



TUBE TWIN (PASSENGER COUNT FORECASTING/GENERAL TUBE ANALYSIS)

Executed by
INTERN TEAM-6 2201

Insight for better decision making

INTERN TEAM-6

2201

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.

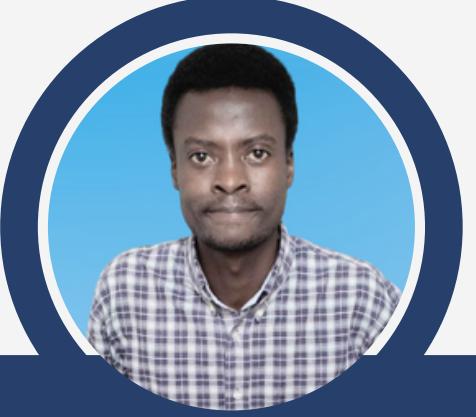
Our Values

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
Administrative Support



Endurance Arienkhe
Data Analyst

TABLE OF CONTENTS

- Introduction
- The Problem
- Our Solution
- Project Components
- Project Demo
- Conclusion
- Recommendation



0 0 0 0



INTRODUCTION

We're going to introduce you to why we are here today.





....

INTRODUCTION

The London Underground (aka “The Tube”) is a network of train stations which connects the city. Understanding the behaviour of this tube network will help the management make key decisions on improving the network and offering better customers experience.





PROBLEM STATEMENT

Lets take a look at what we have





PROBLEM STATEMENT

London has a population of approximately 9 million people, on top of which hundreds and thousands of tourists visit it every day. This presents some challenges to TfL concerning operations of the tube railway network.





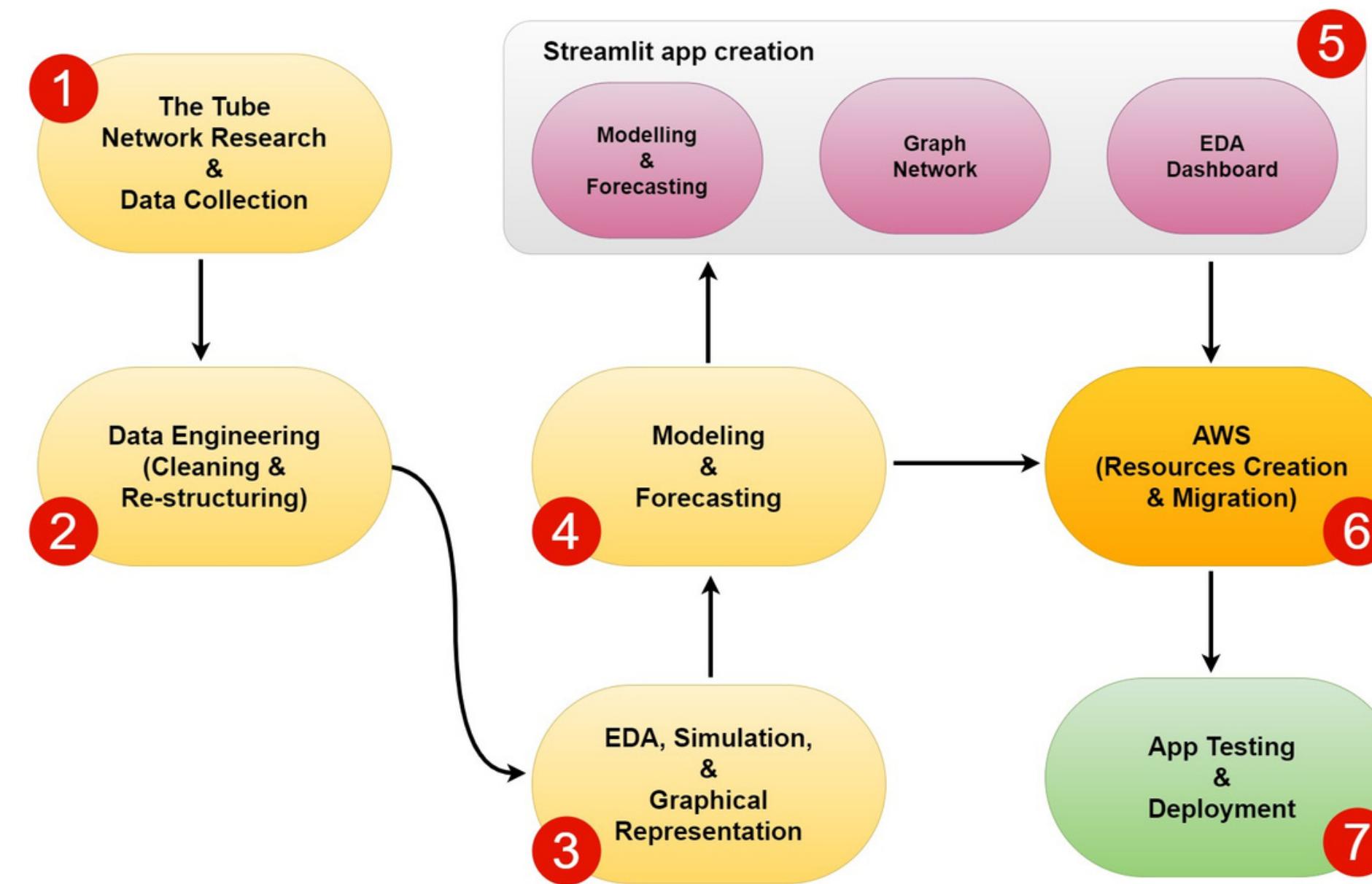
Operational Challenges of the London Tube

