அரசினர் பொறியியல் கல்லூரி, ஈரோடு
Government College of Engineering, Erode
(Approved by AICTE, New Delhi and Affiliated to Anna University, Chennai)

GOVERNMENT COLLEGE OF ENGINEERING ERODE

B.E Electronics and Communication Engineering

PUBLIC TRANSPORTATION ANALYSIS

Name of the Students	**University Reg No:**
LUTHDRAN CS	731121106029

Under the mentor of

Dr.M.Poongothai

Department of Information Technology (IT)

Department of Electronics and Communication Engineering

Government College of Engineering

Erode, PO, near Vasavi

College,TamilNadu-638316, Affiliated

to Anna University, Chennai.

PUBLIC TRANSPORTATION ANALYSIS

INTRODUCTION:

Public transport data analytics is a dynamic field at the intersection of transportation and data science. It involves the collection, processing, and interpretation of vast amounts of data generated by public transportation systems. By harnessing this data, we can optimize routes, improve service reliability, reduce congestion, and enhance the overall commuter experience. In this conversation, we'll explore the various facets of public transport data analytics and its far-reaching impact on urban mobility.

PROJECT OBJECTIVE:

- This analysis seeks to address the complex and pressing issues that plague public transportation networks, which impact the quality of service riders numbers, environmental sustainability, and the economic viability of these systems.

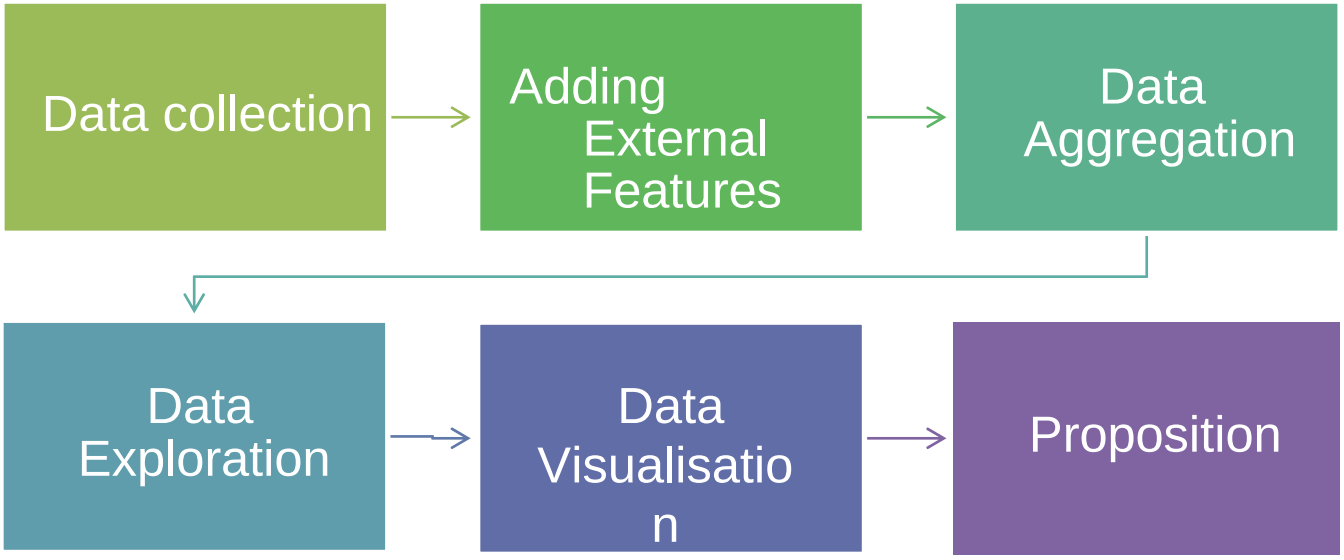
- This analysis have explored how people are travelling from different stops in Adelaide Metropolitan area and the rate at which passengers on each bus route are increasing Finally created a predictive model to find the load of passengers on public Bus transport system in future.

- To visualize on-time performance, passenger feedback, and service efficiency metrics.

- By analyzing data on ridership patterns, route performance, and vehicle health, the project aims to enhance the overall quality of public transportation services. This includes reducing delays, optimizing routes, and improving passenger satisfaction.

- Overall, the objective is to empower transit agencies with actionable insights derived from data. By using data analytics, IBM projects enable agencies to make informed decisions in real-time and plan for the future based on historical trends and predictive models.

BLOCK DIAGRAM:



PROJECT STEPS:

DATA COLLECTION:

- Data collection is the process of collecting and analyzing information on relevant variables in a predetermined, methodical way so that one can respond to specific research questions, test hypotheses, and assess results. Data collection can be either qualitative or quantitative.

- The following raw data is obtained from Southern Australian Government.It includes public transport boarding summary by Route, Trip, Stop and Week of Year. The data which is obtained is in the form of excel sheet.The objective is to provide insights that support transportation improvement initiatives and enhance the overall public transportation experience.

- In addition to this we have add some secondary data set source to get final data set for analysis

The data fields in the data set are:

- TripID Unique identity of trip
- RouteID Value representing public transport route
- StopID Unique identity of stop
- StopName Name of given stop
- WeekBeginning Date representing first day of any week
- NumberOfBoarding Count of all boarding's occurred at this stop for the named trip over the previous week

	TripID	RouteID	StopID	StopName	WeekBeginning	NumberOfBoardings
0	23631	100	14156	181 Cross Rd	2013-06-30 00:00:00	1
1	23631	100	14144	177 Cross Rd	2013-06-30 00:00:00	1

The data fields in the secondary data source are:

- **Latitude** Latitude of the Bus Stop
- **Longitude** Longitude of the Bus Stop
- **Type** Typeof Address
- **PostCode** Post code of the area in which bus stop is present.

	accuracy	formatted_address	google_place_id	input_string	latitude	longitude	number_of_results	postcode	status	type
0	ROOFTOP	181 Cross Rd, Westbourne Park SA 5041, Australia	ChIJKT7I9rbPsGoRVHMHkly-Oyk	181 Cross Rd	-34.966656	138.592148	1	5041	OK	street_address
1	ROOFTOP	177 Cross Rd, Westbourne Park SA 5041, Australia	ChIJ-VFZ87bPsGoRyfVgC5qbPpE	177 Cross Rd	-34.966607	138.592301	1	5041	OK	street_address

ADDING EXTERNAL FEATURES:

Some Important external data fields calculation

- **IsHoliday** Number of public holidays within that week
- **DistanceFromCentre** Distance measure from the city centre

For Calculating Distance between centre with other bus stops by using Longitude and Latitude we have used the Haversine formula.

DATA AGGREGATION:

- ❖ Data aggregation is the process of collecting data to present it in summary form. This information is then used to conduct statistical analysis and can also help company executives make more informed decisions about marketing strategies, price settings, and structuring operations, among other things.

- ❖ **Data aggregation** is the process where raw data is gathered and expressed in a summary form for statistical analysis.

- ❖ In this project we combine gelocation data and given data set to get desired output
  .
- ❖ In this project Python library named Pandas is used to clean data.The Python Pandas library is a powerful data manipulation and analysis library that is widely used for data cleaning and preparation. It offers functions and methods for tasks like handling missing values, data transformation, and filtering

Aggregate the Data According to Weeks and Stop names

- **NumberOfBoardings_sum** Number of Boardings within particular week for each Bus stop
- **NumberOfBoardings_count** Number of times data is recorded within week
- **NumberOfBoardings_max** Maximum number of boarding done at single time within week

DATA EXPLORATION:

- ❖ Data exploration refers to the initial step in data analysis in which data analysts use data visualization and statistical techniques to describe dataset characterizations, such as size, quantity, and accuracy, in order to better understand the nature of the data.
- ❖ In this data we have 1 year of data from date 2013-06-30 till 2014-07-06 in a weekly interval based and having total of 4165 Stops in South Australian Metropolitan Area.

DATA VISUALISATION:

- ❖ Data visualization is the graphical representation of data to help people understand the information contained in the data more easily. Visualizations make patterns, trends, and

insights in the data more apparent. Here the analysed data can be visualised by using cognos ibm and matplot python library.The data can be visualised in the form of graphs,tables,charts etc.

- IBM Cognos is a business intelligence and performance management software suite developed by IBM. It is designed to help organizations access, analyze, and report on their data to make more informed business decisions. Cognos provides a wide range of tools and features for data integration, reporting, dashboards, and analytics.
- Dashboarding: Users can create interactive and visually appealing dashboards to monitor key performance indicators (KPIs) and track business metrics. Dashboards can include charts, graphs, and other visual elements.
- Data Visualization: Cognos offers data visualization capabilities, allowing users to create compelling charts, graphs, and other visual representations of data to aid in data analysis.

PROPOSITION:

- In the context of data analysis, a "proposition" typically refers to a statement or hypothesis that is made based on the data being analyzed. Propositions play a crucial role in the data analysis process as they guide the analysis and help in drawing meaningful conclusions from the data.

- **Predictive Modeling**: In machine learning and predictive analytics, propositions can take the form of predictive statements. For example, "Based on historical customer data, we propose that customers with certain characteristics are more likely to purchase our product."

