



Analysis of U.S. Car Accident

Team Members: Tianchen Wang, Zitao Cheng, Yu Li



Introduction

- This project goes over the dataset about the U.S. car accident between 2016 and 2020. It basically has these attributes: temperature, precipitation, wind speed, pressure, humidity, etc. To play around these attributes gets the relationship along with the occurrence rate of a car accident in the United States.
- Whether the occurrence rate is related to the day or night.
- Whether the time zone would affect the car accident.
- What range of visibility would cause a car accident more frequently?
- Which time period has the highest occurrence rate of a car accident in a day?



Related Work

1. The description above has indicated the directions we want to conduct data mining to analyze the U.S Car accident.
2. The dataset has 3.5 million objects with 49 attributes.
3. We have downloaded the dataset and discussed what types of data we can mine and what knowledge we learned in the class can utilize in data mining.



Datasets

- We found data from

<https://www.kaggle.com/sobhanmoosavi/us-accidents>

- Our dataset has around 3.5 millions objects and 49 different attributes.
- Moosavi, Sobhan, Mohammad Hossein Samavatian, Srinivasan Parthasarathy, and Rajiv Ramnath. “[A Countrywide Traffic Accident Dataset](#).”, 2019.
- Moosavi, Sobhan, Mohammad Hossein Samavatian, Srinivasan Parthasarathy, Radu Teodorescu, and Rajiv Ramnath. “[Accident Risk Prediction based on Heterogeneous Sparse Data: New Dataset and Insights](#).” In proceedings of the 27th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, ACM, 2019.

