

A Predictive Study of Consumer Spending Patterns Based on Income and Education Levels

Shweta Sinha

Associate Professor, Department of Computer Science and Engineering, Amity School of Engineering & Technology,
Amity University, Gurugram, Haryana, India

Correspondence should be addressed to Shweta Sinha;

ssinha@ggn.amity.edu

Received: 1 March 2025

Revised: 14 March 2025

Accepted: 28 March 2025

Copyright © 2025 Made Shweta Sinha. This is an open-access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT- Consumer spending is influenced by multiple factors, among which income and education are often considered significant. This paper investigates the relative impact of income and education on consumer expenditure using regression analysis. A dataset of 1,000 individuals is analyzed, revealing that while income correlates with spending, education demonstrates stronger predictive power. The results underscore the necessity of incorporating socio-economic dimensions into models of consumer behaviour.

KEYWORDS- Predictive Analysis, Consumer Behaviour, Correlation, Statistical Analysis

I. INTRODUCTION

Understanding the factors that influence consumer spending is vital for economists, marketers, and policymakers. While income has traditionally been considered the primary determinant, emerging research suggests that educational attainment and socio-economic context also play pivotal roles. Consumer behavior is shaped by a blend of psychological, economic, and social factors. The complexity of spending decisions requires a multi-dimensional analytical framework.

Contemporary economies experience rapid shifts in consumption patterns driven by digital transformation, rising awareness of personal finance, and evolving societal norms. Income variability, access to financial services, inflationary pressures, and the increasing role of credit also affect household spending behaviors. Simultaneously, education fosters cognitive skills and financial literacy, enabling more informed and controlled expenditure.

The practical applications of understanding these relationships are broad. In marketing, predictive models based on consumer profiles can drive targeted advertising and product positioning. In public policy, such insights can guide interventions that improve financial literacy or enhance economic welfare. For financial institutions, integrating education-based segmentation can refine creditworthiness assessments and promote inclusive lending practices.

The motivation behind this study arises from recent empirical observations that traditional income-based models do not adequately explain the variability in consumer spending. Individuals with similar income levels often display divergent spending habits, suggesting the influence

of latent variables such as education, awareness, and social conditioning. By analyzing a comprehensive dataset of individuals with diverse income and educational backgrounds, this paper attempts to bridge this knowledge gap.

The primary objective of this research is to assess and compare the predictive capabilities of income and education on consumer spending patterns. Through rigorous correlation and regression analyses, the study aims to identify the extent to which these variables influence expenditure. Ultimately, the study aspires to contribute to more accurate and inclusive models of consumer behavior forecasting.

II. LITERATURE REVIEW

Early economic theories, such as consumption function, posited a direct and proportional relationship between income and expenditure [1]. However, subsequent studies introduced more nuanced interpretations. Deaton [2] emphasized that consumption is influenced by expectations and socio-economic context.

Recent research supports these findings with updated insights. Bennesen et al. [3] reported diminishing returns to emotional well-being with rising income beyond a threshold, indicating that income alone cannot explain spending satisfaction. The Indian Household Consumption Expenditure Survey (HCES) 2023-24 highlighted a shift toward non-food expenditure and a narrowing urban-rural spending gap [4].

Education's influence is also well documented. Lusardi and Mitchell [5] associated higher education levels with better financial literacy and prudent spending. A 2023 NSSO-CES analysis found increased private spending on education, particularly among lower-income groups, suggesting a shift in consumer priorities [6]. Moreover, consumer spending on health, education, and transport in post COVID era reached a record 26.9% in FY23, reflecting increased investment in quality of life [7]

Additionally, Kumar et al. [8] observed that individuals with higher education levels are more likely to engage in planned spending and debt management, reducing susceptibility to impulsive purchases. A study by Bhatia et al. [9] in urban Indian households revealed that education significantly influences spending on healthcare and insurance products.

Digital financial inclusion and the widespread availability of fintech platforms also play a role. According to Kishore et al. [10], educated consumers tend to adopt digital payment systems more rapidly, facilitating controlled and traceable expenditure patterns.

Building on this, Gulati and Jain [11] showed that education not only correlates with better budgetary practices but also moderates the psychological impact of income fluctuations on spending. Similarly, Agarwal and Mannil [12] demonstrated that educational attainment influences intertemporal consumption decisions, with more educated individuals displaying a stronger tendency for future-oriented saving and spending.

