Databases Final Project

Sayuri Monarrez Yesaki(sdm3465)
Anson Cooper Fryar(acf2522)

Data Sets

Iowa Alcohol Sales

The original dataset contains the spirits purchase information of Iowa Class "E" liquor licensees by product and date of purchase from January 1, 2012 to current. We used PostgreSQL to separate years 2014 and 2015 from the original data into their own tables. After that, we ordered the data by County, Item Number and Sale Dollars then added a Column named "TempKey" to both 2014 and 2015 Sales table to add a primary key.

Driver License Revocations

This dataset contains data about the number of driver license revocations for operating while intoxicated from 2000 to 2015.

Alcohol Impaired Driving Deaths

This dataset contains data about motor vehicle crash deaths with alcohol involvement.

These three datasets are able to be connected through the year 2014, 2015 and Counties.

Data Description

Iowa Gov

2014/2015 Stores (2 Tables)
- Temp Key
- Store
- Name
- County
- Address
- Zip Code

2014/2015 Sales (2 Tables)
- Store
- County
- Item
- Liters
- Bottle_Qty
- Total

Country Health Rankings

2014
- County
- # Alcohol-Impaired Deaths
- # Driving Deaths
- % Alcohol Impaired
- Z-Score

2015
- County
- # Alcohol-Impaired Deaths
- # Driving Deaths
- % Alcohol Impaired
- Z-Score

Des Moines Register

Year By Year Table
- County
- 2000 -> 2015 (16 Different Columns)

Purpose

- ❏ We will be trying to see the relationship between the purchases of alcohol, driver license revocations for operating while intoxicated and motor vehicle crash deaths with alcohol involvement, in Iowa separated and compared by their counties.

Big Questions

- ❏ Do the ratios between Population, Alcohol Sales, and Motor Vehicle Crash Deaths with alcohol involvement per county differ by a wide range or are they consistent throughout the state of Iowa?

- ❏ Are the ratios between Vehicle Crash Deaths with alcohol involvement and driver license revocations per county differ by a wide range or are they consistent throughout the state of Iowa?

- ❏ Are there differences between the years 2014 and 2015?

- ❏ Is there a correlation between alcohol sale and driving deaths?

- ❏ Is there a correlation between alcohol sale and driver license revocation?

Solutions: Beam Transforms

Alcohol Impaired Driving Deaths:
- Normalization of County names
- Edit table names

Driver License Revocations:
- Normalization of County names and Revocations.
- Select only 2014 and 2015 revocations per county.
- Remove "NA" and "Non-Resident" values.

Alcohol Sales:
- Normalization of County names
- Group Sales by County.

Source code: https://github.com/cs327e-fall2018/Data-Chefs/blob/master/ParDo_AlcoholImpairedDrivingDeaths14_cluster.py
https://github.com/cs327e-fall2018/Data-Chefs/blob/master/ParDo_DriverLicenseRevocations_cluster.py
https://github.com/cs327e-fall2018/Data-Chefs/blob/master/Summary_AlcoholSales14_cluster.py