	R	S	T	U	V	W	X	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
County	State	Zipcode	Country	Timezone	Airport	Weather	Time	Temperature	Wind	Chill	Humidity	Pressure	Visibility	Wind	Dir	Wind	Speed	Precip	Cloud	Condition	Altitude	Station	Stop	Traffic	Accident	Turn	Surprise	Scenic	View	Nautical	Astronomical	Twilight
2	Montgomery	OH	45424 US	US/Eastern/KFTO	2016/2/8 5:58	36.9	91	29.68	10 Calm	0.02 Light Rain	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	Night	Night	Night	Night	Night	
3	Franklin	OH	43068-34 US	US/Eastern/KCMH	2016/2/8 5:51	37.9	100	29.65	10 Calm	0.02 Light Rain	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	Night	Night	Night	Night	Night	
4	Clermont	OH	45176 US	US/Eastern/KT19	2016/2/8 6:56	36	33.3	100	29.67	10 SW	3.5	Overcast	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE	FALSE	Night	Night	Day	Day	Day	
5	Montgomery	OH	45417 US	US/Eastern/KDAY	2016/2/8 7:38	35.1	31	96	29.64	9 SE	4.6	Mostly Cloudy	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	Night	Night	Day	Day	Day	
6	Montgomery	OH	45459 US	US/Eastern/KMGV	2016/2/8 7:53	36	33.3	89	29.65	6 SW	3.5	Mostly Cloudy	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE	FALSE	Day	Day	Day	Day	Day	
7	Franklin	OH	43081 US	US/Eastern/KCMH	2016/2/8 7:51	37.9	35.5	97	29.63	7 SSW	3.5	0.03 Light Rain	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	Day	Day	Day	Day	Day	
8	Montgomery	OH	45417-2470 US	US/Eastern/KDAY	2016/2/8 7:56	34	31	100	29.66	7 WSW	3.5	Overcast	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	Day	Day	Day	Day	Day	
9	Montgomery	OH	45405 US	US/Eastern/KDAY	2016/2/8 7:56	34	31	100	29.66	7 WSW	3.5	Overcast	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	Day	Day	Day	Day	Day	
10	Montgomery	OH	45404-1920 US	US/Eastern/KFTO	2016/2/8 7:58	33.3	99	29.67	5 SW	1.2	Mostly Cloudy	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	Day	Day	Day	Day	Day	
11	Franklin	OH	43081 US	US/Eastern/KCMH	2016/2/8 8:28	37.4	33.8	100	29.62	3 SSW	4.6	0.02 Light Rain	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	Day	Day	Day	Day	Day	
12	Franklin	OH	43228 US	US/Eastern/KITZ	2016/2/8 7:50	35.6	30.7	93	29.64	5 WNW	5.8	Rain	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	Day	Day	Day	Day	Day	
13	Franklin	OH	43068 US	US/Eastern/KCMH	2016/2/8 8:28	37.4	33.8	100	29.62	3 SSW	4.6	0.02 Light Rain	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	Day	Day	Day	Day	Day	
14	Montgomery	OH	45420-1800 US	US/Eastern/KFTO	2016/2/8 8:28	33.8	89	29.63	3 SW	2.3	Overcast	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	Day	Day	Day	Day	Day	
15	Montgomery	OH	45406-2700 US	US/Eastern/KDAY	2016/2/8 8:56	36	31.1	89	29.65	10 NW	5.8	Mostly Cloudy	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	Day	Day	Day	Day	Day	
16	Franklin	OH	43213-1000 US	US/Eastern/KCMH	2016/2/8 8:28	37.4	33.8	100	29.62	3 SSW	4.6	0.02 Light Rain	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	Day	Day	Day	Day	Day	
17	Montgomery	OH	45410-1720 US	US/Eastern/KCMH	2016/2/8 8:28	33.8	100	29.63	3 SW	2.3	Overcast	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	Day	Day	Day	Day	Day	
18	Montgomery	OH	45402 US	US/Eastern/KFTO	2016/2/8 8:58	35.6	89	29.65	7 WSW	2.3	Mostly Cloudy	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	Day	Day	Day	Day	Day	
19	Montgomery	OH	45417-1720 US	US/Eastern/KDAY	2016/2/8 8:56	36	31.1	89	29.65	10 NW	5.8	Mostly Cloudy	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	Day	Day	Day	Day	Day	
20	Montgomery	OH	45409-2600 US	US/Eastern/KFTO	2016/2/8 9:38	37.4	32.1	93	29.63	10 WSW	6.9	Overcast	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE	FALSE	Day	Day	Day	Day	
21	Montgomery	OH	45406-2640 US	US/Eastern/KDAY	2016/2/8 9:56	36	30.3	89	29.65	10 NW	6.9	Mostly Cloudy	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	Day	Day	Day	Day	Day	
22	Franklin	OH	43230-1700 US	US/Eastern/KCMH	2016/2/8 10:14	33.8	29.6	100	29.62	2 WNW	4.6	0.01 Light Snow	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	Day	Day	Day	Day	Day	
23	Montgomery	OH	45406-4700 US	US/Eastern/KDAY	2016/2/8 9:56	36	30.3	89	29.65	10 West	6.9	Mostly Cloudy	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	Day	Day	Day	Day	Day	
24	Montgomery	OH	45432 US	US/Eastern/KMGV	2016/2/8 11:53	35.1	28.6	89	29.65	6 WSW	8.1	0.02 Overcast	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	Day	Day	Day	Day	Day	
25	Franklin	OH	43211 US	US/Eastern/KCMH	2016/2/8 11:51	37	32.4	90	29.63	8 West	5.8	0 Overcast	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	Day	Day	Day	Day	Day	
26	Montgomery	OH	45417-1100 US	US/Eastern/KDAY	2016/2/8 12:20	37.9	30.9	76	29.62	10 WNW	10.4	Overcast	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	Day	Day	Day	Day	Day	
27	Licking	OH	43031-9100 US	US/Eastern/KCMH	2016/2/8 12:51	37.9	34.4	100	29.6	5 West	4.6	0.01 Light Snow	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	Day	Day	Day	Day	Day	
28	Montgomery	OH	45419-4100 US	US/Eastern/KMGV	2016/2/8 12:53	36	29	86	29.63	7 West	9.2	0 Light Snow	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	Day	Day	Day	Day	Day	
29	Montgomery	OH	45405-4300 US	US/Eastern/KDAY	2016/2/8 13:56	39.9	32.9	70	29.61	10 WNW	11.5	Mostly Cloudy	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	Day	Day	Day	Day	Day	
30	Montgomery	OH	45406-2700 US	US/Eastern/KDAY	2016/2/8 14:56	39.9	34.5	65	29.61	10 WNW	8.1	Mostly Cloudy	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	Day	Day	Day	Day	Day	
31	Montgomery	OH	45402 US	US/Eastern/KFTO	2016/2/8 15:46	39.2	35.9	75	29.62	10 WNW	4.6	Mostly Cloudy	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	Day	Day	Day	Day	Day	
32	Franklin	OH	43017-2500 US	US/Eastern/KOSU	2016/2/8 15:53	37	31.6	89	29.61	10 West	6.9	0 Overcast	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	Day	Day	Day	Day	Day	
33	Montgomery	OH	45424-4100 US	US/Eastern/KFTO	2016/2/8 17:58	34.2	31.2	97	29.66	10 NW	3.5	0.02 Light Rain	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	Day	Day	Day	Day	Day	
34	Franklin	OH	43224 US	US/Eastern/KCMH	2016/2/8 18:43	33.8	28.6	100	29.62	6 South	5.8	Overcast	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	Night	Day	Day	Day	Day	
35	Franklin	OH	43213 US	US/Eastern/KCMH	2016/2/8 19:51	34	100	29.65	6 Calm	0.07 Overcast	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	Night	Night	Night	Night	Night	
36	Franklin	OH	43212 US	US/Eastern/KOSU	2016/2/8 20:27	33.1	26.1	92	29.63	7 WNW	8.1	0 Light Snow	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	Night	Night	Night	Night	Night	
37	Franklin	OH	43210 US	US/Eastern/KOSU	2016/2/8 20:39	33.1	25.5	92	29.63	9 WNW	9.2	0 Overcast	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	Night	Night	Night	Night	Night	
38	Montgomery	OH	45409-2610 US	US/Eastern/KFTO	2016/2/8 4:28	23	11.7	93	29.65	5 SW	11.5	Light Snow	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	Night	Night	Night	Night	Night	
39	Fairfield	OH	43110 US	US/Eastern/KLCK	2016/2/8 4:55	23	11.7	93	29.63	2.5 WSW	11.5	Light Snow	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	Night	Night	Night	Night	Night	
40	Montgomery	OH	45404-1670 US	US/Eastern/KCMH	2016/2/8 4:58	22.8	11.5	89	29.69	4 SW	11.5	0 Light Snow	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	Night	Night	Night	Night	Night	
41	Franklin	OH	43219-5000 US	US/Eastern/KCMH	2016/2/8 5:27	26.6	18.2	86	29.62	1.5 West	8.1	0.01 Light Snow	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	Night	Night	Night	Night	Night	
42	Warren	OH	45036 US	US/Eastern/KMGV	2016/2/8 5:45	21	10.6	85	29.67	0.8 SW	9.2	0.01 Light Snow	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	Night	Night	Night	Night	Night	
43	Montgomery	OH	45404-1200 US	US/Eastern/KFTO	2016/2/8 5:58	23	12.4	86	29.71	3 SW	10.4	0 Light Snow	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	Night	Night	Night	Night	Night	
44	Clinton	OH	45169 US	US/Eastern/KILN	2016/2/8 5:54	19.9	6.7	89	29.75	1.8 WSW	13.8	0 Light Snow	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	Night	Night	Night	Night	Night	
45	Franklin	OH	43213 US	US/Eastern/KOSU	2016/2/8 5:51	25	15.5	88	29.65	1 SW	9.2	0.02 Light Snow	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	Night	Night	Night	Night	Day	
46	Montgomery	OH	45414 US	US/Eastern/KDAY	2016/2/8 6:03	23	10.1	88	29.66	2 West	15	0 Light Snow	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	Night	Night	Night	Night	Day	
47	Franklin	OH	43205 US	US/Eastern/KCMH	2016/2/8 5:51	25	15.5	88	29.65	1 SW	9.2	0.02 Light Snow	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	Night	Night	Night	Night	Day	
48	Franklin	OH	43220 US	US/Eastern/KOSU	2016/2/8 6:22	26.1	15	84	29.63	2 West	12.7	0 Light Snow	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	Night	Night	Night	Night	Day	
49	Franklin	OH	43232 US	US/Eastern/KCMH	2016/2/8 6:36	24.8	16.1	86	29.64	1 SW	8.1	0.01 Light Snow	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	Night	Night	Night	Night	Day	
50	Montgomery	OH	45404 US	US/Eastern/KFTO	2016/2/8 6:51	23	11.7	86	29.68	4 SW	11.5	Light Snow	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE	Night	Night	Night	Night	Day
51	Montgomery	OH	45404 US	US/Eastern/KFTO	2016/2/8 6:51	23	11.7	86	29.68	4 SW	11.5	Light Snow	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	Night	Night	Day	Day	



