

Step 4. Exploratory data analysis (Python)

In addition to the data you retrieved in the previous tasks, you've been given a second file. You now have these two CSVs:

/datasets/project_sql_result_01.csv. It contains the following data:

company_name: taxi company name

trips_amount: the number of rides for each taxi company on November 15-16, 2017.

/datasets/project_sql_result_04.csv. It contains the following data:

dropoff_location_name: Chicago neighborhoods where rides ended

average_trips: the average number of rides that ended in each neighborhood in November 2017.

For these two datasets you now need to

- import the files
- study the data they contain
- make sure the data types are correct
- identify the top 10 neighborhoods in terms of drop-offs
- make graphs: taxi companies and number of rides, top 10 neighborhoods by number of dropoffs
- draw conclusions based on each graph and explain the results

Step 5. Testing hypotheses (Python)

/datasets/project_sql_result_07.csv — the result of the last query. It contains data on rides from the Loop to O'Hare International Airport. Remember, these are the table's field values:

- *start_ts*
 - pickup date and time
- *weather_conditions*
 - weather conditions at the moment the ride started
- *duration_seconds*
 - ride duration in seconds

Test the hypothesis:

"The average duration of rides from the Loop to O'Hare International Airport changes on rainy Saturdays."

Decide where to set the significance level (alpha) on your own.

Explain:

- how you formed the null and alternative hypotheses
- what criterion you used to test the hypotheses and why

How will my project be evaluated?

Here are the project assessment criteria. Read them over carefully before you get to work.

Here's what the project reviewer will look for when assessing your project:

- how you retrieve data from the website
- how you make data slices
- how you group data
- whether you use the various methods for combining data correctly
- how you formulate hypotheses
- what criteria you use to test the hypotheses and why
- what conclusions you reach
- whether you leave comments at each step

The cheat sheets and summaries from previous lessons have everything you need to complete the project.

Good luck!