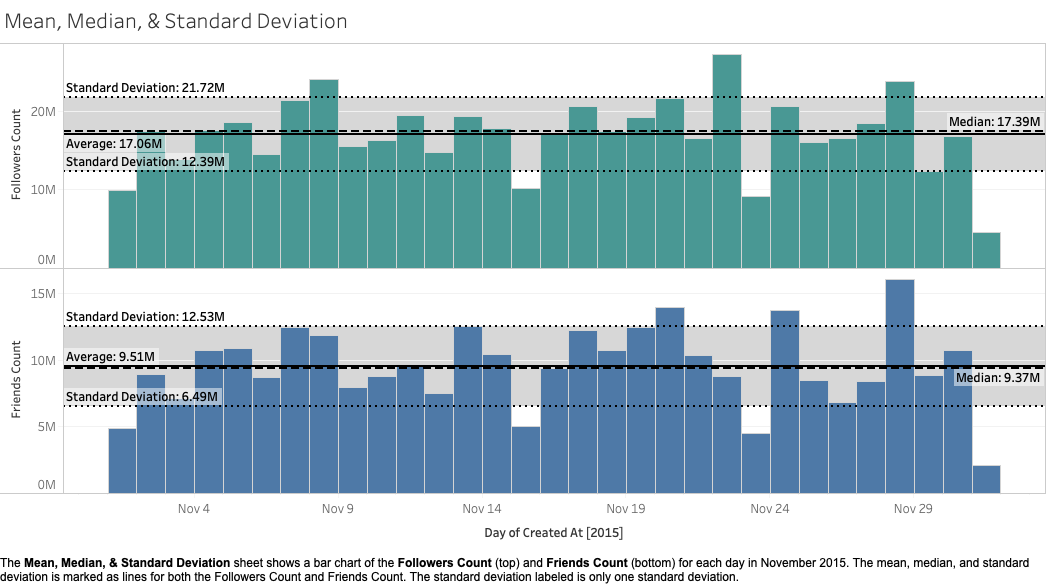
# Web Exercise 3: Tableau Visualization Exercise

1. Please use Tableau to compare the mean, median, and standard variation of “Followers Count” and “Friends Count” of SD\_alltweets\_Original excel file in graphs (Use the panel called “Analytics” next to the “Data” panel, then you can create the average, median, or standard variation lines in graphs). Write a short description with a screenshot of your results.

**Mean, Median, & Standard Deviation.png**



The caption of the image is the description of the screenshot. Here’s what the caption displays:

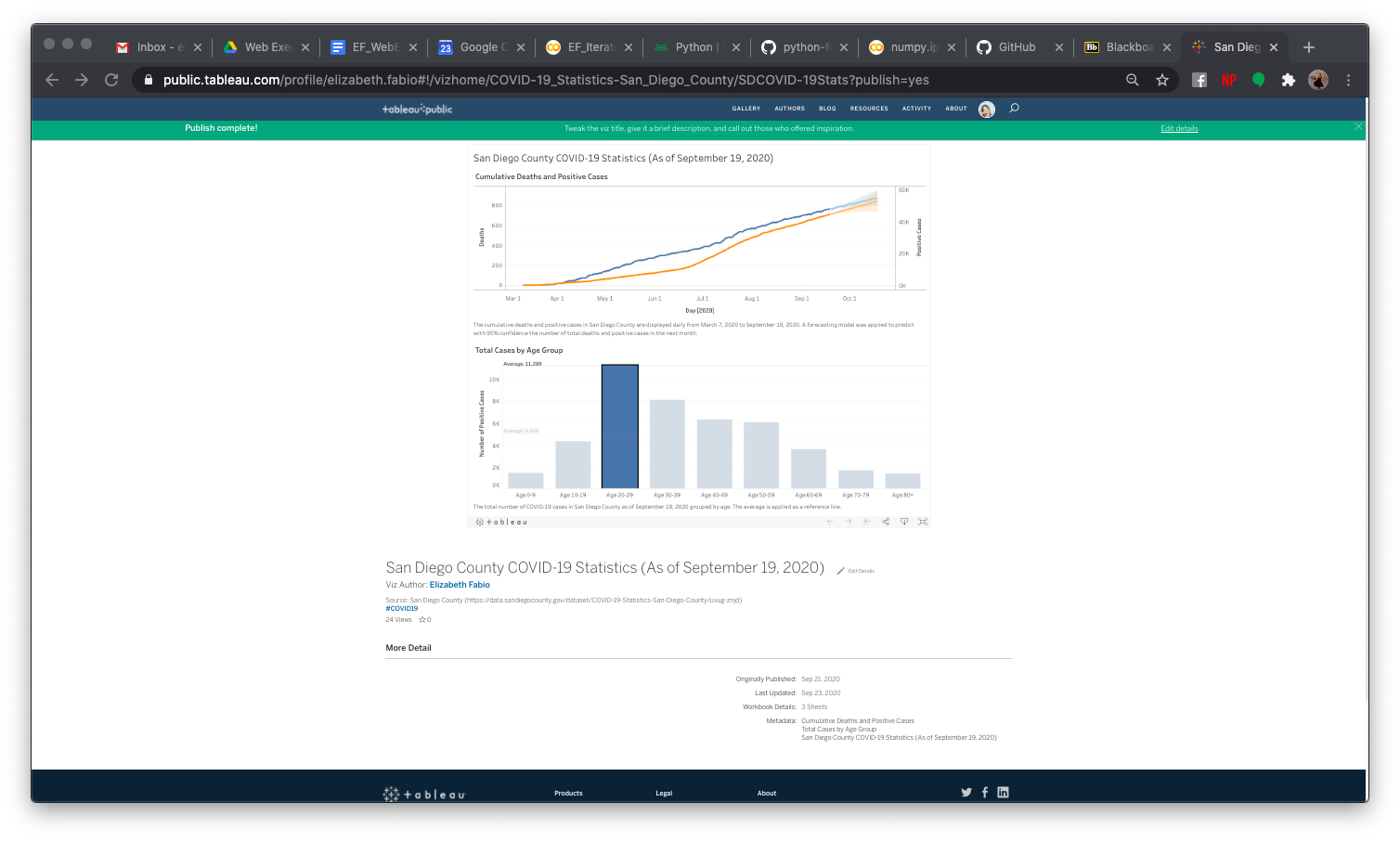
“The **Mean, Median, & Standard Deviation** sheet shows a bar chart of the **Followers Count** (top) and **Friends Count** (bottom) for each day in November 2015. The mean, median, and standard deviation is marked as lines for both the Followers Count and Friends Count. The standard deviation labeled is only one standard deviation.”