ERD
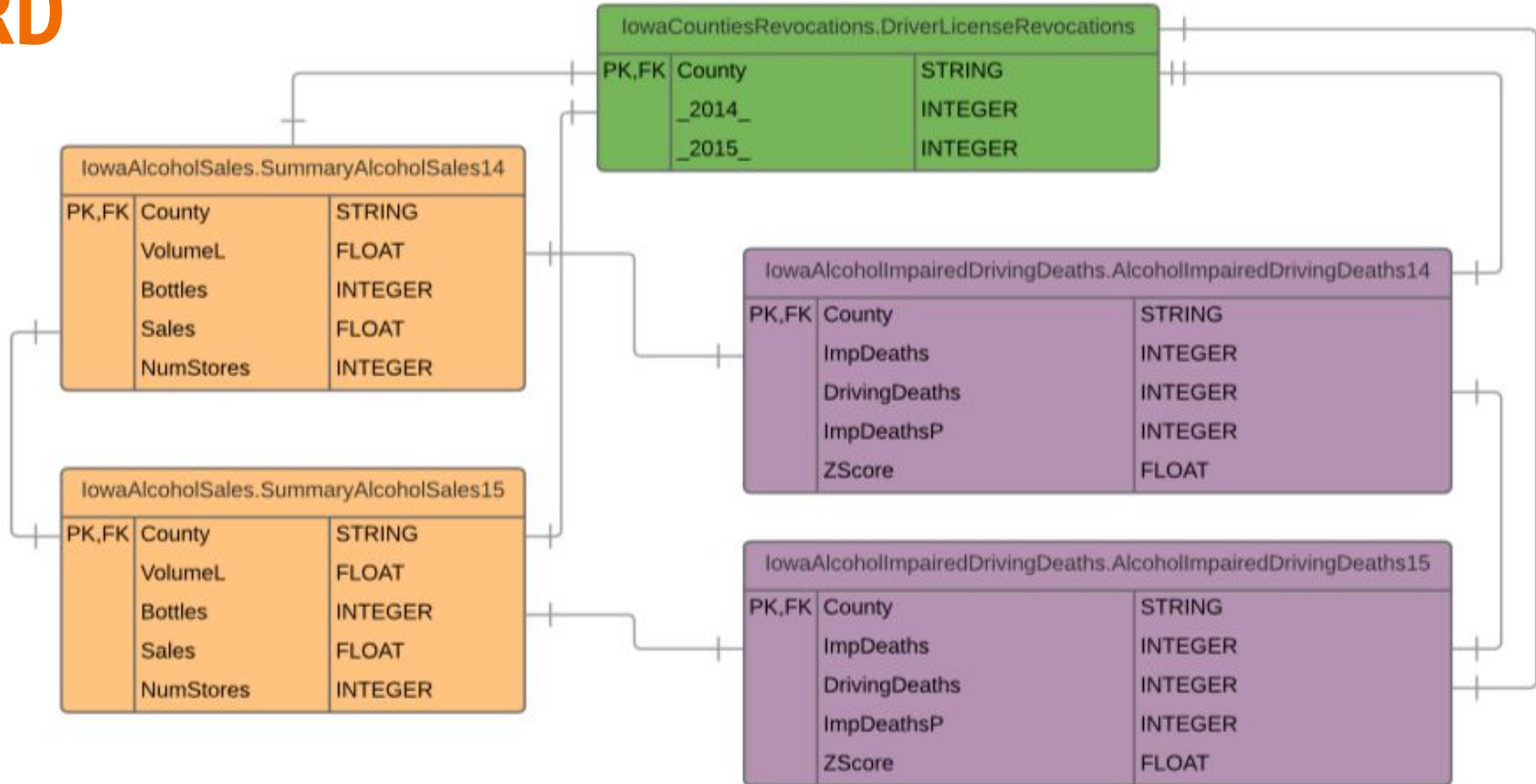
Alcohol was involved in more than 35% of motor vehicle fatalities in 14% Counties in 2014 and 18% Counties in 2015 in Iowa.

Counties in Iowa with more than 35% Alcohol Impaired Driving Deaths in 2014

	County	Impaire...	Vol Alcoho...	Num Stores	Num Revocat...
1.	CRAWFORD	35	11,261.76	5	82
2.	ALLAMAKEE	37	3,591.47	8	30
3.	PLYMOUTH	37	6,602.94	8	129
4.	WOODBURY	37	7,305.58	38	781
5.	KOSSUTH	38	7,330.72	8	32
6.	DAVIS	38	2,735	1	23
7.	GUTHRIE	38	2,117.49	4	49
8.	MADISON	40	2,997.88	7	26
9.	GREENE	42	2,116.36	6	40
10.	JACKSON	50	3,158.09	9	86
11.	LOUISA	50	1,370.07	5	58
12.	HUMBOLDT	56	6,902.4	4	23
13.	POCAHONTAS	67	2,095.31	7	54

