Crash Trends and Locations

(City of Cambridge, MA)
(2010-2016)

ALY 6070 - Communication and Visual Data Analytics

Professor Jack Bergersen
Northeastern University

Proposed by:
Pratibha Sridhar Iyengar
Siddhant Sanjiv Sapte
Rahul Tarachand Rewani

Introduction

The following white paper describes further analysis and insights on the Crash trends dataset of the city of Cambridge, Massachusetts. The dataset includes distinct variables about the crashes that occurred from the year 2010 to 2016.
In this white paper, we have conducted research in two parts

- Crash Summary Report- Part 1
  This report summarizes the trend of crashes in Cambridge on an Year by Year basis, drills-down the monthly crashes for the highest crash-occurrence year. It also analyzes the relation between the elements involved in the accidents and the most affected objects in crash. From these reports we can understand the trend and check if the crashes are seasonal and which categories of vehicles cause accidents.
- Crash Summary report- Part 2
  The report is used for Street-wise analysis of the data. Here we can analyze, for each street, the intensity of crash-occurrences in terms of number of crashes and also analyze whether the busiest streets are also the most accident prone.

The Dataset and Data Cleansing

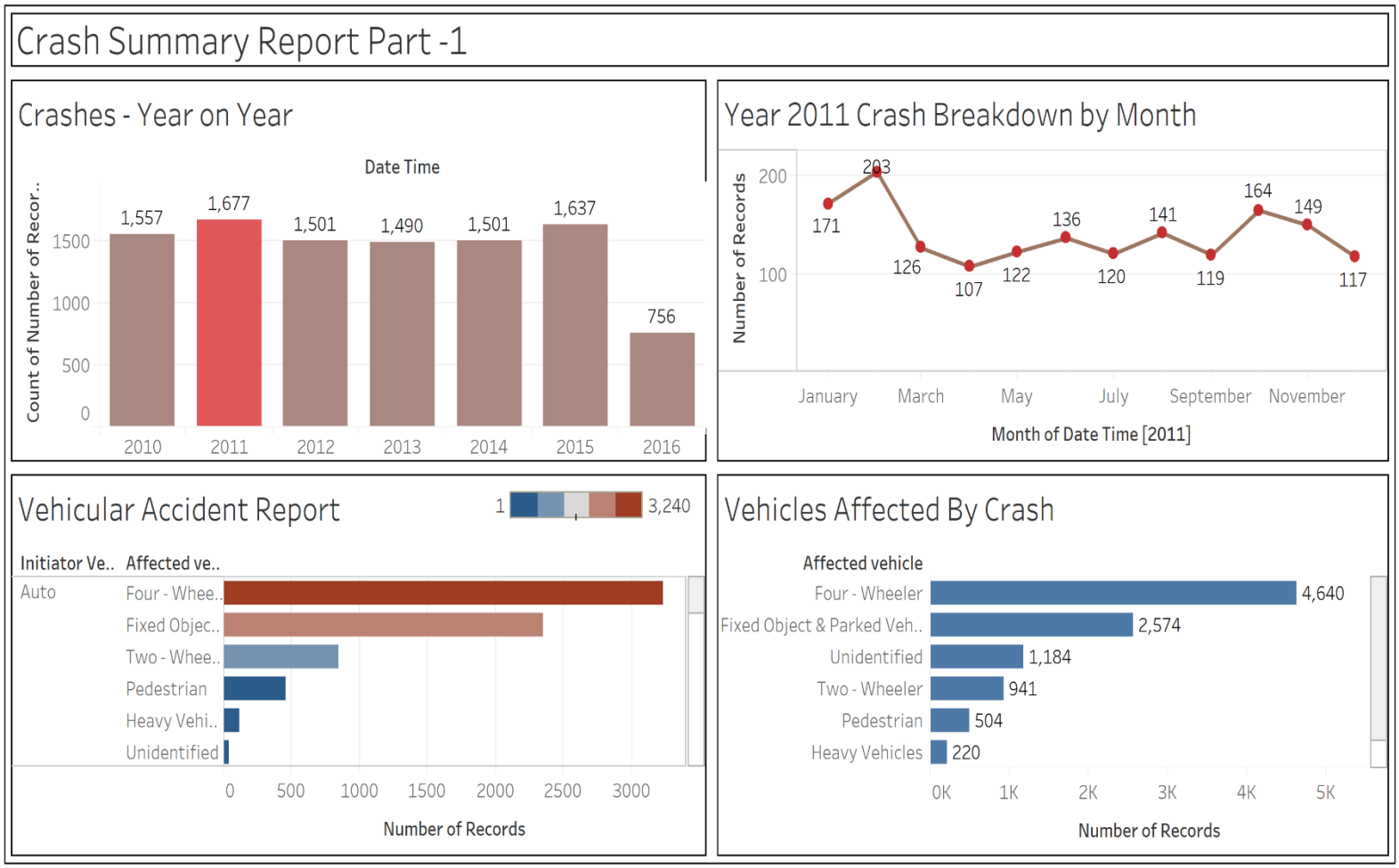
The datasets used are

- Cambridge Police Department Crash Historical Dataset (Jan 2010-Jun 2016)
  The Police Department Crash dataset provides historical information about the registered accidents in Cambridge between 2010 and 2016 through crash number, date and time of incident, objects(elements) involved in the crash, the location and co-ordinates of the location. The location is broken down as Street number, street name and cross street. Since the data may have been collected from different sources, the data in street names is very inconsistent in terms of cases and formats referring to the same street. For example, Street Name has multiple values pointing to the same street- “Mass Ave”, ”Massachusetts Avenue”, “Massachusetts Ave”, “MASs Ave”.
