



# TUBE TWIN (PASSENGER COUNT FORECASTING/GENERAL TUBE ANALYSIS)

Executed by  
INTERN TEAM-6 2201AC

Insights for Better Decision Making

# INTERN TEAM-6 2201AC

## Our Philosophy

Our emotional connection to the world reflects a blended passion for analytics and advocacy.

## Our Mission

To empower and inspire with the most trusted analytics.

## Our Vision

To transform a world of data into a world of intelligence.

## Our Values

The collective power of our values influences everything we do.

# MEET INTERN TEAM-6 2201



**Fielami Emmanuel  
David**  
Project Lead



**Hakim Balogun**  
Administrative Lead



**Harmony Odumuko**  
Technical Support



**Michael Mamah**  
Technical Lead



**Kelvin Mwaniki**  
Data Scientist



**Endurance Arienkhe**  
Data Analyst

# TABLE OF CONTENTS

- Introduction
- The Problem
- Our Solution
- Project Components
- Project Demo
- Conclusion & Recommendations



0 0 0 0



# INTRODUCTION

**Let's get into it!**





# The Company

- Transport for London (TfL)
- The London Underground (aka “The Tube”)





---

# PROBLEM STATEMENT

**What are we trying to solve?**





# PROBLEM STATEMENT

- Network operations and surveillance.
- Traffic flow, analyses, and control.
- Infrastructure management and maintenance.





# OUR SOLUTION

**Let's take a look at what we did.**





## OUR SOLUTION

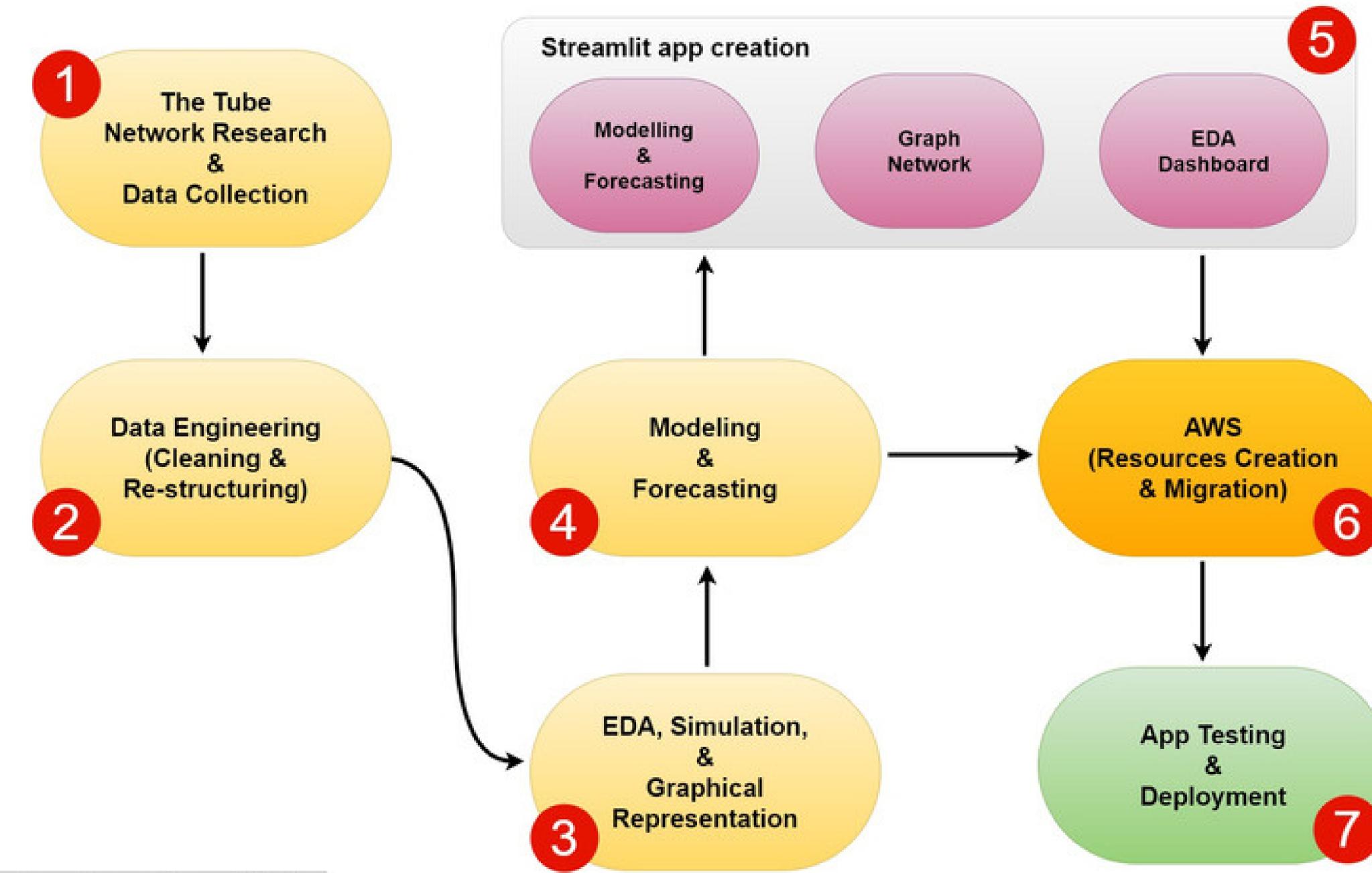
- Graph representation and simulations of the tube network.
- Model to forecast passenger count at each station based day of the week and time of the day.
- Dashboard for tube analyses to provide insights for business decision-making.





# THE PROJECT ROADMAP

TubeTwin Project Flowchart



Explore Intern Team-6 2201



# PROJECT COMPONENTS

A deeper dive into our solutions.





# DATA SOURCING & ENGINEERING

**Data sources** - TfL, Wikipedia.

## Available Data

- Historical Tube Data (2017 - 2021)
- Train Loading Data
- LU Lines and Stations Data

