

Factors that cause traffic accidents in DC

—

Yuka & Connie

Data Collection

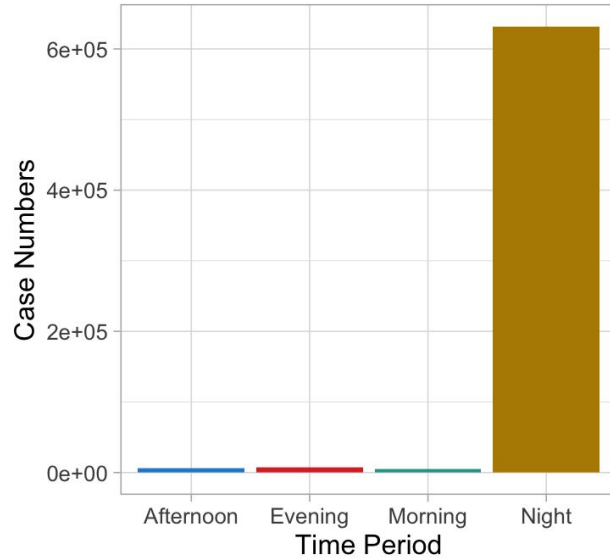
DC Govt : Traffic Data from 2000 - 2021

- Pulled the data using the API
- We split the data 2000 - 2010/2010-2021
- Number of variables and observations

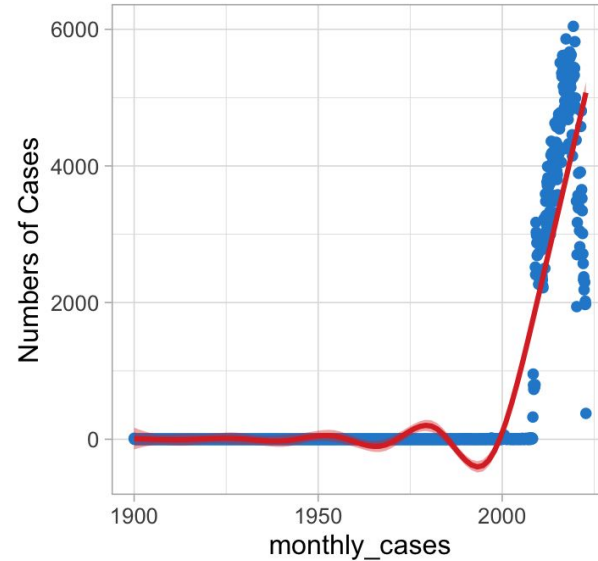
Data Exploratory Analysis

Car Crashes by Time Period

Time Period: ■ Afternoon ■ Evening ■ Morning

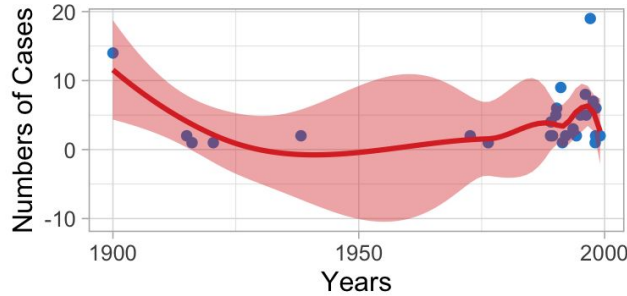


Car Crashes Cases by Months Before 2000

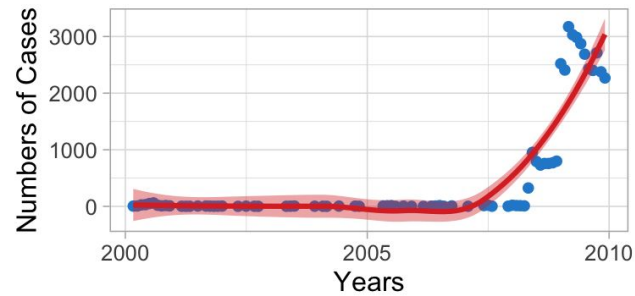


Data Exploratory Analysis

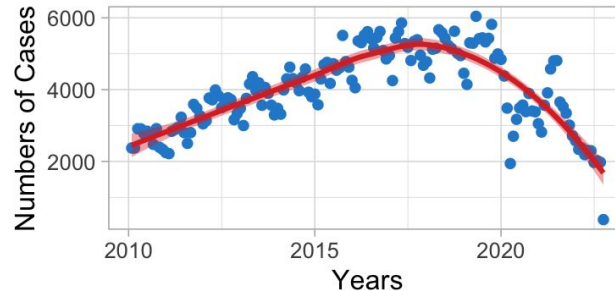
Car Crashes Cases by Months
between 2000 to 2010