  In order to make this data consistent, we used VLOOKUP() function in MS Excel Worksheet and saved the result in a new column “Street1 New”.
  This step will help reduce redundancy while plotting charts and graphs in Tableau.
- Average Daily Traffic Data of Cambridge (1972-2017)

The Average Daily Traffic Data of Cambridge provides average yearly traffic data for streets in Cambridge between 1972 and 2017. The data provides information about the streets, the average Latitude and Longitude, traffic (in numeric) for data available for each year and Peak AM and Peak PM traffic. The dataset provided was very clean but inconsistent with our primary data source, the crash data. In order to bring integrity to make joins possible for analysis, the street name in this dataset was mapped with the Police Department Crash Data “Street1 New” column using VLOOKUP() function in MS Excel and the result was stored in new column “StreetMap”. The unmatched streets were stored as it available in original dataset.
Post these steps, in order to handle case-sensitivity, after add data source connections in Tableau, two calculated fields, named “STREET” in both cases, was created for “Street1 New” and “StreetMap” with values in uppercase using UPPER() function in Tableau.
This way, the data integrity was achieved and the relationship between the two datasets will be more consistent.

Crash Summary Report – Part 1

The above Dashboard accommodates four graphs. The first two graphs



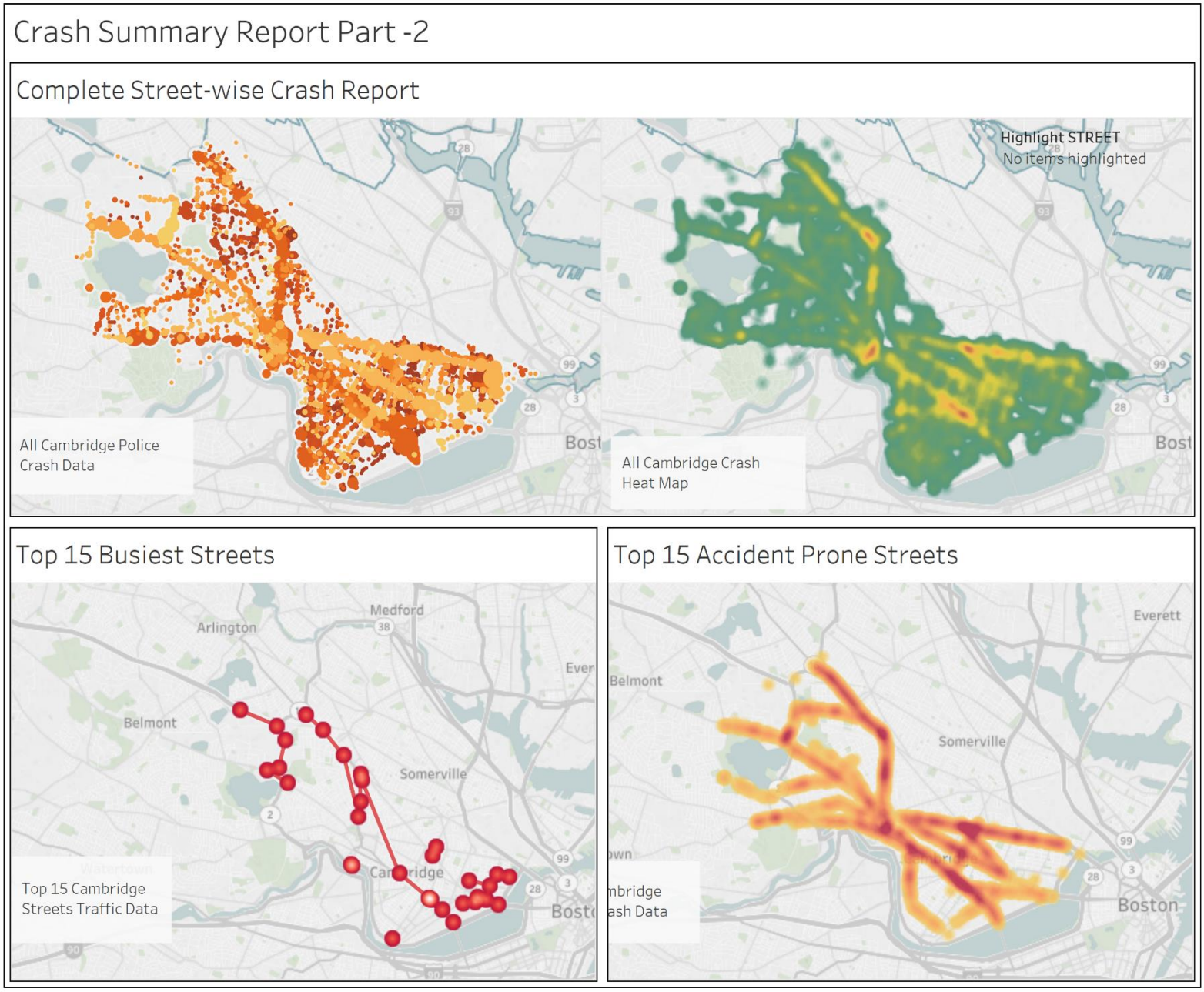
share the same colour palette as both the graphs are interrelated. The first graph titled as “Crashes - Year on Year” represents the number of crashes that occurred each year from 2010 to 2016 and the second graph titled as “Year 2011 Crash Breakdown by Month” represents the further breakdown of the year 2011 which saw the most number of crashes. The line chart shows the monthly trend of accidents in the year 2011. This trend seems to be seasonal with the highest number of accidents occurring in February which is generally peak winter time in Cambridge, Massachusetts and also the range of accidents lies approximately between 125-205 covering both fall and winter months from September through March.