PROGRAM:

```
%matplotlib inline

import numpy as np # linear algebra

import pandas as pd # data processing, CSV file I/O (e.g. pd.read_csv)

import matplotlib.pyplot as plt

import os


print(os.listdir(r'C:\project'))

# Any results you write to the current directory are saved as output.
```

```
## For Multiple Output in single cell
from IPython.core.interactiveshell import InteractiveShell
```

```
InteractiveShell.ast_node_interactivity = "all"


## For Multiple Output in single cell
import warnings
from IPython.core.interactiveshell import InteractiveShell
InteractiveShell.ast_node_interactivity = "all"
warnings.filterwarnings('ignore')


data = pd.read_csv(r'C:\project\20140711.CSV')
out_geo = pd.read_csv(r'C:\project\output_geo.csv')
data.shape
data.head(2)
out_geo.head(2)


from math import sin, cos, sqrt, atan2, radians
def calc_dist(lat1,lon1):
## approximate radius of earth in km
R = 6373.0
dlon = radians(138.604801) - radians(lon1)
dlat = radians(-34.921247) - radians(lat1)
a = sin(dlat / 2)**2 + cos(radians(lat1)) * cos(radians(-34.921247)) * sin(dlon / 2)**2
c = 2 * atan2(sqrt(a), sqrt(1 - a))
return R * c
out_geo['dist_from_centre'] = out_geo[['latitude','longitude']].apply(lambda x: calc_dist(*x), axis=1)
##Fill the missing values with mode
out_geo['type'].fillna('street_address',inplace=True)
out_geo['type'] = out_geo['type'].apply(lambda x: str(x).split(',')[-1])
out_geo['type'].unique()
```

'''Holidays--

2013-09-01,Father's Day

2013-10-07,Labour day

2013-12-25,Christmas day

2013-12-26,Proclamation Day

2014-01-01,New Year

2014-01-27,Australia Day

2014-03-10,March Public Holiday

2014-04-18,Good Friday

2014-04-19,Easter Saturday

2014-04-21,Easter Monday

2014-04-25,Anzac Day

2014-06-09,Queen's Birthday'''

```
def holiday_label (row):
if row == datetime.date(2013, 9, 1) :
return '1'
if row == datetime.date(2013, 10, 6) :
return '1'
if row == datetime.date(2013, 12, 22) :
return '2'
if row == datetime.date(2013, 12, 29):
return '1'
if row == datetime.date(2014, 1, 26):
return '1'
if row == datetime.date(2014, 3, 9):
return '1'
if row == datetime.date(2014, 4, 13) :
```

```
return '2'

if row == datetime.date(2014, 4, 20):

return '2'

if row == datetime.date(2014, 6, 8):

return '1'

return '0'


data['WeekBeginning'] = pd.to_datetime(data['WeekBeginning']).dt.date

import datetime

data['holiday_label'] = data['WeekBeginning'].apply (lambda row: holiday_label(row))

data= pd.merge(data,out_geo,how='left',left_on = 'StopName',right_on = 'input_string')

col = ['TripID', 'RouteID', 'StopID', 'StopName',
'WeekBeginning','NumberOfBoardings','formatted_address',

'latitude', 'longitude','postcode','type','dist_from_centre',]


data = data[col]

##saving the final dataset

data.to_csv('Weekly_Boarding.csv',index=False)


## getting the addresses for geolocation api.

# Address data['StopName'].unique()

# sub = pd.DataFrame({'Address': Address})

# sub=sub.reindex(columns=["Address"])

# sub.to_csv('addr.csv')


# st_week_grp1 =
pd.DataFrame(data.groupby(['StopName','WeekBeginning','type']).agg({'NumberOfBoarding
s': ['sum', 'count']})).reset_index()

grouped = data.groupby(['StopName','WeekBeginning','type']).agg({'NumberOfBoardings':
['sum', 'count','max']})

grouped.columns = ["_".join(x) for x in grouped.columns.ravel()]
```

```
st_week_grp = pd.DataFrame(grouped).reset_index()
st_week_grp.shape
st_week_grp.head()
st_week_grp1 = pd.DataFrame(st_week_grp.groupby('StopName')['WeekBeginning'].count()).reset_index()
aa=list(st_week_grp1[st_week_grp1['WeekBeginning'] == 54]['StopName'])
bb = st_week_grp[st_week_grp['StopName'].isin(aa)]
## save the aggregate data
bb.to_csv('st_week_grp.csv', index=False)
data.nunique()
data.shape
data.columns
data.head(3)

data.isnull().sum()
data['WeekBeginning'].unique()
##can assign the each chart to one axes at a time

##VISUALISATION

fig,axrr=plt.subplots(3,2,figsize=(18,18))

data['NumberOfBoardings'].value_counts().sort_index().head(20).plot.bar(ax=axrr[0][0])
data['WeekBeginning'].value_counts().plot.area(ax=axrr[0][1])
ax=axrr[1][0]
ax.set_title("most Busiest Route")
data['RouteID'].value_counts().head(20).plot.bar(ax=axrr[1][0])
```

```
ax=axrr[1][1]

ax.set_title("least Busiest Route")

data['RouteID'].value_counts().tail(20).plot.bar(ax=axrr[1][1])

data['type'].value_counts().head(5).plot.bar(ax=axrr[2][0])

data['type'].value_counts().tail(10).plot.bar(ax=axrr[2][1])

data['postcode'].value_counts().head(20).plot.bar()

# data['dist_from_centre'].nunique()

bb_grp = data.groupby(['dist_from_centre']).agg({'NumberOfBoardings':
['sum']}).reset_index()

bb_grp.columns = bb_grp.columns.get_level_values(0)

bb_grp.head()

bb_grp.columns


from plotly.offline import download_plotlyjs, init_notebook_mode, plot, iplot

trace0 = go.Scatter(

x = bb_grp['dist_from_centre'],

y = bb_grp['NumberOfBoardings'],mode = 'lines+markers',name = 'X2 King William St')


data1 = [trace0]

layout = dict(title = 'Distance Vs Number of boarding',

xaxis = dict(title = 'Distance from centre'),

yaxis = dict(title = 'Number of Boardings'))

fig = dict(data=data1, layout=layout)

iplot(fig)


## for finding highest number of Boarding Bus stops

bb_grp = bb.groupby(['StopName']).agg({'NumberOfBoardings_sum':
['sum']}).reset_index()['NumberOfBoardings_sum'].sort_values('sum')

bb_grp[1000:1005]
```

```
bb.groupby(['StopName']).agg({'NumberOfBoardings_sum':
['sum']}).reset_index().iloc[[2325,1528,546,1043,1905]]

# bb_grp.iloc[[3054]]


source_1 = bb[bb['StopName'] == 'X2 King William St'].reset_index(drop = True)

source_2 = bb[bb['StopName'] == 'E1 Currie St'].reset_index(drop = True)

source_3 = bb[bb['StopName'] == 'I2 North Tce'].reset_index(drop = True)

source_4 = bb[bb['StopName'] == 'F2 Grenfell St'].reset_index(drop = True)


trace0 = go.Scatter(

x = source_1['WeekBeginning'],

y = source_1['NumberOfBoardings_sum'],mode = 'lines+markers',name = 'X2 King William
St')

trace1 = go.Scatter(

x = source_2['WeekBeginning'],

y = source_2['NumberOfBoardings_sum'],mode = 'lines+markers',name = 'E1 Currie St')

trace2 = go.Scatter(

x = source_3['WeekBeginning'],

y = source_3['NumberOfBoardings_sum'],mode = 'lines+markers',name = 'I2 North Tce')

trace3 = go.Scatter(

x = source_4['WeekBeginning'],

y = source_4['NumberOfBoardings_sum'],mode = 'lines+markers',name = 'F2 Grenfell St')

trace4 = go.Scatter(

x = source_5['WeekBeginning'],

y = source_5['NumberOfBoardings_sum'],mode = 'lines+markers',name = 'D1 King William
St')



data = [trace0,trace1,trace2,trace3,trace4]

layout = dict(title = 'Weekly Boarding Total',

xaxis = dict(title = 'Week Number'),
```

```
yaxis = dict(title = 'Number of Boardings'),
shapes = [{# Holidays Record: 2013-09-01
'type': 'line','x0': '2013-09-01','y0': 0,'x1': '2013-09-02','y1': 18000,'line': {
'color': 'rgb(55, 128, 191)','width': 1,'dash': 'dashdot'},},
{# 2013-10-07
'type': 'line','x0': '2013-10-07','y0': 0,'x1': '2013-10-07','y1': 18000,'line': {
'color': 'rgb(55, 128, 191)','width': 1,'dash': 'dashdot'},},
{# 2013-12-25
'type': 'line','x0': '2013-12-25','y0': 0,'x1': '2013-12-26','y1': 18000,'line': {
'color': 'rgb(55, 128, 191)','width': 3,'dash': 'dashdot'},},
{# 2014-01-27
'type': 'line','x0': '2014-01-27','y0': 0,'x1': '2014-01-28','y1': 18000,'line': {
'color': 'rgb(55, 128, 191)','width': 1,'dash': 'dashdot'},},
{# 2014-03-10
'type': 'line','x0': '2014-03-10','y0': 0,'x1': '2014-03-11','y1': 18000,'line': {
'color': 'rgb(55, 128, 191)','width': 1,'dash': 'dashdot'},},
{# 2014-04-18
'type': 'line','x0': '2014-04-18','y0': 0,'x1': '2014-04-19','y1': 18000,'line': {
'color': 'rgb(55, 128, 191)','width': 3,'dash': 'dashdot'},},
{# 2014-06-09
'type': 'line','x0': '2014-06-09','y0': 0,'x1': '2014-06-10','y1': 18000,'line': {
'color': 'rgb(55, 128, 191)','width': 1,'dash': 'dashdot'},},])
fig = dict(data=data, layout=layout)
iplot(fig)

source_6 = bb[bb['StopName'] == '57A Hancock Rd'].reset_index(drop = True)
source_7 = bb[bb['StopName'] == '37 Muriel Dr'].reset_index(drop = True)
source_8 = bb[bb['StopName'] == '18B Springbank Rd'].reset_index(drop = True)
source_9 = bb[bb['StopName'] == '27E Sir Ross Smith Av'].reset_index(drop = True)
```

```
source_10 = bb[bb['StopName'] == '46A Baldock Rd'].reset_index(drop = True)


trace0 = go.Scatter(
x = source_6['WeekBeginning'],
y = source_6['NumberOfBoardings_sum'],mode = 'lines+markers',name = '57A Hancock Rd')
trace1 = go.Scatter(
x = source_7['WeekBeginning'],
y = source_7['NumberOfBoardings_sum'],mode = 'lines+markers',name = '37 Muriel Dr')
trace2 = go.Scatter(
x = source_8['WeekBeginning'],
y = source_8['NumberOfBoardings_sum'],mode = 'lines+markers',name = '18B Springbank Rd')
trace3 = go.Scatter(
x = source_9['WeekBeginning'],
y = source_9['NumberOfBoardings_sum'],mode = 'lines+markers',name = '27E Sir Ross Smith Av')
trace4 = go.Scatter(
x = source_10['WeekBeginning'],
y = source_10['NumberOfBoardings_sum'],mode = 'lines+markers',name = '46A Baldock Rd')


data = [trace0,trace1,trace2,trace3,trace4]
layout = dict(title = 'Weekly Boarding Total',
xaxis = dict(title = 'Week Number'),
yaxis = dict(title = 'Number of Boardings'),
shapes = [{# Holidays Record: 2013-09-01
'type': 'line','x0': '2013-09-01','y0': 0,'x1': '2013-09-02','y1': 80,'line': {
'color': 'rgb(55, 128, 191)','width': 1,'dash': 'dashdot'},},
{# 2013-10-07
'type': 'line','x0': '2013-10-07','y0': 0,'x1': '2013-10-07','y1': 80,'line': {
```

```
'color': 'rgb(55, 128, 191)','width': 1,'dash': 'dashdot'},},
{# 2013-12-25
'type': 'line','x0': '2013-12-25','y0': 0,'x1': '2013-12-26','y1': 80,'line': {
'color': 'rgb(55, 128, 191)','width': 3,'dash': 'dashdot'},},
{# 2014-01-27
'type': 'line','x0': '2014-01-27','y0': 0,'x1': '2014-01-28','y1': 80,'line': {
'color': 'rgb(55, 128, 191)','width': 1,'dash': 'dashdot'},},
{# 2014-03-10
'type': 'line','x0': '2014-03-10','y0': 0,'x1': '2014-03-11','y1': 80,'line': {
'color': 'rgb(55, 128, 191)','width': 1,'dash': 'dashdot'},},
{# 2014-04-18
'type': 'line','x0': '2014-04-18','y0': 0,'x1': '2014-04-19','y1': 80,'line': {
'color': 'rgb(55, 128, 191)','width': 3,'dash': 'dashdot'},},
{# 2014-06-09
'type': 'line','x0': '2014-06-09','y0': 0,'x1': '2014-06-10','y1': 80,'line': {
'color': 'rgb(55, 128, 191)','width': 1,'dash': 'dashdot'},},])
fig = dict(data=data, layout=layout)
iplot(fig)


d=[]
for i in bb['StopName'].unique():
d.append({'StopName': i,'Boarding_sum':np.sum(bb[bb['StopName'] ==
i]['NumberOfBoardings_sum'].pct_change())/54,
'Boarding_count':np.sum(bb[bb['StopName'] ==
i]['NumberOfBoardings_count'].pct_change())/54,
'Boarding_max':np.sum(bb[bb['StopName'] ==
i]['NumberOfBoardings_max'].pct_change())/54})
pct_chng = pd.DataFrame(d)


#pct_chng.head()
pct_chng['Boarding_sum'].nlargest(5)
```

