**Become a General Manager Part 3**

**Jackson, Ben, Will & Cam**

**MYTH #1 – Players from the West clearly eat AAA Beef**

A. Using a 2-sample t-test will provide a statistical way to check whether the assumptions that Western Conference NHL teams are "bigger" and "taller" than Eastern Conference teams are valid. The parameters needed to do the calculations are the two samples: Western and Eastern players. Also, the variables: height and weight, the standard deviation from those variables and a significance level which will be 0.05 for this one. After that we have all our information, if the p-value of the t-test is less than the significance level, we reject the null hypothesis and conclude that there is a significant difference in the mean height and weight between the two conferences.

B.

| Conference | Height average | Weight average | Standard dev. Height | Standard dev. Weight | # of player |
|---|---|---|---|---|---|
| Eastern | 73.06 | 198.37 | 2.13 | 15.1 | 373 |
| Western | 73.03 | 197.95 | 2.11 | 15.37 | 355 |
| For height, t-value is 0.20. The p-value is 0.42. The result is not significant at $p < 0.05$. | | | | | |
| For weight, t-value is 0.37. The p-value is 0.35. The result is not significant at $p < 0.05$. | | | | | |

C. As seen in the table, we obtain a t-value of 0.20 and a p-value of 0.42 for height and a t-value of 0.37 and a p-value of 0.35 for weight. Both of these p-values are greater than the significance level of 0.05, indicating that we fail to reject the null hypothesis and that there is no significant difference in the mean height and weight between the conferences. Actually, fun fact, the Eastern conference players have a tiny higher average in both height and weight…

**MYTH #2 – A world of giants**

A) To shed light on this myth I will calculate the correlation coefficient between the draft rankings and the players height to determine how weak or strong the correlation is between draft ranking and height. I would also create a table to compare the height average of each pick in past NBA draft statistics.

B)

| t-Test: Paired Two Sample for Means | | |
|---|---|---|
| | | |
| | *Variable 1* | *Variable 2* |
| Mean | 30.5 | 200.633333 |
| Variance | 301.171548 | 70.0490934 |
| Observations | 240 | 240 |
| Pearson Corr | -0.1017453 | |
| Hypothesized | 0 | |
| df | 239 | |
| t Stat | -131.65688 | |
| P(T<=t) one- | 2.372E-225 | |
| t Critical one | 1.65125417 | |
| P(T<=t) two- | 4.743E-225 | |
| t Critical two | 1.96993941 | |

| Draft Ranking | Height Average |
|---|---|
| 21 to 30 | 201.775 |
| 1 to 10 | 201.5897436 |
| 11 to 20 | 201.2 |
| 41 to 50 | 200.8 |
| 31 to 40 | 199.375 |
| 51 to 60 | 198.5 |

C) So as you can see after looking at the data and calculating the correlation coefficient for the draft rankings and the players height we got a correlation of -1.102 which means that height is not a significant determinant of draft position. Especially after looking at the table, I compared every 10 draft picks of all 4 years, all the way up to pick 60 as the tallest players are usually drafted from pick 21 to 30 in the first round. However, after looking at the table I realized that the tallest players at an average of 201.52 metres are drafted in the first round (picks 1 to 30). So to conclude a player's height is not very important in the scouting ranking as the tallest players get drafted anywhere from the 21st pick to the 31st. However, the tallest players do appear to get drafted in the first round but since picks 1 to 10 do not hold the tallest players and the correlation appears to be negative then there does not seem like there is a strong correlation between draft rankings and a player's height.

MYTH #3

**A.**

a.)

We decided an appropriate statistical test you can use to answer which test(s) seem to be the most important to predict the players' draft rank would be a regression model. This is the best test to verify which variables have the most impact on our topic of interest.

b.)

