



# Path of Exile Leaderboard Analysis – October 26, 2017

Jimmy Tran  
Georgia State University

## 1. Overview

Originally, I decided on using a dataset about possible Poker hands and the odds of winning with those hands. The simplicity of the data coupled with the extensive math involved turned me away from that though.

As a result, I switched to a dataset about a popular free-to-play massively multiplayer online game called Path of Exile (Figure 1). The data set in question comes from a website called Kaggle and focuses on the top 60000 players on the PoE leaderboards as of October 26, 2017. A link is provided here to that dataset although it may or may not be updated as of recent due to constant daily changes in the PoE leaderboards.

Link: <https://www.kaggle.com/gagazet/path-of-exile-league-statistic/data>

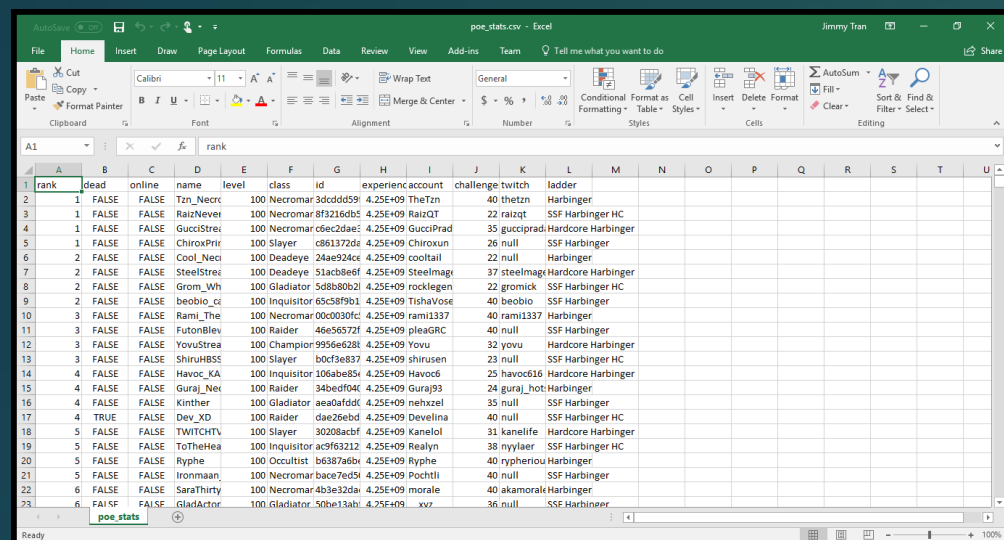


Figure 1. The initial data viewed in Excel.

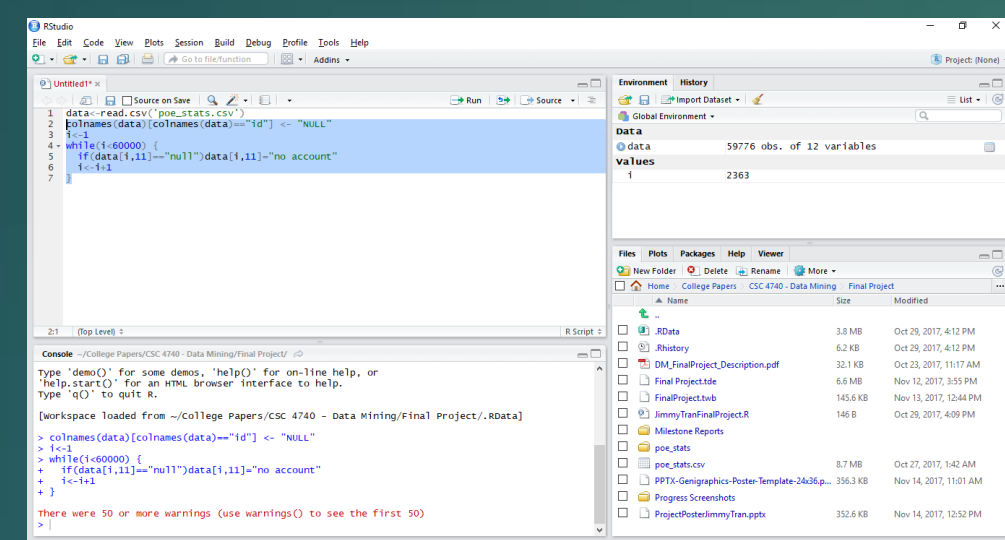


Figure 2. Attempts to clean in R end in failure.

## 2. Cleaning Process

Initial cleaning attempts were based in R script (Figure 2). Twenty minutes after beginning though, I realized that would not end well at all and I also questioned whether I had to technically clean my data since I could simply ignore the attributes in question anyway.

