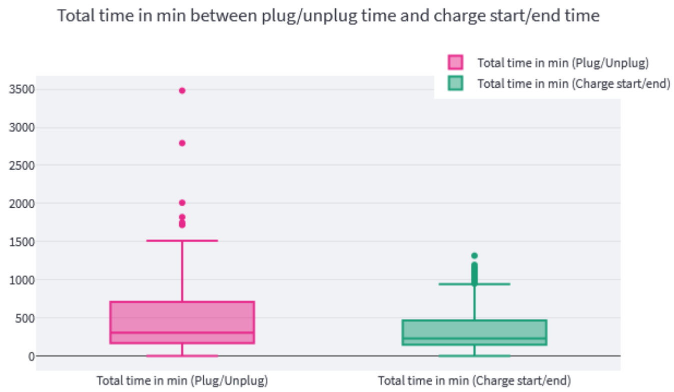


Figure 132: Total time in min between plug/unplug time and charge start/end time

Figure 132: From this comparation, it can be seen that the total time for charging is spread over a smaller data range compared to the total time the vehicle is plugged in. The outliers for the charging are closer to the max than for the plug/unplug. *Min-median* is very concentrated for both times compared to *median-max* which is more spread-out having higher variability.

For plug/unplug time, 50% of the data is between 170 to 707.5. For charging start/end, 50% of the data is between 146 to 466.5. Both boxplots are right-skewed.

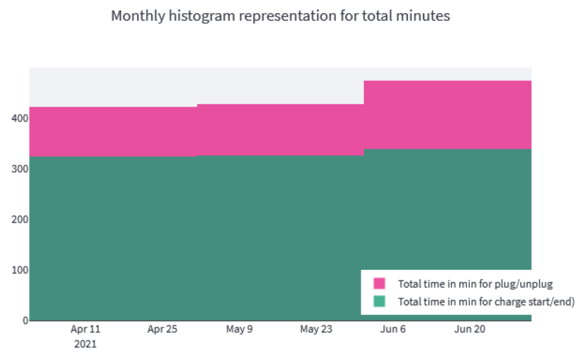


Figure 133: Monthly histogram representation for total minutes

Figure 133: The histogram shows that there is a high difference (~100 to 135 minutes) between the time (in minutes) for plug/unplug and charging start/end. The time becomes bigger for June. For April-May, the difference almost stagnates.

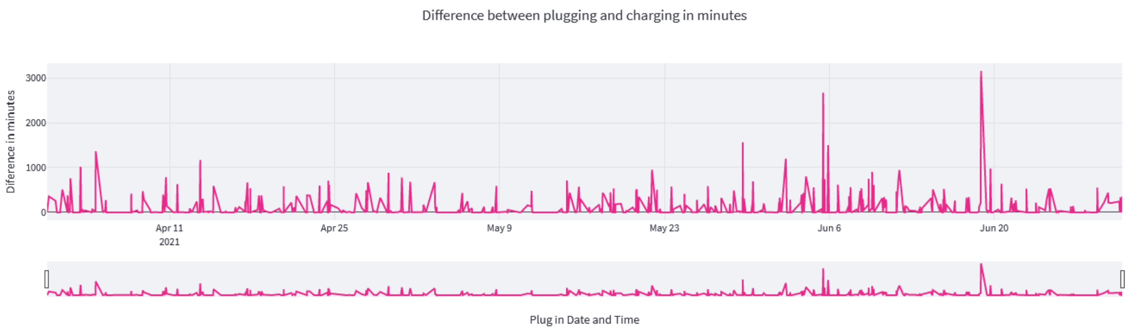


Figure 134: Difference between plugging and charging in minutes

Figure 134: This graph tries to put in perspective the difference based on time. Few spikes can be seen toward the end of the graph. Overall, an objective pattern cannot be fully said. There seems to be a pattern for small spikes followed by very low difference, sort of.

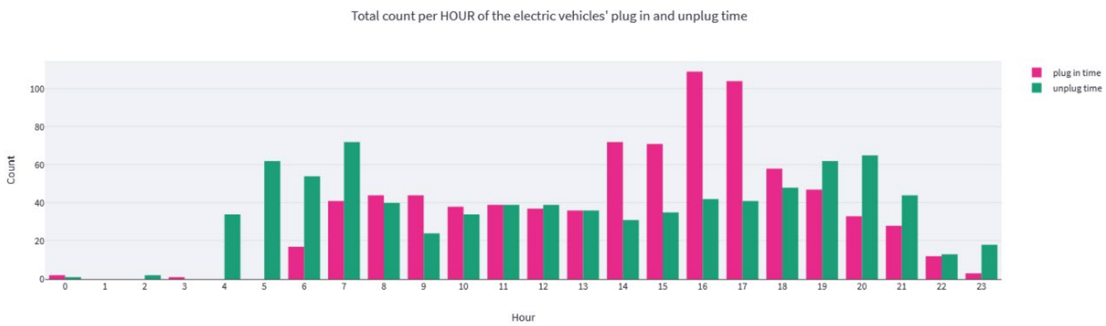


Figure 135: Total count per HOUR of the electric vehicles' plug in and unplug time

Figure 135: Most popular hours for unplugging the vehicles are 4 to 7AM followed by 7PM to 11PM. Most popular hour for plugging in the vehicles is predominantly just after mid-day 2 to 6PM. 2 charging events or below for 12AM-3AM for both events. It seems that for morning to mid-day (8AM-1PM) the 2 charging events are quite similar.

