Practice session Objectives

1) Learn to connect to a remote database through API (Carto, former CartoDB) with Python 2 / 3;

2) Learn to create SQL queries for filtering out data, calculating summary statistics, etc;

3) Develop intuition of SQL queries.

Task 1 – simple queries

1) Sort data by both: start_station_id ascending, and tripduration descending (hint: ASC, DESC);

2) Select last 10 records of the table. Hint: use the table's main id field;

3) List all unique birth years. Hint: distinct;

4) Find minimal, maximal and average trip duration

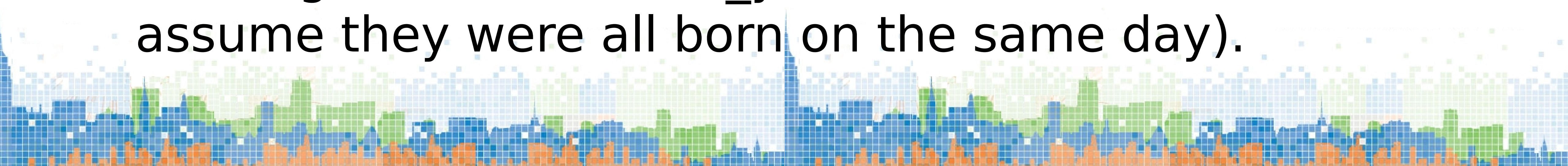hint: min(), max(), avg().

Task 2 – date/time and conditions

1) Select only trips that started at 1 AM

   hint: EXTRACT(HOUR FROM fieldname::time);

2) What is the average birth year of people that ride bikes at 2 AM?;

3) What is the age of the oldest person riding at 3 AM?

   Hint: age = 2018 – birth_year. (For this task we assume they were all born on the same day).

Task 3 – aggregation

1) Find the "start_station_id" that had the highest number of bikes taken from it

   hint: GROUP BY station id, COUNT();

2) Show top 3 "end_station_id" with the largest total "tripduration". Hint: GROUP BY station id, SUM();

3) Find the "start_station_id" with the shortest average trip duration during 1 AM.