**DATASET**

The dataset contains a comprehensive collection of electronic

health records belonging to patients who have been diagnosed with a specific

disease. These health records comprise a detailed log of every aspect of the

patients' medical history, including all diagnoses, symptoms, prescribed drug

treatments, and medical tests that they have undergone. Each row represents a

healthcare record/medical event for a patient and it includes a timestamp for each

entry/event, thereby allowing for a chronological view of the patient's medical history

The Data has mainly three columns

1) Patient-Uid - Unique Alphanumeric Identifier for a patient

2) Date - Date when patient encountered the event.

3) Incident - This columns describes which event occurred on the day.

There are total 27K unique patients present in train.parquet

Similar to drug\_type\_7, there is also drug called ‘Target Drug”, this drug is of interest for the

assignment, there are total 9K patients in train.parquet who have taken “Target Drug” atleast

once.