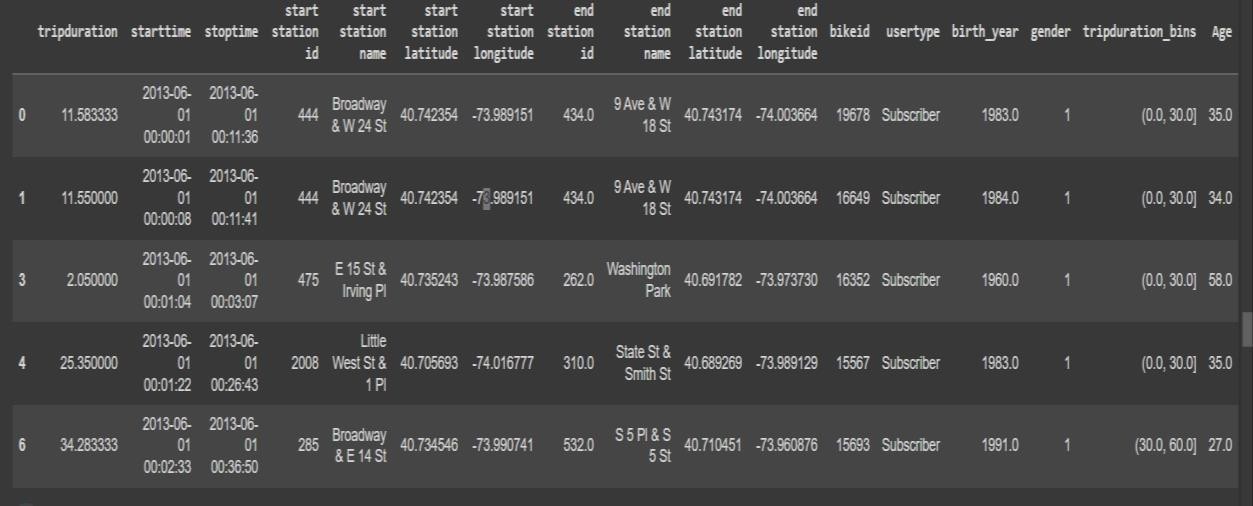
Project Development Phase – Sprint 2

|  |  |
| --- | --- |
| Date | 31 October 2022 |
| Team ID | PNT2022TMID22202 |
| Project Name | A new hint to transportation – Analysis of the NYC bike share system. |
| Maximum Marks | 20 Marks |

# Feature Engineering:

**calculating Age from birth year** from datetime import datetime, date age=2018-df['birth\_year'] df['Age']=age

df.head()

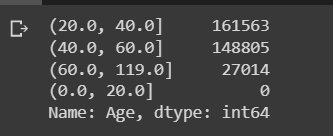


**calculating age group from age** max\_limit = df['Age'].max() max\_limit

bins = [0,20,40,60,max\_limit]

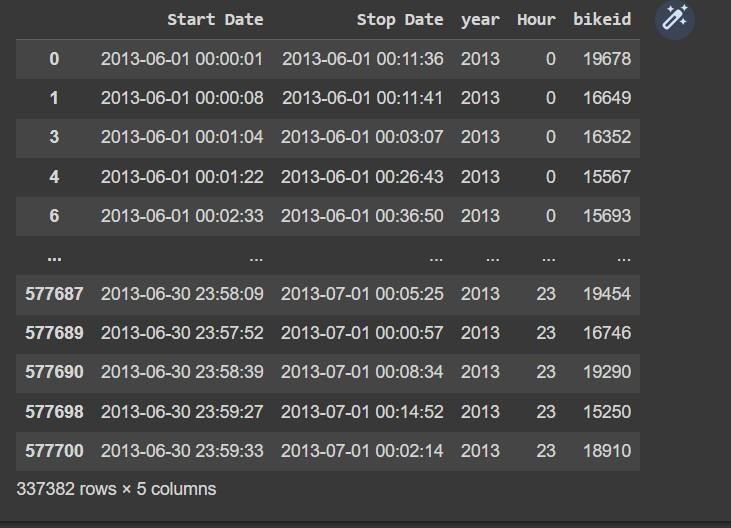
agegroup = pd.cut(df['Age'], bins=bins).value\_counts()