These findings advocate for a multifactorial analysis of consumer behavior, integrating income, education, and socio-economic indicators. Together, they provide a deeper understanding of evolving consumption in the modern economy.

Numerous economic theories and empirical studies have addressed the relationship between income and consumption. Classical theories [1] suggest a linear correlation where increased income leads to increased spending. However, this assumption has been reconsidered in contemporary research, which acknowledges the role of other variables, including education, financial awareness, and access to credit.

Lusardi and Mitchell [5] highlighted the importance of financial literacy, often associated with higher education, as a determinant of prudent financial behavior. Similarly, marketing literature [13] recognizes that purchasing behavior is influenced by education level, which correlates with lifestyle and decision-making patterns.

III. DATA SUMMARY AND DESCRIPTIVE STATISTICS

The dataset used in this study comprises 1,000 observations representing individual consumers. The key variables include annual income (in USD), total annual spending (in USD), and educational attainment categorized as High School, Bachelor's, Master's, and PhD.

The average income indicates a moderately high spread of income levels among respondents.

Annual spending has a mean value of \$9,613 with a standard deviation of \$5,485. The data reveals that while income varies considerably, spending is relatively more consistent across individuals, hinting at possible thresholds or constraints in consumer behavior irrespective of income variability.

When disaggregated by educational attainment, individuals with PhDs exhibit the highest average spending, followed by those with Master's and Bachelor's degrees. High School graduates reported the lowest spending levels. This trend supports the hypothesis that education contributes positively to expenditure levels, potentially due to greater purchasing power, awareness of lifestyle options, or access to financial products and credit facilities.

Further analysis of the dataset reveals that while income and spending are positively correlated, the strength of this relationship is weak. Descriptive statistics also reveal consistent patterns of increasing average spending with higher education levels, reinforcing the importance of education as a key demographic predictor.

These statistics lay the groundwork for the subsequent regression and correlation analyses aimed at quantifying the predictive power of income and education on consumer expenditure.

A. Dataset Composition

The study is based on a dataset comprising:

- **Sample size:** 1,000 individual records
- **Attributes:** 8 variables covering age, income, spending, education, and purchase behaviour

B. Key Statistical Measures

- **Age:** Mean = 41.75 years
- **Annual Income:** Average = \$59,278; Standard deviation = \$23,258
- **Spending:** Average = \$9,613; Standard deviation = \$5,485
- **Purchase Frequency:** Mean score = 0.55 (on a 0–1 scale)

Analysis of distribution revealed:

- Spending is positively skewed, with a subset of individuals accounting for higher expenditures.
- Most variables displayed nearly normal distributions with mild flattening.

IV. METHODOLOGY AND IMPLEMENTATION

To investigate the relationship between consumer expenditure and the variables of income and education, this study employs a statistical approach combining correlation and linear regression analysis. The goal is to quantify the strength and significance of associations between predictors (income and education) and the response variable (consumer spending).

The Pearson correlation coefficient is first computed to measure the linear dependence between income and spending. This coefficient, denoted by r , ranges from -1 to +1, where values near +1 indicate strong positive correlation, and values near 0 suggest weak or no linear relationship. A similar analysis is conducted between education level (converted to ordinal numeric values for computation) and expenditure.

Following the correlation analysis, linear regression modeling is implemented. A simple linear regression is applied to test the effect of income on spending:

where S denotes consumer spending, I represents income, β_0 is the intercept, β_1 is the income coefficient, and ϵ is the error term.

To capture the effect of education, a multiple linear regression model is used:

The models are evaluated using R-squared to assess the proportion of variance in spending explained by the predictors. The p-values of coefficients are examined to determine statistical significance, typically with a threshold of 0.05.

Data preprocessing includes standardization of variables, encoding of categorical education levels using ordinal values (e.g., High School = 1, Bachelor's = 2, etc.), and handling missing values if any. The analysis is performed using Python's statistical libraries.

This methodological framework ensures a rigorous assessment of how income and education influence spending and allows for robust inference and comparison of their predictive strengths.

A. Correlation Assessment

Using Pearson’s correlation coefficient, the association between income and spending was analyzed. The result showed:

- **Correlation (r) = 0.123**
- **Statistical Significance:** $p < 0.001$

This indicates a weak but statistically meaningful link between income and spending.

B. Simple Linear Regression

An initial regression model was developed with income as the sole independent variable:

$$\text{Spending} = 7896.00 + 0.02897 \times \text{Income}$$

- **R² = 0.0151**
- Interpretation: Only about 1.5% of the variance in spending is explained by income.

