PROJECT PART C

Reliable Housewares is a local store that sells many household items and issues its own credit card to its customers. The store manager wants to study the purchasing behavior of its "credit" customers. To that end, he has come to DeVry and asked our MBA students for help. The manager has brought with him data on five variables of 55 randomly selected credit customers.

* **LOCATION** (Rural, Urban, Suburban – Household location of the credit customer)
* **INCOME** (in $1,000's – be careful with this)
* **SIZE** (Household Size - number of people living in the household of credit customer)
* **YEARS** (the number of years that the customer has lived in the current location)
* **CREDIT BALANCE** ($ balance on customer’s store credit card)

Regression and Correlation Analysis

Using Excel perform the regression and correlation analysis for the data on CREDIT BALANCE (Y) and SIZE (X) by answering the following.

1. Generate a **scatterplot** for CREDIT BALANCE vs. SIZE, including the graph of the "best fit" line. Interpret.
2. Determine the **equation of the "best fit**" line, which describes the relationship between CREDIT BALANCE and SIZE. Interpret the values for slope and intercept.
3. Determine the **coefficient of correlation**. Interpret.
4. Determine the **coefficient of determination**. Interpret.
5. Test the **utility of this regression model** (use a two tail test with α =.05). Interpret your results, including the p-value.
6. Based on your findings in 1-5, what is **your opinion** about using SIZE to predict CREDIT BALANCE? Explain.
7. Compute **the 95% confidence interval for β1** (the population slope).  Interpret this interval.
8. What can we say about the credit balance for a customer that has a household size of **10**? Explain your answer.

In an attempt to improve the model, we attempt to do a multiple regression model predicting CREDIT BALANCE based on INCOME, SIZE and YEARS.

1. Using Excel run the multiple regression analysis using the variables INCOME, SIZE and YEARS to predict CREDIT BALANCE. State the equation for this multiple regression model.
2. Perform the Global Test for Utility (F-Test). Explain your conclusion.
3. Perform the t-test on each independent variable. Explain your conclusions and clearly state how you should proceed. In particular, which independent variables should we keep and which should be discarded.
4. Is this multiple regression model better than the linear model that we generated in parts 1-8? Explain.

Summarize your results from 1-12 in a report that is three pages or less in length and explains and interprets the results in ways that are understandable to someone who does not know statistics.

**Submission:** A report in Microsoft Word containing the summary report + all of the work done in 1-12 (Excel Output + interpretations) as an appendix.

**Report Format:**

1. Summary Report
2. Bullets 1-12 addressed with appropriate Excel outputs, graphs and interpretations. Be sure to number each bullet 1-12.

**Project Part C: Grading Rubric**

|  |  |  |
| --- | --- | --- |
| **Category** | **Points** | **Description** |
| Questions 1 - 10 and 12 - 6 pts. each | 66 | addressed with appropriate output, graphs and interpretations |
| Question 11 | 15 | addressed with appropriate output, graphs and interpretations |
| Executive Summary | 19 | writing, grammar, clarity, logic, and cohesiveness |
| **Total** | 100 | A quality paper will meet or exceed all of the above requirements. |

**Optional**

* Using an **interval**, estimate the **average credit balance for customers** that have household size of **5**. Interpret this interval.
* Using an **interval**, predict the credit balance for **a customer** that has a household size of **5**. Interpret this interval.