

Review of 2019 Accident Data

Go Safe Labs
research@gosafelabs.com

Go Safe Labs
San Francisco, CA 94107

Abstract

Official government reports on accidents often don't come out until months or even years after the date of accident. Meanwhile accident patterns can change dramatically, so timely analysis of accident data is critical to keeping today's drivers safe. Here we analyze an 1.8 million accident report sample compiled across the 48 contiguous United States in 2018 and 2019. Preliminary results show that accidents increased in 2019 compared to 2018. 2019 top 10 crash 'hotspots' include heavily trafficked areas such as the George Washington Bridge crossing in NYC, but also reveal a series of unexpected hotspots in South Carolina. Houston continued to see the most accidents in 2019, followed by Charlotte, Los Angeles, Austin, Dallas, Raleigh, Oklahoma City, Baton Rouge, Nashville, and Phoenix.

1 Introduction

Relying on official government reports is the gold standard for accident reporting, but the time lag between accident occurrence and official report can be measured in month or even year timeframes[1]. Analyzing systematically collected data from online accident reporting services is one way to circumvent this lag and provide more timely analysis. Analyzing over 1.8 million accidents in 2018 and 2019, we asked if accidents were increasing, and where they were most concentrated. Accidents increased in 2019 compared to 2018 in this sample. Houston saw the most 2019 accidents, followed by Charlotte, Los Angeles, Austin, Dallas, Raleigh, Oklahoma City, Baton Rouge, Nashville, and Phoenix. These cities were also in the top 10 for 2018, with the exception of Phoenix replacing Atlanta in the #10 spot. Surprisingly, the top 10 2019 crash hotspots were all located outside these high volume cities. We highlight these 2019 Go Safety Hotspots to keep 2020 drivers up to date and vigilant.

2 Methodology

This report draws on data collected by Sobhan Moosavi, Ph.D. during his doctoral work with colleagues and collaborators in the Rajiv Ramnath lab at The Ohio State University's Department of Computer

Science & Engineering [2][3[4]. Briefly, across the contiguous U.S., 3.0M accidents were recorded from the Bing or Mapquest APIs, along with concurrent weather information from February 2016 until the end of December 2019. Note that data is not equally distributed across the US, with larger states (TX, CA, FL, NY) generating a substantial fraction of the data [5].

To find traffic patterns, we isolated 1,846,245 accidents that occurred from January 2018 to the end of December 2019. To find year over year variation, we compared 2018 accident counts to 2019 counts. We then subset by city code and found the 10 highest accident counts by city during 2019.

To find accident hotspots, we subset by latitude/longitude pair for 2019 and found the 10 highest accident count 'hotspots'. To account for errors in geolocating, we clustered latitude/longitude decimal degree accuracy to three places, which creates roughly 110m accuracy 'spots'. We further removed duplicate reports within one hour of each other at the same lat/long pair.

3 2019 Safety Ranking

Rank	Latitude	Longitude	City	Count	Location
1	34.859	-82.260	Greenville	404	1-85 at Pelham Rd Exit
2	40.846	-73.932	New York	208	I-95 at Harlem River Dr, west of Harlem River
3	34.831	-82.299	Greenville	189	I-85/I-385 interchange
4	44.968	-93.270	Minneapolis	144	Old Highway 65/I-94 (under construction)
5	40.845	-73.925	Bronx	142	I-95 at Sedgwick Ave, east of Harlem River
6	34.084	-80.939	Columbia	135	Two Notch Rd at Trenholm Rd Extension
7	40.848	-73.943	New York	132	I-95/George Washington Bridge at Riverside Dr
8	45.580	-122.544	Portland	124	I-205 Glenn L Jackson Memorial Bridge over Columbia River
9	25.899	-80.210	North Miami	119	I-95 at Opa-Locka Blvd
10	34.799	-82.372	Greenville	111	I-85 between Maudlin Rd and 291

City	2019 Count	2018 Count
Houston	22,188	25,244
Charlotte	21,818	25,172
Los Angeles	19,660	15,779
Austin	16,635	16,151
Dallas	14,685	14,251
Raleigh	12,846	17,246
Oklahoma City	12,476	11,759
Baton Rouge	11,313	11,982
Nashville	10,091	11,925
Phoenix	9,876	7,998

4 Preliminary Conclusions

We analyzed 1,846,245 accident reports from 2018 and 2019 to find emerging areas of concern to the safe driver. First, we asked if accidents are trending up or down year-over-year. Overall we find a 6.8% increase in accidents in 2019 over 2018 (953,630, up from 892,615). Second, we asked which cities saw the most accidents in 2019. Houston saw the most accidents (22,188) followed by Charlotte (21,818), Los Angeles (19,660), Austin (16,635), Dallas (14,685), Raleigh (12,846), Oklahoma City (12,476), Baton Rouge (11,313), Nashville (10,091), and Phoenix (9,876).

Notably, while overall accidents rose in 2019 over 2018, not all the top 10 cities saw year over year decreases or were flat. Decreases were seen in Houston (-12.1%), Charlotte (-13.3%), Raleigh (-25.5%), Baton Rouge (-5.6%), and Nashville (-15.4%). Small increases were seen in Austin (+3.0%), Dallas (+3.0%), and Oklahoma City (+6.1%). Los Angeles (+24.6%) and Phoenix (+23.5%) were the only top 10 cities to see a sizable increase in accidents.

