Pandemic Implications on Iowa Spirits Sales

——————

Project Proposal

Business Problem

The Covid-19 pandemic is the major event of the 3rd Millennium that changed lives and businesses. We know that some businesses went bankrupt or had to close their businesses overnight, such as bars, restaurant, hotels, or aviation. Some other flourished, such as online shopping, grocery stores or delivery services. Our focus is to use data to create business insight, to reveal patterns and make predictions. And nowadays more than ever we need data to understand how the pandemic impacted different parts of economy. Our project will be directed on analyzing the spirits sales in Iowa since 2012 to present.

This project looks to resolve some issues that are of high importance for a diverse number of involved entities:

- Exploration on what was the impact of Covid-19 on the Alcoholic Beverages Industry.
- Storage capacity management exploratory analysis for Iowa Department of Commerce, Alcoholic Beverages Division.
- Cohort Analysis and Customer Segmentation using RFM(Recency, Frequency and Monetary value) and Unsupervised Learning
- Using time series analysis and predictions to predict sales for next month for a hypothetical liquor store.
- Lastly, we want to assist a hypothetical liquor store owner in Iowa in expanding to new locations throughout the state.

Datasets

The main dataset used is the Iowa Liquor Sales database from ***Data.Iowa.gov.***
It contains more than 24 million records of spirits purchase of Class "E" liquor licenses by product and date of purchase from January 1, 2012 to current, data provided and updated monthly by Iowa Department of Commerce, Alcoholic Beverages Division.
The data contains labels such as Invoice number, Store, Address, Zip Code, Geographical Location, beverage category, vendor name, Item Description, State Bottle Cost, State Bottle Retail, Bottles Sold and Sale.

The fact that the data is exhaustive for all the sales of this kind in the state of Iowa is a great statistical feature of our data because we are working with the whole population of sales of this category of alcoholic beverages and not with just a sample, which allows us to create powerful business insights with great confidence levels.

In addition, we will utilize other datasets regarding Demographics or per Capita Personal Income in the State of Iowa, available on the website mentioned above.

Anticipated Data Science Approach

One of the first issues we will encounter will be dealing with the big number of records, and for this we plan to use the Dask library to reduce the computational time.
In the Data Wrangling part, we will collect the additional data and clean data of missing values and outliers. We will also merge and group our data by different features using SQL, with the purpose of creating different subsets to be used in our EDA or Modeling parts to help us answer the proposed Business Problems.

In our EDA we will also answer 2 Business Problems stated in the beginning of this document.

      1.We will conduct Exploratory Data Analysis on <u>what was the impact of Covid-19 on the Alcoholic Beverages Industry</u>. We will take into consideration the demographic trans and Income per capita, but we will also make our recommendations and analysis considering the ***"regression to the mean"*** statistical principle.
      2. We will also conduct Storage Capacity exploratory analysis with the purpose of <u>predicting how the storage needs will change withing one year</u> from the current date.

Then, after the Pre-processing part, in the Modeling section of the Project we will use <u>time series analysis and predictions to predict Next Month's sales</u> for a hypothetical liquor store.

Lastly, using Regression Algorithms we will answer the last and the main Business Problem of the Project: assisting a hypothetical liquor store owner in Iowa in expanding to new locations throughout the state.

Deliverable:

- All code - Jupiter Notebooks
- Final Report of the Project
- Presentation Slide Deck
- A Tableau Story containing Dashboards and Worksheets relevant for the hypothetical liquor store owner to assist him in making data-driven business decisions.