At this point, I imported my data, conveniently in the form of a .csv file, into Tableau (Figure 3). With no issues arising from the import, I also realized I could actually hide unwanted attributes or change the names of null values by providing an alias. The actual data would not be affected either, only the front-end would be changed and visible. Suffice to say, I saved a lot of time skipping the R cleaning process (Figure 4).

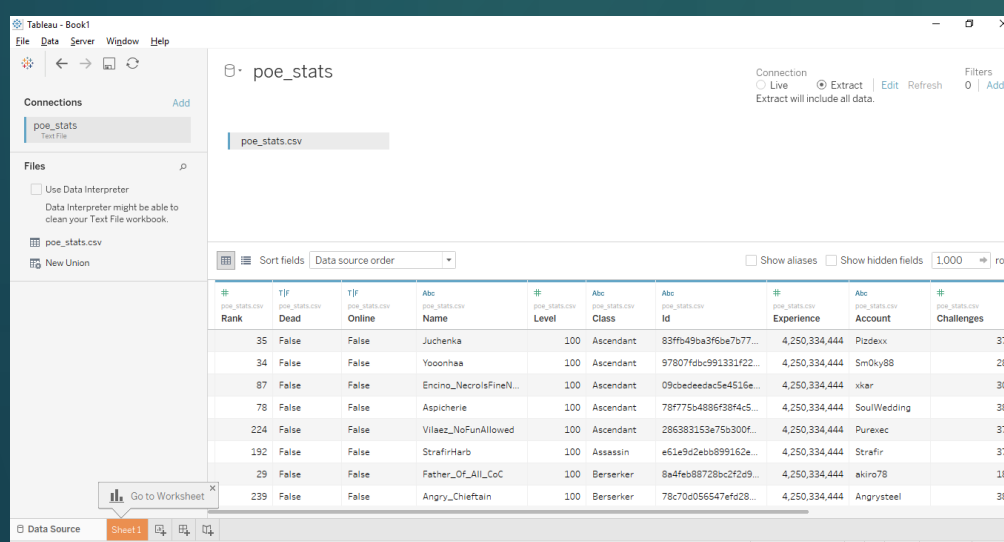


Figure 3. Data before cleaning in Tableau

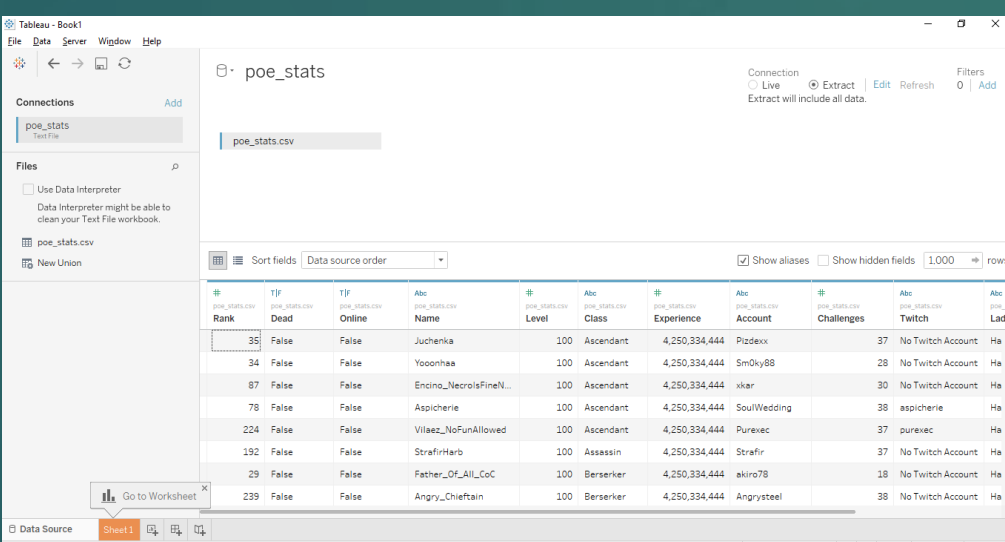


Figure 4. Data after cleaning in Tableau

## 3. Initial Tableau Plotting

With pseudo-cleaning complete, attempts to create the first plot began. As can be discerned from the given images (Figures 5 and 6), this did not tide over well either. Working with dimensions and measures sounded odd given how much time I spent with R. I was not going back to R though, so I continued with Tableau and eventually got a good result.

