

# Proposal For Starting Liquor Business

Group 72: Rainy Chen 500856, Yvie Lin  
491764, Lanlin Su 487930, Tian Liu  
502454, Huina Cao 500775



# Agenda

**1. Data Description**

**2. Why is this big data?**

**3. Problem Statement**

**4. Methodology**

**5. Result Analysis**

**6. Conclusion**

# 01 Description

Information about our  
dataset.





## Context & Content

- Spirits purchase information since 2012-2017
- 25 columns in total
- Collected by The Iowa Department of Commerce in 2017
- Source: Kaggle

# 02

## Why big data?

Characteristics that prove this dataset is an example of big data.



# Six Dimensions of big data



## Volume

12591077 rows and  
3.47 GB



## Velocity

Every row created if  
one order takes place



## Variety

Numerical data and  
categorical data



## Value

Help make better  
decisions for company



## Veracity

Published by the State  
of Iowa



## Tools

Bash server,  
Hive/Spark/Impala



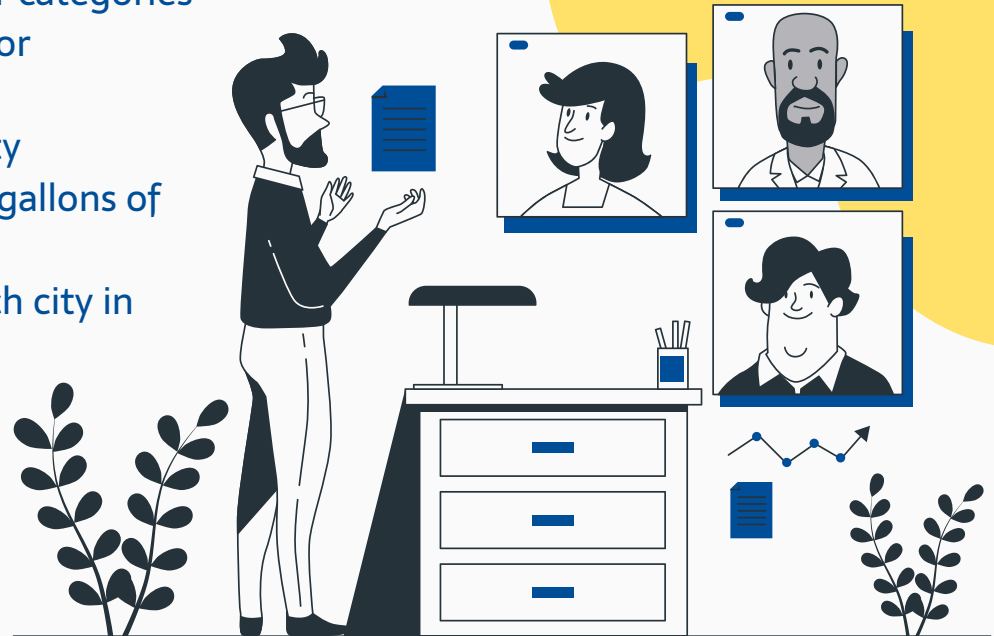
# 03 Problem Statement

Problem we want to solve  
and questions we create.

**“To optimize the storage and distribution problem, elevate the R&D production strategy and earn higher return, we provide region-specific recommendations for liquor entrepreneurs.”**



1. Top 5 cities sold the most each year
2. Highest concentrated sales area on map
3. Best sellers (vendors)
4. Monthly top 10 consumed liquor categories
5. Monthly least 10 consumed liquor categories
6. Liquor stores amount in each city
7. Top 10 stores that sell the most gallons of liquor each month
8. Top 10 consumed liquors for each city in terms of profit



**04**

## **Methodology**

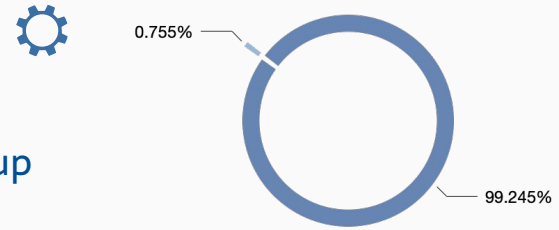
The methods we choose to process the data.



