**A logo for college computing

Description automatically generated**

**Assessment Cover Page**

|  |  |  |
| --- | --- | --- |
| ***Student Full Name: Conor Kirwan*** |  | |
| ***Student Number: 24130*** |  | |
| ***Module Title: DVT*** |  | |
| ***Assessment Title: Summer Repeat CA*** | |  |
| ***Assessment Due Date: 27/July/2024*** |  | |
| ***Date of Submission: 24/July/2024***  ***Word Count: 1625*** |  | |

**Declaration**

By submitting this assessment, I confirm that I have read the CCT policy on academic misconduct and understand the implications of submitting work that is not my own or does not appropriately reference material taken from a third party or other source.

I declare it to be my own work and that all material from third parties has been appropriately referenced.

I further confirm that this work has not previously been submitted for assessment by myself or someone else in CCT College Dublin or any other higher education institution.

**Contents**

[Title**:** 3](#_Toc204262231)

[Introduction: 3](#_Toc204262232)

[Exploratory Data Analysis (EDA): 4](#_Toc204262233)

[Dashboard Design and Visualisation Choices: 6](#_Toc204262234)

[Interactivity and Functionality: 10](#_Toc204262235)

[Interpretation and Justification: 11](#_Toc204262236)

[Future Work: 12](#_Toc204262237)

[References 12](#_Toc204262238)

[Appendix 1 – Github 12](#_Toc204262239)

[Appendix 2 – Old Dashboard 13](#_Toc204262240)

[Appendix 3 – New Dashboard 15](#_Toc204262241)

[Appendix 4 – Student contribution 15](#_Toc204262242)

# Title**:**

CA3 - Fremont Bicycle Data Dashboard

# Introduction:

In this report, an exploratory data analysis (EDA) was performed using bicycle data from the “Fremont Bridge Bicycle Counter” dataset. From this, an interactive user-friendly dashboard was built in order to analysis & visualize some of the trends that were uncovered, with the aim being to inform stakeholders of bike usage patterns over a prolonged time period to help urban planners with decisions in relation to sustainable transport infrastructure.

The aim of this assignment is to reveal underlying patters that may influence hourly commuting behaviour, such as seasonality, weekdays vs weekend, and the direction of the flow of traffic over a prolonged time period. All of the insights were gathered from python generated plots post EDA, and a range of modifications were applied to better showcase the data via a dashboard, with the ability to filter for year, weekend/weekday, and direction of traffic flow completed. The dashboard aimed to be functionally robust, visually appealing and interpretable to a wide range of stakeholders to better provide data insights around bicycle infrastructure planning in the Fremont area.

The research identifies implementation issues, describes each visualisation choice, and makes recommendations for future development in order to address this challenge. Feedback from earlier tasks was given particular consideration, with particular emphasis on avoiding bullet points for justifications and making sure that every explanation was important.

A screenshot of a computer

AI-generated content may be incorrect.

Figure : Interactive dashboard layout deployed via Voilà, showing traffic filters and chart arrangement.

# Exploratory Data Analysis (EDA):

The dataset was first loaded int a Jupyter notebook to gather insights into the shape, info & description. It was noted that there were 4 columns, which included the date/ time, journeys east, journeys west, and total journeys. To better make the data interpretable, it was decided to relabel the column names for interpretability, as seen in Figure 2.

A screenshot of a computer

AI-generated content may be incorrect.

Figure : df loading & initial df exploring

The next step taken was to handle missing values. As can be seen in Figure 3, there were 100 rows which contained null values, and the df contained 111695 rows. It was decided that any row that contained a null value would be removed from the df as this accounts for less than .1% of the rows, so it would not skew the data. The df also only began records at the end of 2012, and the middle of 2025, but this data was decided to be kept. The time column was further separated by the day, week, & month, and further algorithms were performed to determine if a day was a weekend or not.

A screenshot of a computer code

AI-generated content may be incorrect.