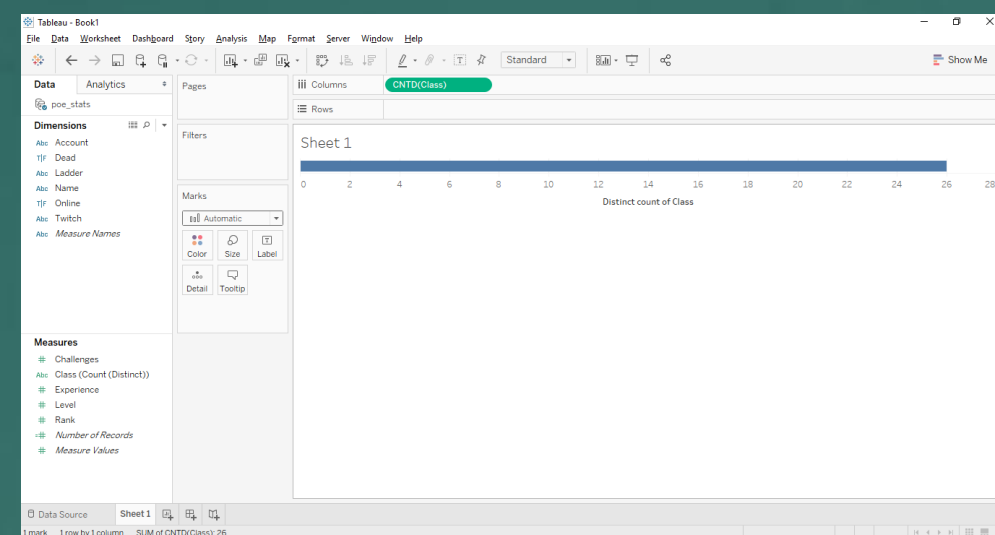


Figure 5. What is a measure?

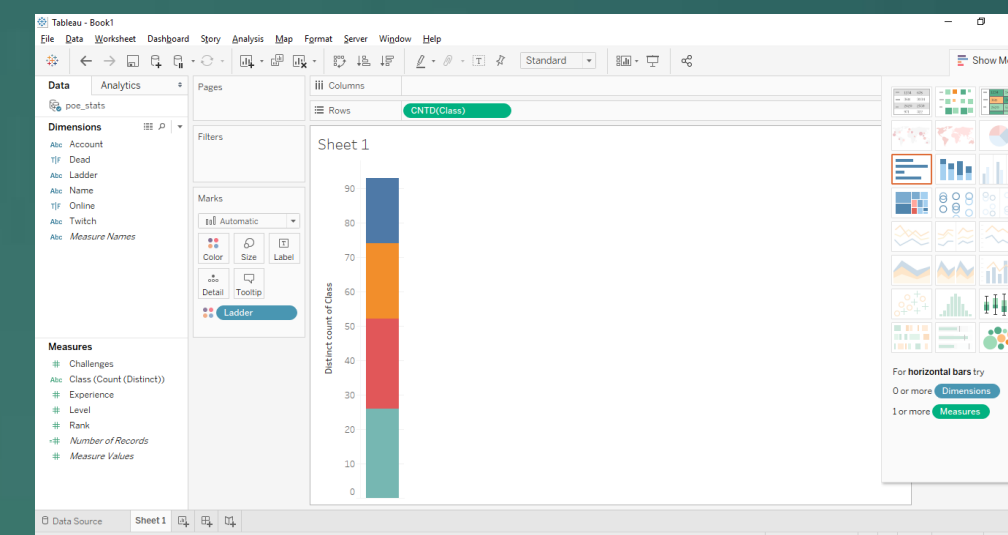


Figure 6. Another bad plot.

## 4. Successful Plotting

I fiddled around with dimensions and measures for about an hour until the first successful plot was achieved. From there, the process is all downhill as plotting dimensions against themselves as a measure worked out a good percentage of the time. A few plots are just duplicates since I wanted to display the percentages rather than the raw numbers (not present on this poster because of lack of space) but there is a fair amount of data that I gleaned from my plots even if a PoE outsider has no clue what he/she is looking at. In some cases, I used program screenshots instead of an exported .png file as I am sure no one wants a 30000 pixel anything.

