

# Traffic, Weather, & Incidents in Nashville, TN

Zakariyya Al-Quran,  
Rithik Reddy,  
Mohamed Sufiyaan





# Introduction

- Incident Response is a challenge faced by communities across the globe.
- Globally, about 3,200 people die every day from road accidents alone, leading to a total of 1.25 million deaths annually
- Furthermore, accidents can cause traffic and congestion
- We want to analyze the incidents, and their relationship with response time, traffic, and weather.





# Data Sets



- Incidents
  - Data of all of the incidents which occurred in Nashville from 2017-2021
- Roads
  - Data of all roadways in Tennessee
  - Gives info such as length of roadway, number of lanes, etc.
- Traffic
  - Gives traffic data across Nashville
- Weather
  - Weather data across Nashville. Weather in each area, precipitation, etc.



# Technologies Used

- Storing and loading data
  - S3 to store original and merged data sets
- Transforming data
  - Spark with EMR to add window column
- Querying data
  - Athena
- Visualization
  - Pandas to load result of queries
  - Plotly to generate maps and graphs
- Machine learning
  - Spark with EMR to run models



plotly

Pandas



**amazon**  
ATHENA



Amazon S3



**amazon**  
EMR



jupyter

**Spark**

# Sample Queries



1. 

```
SELECT xgroup, COUNT(distinct id_original) as cnt, AVG(congestion) as cong
FROM "incidents"."merge" where xgroup != 0
GROUP BY xgroup ORDER BY cnt DESC, cong DESC LIMIT 20;
```
2. 

```
SELECT COUNT(distinct incident_id) as num_incidents, AVG(response_time_sec) as avg_response,
AVG(speed) as avg_speed, hour_of_day
FROM "incidents"."merge" WHERE xgroup in (SELECT xgroup FROM "incidents"."top20" as t20)
GROUP BY hour_of_day ORDER BY hour_of_day;
```
3. 

```
SELECT COUNT(incident_id) as num_incident, temp_range, AVG(avg_response) as response,
AVG(avg_speed) as speed, AVG(avg_congestion) as cong
FROM (SELECT *, CAST(FLOOR(avg_temp/10) as INT) as INT) temp_range FROM "incidents"."weather")
GROUP BY temp_range ORDER BY temp_range;
```
4. 

```
SELECT COUNT(distinct incident_id) as num_incidents, month
FROM (SELECT * FROM "incidents"."merge"
WHERE latitude >= 36.13 AND latitude < 36.15 AND longitude >= -86.81 AND longitude < -86.79)
GROUP BY month ORDER BY month;
```

# Sample Results

Query 1: Top 20 Roadways

1	xdgroup	cnt	cong
2	2858287	623	0.15202679436521624
3	1746789	377	0.08957439740297297
4	1743946	356	0.10160349472945007
5	1621379	340	0.12447910501136183
6	1740471	238	0.10180981537258228

Query 2: Incident Statistics by Hour

7	1681622	237	1	num_incidents	avg_response	avg_speed	hour_of_day
8	2860921	232	2	110	403.53636363636366	49.85473737373736	0
9	3177417	222	3	104	448.44005358338916	54.87695378432684	1
10	3176773	219	4	115	514.7913043478261	50.82140700483088	2
11	1722191	218	5	100	471.0251051893408	54.766370266479655	3
12	1627331	216	6	78	434.69005010737294	56.06633500357912	4
			7	138	420.8027998791419	56.72735220062441	5
			8	209	411.79514143094843	47.26504159733775	6
			9	212	391.31253277399054	47.3891033036182	7
			10	174	406.3735632183908	46.3839750957854	8
			11	159	391.50314465408803	44.568080887491234	9
			12	177	378.01483167229065	46.44047006199475	10
			13	227	391.15859030837004	46.1140363436123	11

# Sample Results (Cont'd.)

Query 3: Incident Statistics by Temperature Range

	num_incident	temp_range	avg_response	avg_speed	avg_congest
1					
2	386	-1	400.97668393782385	37.95402752810969	0.11222674
3	3214	0	381.3453640323584	36.960817155673716	0.11476722
4	4383	1	372.6390600045631	36.50840666420556	0.11988138
5	5294	2	361.5521344918776	37.19889814320872	0.11488573
6	1299	3	367.75365665896845	33.70927520048498	0.16693998