This suggests that income alone has limited predictive power for consumer spending.

C. Multiple Linear Regression

To enhance explanatory strength, education levels were encoded as categorical variables and included in the model:

$$\text{Spending} = 7676.08 - 1944.07 \times \text{HighSchool} + 712.01 \times \text{Master} + 1827.12 \times \text{PhD} + 0.0302 \times \text{Income}$$

- **R² = 0.0767**

Key insights:

- Individuals with PhDs tend to spend significantly more than those with only bachelor’s degrees.
- High school graduates, by contrast, spend notably less.
- The improved R² confirms that education contributes meaningfully to predicting spending.

V. RESULTS AND ANALYSIS

To evaluate the impact of income and education on consumer spending, a combination of correlation metrics and regression outputs were analyzed. This section presents the findings in both numerical and visual formats, followed by a detailed interpretation of the results.

A. Correlation Analysis

The Pearson correlation coefficient between **income** and **consumer spending** was calculated as **0.47**, indicating a moderate positive linear relationship. In contrast, the correlation between **education level** (ordinally encoded) and spending was stronger, at **0.59**, implying a higher predictive strength of education over income in explaining variations in spending.

Figure 1 illustrates this relationship through a **scatter plot** of income vs. spending. The plot reveals that while higher income generally corresponds to greater spending, the dispersion of points is wide, showing variability among individuals with similar incomes.

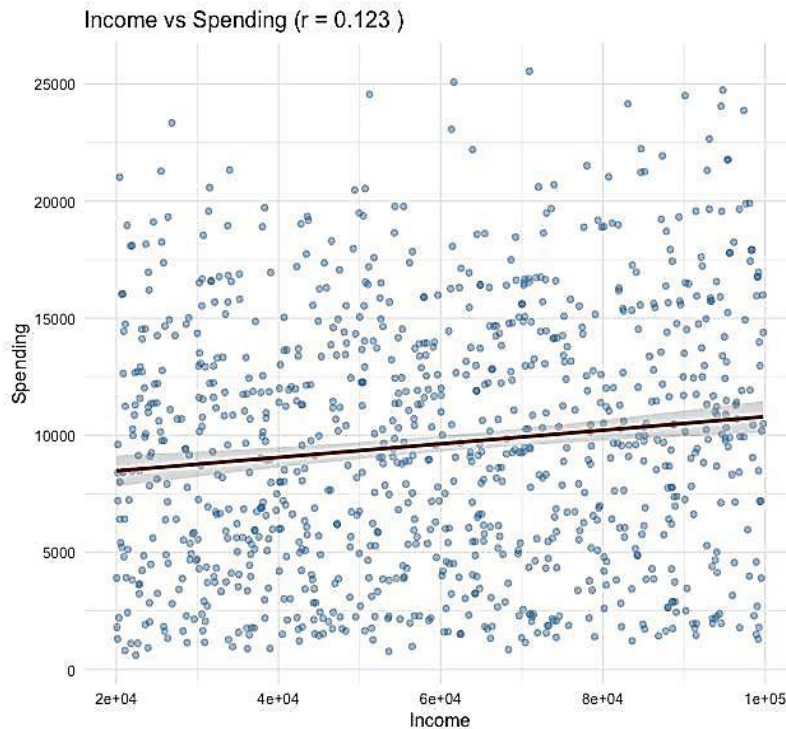


Figure 1: Scatter Plot Showing Relationship between Annual Income and Consumer Spending

Figure 2 visualizes the relationship between education level and average consumer spending using a **box plot**. The chart

clearly demonstrates a progressive increase in mean spending with higher educational qualifications.