Figure : null values in df

The initial insights revealed some interesting trends that were

* Seasonality – during the summer months (May – September) bicycle counts increased, while during winter months (Oct – Apr) bicycle counts decreased significantly.
* COVID – There was a sharp decline at the onset of the pandemic, and in the years following 2020, pre pandemic peaks as seen in 2019 have not yet materialized.
* 24 Hour Peaks – There was an influx in the total number of bicycles that used the bridge during peak times, which was 8am & 5pm, which indicates this is a primary route used by “9-5” / office workers.
* Directional Bias – There was more commuters using the East bound lane rather than the West bound lane, and this is continuous for every year.
* Weekday vs Weekend – Weekday usage was much higher than weekend usage in terms of total bicycles using the bridge. The weekend generally tended to be flatter, i.e. consistent usage over the day, with no real “peak” usage.

These insights provided the basis for selecting visualisation types and configuring interactivity to suit the end-user’s needs. Beyond just confirming expectations, the EDA also revealed year-over-year fluctuations that were not linear — post-pandemic recovery in 2022–2024 did not follow the same trajectory as the pre-pandemic growth. This can allude to behavioural shifts or the long-term effects of working remotely. Additionally, compared to weekdays, weekend traffic varied more, indicating that non-commute cycling was more susceptible to outside influences like the weather or events. Anomalies in hourly trends — such as unexpected midday peaks on certain weekends — also signalled possible one-off events or inconsistencies in counter operation. These nuances, while subtle, highlight the need for city planners to take a layered, flexible approach to interpreting cycling data.

# Dashboard Design and Visualisation Choices:

The dashboard utilized six key visualisations arranged in a 3x2 layout using Matplotlib and Seaborn. A uniform colour scheme that was acceptable to all stakeholders was used throughout to convey meaning. To further ensure uniformity, east traffic is displayed in green (#2a9d8f), west traffic is displayed in red-orange (#e76f51), and all traffic uses the same green. These colours were chosen with care for colour-blind users as well as aesthetics and accessibility (Kirk, 2019).

The monthly traffic over time chart used a line plot to demonstrate cyclical patterns and long-term trends. Line plots are effective for showing seasonality and continuity, and in this case clearly highlighted COVID-era dips. The weekday bar chart illustrated categorical comparisons between days. Early drafts used a gradient, but this was revised to a single tone per category, as differences in height are a more effective and less misleading representation of volume.

A graph of a graph of a bicycle traffic over time

AI-generated content may be incorrect.

Figure :line plot showing monthly bicycle traffic trends and seasonal peaks (above EDA, below dashboard)

A graph of a number of blue and green bars

AI-generated content may be incorrect.

Figure : bar plot comparing total bicycle counts across weekdays and weekends (above EDA, below dashboard)

For directional insights, average hourly traffic charts were used to display East vs West patterns using dual-line plots. This helped establish infrastructure pressure points. A separate plot comparing weekday to weekend usage over a 24-hour period was chosen to highlight behavioural shifts between workdays and leisure days.

A graph of a number of different types of traffic

AI-generated content may be incorrect.

Figure : line plot comparing hourly East and West traffic volumes (above EDA, below dashboard)

A graph of a diagram

AI-generated content may be incorrect.

Figure : comparison of weekday and weekend hourly traffic showing distinct commuting and leisure patterns(above EDA, below dashboard)

The year-over-year total count bar chart allowed stakeholders to visually correlate cycling volume with policy changes or external events. Scientific notation was automatically applied for values over 1 million. While this helped conserve space, further user testing could determine whether this abstraction is suitable for all audiences. The weekday-hour heatmap summarised dense data in an accessible grid, clearly highlighting the 8 AM and 5 PM commute peaks.

A graph of different colored bars

AI-generated content may be incorrect.

Figure : yearly total bicycle counts, highlighting the impact of COVID-19 and post-pandemic recovery (above EDA, below dashboard)

A graph showing a number of rectangles

AI-generated content may be incorrect.

Figure : heatmap visualising peak commuting hours across weekdays

Throughout all visuals, design consistency was emphasised. Fonts, legends, and axis labels were manually formatted for legibility. Visual clutter was avoided, and tick marks were configured to avoid overlap, especially when interacting with filters. Accessibility was also a consideration: contrasting colours, appropriately sized text, and intuitive layout all contribute to usability across age and skill demographics. Dynamic graph titling was also employed depending on which filters were selected.

A screenshot of a graph

AI-generated content may be incorrect.