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



TUBE-TWIN (R&D) PROJECT ROADMAP

TubeTwin Project Flowchart



Explore Intern Team-6 2201



OUR SOLUTION

Lets take a look at what we did.





SOLUTION

Use of digital twins in systems design has the following advantages:

- Graph representation of the network and Simulation
- Passenger count forecast at each station (entry and exit), in 15 minutes increments, based on time of day, and counts at other stations
- Dashboard showing key insights to railway technicians and commuters





PROJECT COMPONENTS

Lets take a look at what we have





Data Engineering





....



Exploratory Data Analysis

What is Exploratory Data Analysis?

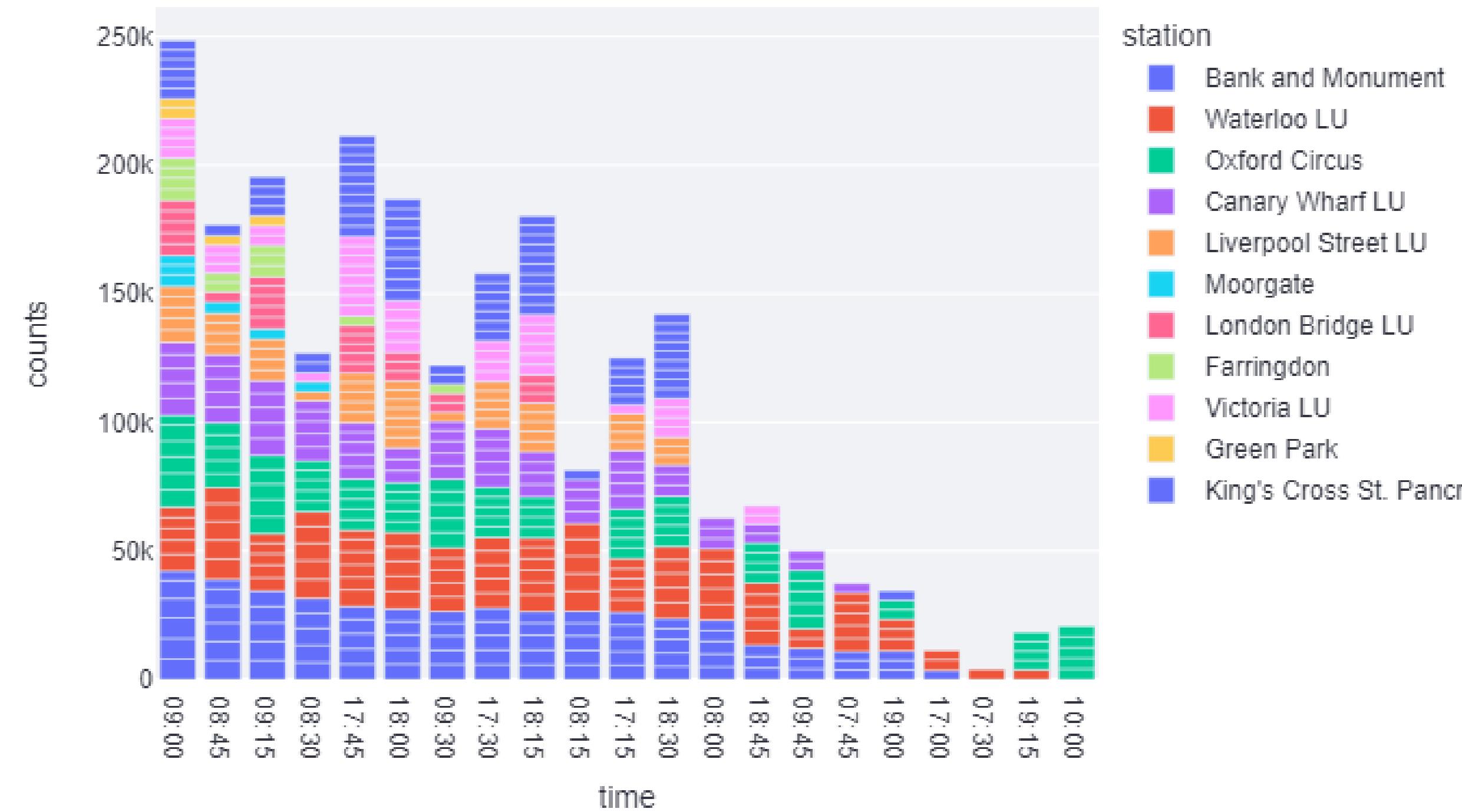
Use statistical tools to investigate datasets