## Re-engineering

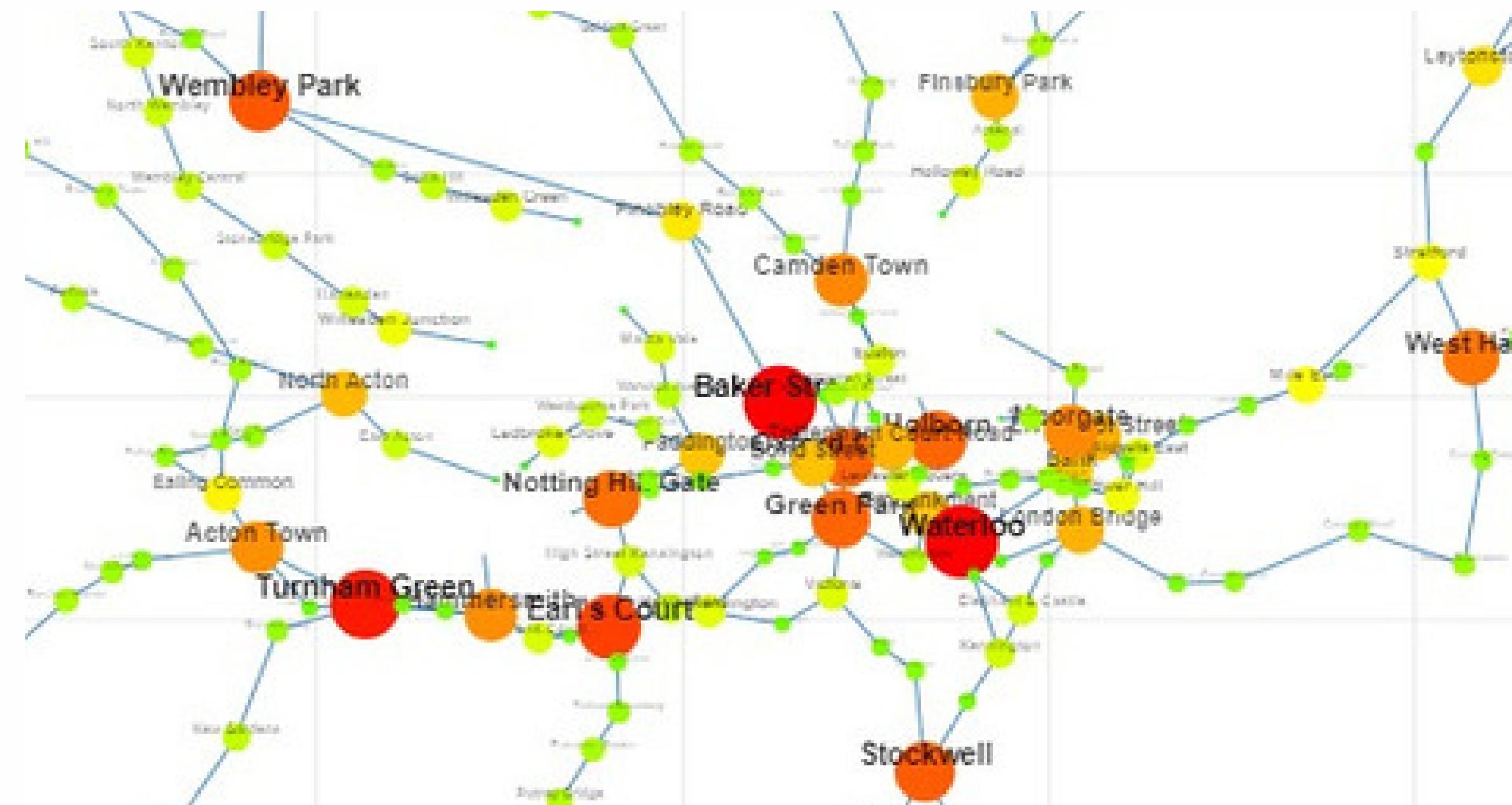
- Merging
- Cleaning
- Re-structuring





# GRAPHICAL NETWORK OF THE TUBE

Graph showing stations based on their rankings. Red nodes are more important and green nodes are the less important ones.





# EXPLORATORY DATA ANALYSIS





## DATASETS EXPLORED

### Passenger Station Entry Counts :

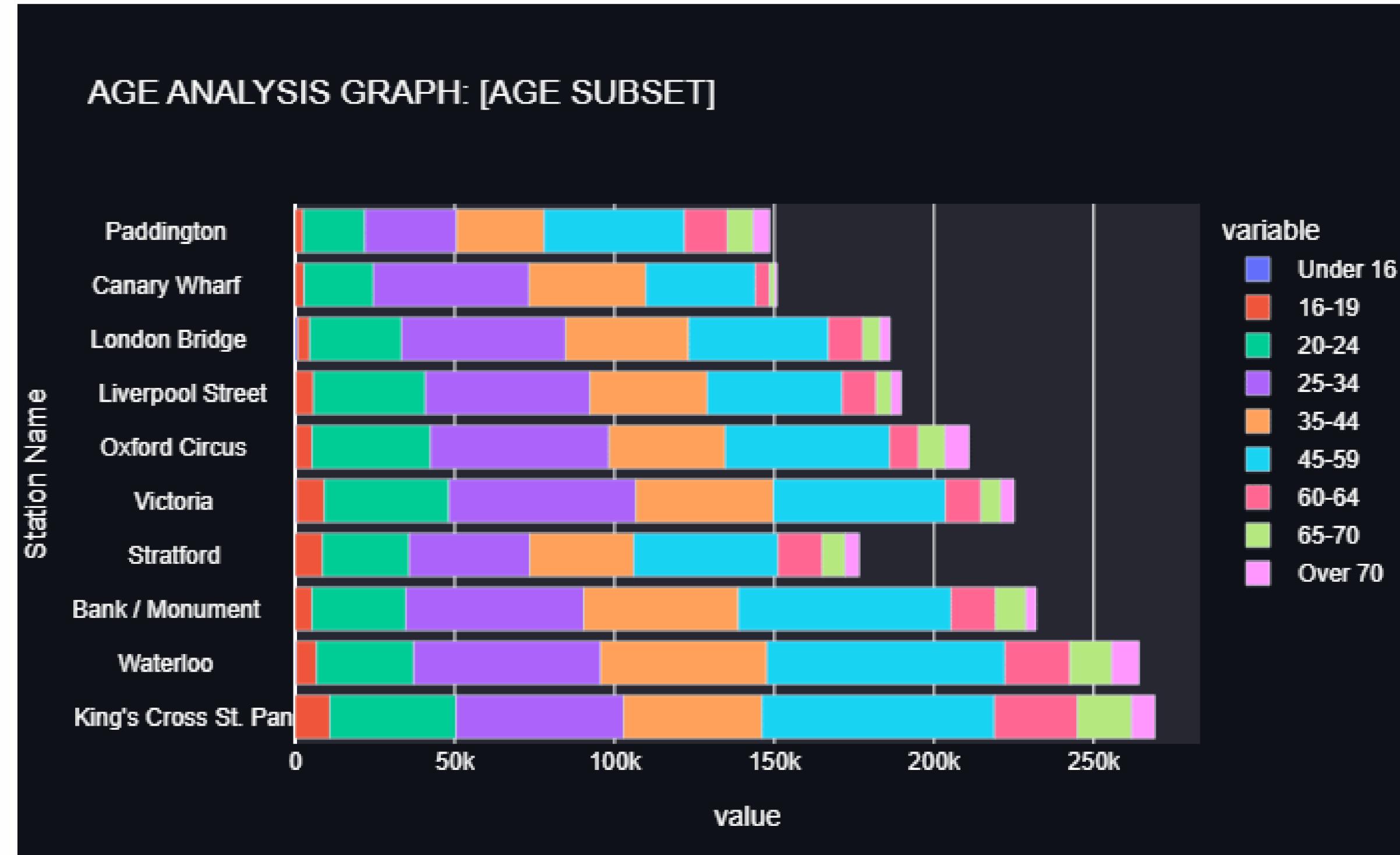
1. Age, Gender & Mobility
2. Average Journey Time
3. Distance Travelled
4. Journey Purpose
5. Ticket Type

