The sales situation of Liquors in different regions

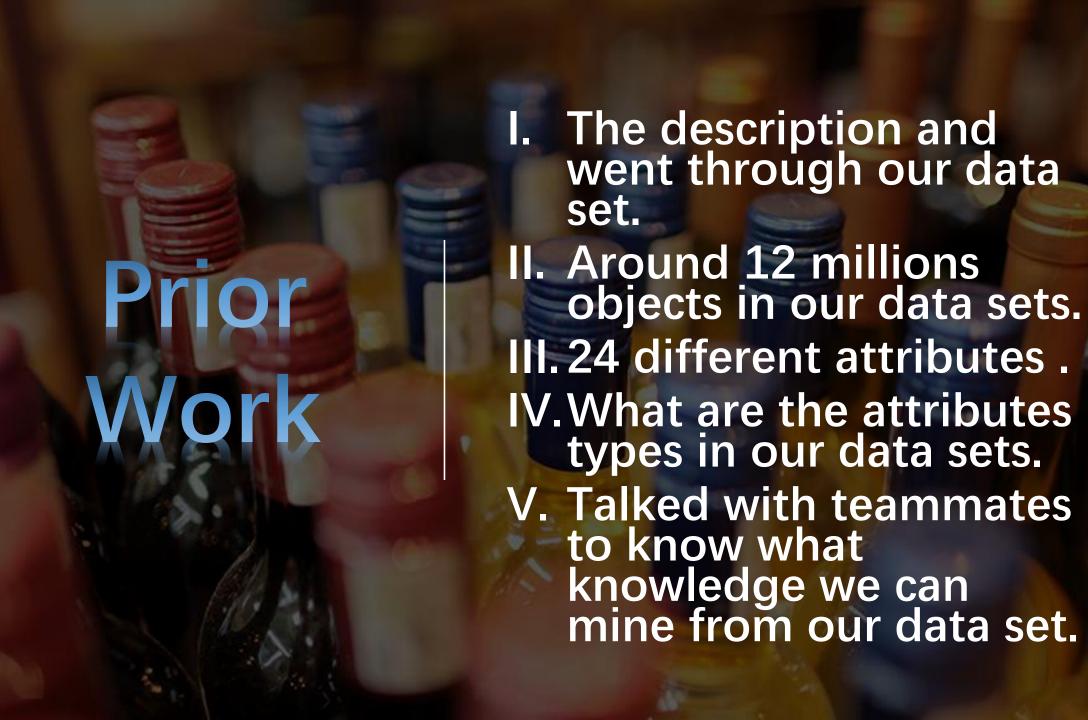# Team members :

- Mingxuan Zhang

- Lei Teng

- Yuxiang Wang

Description:

- Our projects works for analyzing the name, date, kind, price, quantity, and location of sale of sales of individual containers or packages of containers of alcoholic beverages from 2015 to get the sales situation and drinking behaviors of people in different regions.

- Which alcoholic beverage is the most popular in different regions.

- People in which region drink more alcoholic beverage.

- Which kind of alcoholic beverage gains the maximum profit.

- The tendency of sales volume of different alcoholic.

- How much alcoholic beverages sold and consumed.

- The distribution of prices.

Prior Work

I. The description and went through our data set.
II. Around 12 millions objects in our data sets.
III. 24 different attributes .
IV. What are the attributes types in our data sets.
V. Talked with teammates to know what knowledge we can mine from our data set.

Datasets

- Our datasets is from The Iowa Department of Commerce, this is a government agency asked to have a class "E" Liquor license in the bottled form in order to off-the-premises use for individual store which sells alcohol. So each alcoholic sales need registered with Iowa system , then the open data is on the State of Iowa. We find our datasets on Kaggle , but the original published by the State of Iowa.

- Our datasets have around 12 millions objects and 24 different attributes.

- List of datasets to use

  Iowa_Liquor_Sales.csv

- We found data from
  https://www.kaggle.com/residentmario/iowa-liquor-sales
  This datasets original published by the State of Iowa
  https://data.iowa.gov/Economy/Iowa-Liquor-Sales/m3tr-qhgy

