Snizhana Kurylyuk

Abstract

CDC estimates that each year roughly 1 in 6 Americans (or 48 million people) gets sick, 128,000 are hospitalized, and 3,000 die of foodborne diseases. People should feel comfortable going to a food distributing facility knowing they have successfully passed their inspections. The goal of this project was to create a working pipeline that can actively pull and clean data with one run. As well as create a dashboard where community members can easily view inspection results and make informed decisions about where they choose to eat. My app will be deployed to Streamlit where it can be interacted with based on different inspection variables.

Design

The pipeline is going to pull data from the city of chicago inspections website. It will read in the data through a SOAP API. Then the pipeline will save the dataset to SQLlite. Afterwords, the data will be imported into python for analysis and feature engineering. Once the data is clean it will be pulled into Tableau for data visualization and lastly deployed to Streamlit.

Data

The dataset was pulled using a SOAP API from the City of Chicago website. This information is derived from inspections of restaurants and other food establishments in Chicago from January 1, 2010 to the present. Inspections are performed by staff from the Chicago Department of Public Health's Food Protection Program using a standardized procedure. The results of the inspection are inputted into a database, then reviewed and approved by a State of Illinois Licensed Environmental Health Practitioner (LEHP).

Algorithm & Tools

- Numpy and Pandas for data manipulation
- SOAP API for Data Ingestion
- SQL alchemy for extraction into python
- SQLite3 for Data storage
- Scikit-learn for modeling
- Matplotlib and Seaborn for plotting

Snizhana Kurylyuk

Communication

The presentation can be presented via powerpoint. Dashboard created in Streamlit.