### Passenger Station Exit Counts :

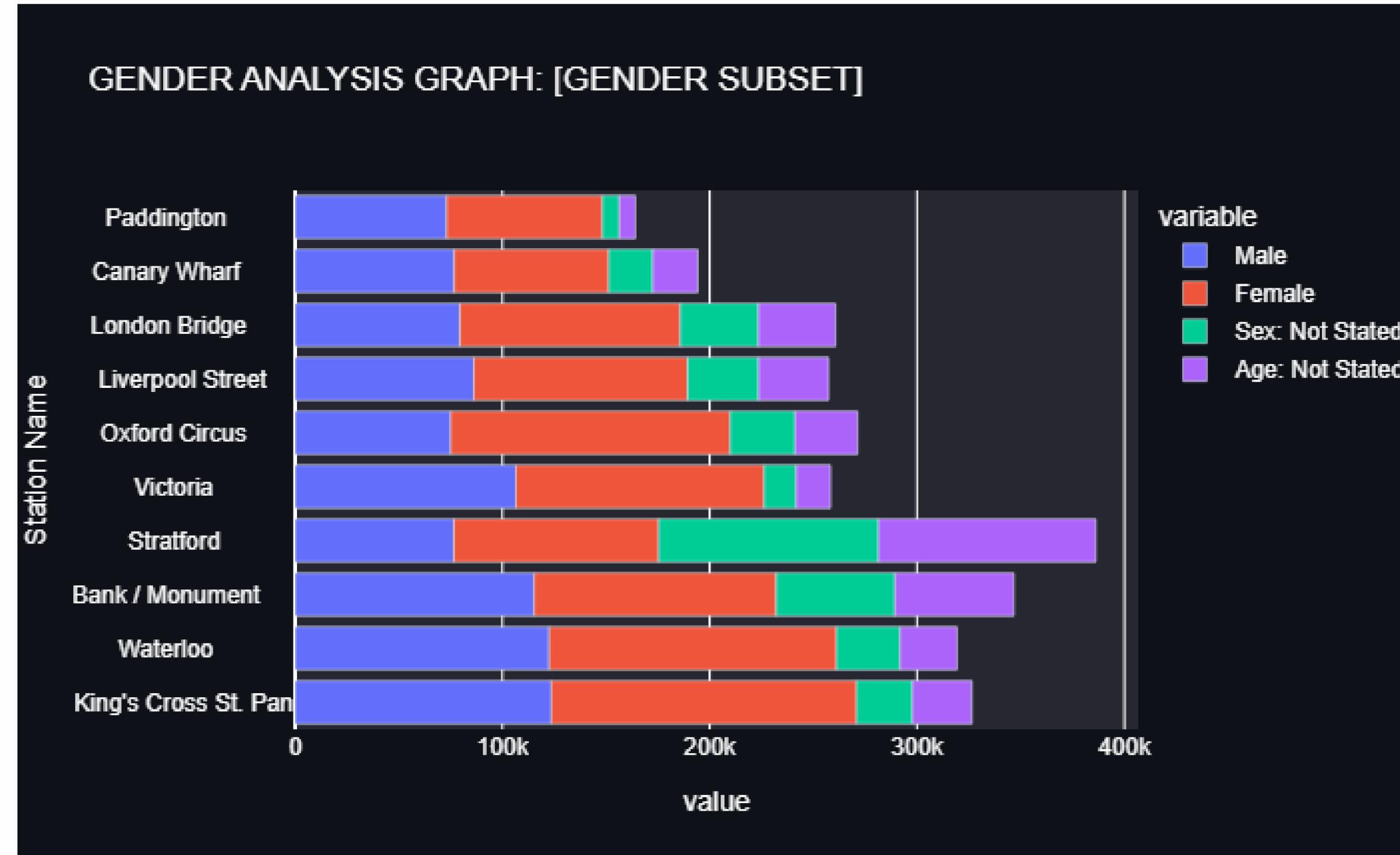
1. Age, Gender & Mobility
2. Average Journey Time
3. Distance Travelled
4. Journey Purpose
5. Ticket Type
6. Egress



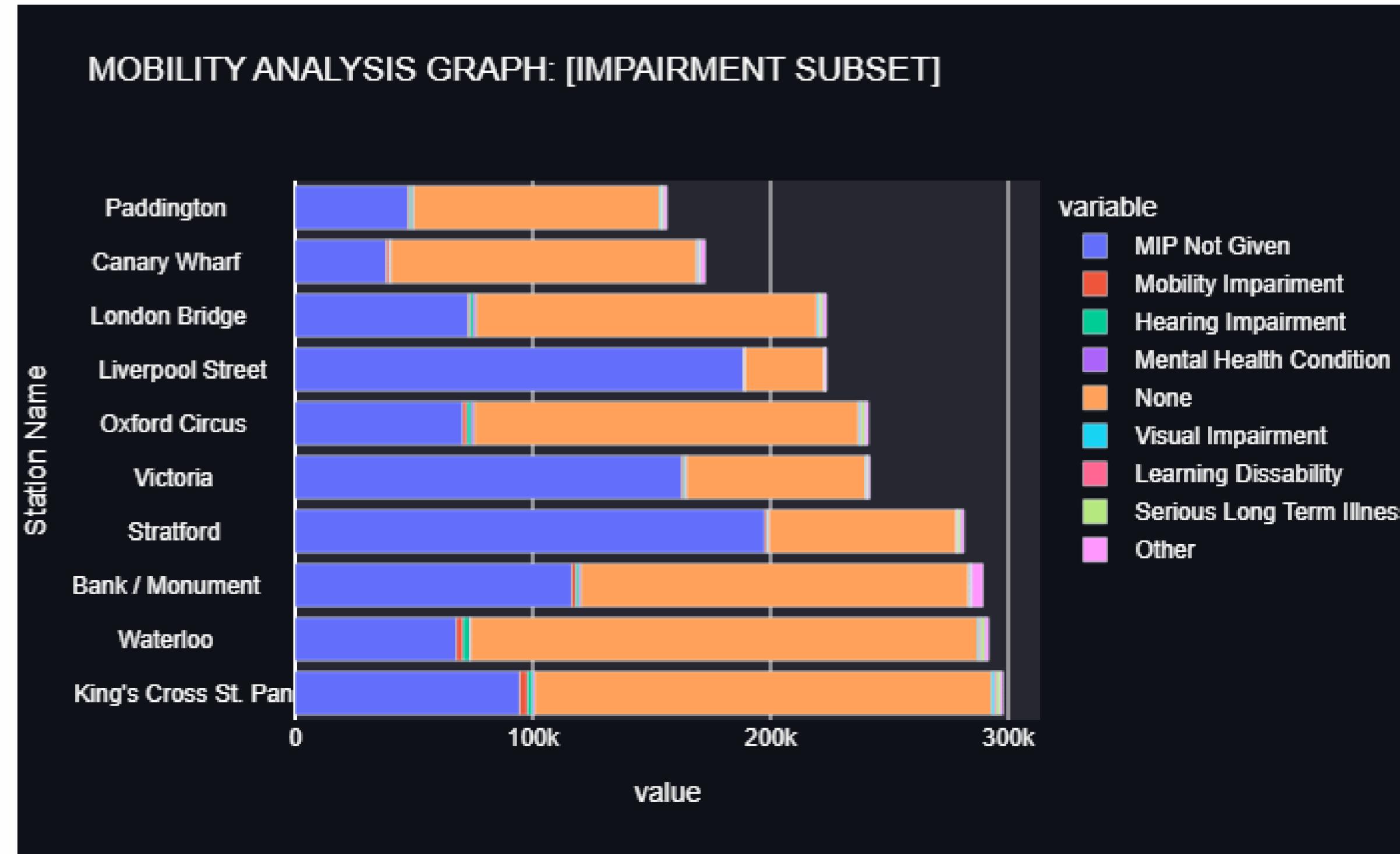
## ENTRY AGE, GENDER AND MOBILITY DATASET: [AGE] SUBSET



## ○ ○ ○ ENTRY AGE, GENDER AND MOBILITY DATASET: [GENDER] SUBSET



# OOOO ENTRY AGE, GENDER AND MOBILITY DATASET: [MOBILITY/IMPAIRMENT] SUBSET





# MODELLING

Let's take a look at the tools we used.

