**Juan I Castaneda**

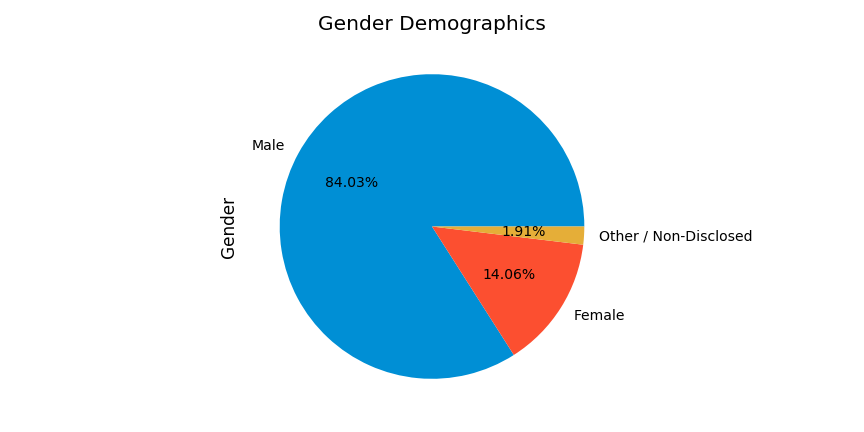
**Homework Assignment #4 Pandas**

**2/15/2021**

**Three observable trends based on the data**

**Gender demographics analysis:**

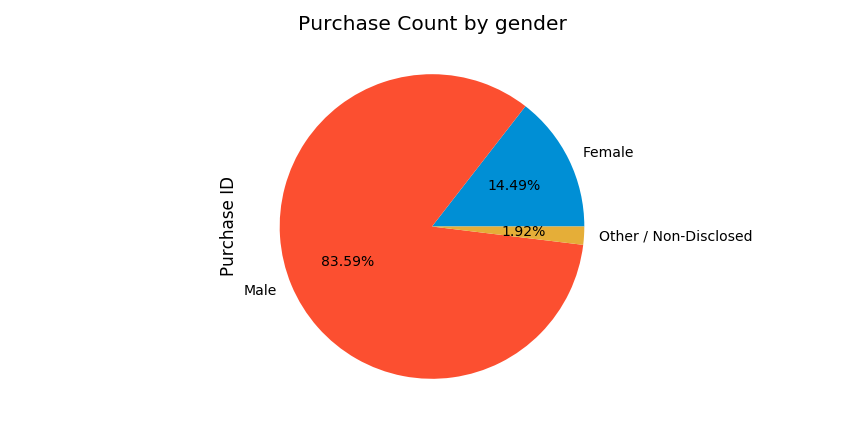
* 1. Based on the data provided we determined the gender demographics. There are 576 players in total; we can clearly see that the predominant gender among the players is “Male”, which represents the 84% of the players who made an in-game purchase.
  2. ” Other / Non-Disclosed” category is by far the category with the lowest presence in the game (1.91%); this might be because the person doesn’t identify with either “Male” or “Female” gender, or because they made a mistake while selecting the category. A further analysis would be necessary if we want to know more about this category.

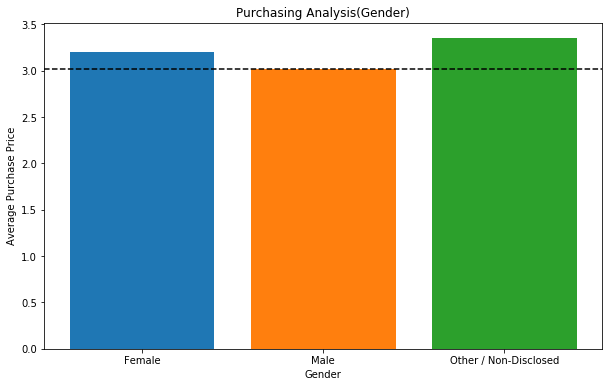


Data Obtained from purchase\_data.csv utilizing pandas

**Purchasing analysis(gender):**

1. There are 780 registered purchases in total from which 83.59% of the purchases were made by males, 14.49% by females and 1.92% by Other / Non-Disclosed



1. Even though the majority of players who made a purchase are males, the *Average Purchase Price* made by “Other / Non-Disclosed” and Females is higher than the Average Purchase Price made by males. This tells us that in average the people in those two categories tend to spend more money per average purchase. If we compare the total purchase value among the categories, we can see that most of the actual revenue came from male players, this only because of the number of male players is clearly predominant. If the all the categories had a similar number of people, Female and other/Non-Disclosed would represent the highest source of revenue.

Data Obtained from purchase\_data.csv utilizing pandas

**Age demographics analysis:**

1. Based on the data provided we determined the age demographics. There are 576 players in total; we can clearly see that the predominant age rage among the players is “20-24 years old”, which represents the 44.79% of the players who made an in-game purchase.
2. The age ranges where the total count individually represents less than 10% of the total are:
   * < 10 (2.08%)
   * 10-14 (2.95%)
   * 30-34 (3.82%)
   * 35-39 (5.38%)
   * 40+ (9.03%)