Car Crashes Cases by Months
between 2000 to 2010

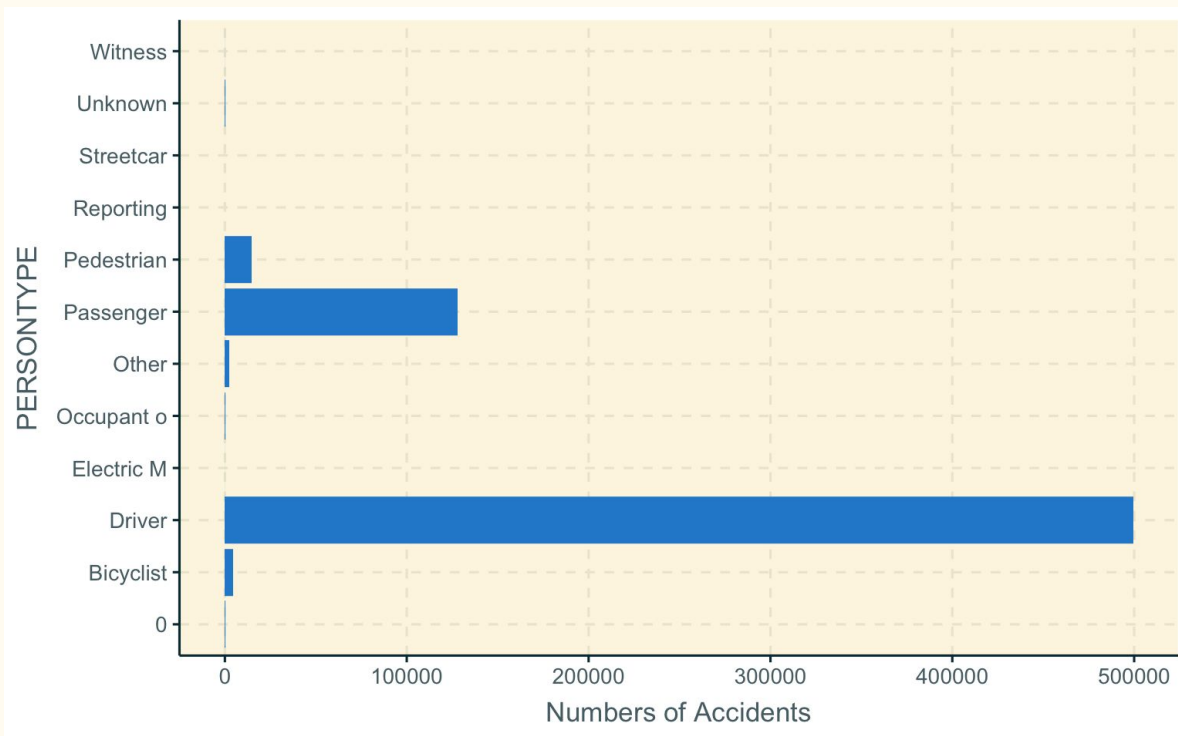


Car Crashes Cases by Months
after 2010



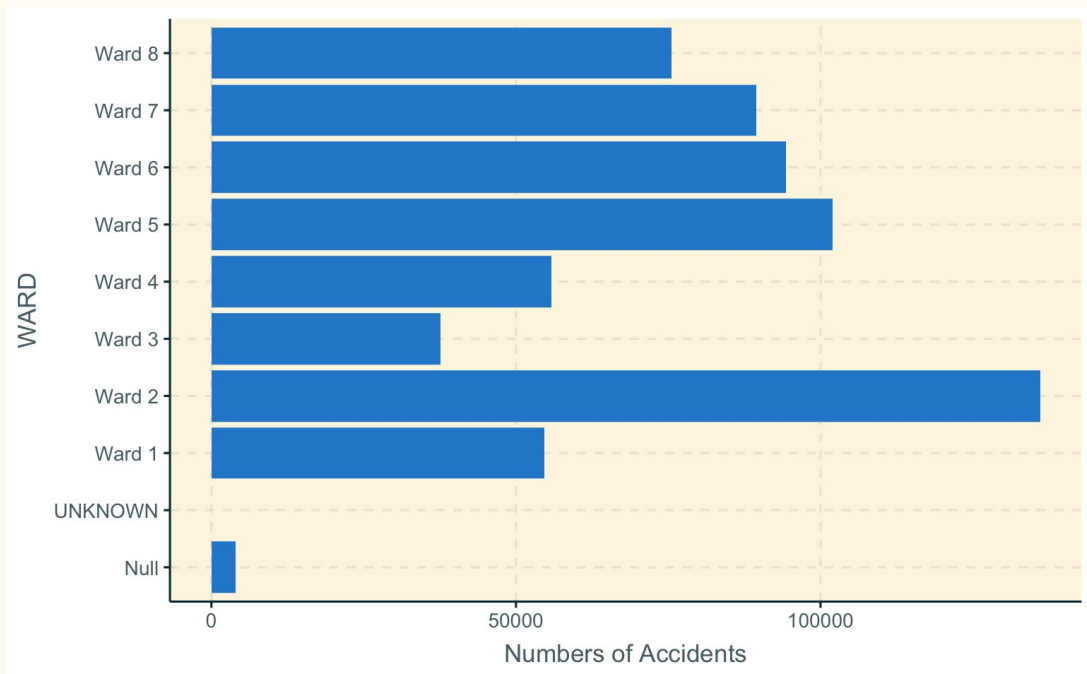
Data Exploratory Analysis

PERSONTYPE	Numbers of Accidents
0	182
Bicyclist	4714
Driver	447058
Electric M	31
Occupant o	338
Other	2200
Passenger	112612
Pedestrian	12438
Reporting	2
Streetcar	4
Unknown	128
Witness	36