The purpose is to find out what is happening in our data.



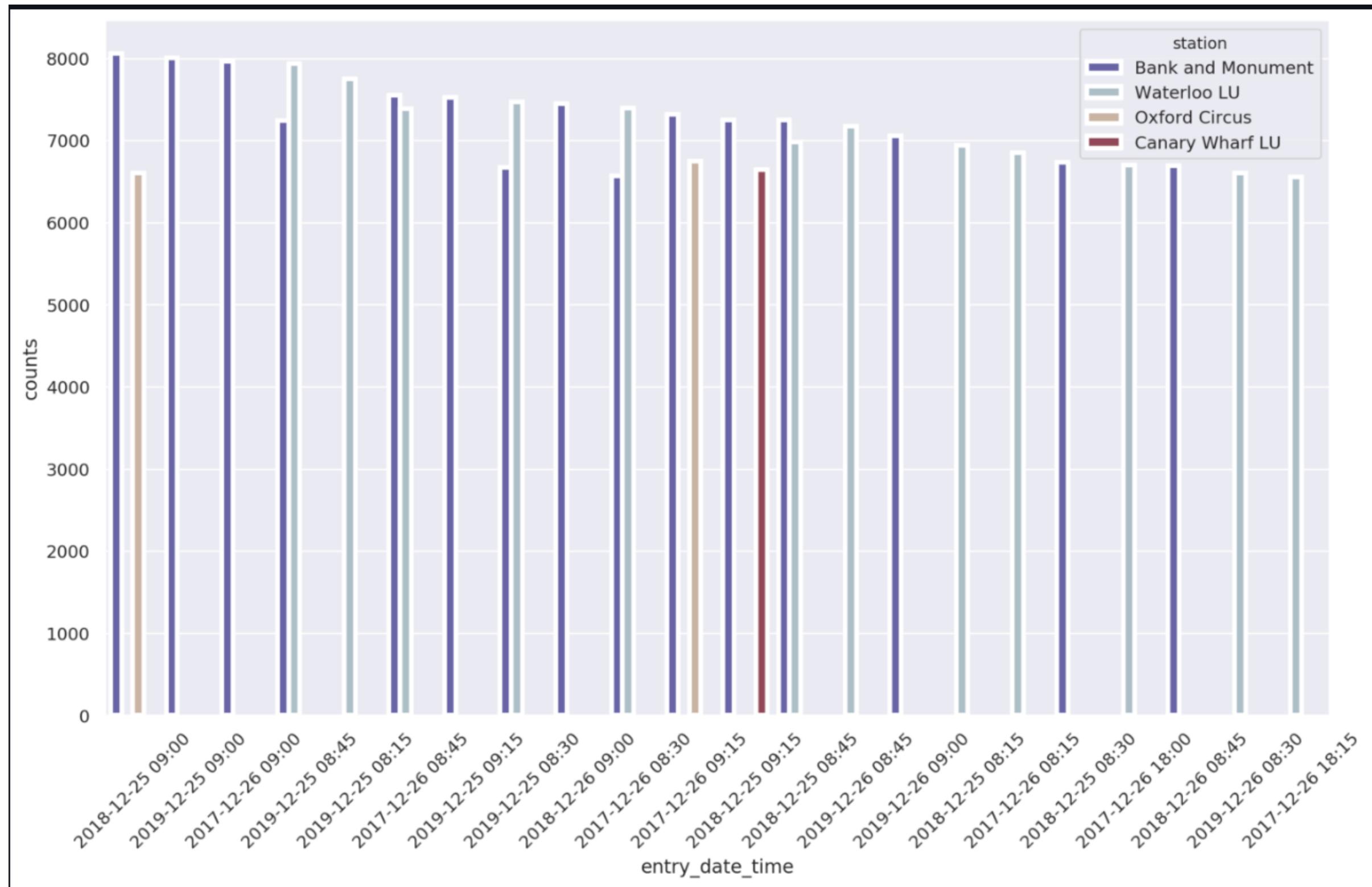


Using an interactive Bar Chart, we can visualize these stations and