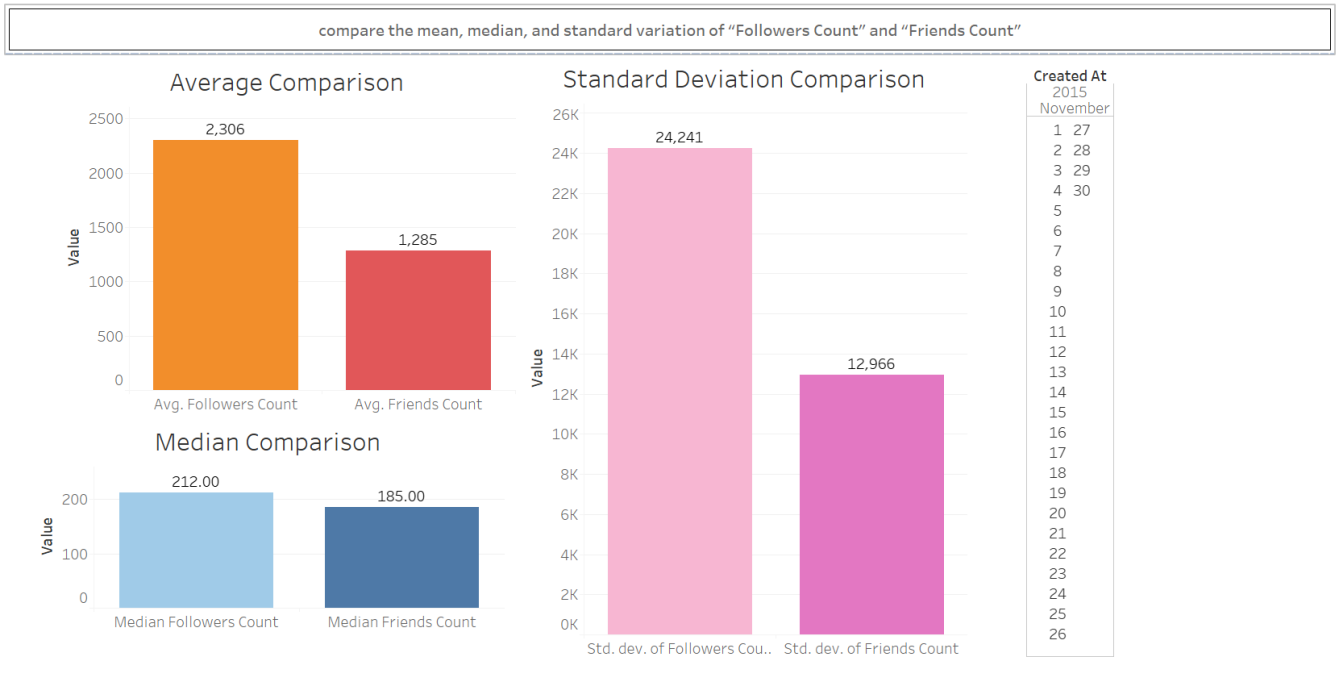
GEOG 594: Tableau Visualization Lab 3
Submitted By: Harshal Sanap
Dated 9/22/2018

Q1. Please use Tableau to compare the **mean, median, and standard variation** of "Followers Count" and "Friends Count" in graphs.