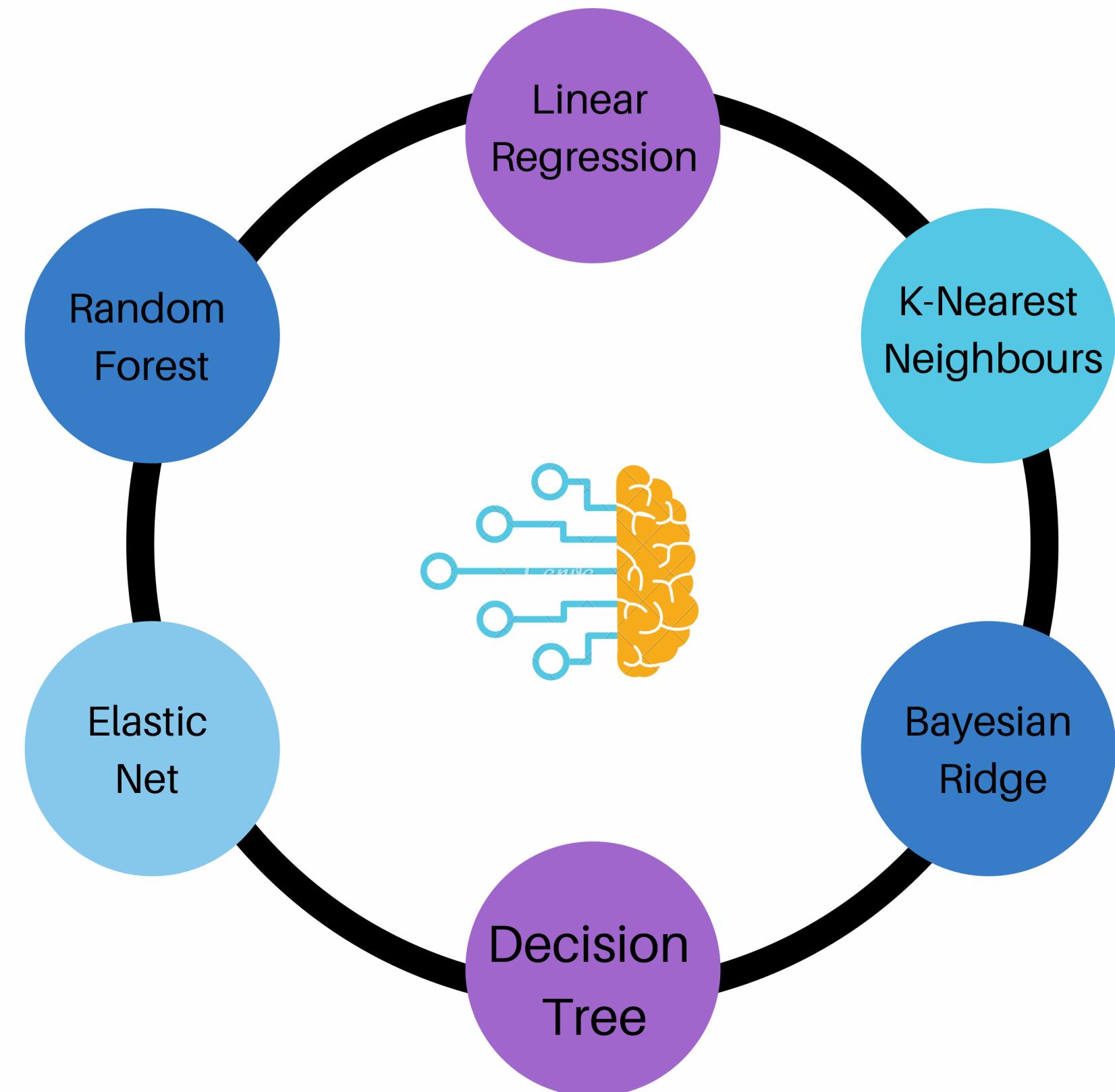


LINEAR REGRESSION



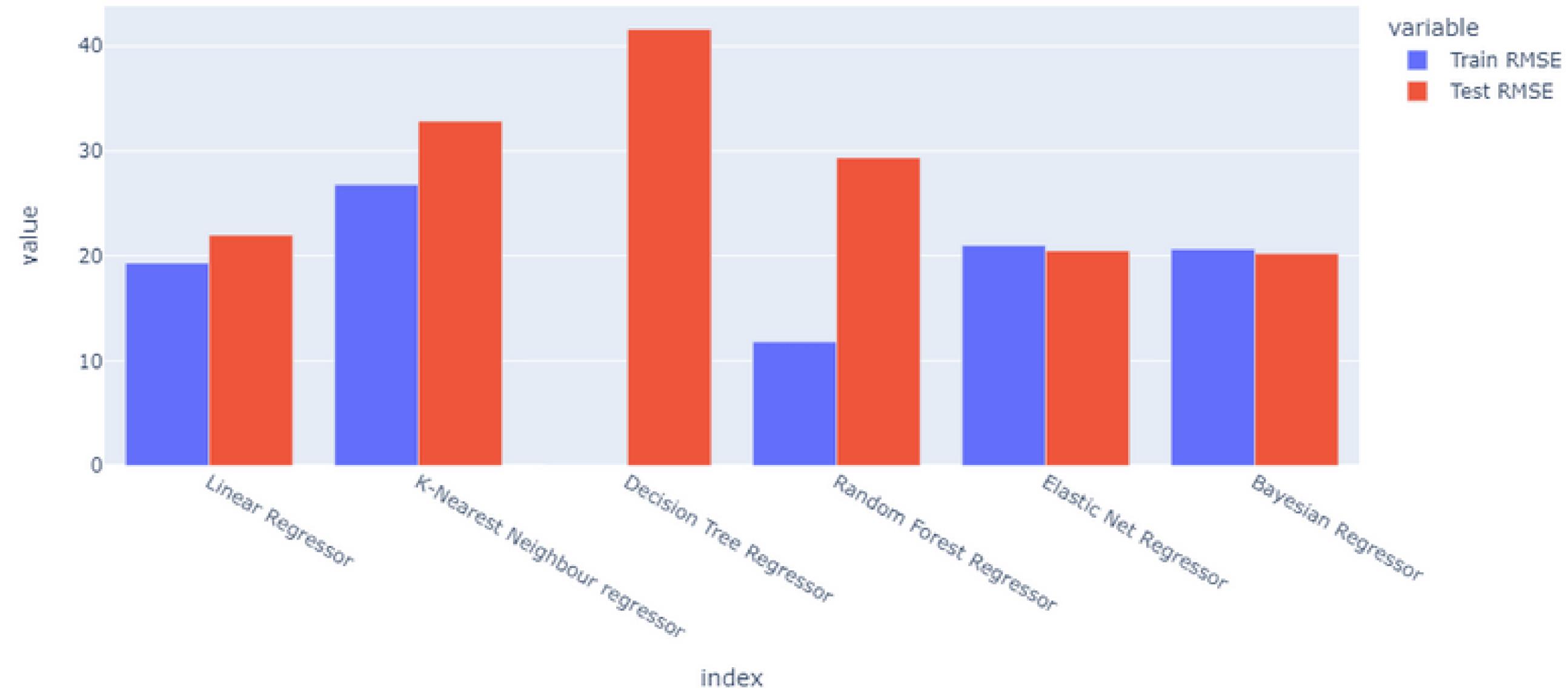
TIME SERIES

# REGRESSION MODELS USED



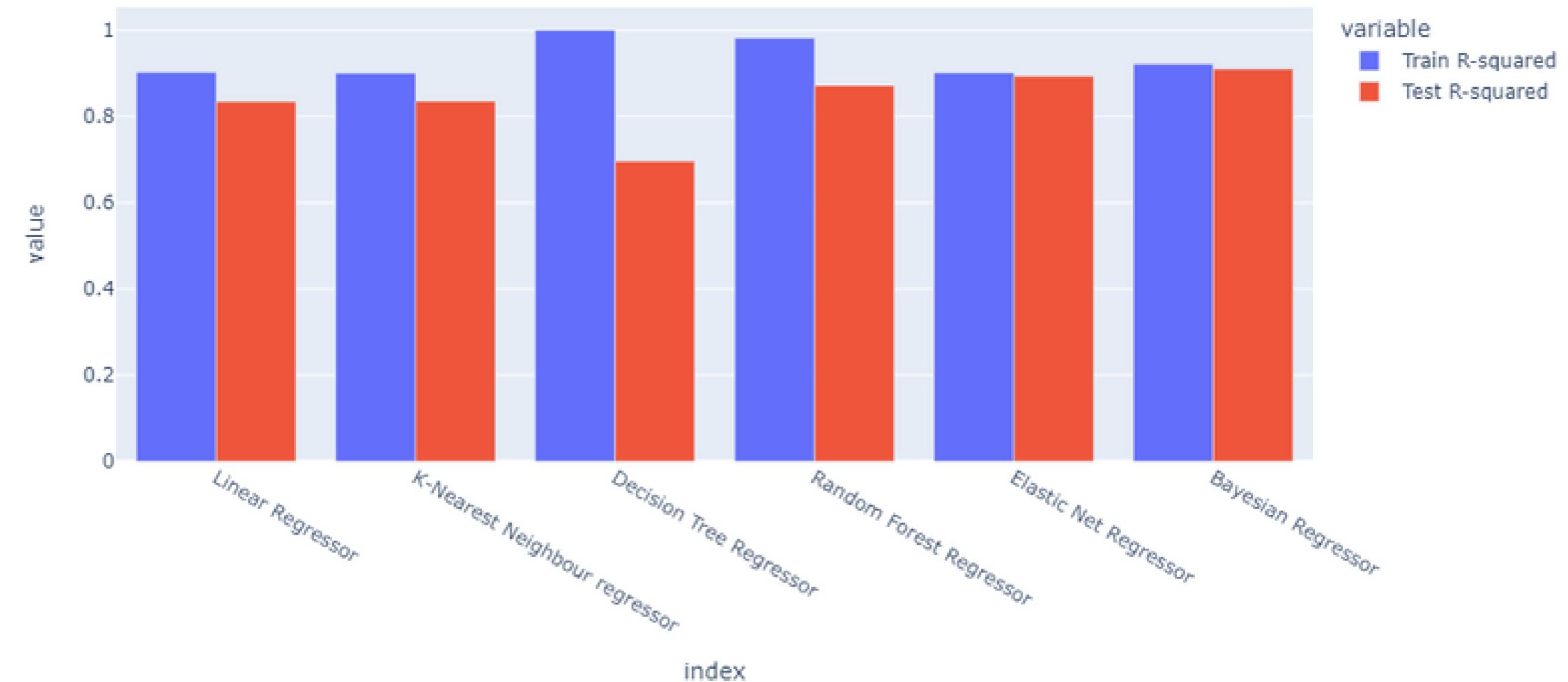
