# CP2403 - Project – Part 2 - ANOVA

First Name: Caleb

Last Name: Webster

|  |
| --- |
| **Step 1: hypothesis** |
| Investigative question: How does the phosphate concentration of the water change as water depth increases? |
| Null hypothesis (Ho): There is no difference between the mean phosphate concentration levels between different water depth categories (phosphate and depth are unrelated).  µ1 = µ2 = µ3 = µ4 |
| Alternative (Ha) hypothesis: There is a difference between the mean phosphate concentration levels for different water depth categories (phosphate and depth are related).  Not all µ are equal |
| **Step 2: Data Selection** |
| * CalCOFI bottle dataset   + Water depths between 0-200m (shallow), 200-400m (kinda shallow), 400-600m (kinda deep), and 600-1000m (deep)   + Phosphate content of water at these depths   + Null values dropped |
| **Step 3: Assess the evidence (ANOVA)** |
| F-statistics: 2.411e+05 |
| Prob(F-statistics): 0.00 |
| Mean values:  DEPTH\_CAT PO4uM  shallow 1.021587  kinda shallow 2.490355  kinda deep 3.070342  deep 3.249309 |
| STD values:  DEPTH\_CAT PO4uM  shallow 0.696238  kinda shallow 0.357878  kinda deep 0.194876  deep 0.197763 |
| **Step 4: Draw Conclusion** |
| The p-value of the ANOVA test is less than 0.05 (0.00), so we reject the null hypothesis and accept the alternative hypothesis. That is, there is a difference between the mean phosphate concentration levels for different water depth categories (phosphate and depth are related).  Implications: the increase in phosphorus concentration of the water with depth may indicate that more sardines are thriving in the shallower water, as sardines are high in phosphorus and the bodies of dead sardines release phosphorus into the water. However, phosphorus, although essential for all life, can be a pollutant that causes eutrophication. Eutrophication leads to algae growth and oxygen depletion, which in turn causes the depletion of fish species. |
| **Box Plot** |
| Figure 1: Box plot of water depth category vs phosphate concentration    It is clear from the box plot that there is an increasing trend between depth categories and phosphate concentration. As water depths increase, the phosphate increases quickly at first, but slows down and curves to form a seemingly logarithmic shape. |
| **Post Hoc Test** |
| Multiple Comparison of Means - Tukey HSD, FWER=0.05  =================================================================  group1 group2 meandiff p-adj lower upper reject  -----------------------------------------------------------------  deep kinda deep -0.179 0.001 -0.1957 -0.1623 True  deep kinda shallow -0.759 0.001 -0.7746 -0.7433 True  deep shallow -2.2277 0.001 -2.2427 -2.2128 True  kinda deep kinda shallow -0.58 0.001 -0.5897 -0.5703 True  kinda deep shallow -2.0488 0.001 -2.0573 -2.0402 True  kinda shallow shallow -1.4688 0.001 -1.475 -1.4625 True  ----------------------------------------------------------------- |
| New p-value: 0.05 / 6 = 0.0083  Mean Differences:   |  |  |  |  |  | | --- | --- | --- | --- | --- | |  | Shallow | Kinda Shallow | Kinda Deep | Deep | | Shallow |  |  |  |  | | Kinda Shallow | -1.4688 |  |  |  | | Kinda Deep | -2.0488 | -0.58 |  |  | | Deep | -2.2277 | -0.759 | -0.179 |  |   P-values:   |  |  |  |  |  | | --- | --- | --- | --- | --- | |  | Shallow | Kinda Shallow | Kinda Deep | Deep | | Shallow |  |  |  |  | | Kinda Shallow | 0.001 |  |  |  | | Kinda Deep | 0.001 | 0.001 |  |  | | Deep | 0.001 | 0.001 | 0.001 |  |   Conclusion from post hoc test: True for difference between all depth categories, so we can say that there is a significant difference in phosphate concentration between all depth categories. |