The third graph titled as “Vehicular Accident Report” represents a summary report of objects in Vehicular accidents. It shows the relation between each

accident-initiating vehicle labelled as “Initiator” with each corresponding affected vehicle. The 4th graph titled as “Vehicles affected by Crash” represents a report on trend of vehicles affected by the crash. From the graph, it was found that the type of vehicles affected the most by the crashes were, Four-wheeler vehicles with 4640 incidents reported which includes Auto, Passenger cars, Taxi, Light Trucks, Van. As there were too many similar kind of vehicles, we grouped them into major categories like “Four wheelers”, “Fixed objects and Parked Vehicles”, “Bus”, “Heavy Vehicles”, “Pedestrians”, “Unidentified Objects” and “Two wheelers”.

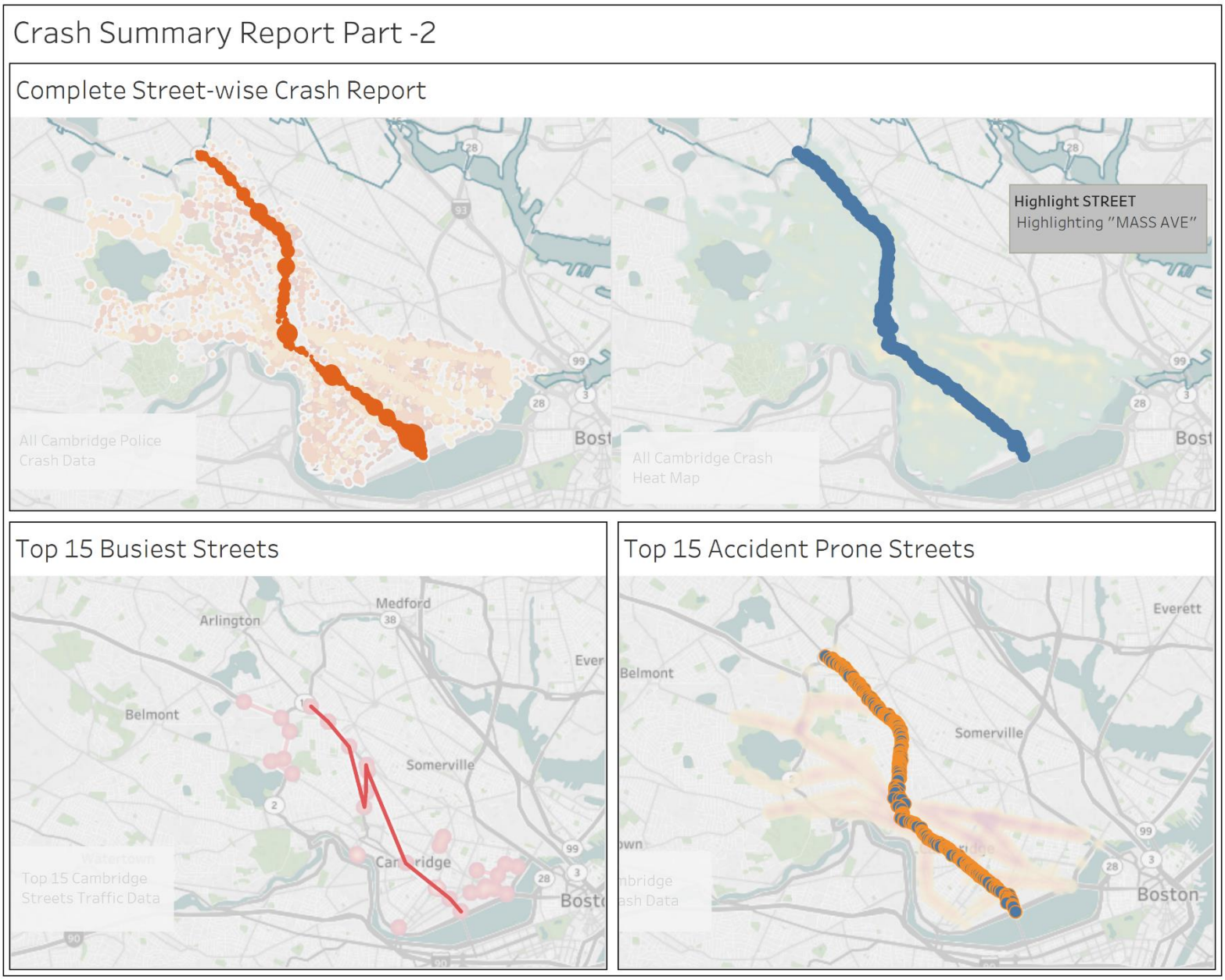
Crash Summary Report – Part 2

The above dashboard contains three graphs namely, “Complete Street-



wise Crash Report”, “Top 15 Busiest Streets”, “Top 15 Accident Prone Streets”. The reason these three graphs have been assembled together is that they indicate incidents based on the street locations. The Complete Street-wise Crash Report represents a thorough analysis of all the crashes occurred in each street in the city of Cambridge, Massachusetts from the year 2010 through 2016. The first map plots location-wise crashes and the

second map is the corresponding Heat-map. Since there are many streets and locations involved in this data, we have included a ‘Highlighter’ to analyze crashes and traffic on a desired street.



Furthermore, the “Top 15 Busiest Streets” and the “Top 15 Accident Prone Streets” represent busiest streets in the City of Cambridge, Massachusetts with “Massachusetts Avenue” being the busiest and the most Accident prone in both the cases. From the graph, we can also determine if the busiest street was the most accident-prone street.