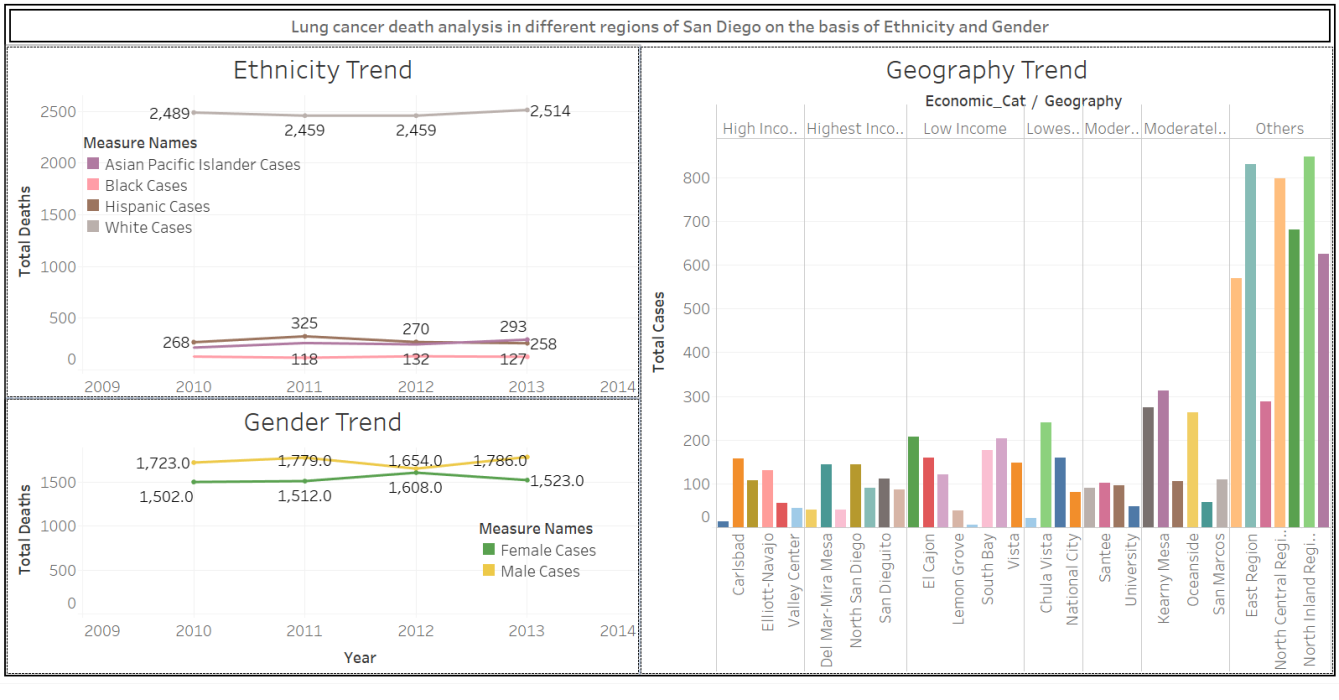
- The below plot helps us to visualize the difference between mean, median and SD for followers & friends count over the entire month of November as well as we can drill down on individual days of the month using the filter given on the right (Created At).



Q2. Please use Tableau to analyze the **Lung_Cancer_Death** data in the Web Exercise 02. Create a Dashboard to show your visualization results and include the Screenshot in the report. Write a short paragraph for each sheet to explain how you create each sheet in the dashboard.

Dashboard 1:

Description: The below dashboard can be used to check the trend of lung cancer deaths in different regions of San Diego along with ethnicity and gender breakdown for that death population.



GEOG 594: Tableau Visualization Lab 3
Submitted By: Harshal Sanap
Dated 9/22/2018

Sheet descriptions:

1. Geography Trend: In order to create this report I first created a hierarchy of economic category and geography. Hierarchy was created simply by right clicking on Economic_category variable and selecting hierarchy, then renamed it as economic_cat. Later I dragged the geography variable over this hierarchy variable thus creating an economic_category -> geography hierarchy. Later I dragged this hierarchy from Dimensions in columns section and total cases from Measures section to rows. Then I clicked on drill down or + sign of the hierarchy. Finally, to add more colors I dragged the geography variable over the color section of Marks.