look at the Time associated with the high traffic per stations.



○ ○ ○ ○

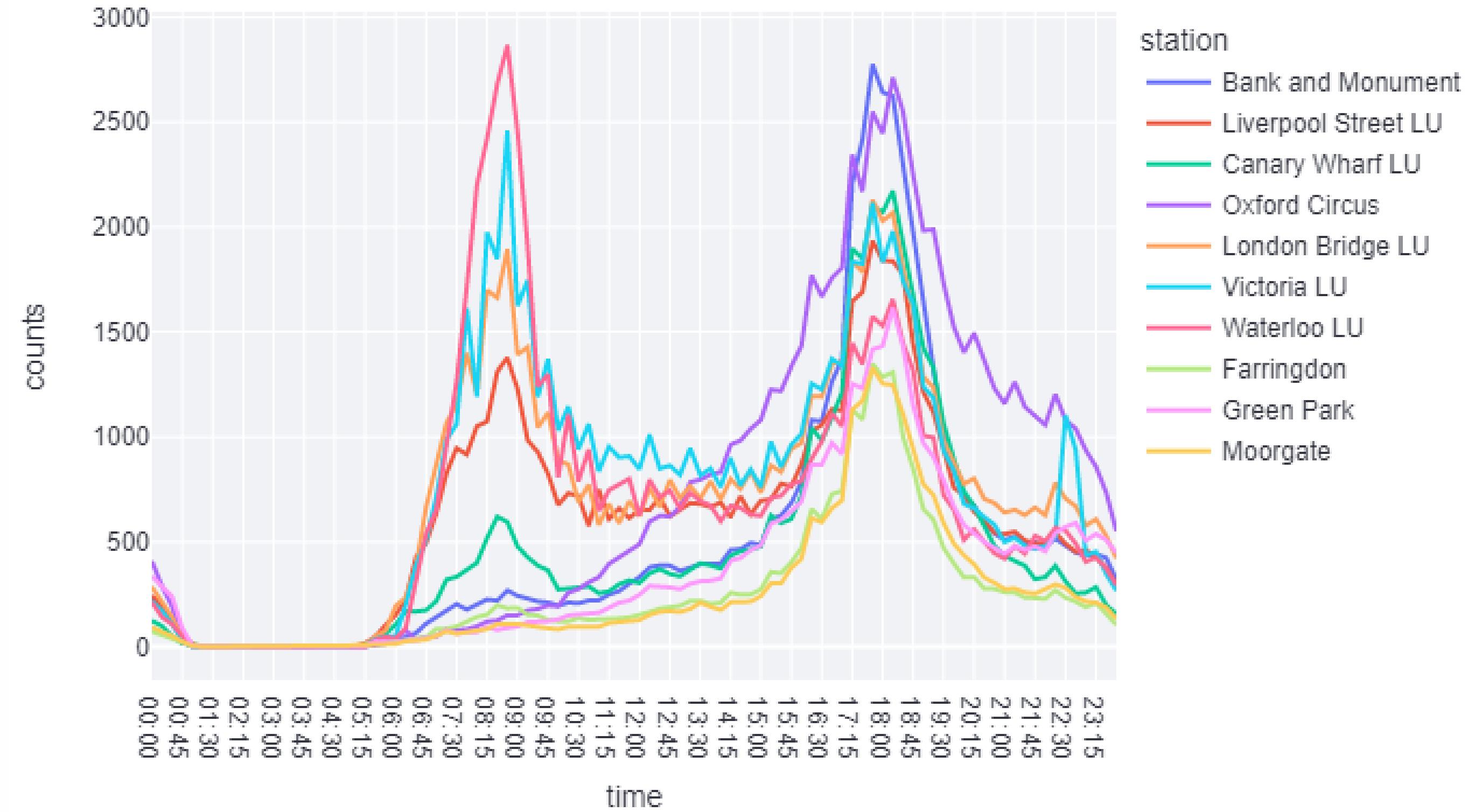
Time and Stations with highest Passenger traffic.



