Collision Avoidance Systems Comparative Report:

Assessing the Value and Efficacy of Sensor Systems in Standardised Vehicles

APRIL 2017

TEAM BUMBLEBEE – DAPT 2018
Virginia Commonwealth University – School of Business
ENGLERT, McCONNELL, RAMIREZ, PATEL

Executive Summary

DriveFactor, a CCC company, recruited our team to investigate an inquiry into the efficacy of vehicle sensor systems. To manage the scope of this inquiry, our team pursued the following analysis into Collision-avoidance Systems (CAS's). Between 2008 and 2016, there have been 85 specific new-model emergences of CAS- standardised vehicles on the market. Within the past two years, Preventative systems have emerged as the new standard in collision-avoidance. To determine the effect of this emerging standard on the number and severity of collisions, our team designed the following study using the following three vehicle-related data-sets: (1) the Polk data set gives us a monthly count of MakesModel&Year registered per state, and (2) the VDot data gives us the total number of daily collisions and their severity in Virginia, but without MakeModel&Year specified. (3) Finally, the DriveFactor data reveals the number and cost of collisions, daily, per state, by VIN number, from which we can derive Make, Model and Year. We hypothesise that by combining these three data-sets, DriveFactor can better inform decision-weights that determine the cost-benefit of Collision Avoidance sensor systems for their insurance customers. DriveFactor can also produce a ranking of CAS-standard vehicles most effective at collision-avoidance, which may benefit both customers and manufacturers.

Beginning with the Polk data, we derived the total number of model_years per state to achieve a count of total vehicle registrations between 2008-2016. For each identified CAS-standard model, we derived a total count of model_years registered in Virginia between 2008-2016. We also calculated a total count for each parent make to represent a totalMake population for 2008-2016. Then, from the VDoT data, we derived a total number of collisions and their severity per year between 2010 and 2015 (the range of the data-set). For each year, we can depict how the number of collisions has changed in relationship with the total number of CAS-standard model years, derived from the Polk data. The DriveFactor repair-cost data provides an essential link between these data sets – each model_year can now be assessed directly with the number and severity of collisions.

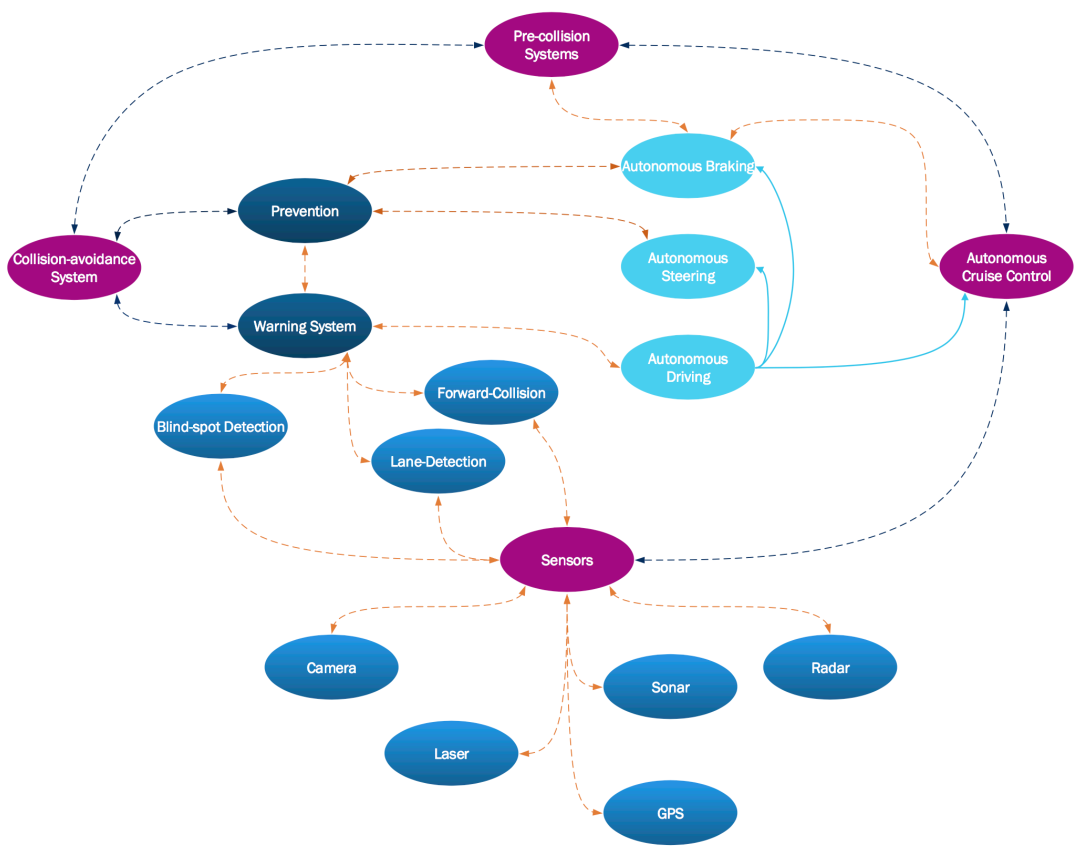
From this exploratory analysis, our team was able to determine that yes, there is a meaningful relationship between CAS standard vehicles and the reduced number of collisions. The annual percent difference in collision totals shows a significant decrease in TotalCrashes after 2013. This corresponds with what we know about the simultaneous total number of CAS-standard warning systems entering the total vehicle pool during this time. Further, certain crash CollisionTypes measure greater percent differences than others. Those most strongly decreasing between 2010 and 2015 include (1)Angle, (2) Rear-end and (3)Head-on collisions. This produces strong evidence for the case of Collision-avoidance systems affecting TotalCrash severity. With DriveFactor data, our study's control-limits can be extended with confidence to include all states and year 2016 collisions. We propose the application of DriveFactor data to both test and refine our initial assessment assumptions, ultimately to produce a prioritised list of sensor systems according to their efficacy in avoiding collisions.

Table of Contents

Vehicle Sensor Systems

Evolving across the span of the last two decades, vehicle sensors have increased in not only complexity and capability, but also in their interdependence as integrated sensor systems (Figure 1). The function of Pre-collision Systems is to reduce the impact of collisions on the vehicle's occupants. Sensing first through driver feedback in the form of break application, and eventually through assimilation into advanced collision-avoidance warning systems, Pre-collision Systems activate when the probability of collision is high, tensioning seat-belts and pressurising break-systems. The catalytic development of Autonomous Cruise Control spurred the innovation of autonomous breaking, first at high and, eventually, low speeds. Simultaneously, three distinct sensor groupings triggered a manufacturing arm's race to ultimate Collision-avoidance System sophistication. Finally, the last two years have seen the bridging of both Collision-avoidance Prevention and Autonomous Cruise Control in the potentially revolutionary emergence of semi-Autonomous Driving.

