

## Yellow Taxi Trip (Final Exam)

Original dataset: <https://www1.nyc.gov/site/tlc/about/tlc-trip-record-data.page>

Fields in the dataset:

VendorID, tpep\_pickup\_datetime, tpep\_dropoff\_datetime, passenger\_count, trip\_distance, RatecodeID, store\_and\_fwd\_flag, PULocationID, DOLocationID, payment\_type, fare\_amount, extra, mta\_tax, tip\_amount, tolls\_amount, improvement\_surcharge, total\_amount, congestion\_surcharge

Field description:

[https://www1.nyc.gov/assets/tlc/downloads/pdf/data\\_dictionary\\_trip\\_records\\_yellow.pdf](https://www1.nyc.gov/assets/tlc/downloads/pdf/data_dictionary_trip_records_yellow.pdf)

Please implement either the first (3 points) or the second (5 points) specification below. Use float to represent tip\_amount. Run your program on VeriFone.csv and upload the python code (not Notebook!) together with the generated output to Moodle. If your code does not generate an output, do not upload anything.

1. Find the average tip [13] for each hour of each day of the week. Take into account the time when the meter was disengaged [2]. The partial output (only Friday) for LLC.csv dataset is shown below:

day	hour	avgTip
Fri	0	2,8562499971
Fri	1	2,7160714226
Fri	2	3,0158620867
Fri	3	2,4014285547
Fri	4	3,5678571633
Fri	5	2,2947368685
Fri	6	2,018939393
Fri	7	2,1328205127
Fri	8	2,1367625966
Fri	9	2,2609150308
Fri	10	2,39755905
Fri	11	2,1913235297

Fri	12	2,4190647452
Fri	13	2,1425503362
Fri	14	2,6980434706
Fri	15	2,3589308174
Fri	16	2,8853147026
Fri	17	2,5497765401
Fri	18	2,5378468908
Fri	19	2,3654010737
Fri	20	2,4272972901
Fri	21	2,4453333318
Fri	22	2,6149444428
Fri	23	2,8026249994

2. For each hour of the day find the day when the average tip is the highest. Sort the results by hour in increasing order (midnight starts at 0). The output for LLC.csv dataset is shown below:

hour	day	avgTip
0	Mon	3,8914893853
1	Mon	4,6004346842
2	Mon	3,7200000763
3	Tue	2,8379310534
4	Wed	4,0125000477
5	Sat	3,9224999952
6	Sun	2,754545434
7	Sun	2,8920000002
8	Sat	2,4712727354
9	Wed	2,7642261837
10	Thu	2,8648780672
11	Tue	2,6273287674

12	Tue	2,779312493
13	Tue	2,4914285651
14	Fri	2,6980434706
15	Sun	2,6959006175
16	Mon	3,0448366115
17	Wed	2,6859624389
18	Wed	2,6928820972
19	Thu	2,8241365621
20	Sun	3,3663636332
21	Sun	3,1088235275
22	Sun	3,0165789402
23	Sun	3,1239215661