You will be dealing with the MIMIC-IV FHIR dataset for the purpose of this assignment.

The link to the dataset is here:
https://physionet.org/content/mimic-iv-fhir-demo/2.0/mimic-fhir/

You can download it via:
https://physionet.org/static/published-projects/mimic-iv-fhir-demo/mimic-iv-clinical-database-demo-on-fhir-2.0.zip

The individual files in the dataset are FHIR resources. Resources are nothing but objects of importance in a hospital system. You can get more familiar with FHIR resources here:

https://hl7.org/FHIR/resourcelist.html

Although being familiar with the resources is not essential in executing the assignment.

The main files of interest for you are:

1. Patient.ndjson
2. Condition.ndjson
3. Encounter.ndjson
4. EncounterICU.ndjson (this file has the same format as Encounter.json)

“Patient.ndjson” file contains all the patients in the mimicIV demo dataset. Each line of this file represents an individual patient json file.
“Condition.ndjson” file contains diseases that were assigned to patients. Each line of this file represents a json referencing a specific disease that was assigned to a patient. The json object also references an encounter that was associated with each such condition.
“Encounter.json” file contains the details of an encounter in each line.

Note that each line in the Condition file references an Encounter and a Patient.

Your goal in this assignment is to:

1. Get familiar with the json object format contained in each line of the 4 files:
    a. Patient.ndjson
    b. Condition.ndjson
    c. Encounter.ndjson
    d. EncounterICU.ndjson
2. For each patient, create an array of conditions associated. The expected output for this is a dictionary with patient_id as key and an array of Condition json as value.
3. For each condition, assign an estimated time for the condition using the corresponding encounter in the Encounter.json or EncounterICU.ndjson. Choose the start_time in the Encounter to associate time to each condition.
4. Finally, create a csv file with the following columns:
    a. Patient_id (Column name: pid)
    b. Timestamp (unix format) (Column name: time)

    c. Condition code (Column name: code)
    d. Condition string (Column name: description)
5. You are required to submit the csv file as well as the jupyter notebook used to generate the csv file

If there are any issues, feel free to message.