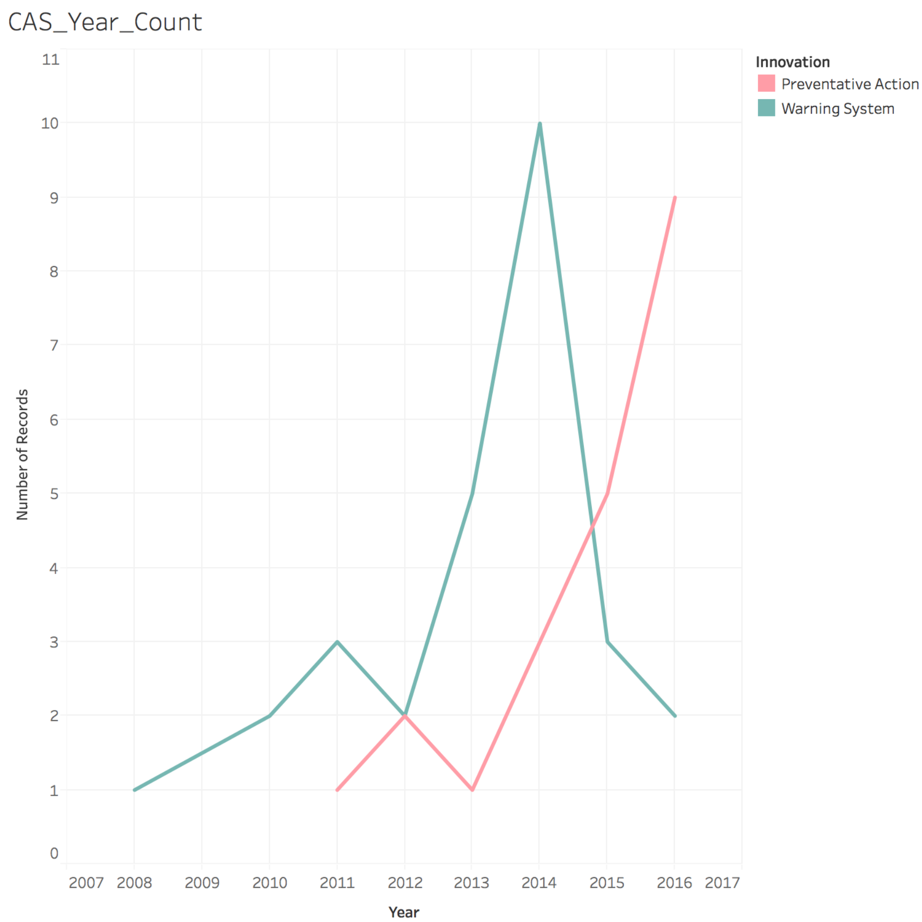
Figure 1



Collision-avoidance Systems

Our researchers group Collision Avoidance Systems into two primary capabilities, with three and two sub-groupings, respectively. Collision Avoidance systems may be (1) Warning and/or (2) Preventative systems. Warning systems provide direct feedback to the driver using visual, auditory or textural alerts, or a combination of these queues. Warning systems rely on sensor-information from (a) Front-collision Avoidance, (b) Lane-keeping, and/or (c) Blind-spot detection systems to predict the high-likelihood of impact and to trigger the alert. Preventative systems build upon the capabilities of Warning systems with the addition of (a) Brake and/or (b) Steering Control. Although often pared with Warning systems, Preventative systems may or may not *alert* the driver before taking autonomous control of braking or steering. Both Warning and Preventative systems translate sensor-information into an action intended to prevent collisions. Since 2014, Preventative systems have emerged as the new standard in collision-avoidance (Figure 2).

Figure 2



La tendance du somme de Number of Records pour le Year1. La couleur affiche des détails associés au/à la Innovation.

Customer:

Founded on 6 September 2011, DriveFactor is a leader in insurance telematics, offering powerful, flexible technology and solutions. The team has built its technology and apps from the ground, up for insurance, and applies their expertise to help clients customize telematics solutions to meet their short- and long-term needs. Integrated into CCC Information Services, Inc in May 2012, the DriveFactor platform is built on a dynamic, scalable architecture that accepts data from numerous input methods and enables easy sharing with business partners. As an industry leader in claims technology and services, CCC works with more than 350 insurance carriers, more than 21,000 repair facilities, and hundreds of other business partners, processing over one million claims-related transactions per day. Together CCC and DriveFactor created the first full-service, open telematics platform for insurance.

Data Points:

In addition to the vehicle-repair data expected from DriveFactor, our team structured the following inquiry testing the comparative efficacy of specific Collision-avoidance Systems (CAS's) using several data sources that include:

- **Polk Registration Data**[1] **–** from which we derive yearly totals of VIN-year makes_models registered
- **Virginia Department of Transportation (VDoT), Collision data –** from which we derive yearly totals of crashes by severity (non-Vin/make_model specific)
- **Insurance Institute for Highway Safety(IIHS), Crash Avoidance Features data –** a public data set from which we identify vehicle make_model_years with standardised CAS; for this study, we exclude both "Adaptive Headlights", as these have become quite common across manufacturers, and some less-abundant vehicle makes (Figure 3)
- **DriveFactor Data –** per make_model_years identified in the IIHS data, our team requested from DriveFactor a healthy data-set reflecting both the number of collisions and their severity for both the specific VIN make_ years of the models identified. This set is therefore designed to also include models very similar to the CAS-standard models for comparison to CAS-standard vehicles within a particular make

[1] R. L. Polk & Company is a provider of automotive information and marketing solutions to the automotive industry, insurance companies, and related businesses. In 2016, Polk provided the DAPT with a vast dataset from which we derive yearly totals of VIN-year makes_models registered each month, across many states.

Figure 3

Year1 pour chaque Make Model. La couleur affiche des détails associés au/à la Innovation.

Data Resources:

In anticipation of receiving data from DriveFactor, our team developed a cloud environment in which to securely host, manage, query and share datasets amongst all the DAPT 2018 academic teams. After much research and experimentation (read "trial and error"), our team successfully established a Relational Database Server (RDS) instance, hosted by Amazon Web Services (AWS). Our team chose to use MySQL for the database we developed because it is a well-supported, open-source relational database. MySQL also comes with the advantage of MySQL Workbench, a database tool that allows multiple users to connect to the Database Server.

