ISOM-670 Business Statistics

Group Assignment 1

Team 2: Yifei, Jie, Eleanor & Susie

Summary Statistics of “Weight” & “Height” - using summary()

```
     Weight
 Min.   :100.0
 1st Qu.:135.0
 Median :160.0
 Mean   :160.3
 3rd Qu.:180.0
 Max.   :225.0
```

```
     Height
 Min.   :60.00
 1st Qu.:65.75
 Median :69.00
 Mean   :68.37
 3rd Qu.:72.00
 Max.   :75.00
```

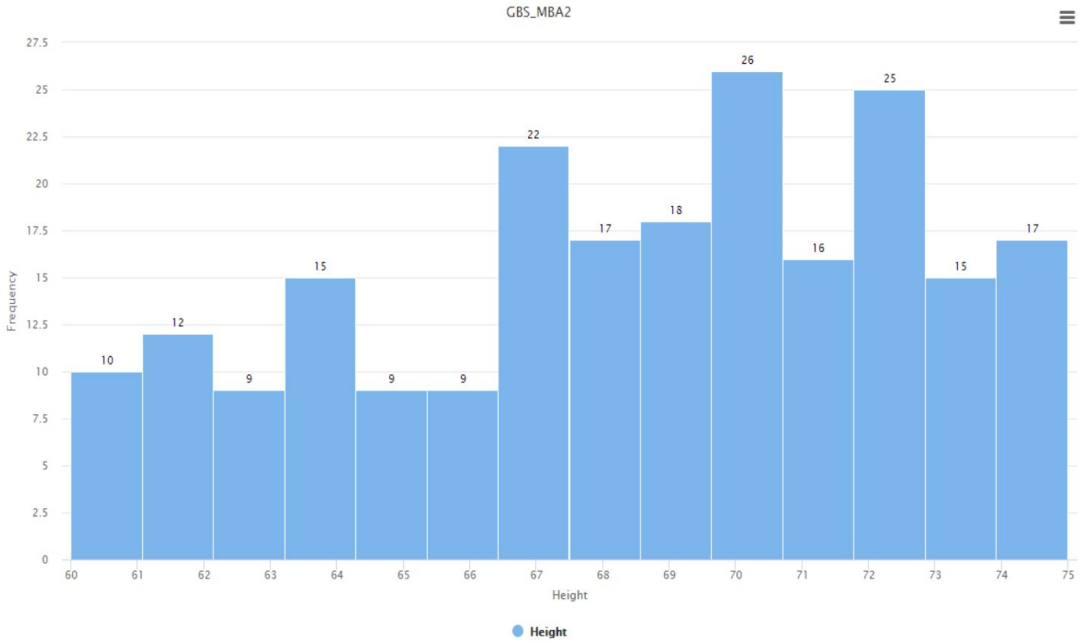
There are no apparent outliers in these two variables - “Weight” and “Height”. No missing values discovered. Both variables seem to be roughly symmetrically distributed; “Weight” is centered around the mean 160.3(lbs) with a minimum of 100(lbs) and a maximum of 225(lbs), while “Height” has a relatively small range between 60(in) and 75(in).

Summary Statistics of “Weight” & “Height” - using describe()

	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
Weight	1	220	160.26	31.04	160	160.00	31.13	100	225	125	0.01	-0.83	2.09
Height	2	220	68.37	3.84	69	68.55	4.45	60	75	15	-0.35	-0.88	0.26

There are 220 observations in this dataset. “Weight” has a mean of 160.26(lbs) and a standard deviation of 31.04(lbs); “Height” has a mean of 68.37(lbs) and a standard deviation of 3.84(lbs). The medians of “Weight” and “Height” are 160(lbs) and 69(lbs), respectively.

