

BI tools



BI Platforms



- Exploratory Data Analysis
 - An easy way to do basic data exploration on databases/datasets
- More advanced analytics possible
 - But it becomes problematic
 - Run scripts in R or Python
 - Output is focused on visualization
- Dashboards!
 - Fast overview of the most important metrics
 - Interactive
 - Set once and forget
 - Excellent tool to communicate with business/customers
- Tableau - probably most popular

Data Sources

- Wide variety of data sources
 - Local files (excel, csv, etc.)
 - Relational databases
 - Cloud dbs (Big Query, Redshift, etc.)
 - Hadoop dbs (Hive, etc.)
- Data is accessed either as live connection or as an extract
 - Extract can be scheduled to be refreshed periodically (every hour, day, etc.)
 - For live connection results seem to be cached for a while

Variables

- Each column is considered as a variable
- There are two main groups of variables
 - Dimension - typically discrete values, cannot be aggregated (e.g. text)
 - Measure - numerical values (can be aggregated sum, max, avg, etc.)
- ... and multiple types (date, text, boolean, numerical, etc.)
- Tableau tries to guess variable type
 - You can reassign the type by right clicking on the

Dimensions	
Abc	Intersection Name
	Latitude
Abc	Location
	Longitude
#	Year
Abc	Measure Names
Measures	
#	Animal Action
#	Backed Unsafely
#	Cyclist Error/Violation
#	Driverless Vehicle
#	Failed To Observe Traffic Si...

Shelves

- Tableau is a drag and drop system
- Dragging pills (variables) to the shelves

The screenshot shows the Tableau interface with the following components:

- Columns Shelf:** Contains a pill for 'Measure Names'.
- Rows Shelf:** Contains a pill for 'Intersection Name'.
- Marks Shelf:** Set to 'Automatic'.
- Measure Values Area:** Contains a pill for 'Measure Values'.

The main view displays a table titled 'Fatalities by intersection' with the following data:

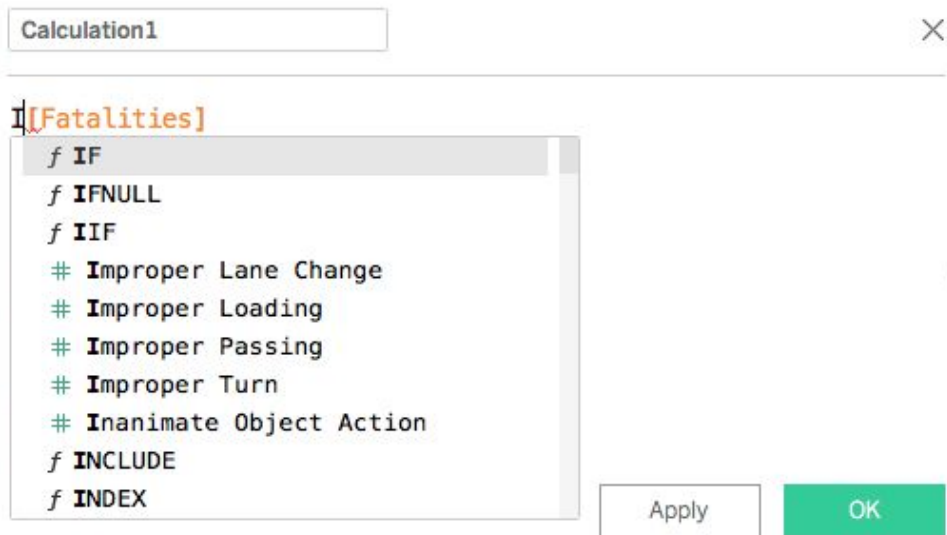
Intersection Name	Avg. To..	Avg. Fa..	Avg. Ma..	Avg. Mi..
YELLOWHEAD TRAIL & 127 STREET NW	83	0	0	13.5
107 AVENUE & 142 STREET NW	80	0	1	9.5
YELLOWHEAD TRAIL & 149 STREET NW	76.5	0	1.5	20.5
118 AVENUE & GROAT ROAD NW	66	0	0	8.5
90 AVENUE & 85 STREET NW	63	0	0	7
23 AVENUE & 91 STREET NW	62.5	0	2.5	10
87 AVENUE & 178 STREET NW	58	0	1.5	9
87 AVENUE & 170 STREET NW	56	0	1	14
34 AVENUE & 91 STREET NW	52.5	0	3	11
137 AVENUE & 97 STREET NW	103	0	1	11
137 AVENUE & 127 STREET NW	50	0	0	10.5
23 AVENUE & 111 STREET NW	48.5	0.5	0.5	7.5
137 AVENUE & 50 STREET NW	47.5	0	2	7.5

