**IS432: Predictive Analytics for Business**

**Project #1: Multiple Regression Analysis & Factor Analysis**

**Dr. Soluade Fall 2024**

1. In the oil industry, waste that mixes with crude oil during production and transportation must be removed. Chemists have found that the water/oil mix electrically. Researchers at a US college conducted a series of experiments to study the factors that influence the voltage (y) required to separate the water from the oil. The seven independent variables investigated in the study are listed in the table below.
2. Determine the appropriate Regression Equation for this set of data.
3. At the 5% significance level, determine which variables significantly impact the voltage.
4. What is the 95% Confidence Interval for “volume”?
5. What is the degree of relationship between the independent variables and the voltage?
6. What is the 90% Confidence Interval for “volume”?
7. What percentage of variation in the voltage is explained by the independent variables?
8. What is the standard error of the regression? Explain.
9. Based on the regression output, suggest a more appropriate expression for the regression equation.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| # | Voltage  (y kw/cm) | Volume  (x1%) | Salinity  (x2%) | Temp-  erature  (x3%) | Time Delay  (x4 hours) | Concentration  (x5 %) | Span  (x6) | Particles  (x7 %) |
| 1 | 0.64 | 40 | 1 | 4 | 0.25 | 2 | 0.25 | 0.50 |
| 2 | 0.8 | 80 | 1 | 4 | 0.25 | 4 | 0.25 | 2.00 |
| 3 | 3.2 | 40 | 4 | 4 | 0.25 | 4 | 0.75 | 0.50 |
| 4 | 0.48 | 80 | 4 | 4 | 0.25 | 2 | 0.75 | 2.00 |
| 5 | 1.72 | 40 | 1 | 23 | 0.25 | 4 | 0.75 | 2.00 |
| 6 | 0.32 | 80 | 1 | 23 | 0.25 | 2 | 0.75 | 0.50 |
| 7 | 0.64 | 40 | 4 | 23 | 0.25 | 2 | 0.25 | 2.00 |
| 8 | 0.68 | 80 | 4 | 23 | 0.25 | 4 | 0.25 | 0.50 |
| 9 | 0.12 | 40 | 1 | 4 | 24 | 2 | 0.75 | 2.00 |
| 10 | 0.88 | 80 | 1 | 4 | 24 | 4 | 0.75 | 0.50 |
| 11 | 2.32 | 40 | 4 | 4 | 24 | 4 | 0.25 | 2.00 |
| 12 | 0.4 | 80 | 4 | 4 | 24 | 2 | 0.25 | 0.50 |
| 13 | 1.04 | 40 | 1 | 23 | 24 | 4 | 0.25 | 0.50 |
| 14 | 0.12 | 80 | 1 | 23 | 24 | 2 | 0.25 | 2.00 |
| 15 | 1.28 | 40 | 4 | 23 | 24 | 2 | 0.75 | 0.50 |
| 16 | 0.72 | 80 | 4 | 23 | 24 | 4 | 0.75 | 2.00 |
| 17 | 1.08 | 0 | 0 | 0 | 0 | 0 | 0 | 0.00 |
| 18 | 1.08 | 0 | 0 | 0 | 0 | 0 | 0 | 0.00 |
| 19 | 1.04 | 0 | 0 | 0 | 0 | 0 | 0 | 0.00 |

1. Re-run the data using Factor analysis and interpret your results.
2. Refer to the Weinberg & Abramowitz e-book and solve problem 16.8 page 531