Figure 136: Hour plug-in/unplug per each charge point ID

Figure 136: As it can be seen from both graphs, some of the charging point locations have a charging event almost every hour (eg: 6101, 6116, 6118, 6132, 6136, 6151, 6169, etc.) whereas some of the locations have a very limited number of charging events (eg: 6103, 6111, 6125, 6146, 6159, 6192, etc.).

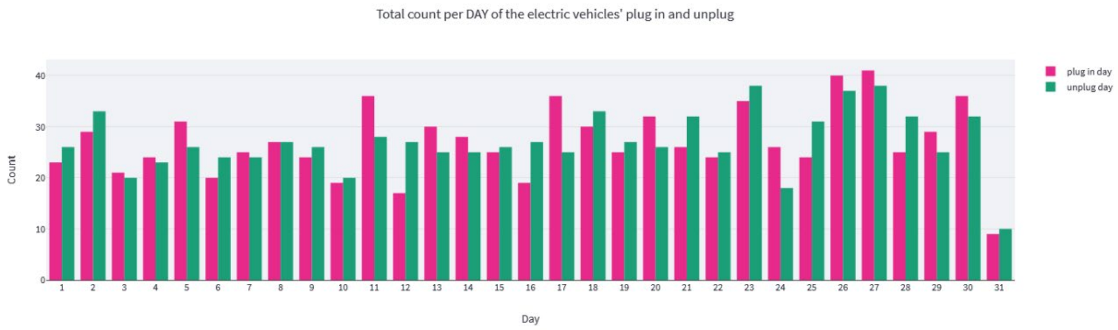


Figure 137: Total count per DAY of the electric vehicles' plug in and unplug

Figure 137: When looking at how many vehicles have been plugged-in/unplugged in a day, it seems that it balances out over a few days. Meaning that if the number of plug-in vehicles was bigger than the number of unplugged vehicles in the beginning, after a few days, the roles reverse (as logically you would expect). Overall, the biggest values for plug-in day seem to be the 11th, 17th, 23rd, 26th, 27th and 30th, whereas for unplugged there are the 2nd, 18th, 23rd, 26th, 27th and 30th. Being almost an overall or very close by.

The most unpopular dates for plug-in are the 6th, 10th, 12th, 16th and 31st whereas for unplugging are the 3rd, 10th, 24th, 31st.

Figure 138: Day plug-in/unplug for each charge point ID grouped by month

Figure 138: For April, charging point ID 6189 has the most transactions overall. For May, charging point IDs 6132 and 6189 have the most transactions overall. For June, charging point IDs 6132, 6155, 6170 and 6189 have the most transactions overall. July and March have only 3 respectively 2 transactions across all charging point IDs.

Overall, for April-June there are charging points IDs with only one transaction overall.

From the presented graphs, it can be concluded that there are 1 to 4 charging point IDs with the most usage overall per month. Charging point ID 6118 looks like it used daily overall, but for individual months, the usage is not that high.

Citizen → Location RR

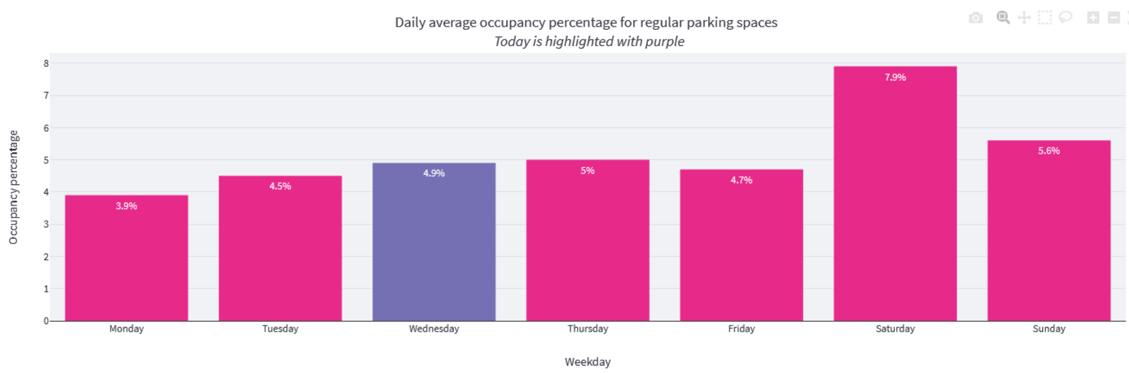


Figure 139: Daily average occupancy percentage for regular parking spaces

Figure 139: Most popular days are Saturday followed by Sunday. The least popular are Monday and Tuesday.