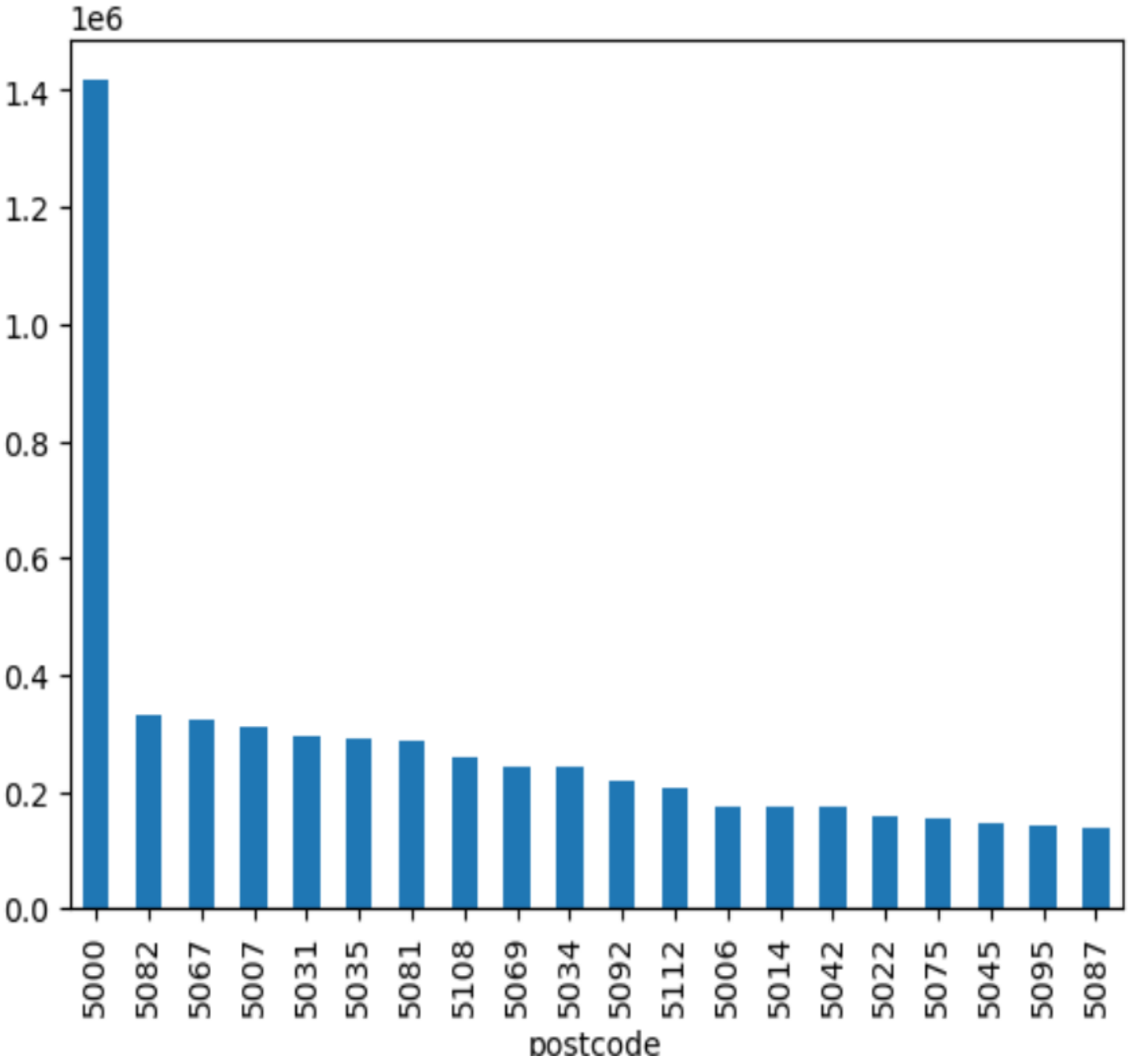
```
pct_chng['Boarding_sum'].nsmallest(5)

pct_chng[pct_chng['Boarding_sum']<0].shape

pct_chng.iloc[[3110,2134,214,1538,1290]]


bb1 = pd.merge(bb, out_geo, how='left', left_on = 'StopName', right_on = 'input_string')

bb1['holiday_label'] = bb1['WeekBeginning'].apply (lambda row: holiday_label(row))

##Final 11 features have been used for the forecastng.

cols=['StopName','WeekBeginning','type_x','NumberOfBoardings_sum','NumberOfBoardings_count','NumberOfBoardings_max','latitude','longitude','postcode','dist_from_centre','holiday_label']

bb1=bb1[cols]

bb1.shape

bb1.head()
```

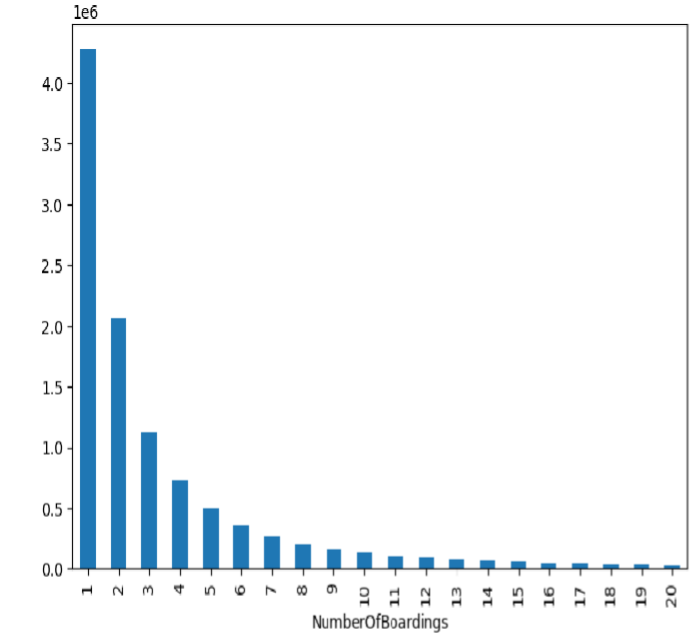
OUTPUT:

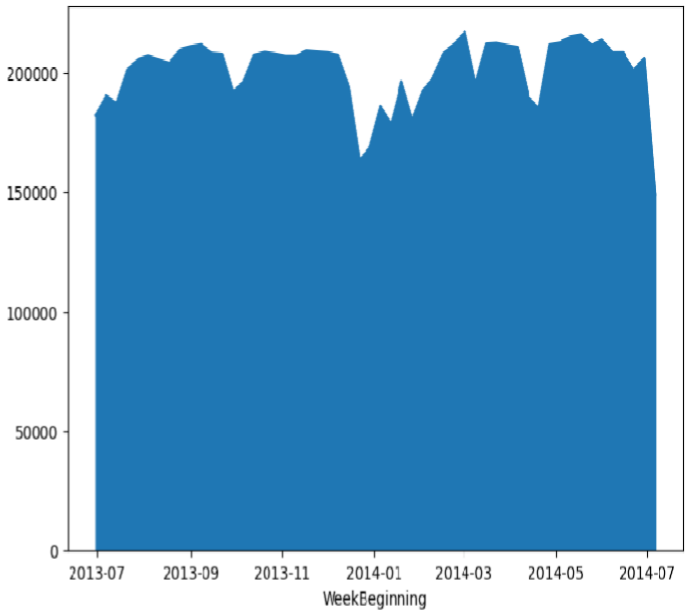
ANALYSIS USING PYTHON LIBRARIRES:

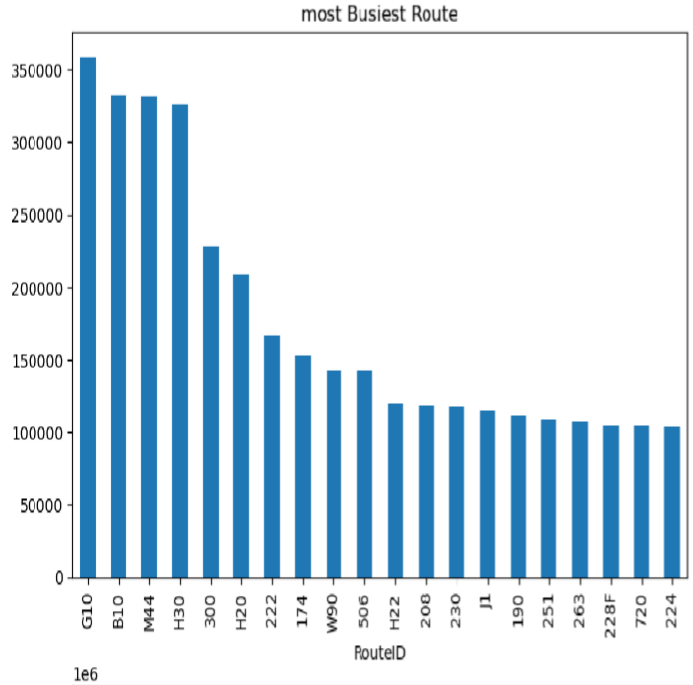
- More than 40 lakhs times only single person board from the bus stop.
- There are average of 1.8 lakhs people travel every week by bus in adelaide metropolitan area.
- G10,B10,M44,H30 are the most busiest routes in the city while FX8,FX3,FX10,FX1,FX2 are the least
- Most of the Bus stops are Street_Address Type while there are very few which are store or post office.
- As we move away from centre the number of Boarding decreases
- There are cluster of bus stops near to the main Adelaide city as oppose to outside.so that's why most of boardings are near to center
- X2 King William St and stop near to that are the most busiest stops in the city. which having number of boardings per week more than 10k.
- Vertical lines are the indicator of holidays which came within that week.

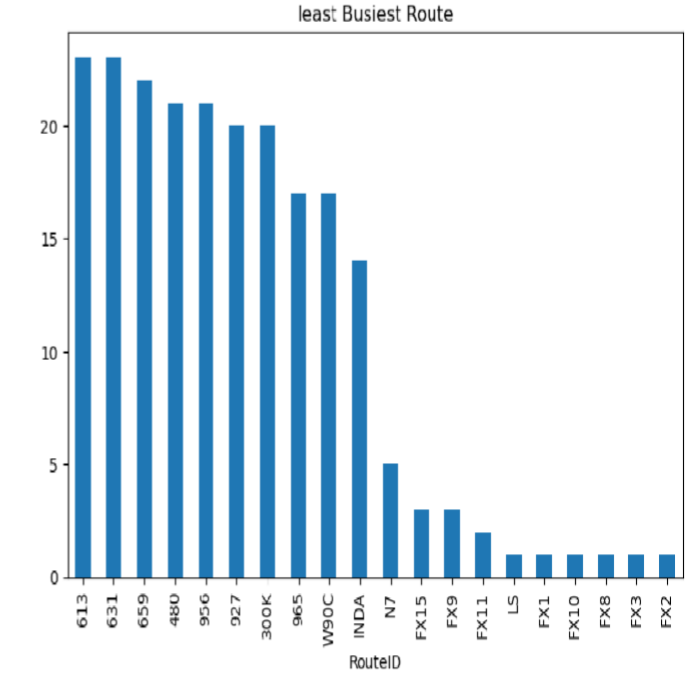
- Whenever there is any Public holiday that week period have less than average number of people travelled from bus.These 5 stops W Grote St,52 Taylors Rd,13 Tutt Av,37A Longwood Rd,32A Frederick Rd having the largest percent increase.
- There are 27 Bus stops where number of boardings have decreased.

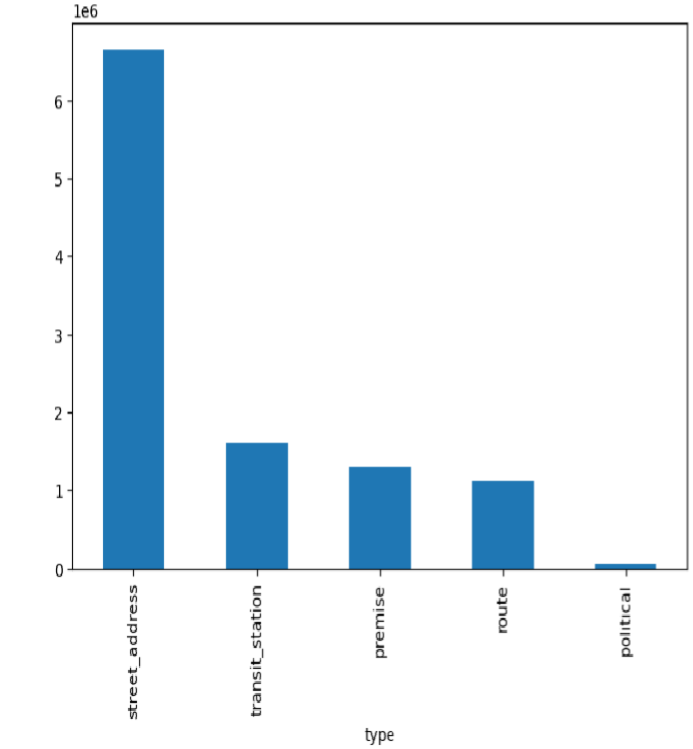


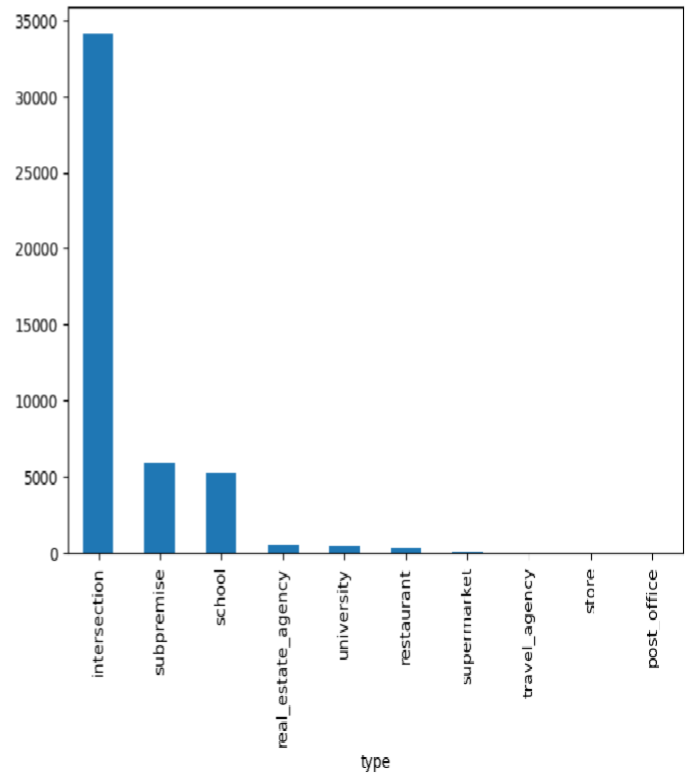

1e6
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
NumberOfBoardings

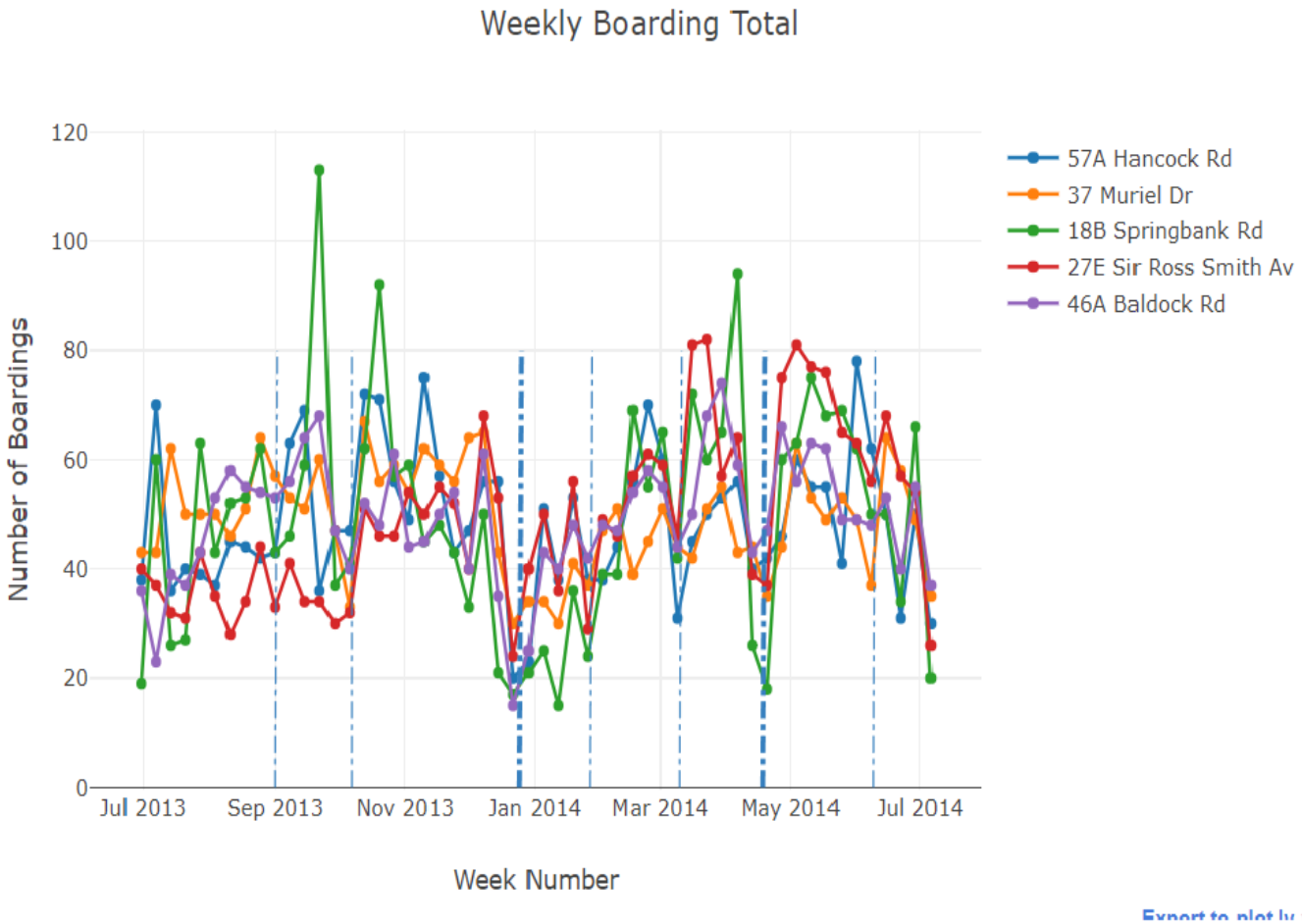

200000
150000
100000
50000
0
2013-07 2013-09 2013-11 2014-01 2014-03 2014-05 2014-07
WeekBeginning


most Busiest Route
350000
300000
250000
200000
150000
100000
50000
0
G10 B10 M44 H30 300 H20 222 174 W90 506 H22 208 230 J1 190 251 263 228F 720 224
RouteID

