Stats 400- Assignment 2(Data Types)

1. Nominal - Gender, Race

Gender: used to make classification of individuals that have participated in the tests. Categories individuals based on the cultural norms and society norms. Also based on the masculism and femineity of the individual.

Race: Categorizing individuals based on their ethnic background, physical characteristics and ancestry. This is a nominal categorical variable, because this variable cannot be ordered, which we can not rank them, or place them in a sorted order.

Ordinal – General Health condition, Monthly Family Income

General Health Condition: Place different individuals into different ranges or ranks, based on their health condition, after carrying out some medical tests, or through gathering information from the health services of the people that have visited them.

Monthly Family Income: Ranking different family incomes and placing them into different groups. This would be a ordinal categorical because it can be ordered either in an ascending or descending order, based on the income that different families make.

Numeric and continuous –Weight and standing height.

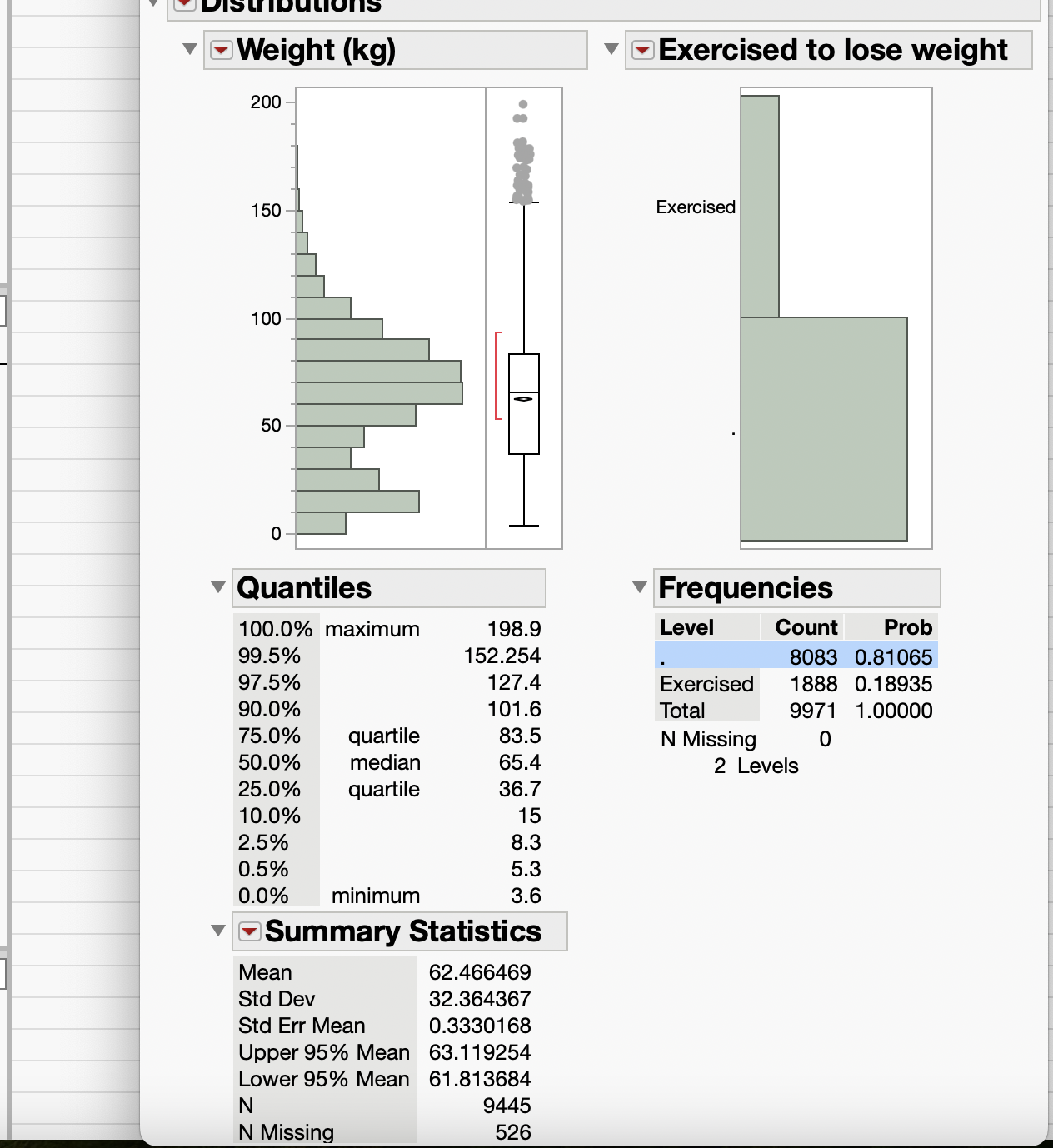
Weight: This would be numeric and continuous, because it is fully numbers and there is no limit in this data. Different individuals can have different weights and they can go up to infinity.

