How To Run, assumptions and outputs

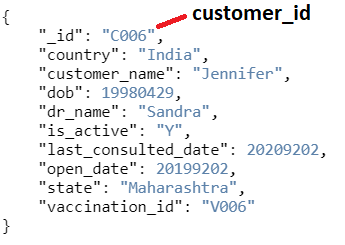
How to run

1. Please install the following python dependencies via pip:
   1. pandas==1.1.4
   2. tabulate==0.8.9
   3. dnspython==2.1.0
   4. pymongo==3.12.0
2. Please place the ‘raw\_data.csv’ and ‘run.py’ in the same folder.
3. Run the python code using the following command: ‘python run.py’
4. You can chk the databse directly by logging into mongodb cloud database.   
   link - <https://www.mongodb.com/cloud/atlas>  
   email – soumyadipauddy.incubyte@protonmail.com

Password - Photo@2021

Assumptions

1. Since mongo is used here, we need to declare a primary key for all the tables. Here ‘customer\_id’ is used as a primary key. The primary key in mongo is named as ‘\_id’. So ‘\_id’ in the tables means ‘customer\_id’
2. Table schemas:

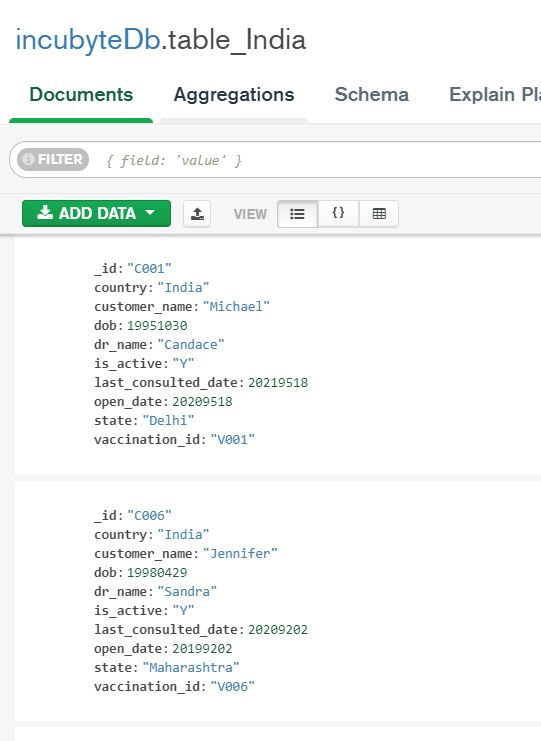


Output

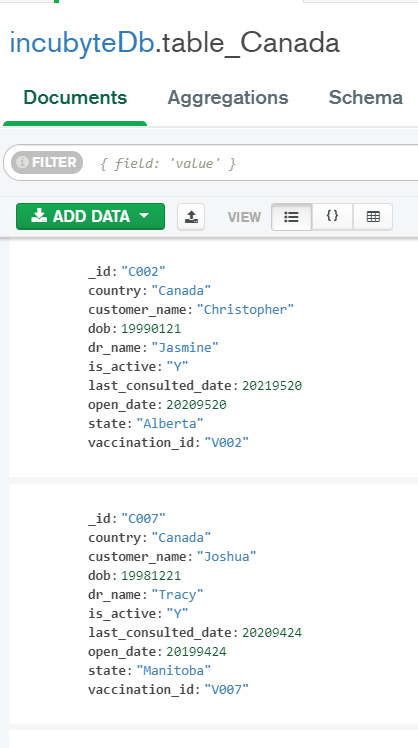
Please login to Mongodb cloud databse to see the outputs.  
link - <https://www.mongodb.com/cloud/atlas>  
email – soumyadipauddy.incubyte@protonmail.com  
Password - Photo@2021

In my raw data there are 5 countries: 'India', 'Canada', 'Uganda', 'Spain', 'Poland'. So 5 tables have been created.

1. India



1. Canada



Similarly, rest of the countries have their separate table.