# Model Performance Metrics

TRAIN-TEST RMSE PERFORMANCE METRICS GRAPH



# Model Performance Metrics

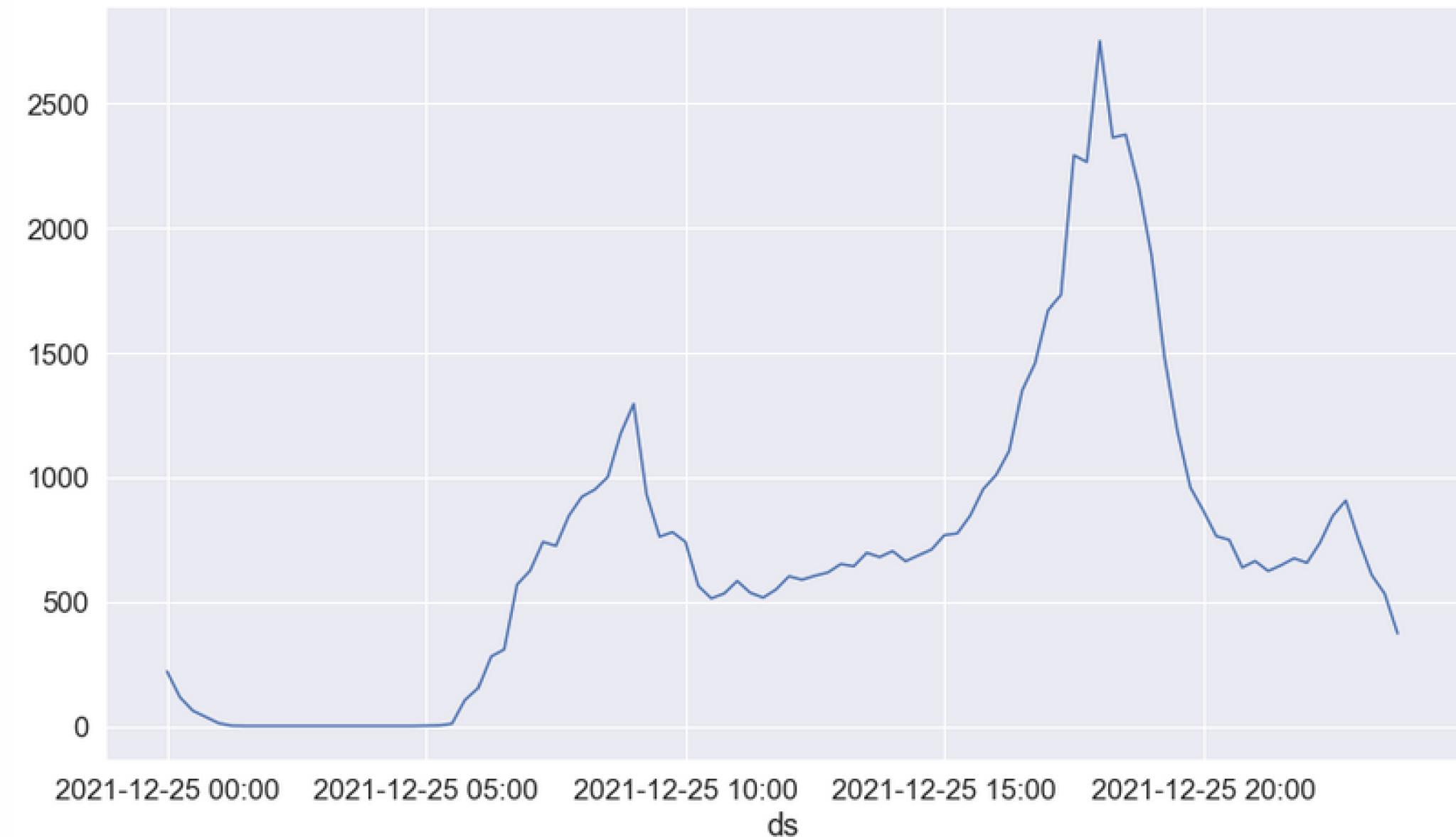
TRAIN-TEST R SQUARED PERFORMANCE METRICS GRAPH





