**EDA - Overview of trip.csv**

#Step 1: glimpse(trips) #here we see all variables are of type int or character

as.data.frame(status(trips)) #here we see metrics: data types, zeroes, infinite/missing values. There are 50 zeroes in zip\_code, no NA or infinity values acorss variables, and subscription\_type with only 2 unique values

#Step 2: Use freq to analyze character/factor variables –

**Frequency of subscription type**

A diagram of a customer

Description automatically generated with medium confidence

**frequency of start stations**

**A rainbow colored line with white text

Description automatically generated**

**frequency of end stations**

**A rainbow colored line with white text

Description automatically generated**

**Step 3: Analyse numerical data**

**A graph of different colored bars

Description automatically generated**

**print(profiling\_num(trips)) #runs for all numerical/integer variables to describe their distributions – MV DESCRIBE ANYTHING YOU SEE OF INTEREST BELOW**

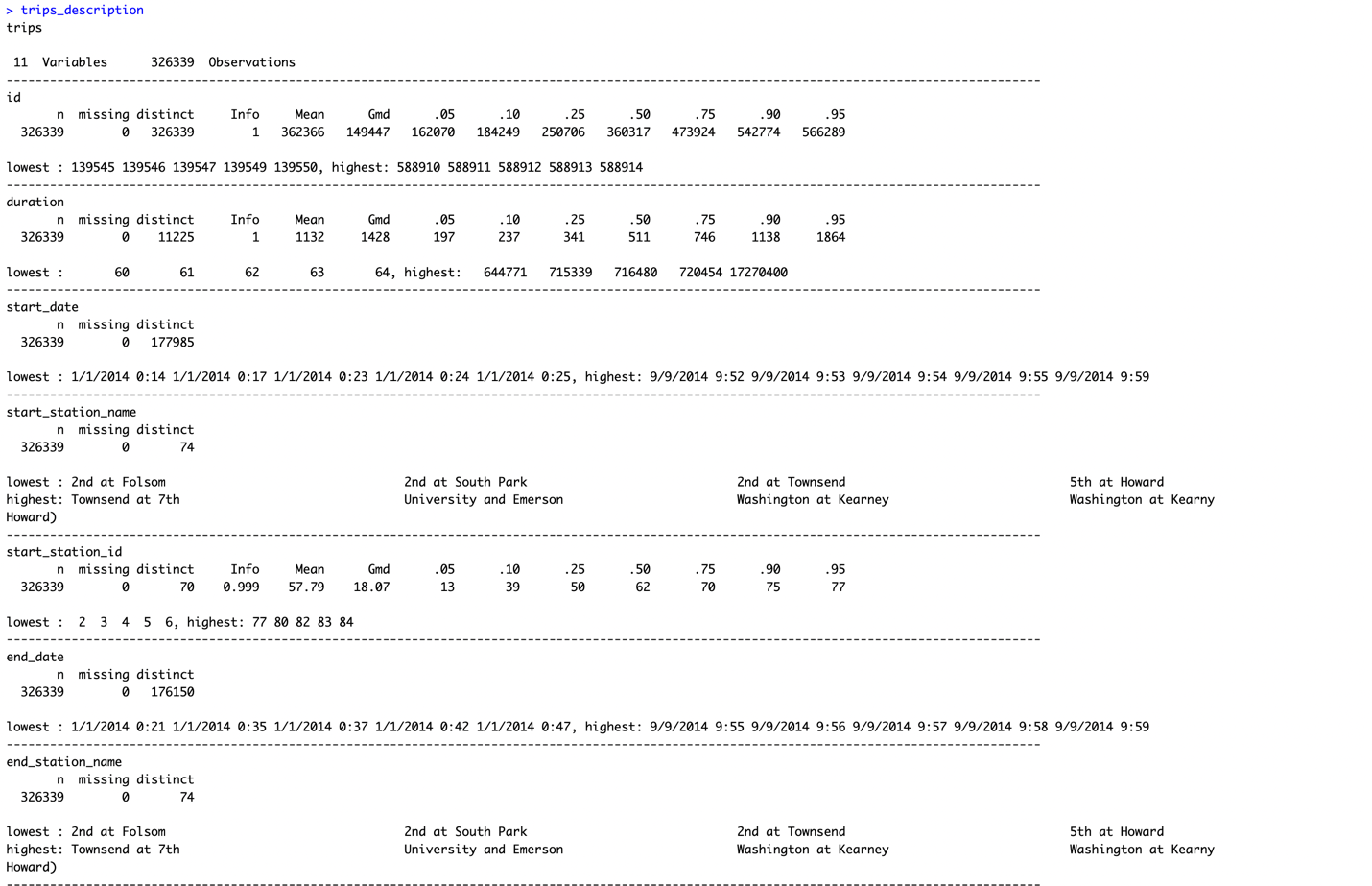
A white background with numbers and numbers

Description automatically generated with medium confidence

#step 4: describe all numerical and categorical data

Zip code has lots of missing! MV LOOK THROUGH IT ALL!

A screenshot of a document

Description automatically generated

**EDA - Overview of weather.csv**