

## Directory:

### 1. shaiwal/Uber/

**ByTimezone** --> contains all the data collected

**dataset** --> contains the NYC Dataset and also locations file for different timezones

**Some links**

#### /ByTimezone

**a ) 0,1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23**

There is folder for each timezone

Each of the above folders have

**logs** --> the dataset is split into many parts and data is collected. This folder has logs for that.

**nyc\_taxi** --> Two file (list of locations(OD) pairs of that timezone) in txt and sqlite

**uber** --> has the json files collected for each split of data

**mergedjsonfile(timezone-uberprice.json.gz)** ---> merging the result.

#### **b) candidates (txt+sqlite)**

this contains the common locations of Timezone 6,10,16,20 around 19 thousand OD pairs.

#### **c)code**

contains all the codes , for more information, see the code\_guide

#### **d)code\_old**

very initial codes

#### **e)code\_scheduling**

contains all the codes using the python scheduler to run the code at a specified time

#### **f)estimator**

it contains the training set for each timezone for our model

For example ,0\_1lat\_long\_popularity\_surcharge.csv --> for timezone 0

#### **g)keys**

**uber** --> **api\_keys.txt** , contains 100 server tokens and **api\_keys1.txt** has 200 server tokens.

**Lyft** --> **uses Outh2.0**, first generates the server tokens(valid for one hour only) and then use them .

#### **h)marker**

the json inside the folder has a set of locations in different clusters, we can plot them on google map

#### **i)mergelog**

when the json files are merged , just a log of that operation

#### **j)other**

contains insights about shell commands

how to find union ? Intersection ? Remove duplicated using many files

#### **k)pdfs**

the pdf files we used information from

## **l)pi\_drop\_profile**

**dropoff.txt** --> list of all dropoff locations

**pickup.txt** ---> list of all pickup locations

**pi\_drop\_union.txt** ---> union of above two files

**/profiles\_json** --> contains the info collected for pi and drop at diff ,timezones into json.

**pi\_drop\_merged.json.gz** --> just a merged file of all the above documents

## **m)plots**

contains all the plots CDF,heatmap files too.

## **n)Project presentation**

all the files for latex report and powerpoint presentation.

## **o)Requirements**

**Miniconda2-latest-Linux-x86\_64.sh**

**geekdon.zip** ---> enviroment is zipped .

After installing, copy to miniconda2/env

## **p) Research Papers**

The papers we referred.

## **q)testing1**

results of testing a random set of 50 locations versus the real surcharge at that location returned by Uber.

## **2. shaiwal/OSM/ (Open Street Map)**

**code**

**roadnetwork\_plot\_v2.py** ---> Useful code for plotting the roadnetwork within the boundary locations of the latitude and longitude

**roadnetwork.py** --> for generating the nodes.txt and edges.txt and visulization using networkx library.

**roadnetwork\_v2.py** ---> for generating the nodes.txt and edges.txt

**(parsed xml using BeautifulSoup)**

**data (input xml file)**

**output**

For a road map, we would have nodes (lat\_long) and edges (lat\_long\_lat\_long),

**edges.txt**

**nodes.txt**

**bs4.zip** --> beautifule soup package file

**install\_beautiflsoup(bs4)**