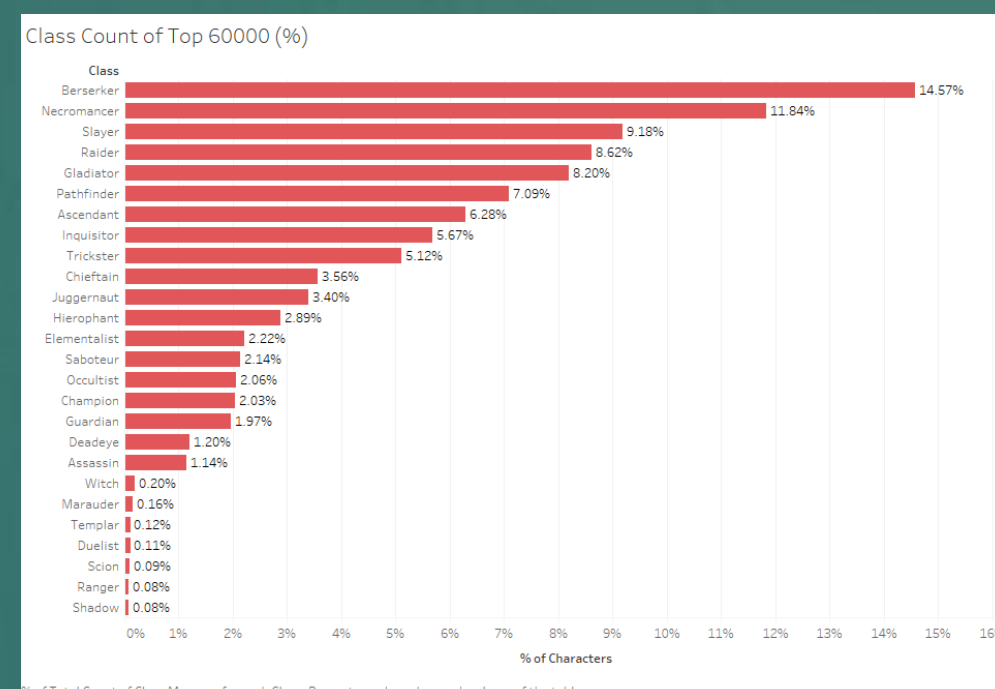


Chart 1. Percentage of characters created by class. Berserker is more popular than necromancer? What is this nonsense?

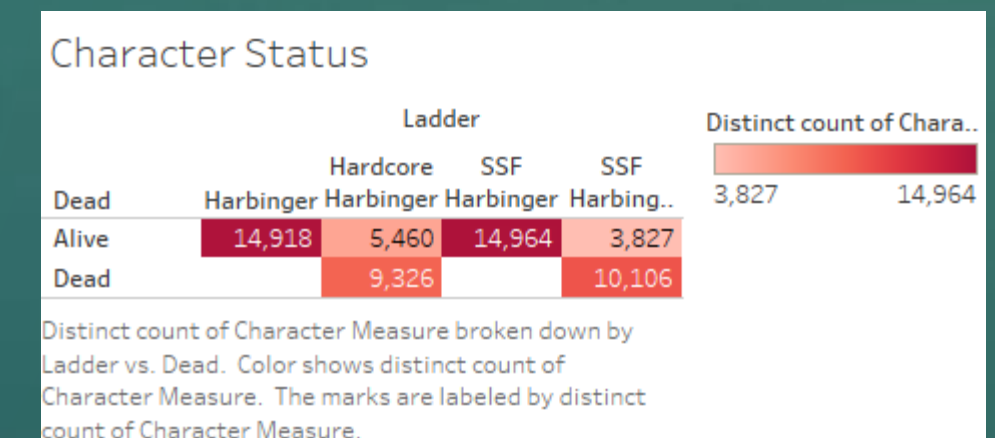


Chart 2. Character status by ladder. Harbinger and SSF Harbinger can be disregarded as characters cannot permanently die in those ladders.

