

# The Numbers behind eSports

HCII Data Visualization, Spring 2017

Neil Bantoc(neilbantoc@gmail.com), Adam Yee(adamy@andrew.cmu.edu)

<https://github.com/adamgyee12/eSportsVis>

## Background and Motivation

eSports are becoming more and more prominent in mainstream media as big sports news organizations such as ESPN dedicate more resources to it. Yet some people argue that eSports can't be seen as "real sports" as traditional sports lean more towards the physical side.

## Project Objectives

With this project, we want to educate an audience unfamiliar with eSports on what it is and how it has grown through the years. We also want to provide an argument to people who don't see eSports as real sports by showing through data visualization just how big the industry is.

We will break down our project into three sections:

- Compare eSport viewership and traditional sport viewership
- Compare and contrast eSport communities to one another over time
- An in-depth case study of popular eSport game, Dota 2

Considering time is a heavy constraint in this project, will consider our project a success if we at least accomplish the first two bullet points.

## Data

The data that we're going to use to fuel our visualizations will come from a variety of sources. When comparing the growth of eSports against more traditional sports, we will use viewership data collected from Twitch.tv

(<https://newzoo.com/insights/rankings/top-games-twitch/>). Twitch.tv is the major platform in which eSport events, matches, and games are streamed on, allowing us to compare these numbers to that of television viewership for other sports (<https://www.statista.com/statistics/251536/average-tv-viewership-of-selected-major-league-baseball-games/>).

When we take a deeper dive into exploring the different communities within eSports, we will use the data sources from eSports Earnings (<http://www.esportsearnings.com/history/2016/games>). This data set has nearly two decades worth of historical data tracking the number of professional players in each eSport game, the amount of monetary prizes for each eSport, and even a breakdown of individual's earnings and earnings by country. This data will allow us to compare and contrast a variety of games with many variables, as well as allow us to track the popularity of certain communities over time.

We also may take a closer look at a single eSport community granted we have enough time. To do so, we will conduct a case study of a popular MOBA eSport: Dota 2. We will use the data provided from the game's community Wikipedia page ([http://wiki.teamliquid.net/dota2/Main\\_Page](http://wiki.teamliquid.net/dota2/Main_Page)) which will give us access to in-depth numbers for tournaments, team earnings, and individual player earnings.

## **Data Processing**

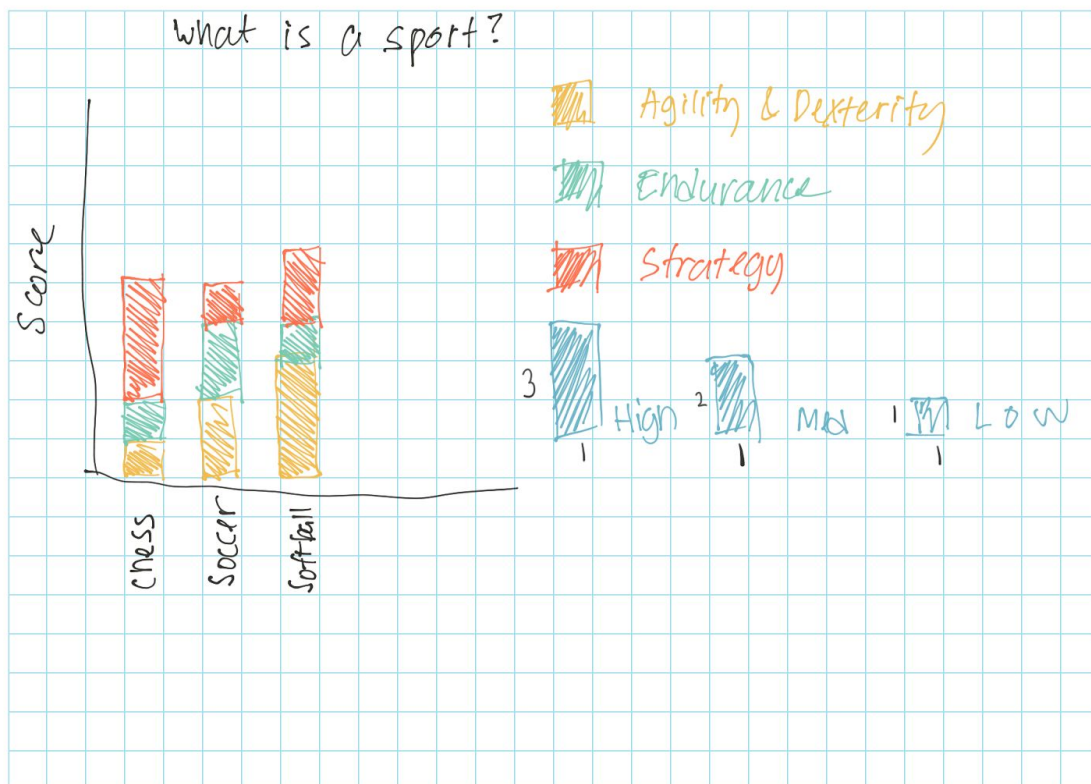
A majority of the data that we will be working with is currently provided in segmented tabular forms. This will require us to do a fair amount of data cleansing before we are able to quickly visualize the data. Due to the fact that most of the data we will be working with will be strictly quantitative (number of people watching, prize pool amounts, etc), the actual cleaning of our data shouldn't be too difficult, since they're all in a similar format.