1 - 13 / 13 ‹ ›

Counties in Iowa with more than 35% Alcohol Impaired Driving Deaths in 2015

	county	Imp Driving ...	Vol Alcohol ...	Num Stor...	Num Revoca...
1.	PLYMOUTH	35	5,567.96	8	132
2.	HANCOCK	36	1,940.72	4	34
3.	FREMONT	36	719.4	1	37
4.	WRIGHT	38	1,720.47	10	73
5.	ALLAMAKEE	38	3,955.61	6	43
6.	GREENE	39	3,734.82	6	35
7.	WOODBURY	39	6,510.61	38	809
8.	LYON	40	1,785.94	10	42
9.	BUCHANAN	44	3,891.04	10	97
10.	SHELBY	45	8,557.28	4	33
11.	JACKSON	47	3,094.09	8	75
12.	DAVIS	50	2,247.75	1	19
13.	TAMA	50	3,952.69	5	79
14.	LOUISA	53	1,058.18	5	40
15.	HUMBOLDT	55	5,279.1	4	23
16.	POCAHONTAS	67	2,102.18	5	52

1 - 16 / 16 ‹ ›

Correlations

ρ : Correlation between Volume of Alcohol Sold per Store and Alcohol Impaired Driving Deaths per County in Iowa 2014