complete date and time on X-axis. and Passenger counts on Y-axis

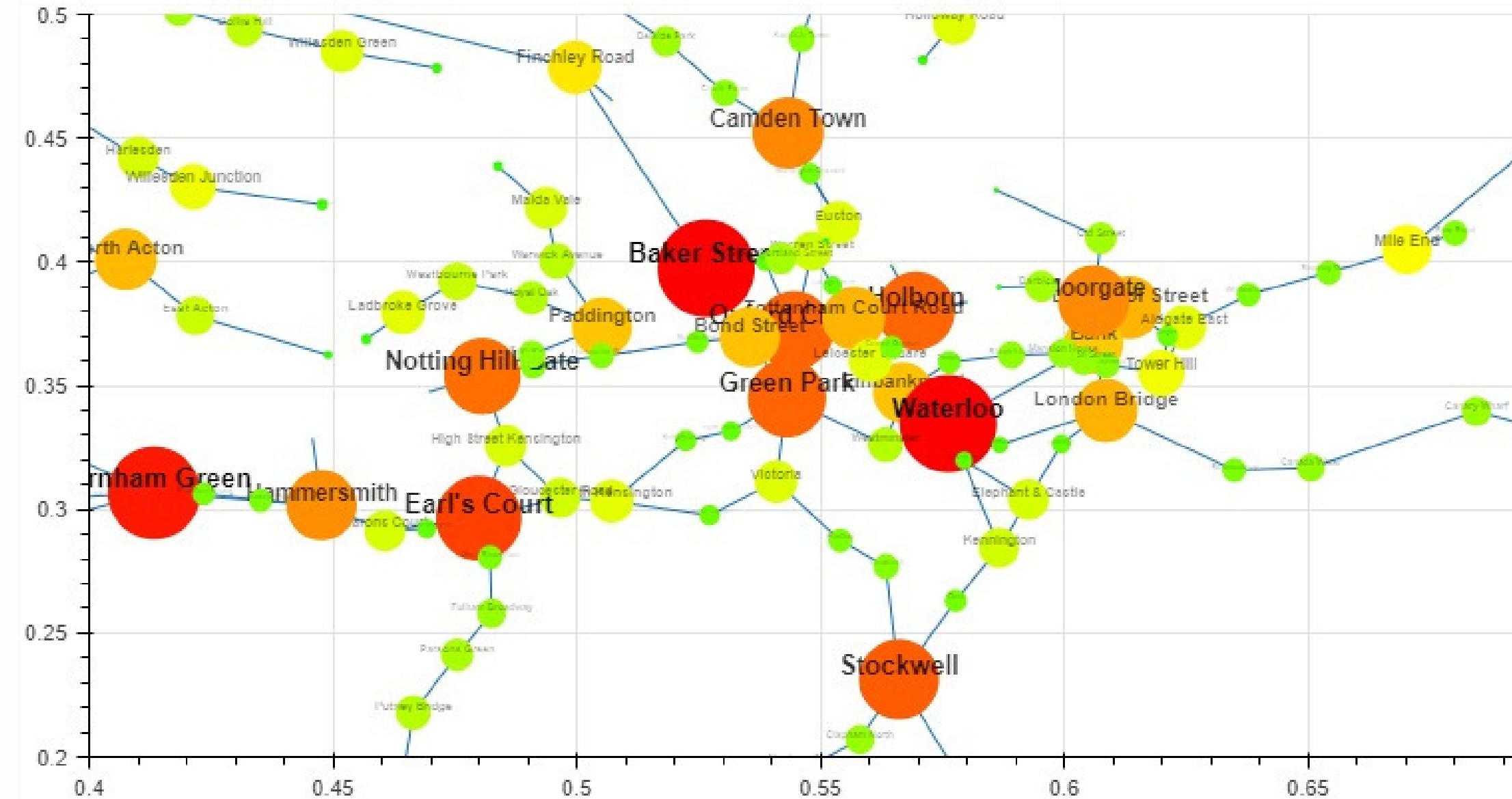


2021 Monday to Thursday inflow dataset





Graphical Network of the Tube



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

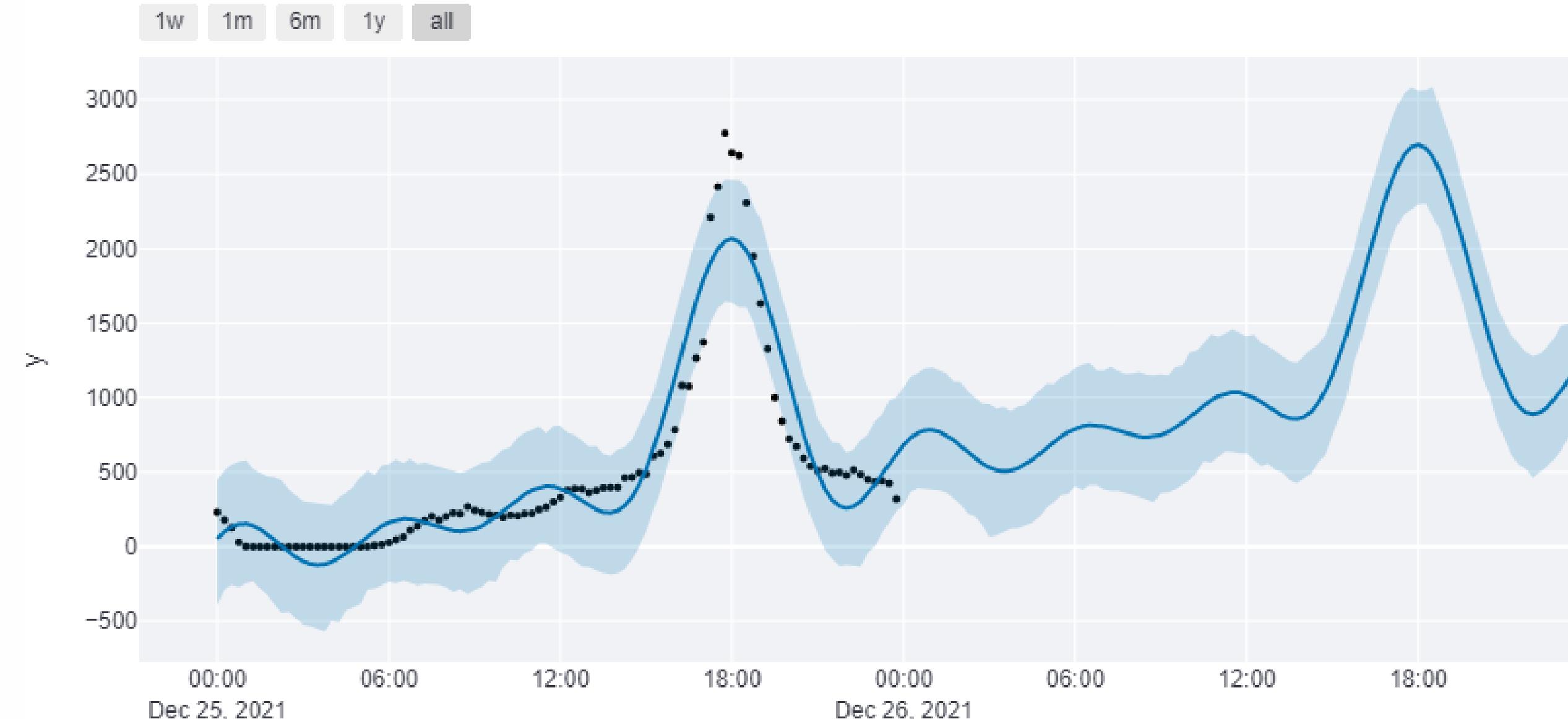


MODELLING

Let's take a look at what we did

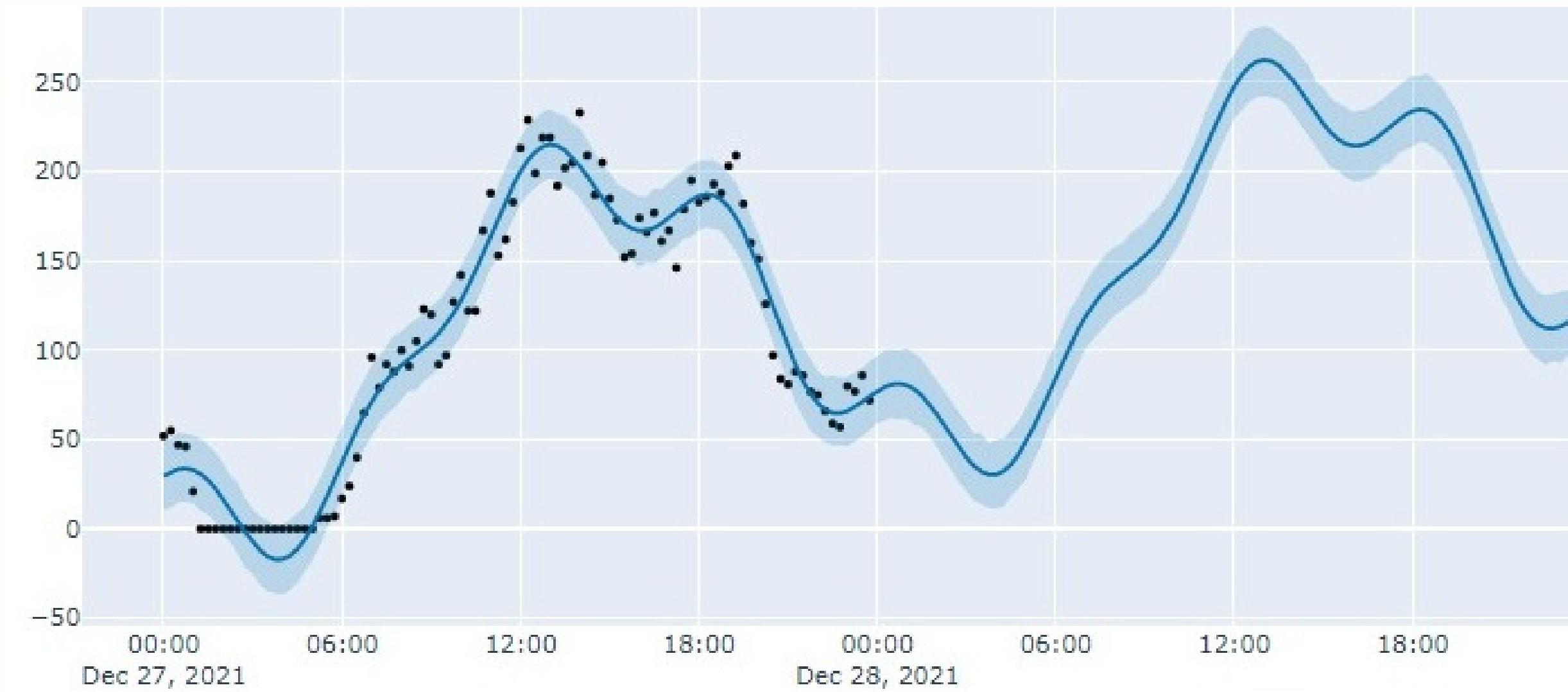
TIME SERIES MODEL

Passenger inflow forecast on MTT for Bank and Monument





FB Prophet time series model was developed by Facebook for fast time series forecasting



15-minute interval passenger outflow forecast on Saturdays for Moorgate station



PROJECT DEMO

Lets take a look at what we did.





•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

Tube Twin Architectural Diagram

AC Intern Team-6 2201

Data Source

TfL Open Data API serving as the main data source, together with other few external sources.



Data Source

Network Simulation



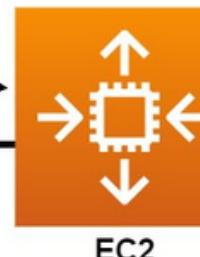
SUMO creates a simulation of traffic between a selected few of the stations in the tube network.

This serves as the storage location for all the codes, files, and data needed to run the whole application.



S3 Bucket

EC2 is the computing power between the S3 and the Streamlit app. It gets input from the app to run scripts from the S3, then it returns the output.