We are not working with any APIs to extract our data, so the method in which we will collect and process our data will be to copy the data from our sources into a CSV file. This will allow us to manipulate the data either through Excel (or Python if required) to create a common format.

## Visualization Design

We plan to display our data in a narrative, scrolling webpage. This will allow us to convey a story describing what eSports is, and provide context for a deeper dive into the statistics behind it.

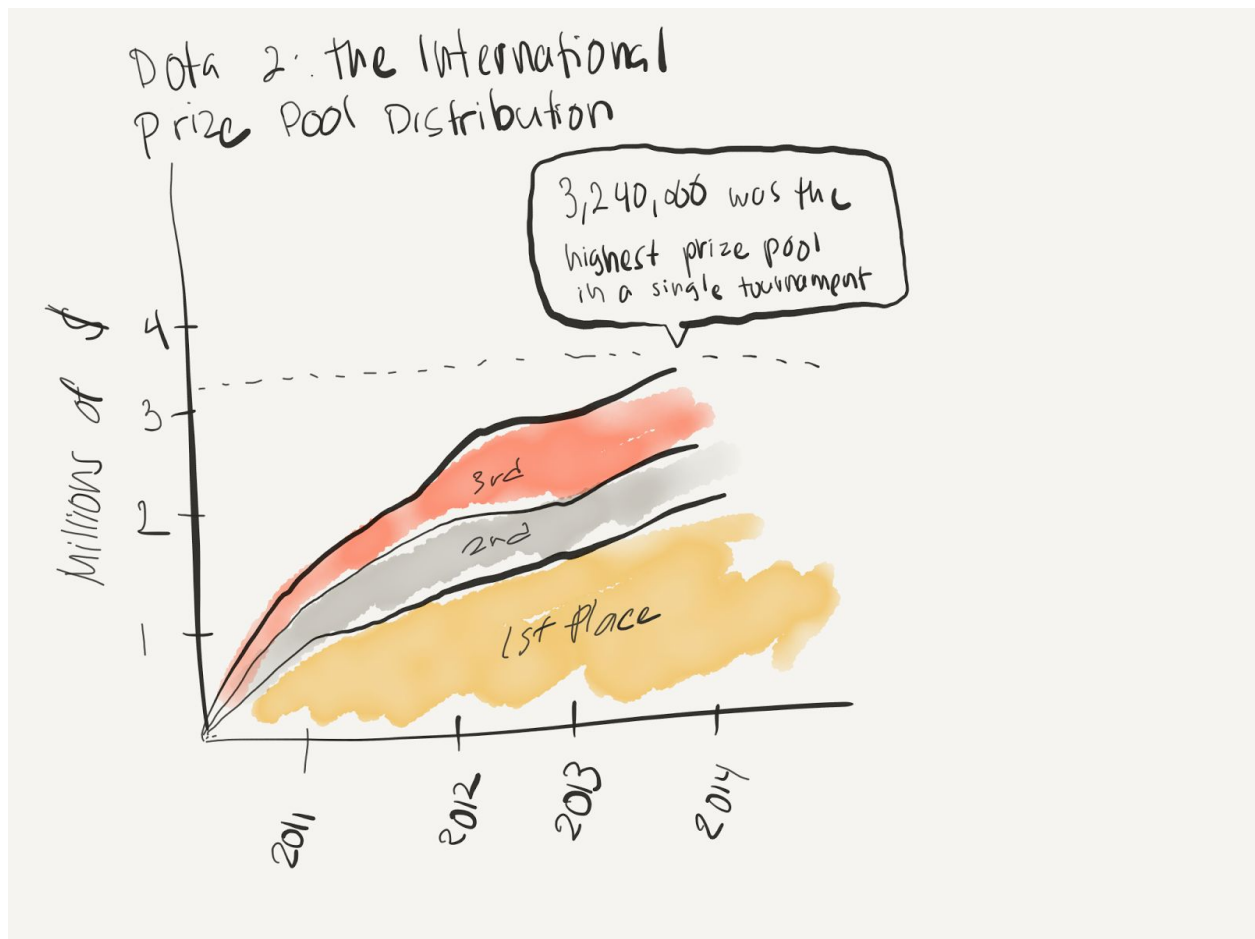
Some initial visualizations we've sketched out are as follows:



## What is a sport?

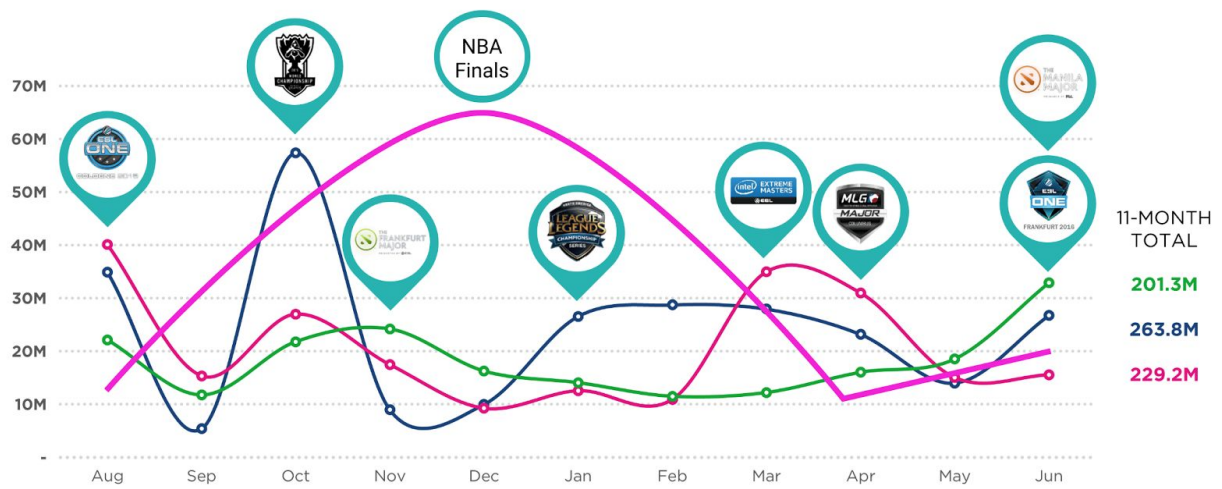
To present an argument that existing sports are not purely a demonstration of physical prowess, we want to use a visualization to show proportionality between the

non-physical aspect of a sport such as strategy making and the physical ones such as agility and dexterity.