## 3 Steps We Use to Deal with Data

# Preprocessing

- Download the raw data from Kaggle
- Use python to standardize the data types, and clean up the rows where the missing values are located



## Analysis

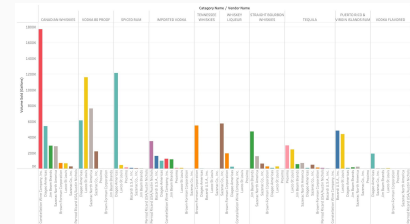
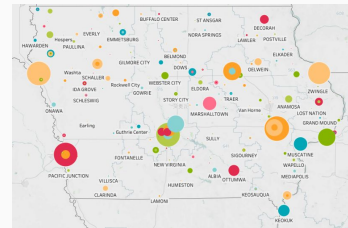
- Upload the data to the server
- Copy the data to HDFS
- Use Mapreduce and Hive to analyze the data



queries

## Visualization

- Use Tableau to visualize our results





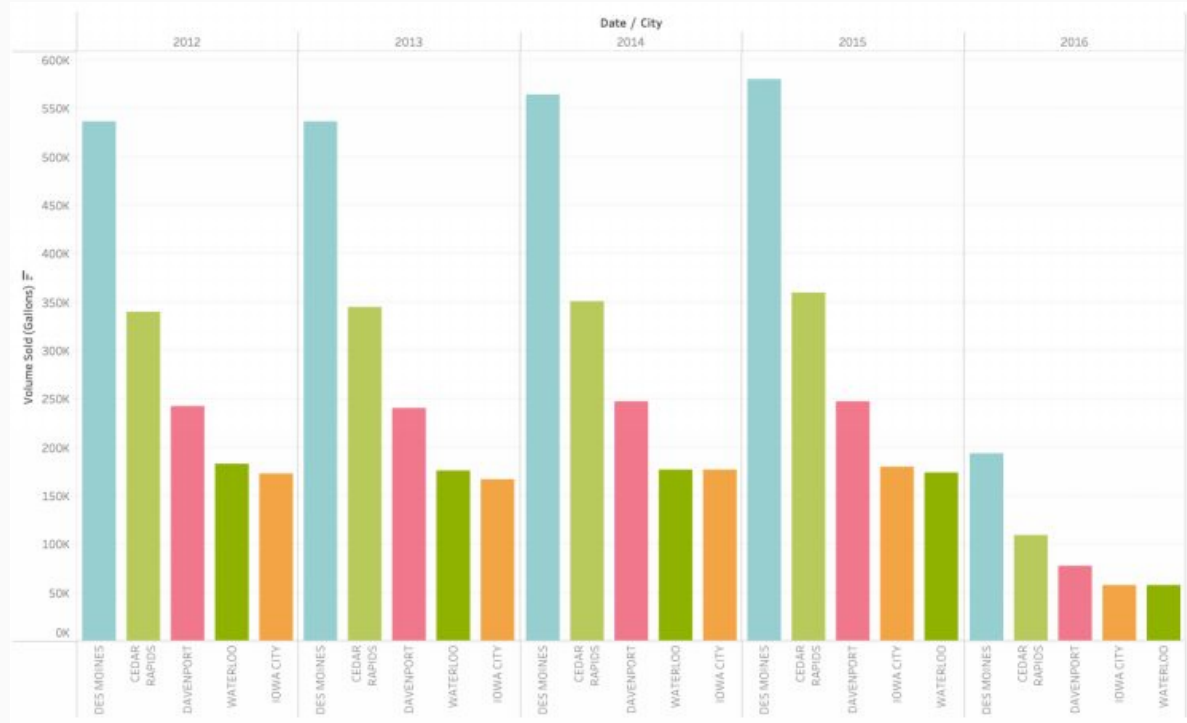
**05**

## **Result Analysis**

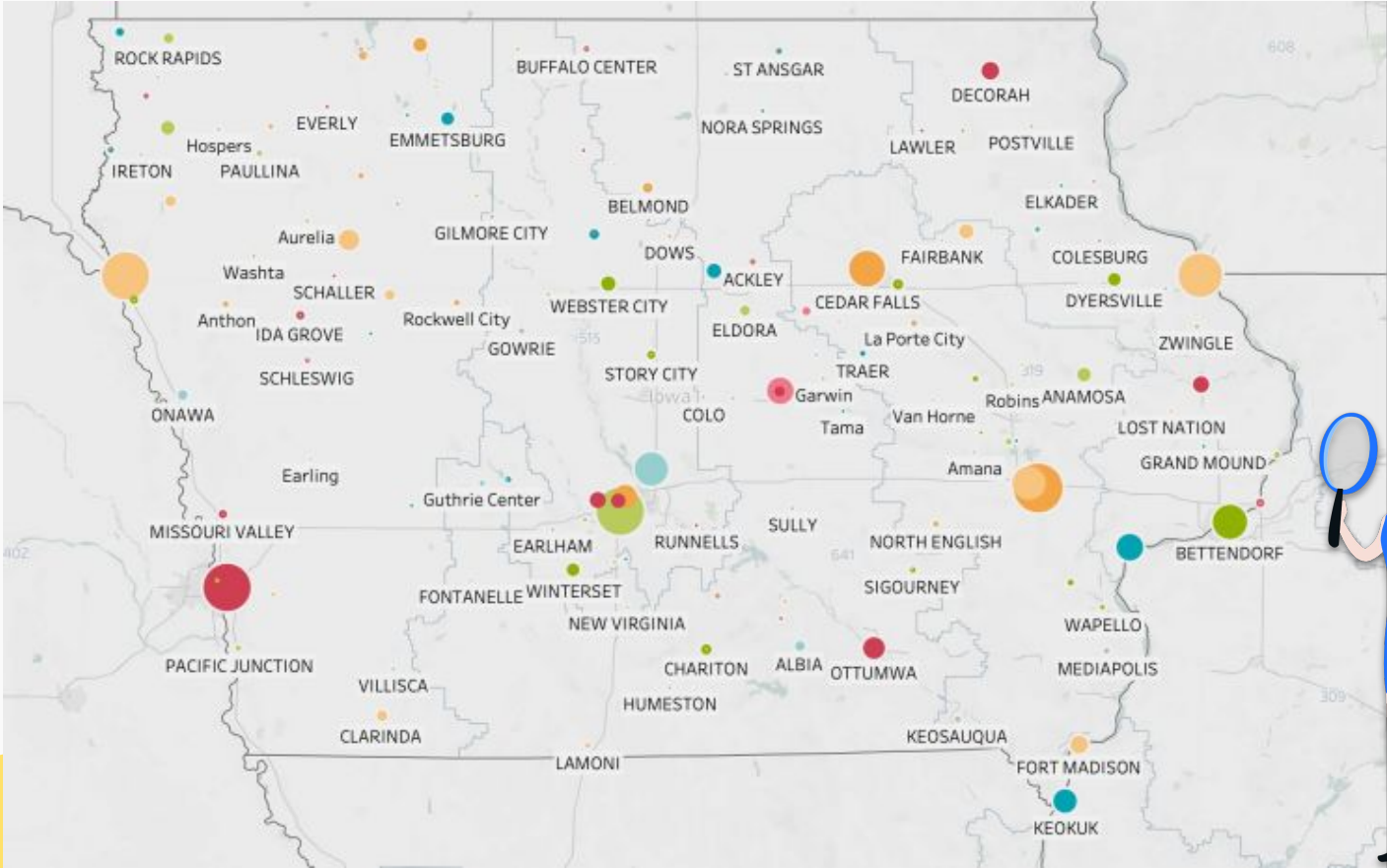
What we get from the Data analysis and how can we put it in business strategy