Temperature Range

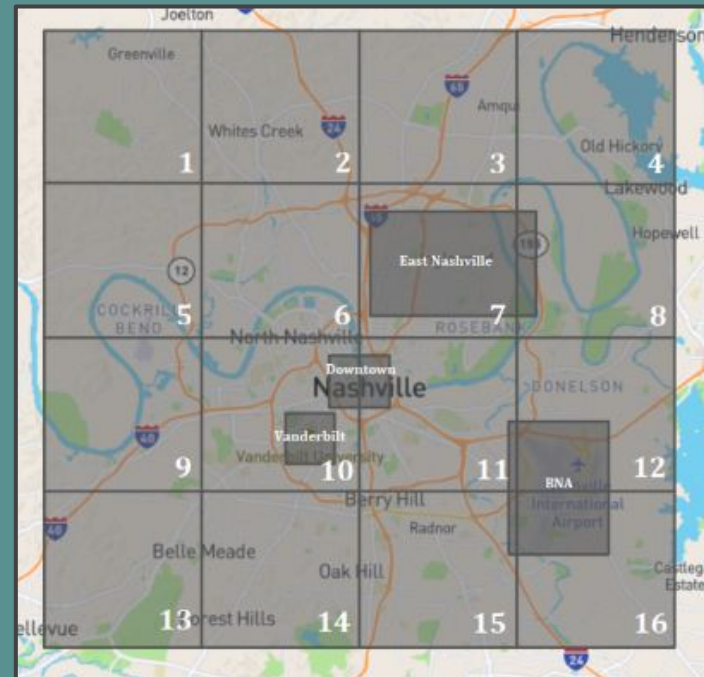
-1	-10 °C to 0 °C
0	0 °C to 10 °C
1	10 °C to 20 °C
2	20 °C to 30 °C
3	30 °C +

Query 4: Vanderbilt Incident Frequency by Month

num_incidents	month
15	1
9	2
7	3
7	4
9	5
13	6
16	7
28	8
12	9
19	10
6	11
10	12

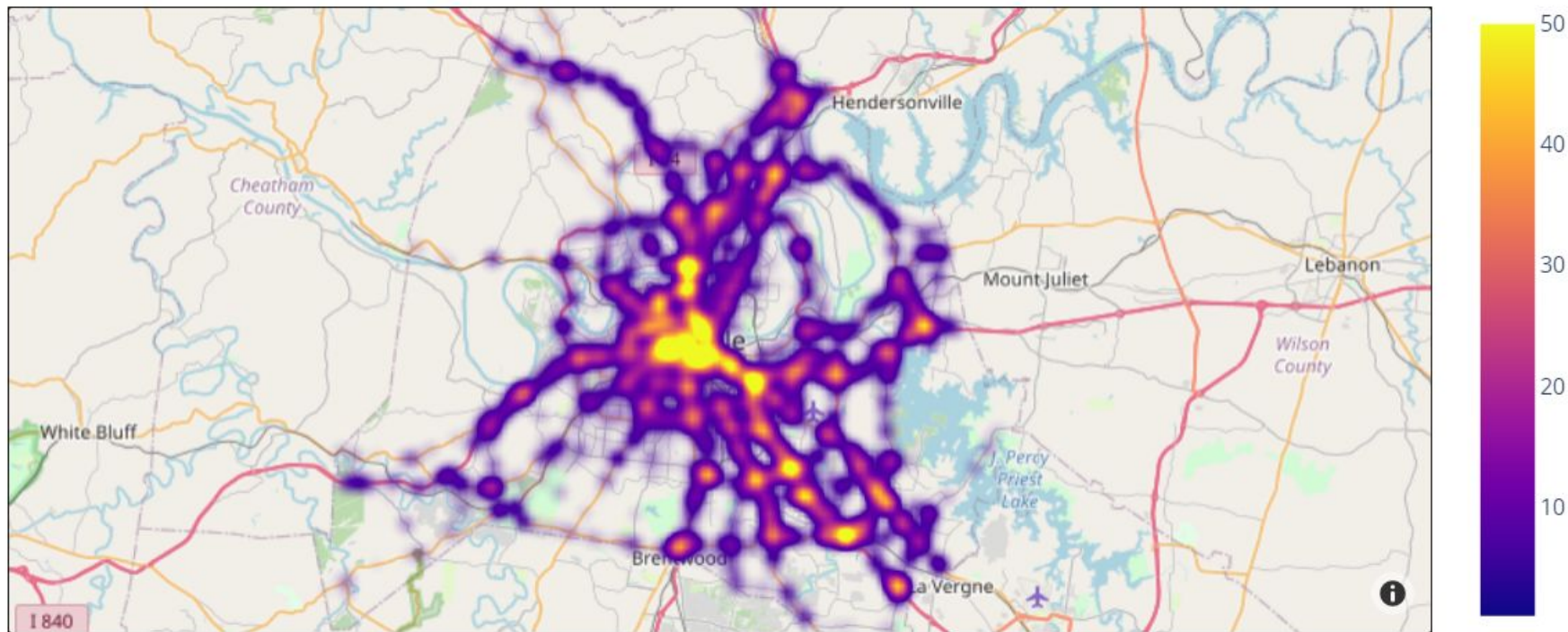
# Machine Learning

Area of grid	Linear Regression RMSE	Random Forest Regression RMSE	Gradient Boosted Tree Regression RMSE	Lowest RMSE Regressor's Prediction
Q1	0.5369	0.8814	<b>0.5</b>	1
Q2	<b>1.6988</b>	1.83868	1.7388	5.33
Q3	10.914	<b>5.7268</b>	11.187	24.2
Q4	<b>2.499</b>	2.848	2.965	6.49
Q5	1.929	<b>1.456</b>	1.732	2.5
Q6	13.559	<b>5.552</b>	8.503	20.6
Q7	19.96	<b>7.439</b>	10.583	40.95