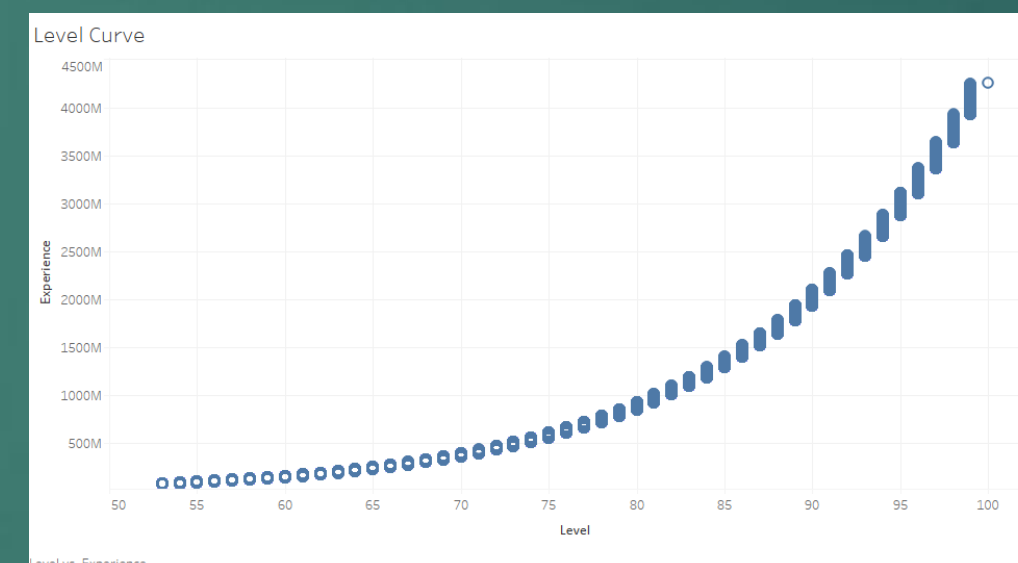


Chart 3. Level curve. Kinda glad I never played to 100. Exponential experience requirement is ugh.

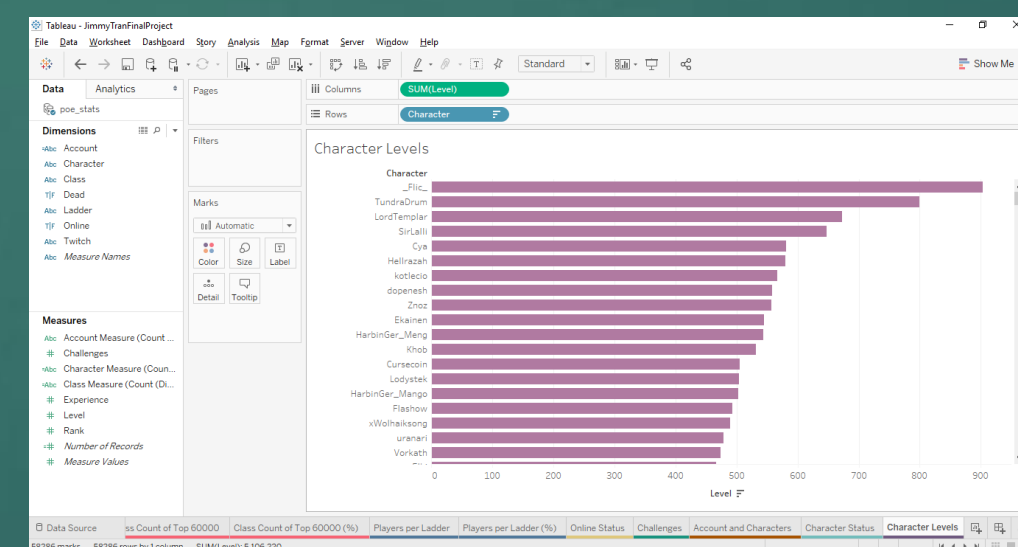


Chart 4. Level cap increase or hackers? Last I checked the cap was 100.

