Written Analysis

Jordan Mullin

February 26, 2020

Heroes of Pymoli

# Three conclusions regarding the data for Pymoli:

1. It is important to analyze the data closely to ensure that the outcome meets the criteria of what was asked, simply having the correct code in some instances is not good enough without additional problem solving. For example, in this data, the same “Item Name” had multiple “Item ID”’s. Additionally the corresponding price to the item ID was also different, despite having the same Item Name.
2. It is also important to consider that the same player (“SN”) may have made multiple purchases, and therefore important to stay mindful of this when sifting through data to get the requested results. This was especially important when obtaining the player count.
3. It was necessary to continuously check that the when finished writing a portion of code, that the result is indeed correct, even if the initial top ~20 rows appeared correct. One instance that this made a difference was when constructing the age bins. I found that players with an age of 10 were included in the “< 10” category, when in fact they should have been included in the “10-14” age bin.
4. Writing the code for this assignment was very tedious, and it was helpful and almost necessary to continuously copy and paste previous code (after confirming that it was correct) and then make the minor changes necessary.
5. When concatenating data frames together, it is important to be mindful of similar or identical column headers in order to avoid concatenating two tables again that repeat the column header, but have different values in the table. This is helpful to keep in mind because you cannot then change the header name without going back and renaming the column header before concatenating the data frames.