1. Please use Tableau to analyze the Lung\_Cancer\_Death data (or other data from the San Diego County Open Data Portal https://www.sandiegocounty.gov/content/sdc/data.html ) from the Web Exercise 02. Create a Dashboard to show your visualization results. Publish this new Dashboard to the Tableau Public Account and include the Web URL and a Screenshot of the web dashboard in the report. Write a short paragraph to explain how you create each sheet or the dashboard and how to share the dashboard to the public via Tableau Public.

**COVID-19 Statistics-San Diego County Dataset Link**: <https://data.sandiegocounty.gov/dataset/COVID-19-Statistics-San-Diego-County/uvug-znjd>

**Tableau Public Dashboard Link**: <https://public.tableau.com/profile/elizabeth.fabio#!/vizhome/COVID-19_Statistics-San_Diego_County/SDCOVID-19Stats>

**Screenshot\_COVID-19\_Statistics\_SDCounty.png**

****

The top sheet of the dashboard shows the cumulative number of positive cases and deaths in San Diego County. These numbers are displayed over time from March 7, 2020 to September 19, 2020.The total number of deaths were not collected until March 31, 2020 when the first death was recorded. A forecast model was applied to predict the number of positive cases and deaths in the next month (to October 18). The bottom sheet displays the total number of COVID-19 cases for each age group. A reference line was added to view the average. Based on the bar graph, the 20-29 age group has the most number of cases while the 80+ age group has the least number of cases.

After creating a Tableau Public account, I am able to export the dashboard along with the data (this will allow the dashboard to remain interactive) to the website. Tableau Public allows you to share your interactive dashboards and stories without the viewers needing to have Tableau downloaded.

1. Compare the functions between R and Tableau (including Tableau Public). What’s their advantages and disadvantages for each?

One big difference between R and Tableau is R requires scripts to build visualizations. R is a programming language by itself that can be built on top of multiple packages to provide a variety of charts, graphs, and other visualizations. Tableau, on the other hand, has a limited number of visualizations. With R, you are able to apply machine learning to your dataset. As for Tableau, you can only visualize what is provided in your dataset and machine learning cannot be applied. Pretty much with Tableau you are limited to the tools provided. For R, you have the ability to integrate with general applications, like Hadoop, and social media applications, like Twitter.