- All teammates in our group already downloaded that

#	Invoice	Date	Store	Store Name	Address	City	Zip	Store Location	County#	County	Category#	Category Name	Vendor#	Vendor Name	Item#	Item Description	Pack	Bottle Vol	Cost	Retail	Sold	Sale	Liters	Gallons
4	S2905030(	11/16/2015	3549	Quicker Li	1414 48TH	FORT MA	52627	1414 48TH	56	Lee			130	Disaronno	249	Disaronno	20	150	$6.40	$9.60	2	$19.20	0.30	0.08
5	S2886770(	11/04/2015	2513	Hy-Vee Fo	812 S 1ST	IOWA CIT	52240	812 S 1ST	52	Johnson			65	Jim Beam	237	Knob Cree	3	1750	$35.55	$53.34	3	$160.02	5.25	1.39
6	S2905080(	11/17/2015	3942	Twin Town	104 HIGHW	TOLEDO	52342	104 HIGHW	86	Tama			130	Disaronno	249	Disaronno	20	150	$6.40	$9.60	2	$19.20	0.30	0.08
7	S2886920(	11/11/2015	3650	Spirits, Sto	118 South	HOLSTEIN	51025	118 South	47	Ida			65	Jim Beam	237	Knob Cree	3	1750	$35.55	$53.34	1	$53.34	1.75	0.46
8	S2886570(	11/09/2015	2538	Hy-Vee Fo	1422 FLAM	WATERLC	50702	1422 FLAM	07	Black Haw	1701100	DECANTE	962	Duggan's [	238	Forbidden	6	1500	$11.62	$17.43	6	$104.58	9.00	2.38
9	S2886950(	11/10/2015	3942	Twin Town	104 HIGHW	TOLEDO	52342	104 HIGHW	86	Tama			65	Jim Beam	237	Knob Cree	3	1750	$35.55	$53.34	2	$106.68	3.50	0.92
10	S2933930(	11/30/2015	2662	Hy-Vee Wi	522 MULB	MUSCATIN	52761	522 MULB	70	Muscatine	1701100	DECANTE	65	Jim Beam	173	Laphroaig	12	750	$19.58	$29.37	4	$117.48	3.00	0.79
11	S2905090(	11/16/2015	4307	Crossroad	117 IOWA	DUNLAP	712-2	117 IOWA	43	Harrison			130	Disaronno	249	Disaronno	20	150	$6.40	$9.60	2	$19.20	0.30	0.08
12	S2904990(	11/17/2015	2661	Hy-Vee Fo	1989 PARI	SHELDON	51201	1989 PARI	71	O'Brien			130	Disaronno	249	Disaronno	20	150	$6.40	$9.60	20	$192.00	3.00	0.79
13	S2886820(	11/05/2015	2561	Hy-Vee Fo	4605 FLEU	DES MOIN	50321	4605 FLEU	77	Polk			65	Jim Beam	237	Knob Cree	3	1750	$35.55	$53.34	2	$106.68	3.50	0.92
14	S2886960(	11/09/2015	4114	After 5 Sor	704 W 7TH	ATLANTIC	50022	704 W 7TH	15	Cass			65	Jim Beam	237	Knob Cree	3	1750	$35.55	$53.34	3	$160.02	5.25	1.39
15	S2886690(	11/11/2015	3650	Spirits, Sto	118 South	HOLSTEIN	51025	118 South	47	Ida	1701100	DECANTE	962	Duggan's [	238	Forbidden	6	1500	$11.62	$17.43	1	$17.43	1.50	0.40
16	S2905010(	11/19/2015	2806	Osco #881	1307 N SE	CLINTON	52732	1307 N SE	23	Clinton			130	Disaronno	249	Disaronno	20	150	$6.40	$9.60	20	$192.00	3.00	0.79
17	S2904960(	11/17/2015	2624	Hy-Vee #2	2395 NW A	DUBUQUE	52002	2395 NW A	31	Dubuque			130	Disaronno	249	Disaronno	20	150	$6.40	$9.60	20	$192.00	3.00	0.79
18	S2886840(	11/04/2015	2572	Hy-Vee Fo	6301 UNIV	CEDAR FA	50613	6301 UNIV	07	Black Haw			65	Jim Beam	237	Knob Cree	3	1750	$35.55	$53.34	3	$160.02	5.25	1.39
19	S2919630(	11/24/2015	2595	Hy-Vee Wi	1620 4TH	DENISON	51442	1620 4TH	24	Crawford			255	Wilson Da	297	Templeton	6	750	$18.09	$27.14	6	$162.84	4.50	1.19
20	S2913430(	11/18/2015	3723	J D Spirits	1023 9TH	ONAWA	51040	1023 9TH	67	Monona	1081200	CREAM LI	305	MHW Ltd	258	Rumchata	1	6000	$99.00	$148.50	1	$148.50	6.00	1.59
21	S2886900(	11/10/2015	2665	Hy-Vee / V	1005 E HIC	WAUKEE	50263	1005 E HIC	25	Dallas			65	Jim Beam	237	Knob Cree	3	1750	$35.55	$53.34	2	$106.68	3.50	0.92
