

The dataset has three files. The data in the first two files can be used for training the model and the third file contains test data.

- **drivers.csv:** This file contains driver profile data
- **pings.csv:** This file contains the driver pings
- **test.csv:** This file has the test data

The fields present in these files are described below.

drivers.csv:

- **driver_id:** Unique driver id
- **gender:** Gender of the driver
- **age:** Driver's age
- **number_of_kids:** The number of kids the driver has

pings.csv:

- **driver_id:** The id of the driver whose ping has been received. It is guaranteed that the details of this driver will be available in the drivers.csv file.
- **ping_timestamp:** The [Unix epoch](#) timestamp when ping was received by the system. Timezone is "Asia/Jakarta".

test.csv:

- **driver_id:** The id of the driver whose ping has been received. It is guaranteed that the details of this driver will be available in the drivers.csv file.
- **date:** The date in YYYY-MM-DD format for which the online hours needs to be predicted.
- **online_hours:** The number of online hours

Given driver_id and date, you should predict online_hours.