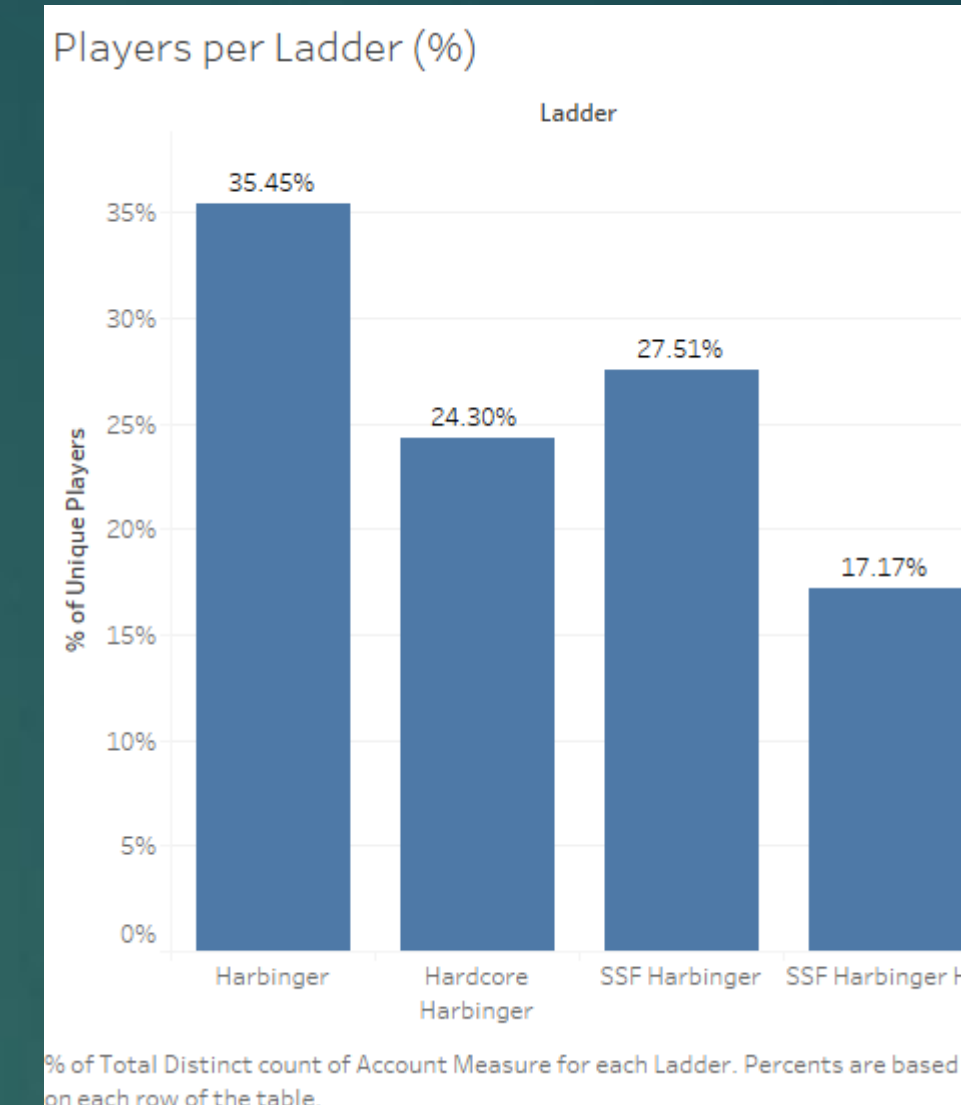


Chart 5. Percentage of players involved in each ladder. Note that it is possible for a player to be in multiple ladders.

