

# IST 718: Big Data Analytics

Traffic Accident Analytics

*Holden Herrell, Karima Hunter, and Vaibhav Mehta*

# TOPICS

---

Background

Data acquisition

Data Prep/Cleansing

Next Steps

# Background

---

**Roles:** City Planners for Charlotte, NC

**Business Question:** Should we approve the construction of a stadium for a third major professional sports team (baseball) in Charlotte, NC?

**Concern:** Will adding a third stadium negatively impact traffic accidents?

# Data Acquisition

---

## Traffic Accidents

**Source:** Kaggle

**Description:** 3M traffic accident from 2016 to 2019. Datasets contains accidents for 49 States

**Key Fields:**

- Location
- Time
- Severity
- Weather Conditions

## Sporting Schedule

**Source:** Various\*

**Description:** Contains sporting event data for NFL, NBA, MLB.

**Key Fields:**

- Game Location
- Time
- Attendance

\*baseball-reference.com

\*pro-football-reference.com

\*basketball-reference.com

# Data clean up

---

## Traffic Accidents:

- Data typing
- Remove Nulls/NAs
- Subset data to target market

## Sporting Dataset:

- Import and consolidate individual datasets
- Utilize game schedules to retrieve attendance values

Cities	NFL	MLB	NBA
Houston	Texans	Astros	Rockets
Atlanta	Falcons	Braves	Hawks
Miami	Dolphins	Marlins	Heat
Dallas	Cowboys	Rangers	Mavericks

# Next Steps

---

- Exploratory Data Analysis
- Traffic Accident correlation analysis to identify relevant variables
- Implement methodology to join datasets

# Accident Breakdown

Frequency Distribution of US-Accidents (December 2019)  
(Hover for breakdown)