The team tested the functionality of the MySQL Server by loading the Polk dataset into the cloud environment. This was an excellent initial test of our team's Database Server because the Polk dataset is very large (>15GB), and is very awkward to manipulate and share across individual computers. Microsoft Excel opens spreadsheets in an all-at-once fashion, and thus cannot support larger file-sizes such as the Polk dataset. Our team found that using a Database Server was the best way to collaborate as a team in real-time.

The team used Python to scrub the Polk data and parse into manageable file sizes. When the team initially tried to load data into the MySQL server, various errors were encountered. Eventually through trial-and-error it was discovered that there were several instances of the "/" symbol scattered throughout the dataset. Example; Ford Bronco 2D/AWD. The presence of the "/" symbol in the dataset was causing the SQL import statements to fail. We used Python to remove all of the "/" characters from our data.

Another major roadblock was the size of the Polk dataset. If we attempted to load the ~10GB file into the MySQL Server in one shot, AWS would require ~30GBs of space. This is 10GB over the free tier limit. The team also discovered that files exceeding 1GB can cause MySQL to act unpredictably. To solve these issues the team again used Python to break up the large file into several smaller "chunks" to load into the MySQL database.

Table 1: AWS Data Load Time

250 Million Records	Approximate # of records in the Polk dataset
5 Records	# of records per second imported to the database with MySQL Import Wizard
30 Records	# of records per minute imported to the database with MySQL Import Wizard
18,000 Records	# of records per hour imported to the database with MySQL Import Wizard
~13,888 Hours	# of hours required to load entire Polk dataset
~83 Weeks	# of hours required to load / hours in a week

However, even after the Polk dataset was broken up into smaller chunks, we still encountered issues loading the data into the MySQL server. The team's initial attempts were to load the data into the server using the MySQL Workbench Import Wizard. The data load times using MySQL were unacceptable for larger data files, as described in Table 1.

The team decided that DriveFactor would likely not be happy with waiting 83 weeks for data to load, so the team researched alternatives. Eventually the team came across an open-source SQL tool called HeidiSQL. HeidiSQL was able to connect to the MySQL Server and enable inline bulk data imports (versus importing one record at a time). HeidiSQL reduced the loading time of data to 20-minutes per ~1GB chunk. This allowed the team to cut the load time from >1 year to less than a day.

For more information, both our AWS/RDS process and our python script are included in *Appendix A*.

Hypothesis:

Our team developed the following hypothesis to guide the comparative assessment of Collision-avoidance systems:

> *If we examine the functional systems within collision-avoidance system classes, we will find differing levels of efficacy in preventing both the number of crashes and the severity of crashes. This finding is important because the ability to measure efficacy of systems will lend to a prioritised ranking of safety systems, potentially optimised from insurance, consumer and manufacturer perspectives.*

From the Polk data, we can derive the number of model_years per state. We can therefore use Polk data to achieve a count of CAS-standard models between 2007-2016. For each model_year, we are able to derive a percentage of the total number of model_years registered in Virginia. From the VDoT data, we can derive a total number of collisions and their severity per year. Thus, for each year, we can depict how the number of collisions has changed in relationship with the total number of CAS-standard model_years, as derived from the Polk data.

The DriveFactor repair-cost data provides an essential link between these data sets, allowing each model_year can be assessed directly with the number and severity of collisions. The ratio of the number of CAS-collisions and total collisions can further be broken down into CAS-standard sub-groups, thus informing the efficacy and value of each sub-group. We therefore use this data to both test and validate our initial assessment assumptions, and ultimately produce a prioritised list of sensor systems according to their efficacy in avoiding collisions.

Method:

Using both the Business Analytics Book of Knowledge (BABOK) and Cross Industry Standard Process for Data Mining (CRISP) frameworks, our team defined a methodology to support an iterative practice of defining needs and recommending solutions driving competitive change for our customer. Our adopted methodology can be articulated as six primary steps, each with supporting tasks, that together energise a knowledge life-cycle within an enterprise (*Appendix B*). Our first three sub-processes function to establish both a discovery and a need. In pursuing both, we hypothesise and validate the value of pursuing an analytic inquiry into vehicle safety-sensor systems. The final three sub-processes function to align identified requirements and business strategy to the development of a solution capability, and to gather analytic feedback to test and refine that capability over time. It is the view of this research team that analytic maturity within an enterprise is an essential means by which companies innovate competitive solutions to increase their industry advantage.

In the first, *Planning and Engagement*, our team established relationships with DriveFactor stakeholders and classmates, outlined its business analysis approach, and established secure and shared platforms for information and data management. Our second step defines the process of *Elicitation and Collaboration*. Within this step, the CRISP method served as an outline for conducting our primary inquiry into the underlying question of safety-sensor efficacy. Our team prepared for elicitation by researching vehicle sensor systems, developing a testable hypothesis and submitting a formal request for DriveFactor data. We conducted an exploratory analysis using pre-established data sets to confirm and refine our hypothesis.[2] Throughout this second step, our team worked diligently to maintain stakeholder collaboration through consistent communication efforts. Next, our team defined a process of *Requirements Lifecycle Management*, which, when combined with the formal analytic investigation of safety-sensor efficacy, identifies, defines and prioritises the requirements driving our customer's business processes.

In step four, *Strategy Analysis*, the analytics team works with stakeholders to align analytics with requirements in formal business strategy. Here, the team assesses the current state, defines a potential future state, assesses risks and benefits of adopting a strategic change, and defines a change strategy. In *Requirements Analysis and Design*, the analytic team further investigates a solution space by verifying requirements and validating materiel capabilities in quantifiable terms. During this phase, the team may assess one or several proof-of-concept implementations in order to generate data-driven design options and to assess potential business value. Ultimately, the team integrates these results to prescribe a solution for implementation. In our final step, the team returns to evaluate the solution's performance and limitations, as well as the performance and limitations of the enterprise in its application of the solution. Here, the team generates a prescriptive analysis of actions to increase the solution's value to the enterprise.

