# CP2403 - Project – Part 2 - ANOVA

First Name: Kaung Khant

Last Name: Naing

|  |
| --- |
| **Step 1: hypothesis** |
| Null hypothesis (Ho):  Record indicator (Recind) is equal to Reported salinity of water (R\_SALINITY) – NO DIFFERENCE |
| Alternative (Ha) hypothesis:  Record Indicator (Recind) is not equal to Reported Salinity (R\_SALINITY) – THERE IS DIFFERENCE |
| **Step 2: Data Selection** |
| Data is taken from CalCOFI database – bottle table:  (Cast count <= 1)   * Record Indicator (RecInd) * Reported Salinity (R\_SALINITY) |
| **Step 3: Assess the evidence (ANOVA)** |
| F-statistics:  F-statistic: 0.4950 |
| Prob(F-statistics):  Prob (F-statistic): 0.488 |
| Mean values:  Means for reported salinity by recorded indicator  R\_SALINITY  RecInd  3 33.721667  7 33.807250 |
| STD values:  standard deviation for reported salinity by recorded indicator  R\_SALINITY  RecInd  3 0.329871  7 0.309932 |
| **Step 4: Draw Conclusion** |
| Based on the analysis above, it can be concluded that because the p value is 0.488 which is more than 0.05 (p>0.05) thus the null hypothesis is accepted. Therefore, Record indicator (Recind) is equal to Reported salinity of water (R\_SALINITY) (there is no difference). |
| **Box Plot** |
| Based on the box plot above it can be seen that the data has no outliers. Moreover, observed data (3) has higher reported salinity value than interpolated to standard depth value (7). However, interpolated to standard depth value (7) has higher median than observed data (3). |