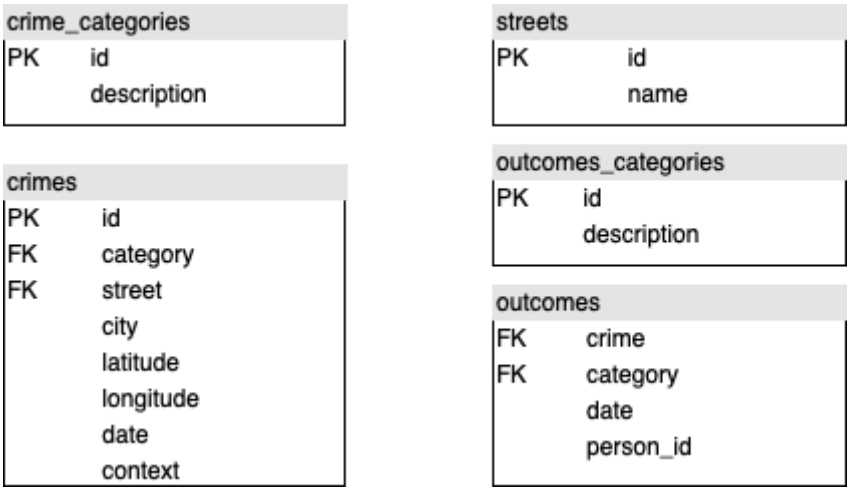
faire.ai's Data engineer case study

Hi and welcome to the faire.ai's case study that we hope you'll enjoy to develop with your best skills and creativity.

CrimeDB

Create an integration layer for data that can be retrieved by means of the Police API. (See here the API documentation at https://data.police.uk/docs/.)

- This integration will populate the following tables:



- Create the Python ETL script that contains the processes to update the tables where:
    i. a first job creates the tables and populates them with the historical data from T_0 to $T_{Current}$;
    ii. a second job that updates the tables with only the new events from $T_{LastUpdate}$ to $T_{Current}$;
    iii. a last job that receives as input the $T_{Current}$, (for example: *"03-2019"*).
- Create a Docker compose file that allows the running of the ETL script.
- Put the project documentation in a README.md in the project root directory.

Case study submission

For submitting the case study and, for you, to keep track of your work, we'd like to get the code from a GIT repository: you can choose GitHub or GitLab or other providers that offer free repositories.

How is it evaluated

During the evaluation of the case study we put the attention on:

- clear code/documentation;
- solutions found.