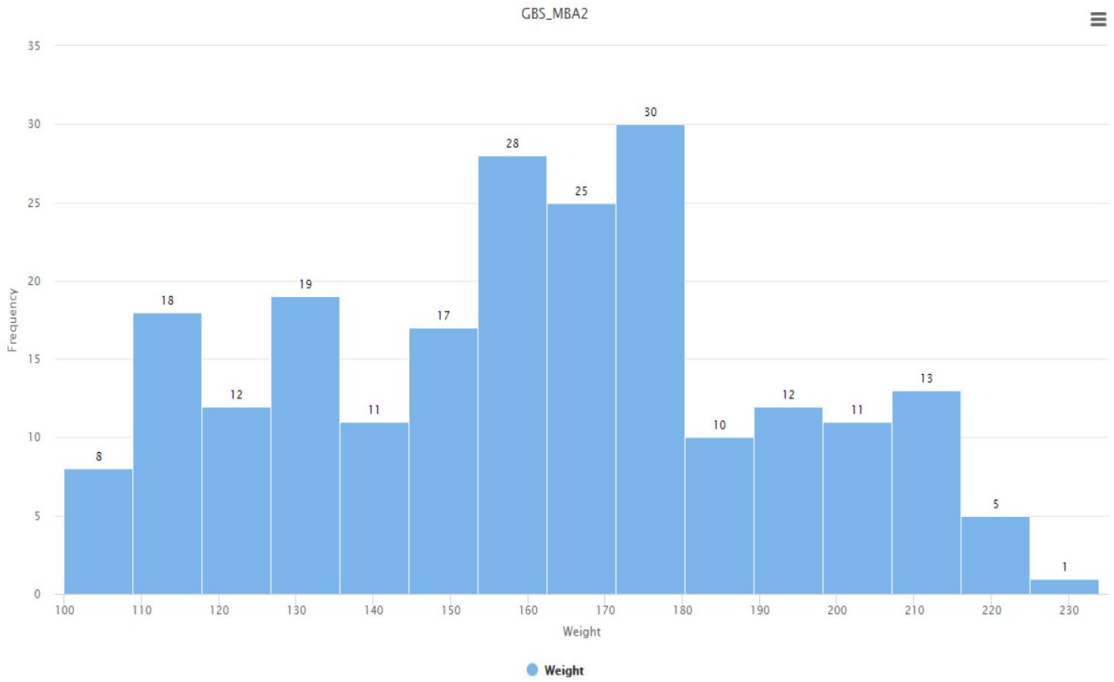
Histogram of “Height”



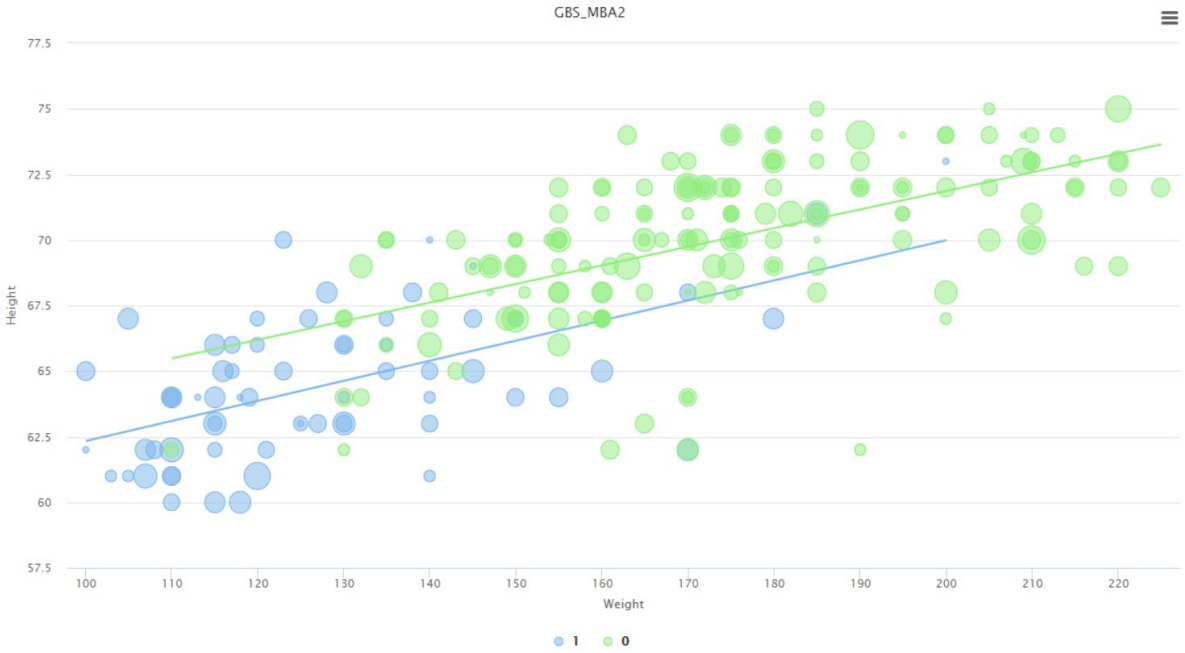
From the histogram, we can see that the distribution of the “Height” variable is skewed to the left, but only slightly. The bucket with the highest frequency of data is around 70 (inches) with 26 observations.

Histogram of “Weight”

From this histogram, it could be observed that the distribution of the “Weight” variable is nearly symmetrically distributed. The bucket including the most data is roughly 170-180 (lbs) with 30 entries.



Scatter plot of “Height” versus “Weight”



For the entire dataset, the “Height” variable is generally positively correlated with the “Weight” variable. The blue dots in the graph represent females, and the green dots in the graph represent males. Most data about females gather around the bottom left corner, while the data about males have a wider spread and lean towards the top right corner.