# Hypothesis

How much are the Citi Bikes used in Winter versus Summer between New York and Jersey City, NJ.

# Data Source

Exported the below CSV files from <https://s3.amazonaws.com/tripdata/index.html>

#### Summer 2022, NY and NJ

* 202206-citbike-tripdata.csv
* JC-202206-citibike-tripdata.csv

#### Winter 2022, NY and NJ

* 202201-citibike-tripdata.csv
* JC-202201-citibike-tripdata

# Data Cleaning

1. After appending all files, I split the columns “started\_at” and “ended\_a” to contain dates separate from time.
   * This would later allow me to calculate the start/end time for an average ride time in Tableau
2. I removed columns that were string or text heavy to reduce processing time when loading into Tableau.
   * I created a “Key” table to use during joining to return the values for station names.
3. I changed string/text values for bike type and subscriber status to integers to reduce file size.
   * The values were also added to the “Key” table for lookup in Tableau.
4. I dropped any rows that contained na’s to reduce the overall file size and eliminate null values.
5. Exported as CSV to load to Tableau.

# Analysis Summary

* Overall Sunday is the most popular day for casual riders to ride a Citi Bike. Travel time is less on electric bikes and they travel at greater distances.
* Summer is the most popular time to ride a Citi Bike, Wednesday and Thursday being the busiest days at the Central park stations. Classic bikes are the most popular type, majority of users are members.
* The stations below are the top 20 most popular stations used. Seeing a trend in and around Time Square and Columbus Circle.