# #1. The steady Top 5 cities sold the most each year

Des moines,  
Cedar rapids,  
Davenport,  
Iowa city,  
Waterloo



## #2. Visualization Map of Distribution of Cities in term of volume sold





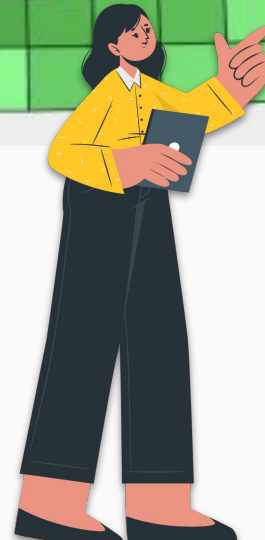
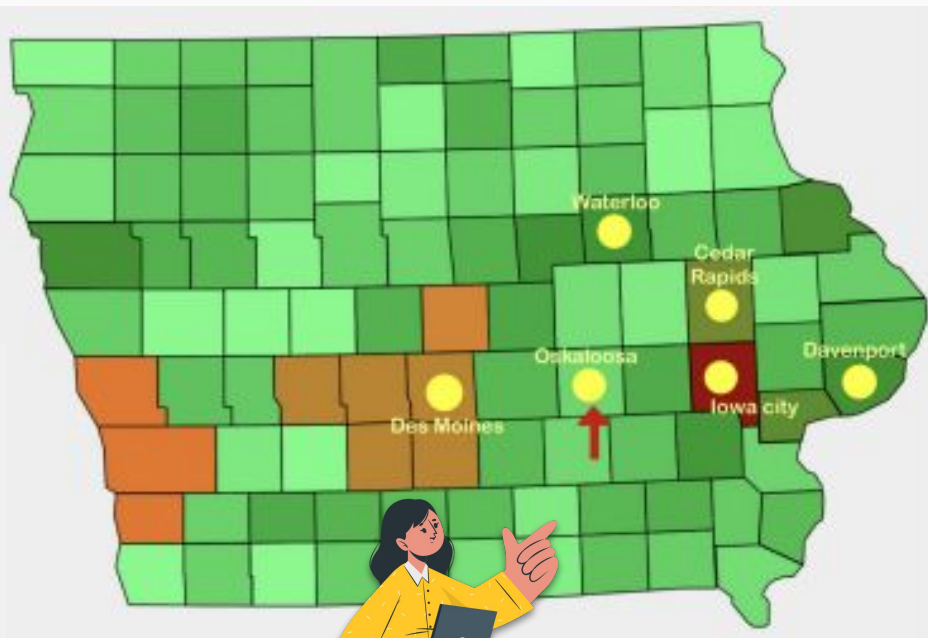
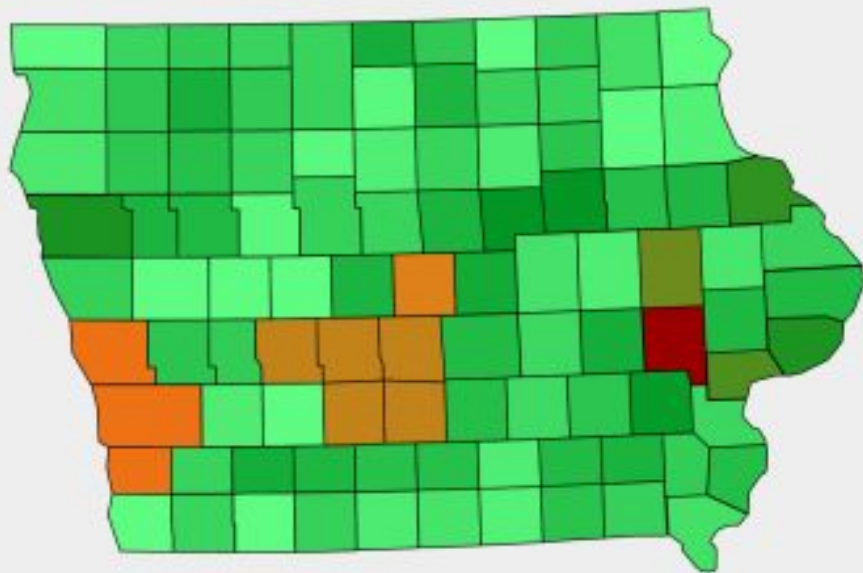
# Top 5 cities sold the most each year

Des moines  
Cedar rapids  
Davenport  
Iowa city  
Waterloo



# Iowa Fair Market Rent Map

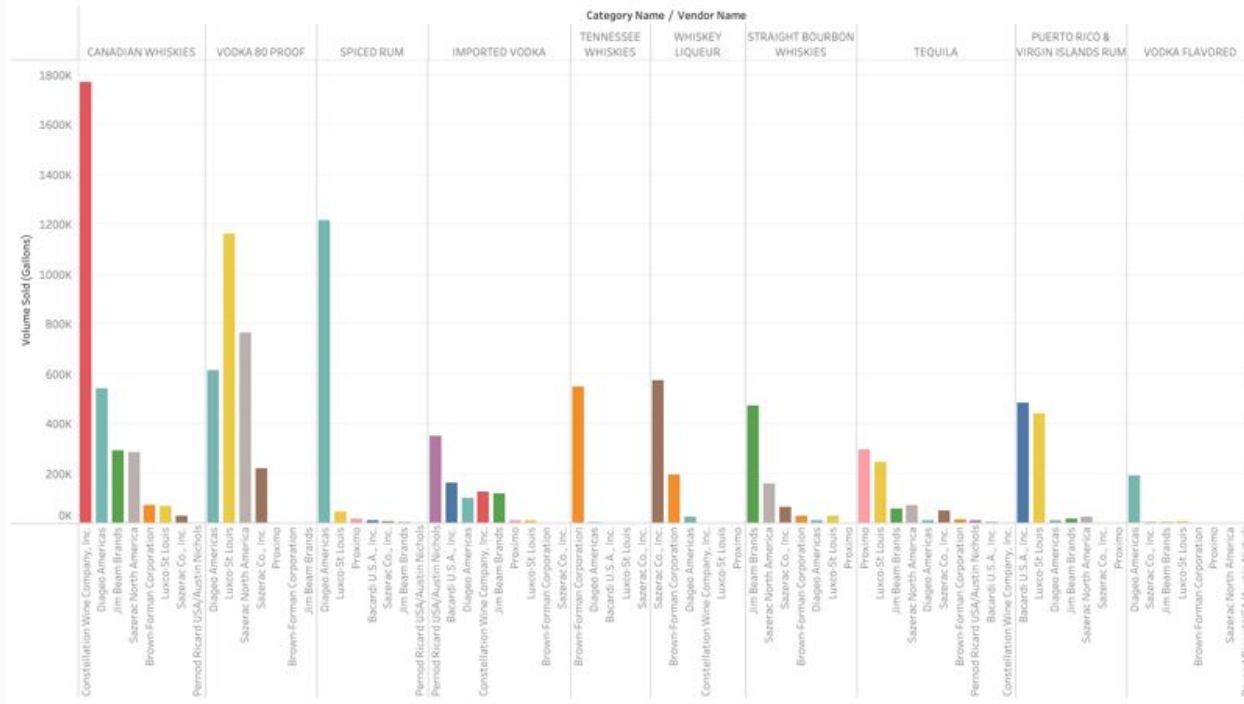
The shade of color indicates the  
price of rent