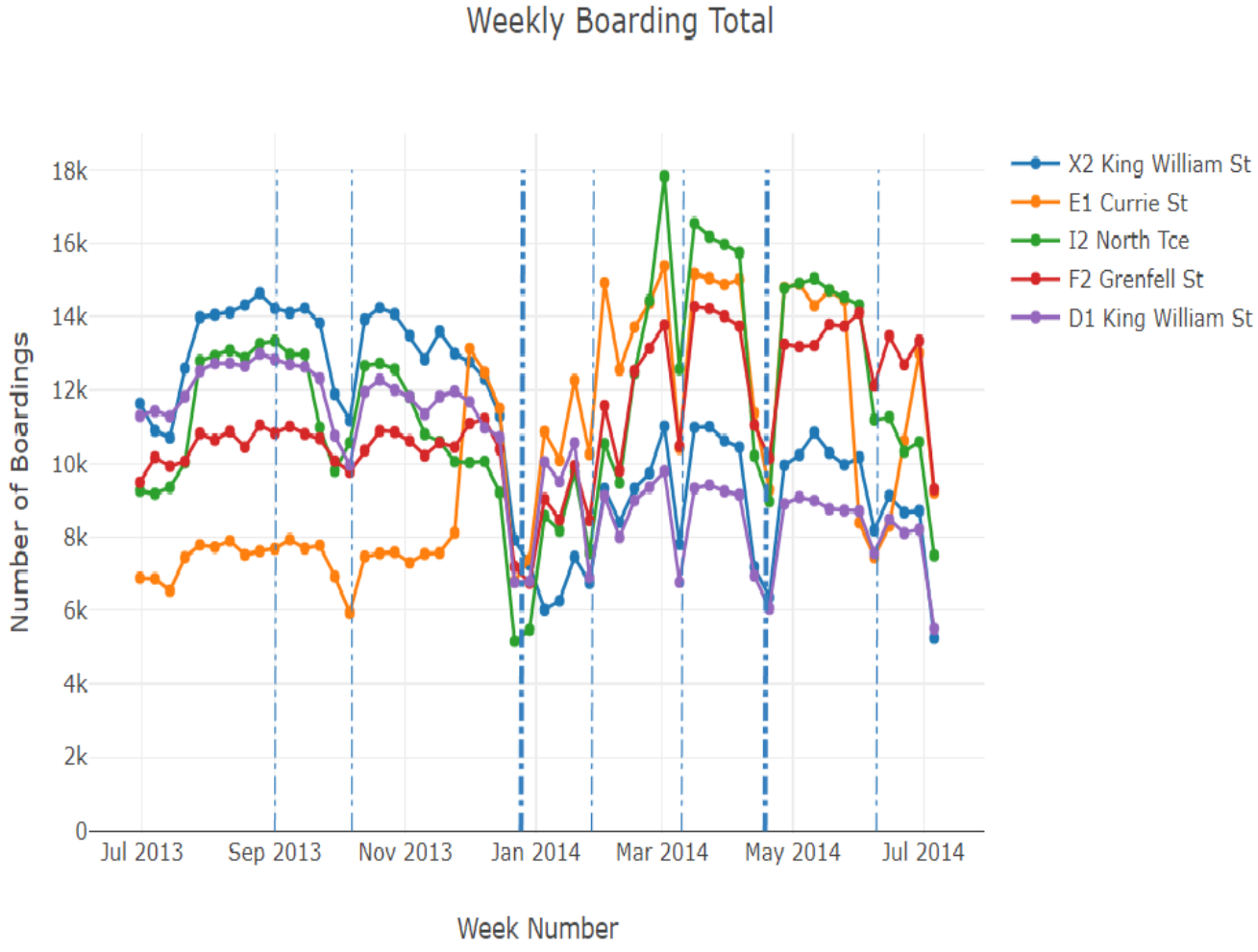

least Busiest Route
20
15
10
5
0
613 631 659 480 956 927 300K 965 W90C INDA N7 FX15 FX9 FX11 LS FX1 FX10 FX8 FX3 FX2
RouteID


1e6
6
5
4
3
2
1
0
street_address transit_station premise route political
type


35000
30000
25000
20000
15000
10000
5000
0
intersection subpremise school real_estate_agency university restaurant supermarket travel_agency store post office
type



Export to plot.ly »



DATA VISUALIZATION USING COGNOS IBM:

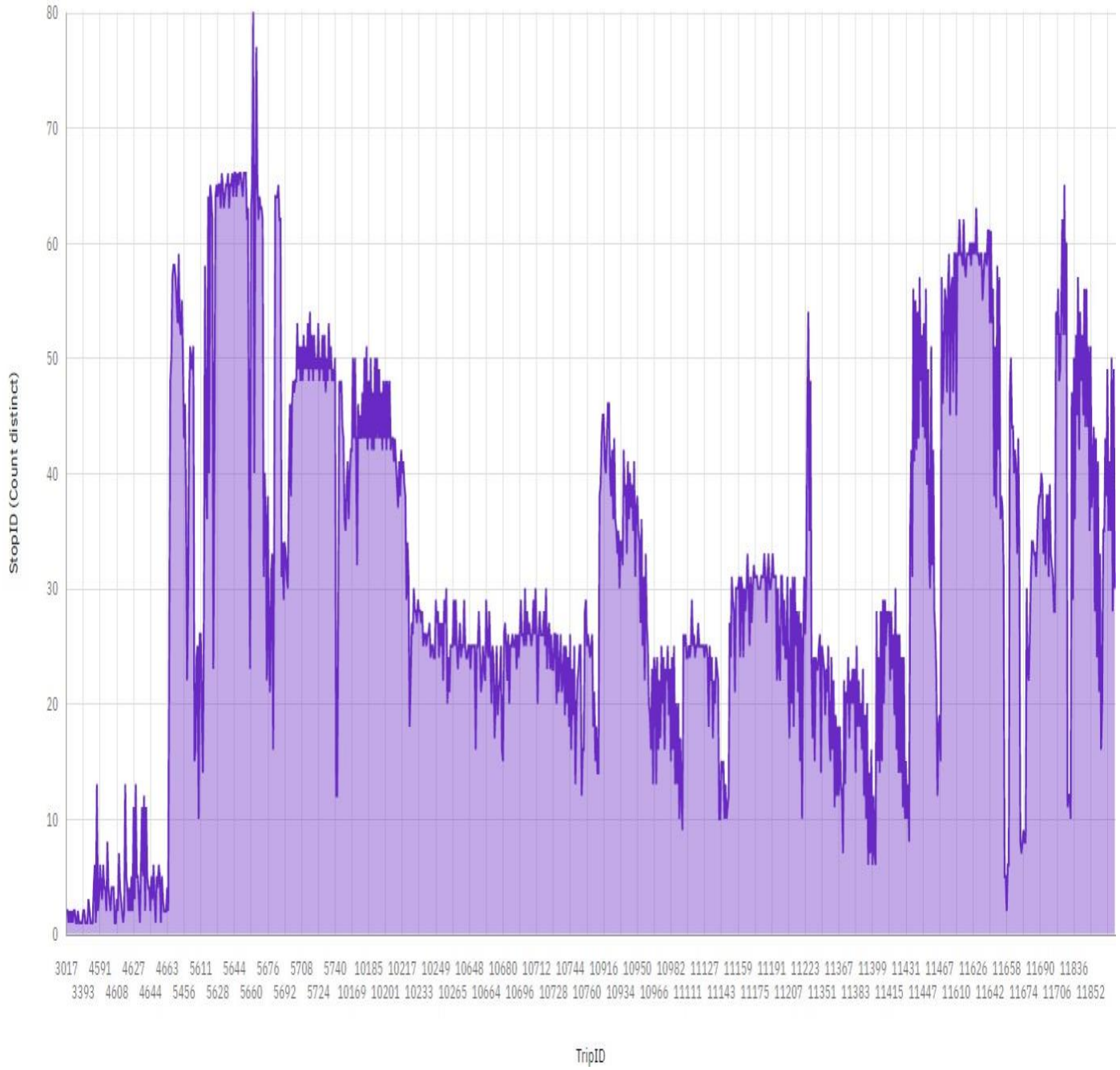
- Data visualization is the graphical representation of data to help people understand the information contained in the data more easily. Visualizations make patterns, trends, and insights in the data more apparent. Here the analysed data can be visualised by using cognos ibm and matplot python library.The data can be visualised in the form of graphs,tables,charts etc,

  - IBM Cognos Analytics is a specific product within the IBM Cognos suite, focusing on analytics and business intelligence. Cognos Analytics is designed to help organizations transform their data into meaningful insights and actionable information.