Elicitation and Collaboration:

During the Elicitation and Collaboration phase, our team researched vehicle sensor systems to develop a hypothesis. Because we understood that the DriveFactor make_model information could only be obtained from vehicle VIN numbers, we also researched a comprehensive list of vehicles to query for the study. While awaiting data, our team then constructed a secure cloud-based environment in which to host the incoming data. Finally, we performed an exploratory analysis of pre-existing data sets to investigate the hopefulness of our hypothesis: *is there really an indication that CAS-standard vehicles are affecting the numbers and/or severity of collisions?* For this result, we

[2] Currently, as we are without access to DriveFactor data, we are unable to proceed beyond this task to the execution of a formal analysis validating our initial hypothesis and exploratory analysis results to our proposed CAS-prioritisation deliverable.

encourage you to proceed through the following brief sections.

Data Understanding

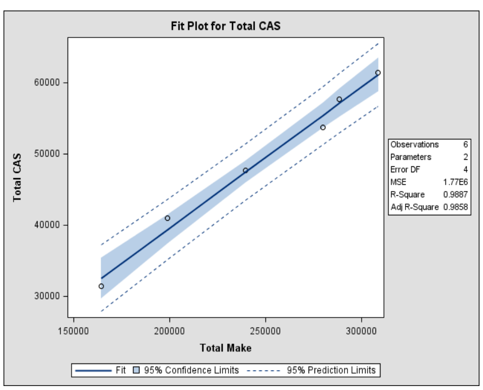
Using the Polk-registrations-per-month-per-state data and the collisions-per-day-within-Virginia data sets, our analysis team tested the validity of the following four assumptions:

1) There is a relationship between the total number of CAS-standard vehicles and the total number of accidents
2) There is a relationship between the total number per Make and the total number of CAS-standard vehicles
3) There is a reduction in accident severity, which can be measured by the reduced number of accidents with injuries and, especially, fatalities as the number of CAS-standard vehicles increases
4) There is a shift in the dominant types of collisions occurring, for which type of intersection and collision types might be good monitors

Exploratory Linear Regression

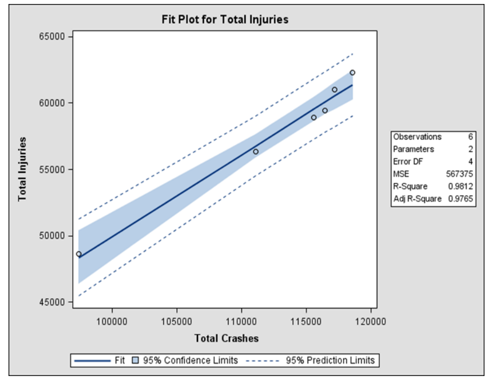
To test the high-level assumption that there is a correlative relationship between the changing number of CAS-standard vehicles and the changing number of collisions per year, our team performed a series of exploratory linear regression analyses. Of these, two sets stand out as most notable. The first of these include the following: (1) Linear Regressions assessing the relationships between CAS vs. Make (Figure 4) and Crashes vs. Injuries (Figure 5). Both Make and CAS appear highly correlated with the other as denoted by the value of the squared correlation coefficient of 0.9887. This value indicates that ~99% of the variance in Make is explained by CAS (and vice versa). Both Crashes and Injuries appear closely correlated with the other as denoted by the value of the squared correlation coefficient of 0.9812. This means that ~98% of the variance in Crashes is explained by Injuries (and vice versa).

Figure 4



The conclusion we draw from this is that either variable in either relationship may be substituted for the other variable of their respective relationships. In the case of TotalMake vs. TotalCAS, this might actually be a problem. Here we see nearly identical relationships when we assess the influence of vehicle type on total crashes: as TotalMake increases, TotalCrahes decrease; as TotalCAS increases, TotalCrashes decrease.

Figure 5



So,then... Which relationship is most impacting crashes?

This question leads us to the second discovery of interest: TotalMake and TotalCAS are both predictors of TotalCrashes and TotalInjuries. To predict total crashes, either variable may be sufficient. The relationship between TotalMake and TotalCrashes is nearly the same as the relationship between TotalCAS and Total Crashes. This can be read in two ways: (1) as the total number of vehicles increase, the total number and severity of collisions decrease; (2) as the number of CAS-standard vehicles increase, the number and severity of collisions decrease. The first does not make intuitive sense whereas the second sounds potentially more likely to be true and produces a slightly stronger R^2 (squared correlation coefficient). The slightly better correlation coefficient in (R) indicates that Total CAS is a slightly better fit at explaining the variance in Total Crashes. However, to understand the effect of CAS-standard vehicles in the total vehicle pool on the number and severity of collisions, <u>we need to collect more evidence</u> of a shifting pattern corresponding to CAS-standard vehicle market emergence.

Figure 6

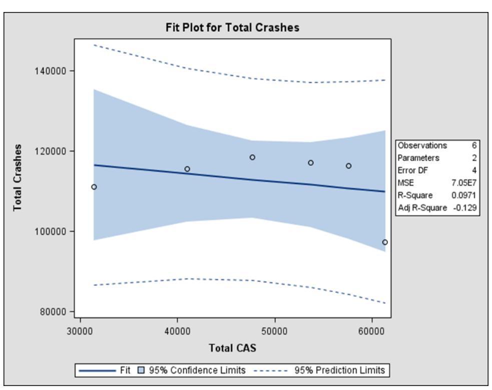
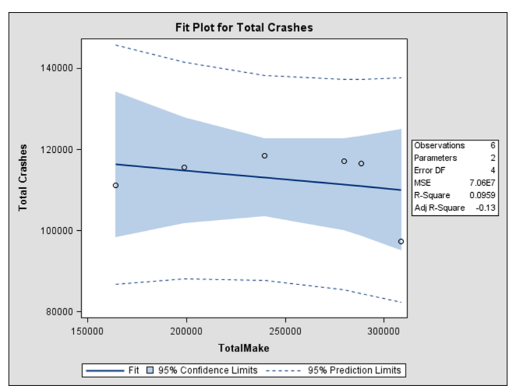


Figure 7



Findings

The team used MS Excel, Tableau and SAS to analyse data and create data visualizations. The following are the data visualizations that we feel strongly support our hypothesis. The reader will find the full analysis presented in *Appendix C*.

CAS-standard Vehicle Ratio

Between 2014-2015, total crashes decreased significantly. This corresponds with what we know about the simultaneous total number of CAS-standard warning systems in the total vehicle pool during this time.

Collision Type

Those most strongly decreasing between 2010 and 2015 include (1) Angle, (2) Rear-end and (3)Head-on collisions. This produces strong evidence for the case of Collision-avoidance systems affecting Total Crash severity

Intersections

Most collisions do not occur at intersections; between 2010-2015, our team finds the most significant decrease in the (1) Not-at-intersection category. Of collisions at intersections, those most strongly decreasing between 2010 and 2015 include (2) Two-way and (3) Four-way approaches.