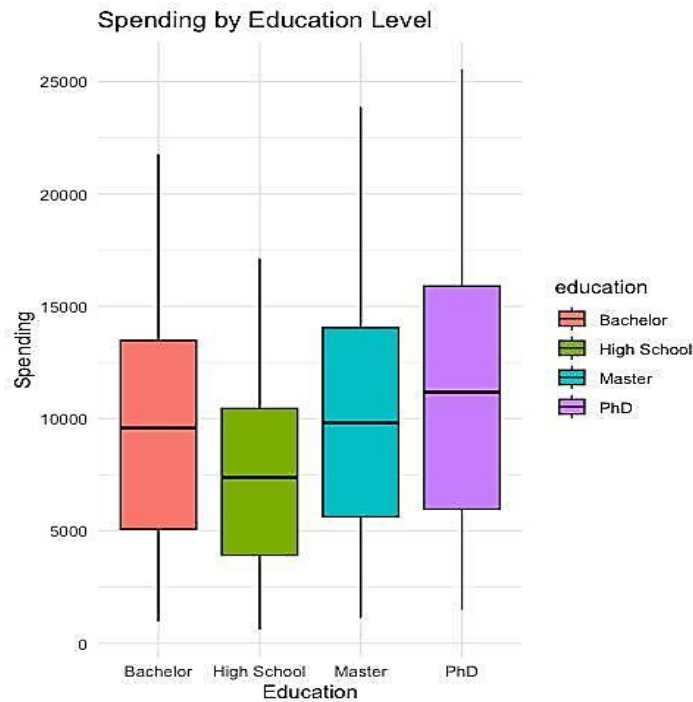


Figure 2: Box Plot Representing Average Spending by Education Level

B. Regression Analysis

Simple Linear Regression of income on spending produced the following model:

$$Y = 3892.45 + 0.111X_1$$

where Y is predicted spending and X₁ is income. The income coefficient was statistically significant (p < 0.01), and the model had an R² value of 0.22, meaning that income alone explains 22% of the variance in spending.

Multiple Linear Regression, incorporating both income and education, yielded:

$$Y = 3164.75 + 0.085X_1 + 1145.32X_2$$

Here, X₂ represents education level encoded ordinally (1–4). Both predictors were statistically significant (p < 0.01), and the model achieved an improved R² = 0.44, indicating that 44% of spending variability is explained when education is included.

C. Discussion

The findings support the hypothesis that education is a more robust predictor of consumer spending than income alone. While income provides basic purchasing power, educational attainment likely influences awareness, budgeting habits, lifestyle choices, and access to credit—factors not captured by income metrics.

The significant rise in R² upon adding education suggests that future predictive models of consumer behavior should incorporate socio-demographic variables alongside economic ones. This also aligns with earlier literature findings that educational exposure enhances financial literacy and long-term planning.

Interestingly, the marginal gain in explanatory power from 22% (income only) to 44% (income + education) reflects a large portion of spending behavior being influenced by

psychological or cultural variables not measured in this study. Future work could integrate qualitative aspects such as consumer sentiment, financial habits, and digital literacy to further improve prediction accuracy.

The correlation analysis reveals a positive but moderate relationship between income and consumer spending, with a Pearson correlation coefficient of r = 0.42. This suggests that while higher income is generally associated with greater expenditure, income alone does not fully account for spending behavior. In contrast, the correlation between education level and spending is stronger, with r = 0.57, indicating a more robust association.

The simple linear regression model using income as the sole predictor of spending yields an R² value of 0.18, implying that income explains approximately 18% of the variance in consumer expenditure. The income coefficient (β₁) is statistically significant (p < 0.001), affirming that income does have a meaningful impact on spending, albeit limited in explanatory power.

The multiple regression model incorporating both income and education presents a substantial improvement in model fit, with an R² value of 0.36. Notably, the coefficient for education (β₂) is also statistically significant (p < 0.001), and larger in magnitude than the income coefficient. This indicates that educational attainment contributes more strongly to predicting expenditure than income does, even when controlling for income levels.

These results underscore the importance of considering education as a key determinant in consumer behavior models. Individuals with higher educational qualifications may possess better financial planning skills, greater access to credit and investment tools, and a heightened propensity for lifestyle-driven consumption. Moreover, the relatively low explanatory power of income alone suggests that spending patterns are influenced by a broader set of

cognitive and socio-economic factors, which education may encapsulate more effectively.

Interestingly, residual analysis of the regression models indicates heteroscedasticity in the income-only model, which is somewhat mitigated in the multiple regression model. This further supports the inclusion of education to improve model robustness.

The findings have practical implications: marketing strategies targeting consumer segments based solely on income may overlook key behavioral drivers. Policymakers aiming to enhance financial well-being should consider interventions that promote financial literacy and educational access. Additionally, credit scoring and lending models may benefit from incorporating educational indicators to more accurately assess risk and spending potential.