22	S2919870(	11/24/2015	5093	Cody Mart	1220 N CC	LE CLAIRE	52753	1220 N CC	82	Scott			255	Wilson Da	297	Templeton	6	750	$18.09	$27.14	3	$81.42	2.25	0.59
23	S2928280(	11/23/2015	2642	Hy-Vee Wi	512 E OSH	PELLA	50219	512 E OSH	63	Marion	1701100	DECANTE	962	Duggan's [	238	Forbidden	6	1500	$11.62	$17.43	6	$104.58	9.00	2.38
24	S2886800(	11/04/2015	2548	Hy-Vee Fo	100 8TH S	ALTOONA	50009	100 8TH S	77	Polk			65	Jim Beam	237	Knob Cree	3	1750	$35.55	$53.34	3	$160.02	5.25	1.39
25	S2919960(	11/20/2015	2558	Hy-Vee Fo	1700 E WA	MOUNT PI	52641	1700 E WA	44	Henry			255	Wilson Da	297	Templeton	6	750	$18.09	$27.14	18	$488.52	13.50	3.57
26	S2905050(	11/18/2015	3735	C B Liquor	1202 A AVI	OSKALOO	52577	1202 A AVI	10	Buchanan			130	Disaronno	249	Disaronno	20	150	$6.40	$9.60	3	$28.80	0.45	0.12
27	S2886700(	11/04/2015	3842	Bancroft Li	107 N POF	BANCROF	50517	107 N POF	55	Kossuth	1701100	DECANTE	962	Duggan's [	238	Forbidden	6	1500	$11.62	$17.43	3	$52.29	4.50	1.19
28	S2886860(	11/09/2015	2650	Hy-Vee Wi	1808 23RD	HARLAN	51537	1808 23RD	83	Shelby			65	Jim Beam	237	Knob Cree	3	1750	$35.55	$53.34	3	$160.02	5.25	1.39
29	S2886910(	11/10/2015	2666	Hy-Vee #2	2510 SW S	ANKENY	50023	2510 SW S	77	Polk			65	Jim Beam	237	Knob Cree	3	1750	$35.55	$53.34	3	$160.02	5.25	1.39
30	S2919790(	11/23/2015	3842	Bancroft Li	107 N POF	BANCROF	50517	107 N POF	55	Kossuth			255	Wilson Da	297	Templeton	6	750	$18.09	$27.14	18	$488.52	13.50	3.57
31	S2886580(	11/09/2015	2539	Hy-Vee Fo	HIGHWAY	IOWA FAL	50126	HIGHWAY	42	Hardin	1701100	DECANTE	962	Duggan's [	238	Forbidden	6	1500	$11.62	$17.43	6	$104.58	9.00	2.38
32	S2886870(	11/10/2015	2651	Hy-Vee / V	1311 4 STI	WAVERLY	50677	1311 4 STI	09	Bremer			65	Jim Beam	237	Knob Cree	3	1750	$35.55	$53.34	2	$106.68	3.50	0.92
33	S2905000(	11/17/2015	2666	Hy-Vee #2	2510 SW S	ANKENY	50023	2510 SW S	77	Polk			130	Disaronno	249	Disaronno	20	150	$6.40	$9.60	20	$192.00	3.00	0.79
34	S2919810(	11/23/2015	4162	Fareway S	4220 16TH	CEDAR R/	52404	4220 16TH	57	Linn			255	Wilson Da	297	Templeton	6	750	$18.09	$27.14	6	$162.84	4.50	1.19
35	S2919580(	11/21/2015	2552	Hy-Vee Fo	20 WILSO	CEDAR R/	52404	20 WILSO	57	Linn			255	Wilson Da	297	Templeton	6	750	$18.09	$27.14	90	$2442.60	67.50	17.83
36	S2919780(	11/23/2015	3650	Spirits, Sto	118 South	HOLSTEIN	51025	118 South	47	Ida			255	Wilson Da	297	Templeton	6	750	$18.09	$27.14	18	$488.52	13.50	3.57
37	S2904910(	11/16/2015	2538	Hy-Vee Fo	1422 FLAM	WATERLC	50702	1422 FLAM	07	Black Haw			130	Disaronno	249	Disaronno	20	150	$6.40	$9.60	40	$384.00	6.00	1.59
38	S2904950(	11/18/2015	2594	Hy-Vee Fo	4500 SER	SIOUX CIT	51106	4500 SER	97	Woodbury			130	Disaronno	249	Disaronno	20	150	$6.40	$9.60	20	$192.00	3.00	0.79
39	S2905100(	11/19/2015	4988	Happy's W	5925 UNIV	CEDAR FA	50613	5925 UNIV	07	Black Haw			130	Disaronno	249	Disaronno	20	150	$6.40	$9.60	20	$192.00	3.00	0.79
40	S2886710(	11/09/2015	4604	Pit Stop Li	1324, 1st A	NEWTON	50208	1324, 1st A	50	Jasper	1701100	DECANTE	962	Duggan's [	238	Forbidden	6	1500	$11.62	$17.43	2	$34.86	3.00	0.79
41	S2904980(	11/16/2015	2643	Hy-Vee Wi	2126 KIME	WATERLC	50701	2126 KIME	07	Black Haw			130	Disaronno	249	Disaronno	20	150	$6.40	$9.60	20	$192.00	3.00	0.79