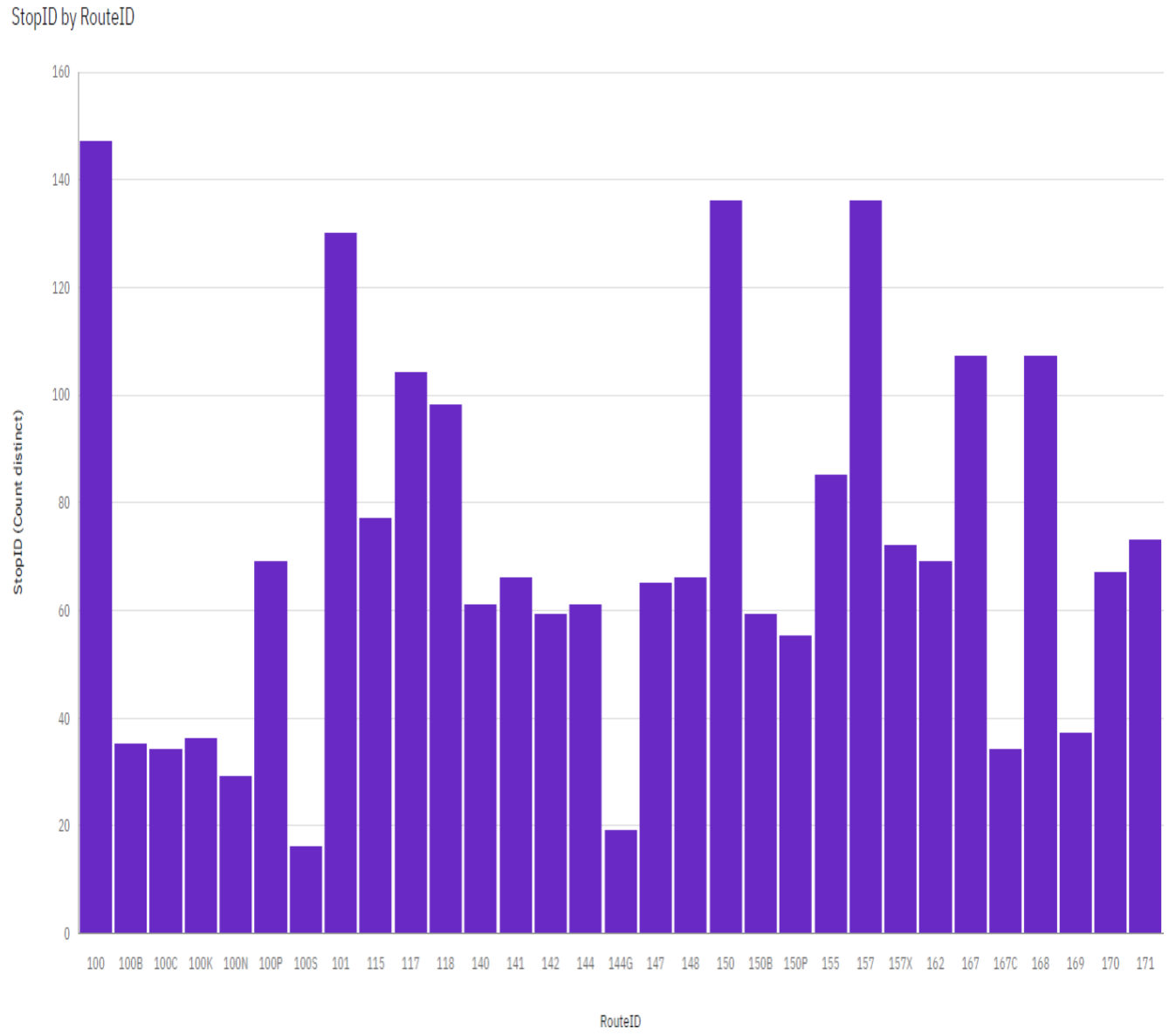
  - Cognos Analytics provides a wide range of data visualization options, including charts, graphs, and maps, to represent data in a visually appealing and informative way.

Tab 2

StopID by TripID



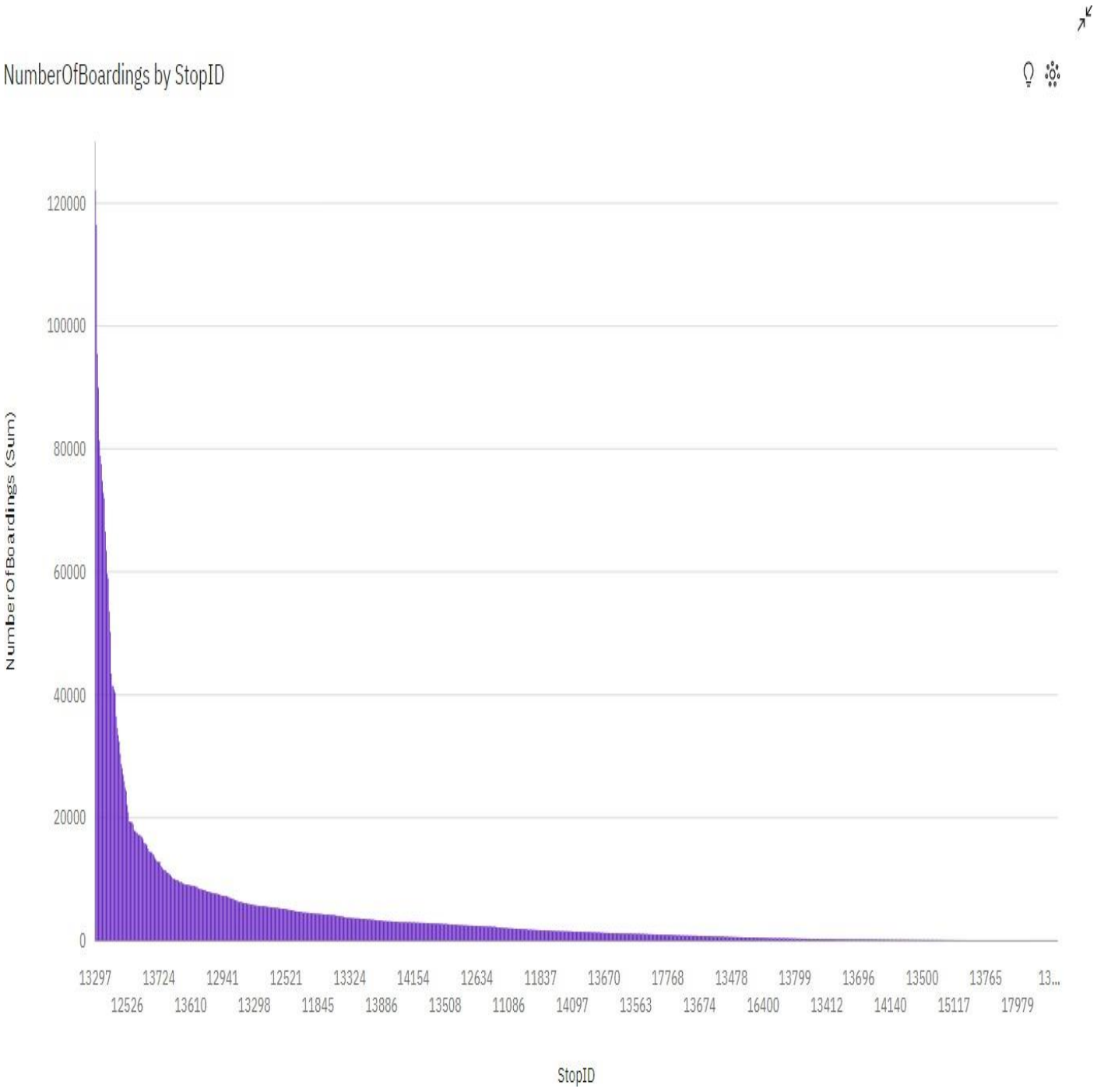
Tab 4



- RouteID 150 has the highest total NumberOfBoardings due to StopID 13297.
- StopID 13297 has the highest NumberOfBoardings at over 122 thousand, out of which RouteID 150 contributed the most at over 45 thousand

Tab 5

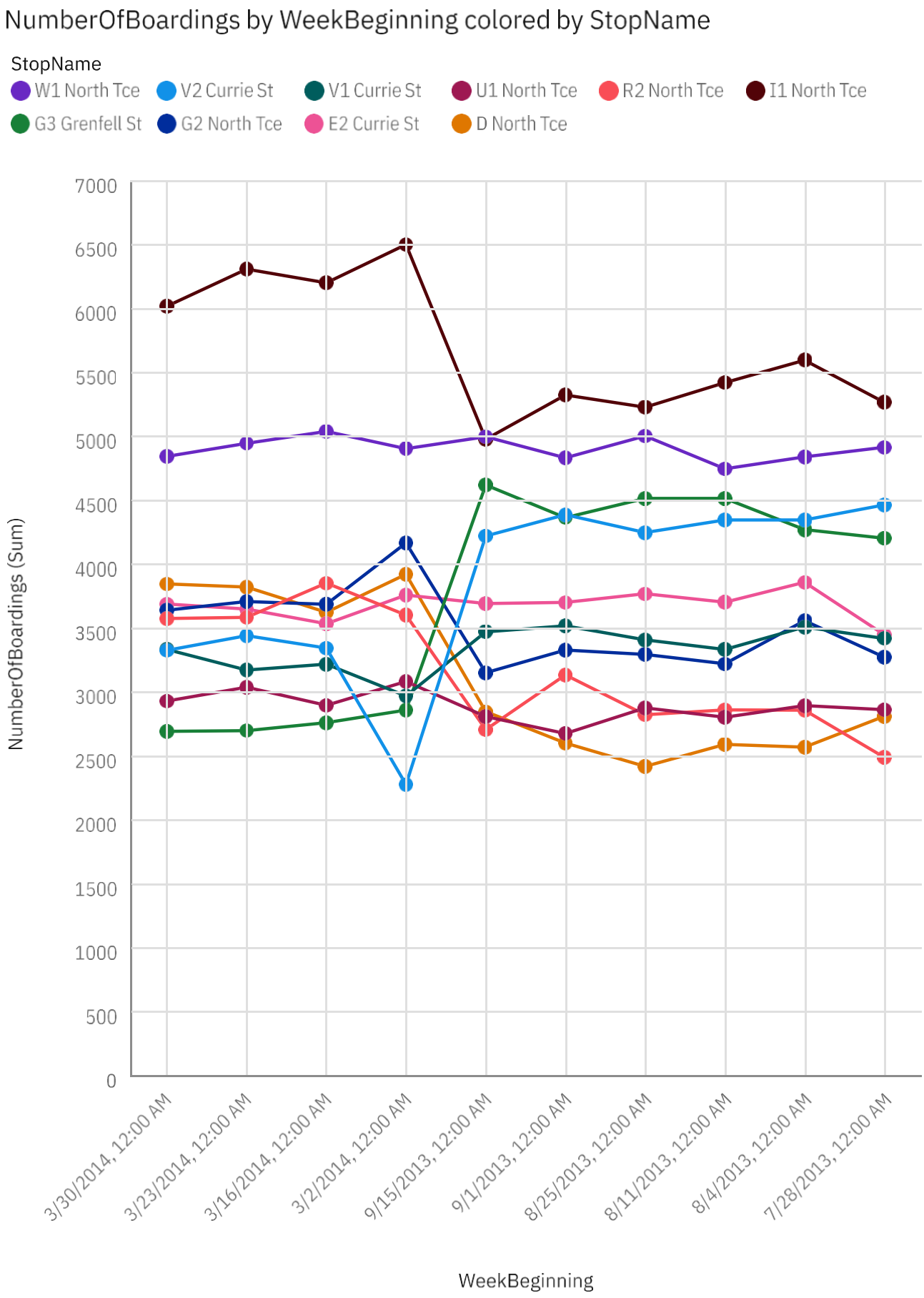
- NumberOfBoardings is unusually high when RouteID is 150.
- It is projected that by 21 Sep 2014, 150 will exceed 118 in NumberOfBoardings by over two thousand.
- From 22 Sep 2013 to 29 Sep 2013, 171's NumberOfBoardings dropped by 90%.
- Across all values of RouteID, the sum of NumberOfBoardings is over 4.3 million.
- NumberOfBoardings ranges from 260, when RouteID is 100S, to nearly 425 thousand, when RouteID is 150.