Figure : dynamic title based off direction selected

One particularly difficult design decision involved selecting the heatmap colour scheme. Multiple pallets were explored however, "YlGnBu" was selected due to its intuitive range from light to dark, which effectively communicates intensity. "Inferno" and "magma" were more explored but lacked the subtle contrast needed for readability for a large range of stakeholders. Likewise, bar widths and spacing were manually changed to ensure all bars were legible regardless of the number of years/ categories being displayed.

# Interactivity and Functionality:

In order to created functionality, interactive elements were created using “ipwidgets”. The following filters were created

* “Year Range” – This can be adjusted by sliding between the range the user desires
* “Traffic Type” – This can filter between distinct traffic patterns (weekend & weekday)
* “Direction” – This can filter between East, West & Total bikes across a multiple visualizations

By controlling the dashboard state across all visualizations, these filters made it possible to extract insights unique to a given scenario. The design of the filters was keeping with the sustainability them, and a consistent colour scheme with shade of green chosen in keeping with the plots colour scheme.

A screenshot of a computer

AI-generated content may be incorrect.

Figure : interactivity features colour scheme(before EDA, after dashboard)

Voilà worked well with Jupyter in terms of responsiveness and interpretability. However, some performance issues arose due to six subplots layout (chosen 3x2). Some of the problems encountered were filters not applying uniformly across all plots, as well as rendering issues. This brought the need to establish global states for filtering and meticulous debugging within the Jupyter environment. This was however fixed upon final dashboard completion.

# Interpretation and Justification:

The reason behind each of the visualizations was based on the findings encountered during the initial EDA and creating a dashboard capable of communicating these findings to a range of stakeholder with a varying level of technical expertise. . Time series line charts revealed long-term and seasonal patterns. Bar charts provided intuitive category comparisons. Heatmaps compactly presented large volumes of time-categorised data, and line plots were used to uncover dual-axis relationships, such as direction or temporal flow.

The choices are in line with best practices as discussed by (Few, 2009) & (Ware, 2013) where representation visually is a function of position, length & colour to represent data dimensions without increasing data misinterpretation. The dashboard shows trends like cycle patterns, how they chang over time, and what differences exist between different parameters e.g. East vs West Lane usage. This allows planners to address items such as demand peaks, utilize underused infrastructure, or align planned works with observed patterns.

An example would be the contrast on weekday vs weekend bike usage. The line plot of the total traffic by day of week suggests a drop-off during the week, however both the heatmap & line plot of the average hourly traffic show much flatter midday peaks on the weekend when compared to the sharp peaks as seen during the weekday. This enforces the narrative that weekday traffic is predominantly commuter driven usage, whereas the weekend would tend to be more leisure-based usage. This trend is significant for both cultural programming and infrastructural development; depending on this usage, cities may choose to increase access to greenways or hold weekend bicycle festivals.

Additionally, the discrepancy between East vs West Lane usage of the bridge raises further queries. Environmental factors such as lighting, surface conditions, access/agress could be reasons for the East side's higher usage during the peaks. This insight highlights additional queries that could prompt a follow-up by city planning agencies, even though the dataset does not explain why. As a result, the dashboard functions as both an endpoint and a tool for generating hypotheses.

# Future Work:

Future versions might have real-time dashboards through PowerBI, weather integration, and hover functionality. Usability and context would be improved with user input and spatial overlays. Planning could be further aided with export capabilities and more precise time filters.

# References

Few, S. (2009) *Now You See It: Simple Visualization Techniques for Quantitative Analysis*. Oakland, Calif: Analytics Press.

Kirk, A. (2019) ‘Data Visualisation : A Handbook for Data Driven Design’, pp. 1–328.

Ware, C. (2013) *Information Visualization: Perception for Design*. Elsevier.

# Appendix 1 – Github

*Please note that an incorrect GitHub link was initially used during development. As the assignment requires a history of meaningful commits, I have retained the original repository for continuity. However, the correct and final version of this project is now available at:*

[*GitHub Project Link*](https://github.com/sba24130/SSBD_Summer_Repeat)

*All files, code updates, and documentation relevant to this assignment are reflected in this updated repository.*

# Appendix 2 – Old DashboardA screenshot of a graph AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

# Appendix 3 – New Dashboard

A screenshot of a computer

AI-generated content may be incorrect.

# Appendix 4 – Student contribution

A blue circle with text

AI-generated content may be incorrect.