**Consumer Spending in Canada: A 12-Year Provincial Analysis of Patterns, Key Drivers, Regional Variations, and Future Trends**

Shannon School of Business, Cape Breton University

A Capstone Project on Business Analyst Research Paper

In Partial Fulfillment of the Requirements in BUSS-5802-11

Instructor: Enayat Rajabi

**Graphs and Charts of EDA and Modeling Part**

Group E

Jereika Aiko Tan- Student ID: 0264312

Luzviminda Deinla- Student ID: 0274950

Shijie Liu - Student ID: 0276875

Simarpreet Kaur- Student ID: 0285168

**Table of Contents**

[1 Exploratory Data Analysis (EDA) 3](#_Toc183777290)

[1.1 Summary Statistics of Key Variables 3](#_Toc183777291)

[1.2 Trend Analysis: Overview of Spending Over Years 3](#_Toc183777292)

[1.2.1 Household Spending Trends by Category (2010-2021) 3](#_Toc183777293)

[1.2.2 Household Spending Trends by Province (2010-2021) 4](#_Toc183777294)

[1.3 Provincial Share of Household Expenditure 4](#_Toc183777295)

[1.4 Trend Analysis: Household Spending by Category Across Provinces 5](#_Toc183777296)

[1.5 Impact of Economic Factors (Disposable Income and Consumer Price Index) 7](#_Toc183777297)

[1.6 Impact of Demographic Factors (Age Groups, Household Type and Income Quintile) 10](#_Toc183777298)

[1.7 Regional Differences on Household Spending 11](#_Toc183777299)

[2 Predictive Model Fitting, Evaluation and Selection 13](#_Toc183777300)

[2.1 Data Transformation and Pre-processing 13](#_Toc183777301)

[2.2 Predictive Models Fitting and Evaluation 16](#_Toc183777302)

[2.3 Predictive Optimizing Linear Regression Performance with Ridge and Lasso Regularization Techniques 20](#_Toc183777303)

# Exploratory Data Analysis (EDA)

## Summary Statistics of Key Variables

A white table with black text

Description automatically generated

## Trend Analysis: Overview of Spending Over Years

### Household Spending Trends by Category (2010-2021)

A graph of a number of people

Description automatically generated with medium confidence

### Household Spending Trends by Province (2010-2021)

A graph showing the number of household spending trends

Description automatically generated

## Provincial Share of Household Expenditure

A pie chart with different colored circles

Description automatically generated

## Trend Analysis: Household Spending by Category Across Provinces

* **Alberta**

**A graph of different colored bars

Description automatically generated**

* **British Columbia**

**A graph of different colored bars

Description automatically generated**

* **Ontario**

**A graph of different colored bars

Description automatically generated**

* **Quebec**

**A graph of different colored bars

Description automatically generated**

* **Saskatchewan**

**A graph of different colored bars

Description automatically generated**

## Impact of Economic Factors (Disposable Income and Consumer Price Index)

* **Alberta**

**A graph of a graph showing the average income

Description automatically generated with medium confidence**

* **British Columbia**

**A graph of a line graph

Description automatically generated with medium confidence**

* **Ontario**

**A graph showing the average income

Description automatically generated**

* **Quebec**

**A graph of a graph showing the average income

Description automatically generated**

* **Saskatchewan**

A graph of a graph showing the average income

Description automatically generated with medium confidence

## Impact of Demographic Factors (Age Groups, Household Type and Income Quintile)

* **Age Group**

A graph of different colored squares

Description automatically generated

* **Household Type**

A graph of different colored squares

Description automatically generated

* **Income Quintile**

A graph of a bar chart

Description automatically generated with medium confidence

## Regional Differences on Household Spending

* **Food**

A graph of food and food

Description automatically generated with medium confidence

* **Housing**

**A graph of a number of people

Description automatically generated with medium confidence**

* **Transportation**

**A graph of different types of lines

Description automatically generated with medium confidence**

* **Discretionary Goods**

**A graph of a chart

Description automatically generated with medium confidence**

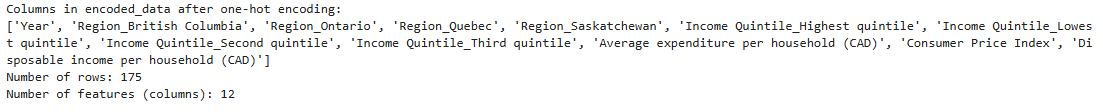
# Predictive Model Fitting, Evaluation and Selection

## Data Transformation and Pre-processing

A screenshot of a graph

Description automatically generated

**2.1.1 One-Hot Encoding of Categorical Variables**

****

**2.1.2 Outlier Detection and Handling**

A group of blue rectangular objects

Description automatically generated with medium confidence

**2.1.3 Multicollinearity Analysis: Variance Inflation Factor (VIF) and Correlation Matrix Heatmap**

A graph with numbers and a number

Description automatically generated with medium confidence

**A screenshot of a graph

Description automatically generated**

**2.1.4 Addressing Multicolliearity and Dimensionality Reduction**

A graph with a number of bars

Description automatically generated with medium confidence

**A screen shot of a graph

Description automatically generated**

## Predictive Models Fitting and Evaluation

* **Linear regression**

A graph with a red line

Description automatically generated

A graph with blue squares

Description automatically generated

* **Random Forest Regressor**

A graph with a red line and blue dots

Description automatically generated

A graph with text on it

Description automatically generated

* **XGBoost Regressor**

A graph with blue dots and red line

Description automatically generated

A graph with text on it

Description automatically generated

* **Neural Network Regressor**

A graph with a red line and blue dots

Description automatically generated

A graph with text and numbers

Description automatically generated with medium confidence

## Predictive Optimizing Linear Regression Performance with Ridge and Lasso Regularization Techniques

* **Actual Vs Predicted Expenditure**

A graph with a red line

Description automatically generated

* **Residuals Vs Predicted values**

A graph with blue dots and red lines

Description automatically generated

* **Distribution of residuals**

A graph with a line going up

Description automatically generated