**Oskaloosa**



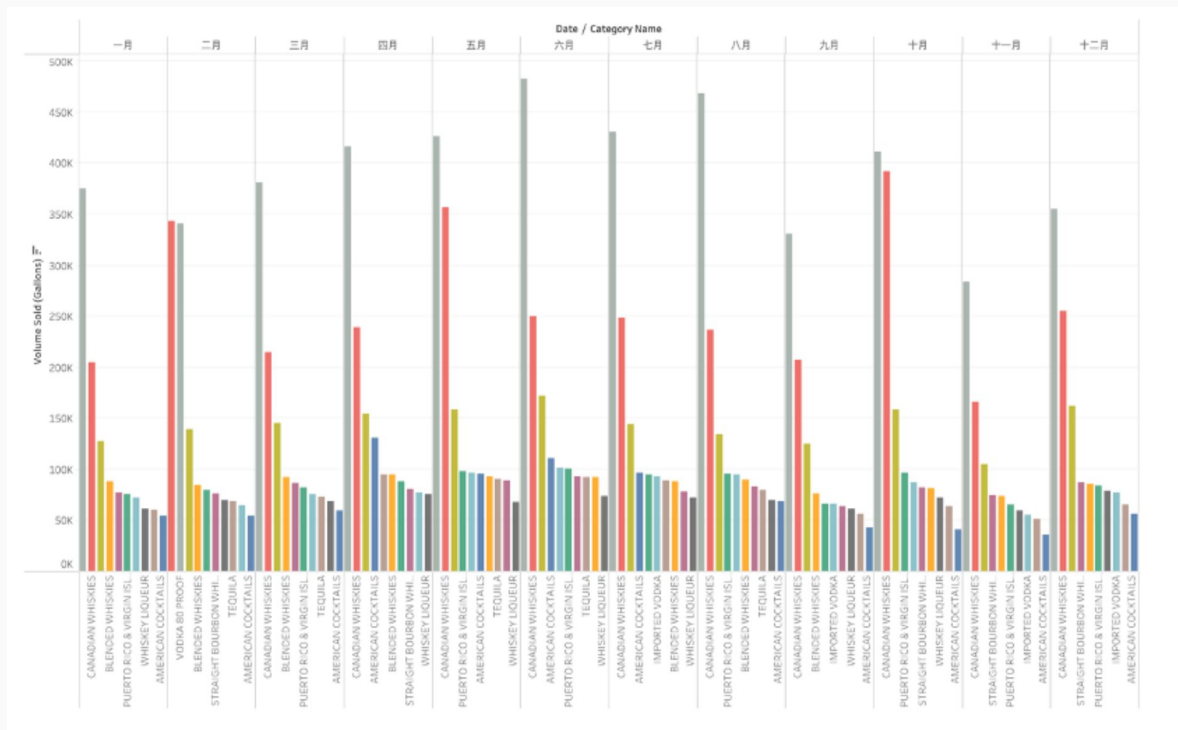
# #3. Potential competitor



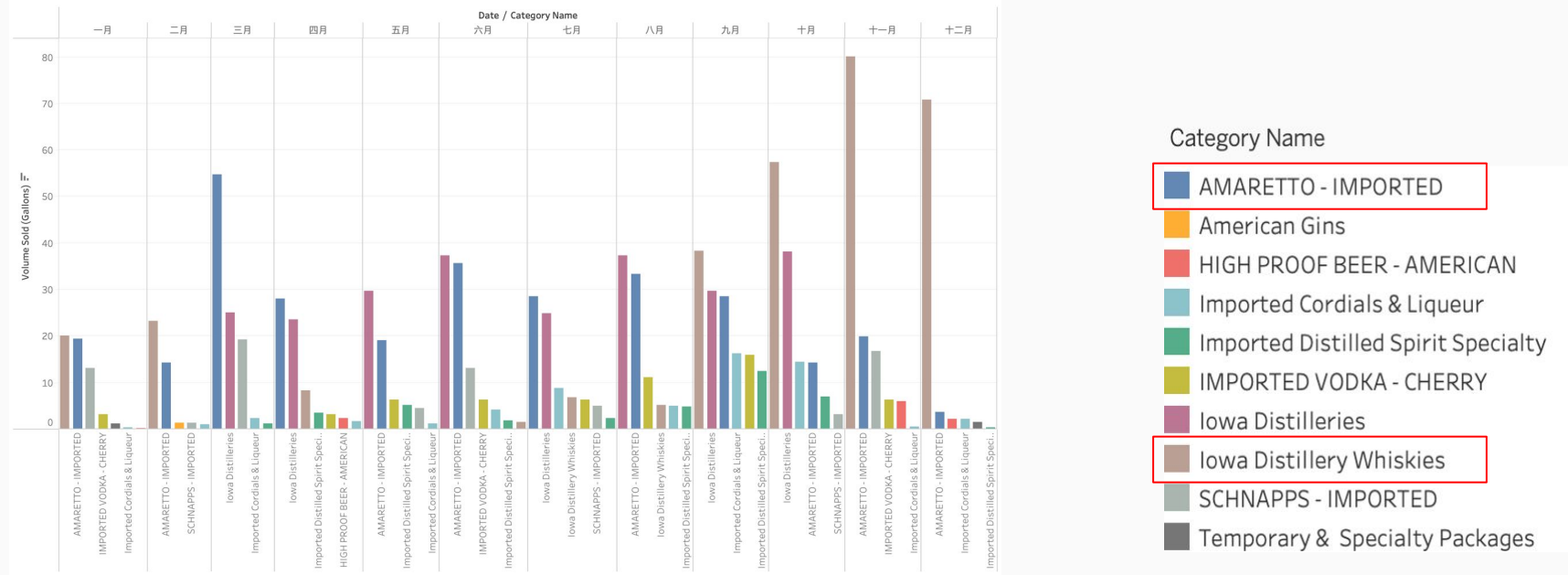
## Vendor Name

- Bacardi U.S.A., Inc.
- Brown-Forman Corporation
- Constellation Wine Company, Inc.
- Diageo Americas
- Jim Beam Brands
- Luxco-St Louis