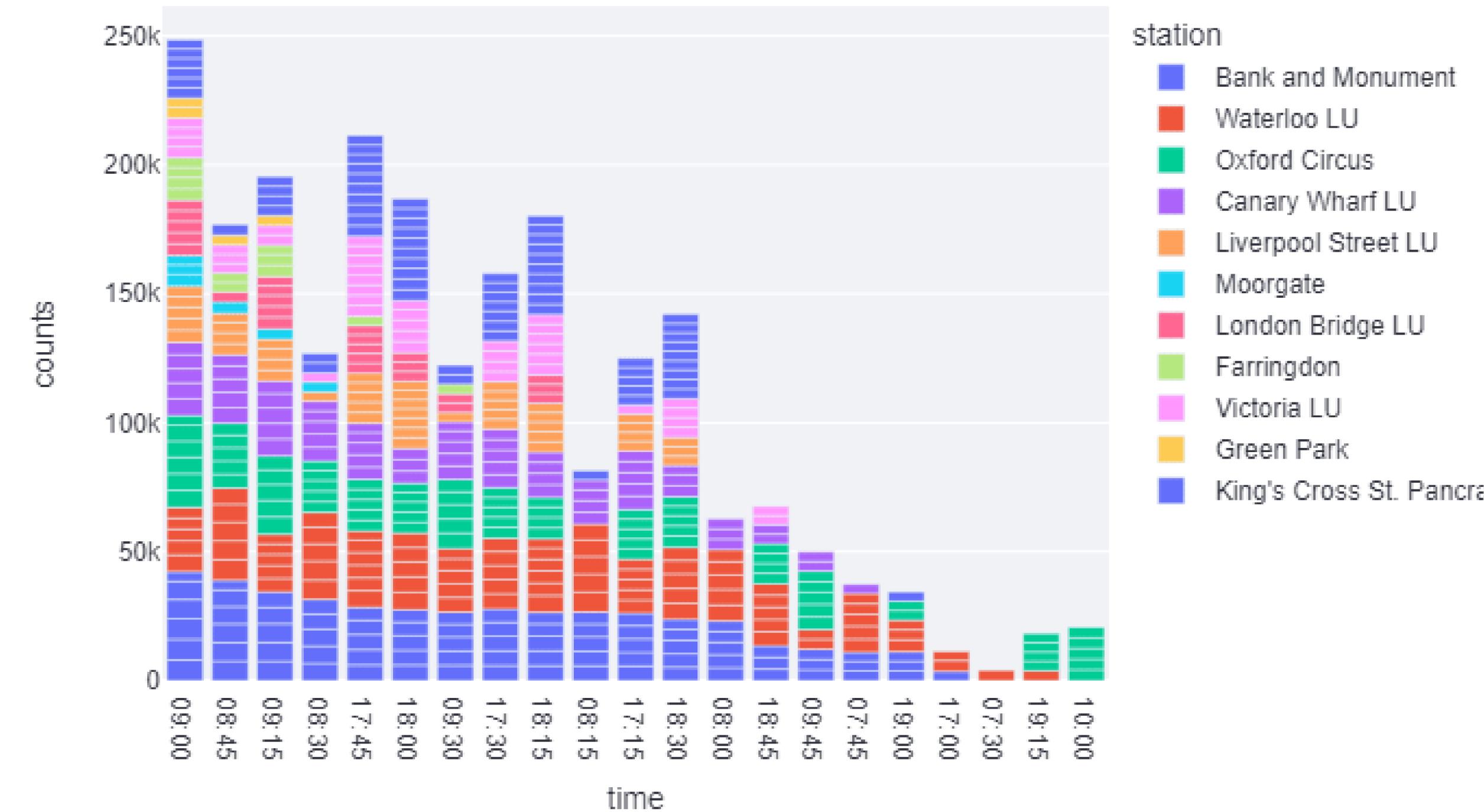
Time series analysis is a specific way of analyzing a sequence of data points collected over an interval of time

# Time Series Analysis and Modelling (FBProphet)



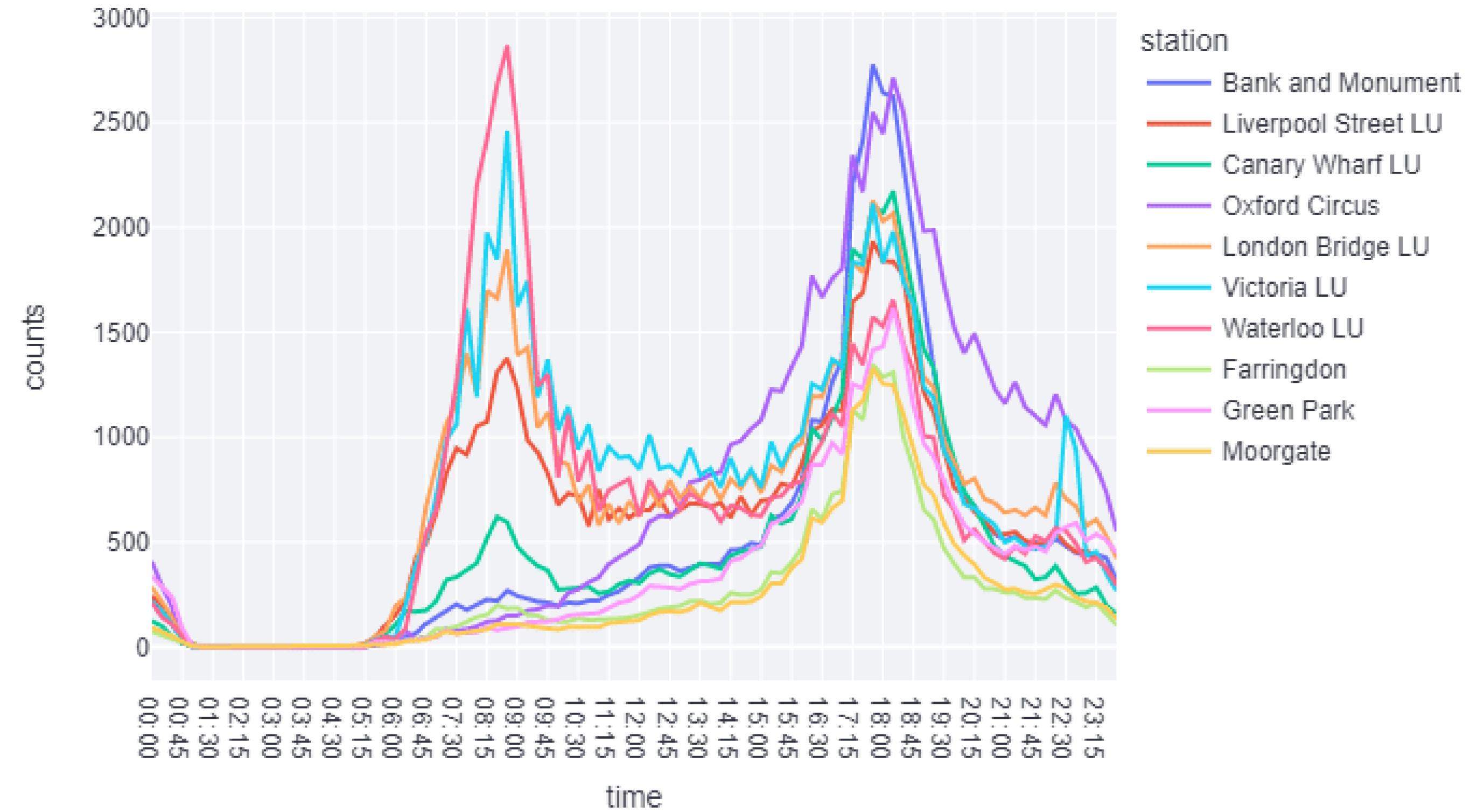


Using an interactive Bar Chart, we can visualize these stations and look at the Time associated with the high traffic per stations.