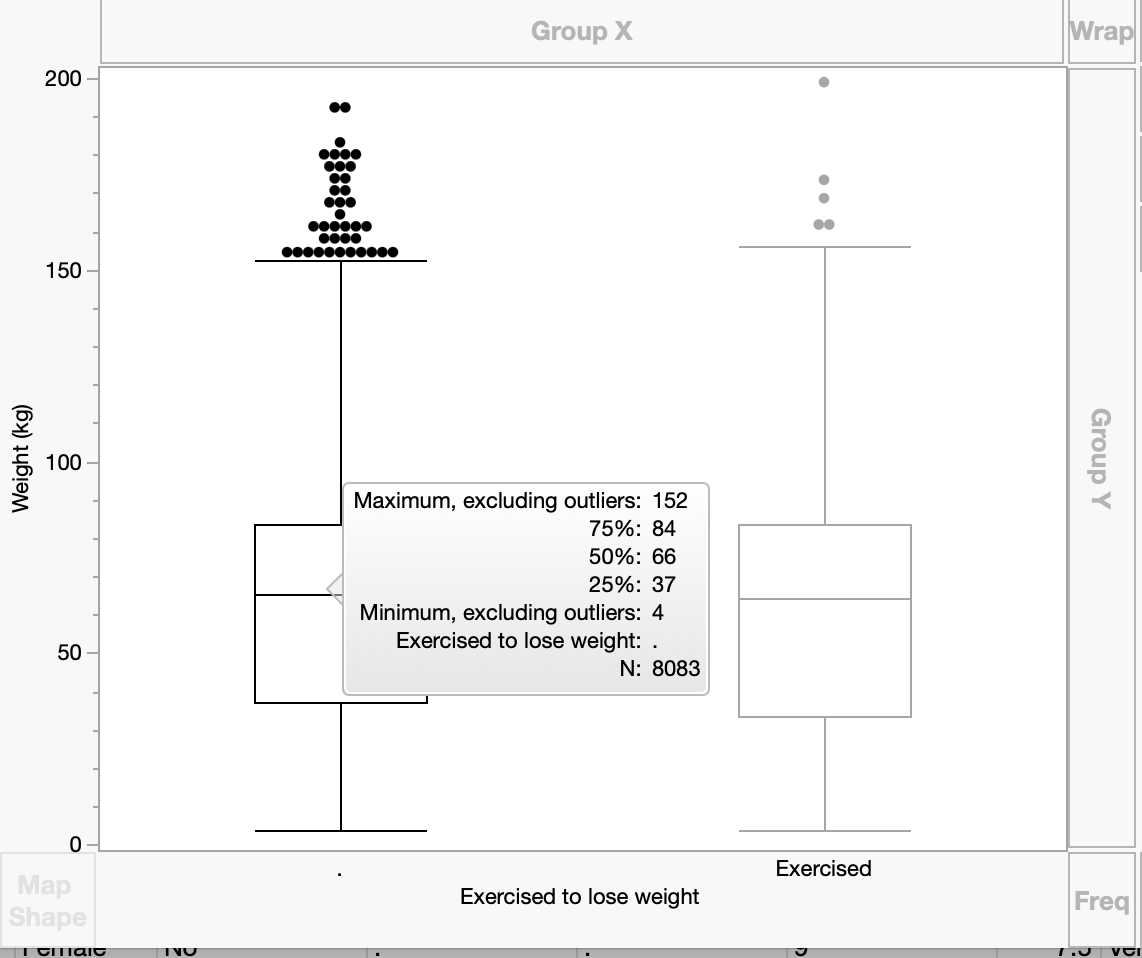
Standing Height: Can also be numeric and continuous, because the data is many made up of numbers and it is hard to place a limit in the grow of humans. People can have varying heights. Standing heights can be represented in decimal formats as well. Hence, they are continuous.

