Bradley Higdon

8/29/20

Pandas Data Trends

The first major data trend that I can detect within this set of data would be the relationship between item price and item popularity. Looking at the top five most popular items, there is a fairly consistent pattern between an item having a high price and being more popular. However, the dataset is still limited in sample size as the top five items have around 10 purchases each, making them susceptible to sampling errors. Additional purchases data may be necessary to improve the accuracy in proving this trend.

The second major data trend I noticed when looking at the results of this exercise was the variance in age and purchases made. The 20-24 age bracket made up just under half of all purchases despite being just one of many age brackets. No other age bracket can even reach half of both the total purchases and value of purchases categories of the 20-24 age bracket. A similar element of age and purchases that stood out was that the two brackets with the highest average purchase price was the <10 and the 35-39 brackets. Both brackets have a small purchase count total, so it could be a sampling error, but it is interesting that both brackets are the highest two in average money spent per purchase.

The third data trend for this data set is the rather large gender gap in purchases. Despite being roughly equal quantities in real life, the player base of the game is mostly male. Not surprisingly, most of the purchases made are by male players. However, this fact should not discourage from the idea of trying to attract more female players. Though men make more purchases due to their larger quantity, the data shows that females will, on average, purchase slightly more expensive items than their male player peers. In the context of in game purchases, female players are therefore slightly more profitable customers to have than male players.