

# Data Management and Visualization Assignment Report

Iowa State Liquor sales and Volume Dashboard

**November 20, 2021**

**Prepared By**

Shubham Rai (Student No: 21261161)

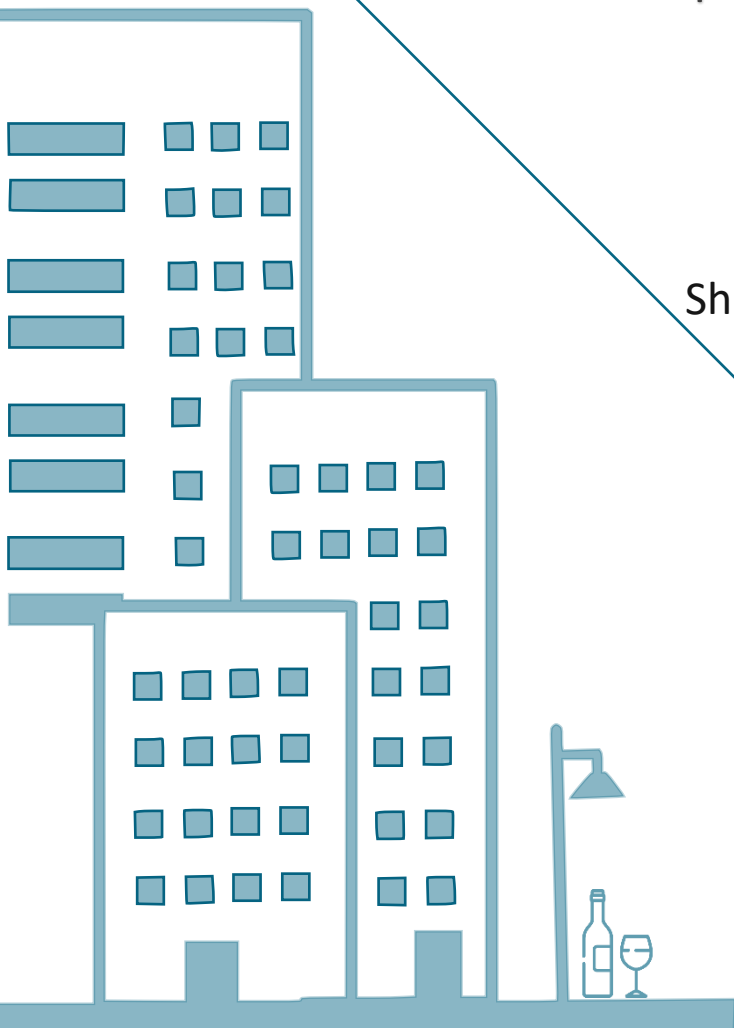
Sarthak Taru

**Group:** Pair M

[Tableau Public Link](#)

[Github Link](#)

[Video Presentation](#)



## ***Abstract***

The aim of this project is to create a dashboard for Iowa state employees to summarise Volume, Sales and Profit for Liquor sold in IOWA state. The focal point for this project is to provide users control to track rise and fall in Sales and profit by regions and cities across Iowa state which can help them to implement or modify existing policies.

## ***Data Collection***

We have collected the Iowa state liquor data from Iowa state government site (<https://data.iowa.gov/Sales-Distribution/Iowa-Liquor-Sales/m3tr-qhgy>). This data is provided by Iowa Department of Commerce, Alcoholic Beverages Division and can be downloaded in CSV format. This dataset contains every individual purchase of liquor in the state of Iowa from retailers since 2014. Each row from dataset contains the individual purchase of the liquor bottle with store number, store location, liquor type, volume, brand etc. The dataset contains almost **8.14 million records** before cleaning and pre-processing. Space utilisation of CSV file is 3.17 GB. This covers the **volume aspect of big data**. We are using this data set to analyse total spirits sales in Iowa of over the years from 2018 to October 2021.

## ***Data Exploration, pre-processing, Cleaning and Transformation***

### **Pre-processing and Exploration**

**Initial Data Exploration** and pre-processing was **performed in Python** with goal to explore fields, reformatting geographic fields and removing nulls. Initially, we explored data types of the field and checked for null values for each field and dropped rows which had null values for fields that we were going to analyse. Later we reformatted Geographical fields to desired format. This file was exported to be cleaned and transformed in Tableau Prep.

## Cleaning and Data transformation

For **cleaning and Data transformation** we used **Tableau Prep** due to its easy-to-use approach and ability to export data in hyper file.

### Steps Performed

- 1) Drop unwanted fields
- 2) Creating Regions field based on city field to divide Iowa state in geographic regions.
- 3) Creating calculated field to assign geographic roles to be used in Tableau Reporting.
- 4) Create calculated field to get State profit based on retail and sales field.
- 5) Data extraction for date range between 2018 and 2021
- 6) Export dataset in **Hyper File instead of CSV** as hyper file is in-memory data engine and **optimised for analytical query processing**.

Outcome of above steps was Hyper file to be used in Tableau for Reporting.

### Screenshot Tableau Prep

The screenshot shows the Tableau Prep interface. At the top, a workflow bar includes steps: cleaned\_data, remove index, assign roles to..., split store loca..., drop point field..., assign correct..., Date and Longi..., and Output. Below this, a table titled 'Save to etl\_output\_hyper.hyper' displays a preview of the data. The table has 12 columns: Invoicenum, Store Location Brackets, Longitude, Latitude, date, storenumber, storename, address, city, zipcode, countynumber, and countyname. The data rows show various store locations across Iowa, including Greenfield, Stuart, Corning, Waukon, and Centerville.

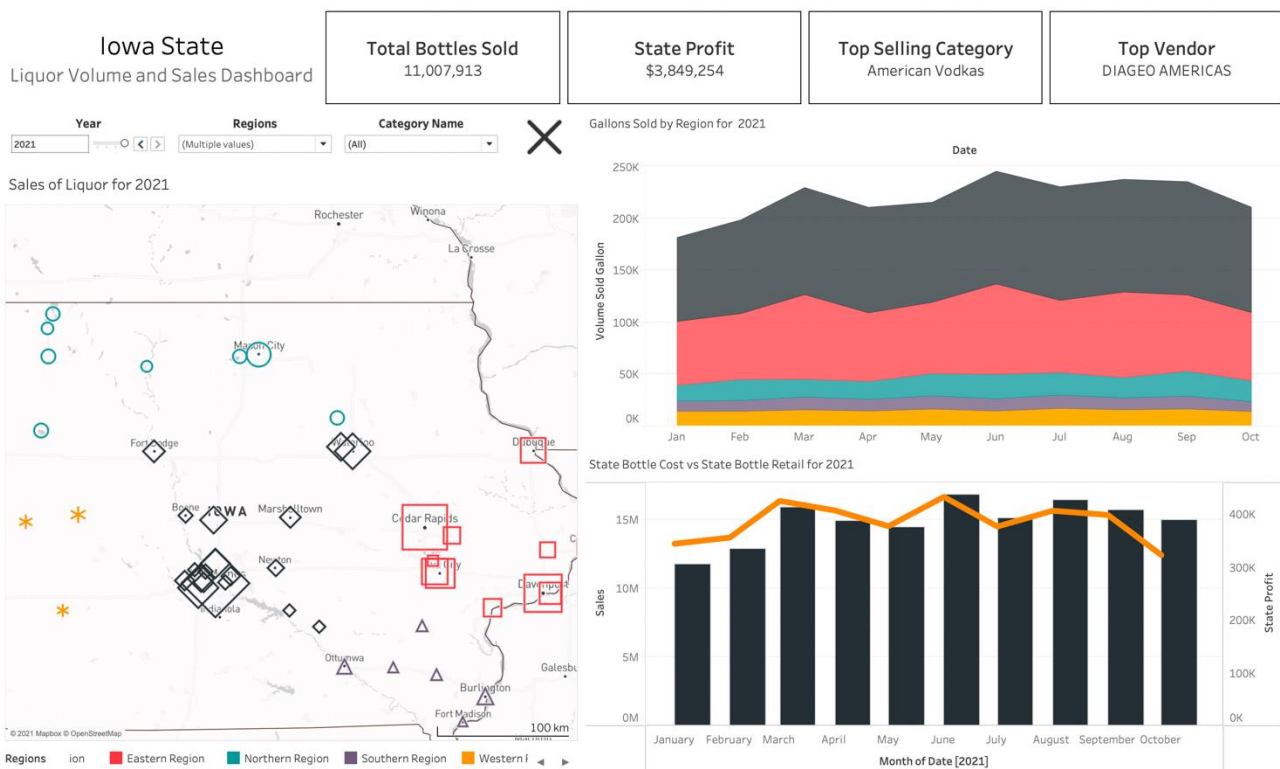
Invoicenum	Store Location Brackets	Longitude	Latitude	date	storenumber	storename	address	city	zipcode	countynumber	countyname
INV-21439800023	-94.462719 41.303167	-94.462719	41.303167	08/23/2019	4,317	Fareway Stores #941 / Greenfield	212 SW Kent	Greenfield	null	1	ADAIR
INV-21449700011	-94.318482 41.496773	-94.318482	41.496773	08/23/2019	4,654	KUM & GO #510 / STUART	629 S Division St	Stuart	null	1	ADAIR
INV-21438900019	-94.731809 40.991861	-94.731809	40.991861	08/23/2019	2,656	Hy-Vee Food Store / Corning	300 10th St	Corning	null	2	ADAMS
INV-21438900024	-94.731809 40.991861	-94.731809	40.991861	08/23/2019	2,656	Hy-Vee Food Store / Corning	300 10th St	Corning	null	2	ADAMS
INV-21438900004	-94.731809 40.991861	-94.731809	40.991861	08/23/2019	2,656	Hy-Vee Food Store / Corning	300 10th St	Corning	null	2	ADAMS
INV-21448700006	-91.484321 43.256623	-91.484321	43.256623	08/23/2019	4,289	Fareway Stores #062 / Waukon	777 11th Ave SW	Waukon	null	3	ALLAMAKEE
INV-21456300016	-92.86793 40.73112	-92.86793	40.73112	08/23/2019	4,271	Fareway Stores #827 / Centerville	305 S 18th St	Centerville	null	4	APPANOOSE

