

Assignment 1 - Data Visualization in Tableau

- Due Jan 25 by 23:59
- Points 12
- Submitting a file upload
- Available Jan 18 at 0:00 - Feb 4 at 23:59

This assignment was locked Feb 4 at 23:59.

^{1/1}Assignment 1 - Data Visualization in Tableau

In this assignment, you will be using Tableau to explore the provided dataset, answer analytical questions about the player behavior represented, and create visualizations.

Format

Please create a .doc, .docx, pdf, or html file with your answers. Identify each answer with the matching question number. Several answers will be in the form of images - embed these under the appropriate question numbers in your response document.

Preparation

Install Tableau. Create a new workbook by connecting to the provided PUBG sample deaths CSV text file. I recommend turning on “extract” in the connection settings to improve performance. If you are having resource issues on your local machine, switch to the “small” sample CSV. It is a 10% set drawn from the same source data, and filtered to only the Erangel map. The difference should not have a meaningful impact on your answers.

Questions

Question 1 (3pts)

Create a calculated, continuous dimension called “Kill Distance”, using the killer and victim X/Y coordinates. Remember your Pythagorean theorem! Use this new field to create Median and Average Kill Distance measures.

Our objective is to compare typical usage by weapon classes.

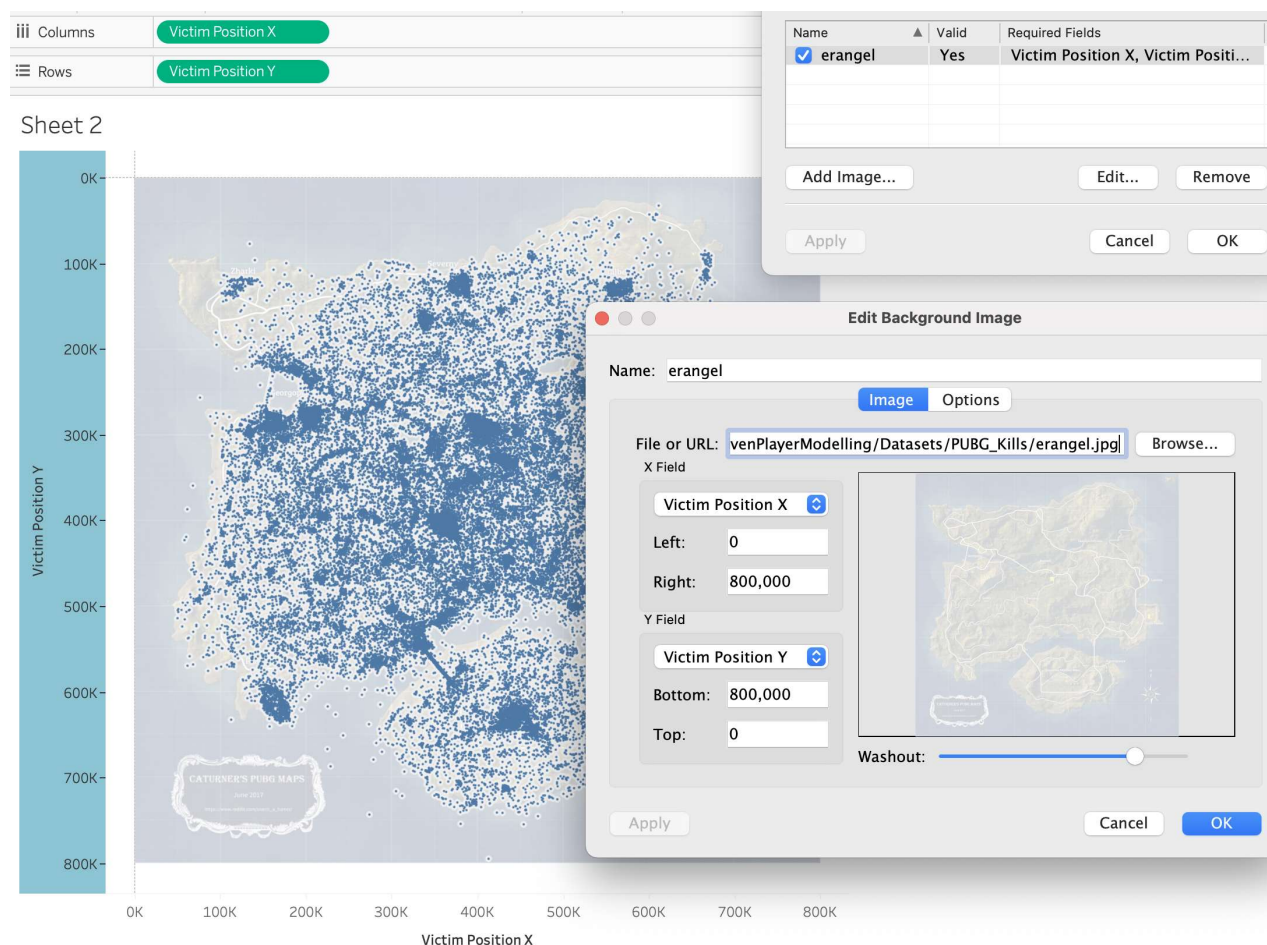
- **1A:** Sniper Rifles have a high effective range. **What other weapon has a comparable range?**
- **1B:** For each weapon, note the differences between the median and average Kill Distance. All of the guns in the Shotgun class share an uncommon property. **What is it? What does it suggest about the *distribution* of kill distances for shotguns vs other weapons?**

Question 2 (5pts)

Using *Map Menu > Background Images*, map erangel.jpg to Victim Position X and Victim Position Y. The origin [0,0] is in the upper left corner of the map. The data dictionary has more information about

the dimensions. Remember to set a reasonable washout, to convert Victim X and Victim Y to continuous dimensions, and to filter the dataset to “Map = E”.

Note: Getting the map axis right can be a bit tricky - as noted in the data dictionary, the origin (0,0) is in the *upper left*, rather than the lower left where you might expect. Thus screenshot may be helpful if your plot, the background image, or both are coming out up-side down.



Our objective is to understand how player deaths are distributed on the map.

- **2A:** Filter to “Time < 300s”. Set the visualization type to “density”. Choose a palette for “color” that seems reasonable. **Export your plot as a .jpg or .png and include it here.**
- **2B:** Change the filter to “Time between 300s and 900s”. Choose an interesting hotspot to zoom in on. Change the visualization type to “shape”. Add “Killed By Group” as the “Shape” parameter. Under “color”, remove halos (this is for performance). **Describe what you see. What does the pattern (or lack of pattern) suggest?**
- **2C:** Adjust the color, size, opacity, and other settings as necessary to make your observations more clear. **Export your plot as a .jpg or .png and include the result.**

Question 3 (4pts)

Create bins for the Time dimension with 60-second intervals. Using the map from Q2, add your time bins as “pages”. Cycle through pages using the controls at the bottom of the Pages card.

Our objective is to understand how a typical match evolves over time.

- **3A:** Note in the 120s bin, you should see a clear ring pattern (we discussed this pattern in class). The ring has effectively “dissolved” by 240s. This indicates a “phase change” as the play pattern in a typical match follows a predictable arc. Step through the time slices and identify another “phase change”. **Describe what you have found, and attach plots showing the before and after states (note: there may be more than one step between the two distinct states).**