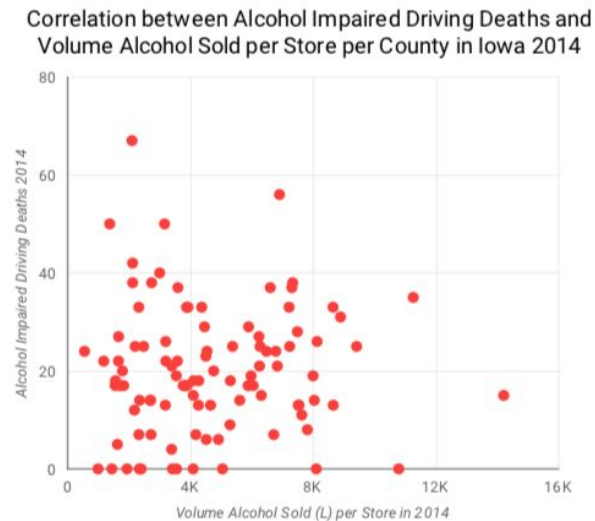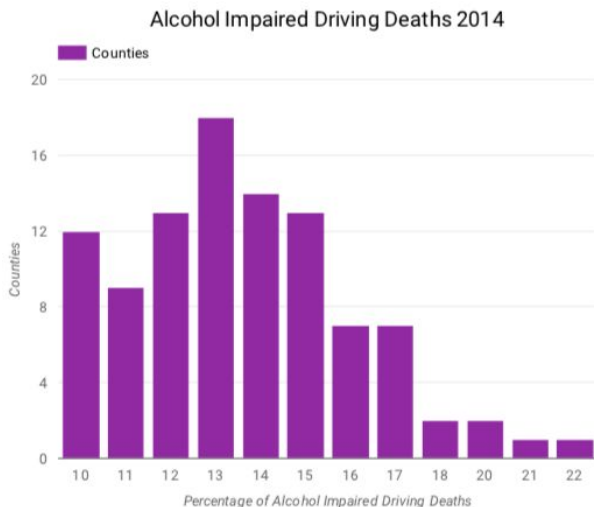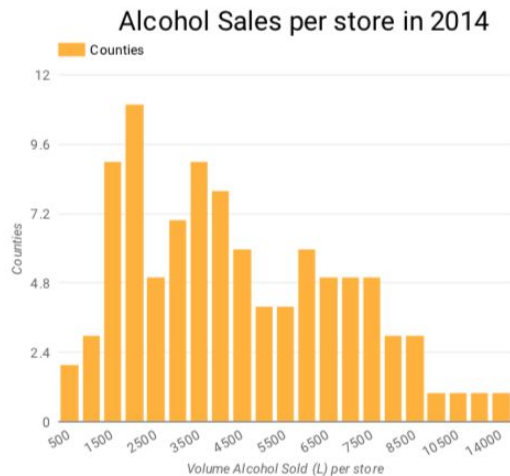
Hypothesis:
Ho: ρ = 0
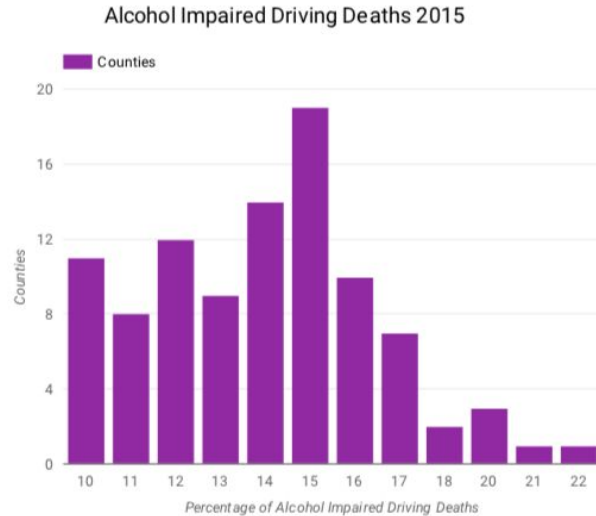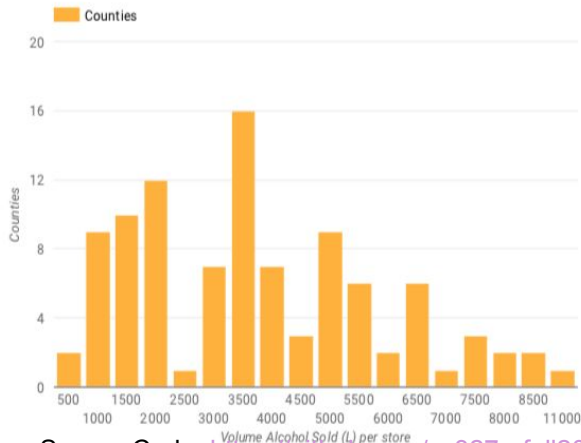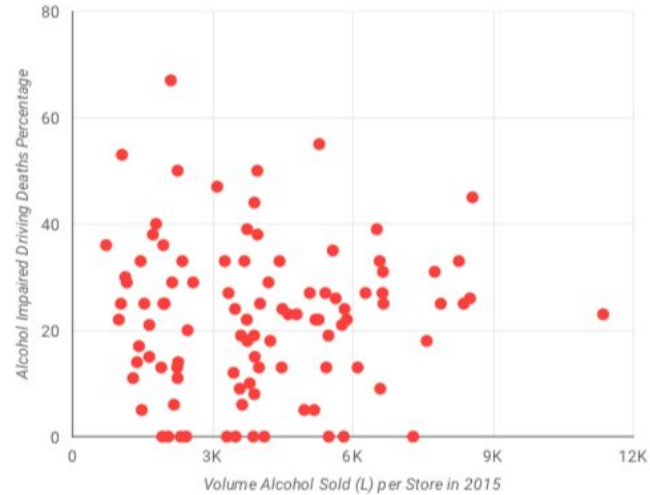Ha: ρ != 0

Results:
ρ = 0.0607

Correlation between Alcohol Impaired Driving Deaths and Volume Alcohol Sold per Store per County in Iowa 2014

Alcohol Sales per store in 2014

Alcohol Impaired Driving Deaths 2014

Source Code: https://github.com/cs327e-fall2018/Data-Chefs/blob/master/cross-dataset-queries.sql

ρ : Correlation between Volume of Alcohol Sold per Store and Alcohol Impaired Driving Deaths per County in Iowa 2015

