

CIS 602 Special Topics: Data Visualization

Final Project Report: Dota 2 Statistics

By

Sujoy Kar (01577342)

Hardik Sankhla (01617980)

Under Guidance of

Dr. David Koop

Table of Content

1.	Introduction
2.	Data-Set Description
3.	Goals?
4.	Evolution of the Project
5.	Output Visualizations
6.	Future Developments and Scope
7.	Conclusion
8.	References

1. Introduction

Dota 2 Statistics provides visualization for any observer to understand how popular the game is all over the world in real-time and by observing data for some period of time. It also provides with the statistics that shows the trend in picking different Heros (Character) that the game has which players use in the game. An average observer should be able to differentiate between why the game is popular in some countries and why people may not be playing in a particular country at a point of time like day or night. It also helps in visualizing the latest trends of picking the hero by players and also provides each hero's statistics about how efficient that specific hero has been.

2. Data-Set Description

The data sets that we selected were available through Steam Dota 2 API database, <http://dev.dota2.com/showthread.php?t=58317>. To access these datasets one requires acquiring a steam developer key like this one that we used, "332C490A9CE67E401FED8FC2DBEBDCFA". The data set contains details about all the match history of games played by players, "GetMatchHistory". The "GetMatchHistory" json file contains details related to player's account_id (player's steam account) which was included under different match_id that is created each time players start playing the game. We used these account_id to call another steam API which was "GetPlayerSummary", which contained the personal details about the player whose account_id we were trying to access & display where these players are playing using the "loccountrycode" predefined Steam attribute. Then to display the stacked bar chart we required hero_id and the images provided by the steam API which we tried to access using details from getHero API and then used "GetMatchHistory" to link both the game details and hero picked attributes to create the visualization. We wanted to display the number of games that were played by picking up a particular hero and provide what were each hero's stats that included Kills, assists, deaths, loss, win, averages, etc. The detailed description of how the datasets were processed will be discussed in the evolution of our project.

Account_id: Unique Numeric Value

Match_id: Unique Numeric Value

Hero_id: Unique Numeric Value

Loccountrycode: Unique Numeric Value

Kills: Quantitative,

Assist: Quantitative

Death: Quantitative

XPM,GPM: Quantitative

3. Goals?

Our main aim while choosing this project was to show the game that we have been playing for about 5 years to the people who haven't heard about this game before and encourage them to at least try the game after checking out the visualization.

The **questions** that we aimed to answer while creating these visualizations what the observer could infer from them were as follows:

1. Where is the game most played both currently (Real-time) and during a phase of time (Recorded)?
2. Was the game successful in gaining interests of players in a particular country to play it?
3. What is the current Trend in picking up Hero by players all over the globe?
4. Which Hero has the most Wins or loses?
5. A Hero maybe amongst the most popular picks, but may have more number of losses compared to wins. Why?

Ans: Hero maybe picked more because it's more Fun to play, probable answer for the above question.

6. You can infer the probable causes for a hero's loss or win based on hero's stats like XPM & GPM.

Ans: Having a Low XPM (experience per minute) or GPM (gold per minute) maybe a reason the player failed to play the hero correctly and ended up losing the game.

4. Evolution of Project

When we first started accessing the data it seemed easy and the data was abundant but once we started creating the choropleth and started plotting the data on the map we were in confusion as we weren't getting any output but when we look at the data sets there were 100 results being provided and so we were expecting something. But then we realized that we couldn't access data that weren't made public by the players who play the game so we weren't able to access their data & so we were getting undefined objects. So to clear this issue we added a loop to exclude the undefined objects and exclude the account_id = 4294967295, i.e. Anonymous account_id that cannot be accessed. There was another issue that we faced was that account_id have to be first converted to a 64-bit format of Steam ID to properly access the player's data. Another problem that we faced was while trying to monitor the data in console after running the script that we had pushed into a new array provided us with an empty array instead. So to solve this problem we added a delay of 15 secs and 45 secs as we needed to use Asynchronous methods to access the database properly.

We first created an orthographic presentation for showing the population of gamers in a particular country but then it looked cool but waiting for the rotation of the globe to complete was kind of a waste of time for the observer, so we changed it to a Mercator presentation in the end for proper

depiction of the data. We added tooltips to both the choropleth map and stacked bar charts. We added legends to the choropleth and bar charts to understand the data being visualized easily. We created drafts for every phase of our projects to keep in touch about our progress. While creating the 3D radar chart we had some issues regarding the countries that we wanted to present on the edges of the circle as the Steam API had different abbreviations of the “locountrycode” compared to the ones that we had in the countries-hires.json that we uploaded on gist.github but eventually we managed to link both of them and present the visualization. We referred <http://www.visualcinnamon.com/2015/10/different-look-d3-radar-chart.html> and <http://bl.ocks.org/nbremer/21746a9668ffdf6d8242> while creating the 3D radar chart.