42	S2904920(	11/17/2015	2544	Hy-Vee Fo	802 SOUT	MARSHAL	50158	802 SOUT	64	Marshall			130	Disaronno	249	Disaronno	20	150	$6.40	$9.60	20	$192.00	3.00	0.79
43	S2919650(	11/21/2015	2607	Hy-Vee Wi	520 SO FF	SHENAND	51601	520 SO FF	73	Page			255	Wilson Da	297	Templeton	6	750	$18.09	$27.14	30	$814.20	22.50	5.94
44	S2919120(	11/19/2015	2248	Ingersoll Li	3500 INGE	DES MOIN	50312	3500 INGE	77	Polk	1701100	DECANTE	65	Jim Beam	173	Laphroaig	12	750	$19.58	$29.37	36	$1057.32	27.00	7.13
45	S2905060(	11/17/2015	3813	CGI Foods	104 NORT	MOUNT A\	50854	104 NORT	80	Ringgold			130	Disaronno	249	Disaronno	20	150	$6.40	$9.60	6	$57.60	0.90	0.24
46	S2919930(	11/20/2015	2539	Hy-Vee Fo	HIGHWAY	IOWA FAL	50126	HIGHWAY	42	Hardin			255	Wilson Da	297	Templeton	6	750	$18.09	$27.14	12	$325.68	9.00	2.38
47	S2919730(	11/21/2015	3443	Super Sav	1141 N BR	COUNCIL	51503	1141 N BR	78	Pottawatta			255	Wilson Da	297	Templeton	6	750	$18.09	$27.14	6	$162.84	4.50	1.19
48	S2886890(	11/10/2015	2661	Hy-Vee Fo	1989 PARI	SHELDON	51201	1989 PARI	71	O'Brien			65	Jim Beam	237	Knob Cree	3	1750	$35.55	$53.34	3	$160.02	5.25	1.39
49	S2919750(	11/20/2015	3525	Wines and	106 W 2NI	WASHING	52353	106 W 2NI	92	Washingto			255	Wilson Da	297	Templeton	6	750	$18.09	$27.14	6	$162.84	4.50	1.19
50	S2904930(	11/19/2015	2567	Hy-Vee Dr	2200 WES	DAVENPC	52806	2200 WES	82	Scott			130	Disaronno	249	Disaronno	20	150	$6.40	$9.60	6	$57.60	0.90	0.24
51	S2919620(	11/20/2015	2591	Hy-Vee Wi	1602 E. 7T	ATLANTIC	50022	1602 E. 7T	15	Cass			255	Wilson Da	297	Templeton	6	750	$18.09	$27.14	24	$651.36	18.00	4.76
52	S2913720(	11/18/2015	2566	Hy-Vee Fo	813 N LIN(	KNOXVILL	50138	813 N LIN(	63	Marion	1701100	DECANTE	962	Duggan's [	238	Forbidden	6	1500	$11.62	$17.43	12	$209.16	18.00	4.76

Proposed work:

- Data cleaning and Data preprocessing
  - :Fill in the missing value manually
  - Fill in it automatically with
    - global constant
    - attribute mean
    - attribute mean of same class
    - inference-based
- Graphic
  - (1)Histogram for different liquors.
- Analyzing the data
  - (1)For example we have to calculate the support for different liquors.
  - (2)We have to calculate the confidence for different liquor and differnet cities in Iowa.
- Get results:
  - (1)From the result, we can find out for different cities what are the popular liquors.
  - (2)What liquor is most popular.

List of tools

- Microsoft Excel
- Pandas
- Python
- Numpy
- Matplotlib
- Delimit
- Jupyter Notebook
- Overleaf

Evaluation:

- We can use refined calculation to evaluate our result like distribution and profit.

- We can use official report on to verify evaluate the result, like Which alcoholic beverage is the most popular in different regions.

  https://www.desmoinesregister.com/story/news/2016/12/28/most-popular-_liquor-iowa-2016-black-velvet/95916046/

- We also can go to State of Iowa , there are some report for sales situation of Liquors in difference regions, they we can evaluate the results like tendency and how much alcoholic beverages sold and consumed