Jimmy Tran

Ahmadzadeh

CSC 4470

9 November 2017

<p style="text-align:center">Mini-Project Milestone 3</p>

Query: A pdf file with 1-3 paragraphs explaining the problems you are dealing with. Avoid

listing general problems. I would like to know the exact problems you have to tackle, and your

raw ideas of attacking them.

As mentioned in the last milestone report, there are null values present under the Twitch

account attribute. This makes organizing the data by that section a bit tricky by default. I will

probably replace the null values with the string "No account" and then consider them as such

when I analyze that particular column. Such accounts will be moved to the bottom of the list if

sorted or organized. The ID column is utterly useless for data analysis and will need to be

removed as well or outright ignored. Its only purpose I could find was for player identification

from a security standpoint. Likewise, experience is another pointless column. It is almost another

measure of level but more difficult to understand from an outside perspective. It needs removal

as well. Lastly, the online attribute is a bit odd. I will keep it as it can be used if I wanted to see

how many players were active at a given time but would probably need more data from other

files for comparison for such an attribute to be useful. Aside from that, it pretty much goes the

way of id and experience.

As mentioned before (again), the current data is organized rather oddly. Rank is the first

column and is the only attribute used to organize the data. It makes more sense to me to sort it by

ladder first as that is more important since the rank 1 characters are not exactly equal in skill and effort, which the current organization implies. I will move the ladder column to the left for easier viewing (at least from my end if not for the audience) and sort it according to ladder name first, then rank if possible. It also makes little sense that the player account name and twitch account attributes are located so far from the character name. I will probably try and move those to the left of the character name column since most people would probably be more interested in who the player is rather than what their character name is.

The rest of the data is not much of an issue and can freely be resorted and organized as I wish. It also all has relevant information that need not be removed or altered in any way like the attributed mentioned above. Most of my analysis was based around the assumed usage of R though. I may try and learn to use Tableau instead, which for all I know, could invalidate, complicate, or even simplify some of the above solutions.