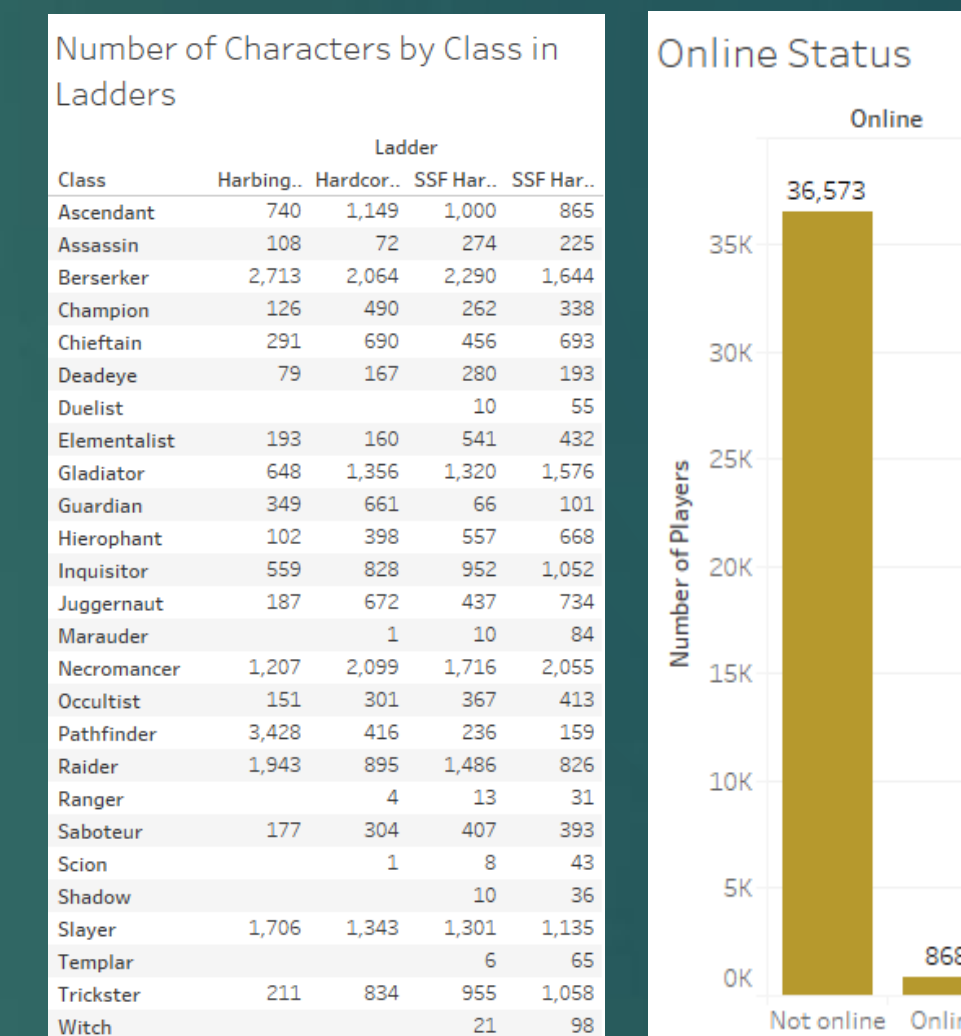


Chart 6. More detailed numbers for Chart 5. This shows class breakdown.

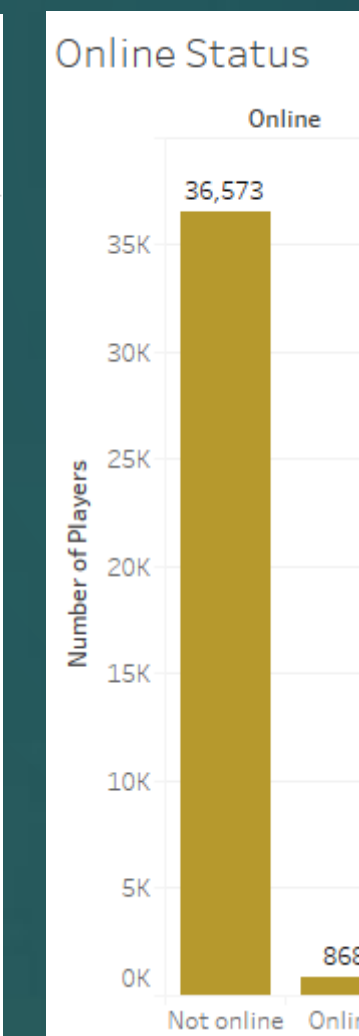


Chart 7. Online status of players at the time the data was mined.

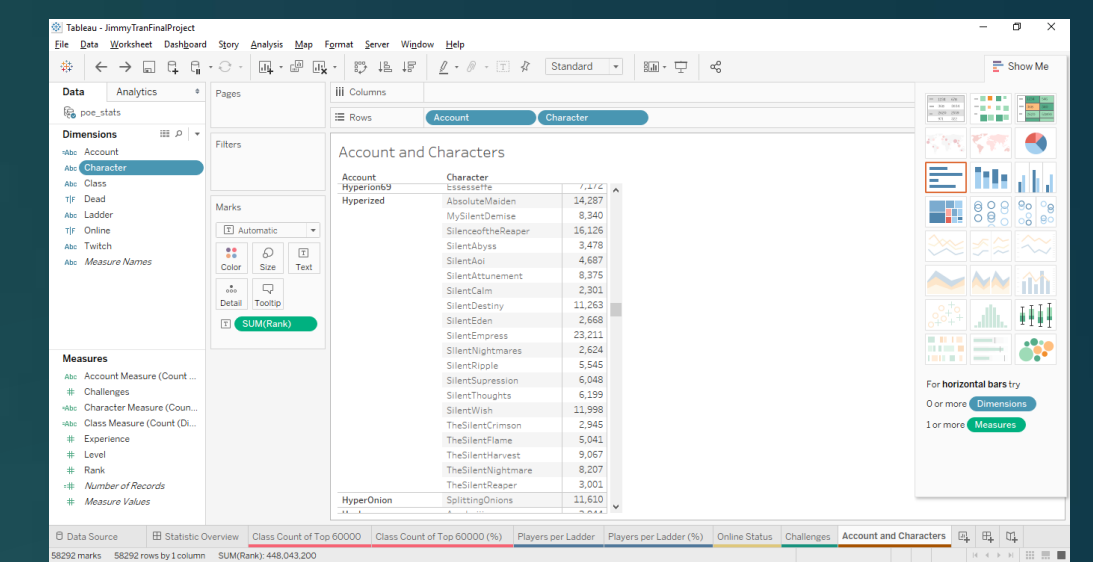


Chart 8. Hyperized seems to play a bit too much.