Injury Severity

Overall, injuries are decreasing in both severity and count

- Fatalities for Other Animal, Rear-end, and Pedestrian collisions have all decreased
- Side-swipe fatalities have increased, particular in same-direction collisions
- Backed into injuries have increased in Incapacitating injury severity while decreased in non-visible injury severity

Potential Recommendations:

We recommend that DriveFactor make use of state-reported DoT collisions data in conjunction with the data that DriveFactor currently possesses. DriveFactor collision totals may be assessed against DoT data collision counts as percentages of a total collision population. Collision and severity types within this transportation data provide a vital piece of the puzzle. A quarterly-published, prioritised ranking of vehicle safety vs repair-cost trade-off will benefit both vehicle buyers as well as manufacturers while also establishing DriveFactor as an industry leader. To understand the effect of CAS-standard vehicles in the total vehicle pool on the number and severity of collisions, we recommend that DriveFactor collect more data to help provide evidence of a shifting pattern corresponding to CAS emergence.

Appendix A

Create Amazon Web Services (AWS) Relational Database Serves (RDS) Instance

1) **Step 0 – Navigate to AWS landing page**
    a. Create an account and sign in
    b. Select the RDS option (relational database) in the Database cluster
    c. Click on the blue "getting started" button.

2) **Step 1 - Select Engine.**
    a. Click the "Free tier eligible only".
    b. For this example I chose MySQL Community Edition

3) **Step 2 - Specify DB Details.**
    a. Instance Specifications
        i. DB Instance Class -> db.t2.micro - 1 vCPU, 1 GB RAM option (only free tier option available)
        ii. Allocated Storage -> 20 GB
    b. Settings
        i. DB Instance Identifier: sandboxvcu
        ii. Master Username: tommy
        iii. Master Password: bumblebee
        iv. Confirm Password: bumblebee

4) **Step 3 - Configure Advance Settings**
    a. Network & Security
        i. Note that the "Create new Security Group" is the default for VPC Security Group(s).
        ii. This is fine for now, but we will need to modify these settings later so that other users can access the database.
        iii. Eventually we will configure our security settings to allow allow connections from:
            1. HTTPS (Port 443)
            2. HTTP (Port 80)
            3. SSH (Port 22)
            4. MySQL/Aurora Port 3306
    b. Database Options
        i. Database Name: sandboxvcu
    c. Backup
        i. Backup Retention Period: Set to zero.
            1. This will disable automated backups.
            2. This setting can be modified after we import data into our database instance.
    d. "Launch DB Instance"
        i. Your AWS db instance is now being created

ii. Click on the blue "View Your DB Instances" to check out your new instance.

5) **RDS dashboard**
   a. You are now located in the RDS dashboard where we can view all of the details for this and other DB instances created
   b. Recommend you copy & paste the EndPoint details into a text document.
   c. This is the address you will use to connect to your DB server.
      i. In this case our EndPoint is: sandboxvcu.cvuz5ugobij2.us-east-1.rds.amazonaws.com:3306
   d. Whenever an application is asking you for a "Host" name they are just asking for the EndPoint information without the port details ':3306'
      i. So "Host" would be sandboxvcu.cvuz5ugobij2.us-east-1.rds.amazonaws.com
   e. If you click on the icon that has paper and a magnifying glass you can see all of the details regarding your DB instance that you may need to refer to.

6) **Launch MySQL Workbench**
   a. Click on the "+" symbol in the upper left-hand corner to create a new connection.
      i. The Setup New Connection window will appear
      ii. Connection Name: sandboxvcu (you can name this connection whatever you want)
   b. Parameters
      i. Hostname: sandboxvcu.cvuz5ugobij2.us-east-1.rds.amazonaws.com
      ii. Port: 3306
      iii. Username: tommy
      iv. If you want: Click on store in vault and enter your password.
   c. Click Test Connection.
      i. If you didn't store password in vault you will now be prompted for your password.
      ii. Hopefully a "Successfully made the MySQL connection" alert pops up. Click ok. Click okay again.
      iii. The sandboxvcu connection you setup should now be visible on the MySQL main view.
   d. Click on the gray square titled vcusandbox to connect to the database.

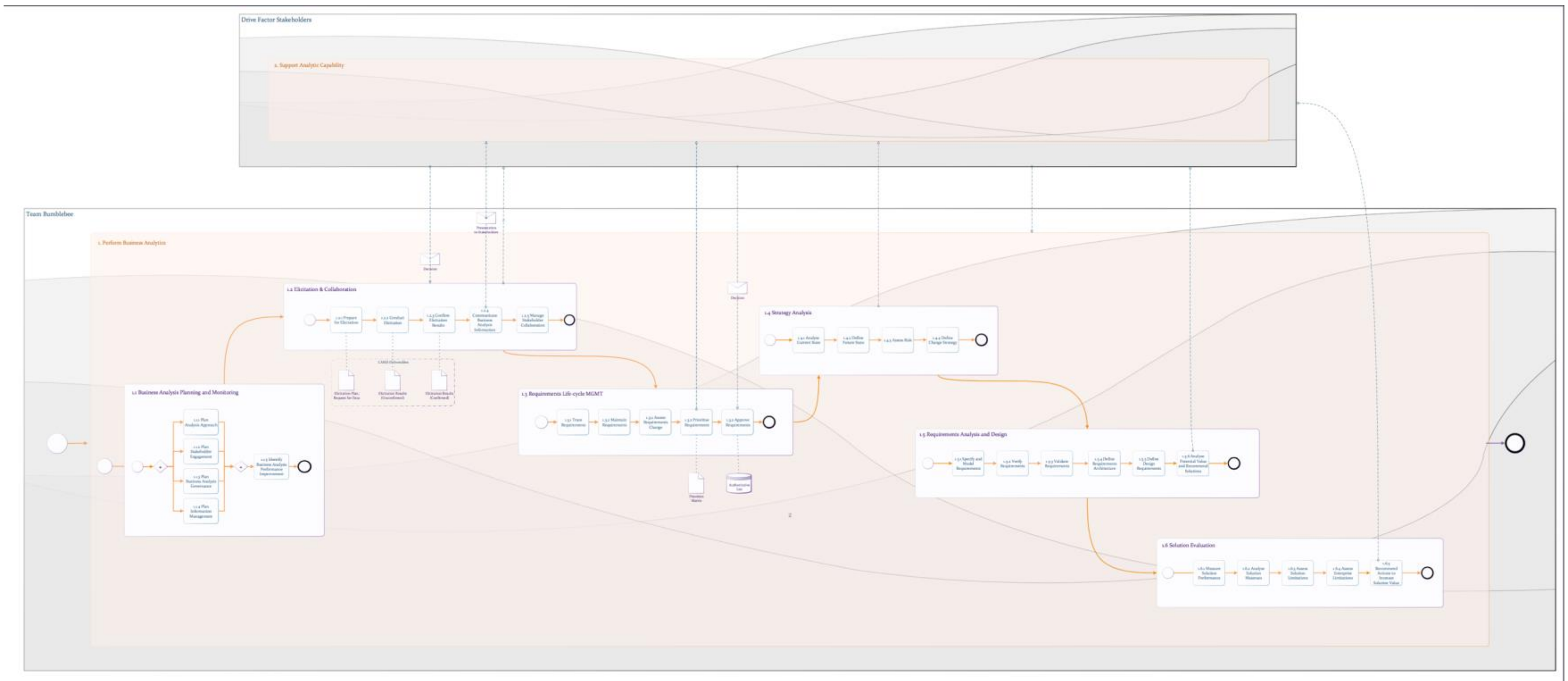
7) **Importing data into the MySQL Server**
   a. Download HeidiSQL @ https://www.heidisql.com/
   b. Establish a connection to the MySQL server using the same info we used to connect with MySQL Workbench.
   c. This link provides a screenshot of the data import tool https://www.heidisql.com/screenshots.php?which=import_textfile
   d. HeidiSQL can be used to quickly import our data into the MySQL server.