## Visualisation

For Dashboard Design reference, we followed **Visual Analysis Best Practices: A Guidebook** **whitepaper** from Tableau.

### Purpose

We aimed to create a dashboard for IOWA state employees to get all necessary information about Sales, Profit and Volume for Liquor sold in single view.



## Design Choices

### Creating effective Views

- 1) **Avoided overloading views:** We have tried to keep the dashboard as minimalistic as possible. Only relevant mark types were included to be displayed. Also, not all users will need filters to be displayed while printing or exporting dashboard as report, hence we have added close button to hide filters.
- 2) **Limit the number of colours and shapes in a single view:** We tried to limit the number of marks on the charts. Limited number of colours were chosen to keep the dashboard clean and minimalistic.
- 3) **Orient views for legibility:** Most of the dashboard labels are kept horizontally oriented to provide ease to users while reading.

### Colour Choices

- 1) **Primary Theme:** The main theme we chose for our dashboard is black and white with custom colour palate for regions. Symbol map displayed was created with custom colour theme with help of MapBox to match overall theme
- 2) **Colour Palette:** We used **custom colour palette** for our dashboard to create contrasting colours. We have considered colour blind audience while choosing

colours and used online tools to visualise our dashboard in **simulated colour-blind environment**. All the colours were distinguishable in Area Chart.