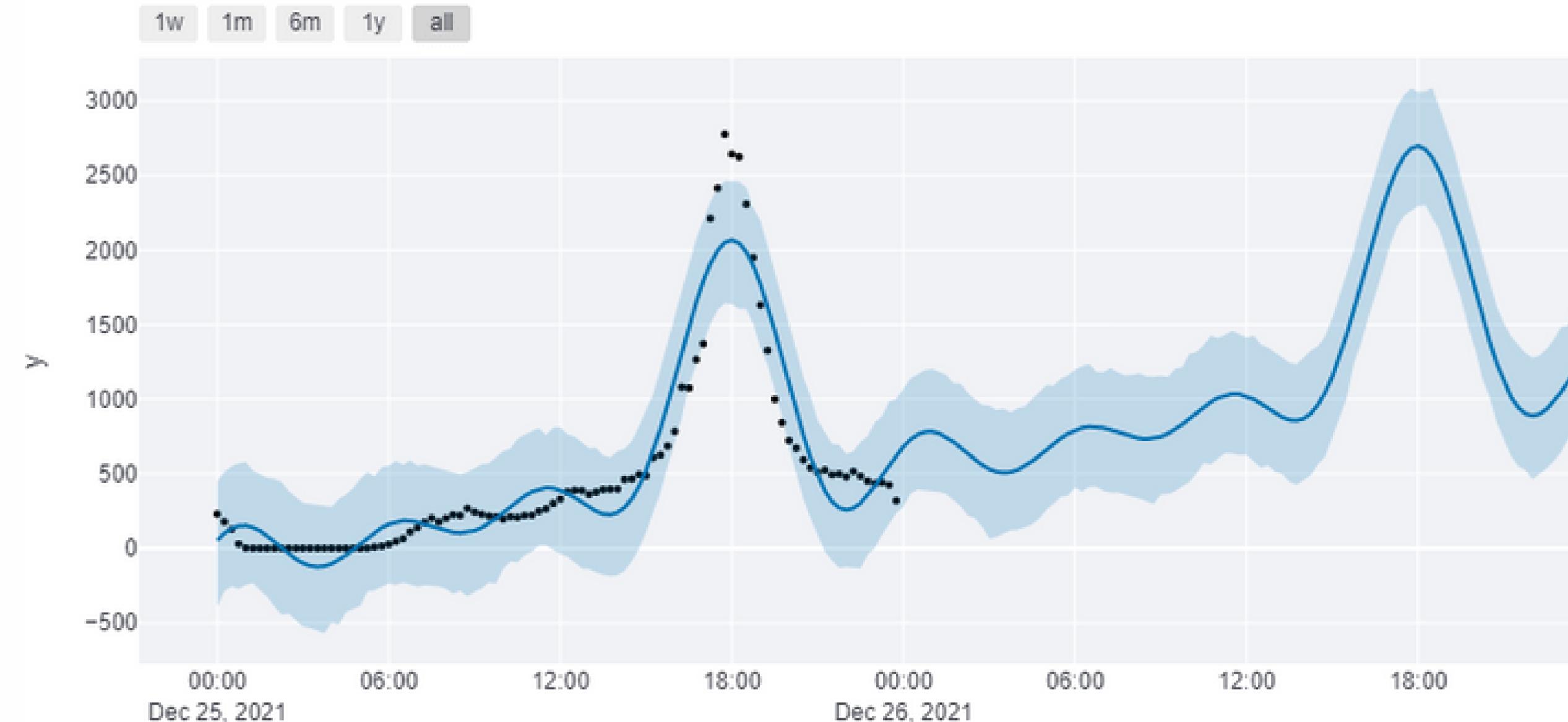


## 2021 Monday to Thursday inflow dataset



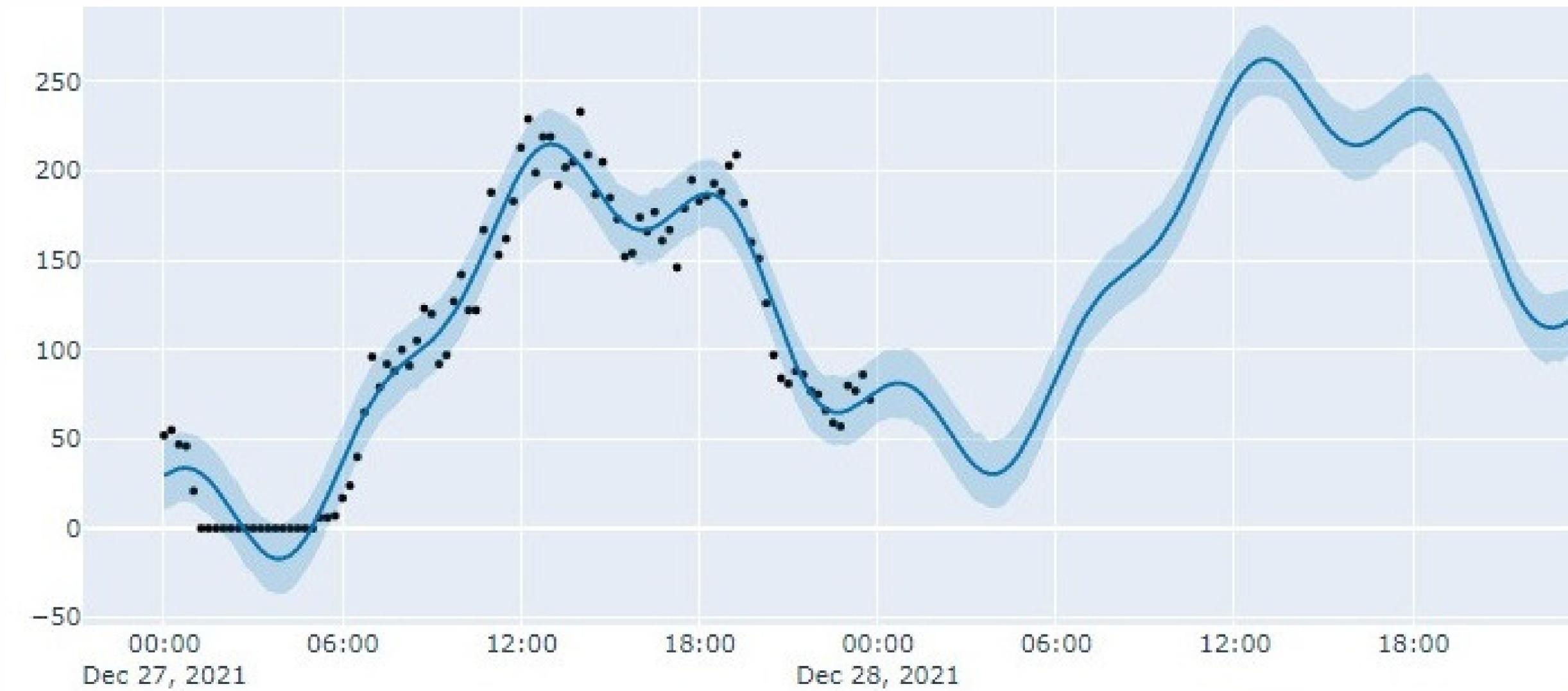


## Passenger inflow forecast on MTT for Bank and Monument





## 15-minute interval passenger outflow forecast on Staurdays for Moorgate station





# Time Series Model

## Reasons

- Forecasting feature
- Time based model analysis
- Greater potential accuracy





# PROJECT DEMO

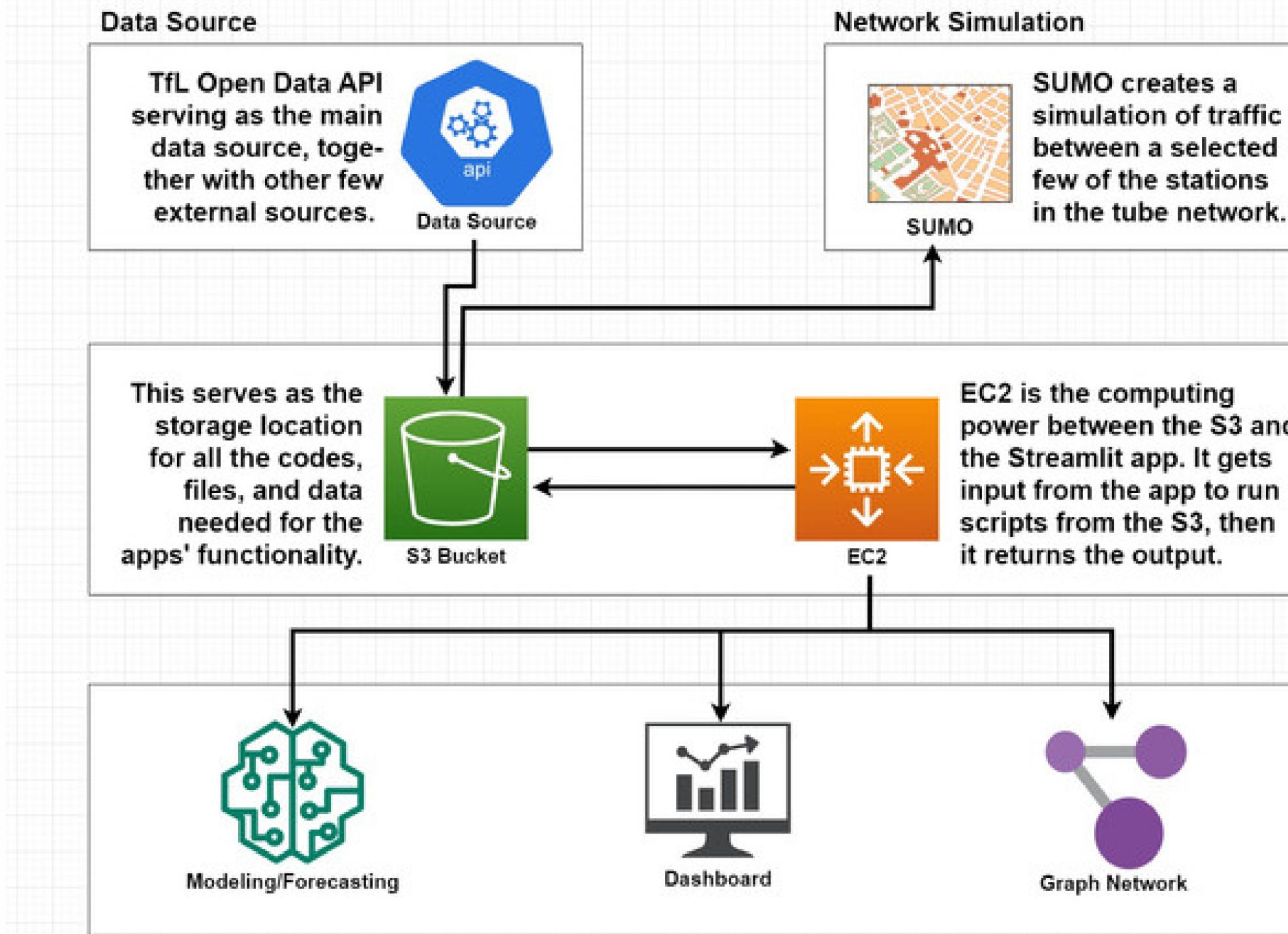
**Telling time is over, now it's show time!**





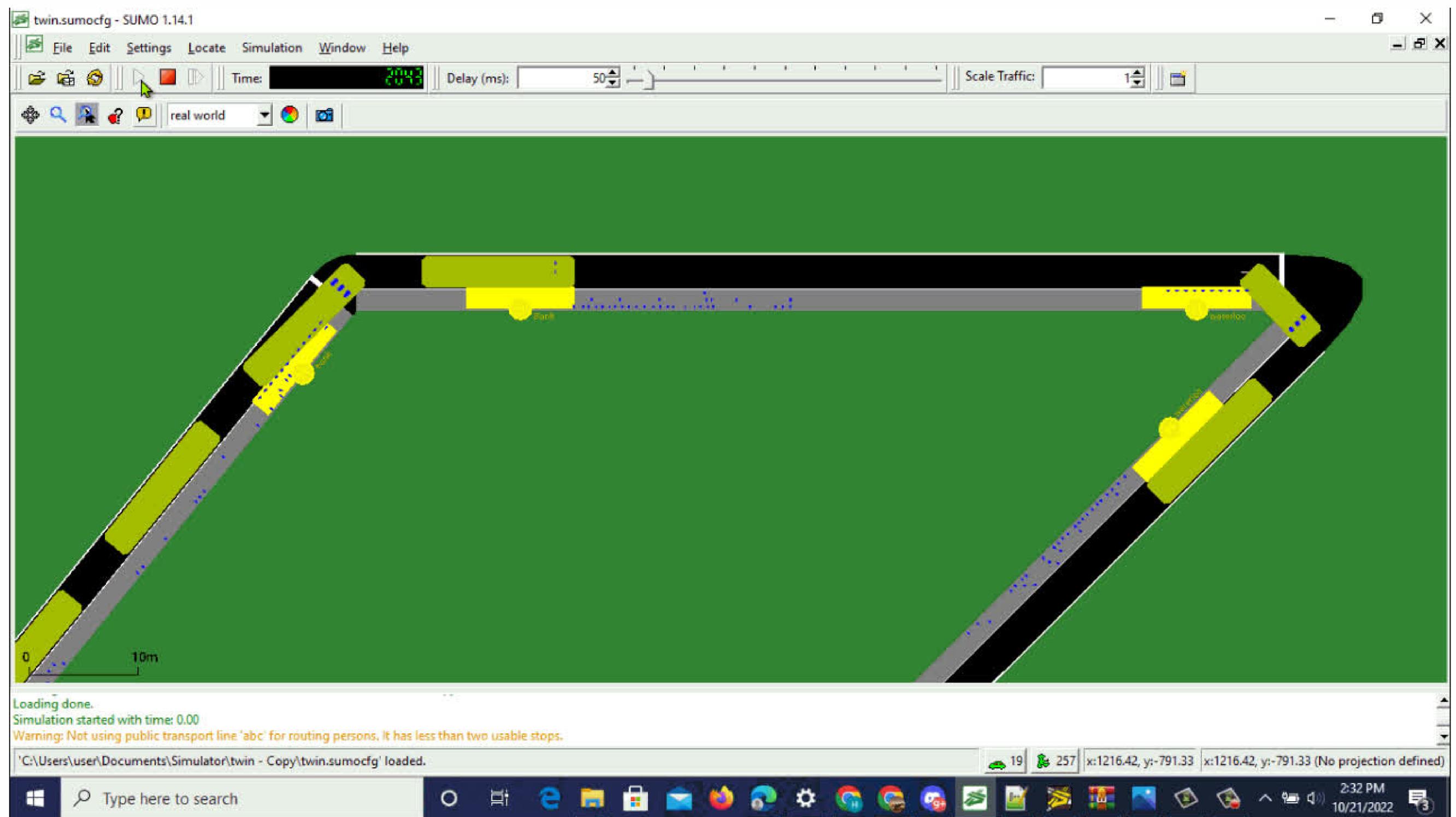
## Tube Twin Architectural Diagram

Explore Intern Team-6 2201





# TRAIN SIMULATION





# APP DEMO

The screenshot shows a mobile application window titled "THE LONDON UNDERGROUND". On the left, there is a sidebar with a red London Underground circular logo and a blue bar containing the text "THE LONDON UNDERGROUND". Below this, a "PAGE SELECTION" dropdown menu is open, showing the following options:

- Landing Page (selected)
- Landing Page
- Live Feed
- Modelling
- Explorative Data Analysis (EDA)
- Train Simulation
- About Team

The main content area displays a photograph of a London Underground train at a station platform. The train is silver and red, with "Heathrow" written on its front. Above the train, a digital display board shows "Piccadilly Line" and the route "1. Heathrow 2. Uxbridge" with travel times "1 min" and "2 min". The background features a blurred view of the station platform and ceiling.



## CONCLUSION

- ◆ The deployed web application adequately covered the top 14 stations for passenger forecasting.
- ◆ The general time series data analysis and the graphical network representation provided an overall analysis.
- ◆ This app can support TFL in achieving many of its Tube management objectives which includes traffic control.



## RECOMMENDATION

- ◆ Further research and improvement of passenger flow and traffic analyses on the London Tube.
- ◆ Expansion in the scope of passenger forecasting .
- ◆ Live streaming data collection of the tube for a minimum period of one year.



**THANK YOU!**  
**IT'S Q & A TIME!**