Python Script: Scrub and Chunk Large DataSet for Easy(-ier) Upload

```
    #Python 2.7

import pandas as pd

    #specify the file to read from and the file to write to

csvfileIn = open("C:/Users/file/Path/file.csv", "r")
csvOut = open("C:/Users/file/Path/file_Scrubbed.csv", "w")

    #remove "/" from the data

for i, row in enumerate(csvfileIn):
    row = row.replace("/", "_")
    csvOut.write(row)

    #close both files. Data will not appear until you close the file

csvOut.close()
csvfileIn.close()

    #chunk the data into manageable sizes

for i,chunk in
enumerate(pd.read_csv("C:/Users/file/Path/file_Scrubbed.csv",
chunksize=15000000)):

chunk.to_csv("C:/Users/file/Path/file_Scrubbed_Chunk{}.csv".format(i))
```

Appendix B

Plan Business Analytic Process Supporting Continuous Innovation



Appendix C

Exploratory Analysis Findings

Data Exploration

Number of Registrations vs Number of Collisions and Severity, by Make

Hypothesis

Collision avoidance systems increase accident prevention and decrease accident severity

Purpose

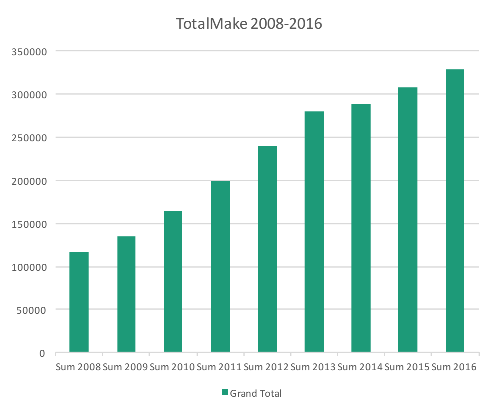
- Understanding the relationship between Collision-avoidance Systems (**CASnumber**) to **CollisionNumber** and **CollisionSeverity** is essential to an understanding of the VALUE of CAS-standard vehicles
- Insurance companies (decision weights in rate-determining models)
  - Value of accident avoidance
  - Cost of repairs
- Vehicles operators (value-in-safety correspondence to vehicle cost)
  - Occupant safety
  - Cost to own/insure/repair
- Manufacturers (maintaining competitive advantage, brand loyalty)
  - Demographic targeting
  - Customer Trust

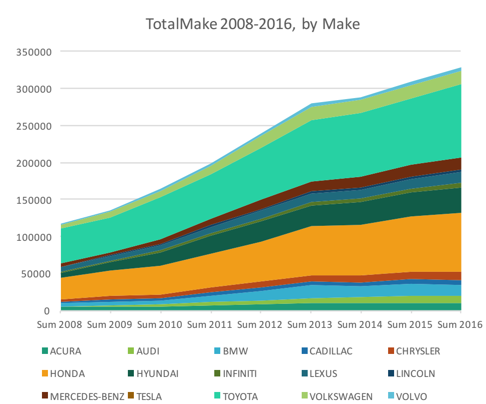
We expect to see:

1) A **relationship** between the total number of CAS-standard vehicles and the total number of accidents
2) A **relationship** between the total number per Make and the total number of CAS-standard vehicles
3) A **reduction** in accident severity, which can be measured by the reduced number of accidents with injuries and, especially, fatalities as the number of CAS-standard vehicles increases
4) A **shift** in the dominant types of collisions occurring, for which type of intersection and collision types might be good monitors

