# CP2403 - Assignment – Part 2 – Task 1: ANOVA Analysis

First Name:

Last Name:

|  |
| --- |
| **Step 1: hypothesis** |
| Null hypothesis (Ho):  There is no significant difference in the mean of the selected quantitative variable among the different categories of the selected categorical variable. |
| Alternative (Ha) hypothesis:  There is a significant difference in the mean of the selected quantitative variable among the different categories of the selected categorical variable. |
| **Step 2: Data Selection** |
| For this analysis, I selected   1. categorical variable "T\_qual" (Temperature Quality) 2. quantitative variable "T\_degC" (Temperature in degrees Celsius) |
| **Step 3: Assess the evidence (ANOVA)** |
| F-statistics: 93.4222200009212 |
| Prob(F-statistics): 1.1499626994142666e-151 |
| Mean values:  Mean values:  T\_qual  0.0 6.811818  1.0 8.247703  2.0 9.312022  3.0 10.206317  4.0 11.401623  5.0 13.576240  6.0 9.925842  7.0 17.408857  8.0 12.381496  Name: T\_degC, dtype: float64 |
| STD values:  STD values:  T\_qual  0.0 0.762674  1.0 0.711415  2.0 0.580184  3.0 0.673224  4.0 0.873658  5.0 1.726851  6.0 3.108433  7.0 0.729064  8.0 5.696665  Name: T\_degC, dtype: float64 |
| **Step 4: Draw Conclusion** |
| The ANOVA analysis resulted in a highly significant F-statistic (F = 93.42, p < 0.05), indicating that there is a significant difference in the mean temperature (T\_degC) among different quality categories (T\_qual). The mean values and standard deviations for each category suggest variations in temperature across quality categories.  Conclusion: We reject the null hypothesis. There is a significant difference in the mean temperature among different quality categories |
| **Box Plot** |
|  |