Hypothesis:
Ho: ρ = 0
Ha: ρ != 0

Results:
ρ = 0.0274

Alcohol Sales per store in 2015

Alcohol Impaired Driving Deaths 2015

Correlation between Alcohol Impaired Driving Deaths and Volume Alcohol Sold per Store per County in Iowa 2015

Source Code: https://github.com/cs327e-fall2018/Data-Chefs/blob/master/cross-dataset-queries.sql

ρ : Correlation between Volume of Alcohol Sold and Driver License Revoked per County in Iowa 2014
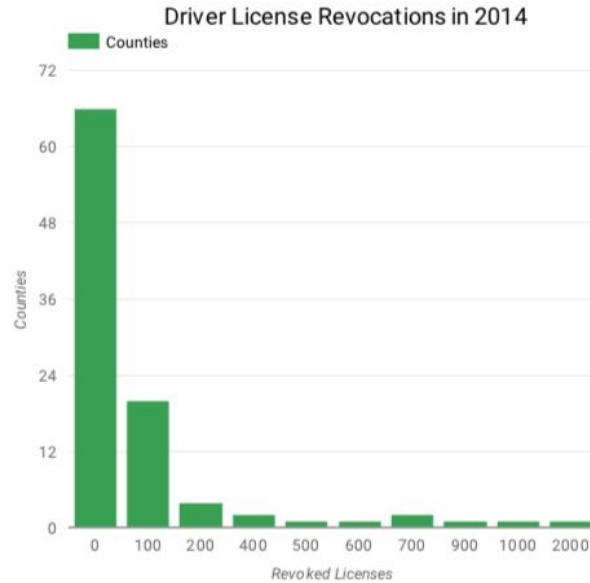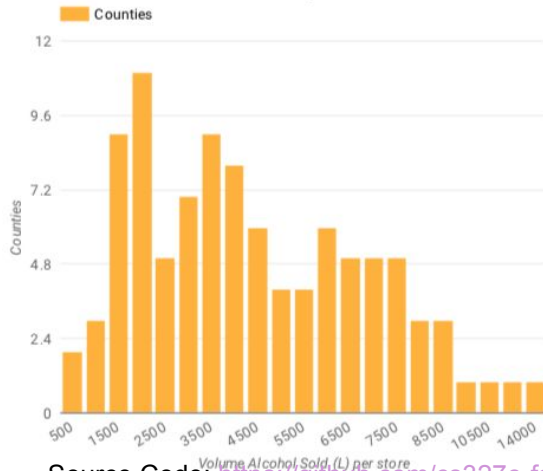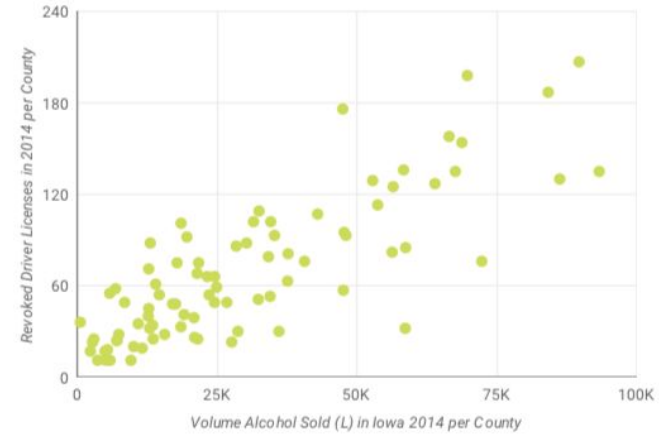
Hypothesis:
Ho: ρ = 0
Ha: ρ != 0

Results:
ρ = 0.8654

Correlation between Revoked Driver Licenses and Volume Alcohol Sold in Iowa 2014 per County

Alcohol Sales per store in 2014

Driver License Revocations in 2014

Source Code: https://github.com/cs327e-fall2018/Data-Chefs/blob/master/cross-dataset-queries.sql