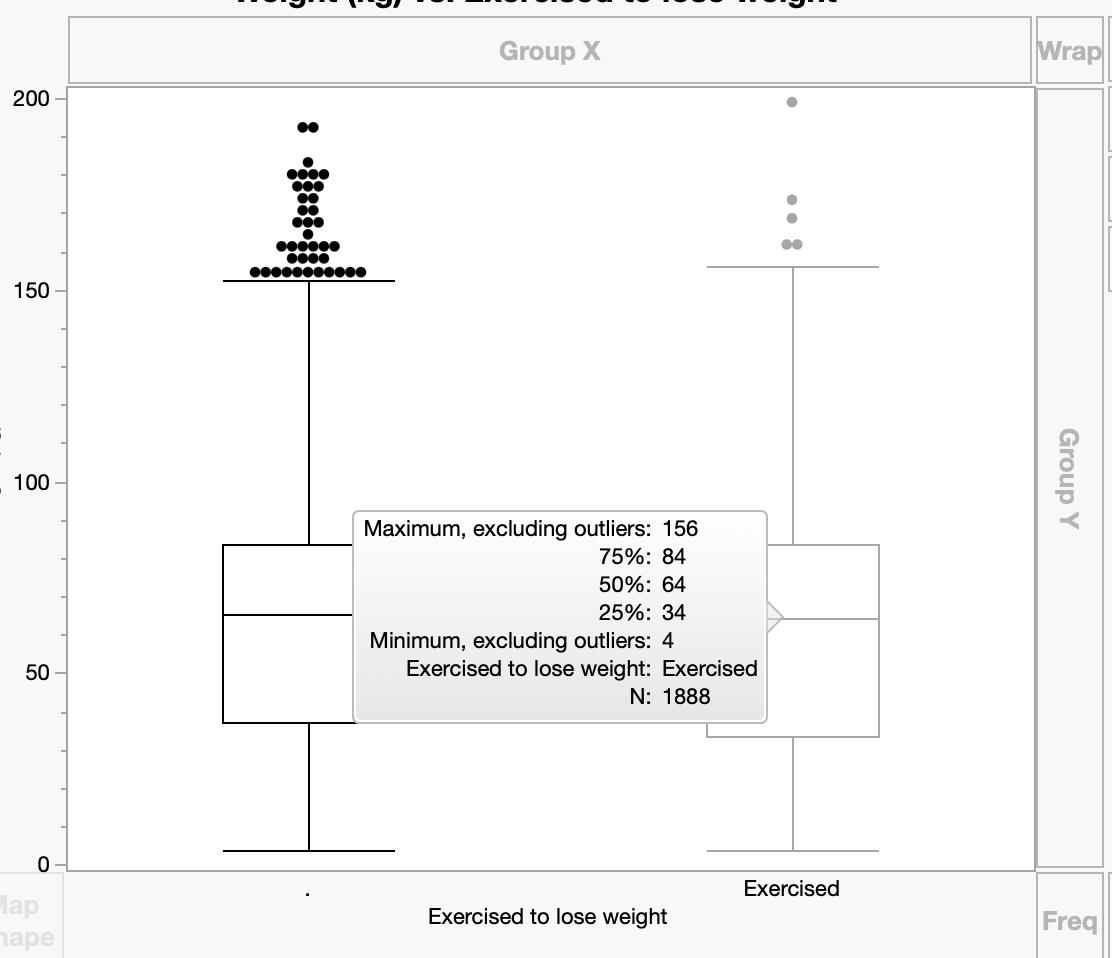
Numeric and discrete – “Number of prescription medicines taken,” “Number of days for moderate workout”

Number of prescription medicines taken: This is numeric and discrete

1. Examining the variables “Weight” and “Exercised to lose weight,” it is noticed that the upper quartile for both variables is 84. However, the lower quartile and mean are different for both variables. The lower quartile for “non exercised” is 37 and the median is 66. For the “Exercised” the lower quartile is 34 and the median is 64. Also, through the boxplot we can see that “Exercised” have less outliers compared to the other factor “Not Exercised” which is represented through the “.”







1. 2 variables selected are: Weight and Standing Height. There is a direct proposition between these variables. As the standing height increases the weight tends to increase as well. However, when comparing the quartiles, it seems like “Standing Weight” quartiles are double the quartile for “Weight”. Looking at the box blot, we see that for “Weight” it is left skewed whereas for “Standing Height” it is right skewed.