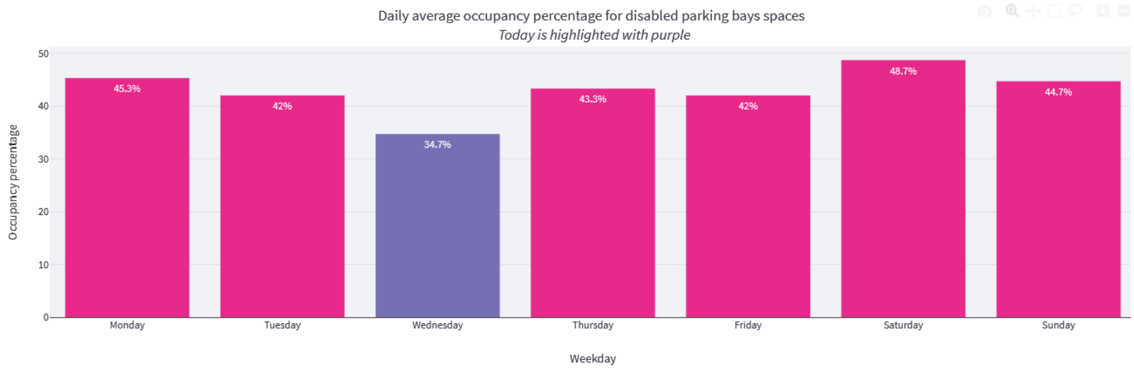


Figure 140: Daily average occupancy percentage for disabled parking bays spaces

Figure 140: A true reflection cannot be told because the dataset is based on *almost randomly* generated data. But the graph shows that Wednesday is the least popular with Saturday and Monday being the most popular.

Citizen → Location A

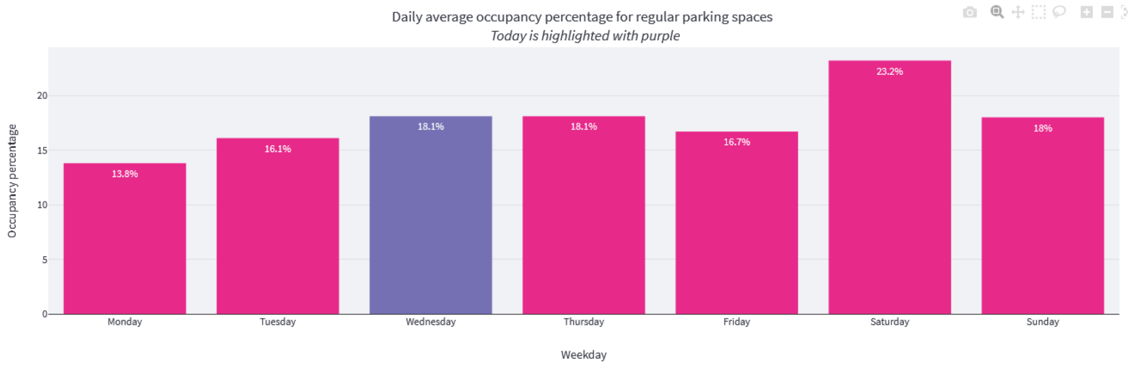


Figure 141: Daily average occupancy percentage for regular parking bays spaces

Figure 141: Saturday is the most popular followed by Wednesday and Thursday. Monday and Tuesday are the least popular.

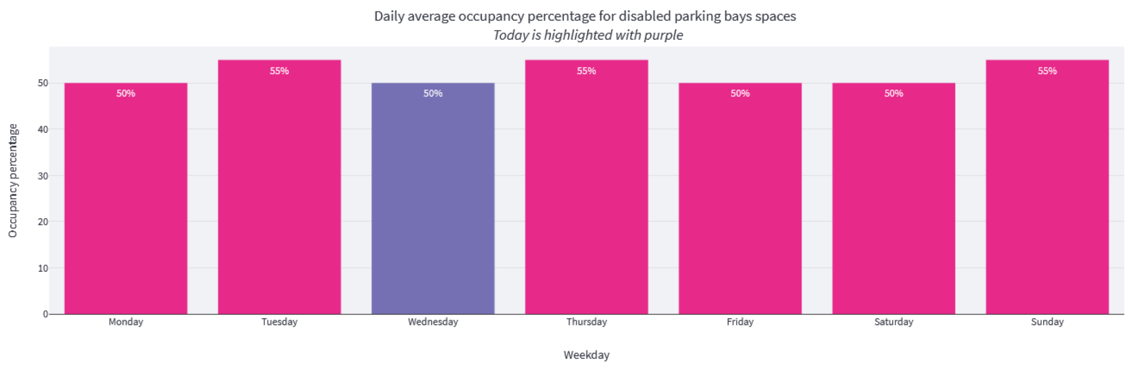


Figure 142: Daily average occupancy percentage for disabled parking bays spaces

Figure 142: A true reflection cannot be told because the dataset is based on *almost randomly* generated data. But the graph shows that Tuesday, Thursday and Sunday are the most popular with the remaining of the days being the least popular.