# Title: Predicting Monthly Credit Card Expenditures Using Multiple Regression Analysis

## Introduction

This report shows the results of a multiple regression analysis performed to predict monthly credit card spending. The purpose of this study is to identify the most important factors affecting monthly credit card spending using a data set containing a variety of variables including socio-demographic, economic and behavioral characteristics. is. The data were analyzed using a four-step analysis procedure and the results of the regression analysis are presented in this report. The report is structured as follows:  
Methodology overview, presentation of regression results, discussion of results, and conclusions.

## Methodology

Multiple regression analysis was performed on a dataset containing 4967 observations and many variables. From these variables, seven factors were selected to explain the variation in monthly credit card spending. These factors include:  
  
  
1. Retired  
2. Household Income  
3. CreditDebt  
4. Credit card  
5. CardTenure  
6. Card ItemsMonthly  
7. Internet  
  
The four-step analytical process used for this analysis is:  
  
  
1. Problem formulation:  
Identify the key factors that affect your monthly credit card spending.  
2. Model specifications:  
Create a multiple regression model using the selected factors.  
3. Model estimation:  
Estimate the model coefficients using ordinary least squares (OLS). 4. Model evaluation:  
Assess the goodness of fit of the model and the statistical significance of the estimated coefficients.

## Regression Output

The regression output provides various statistics that help in evaluating the model's performance and understanding the relationships between the independent variables and the dependent variable (monthly credit card expenditures). The output includes the following key statistics:

Multiple R (coefficient of determination): 0.6448

R Square: 0.4158

Adjusted R Square: 0.4150

Standard Error: 187.7454

Observations: 4967

F-Statistic: 504.2215

Significance F: < 0.0001

The coefficients for each independent variable, their standard errors, t-statistics, p-values, and 95% confidence intervals are also provided in the output.

## Discussion of Results

An R-squared value of 0.4158 indicates that the model explains approximately 41.58% of the variation in monthly credit card spending, indicating moderate explanatory power. The adjusted R-squared value, which accounts for the number of predictors in the model, is also close to the R-squared value of 0.4150. This means that the model is not overfitting and the variables selected are adequate to explain the variation in monthly credit card spending.  
  
The F statistic (504.2215) and corresponding significance level (<0.0001) indicate that the overall model is statistically significant. This means that at least one of the independent variables has a significant relationship with the dependent variable. The estimated coefficient for each independent variable provides information about the relationship between the predictor variable and monthly credit card spending. A simple interpretation of the coefficients is:  
  
  
1. Retired:  
A negative coefficient (-44.1349) indicates that retirees tend to spend less on credit cards each month than non-retirees. All other factors are constant. This relationship is statistically significant (p<0.0001).  
2. Household Income:  
The positive coefficient (1.2659) indicates that the higher the household income, the higher the monthly credit card spending, all other factors held constant. This relationship is statistically significant (p<0.0001).  
3. Credit Debt:  
The positive coefficient (2.9322) indicates that people with higher levels of credit debt tend to spend more on credit cards each month when all other factors are held constant. This relationship is statistically significant (p=0.0052). 4. Credit Card:  
A negative coefficient (-43.8210) indicates that credit card holders tend to spend less on credit cards each month than non-credit card holders. All other factors are constant. This relationship is statistically significant (p<0.0001).  
4. Card term:  
A positive coefficient (0.5900) indicates that people who have used credit cards longer tend to spend more on their monthly credit cards when all other factors are held constant. This relationship is statistically significant (p=0.0264).  
5. CardItemsMonthly:  
A positive coefficient (35.1607) indicates that population who buy more items with their credit cards in a month spend more on their credit cards in a month, all other factors held constant. indicates a tendency. This relationship is statistically significant (p<0.0001).  
6. The Internet:  
A positive coefficient (4.7663) indicates that internet users tend to spend more on their monthly credit card than non-internet users, all other factors being constant. This relationship is statistically significant (p=0.0106).

## Conclusion

A multiple regression analysis performed to predict monthly credit card spending shows that the model is statistically significant and moderately explanatory. Key factors identified to affect monthly credit card spending were retirement status, household income, loan debt, credit card ownership, credit card ownership, and number of items purchased with credit cards each month, and Internet use. These factors provide valuable insight into consumer behavior and can be used by financial institutions and credit card companies to develop targeted marketing strategies and improve customer segmentation. Further research may involve including additional variables or using alternative modeling techniques such as logistic regression or machine learning algorithms to improve the predictive performance of the model. Additionally, models can be validated against out-of-sample data to assess their generalizability and robustness.