Conclusion

In a nutshell, it was found that from the year 2010 through 2016, the number of accidents reported in the year 2011 were 1677 which was the highest. It was also found that the month of February in that particular year had the most number of crashes i.e. 203. The Average number of accidents in the Year 2011 was approximately 5 accidents a day. Also, four-wheelers were the most affected kind of vehicles in the crashes.

Moreover, from the report “Crash Summary Report - II” we can find out that the busiest street in the city was Massachusetts Avenue and it was also the most accident prone street. The reason why it was the busiest and the most accident prone is, it is the longest street stretched across the city of Boston and Cambridge.

References:

- Instantmagazine.com. (2019). *2018 Ultimate Guide: How to Write and Format a White Paper*. [online] Available at: https://www.instantmagazine.com/topics/how-to-write-and-format-a-white-paper [Accessed 28 Jan. 2019]
- Cullen, M. (2019). *How to Write and Format a White Paper: The Definitive Guide (2018)*. [online] Instructionalsolutions.com. Available at: https://www.instructionalsolutions.com/blog/how-to-write-white-paper [Accessed 28 Jan. 2019]
- Cambridgema.gov. (2019). [online] Available at: https://www.cambridgema.gov/~/media/Files/policedepartment/SpecialReports/CrashAnalysisReport_Final_05252017.pdf [Accessed 28 Jan. 2019]
- "Average Daily Traffic Counts Map | Open Data." City of Cambridge. Accessed January 29, 2019. https://data.cambridgema.gov/Traffic-Parking-and-Transportation/Average-Daily-Traffic-Counts-Map/tdg2-57af