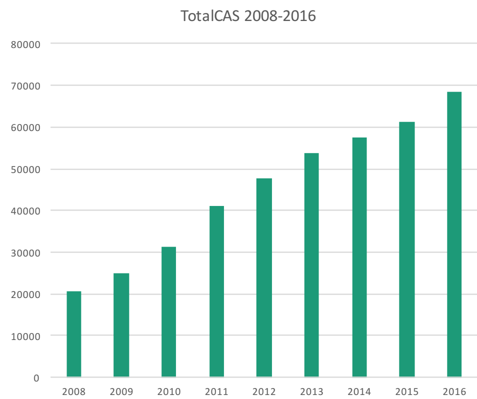
Findings: Relationships

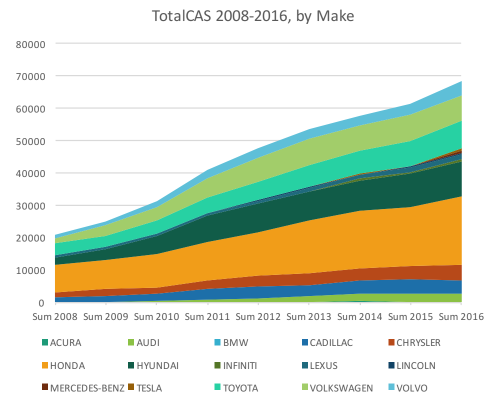
Total Make



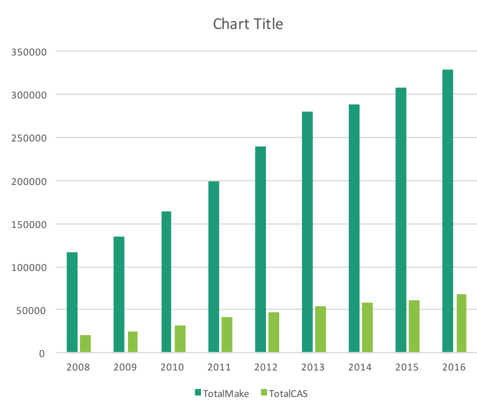


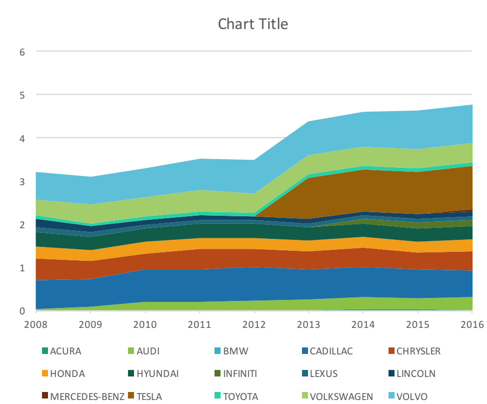
TotalCAS



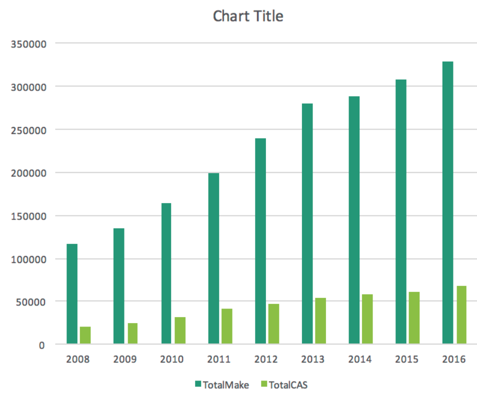


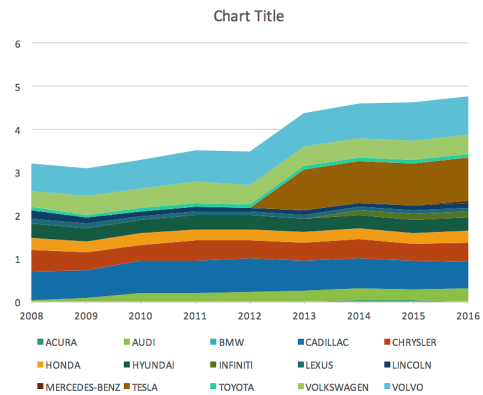
TotalMake vs TotalCAS



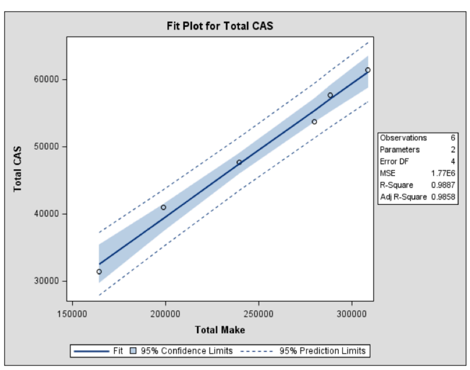


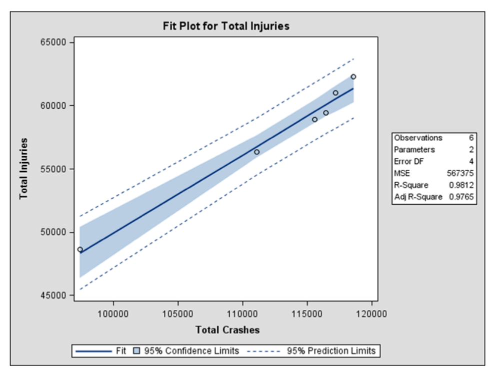
TotalMake vs TotalCAS



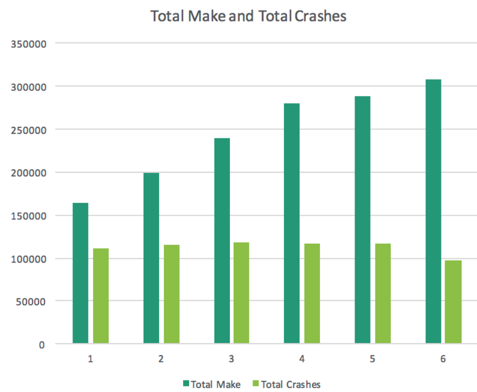


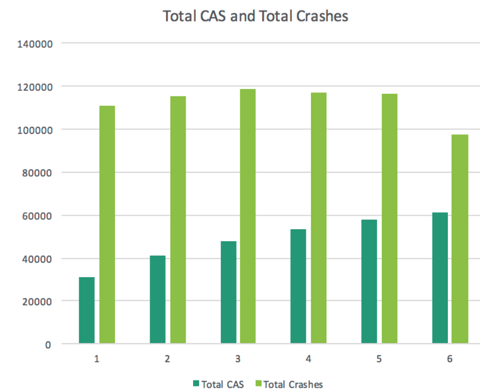
Totals-to-Parts Relationships



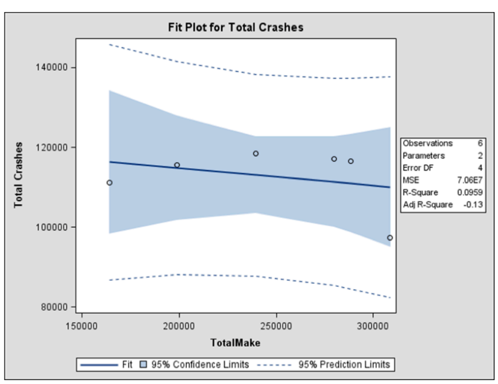


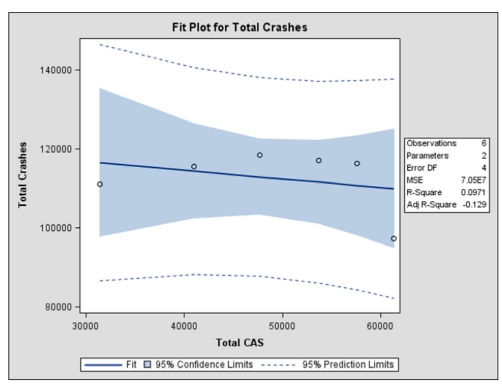
Crashes: TotalMake vs TotalCAS





Crashes and Injuries to TotalCAS





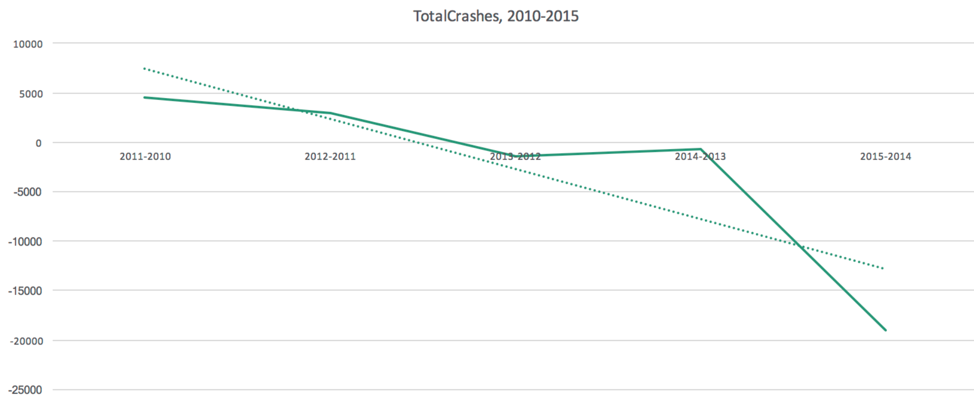
Discussion

- There is a marked increase in the percent frequency of CAS-standard vehicles represented between 2012-2014, which corresponds with the emergence of Tesla-brand CAS-standard S-models
- TotalMake and TotalCAS are both predictors of TotalCrashes and TotalInjuries
  - To predict total crashes, either variable is sufficient
  - The relationship between TotalMake and TotalCrashes is nearly the same as the relationship between TotalCAS and Total Crashes
- This can be read in two ways:
  - As the total number of vehicles increase, the total number and severity of collisions decrease
    - This does not make intuitive sense,
  - As the number of CAS-standard vehicles increase, the number and severity of collisions decrease
    - This makes more sense and produces a slightly stronger R^2 (squared correlation coefficient)

Findings: Effects

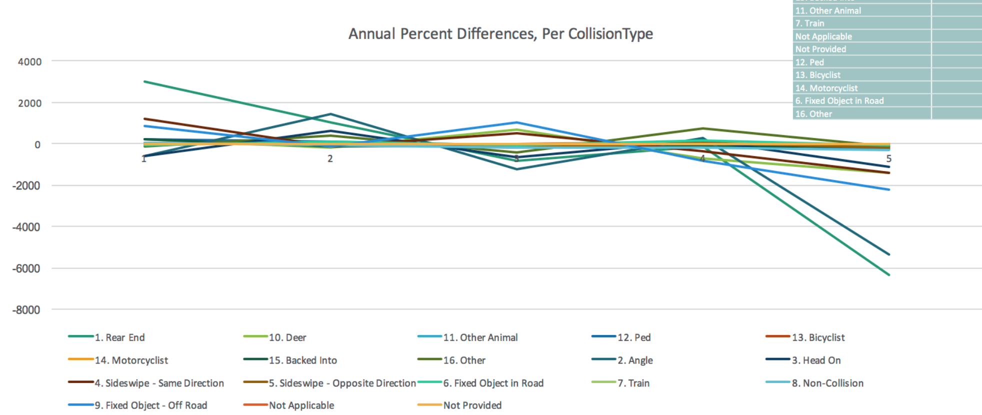
