**PHASE 3 – PUBLIC TRANSPORT EFFICIENCY ANALYSIS**

**PROJECT DEFINITION:**

The project involves analysing public transportation data to assess service efficiency, on time performance, and passenger feedback. The objective is to provide insights that support transportation improvement initiatives and enhance the overall public transportation experience. This project includes defining analysis objectives, collecting transportation data, designing relevant visualizations in IBM Cognos, and using code for data analysis.

**OBJECTIVE:**

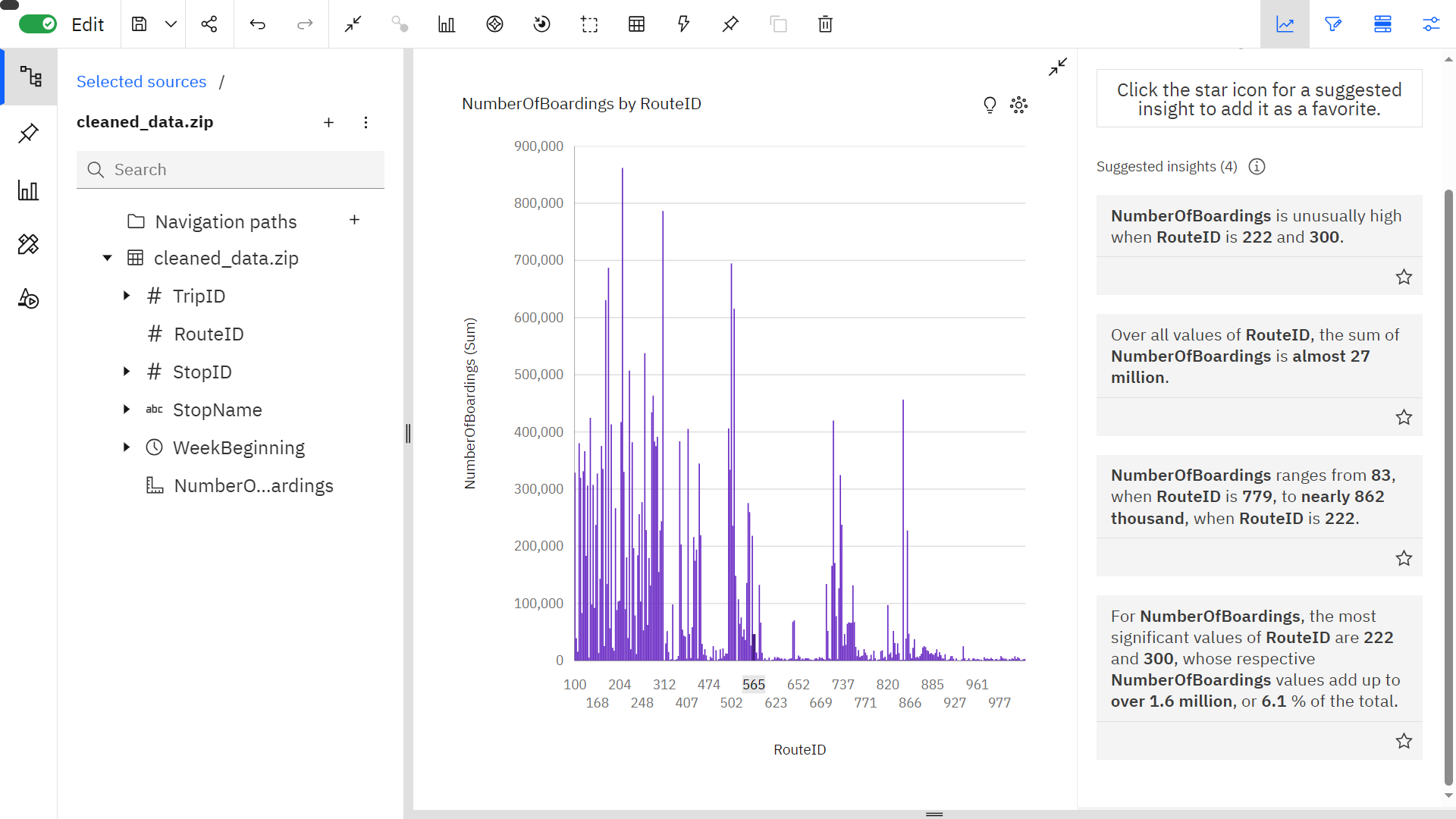
To pre-process and clean the dataset to make it more meaningful for analysis, and perform the same using the visualisation tool IBM Cognos.