Now we needed a different idea and wanted to answer multiple questions in a single visualization itself so we decided to create a stacked bar chart which would include the Hero icons with its respective win/loss data being displayed. We didn’t know about how to use tooltips at that moment so we just focused on presenting the number of wins and losses of a particular hero. Later on we added a tooltip which provided us with the wins and loss & also each loss/wins represents the number of kills and death that happened during that loss/win period and we also provided what was the average KDA (kills, Death, Assist), XPM(experience per min) and GPM (gold per min). We also added the legend in the end for the observer to understand the stacked bar chart easily. In the end we linked all the html files together for the observer to access all data from the index.html itself. So in the end we used hue to differentiate the number of players playing currently all over the globe as well as for the stacked bar chart which represents the wins/loss. The Techniques that we have used are: Choropleth, 3D radar chart, stacked bar chart, tooltip; we tried brushing (but aren’t working properly).

5. Output Visualization:

- Visualization 1 (Comparison between Real-Time and Recorded game Data using Choropleth)

Choropleth visualization for the Recent Games played



(a) Real-Time Data

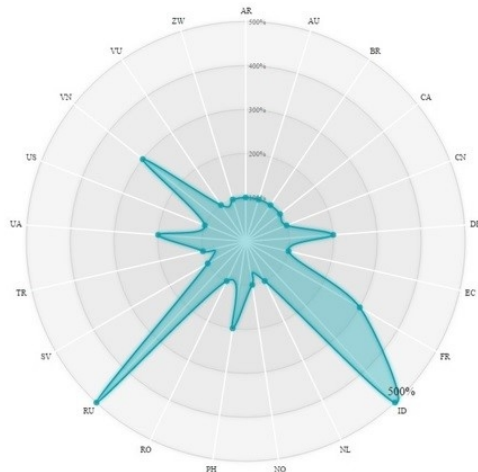
Choropleth Visualization for Games recorded for a Day at different intervals



(b) Recorded Game Data

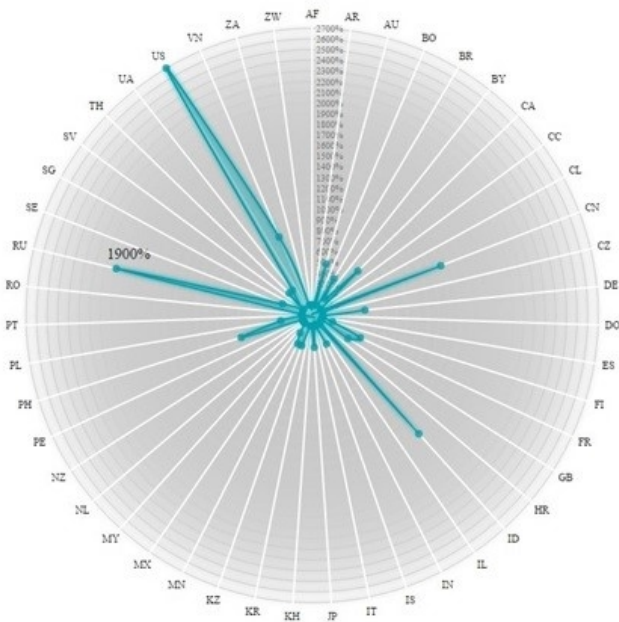
- Visualization 2 (Comparison provides better understanding of the game's popularity using 3D Radar Chart)

3D Radar Chart for the Recent Games played



(a) Real-Time Data - Relation between Countries with at least 1 player playing

3D Radar Chart for Games recorded for a Day at different intervals



(b) Recorded Game Data – Relation between Countries with at least 1 player playing

- Visualization 3 (Trend of Hero most and least Picked using stacked bar chart)

Hero Trend



(c) Hero Trend Data

6. Future Development and Scope

- To complete the visualization regarding using a bar or line chart in a tooltip while hovering over the bar chart which provides the KDA, XPM, GPM representations.
- Creating hero icons in tooltips itself and applying the concept of fisheye to the bar chart to make the chart more compact.
- Provide more data about professional players

7. Conclusion

Our visualization provides where players play the most and least in a particular country. We can compare 2 countries with respect to real-time and previous record data to get a proper analysis as to why there was an increase or decrease in the number of players in a country and understand if the game was able to get more players into playing the game even more. The Bar chart gives an overview detail about the hero that people picked while playing the game. Did the hero fail to win and what were their specific stats recorded while these players picked these heros. It provided the observer an idea as to if the hero is more popular in terms of wins or in terms of loss which can be depicted by more number of losses & also being the highest picked hero amongst the hero pool. Our main aim was to answer the questions that we proposed to the observers so that they can understand the game even much better as to how popular the game is and will they want to try the game once by playing it.

8. References

1. <http://www.visualcinnamon.com/2015/10/different-look-d3-radar-chart.html>
2. <http://bl.ocks.org/nbremer/21746a9668ffdf6d8242>
3. <http://bl.ocks.org/Caged/6476579>