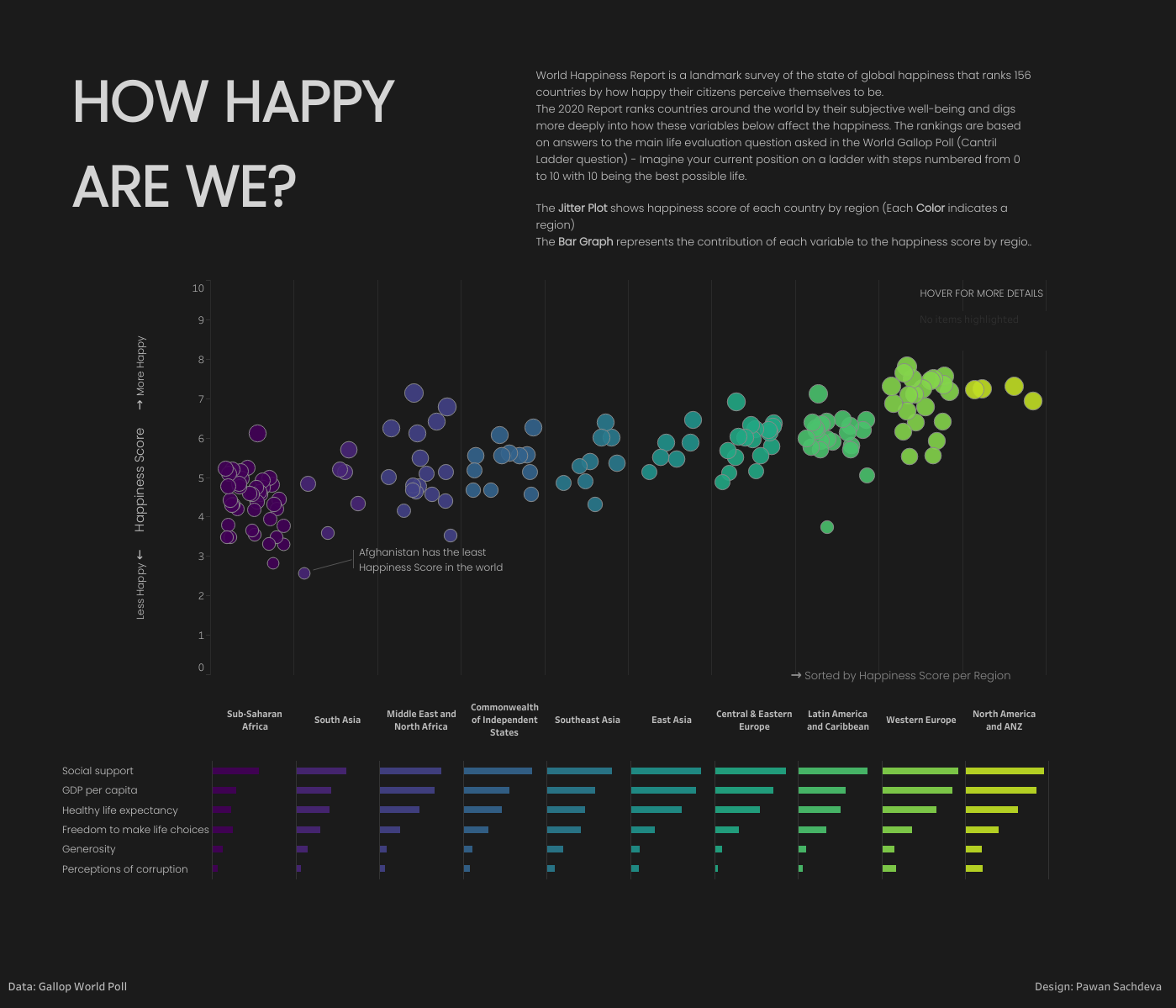
In relation to publishing your data visualization, Tableau Public is a free and easy online software that allows you to share your Tableau workbooks and preserves the interactive capabilities of your data visualization (especially for those who do not have Tableau). However, you would not be able to view your Tableau workbooks outside the Tableau sphere. As for R, you can share your R script and data for others to build the visualization for themselves. Packages like R Shiny can be used to create web apps to share your visualizations online. The difference between R Shiny and Tableau Public is that Tableau Public is easy to create. It only requires publishing your workbook from your Tableau Desktop. As for R Shiny, this requires building a script and requires more time to set up.

|  |  |
| --- | --- |
| **Tableau** | **R** |
| **Pro**:   * Easy to create * No programming knowledge required | **Pro**:   * Different packages to create a variety of visualization * Can be integrated with different apps * Machine learning packages available |
| **Con**:   * Limited data visualization * Tableau workbooks limited within Tableau * No machine learning capabilities | **Con**:   * More time needed to create visualization * Programming knowledge required |

1. Select one good example of data visualization from Tableau Public ([https://public.tableau.com/en-us/gallery/?tab=viz-of-the-day type=viz-of-the-day](https://public.tableau.com/en-us/gallery/?tab=viz-of-the-day&type=viz-of-the-day)). Write a short description (100 words) with a screenshot of your example.

**How Happy Are We?**: <https://public.tableau.com/en-us/gallery/how-happy-are-we?tab=viz-of-the-day&type=viz-of-the-day>

**How Happy Are We.png**



The main visualization is a scatter plot. It is beautifully organized by each region (e.g. Sub-Saharan Africa) using color as the differentiating factor. The secondary visualization is a barplot that lists the region’s happiness score for different variables like social support and GDP per capita. Specifically, the scatterplot is a Jitter Plot. Jitter Plots are used to separate dots so that they are not plotted directly on top of each other. With 156 countries listed, the Jitter Plot makes it easy to find each happiness score. There’s also a filter to search for specific countries as well. The happiness score comes from the World Gallop Poll and reflects Finland being the happiest country and Afghanistan as the least happiest country. Overall, the simplicity of the design makes the visualization easy to read and highlights not only each country’s happiness score but the significance within each region as well.