**Real-Time Traffic Incident Reports:**

* A collection of traffic incidents from Sep 26, 2017 to Sep 11, 2018
* Rows: 66,6K
* Fields:
  + Traffic Report ID
  + Published Date
  + Issue Reported
  + Location
  + Latitude
  + Longitude
  + Address
  + Status
  + Status Date

**Travel Sensors:**

* A collection of all the Bluetooth sensors in Austin with the earliest being activated in Jan 2012
* Rows: 193
* Fields:
  + Reader ID
  + ATD Sensor ID
  + KITS ID
  + ATD Location ID
  + Modified Date
  + Sensor Status
  + Turn On Date
  + Sensor Type
  + COA Intersection ID
  + Primary St Segment ID
  + Cross St Segment ID
  + Landmark
  + Primary St AKA
  + Cross St AKA
  + Signal Eng Area
  + Council District
  + Jurisdiction
  + Location Type
  + Location Name
  + Primary St
  + Cross St
  + Primary St Block
  + Cross St Block
  + Location
  + IP Comm Status
  + Comm Status Datetime UTC
  + Location Latitude
  + Location Longitude

**Individual Address Files:**

* Information about individual devices detected starting in Jan 2016
* Rows: 81.9 M
* Fields:
  + Record ID
  + Host Read Time
  + Fields Device Read Time
  + Reader Identifier
  + Device Address

**Individual Traffic Match Files:**

* Information about individual trips as recorded by two separate sensors
* Rows: 39M
* Fields:
  + Record ID
  + Device Address
  + Origin Reader Identifier
  + Destination Reader Identifier
  + Travel Time Seconds
  + Speed (Miles per Hour)
  + Match Validity
  + Filter Identifier
  + Start time
  + End Time
  + Day Of Week

**Traffic Match Summary Records:**

* Aggregate data about travel times on roadways
* Rows: 12M
* Fields:
  + Record ID
  + Origin Reader Identifier
  + Destination Reader Identifier
  + Origin Roadway
  + Origin Cross St
  + Origin Direction
  + Destination Roadway
  + Destination Cross St
  + Destination Direction
  + Segment Length Miles
  + Timestamp
  + Average Travel Time Seconds
  + Average Speed MPH
  + Summary Interval Minutes
  + Number Samples
  + Standard Deviation

**Proposed Aggregate Data Point:**

* My own collection of traffic incidents with data from the nearest sensor
* Expected Rows: 66K
* Fields:
  + Traffic Report ID
  + Published Date
  + Issue
  + Location
  + Nearest Sensor
  + Speed at Nearest Sensor
  + Day of Week
  + Device Address
  + Distance to Nearest Sensor
  + Time Difference to Nearest Reading
  + Average speed at Nearest Sensor
  + Standard Deviation of Speeds at Nearest Sensor

Possible Problems:

1. At what speed is an accident most likely to occur?
2. Do traffic rates and traffic incidents vary by day of week?
3. Does higher variance in traffic speeds correlate to more accidents?