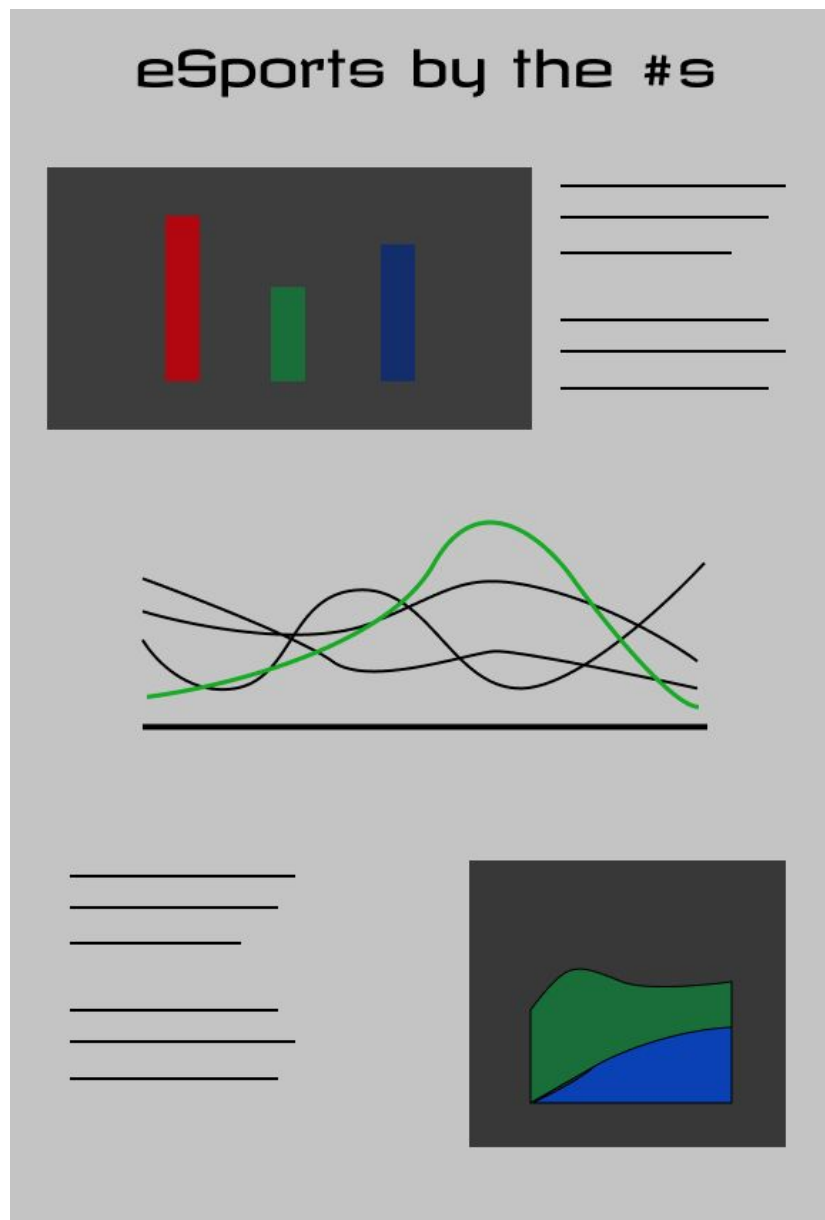
## How big: Prize pool?

Using a stacked line graph, we can show not only how big the overall prizepool is for a specific tournament, but also how much each participant can potentially earn if they compete in the tournament.



A line graph over time will allow us to compare how the viewership of major eSport tournaments compare against the major events of traditional sports, such as the NBA finals or the World Cup.

Final design example:



This example of a final design combines all the visualizations we have ideas for, and places them in a vertical scrolling page. It would be ideal to make this page fluid with animations so that the experience of exploring the data is intriguing. Using a common color scheme throughout the entire web page is important for our project so that users have a lighter cognitive load when swapping contexts for a new visualization.

## Must-Have Features

As mentioned above, we will use a scrollytelling format to progressively tell a story about esports. Graphs and visualizations that animate according to the scroll offset will also be present in multiple points of the webpage. This will be our main form of interactivity.

A demo of the library we will be using to implement this can be found here:

<http://1wheel.github.io/graph-scroll/>

## Optional Features

Other forms of interactivity we can add in include combining multiple graphs that relate a single factor with multiple other factors into one graph and enabling the user to switch between these categories. For example, we can combine a scatter plot that relates total earnings with years playing and another graph that relates total earnings with age. As our visualizations become more concrete, we will be able to explore a wider range of optional features in addition to the ones listed here.

## Project Schedule

Week 1 (March 27 - March 31)

- ☐ Process data from sources into CSV files
  - ☐ eSport viewership
  - ☐ Traditional sport viewership
  - ☐ eSport community earnings per player/country/tournament
- ☐ Sketch iterative designs for potential visualizations

Week 2 (April 3 - April 7)

- ☐ Create first digital iterations of visualizations
  - ☐ Compare eSports and traditional sports
  - ☐ Compare eSport communities against one another

Week 3 (April 10 - April 14) Milestone Due April 10 - Functional Design Prototype

- ☐ Iterate on visualizations
- ☐ Begin putting visualizations into a scrolling layout

- ☐ Process Dota 2 case study data

Week 4 (April 17 - April 21)

- ☐ Iterate on scrolling layout design, refine story we want to tell
- ☐ Case study visualizations

Week 5 (April 24 - April 28)

- ☐ Integrate case study with other visualizations
- ☐ Iterate on scrolling design and story

Final Presentation Due May 3 - Website Showcase