SUMMARY OUTPUT

| *Regression Statistics* | |
|---|---|
| Multiple R | 0.537673722 |
| R Square | 0.289093031 |
| Adjusted R Square | 0.196366035 |
| Standard Error | 63.19819011 |
| Observations | 53 |

ANOVA

| | *df* | *SS* | *MS* | *F* | *Significance F* |
|---|---|---|---|---|---|
| Regression | 6 | 74712.27572 | 12452.05 | 3.117679 | 0.01199652 |
| Residual | 46 | 183724.5167 | 3994.011 | | |
| Total | 52 | 258436.7925 | | | |

| | *Coefficients* | *Standard Error* | *t Stat* | *P-value* | *Lower 95%* | *Upper 95%* | *Lower 95.0%* | *Upper 95.0%* |
|---|---|---|---|---|---|---|---|---|
| Intercept | -1704.881061 | 590.4505178 | -2.88742 | 0.0059 | -2893.396309 | -516.3658118 | -2893.39631 | -516.3658118 |
| 40 yard | 381.317606 | 103.4060109 | 3.687577 | 0.000596 | 173.1721017 | 589.4631103 | 173.1721017 | 589.4631103 |
| vertical | 0.573644284 | 4.31986191 | 0.132792 | 0.894937 | -8.121786741 | 9.26907531 | -8.12178674 | 9.26907531 |
| bench | -3.819879778 | 2.594361828 | -1.47238 | 0.147729 | -9.042059284 | 1.402299727 | -9.04205928 | 1.402299727 |
| broad jump | 0.398569691 | 1.791358596 | 0.222496 | 0.824913 | -3.207248142 | 4.004387525 | -3.20724814 | 4.004387525 |
| 3 cone | 20.38669244 | 51.8466807 | 0.393211 | 0.69598 | -83.97526296 | 124.7486478 | -83.975263 | 124.7486478 |
| shuttle | -8.408598381 | 71.10261074 | -0.11826 | 0.906377 | -151.5307306 | 134.7135339 | -151.530731 | 134.7135339 |

c.)

We conclude that the 40-yard dash is the most important stat to predict the players' draft rank because the p-value is below our assumed alpha of .05, it is also the only variable who's p-value is below our assumed alpha.

**B.**

a.) We can use a correlation test to answer the question of which test(s) are strongly associated with one another. This test assesses the strength of the association between the variables tested which are the 40-yard dash, vertical, bench, broad jump, 3 cones and shuttle.

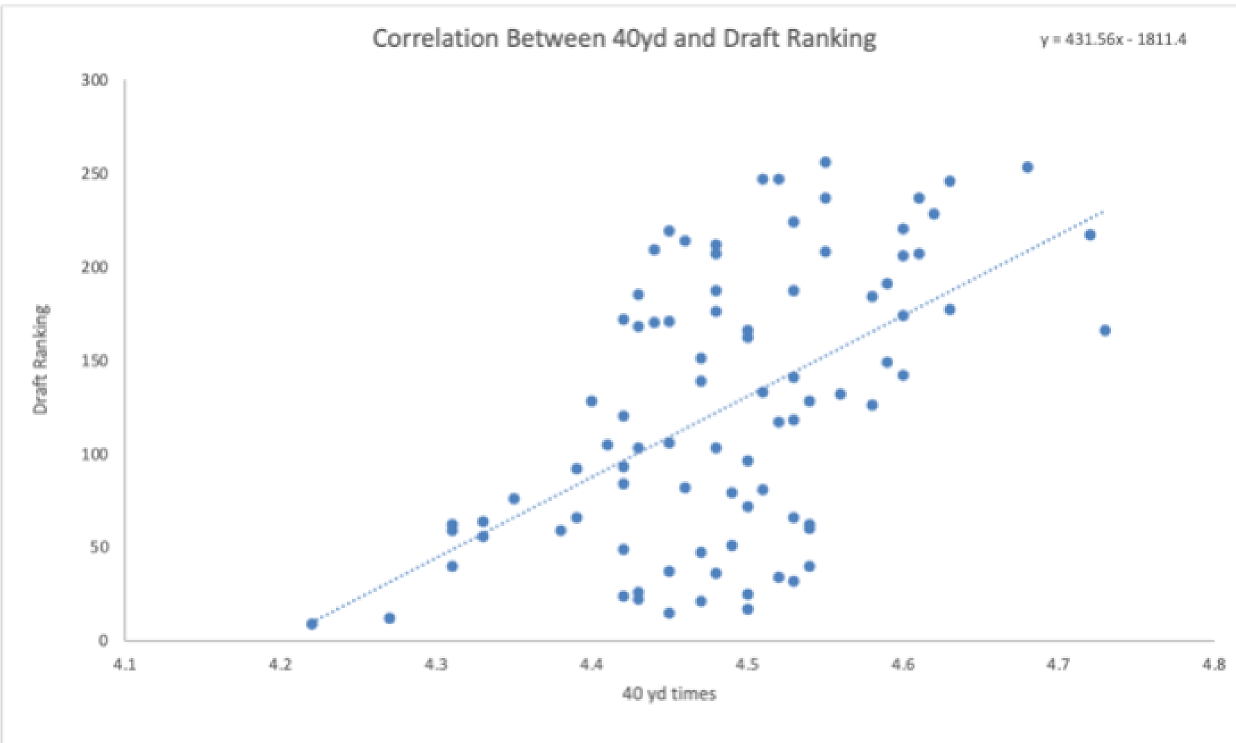
b.)