ρ : Correlation between Volume of Alcohol Sold and Driver License Revoked per County in Iowa 2015
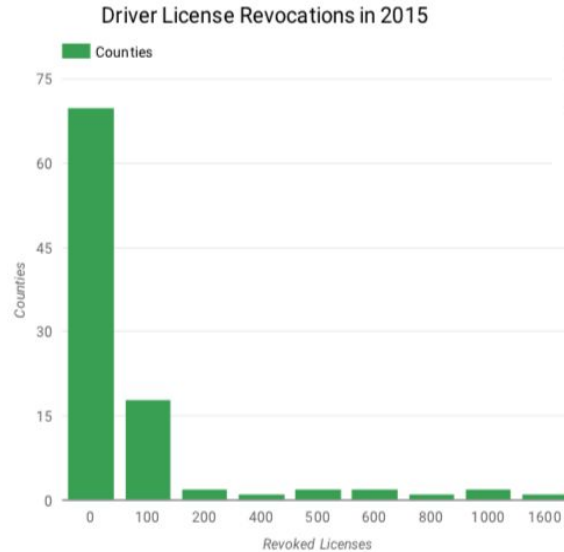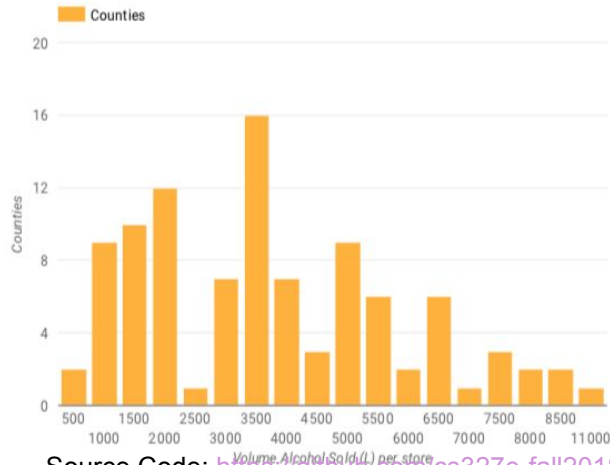
Hypothesis:
Ho: ρ = 0
Ha: ρ != 0

Results:

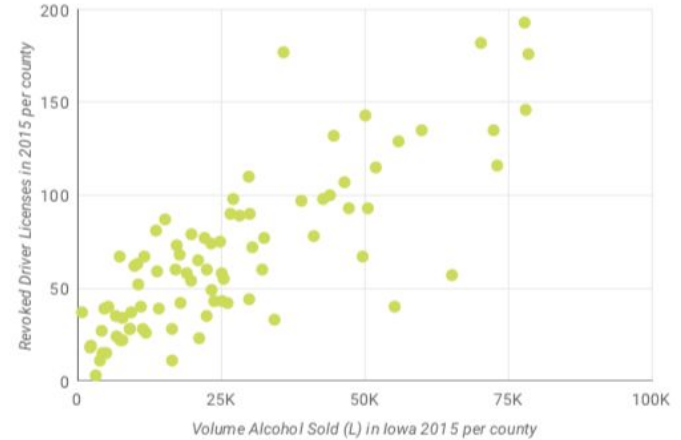ρ = 0.8621

Correlation between Revoked Driver Licenses and Volume Alcohol Sold in Iowa 2015 per County

Alcohol Sales per store in 2015

Driver License Revocations in 2015

Source Code: https://github.com/cs327e-fall2018/Data-Chefs/blob/master/cross-dataset-queries.sql

IN THE FUTURE...

- ❏ More Specific Data Points
  - ❏ County Population

- ❏ More Years

- ❏ Different Data Points
  - ❏ Ethnicity/Religion
  - ❏ Age Ranges
  - ❏ Different Data Points

REFERENCES

https://data.iowa.gov/Economy/Iowa-Liquor-Sales/m3tr-qhgy

http://db.desmoinesregister.com/owi-revocations-by-county

http://www.countyhealthrankings.org/app/iowa/2014/measure/factors/134/data

http://sites.utexas.edu/sos/guided/inferential/numeric/bivariate/rankcor/