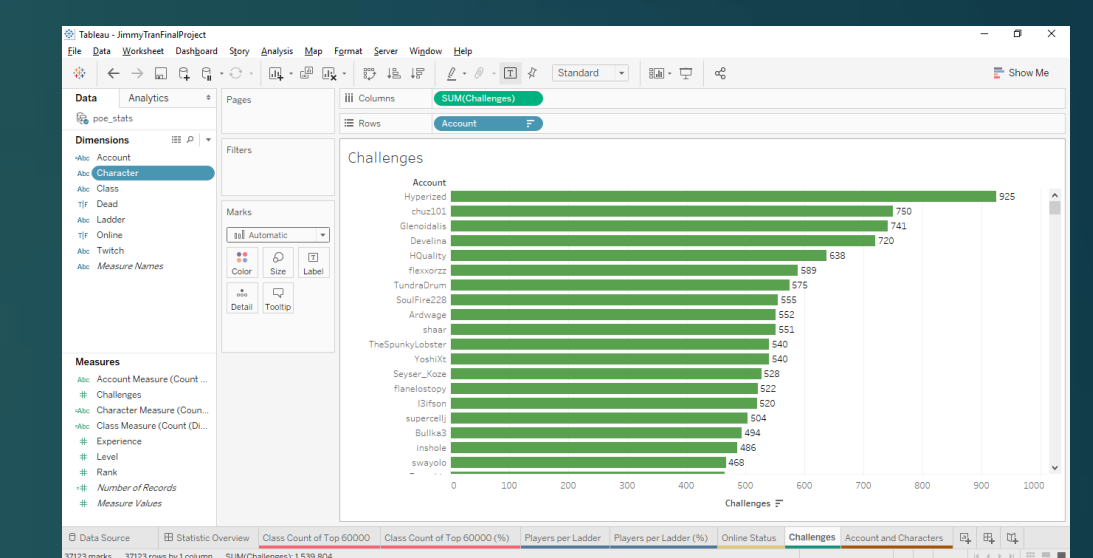


Chart 9. Did I say "a bit"? I meant he probably needs a life.

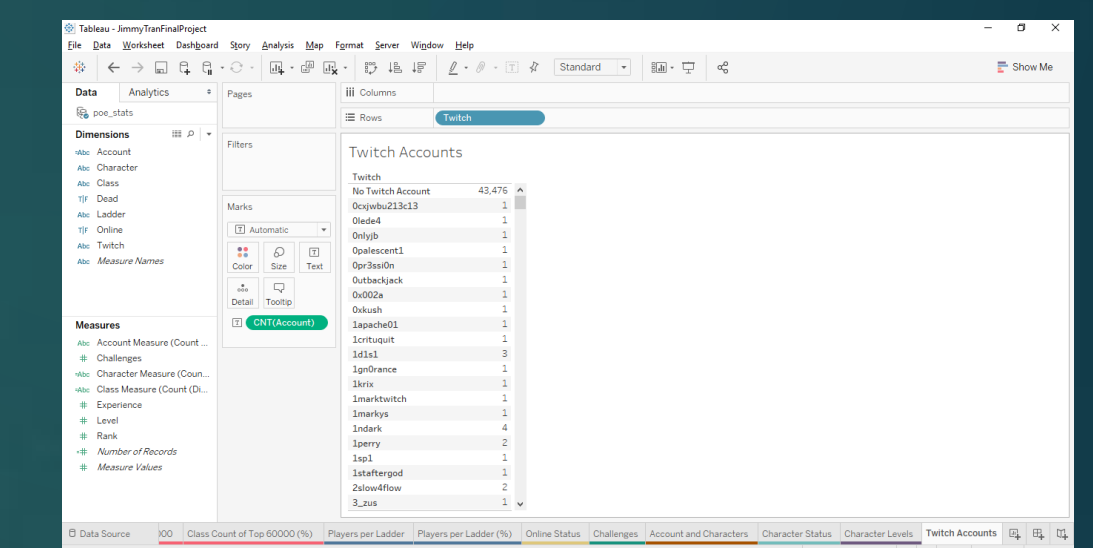


Chart 10. Twitch accounts. I could not quite get this to display right but the math should be obvious since the count for "No Twitch account" is there.

## 5. Conclusions

I enjoyed seeing the data of the leaderboards of Path of Exile. It definitely put some things into better perspective for me personally.

From a data mining standpoint, R is still terrible, and I would have no regrets if I never touched it again. Tableau is difficult to understand with all its additional terminology but honestly was far more user-friendly than I suspected. I had a fun time when all was said and done.

## Contact

Jimmy Tran  
Georgia State University  
Email: [jtran25@student.gsu.edu](mailto:jtran25@student.gsu.edu)  
Phone: 678-644-2270

## References

- Kaggle: <https://www.kaggle.com/gagazet/path-of-exile-league-statistic/data>
- Path of Exile Website: <https://www.pathofexile.com/forum>