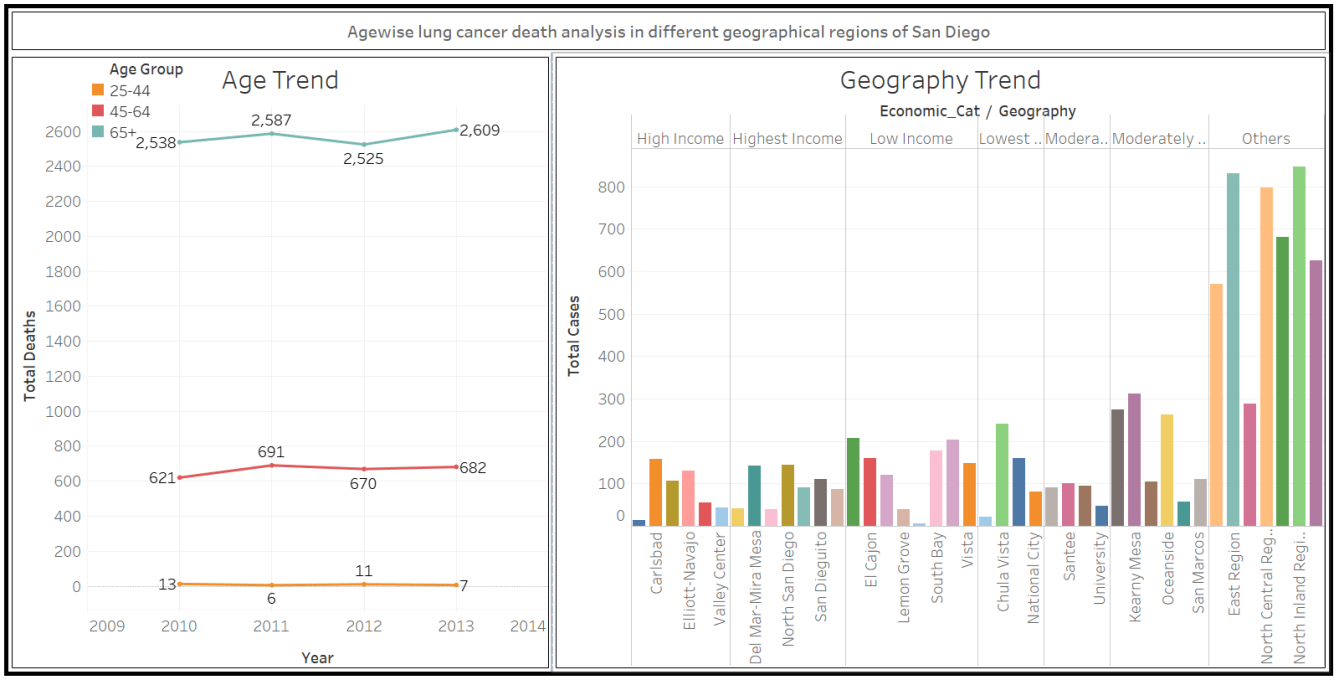
2. Ethnicity Trend: For this I dragged Year in columns then dragged measure names over color of Marks. Later filter the measure names for given 4 ethnicities. Later I dragged the measure values from measures to rows section to get the visual. Kept the default aggregation of the measure values as SUM. Finally, I dragged the measure values over Label of Marks to get the labels.

3. Gender Trend: For this I dragged Year in columns then dragged measure names over color of Marks. Later filter the measure names for given 2 genders. Later I dragged the measure values from measures to rows section to get the visual. Kept the default aggregation of the measure values as SUM. Finally, I dragged the measure values over Label of Marks to get the labels.

Dashboard 2:

Description:

The below dashboard gives age wise trend of total deaths occurred in different regions of San Diego.



Sheet Descriptions:

1. Geography Trend: In order to create this report I first created a hierarchy of economic category and geography. Hierarchy was created simply by right clicking on Economic_category variable and selecting hierarchy, then renamed it as economic_cat. Later I dragged the geography variable over this hierarchy variable thus creating a economic_category -> geography hierarchy. Later I dragged this hierarchy from Dimensions in columns section and total cases from Measures section to rows. Then I clicked on drill down

or + sign of the hierarchy. Finally, to add more colors I dragged the geography variable over the color section of Marks.

2. Age Trend: For this I dragged Year in columns then dragged measure names over color of Marks. Later filtered the measure names for given 3 age groups. Later I dragged the measure values from measures to rows section to get the visual. Kept the default aggregation of the measure values as SUM. Later, I dragged the measure values over Label of Marks to get the labels. Finally, I renamed the age groups as shown in the visual.

Q.3. Compare the functions between R and Tableau. What's their advantages and disadvantages for each?

Comparison Between R and Tableau:

R	Tableau
1. It is a programming language used mostly by statisticians and data science professionals to find patterns or insights in the data.	1. It is a software tool used by data analytics professional to better visualize their historical data.
2. In R, we can use predefined packages or libraries to make our job easy and fast. (Advantage)	2. Tableau provides a set of predefined functions that allow creation of calculated fields to make our job fast.
3. R is open source (Advantage)	3. Tableau is a software whose license is chargeable whereas students can get free access. (Disadvantage)
4. R provides flexibility to make our own visualizations the way we want. (Advantage)	4. In Tableau we need to stick to the predefined visuals and doesn't give any option to make user defined or custom visuals (Disadvantage)
5. R can be extensively used for predictive modeling (Advantage)	5. Tableau is not such a good tool for predictive though it is amazing for business intelligence or data analysis (Disadvantage)
6. We can not create dashboards using R (Disadvantage)	6. Tableau can be used for creating dashboards (Advantage)
7. Not so good with geographic data containing longitude and latitude information (Disadvantage)	7. Easy to interpret geographic data like longitude & latitude (Advantage)
8. Easy to interpret files with type JSON or XML(Advantage)	8. Difficult to interpret JSON or XML file types (Disadvantage)
9. R visuals can not be as good as that of Tableau	9. Tableau is apparently known as the one with most amazing visualizations in the market (Advantage)
10. Machine Learning functionalities can be implemented using R (Advantage)	10. Machine Learning can not be implemented though minor predictive modeling is possible with Tableau