**SOURCE CODE:**

<https://github.com/zzidd/PTEA/blob/main/DAC_Phase3.ipynb>

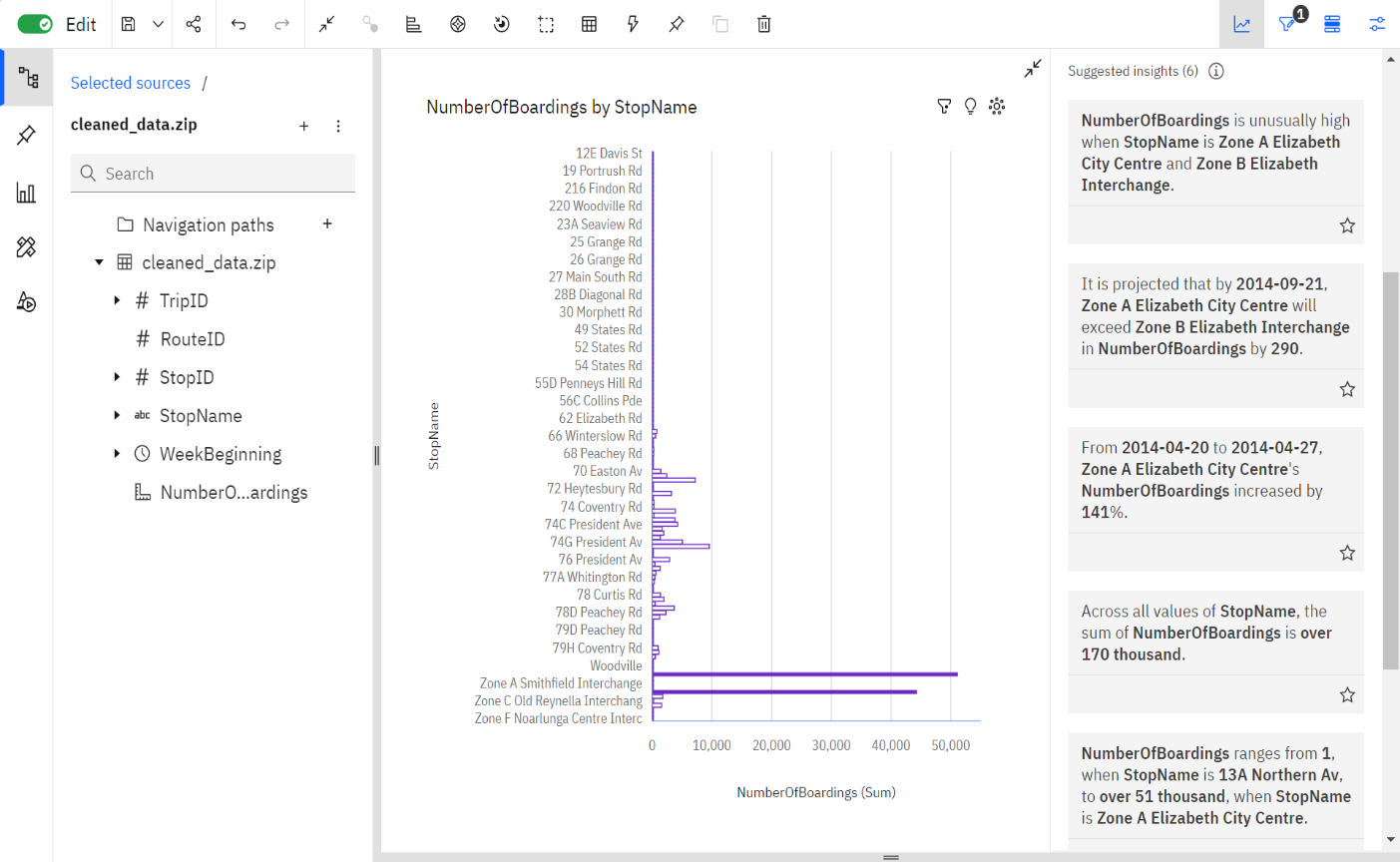
The above GitHub link consists of the source code which was successfully run and the resultant csv file was obtained.

The data was then imported to IBM Cognos for visualisation and the following inferences were noted:



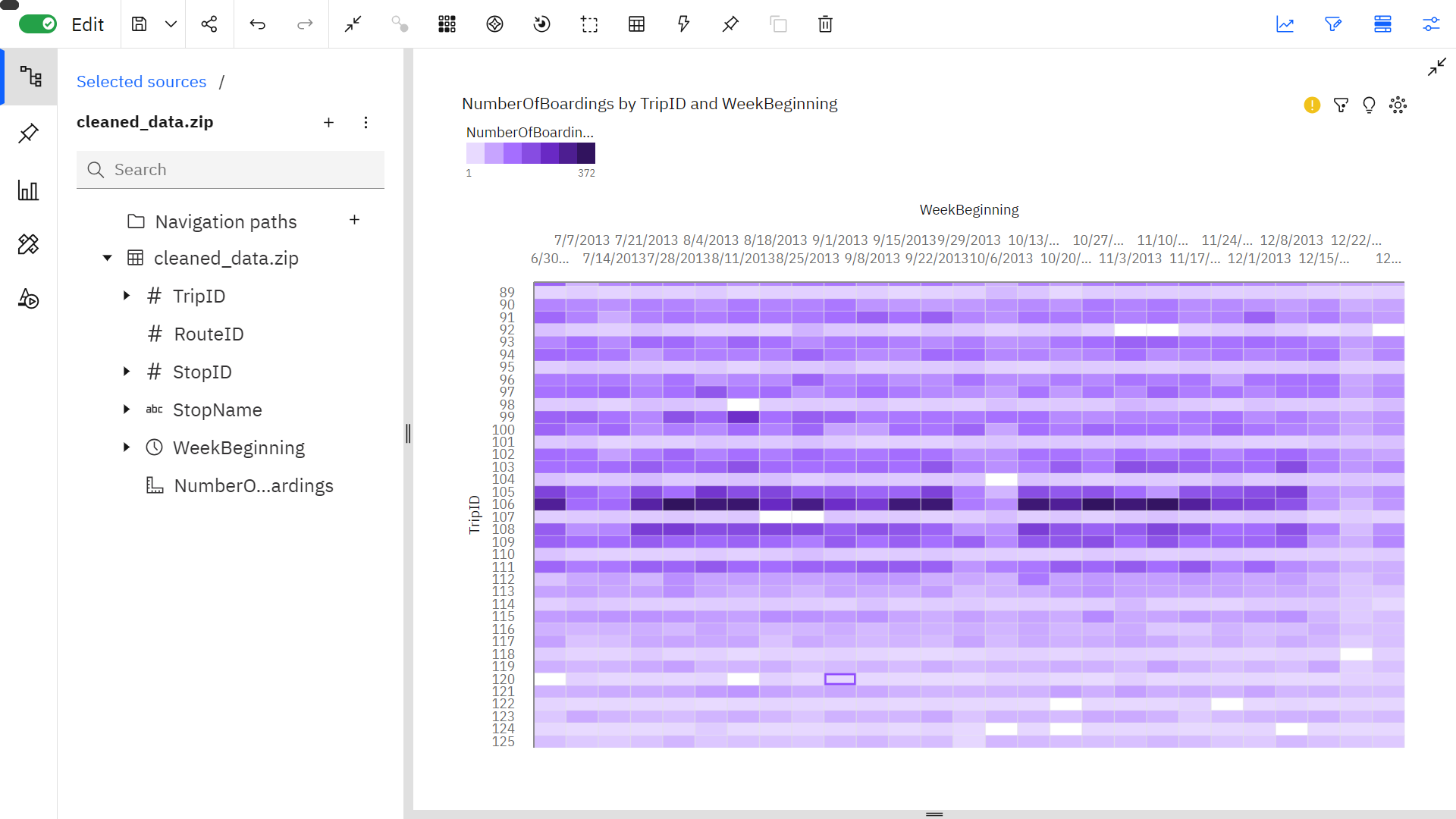
Insights:

* The Number of Boardings is unusually high when RouteID is 222 and 300.
* The sum of Number of Boardings is almost 27 million.
* The RouteID’s 222 and 300 are the most significant values, whose respective Number of Boarding values make up to 1.6 million, i.e., 6.1 % of the total.
* Number of Boardings range from 83, when RouteID is 779, to nearly 862K, when RouteID is 222.



Insights:

* Number of Boardings are unusually high when StopName is “Zone A Elizabeth City Centre” and “Zone B Elizabeth Interchange”.
* It is projected that by 21st September 2014, “Zone A Elizabeth City Centre” will exceed “Zone B Elizabeth Interchange” in Number of Boardings by 290.
* From 20th April to 27th April 2014, the Number of Boardings of “Zone A Elizabeth City Centre” increased by 141%.
* Number of Boardings ranges from 1, when StopName is “13A Northern Av., to over 51K, when StopName is “Zone A Elizabeth City Centre”.
* The sum of Number of Boardings is over 170K.



Insights:

* TripID 106 has the greatest Number of Boardings out of all ID’s, topping in various dates evenly.
* TripID 120 is the least popular TripID in terms of Number of Boardings.

**DATABASE LINK**

[Public Bus Transport Dataset (kaggle.com)](https://www.kaggle.com/datasets/rednivrug/unisys?select=20140711.CSV)

**CONCLUSION**

Therefore, the given data set was cleaned, pre-processed, and was analysed and visualized using IBM Cognos.

**TEAM MEMBERS**

Siddesh S - 2021504545

Vishal K - 2021504558

Mohamed Samasdeen - 2021504301

Prasannah SKS - 2021504538