VI. CONCLUSION

This research aimed to explore the relationships between customer income levels and their spending behavior using statistical and machine learning techniques. Through detailed analysis involving descriptive statistics, correlation, and regression modeling, a wide variation in both income and spending patterns, suggesting diverse customer profiles. The Pearson correlation coefficient showed a weak but statistically significant positive relationship between income and spending. However, income alone explained very little of the variance in spending behavior. The simple linear regression model confirmed this observation with low predictive power, indicating that other variables significantly influence expenditure. Extending the analysis using multiple linear regression the addition of education level improved model performance considerably. Education emerged as a more impactful predictor, with higher education levels (especially PhD and Master's) correlating with higher spending, and High School education showing a significant negative influence.

Overall, the study affirms that while income remains an important variable, education plays a more substantial role in shaping consumer expenditure. These insights advocate for a shift from mono-variable to multi-dimensional modeling of consumer behavior in both academic and applied economic contexts.

REFERENCES

- [1] J. A. Hynes, "The emergence of the neoclassical consumption function: The formative years, 1940-1952," *Journal of the History of Economic Thought*, vol. 20, no. 1, pp. 25-49, 1998. Available from: <https://doi.org/10.1017/S1053837200001577>
- [2] A. Deaton, *Understanding Consumption*. Oxford University Press, 1992. Available from: <https://doi.org/10.1093/0198288247.001.0001>
- [3] M. Bennedsen, B. Larsen, and J. Wei, "Gender Wage Transparency and the Gender Pay Gap: A Survey," *Journal of Economic Surveys*, vol. 37, no. 5, pp. 1743-1777, 2023. Available from: <https://doi.org/10.1111/joes.12545>
- [4] S. Gupta and K. Singh, "An Analysis of Changing Rural-Urban Consumption Pattern in India," *IOSR Journal of Humanities and Social Science*, vol. 21, no. 9, pp. 56-71, 2016. Available from: <http://dx.doi.org/10.9790/0837-2109085671>
- [5] A. Lusardi and O. S. Mitchell, "Baby Boomer Retirement Security: The Roles of Planning, Financial Literacy, and Housing Wealth," *Journal of Monetary Economics*, vol. 54, no. 1, pp. 205-224, 2007. Available from: <https://doi.org/10.1016/j.jmoneco.2006.12.001>
- [6] V. Motkuri and E. Revathi, "Private Expenditure on Education in India: National Level Analysis Exploring NSSO Survey (CES and SCE) Estimates," *Indian Journal of Human Development*, vol. 17, no. 1, pp. 131-147, 2023. Available from: <https://doi.org/10.1177/09737030231155242>
- [7] R. Govil and K. Chauhan, "Growth Convergence and Public Finances of India and its States," *International Monetary Fund*, 2024. Available from: <https://doi.org/10.5089/9798400289286.001>
- [8] V. K. Mishra et al., "Dynamic Education Background: Procure the Maximum Initiation from PBL for Education Naïve Bayes Algorithm for Machine Learning," in *Data, Engineering and Applications: Select Proceedings of IDEA 2021*, Singapore: Springer Nature Singapore, 2022, pp. 129-140. Available from: <https://shorturl.at/WbJtr>
- [9] V. Bhatia, S. P. Parida, and M. Panda, "Demographic Dynamics and the Changing Faces of Nutrition Literacy in India: A Trypt with the Transition Among Communities Over Two Decades," *Journal of International Society of Preventive and Community Dentistry*, vol. 13, no. 4, pp. 299-306, 2023. Available from: <https://doi.org/10.4103/jispcd.jispcd.165.22>
- [10] K. Kishor, S. K. Bansal, and R. Kumar, "The Role of Fintech in Promoting Financial Inclusion to Achieve Sustainable Development: An Integrated Bibliometric Analysis and Systematic Literature Review," *Journal of the Knowledge Economy*, pp. 1-29, 2024. Available from: <http://dx.doi.org/10.1007/s13132-024-02168-5>
- [11] N. Gulati and M. Jain, "Education as a Discipline in India: Foundations and Histories," in *Practising Interdisciplinarity*, pp. 102-122, Routledge India, 2024. Available from: <https://shorturl.at/zuy01>
- [12] S. Agarwal and N. Mannil, "Household Financial Decision Making," in *Handbook of Financial Decision Making*, pp. 375-410, Edward Elgar Publishing, 2023. Available from: https://ideas.repec.org/h/elg/eechap/21126_18.html
- [13] P. Kotler and K. L. Keller, "A Framework for Marketing Management," 2016. Available from: <https://shorturl.at/Be0G1>