Data Exploratory Analysis

WARD	Numbers of Accidents
Null	3798
UNKNOWN	2
Ward 1	48351
Ward 2	120658
Ward 3	31802
Ward 4	48680
Ward 5	91812
Ward 6	84594
Ward 7	81096
Ward 8	68950



Project So Far

Milestones

- Initial Data Set
- Models and Initial Analysis
- Great Communication among Group

Limitations

- Time Management
- Data Integration and Cleaning



General Data Analysis

- Logistic Regression

term	estimate	std.error	statistic	p.value
SPEEDINGY	2.7192910	0.1631818	16.664176	0.0000000
time_periodNight	1.3488187	0.7178111	1.879072	0.0602347
TOTAL_BICYCLES	0.7841557	0.2497068	3.140306	0.0016877
TOTAL_PEDESTRIANS	0.6434274	0.0433031	14.858703	0.0000000
month10	0.5676182	0.2420357	2.345183	0.0190177
month7	0.4368960	0.2476674	1.764043	0.0777247
AGE	0.0107488	0.0022153	4.852089	0.0000012

- According to the 2020 census data Ward 7 and 8 have the lowest income levels and very high poverty level in the District of Columbia. However, our model do not show ward has as much influences compared to other predictors.
- Pedestrians and bikers are at high risk of being involved in fatal accidents - there is a need for DC to invest in sidewalks and bicycle lanes.
- Age also has strong effect on cause of fatal accidents.
- Speeding is the strongest predictor for probabilities of be involved in Fatal accidents, for every accident that take place there is a probability of 2.71 if speeding.
- Night Driving isn't safe in DC with probabilities of 1.34 of being involved in a fatal accident.

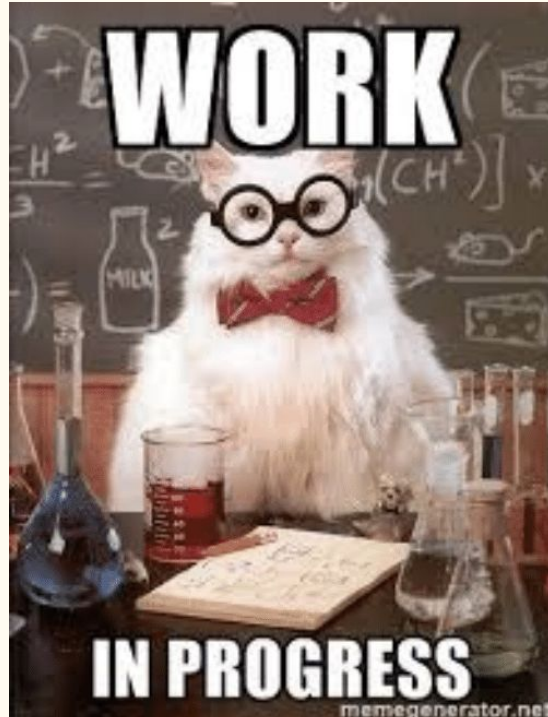
General Data Analysis

- Decision Trees



General Data Analysis

- KNN



References

- Barba, Carolina Tripp, Miguel Angel Mateos, Pablo Reganas Soto, Ahmad Mohamad Mezher, and Mónica Aguilar Igartua. 2012. “Smart City for VANETs Using Warning Messages, Traffic Statistics and Intelligent Traffic Lights.” In *2012 IEEE Intelligent Vehicles Symposium*, 902–7. IEEE.
- Hayakawa, Hiroshi, Paul S. Fischbeck, and Baruch Fischhoff. 2000. “Traffic Accident Statistics and Risk Perceptions in Japan and the United States.” *Accident Analysis & Prevention* 32 (6): 827–35.
- Keep, Matthew, and Tom Rutherford. 2013. “Reported Road Accident Statistics.” *Commons Library Standard Note, SN/SG/2198*.
- Kingham, Simon, Clive E Sabel, and Phil Bartie. 2011. “The Impact of the ‘School Run’ on Road Traffic Accidents: A Spatio-Temporal Analysis.” *Journal of Transport Geography* 19 (4): 705–11.
- Marzoug, R, N Lakouari, O Oubram, H Ez-Zahraouy, A Khallouk, M Limón-Mendoza, and JG Vera-Dimas. 2018. “Impact of Traffic Lights on Car Accidents at Intersections.” *International Journal of Modern Physics C* 29 (12): 1850121.
- Renouf, MA. 1991. “A Car Accident Injury Database: Overview and Analyses of Entrapment and Ejection.”

Our Repo:

[https://github.com/yjchen9596/
AU-GOVT670-F22-Project](https://github.com/yjchen9596/AU-GOVT670-F22-Project)

