***Udacity Data Analyst Track***

**I. Into to Inferential Stats**

8. 1-Way ANOVA, Continued

* Need to do additional testing to figure out which means differ from each other and why = **multiple comparisons tests**
* multiple comparisons tests are not done until 1-way ANOVA is completed
* Common test = Tukey’s Honestly Significant Difference (HSD) 🡪 evaluates significance of difference between any 2 group’s means to allow us to make pair-wise comparisons
* Calculated similarly to Margin of Error (Z \* SE, or Z \* Sigma/Sqrt(n))
* Now comparing 3 or more samples, need new statistic, the **studentized range statistic** (**q) \* Sqrt[MS(w)/n]**
* Mean square for WG variability = Pooled variance = average squared deviation of each value from its respective group’s mean
* Therefore it’s really **q = S(p) / Sqrt(n)**
* **studentized range statistic** is from a table and adjusts the whole HSD so that’s is less likely we commit a Type I Error (reject h(0) when actually true)
* As # of samples increases, q adjusts upwards (increases when more groups are being compared) = less likely to make Type I error
* Remember when sample mean was further from population mean than the margin of error, it’s unlikely to have happened by chance
* Now, if 2 samples means are further apart than the HSD, the difference is honestly significant