- Pernod Ricard USA/Austin Nichols
- Proximo
- Sazerac Co., Inc.
- Sazerac North America

# Top 10 Popular Consumed Liquors for Each Month



# Least 10 Popular Consumed Liquors for Each Month

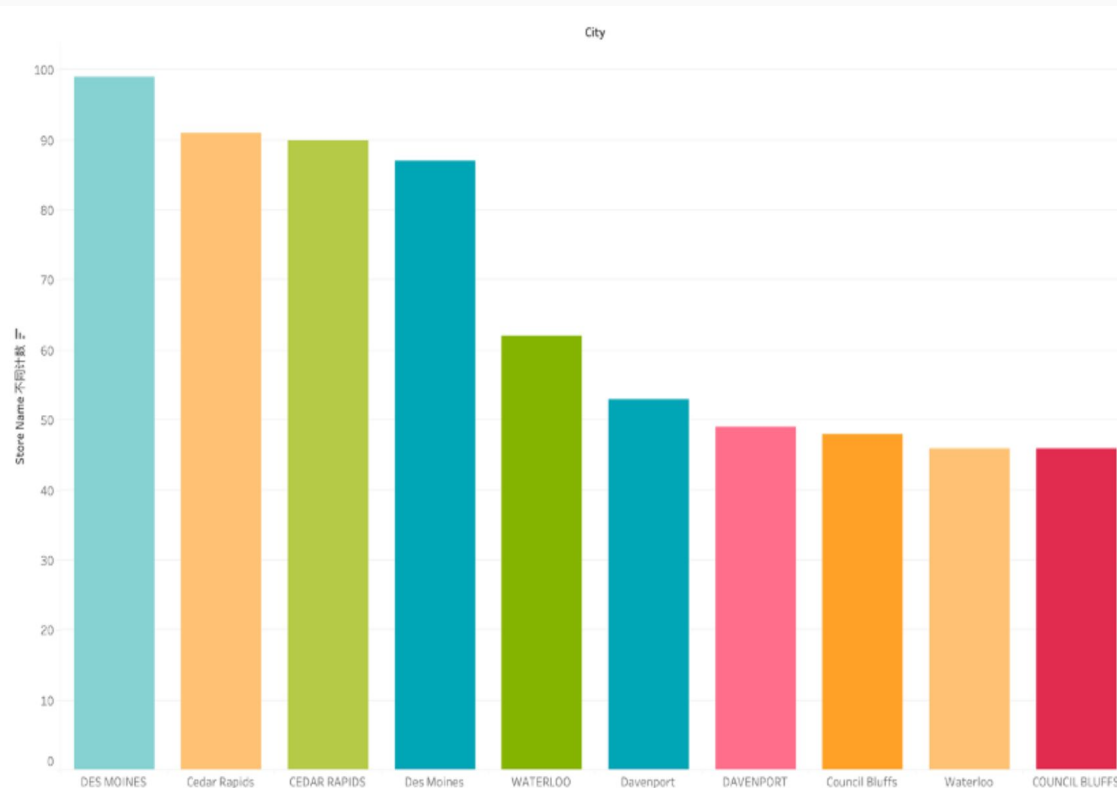
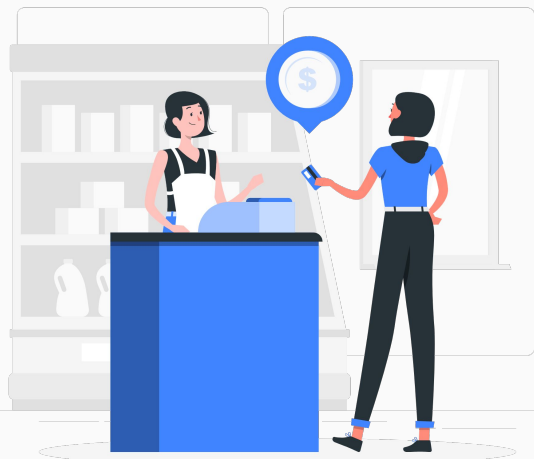