Hotspots could be split into roughly two bins. On the one hand, there are areas of high absolute traffic volume eg three of the top 10 occur along the I-95 crossing from New Jersey, across Upper Manhattan and into the Bronx. Similarly, high traffic areas in Minneapolis, Portland, and Miami appear. On the other hand, the top 10 data includes several medium population areas in South Carolina, namely Greenville and Columbia. One possible explanation for this anomaly may be data-related, for example that South Carolina traffic reporting may use more lenient geolocations than other similarly sized DOTs, or may be mis-reporting or re-reporting accidents (we took reasonable measures to eliminate multiple re-reporting based on data available, see above Methodology). Changes in road use, like construction, could also be to blame for a hotspot.

As this is a broad-based sample study, we did not correct for confounding factors such as city area, traffic volume, population, or vehicle miles traveled [VMT]. We also did not account for the severity of accidents, just the volume of accidents occurring. Future work is needed to account for reporting bias and adjust for additional factors. Our aim in this report is straightforward: find recent high volume accident areas across the US, and relay this to the safe driver.

Exhibits A & B

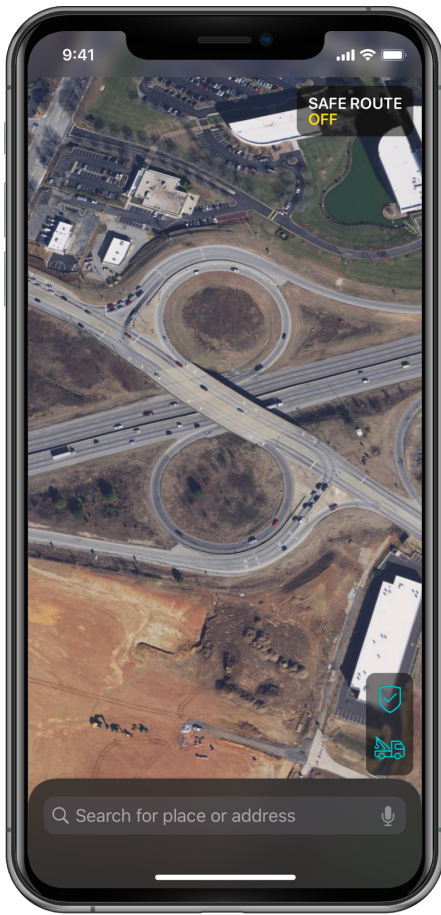


Exhibit A

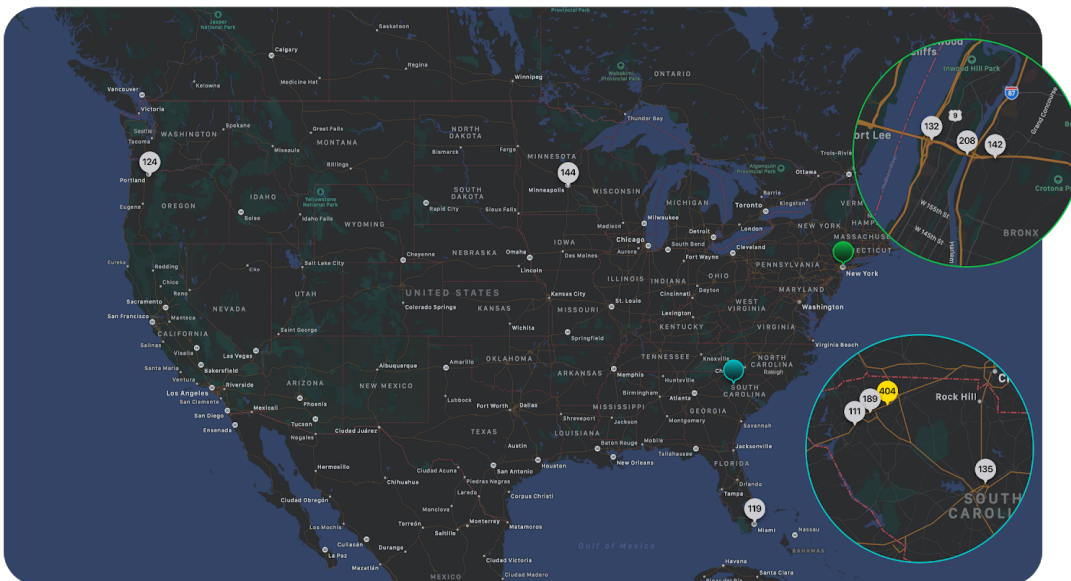


Exhibit B

References

[1] <https://www.nhtsa.gov/press-releases/roadway-fatalities-2018-fars>, publication date: Oct 22, 2019

[2] Moosavi, Sobhan, Mohammad Hossein Samavatian, Srinivasan Parthasarathy, and Rajiv Ramnath. "[A Countrywide Traffic Accident Dataset.](#)", arXiv preprint arXiv:1906.05409 (2019).

[3] 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.

[4] <https://cse.osu.edu/people/ramnath.6>

[5] https://smoosavi.org/datasets/us_accidents