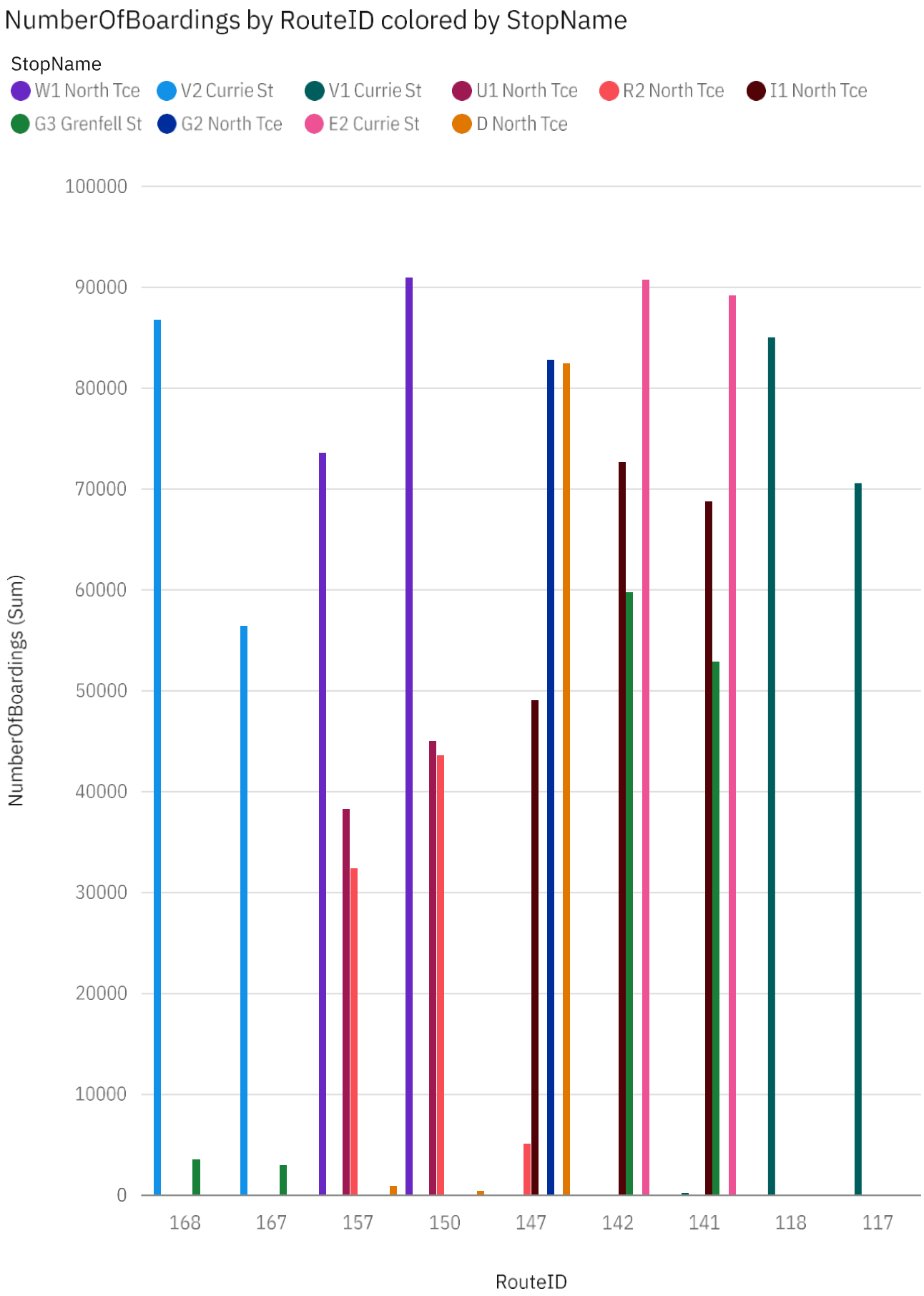
More than 1.2 lakhs boardings use STOP ID 13297

- NumberOfBoardings is unusually high when StopID is 13297 and 13278.
- Across all values of StopID, the sum of NumberOfBoardings is over 4.3 million.
- NumberOfBoardings ranges from 1, when StopID is 13277, to over 122 thousand, when StopID is 13297.
- NumberOfBoardings is unusually high when StopID is 13297 and 13278.
- Across all values of StopID, the sum of NumberOfBoardings is over 4.3 million.
- NumberOfBoardings ranges from 1, when StopID is 13277, to over 122 thousand, when StopID is 13297.
- For NumberOfBoardings, the most significant values of StopID are 13297 and 13278, whose respective NumberOfBoardings values add up to over 238 thousand, or 5.5 % of the total.

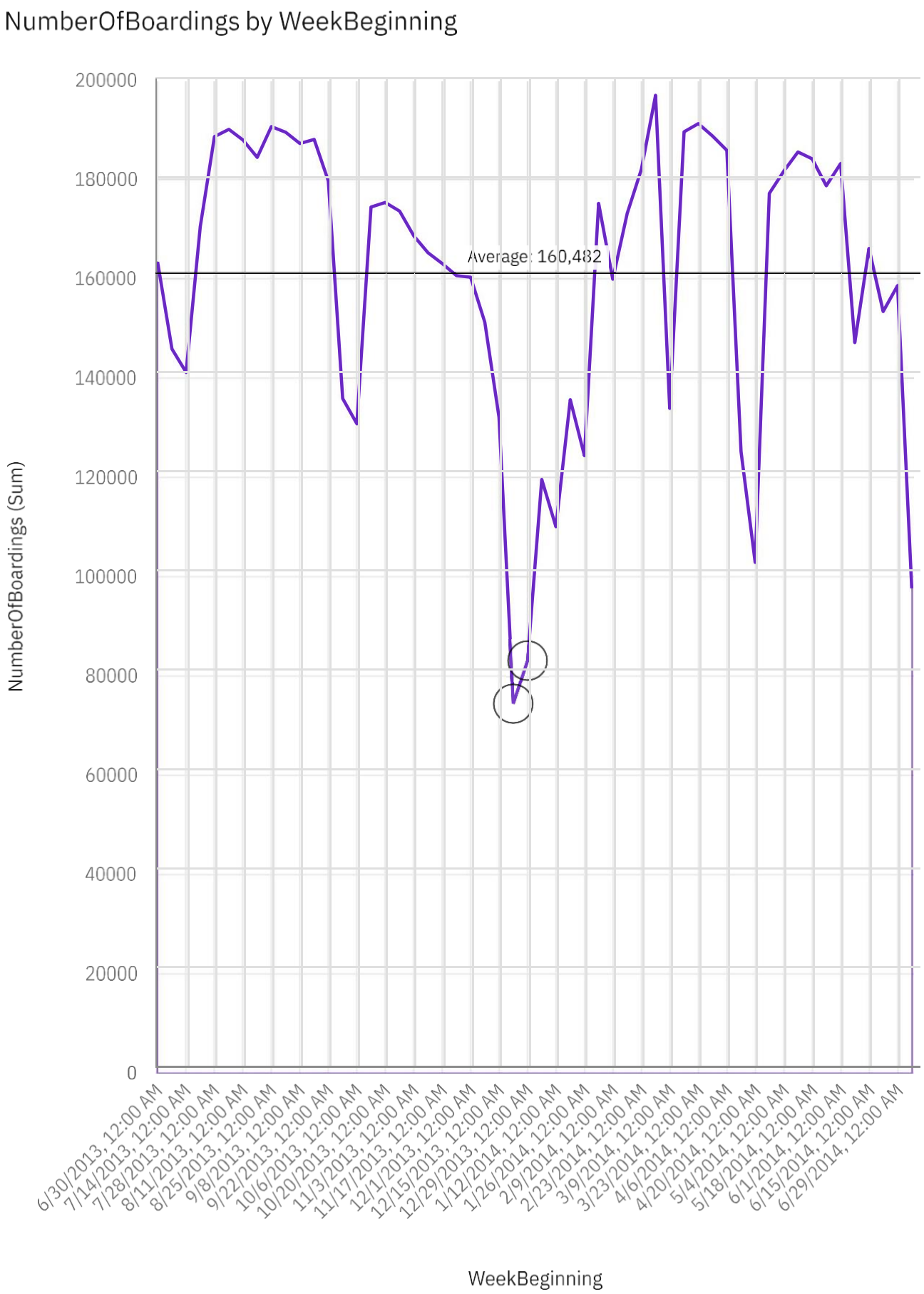
Tab 1


NumberOfBoardings by WeekBeginning colored by StopName
StopName
W1 North Tce
V2 Currie St
V1 Currie St
U1 North Tce
R2 North Tce
I1 North Tce
G3 Grenfell St
G2 North Tce
E2 Currie St
D North Tce
NumberOfBoardings (Sum)
7000
6500
6000
5500
5000
4500
4000
3500
3000
2500
2000
1500
1000
500
0
3/30/2014, 12:00 AM
3/23/2014, 12:00 AM
3/16/2014, 12:00 AM
3/2/2014, 12:00 AM
9/15/2013, 12:00 AM
9/1/2013, 12:00 AM
8/25/2013, 12:00 AM
8/11/2013, 12:00 AM
8/4/2013, 12:00 AM
7/28/2013, 12:00 AM
WeekBeginning