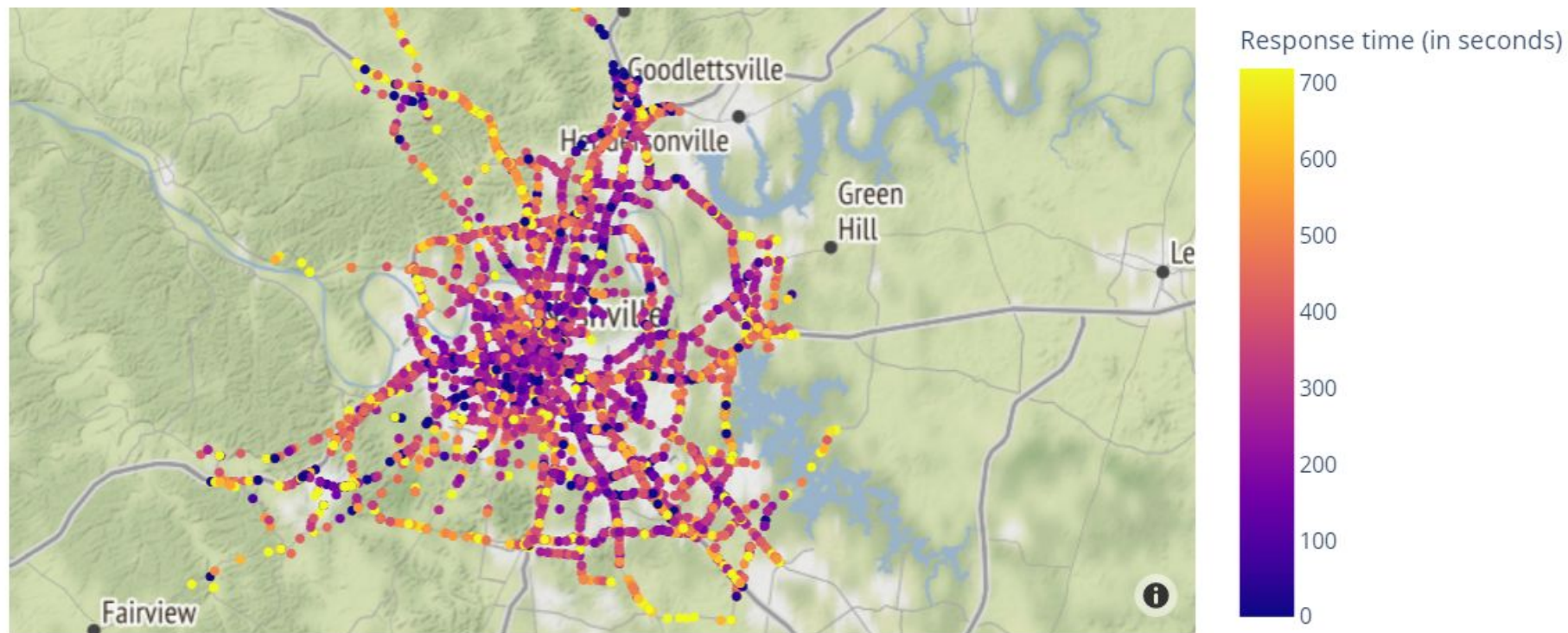




## Heatmap of All Incidents in Nashville, TN



## Scatter Map of All Incidents Response Time in Nashville, TN



Average Speed and Congestion by Hour in Davidson County



Plots

# Average Congestion per Month in Davidson County from 2019-2021



## Plots

## Average Speed by Hour in Davidson County in 2019 vs 2020



# Plots



# Implications

- **Time Implications**
  - Congestion/speed depends on time of day, day of week, etc
  - COVID more or less *didn't* affect the number of incidents on Nashville roadways (2019 vs 2020)
  - Most accidents happen at 5pm (rush hour), least happen at 6am
  - Friday and Saturday are the busiest (and most accident-prone) days to be driving!
- **Weather Implications**
  - Most accidents happen in 20-30 degrees Celsius
  - Counterintuitively, lower visibility does *not* lead to more accidents
    - Perhaps because Nashville has good visibility year-round and/or drivers are more careful when visibility is low, since this is when congestion is *highest*
  - Response Time inv. prop. to precipitation; prop. to temperature & visibility
- **Location-Specific Implications**
  - Certain areas in Nashville are more prone to incidents
  - **Downtown** is ¼ the size of **East Nashville**, but has twice as many accidents
  - In **Vanderbilt**, the first semester of school (months) have the highest incident frequency
  - During summer vacation months and winter break, **BNA** has the most accidents
  - In the afternoon hours (12-2pm) and after 7pm, **Downtown** has the most accidents
    - Lunch break? College nightlife?



# References

- [1] Learning Incident Prediction Models Over Large Geographical Areas for Emergency Response Systems
- [2] A Review of Incident Prediction, Resource Allocation, and Dispatch Models for Emergency Management