**Take Home Project**

The following is a lightweight project to get to know you as a data analyst and your level of expertise as an data/ETL engineer.

We highly recommend that you use NodeJS/Python in your development, however, any programming language is accepted.

You will be writing an ETL process. The goals of the project are twofold:

1. Have final, workable data
2. Write a good, clean code.

You will be coding the ETL process according to the specifications below. Please include all the required aspects, as well as how you work around problems in the data (if there are any).  
  
You have received an excel document titled “Mapping”, that has sheets of the relevant IDs that we know how to handle for each of our entities (Drugs/ labs/ conditions).

You have also received other excel documents with patient data for each of the entities (Drugs/ labs/ conditions).

If we know how to handle a given entity row (all the data is there), you need to insert them into a new excel specifically for the relevant patient.

Create a new sheet in the relevant excel for each entity.

In the end we expect the same amount of excel documents as there are patients. In each of the excel documents we will have only the data relevant to that patient (according to the specifications below).

These are the specifications of which data to save, as well as the format in which you should save them:

**Drugs**

We only want to take drugs which meet the following criteria:

* Their date is within the last 12 months (from today)
* The drug is approved
* We know how to handle the NDC
  + If we don’t know, please write on the MDI ID column that it’s “NEW”

For each drug include: NDC, date, Strength, Unit, Dosage Form, days supply, MDI ID

**Labs**

* We know how to handle the LOINC
  + If we don’t know, please write on the MDI ID column that it’s “NEW”

For each lab include: LOINC, date, value, name, MDI ID

**Conditions**

* We know how to handle the ICD
  + If we don’t know, please write on the MDI ID column that it’s “NEW”
* For each condition include: ICD, date, name, MDI ID