Proposed work

- Data Cleaning and Data Reduction
 - To deal with the missing value, we can put each attribute mean in these missing blocks.
 - More importantly, we have to make data reduction to compress our data since there are too many car accidents in the dataset.
- Graphic Analysis
 - Histogram, scatter plot, line chart, and boxplot for different attributes, e.g., wind, humidity, temperature, time period, etc.
 - Do some calculations to analyze the relationship. For example, get the mean of visibility and see what range of visibility would have the most severe car accident.
- Analyzing the data
 - By using the result we got from the previous part, we could examine the relationship between accident and attributes we chose



Evaluation

- We can use official traffic accident analysis from the authority to evaluate our analysis result.
 - <https://www.resourcesmag.org/archives/how-climate-change-affects-traffic-accidents/>
 - <https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2019GH000212>
- Besides, we can find which one factor we do not discover from the dataset and which one does not exist in authority analysis, but we find it.



Milestones

- Milestone 1 (due: Oct. 11)
 - Data cleaning work
- Milestone 2 (due: Oct. 25)
 - Data reduction work
- Milestone 3 (due: Nov. 8)
 - Graphic analysis
- Milestone 4 (due: Nov.15)
 - All code of our project
 - Start first draft of our final report
- Milestone 5 (due: Nov. 22)
 - Finish the final version of our project report
- Milestone 6 (due: Nov. 29)
 - Finish the slides for final presentation and prepare to present