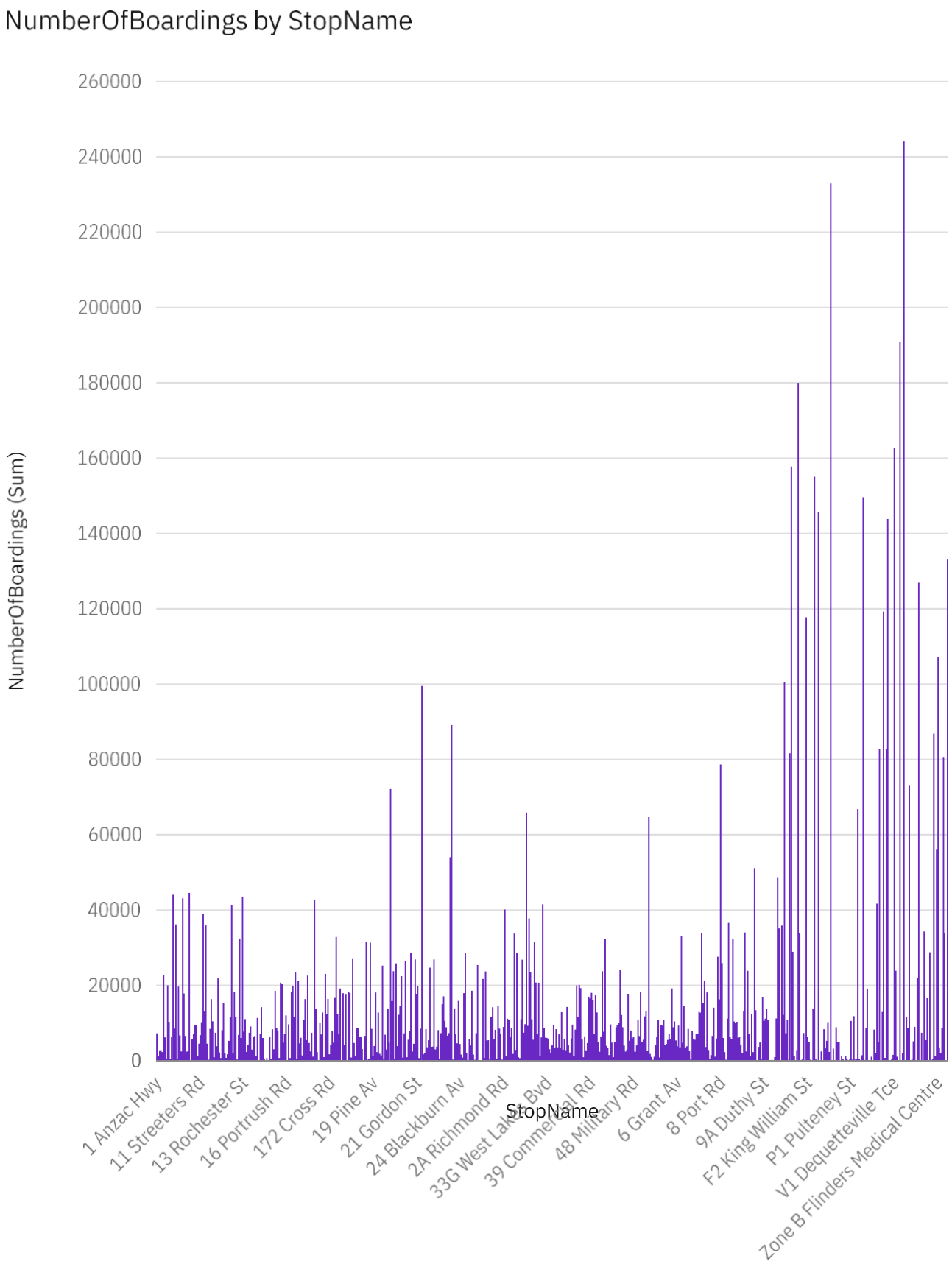
Tab 2


NumberOfBoardings by RouteID colored by StopName
StopName
W1 North Tce
V2 Currie St
V1 Currie St
U1 North Tce
R2 North Tce
I1 North Tce
G3 Grenfell St
G2 North Tce
E2 Currie St
D North Tce
NumberOfBoardings (Sum)
100000
90000
80000
70000
60000
50000
40000
30000
20000
10000
0
168
167
157
150
147
142
141
118
117
RouteID

Tab 3


NumberOfBoardings by WeekBeginning
NumberOfBoardings (Sum)
200000
180000
160000
140000
120000
100000
80000
60000
40000
20000
0
Average: 160,482
6/30/2013, 12:00 AM
7/14/2013, 12:00 AM
7/28/2013, 12:00 AM
8/11/2013, 12:00 AM
8/25/2013, 12:00 AM
9/8/2013, 12:00 AM
9/22/2013, 12:00 AM
10/6/2013, 12:00 AM
10/20/2013, 12:00 AM
11/3/2013, 12:00 AM
11/17/2013, 12:00 AM
12/1/2013, 12:00 AM
12/15/2013, 12:00 AM
12/29/2013, 12:00 AM
1/12/2014, 12:00 AM
1/26/2014, 12:00 AM
2/9/2014, 12:00 AM
2/23/2014, 12:00 AM
3/9/2014, 12:00 AM
3/23/2014, 12:00 AM
4/6/2014, 12:00 AM
4/20/2014, 12:00 AM
5/4/2014, 12:00 AM
5/18/2014, 12:00 AM
6/1/2014, 12:00 AM
6/15/2014, 12:00 AM
6/29/2014, 12:00 AM
WeekBeginning

Tab 4


NumberOfBoardings by StopName
NumberOfBoardings (Sum)
260000
240000
220000
200000
180000
160000
140000
120000
100000
80000
60000
40000
20000
0
StopName
1 Anzac Hwy
11 Streeters Rd
13 Rochester St
16 Portrush Rd
172 Cross Rd
19 Pine Av
21 Gordon St
24 Blackburn Av
2A Richmond Rd
33G West Lakes Blvd
39 Commercial Rd
48 Military Rd
6 Grant Av
8 Port Rd
9A Duthy St
F2 King William St
P1 Pulteney St
V1 Dequetteville Tce
Zone B Flinders Medical Centre

ANALYSIS FROM COGNOS IBM:

- Over all values of WeekBeginning and StopName, the sum of NumberOfBoardings is almost 383 thousand.
- The summed values of NumberOfBoardings range from almost 2500 to over 6500.
- NumberOfBoardings is unusually high when the combinations of WeekBeginning and StopName are 02 Mar 2014 and I1 North Tce, 23 Mar 2014 and I1 North Tce, 16 Mar 2014 and I1 North Tce and 30 Mar 2014 and I1 North Tce.
- NumberOfBoardings is unusually high when StopName is I1 North Tce.
- For NumberOfBoardings, the most significant values of WeekBeginning are 23 Mar 2014, 04 Aug 2013, 16 Mar 2014, 02 Mar 2014, and 30 Mar 2014, whose respective NumberOfBoardings values add up to almost 193 thousand, or 50.4 % of the total.
- For NumberOfBoardings, the most significant values of StopName are I1 North Tce and W1 North Tce, whose respective NumberOfBoardings values add up to almost 107 thousand, or 27.9 % of the total.

- NumberOfBoardings is unusually high when StopName is I1 North Tce.
- Over all values of RouteID and StopName, the sum of NumberOfBoardings is almost 1.3 million.
- The summed values of NumberOfBoardings range from 2 to nearly 91 thousand.
- NumberOfBoardings is most unusual when the combinations of RouteID and StopName are 150 and W1 North Tce, 142 and E2 Currie St, 141 and E2 Currie St, 168 and V2 Currie St, 118 and V1 Currie St and more.
- NumberOfBoardings is most unusual when RouteID is 167, 142 and 147.
- For NumberOfBoardings, the most significant values of StopName are I1 North Tce, E2 Currie St, W1 North Tce, V1 Currie St, and V2 Currie St, whose respective NumberOfBoardings values add up to over 834 thousand, or 65 % of the total.
- For NumberOfBoardings, the most significant values of RouteID are 142, 147, 141, 150, and 157, whose respective NumberOfBoardings values add up to almost 979 thousand, or 76.2 % of the total.

- Over all values of WeekBeginning, the sum of NumberOfBoardings is almost 8.7 million.
- NumberOfBoardings ranges from over 74 thousand, when WeekBeginning is 22 Dec 2013, to 197 thousand, when WeekBeginning is 02 Mar 2014.
- NumberOfBoardings is unusually low when WeekBeginning is 22 Dec 2013 and 29 Dec 2013.

- Over all values of StopName, the sum of NumberOfBoardings is almost 8.7 million.
- NumberOfBoardings ranges from 2, when StopName is 11 East Av, to over 244 thousand, when StopName is W1 North Tce.
- For NumberOfBoardings, the most significant values of StopName are W1 North Tce and I1 North Tce, whose respective NumberOfBoardings values add up to nearly 477 thousand, or 5.5 % of the total

CONCLUSION:

- This analysis have explored how people are travelling from different stops in Adelaide Metropolitan area and the rate at which passengers on each bus route are increasing Finally created a predictive model to find the load of passengers on public Bus transport system in future.
- X2 King William St and stop near to that are the most busiest stops in the city. which having number of boardings per week more than 10k.