| | *40 yard* | *vertical* | *bench* | *broad jump* | *3 cone* | *shuttle* |
|---|---|---|---|---|---|---|
| 40 yard | 1 | | | | | |
| vertical | -0.2372034 | 1 | | | | |
| bench | -0.1434718 | 0.37553216 | 1 | | | |
| broad jump | -0.1970006 | 0.67119902 | 0.20114294 | 1 | | |
| 3 cone | 0.07712316 | 0.03541155 | 0.38334169 | 0.07535371 | 1 | |
| shuttle | 0.03467049 | -0.0031293 | 0.28606199 | 0.02430289 | 0.52693863 | 1 |

c.)
After looking at the correlation analysis table we conducted using Excel, we have determined that the two sets of tests that are the strongest associated with one another are vertical and broad jump (correlation of 0.67), the shuttle test and the 3 cone test (correlation of 0.53) and vertical and bench (correlation of 0.375). Since their correlations are positive that indicates that as one variable increases, the other variable tends to increase as well so for example if a player's vertical increases then it is likely that their bench increases. The tests that are not at all associated with one another are the 40-yard dash and vertical, bench and 40-yard dash, and broad jump and 40-yard dash, as they all have a negative correlation which means they have a very weak linear relationship. Also, a negative correlation coefficient indicates that as one variable increases, the other variable tends to decrease which means that if a player's 40-yard increases it is likely that their vertical, bench and broad jump could all decrease.

**C.**

To find the equation we decided to complete a regression analysis between the number the wide receivers were drafted and their 40 yard dash times to find the appropriate numbers for the equation. We also decided to make a regression table. Both are shown below.




Y = A + Bx
Y = Draft ranking

A = intercept

B = slope of line

X = 40 yard dash

Y = -1820.44 + ( 431.56X)

**D.**

While using the equation found in Question C we are able to determine that…

Y = A +BX

Y = -1820.44 + (431.56X)

Y = - 1820.44 + (431.56 x 4.7)

Y = 207.89200

Y = 207

A wide receiver with a 40 yard dash time of 4.7 seconds should approximately be drafted at 207th overall in the NFL draft.

**E.**

After creating a Z Score table, we found that for our 50th probable ranking receiver, he would have ran a 4.5 40 yd dash (z score = 0.12628). We calculated this by selecting each person's 40 yd time, subtracting it by the mean and dividing it by the standard deviation of the dataset. This 4.5 40 yd dash would ultimately give us Shelton Gibson as our 50th ranked receiver.