To understand the effect of CAS-standard vehicles in the total vehicle pool on the number and severity of collisions, we need to collect more evidence of a shifting pattern corresponding to CAS emergence

Percent Change of TotalCrashes…

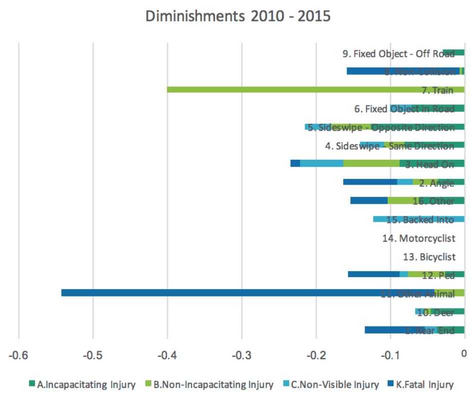


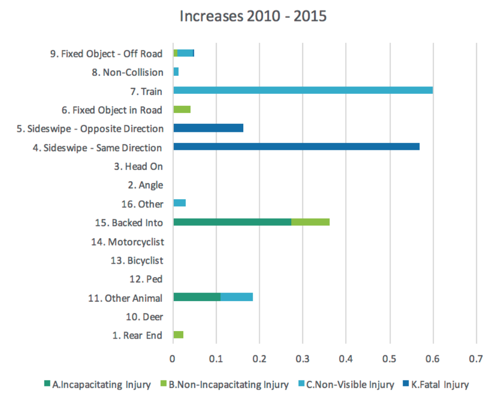
…by Collision Types

CollisionType	2015-2010
2. Angle	-5411
1. Rear End	-3277
3. Head On	-1494
10. Deer	-1374
9. Fixed Object - Off Road	-1259
8. Non-Collision	-681
5. Sideswipe - Opposite Direction	-266
4. Sideswipe - Same Direction	-219
15. Backed Into	-181
11. Other Animal	-21
7. Train	-3
Not Applicable	-1
Not Provided	0
12. Ped	1
13. Bicyclist	3
14. Motorcyclist	4
6. Fixed Object in Road	37
16. Other	475



Injuiries:Crashes Ratio, by Severity





Discussion:

- The annual percent difference in collision totals shows a significant decrease in TotalCrashes after 2013
  - This corresponds with what we know about the simultaneous total number of CAS-standard warning systems entering the total vehicle pool during this time
- Certain crash CollisionTypes measure greater percent differences than others
  - Those most strongly decreasing between 2010 and 2015 include (1)Angle, (2) Rear-end and (3)Head-on collisions
  - This produces strong evidence for the case of Collision-avoidance systems affecting Total Crash severity
- Overall, injuries are decreasing in both severity and count
  - Fatalities for Other Animal, Rear-end, and Pedestrian collisions have all decreased
  - Side-swipe fatalities have increased, particular in same-direction collisions
  - Backed into injuries have increased in Incapacitating injury severity while decreased in non-visible injury severity