Screenshot to **simulate dashboard for colour blindness**



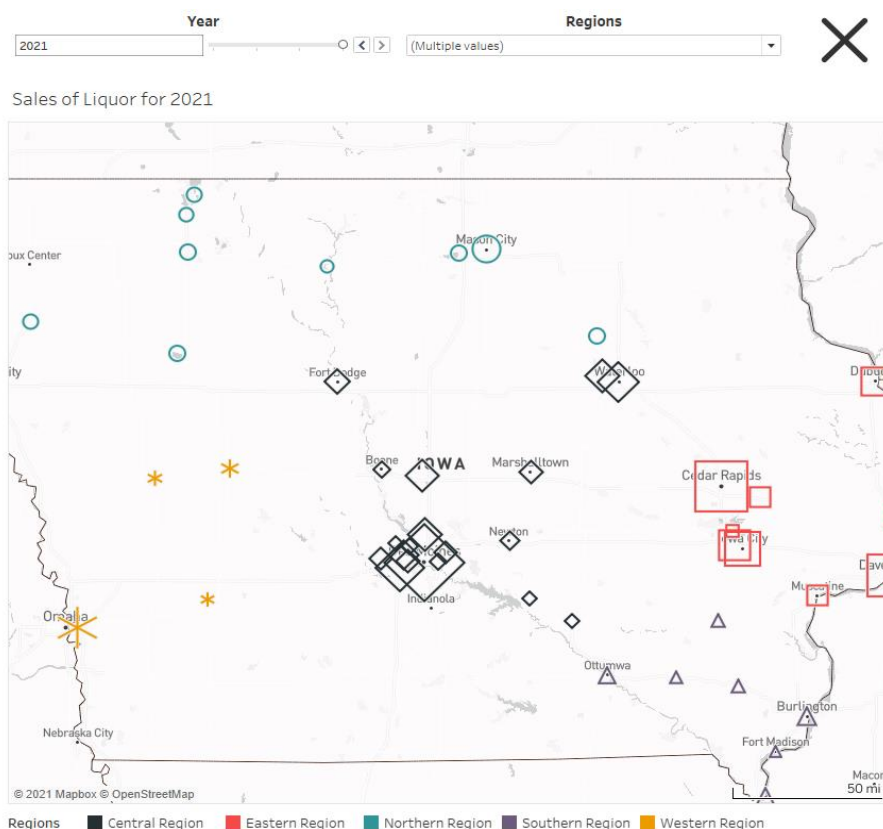
**Animation Choice:** We have added animations which is in sync with year slider filter to show smooth transitions and users can notice changes while applying filter. Also animations were added while filtering data when choosing city mark on map chart.

## Choosing Chart type

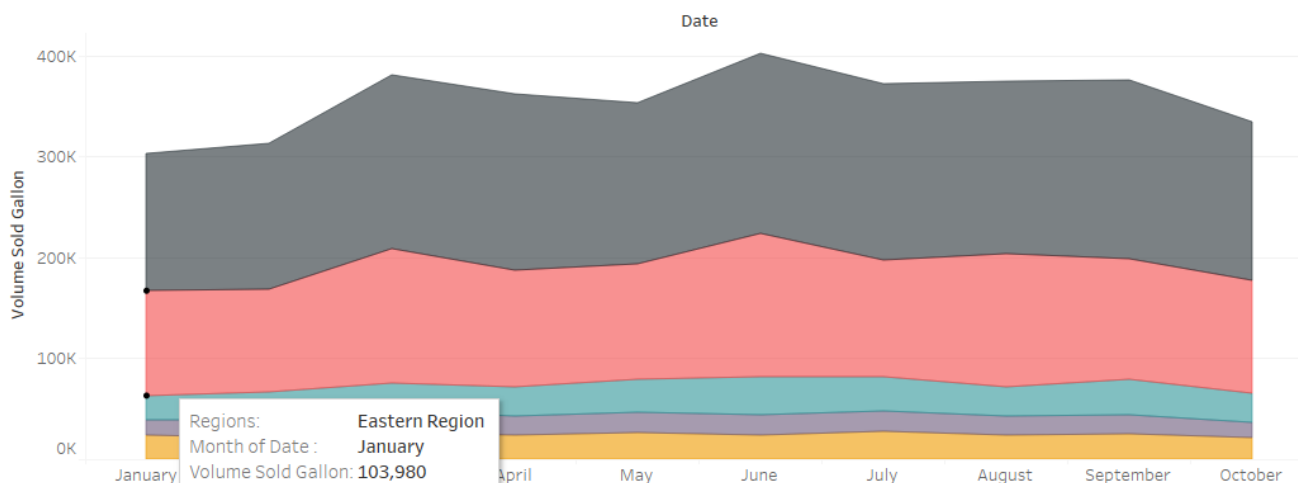
- 1) **Sales Symbol Map:** We created a Symbol Map to highlight sales based on region. **Marks** on chart **indicates cities** and size of marks grows and decrease based on sales made in city. **Custom designed map** was used which was created with help of MapBox. We also used **shape card** in Tableau to distinguish between different regions as it will be easier for **colour blind** audience. Colour choices were also made considering colour blind audience. To view sales by year we have provided **Year Filter** for user interaction along with **search button to filter by city**. Besides above we have also added **zoom in/out** ability to users **for cases of overlapping marks**.
- 2) **Area Chart:** We chose **area chart** to show progression of **volume of liquor sold** in gallons **over time**. We have used colour card to distinguish different regions in chart.

