Technical Write-Up Iowa Liquor Sales Project

1. Introduction: A widely used dataset in DA bootcamps is the Iowa Liquor Sales dataset which at the time this project was done (June 2025) consisted of 31M rows of unique, single line transactions of liquor sales in the state of Iowa. My objective was to acquire the data, clean and format it, understand the “Story of a row”, process and analyze the data, and reveal any interesting and actionable insights.

2. Background: This project was used as a practice/learning tool for understanding my technical competency with SQL, Python, and data visualization tools.

3. Technical details: I prefer to clean the data in Python since there are many libraries such as Pandas and Numpy that make identifying and modifying information in the dataset more efficient. It is also nice to be able to spot check work after making any changes with df.head or df.columns.

For querying the dataset, I found SQL to be a useful tool in terms of the syntax breaking down each line of code into more specific arguments. I learned how to pull very specific statistics from my data that could reveal more insights that just summary statistics could. For instance, YEAR(Date) AS Year, AVG(Sale\_Dollars / NULLIF(Bottles\_Sold, 0)) AS Avg\_Retail\_Transaction for 2019 and 2025 showed that the average transaction (per bottle) went from $15 to $20. From here I could have gone into further analysis as to what caused this; socioeconomic, inflation, price increases, world events, etc.

4. Implementation: Based on my training at General Assembly I knew I would get some meaningful data based on the sheer volume of data. The goal however was more so to demonstrate technical competency rather than analytical prowess. This project has been completed by hundreds of students and has been dissected extensively. However, the ability to consolidate 16M (cleaned) rows of data into a single dashboard and provide intuitive controls and descriptions is what I focused on. The goal was to allow technical and non-technical audiences to draw insights without my direction.

5. Results: I designed the dashboard in an ergonomic, intuitive manner while providing what I believe to be the most insightful information. The goal was to facilitate rapid, accurate information gain to the end-user. Utilizing buttons, bookmarks, and tooltips I was able to share what took hours to gather into informative, condensed paragraphs.

6. Conclusion: I was able to learn a lot with this project and am beginning to discover what style I prefer, and by limiting myself to one dashboard, how to minimize clutter yet display the most relevant information possible.

7. References: https://data.iowa.gov/Sales-Distribution/Iowa-Liquor-Sales/m3tr-qhgy/about\_data

Power BI Desktop, SSMS, Jupyter Notebook, Python, SQL