# Number of Stores for Each City

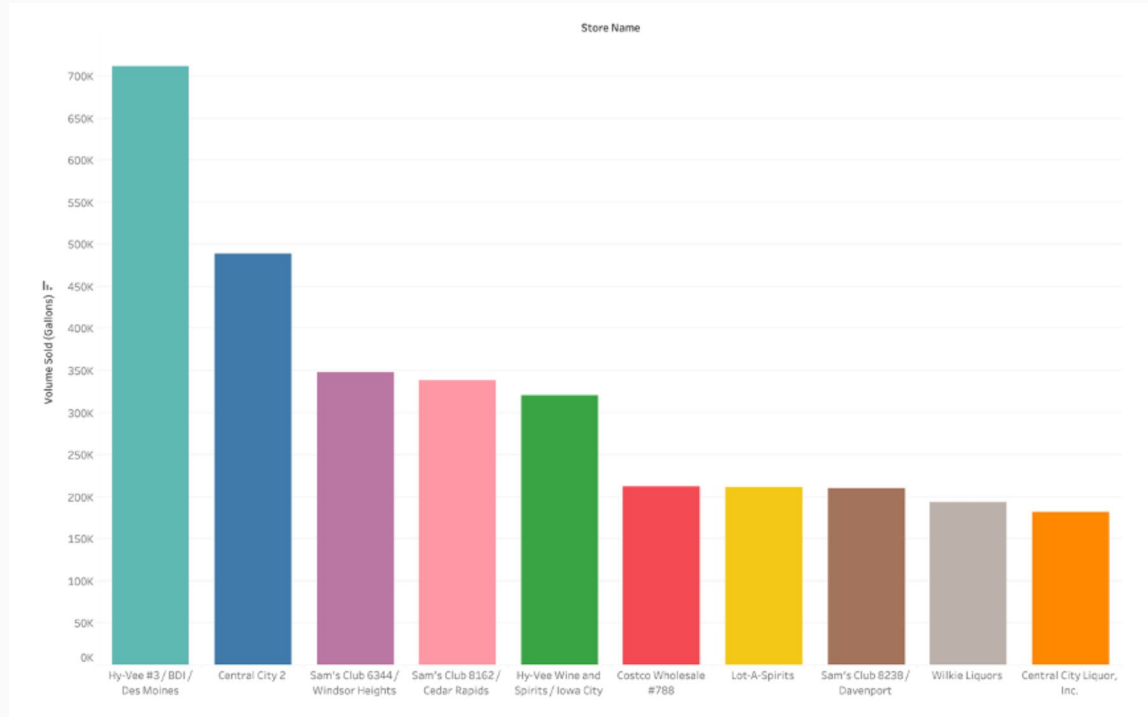
DES MOINES

Cedar Rapids

WATERLOO



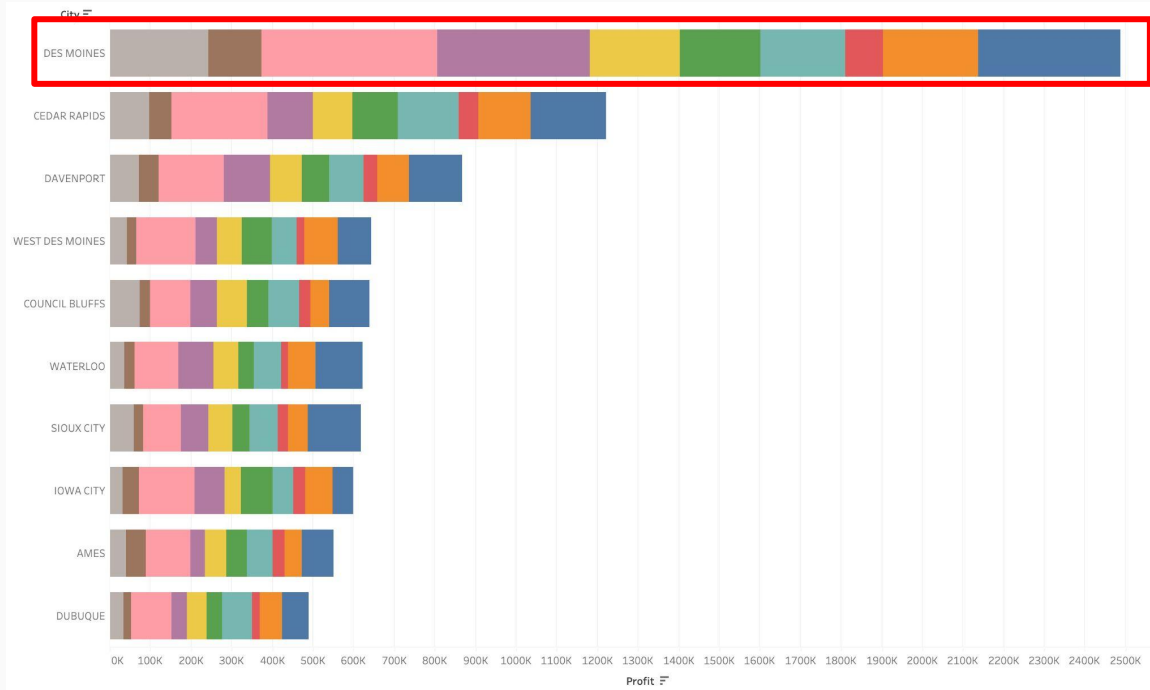
# Top 10 Stores that Sell the Most Gallons of Liquor



## Store Name

- Central City 2
- Central City Liquor, Inc.
- Costco Wholesale #788
- Hy-Vee #3 / BDI / Des Moines
- Hy-Vee Wine and Spirits / Iowa City
- Lot-A-Spirits
- Sam's Club 6344 / Windsor Heights
- Sam's Club 8162 / Cedar Rapids
- Sam's Club 8238 / Davenport
- Wilkie Liquors

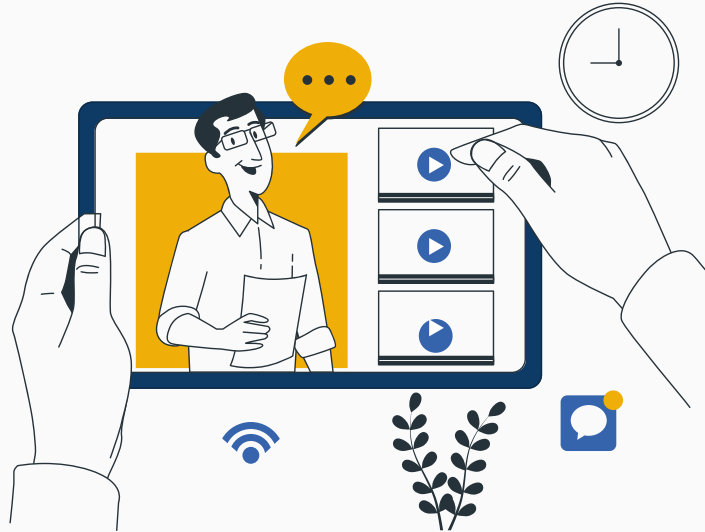
# Trend lines of Top 10 Popular Consumed Liquors for Each City



- CANADIAN WHISKIES
- IMPORTED VODKA
- SPICED RUM
- VODKA 80 PROOF
- VODKA FLAVORED

# 06 Conclusion

You can enter a subtitle  
here if you need it



To make better decisions for the new company to maximize profit, there are **various** factors to be considered, such as: location, popularity of brands, seasons, competitors and so on.

**-recommendation-**



# Factors to be considered for the company:



## Location

Selecting the location of the warehouse in Des Moines can minimize transportation costs.



## Cities

The right part of the state have the best sellers cities. Also, Des Moines and WATERLOO could be the target markets.



## Seasons/Brands

Entrepreneurs can make more production on VODKA 80 PROOF, CANADIAN WHISKIES and SPICED RUM. We recommend increasing the VODKA 80 PROOF in stock in June.



## Competitors

The new company could make marketing searches to the top and least popular competitors, analyze the reason why they are popular/unpopular in such specific seasons/categories/cities



# Q&A

Thank you for listening!  
Any questions?