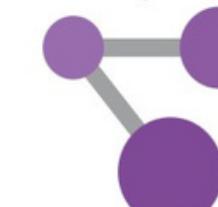
EC2



Modeling/Forecasting



Dashboard



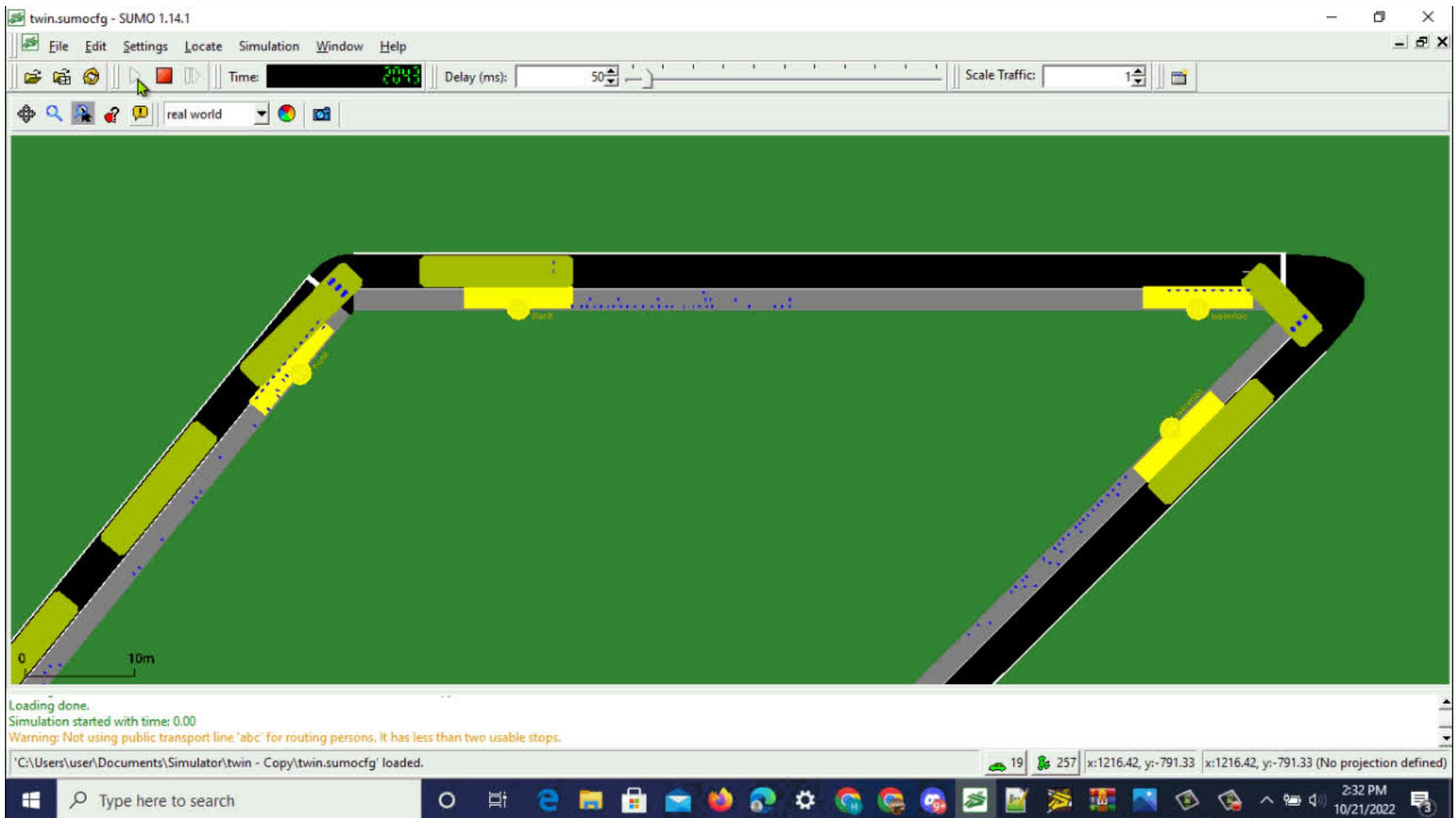
Graph Network

AWS platform providing the compute and storage resources required in the applications processes.

Streamlit web-app with different sections/pages for model/forecasting, dashboard, and a graph-network. They, altogether, provide insights on the tube data.



TRAIN SIMULATION





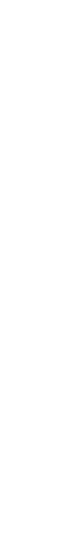
APP DEMO

Made with Streamlit



CONCLUSION

The Tube Twin analysis and passenger forecasting conducted in this project provides the solution to the traffic issues in the heavily-weighted Tube station lines. The deployed web application adequately covered the top 14 stations for passenger forecasting, while the general time series data analysis and the graphical network representation provided an overall analysis. The project outcome gives a reliable prototype system that, if deployed and interactively developed, can support TFL in achieving many of its Tube management objectives which includes traffic control.



RECOMMENDATION

- The Tube Twin project by Explore AI (2022) team 6 provides a good ground for further research and improvement of passenger flow and traffic analyses on the London Tube. We recommend expansion in the scope of passenger forecasting (i.e. more stations) taking into account certain social activities and environmental factors (e.g. social events and weather).
- Lastly, The next research on the project should consider a live streaming data collection of the tube for a minimum period of one year as this will allow for collection of comprehensive real-time data for higher forecasting accuracy.

PROJECT MILESTONE





THANK YOU!

IT'S Q & A TIME!