- 3) **Dual Axis chart:** To provide comparison of sales vs state profit over time, Dual axis was the best choice. Dual axis charts provide an easy-to-read chart and compares 2 measures effectively over time. X axis represents time progression for each year.
- 4) Additionally, we have also added **boxes that showcases important aspects such as top seller, bottles sold, total sales and top selling category.**

## Charts

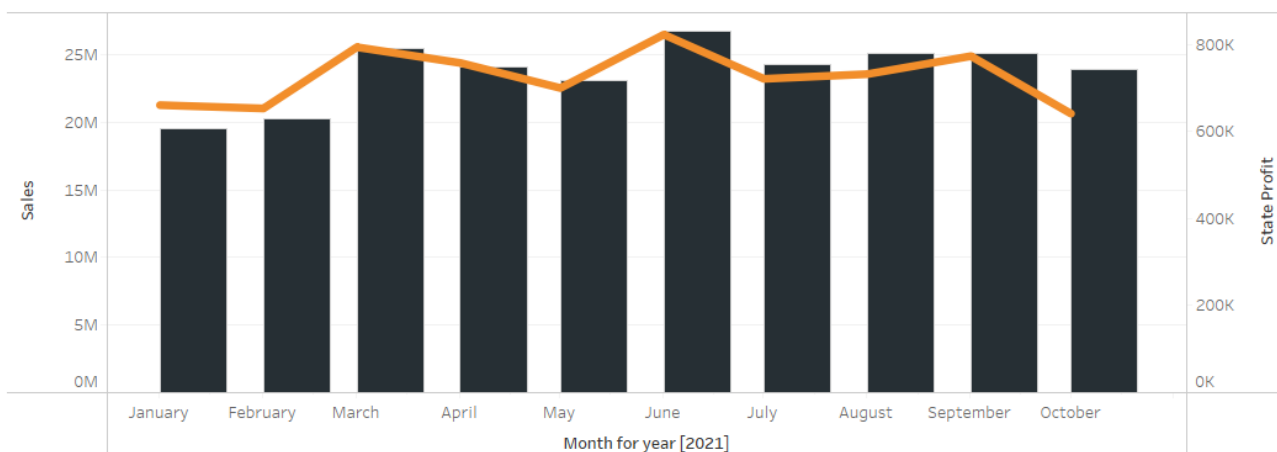


Gallons Sold by Region





State Bottle Cost vs State Bottle Retail for 2021



## Tools and Libraries used

**Data Pre-processing and Exploration:** We have used Python to perform initial pre-processing and Data exploration.

**Data Cleaning and Transformation:** Tableau prep was used for cleaning, transformation and filtering out data.

**Visualisation:** Tableau 2021.3 was used for creating dashboard and charts.

**Colour Blind simulation:** Pilestone.com was used to simulate dashboard for colour blindness.

**Custom Map:** MapBox was used to create custom Symbol Map to be integrated with Dashboard.

The whole dashboard is designed in Tableau Desktop and published to Tableau server

Dashboard Link: Tableau Public

<https://public.tableau.com/app/profile/shubham.rai7724/viz/IowaSalesConsumptionDashboard/Dashboard?publish=yes>

## Conclusion

Dashboard was created with aim to provide IOWA state employees to glance through sales and volume sold figures by year. Most of the time was spent in colour choices and dashboard creation. The charts are simple, easy to read, and interactivity is added in each chart with addition filters. The most interesting observation in sales chart was although each year October had highest sales of Liquor, but state profit did not grow in ratio with sales which is case with other months. We can also see gradual increase in sales and volume of alcohol sold from January until December. Most probably the sales of liquor are high due to holiday season. Highest revenue generation is from Central Region of IOWA state and sales of alcohol is highest in capital city of Des Moines. Top Vendor in IOWA state in all regions is Diageo Americas. For future development, A YOY and MOM chart can be added to get better comparison between dates.

## References

[Tableau Visual Guidebook whitepaper](#)

*For integration of custom maps style we followed following tutorial*  
[MapBox style in Tableau](#)

## Important Links

[Tableau Public Dashboard Link](#)

[Github Link](#)

[Video Presentation](#)