Agegroup

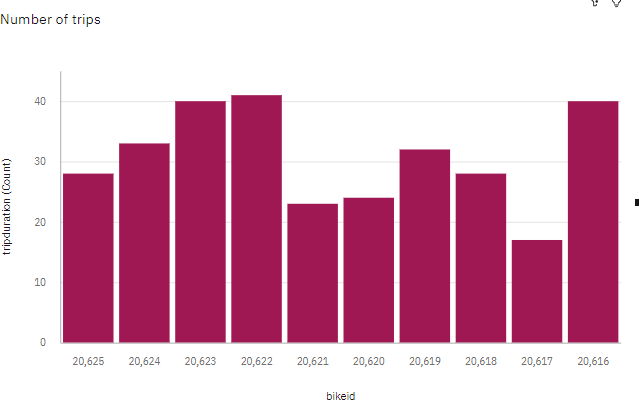


# calculating hour

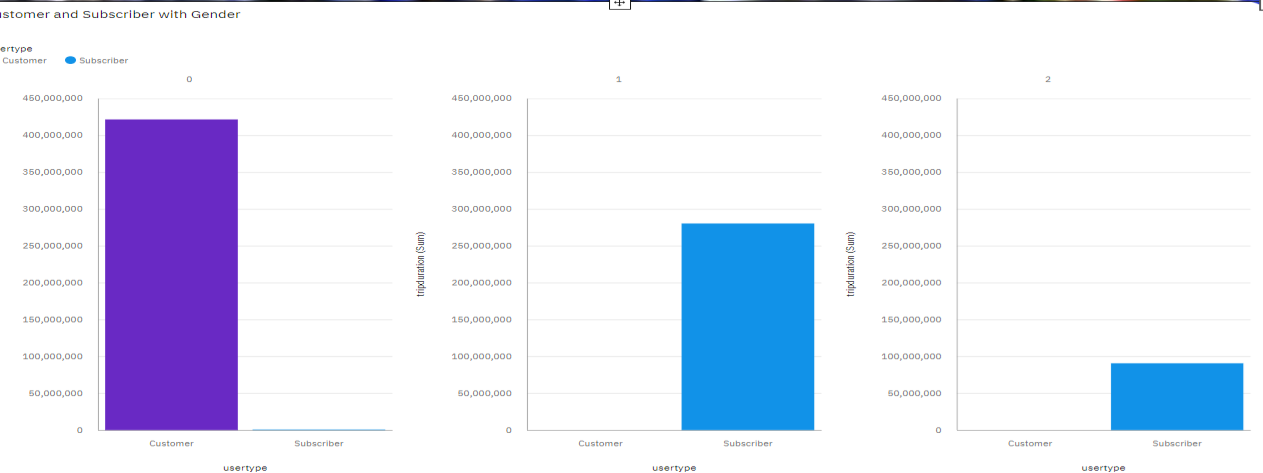
peak\_hour['Start Date'] = pd.to\_datetime(df['starttime']) peak\_hour['Stop Date'] = pd.to\_datetime(df['stoptime']) peak\_hour['year'] =peak\_hour["Start Date"].dt.year peak\_hour["Hour"] = peak\_hour["Start Date"].dt.hour



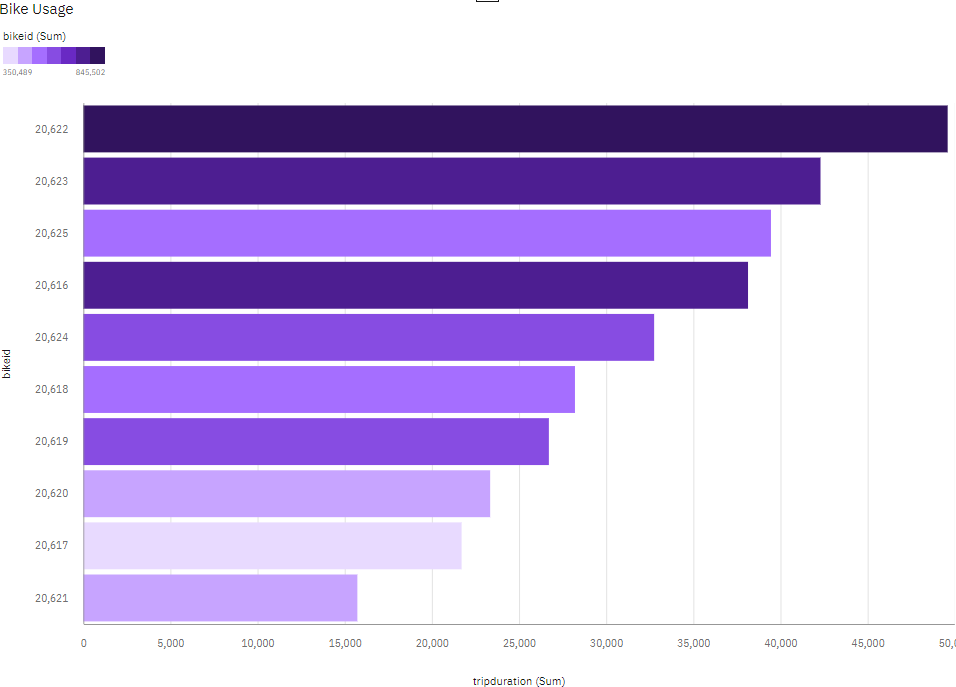
# Visualization of the dataset in COGNOS Platform: Finding the number of trips per each bike:



**Finding the percentage of customers and subscribers**



**Bike Usage - Bike Id Vs Trip Duration:**



# Age Group Differentiation by BikeId: Calculation:

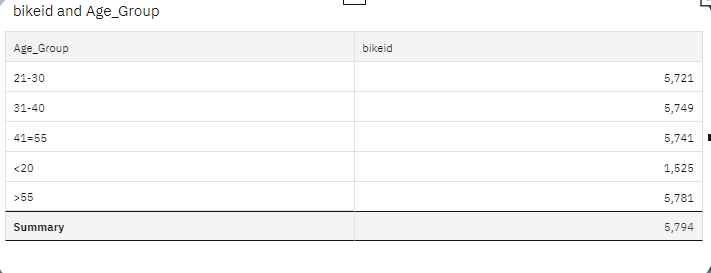
if(age<=20) then ('<20')

else if(age>=21 and age<=30) then ('21-30')

else if(age>=31 and age<=40) then ('31-40')

else if(age>=41 and age<=55) then ('41-55')

else('>55')



# Finding the top 10 start stations with customer age group:

