Read me file for Richard Soos Data Visualisation Module 18 Challenge 07Jul2023 Monash

Acknowledgements to <https://www.citibikenyc.com/system-data> New York Citi Bike

Richard Soos Tableau link is at https://public.tableau.com/views/RichSoosTableauWk18Challenge/EndptsHeight?:language=en-US&publish=yes&:display\_count=n&:origin=viz\_share\_link

I intend to see if elevation has an effect on bicycles being docked (I do not know yet how to identify that a bike has naturally been returned or has been returned by the company itself or a bicycle angel incentivised to return the bicycle).

I have been told that these bikes have a shortfall of returns (Wikipedia link)

Prospect Heights Brooklyn Elevation 33m / 108 feet

Crown Heights Brooklyn Elevation 47m / 155 feet

Jackson Heights Queens Elevation 24m / 79 feet

And I would like to check

Cobble Hill Brooklyn Elevation 336m / 1102 feet

Clinton Hill Brooklyn Elevation 29m / 95 feet (not likely – but it had the word hill in it)

To obtain elevation data for my mapping I wanted an API key for Google Elevation API

<https://console.cloud.google.com/getting-started>

API key , new project SoosProject5005

API and Services 🡪 Credentials Create Credentials

My API key AIzaSyB5uqam59EqOQ2cxEppCddBVvN-BlC1sAA

<https://maps.googleapis.com/maps/api/elevation/json?locations=40.72365896%2C-74.0641943&key=AIzaSyB5uqam59EqOQ2cxEppCddBVvN-BlC1sAA>

{

"results" : [

{

**"elevation" : 20.38863563537598,**

"location" : {

"lat" : 40.72365896,

"lng" : -74.0641943

},

"resolution" : 4.771975994110107

}

],

"status" : "OK"

}

Enriching the data. All procedures Below are to create a list of heights by lon and lat, to the 3rd decimal place. For our data this is beyond simple stations. It is across 5573 locations.

The heights are then placed to each of the bicycle movement rows and a calculated field of difference in height is also added. We should be able to see of people avoid climbs. We will also put in a flag to show whether the journey was at least 2 minutes in length.

We have 2 data files JC-202305-citibike-tripdata.csv and 202305-citibike-tripdata.csv (JC stands for Jersey City).

I make api calls to find the height at each point, but first a create a distinct list of lat and lon; otherwise the number of calls will be so great I will be denied service. Also, I had to restrict the decimal points of lat and lon to just 3, as otherwise there are very few duplicates.

1 decimal place: Approximately 10 kilometers

2 decimal places: Approximately 1 kilometer

3 decimal places: Approximately 100 meters

4 decimal places: Approximately 10 meters

5 decimal places: Approximately 1 meter

6 decimal places: Approximately 0.1 meter or 10 centimeters

As we are interested in hills, being within 100 metres at most of where the height is taken is the best we can do.

Run pointloc.ipynb which reads the 2 mentioned csv files, placing all location (lon, lat) cords into a new file All\_loc.csv

Run heightextract.py which will read from All\_loc.csv and create a similiar file All\_height\_loc.csv but with topographic height as well.

Run JCMerge to create JCCLEANED\_data (reads All\_height\_loc.sc and JC-202305-citibike-tripdata.csv) This will be one of the two readable files

Run plainmergescript to create NYCLEANED\_data.csv (reads 202305-citibike-tripdata.csv)

To create the other readable file.

To put in a column height\_dif

Run jcheightdif.py which will read JCCLEANED\_data.csv and create JCCLEANED\_heightdata.csv

Run nyheightdif.py which will read NYCLEANED\_data.csv and create NYCLEANED\_heightdata.csv

To put in a column RealJourney to tell whether the trip was at least 2 minutes

Run jcreal.py which will read JCCLEANED\_heightdata.csv and create JCReal\_cleaned.csv

Run nyreal.py which will read NYCLEANED\_heightdata.csv and create NYReal\_cleaned.csv

Dashboard 1 Height effect Dashboard

Members and casual pass holders follow the same aversion to travelling a net trip with an elevation in height. The sites where incentives are offered to return bikes are elevated sites, but are more generally out of the way of bicycle travel. More people depart from elevation sites.

Dashboard 2 Circuit Home Dashboard

2 min trips are not real trips and no height is climbed or lost.

Time spent in real trips is not a predictor of net height climbed/lost

For any chance of bikes being returned within 100 metre grid of collection, it must be done Pronto. Overwhelmingly trips are not return to complete trips on same bicycle excursion.

Dashboard 3 Bike Type Time

Members take trips of shorter duration (half)

Neither members nor casuals travel long duration based on bike type

There is no preference for electric bikes for trips with net climb in elevation

Dashboard 4 Weekday Trips Dashboard

One particular day aside (Sat 20 May 2023 that experienced Heavy Intensity rain) , generally Casuals travel more on Saturdays on these bicycles, whereas members’ use drop of on this day. The summed data is difficult to find any trend in.

Dashboard 5 False docking areas Dashboard

False dockings are often attempts of a rider to return their bicycle again as they are unsure that the end of transaction service has been recorded. By gathering numbers we can send technicians to see if the stations are well maintained, or member or casual user needs education. There is some commonality of stations where false docking occurs in the profile of members and casual users.