Shelves

- **Row & Column** - defines how data is displayed
- **Pages** - creates pagination
 - Instead of displaying multiple plots for discrete values it displays one value at a time
 - Can be used to create animation
- **Filters** - defines which data is being displayed
- **Marks** - defines how data is displayed
 - What should be used to define color/shape
 - What information should be in tooltip/label
- More information:
https://onlinehelp.tableau.com/current/pro/desktop/en-us/buildmanual_shelves.htm

Calculated Fields

- Often a metric needs to be calculated based on underlying data
 - E.g. Conversion Rate = $\text{SUM}([\text{Transactions}]) / \text{SUM}([\text{Sessions}])$
- Tableau enables to use a lot of logic and aggregation functions
- Access other measures/dimensions using square brackets []
 - You can use other calculated fields in calculation
- To create calculated field right click on measures/dimensions and select Create -> Calculated Field...



Visualizations

- For any measures dropped on row/column shelves Tableau will suggest some visualization
 - Typically it will be either table with raw text or some sort of bar plot
- Various visualizations are available under “Show Me” menu
- To create more fancy visualizations data of specific type needs to be included
 - There is a hint what needs to be included once you hover over visualization



For maps try

1 geo  Dimension

0 or more Dimensions

0 or 1 Measure

May use spatial measure in place of geo dimension

Dashboards

- So far we were talking about worksheets
 - Single view that covers a single visualization
- Dashboards are a place to display multiple worksheets at once
- To create dashboard you drag and drop sheets on it
- Sheets on dashboard can be connected to make dashboard interactive
- There is also an option to create a data story
 - You can drag and drop sheets and dashboards
 - Save the state of sheet/dashboard
 - And add multiple story points (each point being a sheet or dashboard)

Dashboard Actions

- Sheets on dashboard may be connected with each other
 - Selecting one point on a map/scatter plot can update other sheets to reflect data specific to that point
- Action types: filter, highlight, go to URL
- Define source and target sheets
- Define specific event (hover, select, menu) that action should run on

Name:

Source Sheets

☐ Overview

☐ Fatalities by intersection

☒ Map

☐ Violation per year

☐ Violation type

Run action on:

☐ Run on single select only

Target Sheets

☐ Overview

☒ Fatalities by intersection

☐ Map

☐ Violation per year

☐ Violation type

Clearing the selection will:

☐ Leave the filter

☐ Show all values

☒ Exclude all values

Target Filters

☐ Selected Fields ☒ All Fields

Source Field	Target Field	Target Data Source

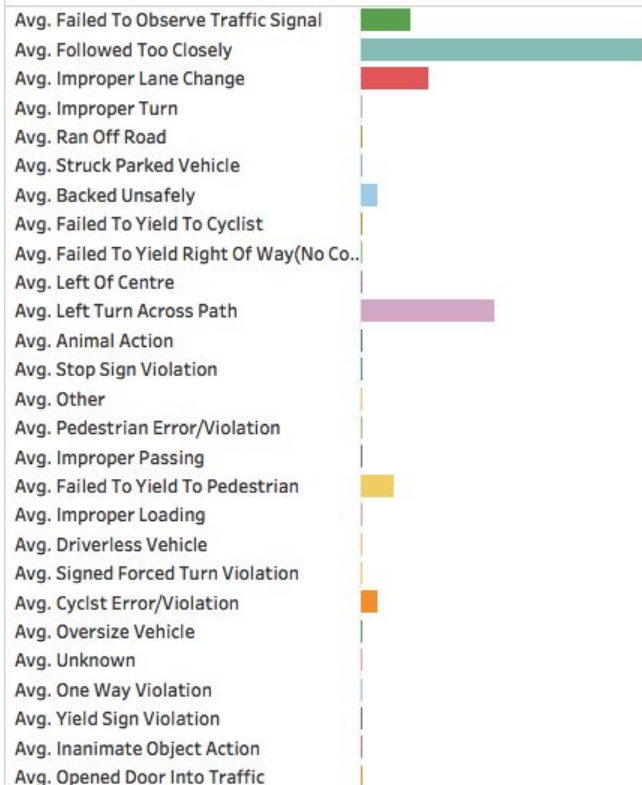
Other tools

- Domo
 - Quite expensive but also comprehensive solution, hosts data
- PowerBI
- Qlik
- Google Data Studio

Assignment - recreate a dashboard

- Open image with Tableau
 - Password is: BIA@CTC
- Open Tableau
- Open “Intersection_Collision_Summary__2010_-_2011_.csv”
 - This is an actual data from Edmonton Open Data Portal
- Play around and create some various visualizations to get familiar with interface
- Recreate the dashboard on next slide
 - Create calculated field Injuries or fatalities (sums minor and major injuries and fatalities)
 - Hovering over Fatalities by intersection should highlight point on the map
 - Selecting a point on the map should display data in Violation type and Violation per year
 - Change default colors to custom ones

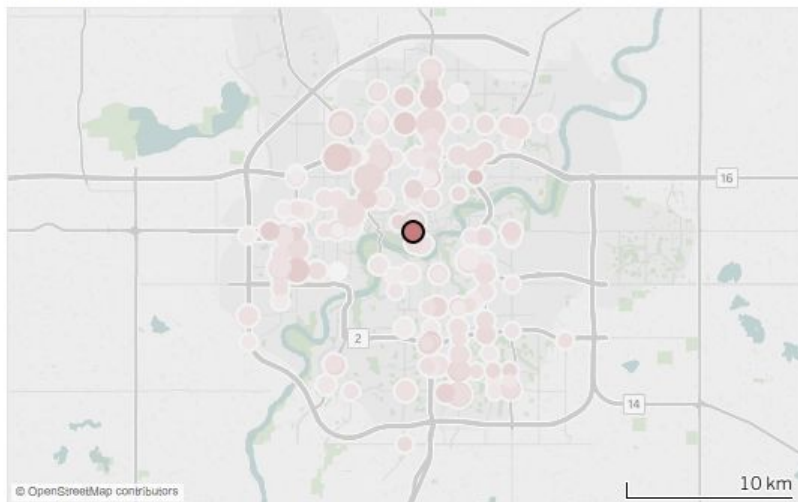
Violation type



Violation per year



Map



Fatalities by intersection

Intersection Name	Avg. To..	Avg. Fa..	Avg. Ma..	Avg. Mi..
153 AVENUE & 127 STREET NW	35.5	0	0.5	8
ARGYLL ROAD & 83 STREET NW	32.5	0	0	4.5
63 AVENUE & 91 STREET NW	32.5	0	0	7.5
STONY PLAIN ROAD & 149 STREET NW	32	0	0.5	5
82 AVENUE & 109 STREET NW	32	0	0.5	5
107 AVENUE & MAYFIELD ROAD NW	32	0	0.5	7.5
95 AVENUE & 170 STREET NW	63	0	0	11
23 AVENUE & PARSONS ROAD NW	31.5	0	0.5	10.5
153 AVENUE & 97 STREET NW	63	0	1	17
137 AVENUE & 66 STREET NW	31.5	0	0.5	8
97 AVENUE & 105 STREET NW	31	0	1.5	5.5