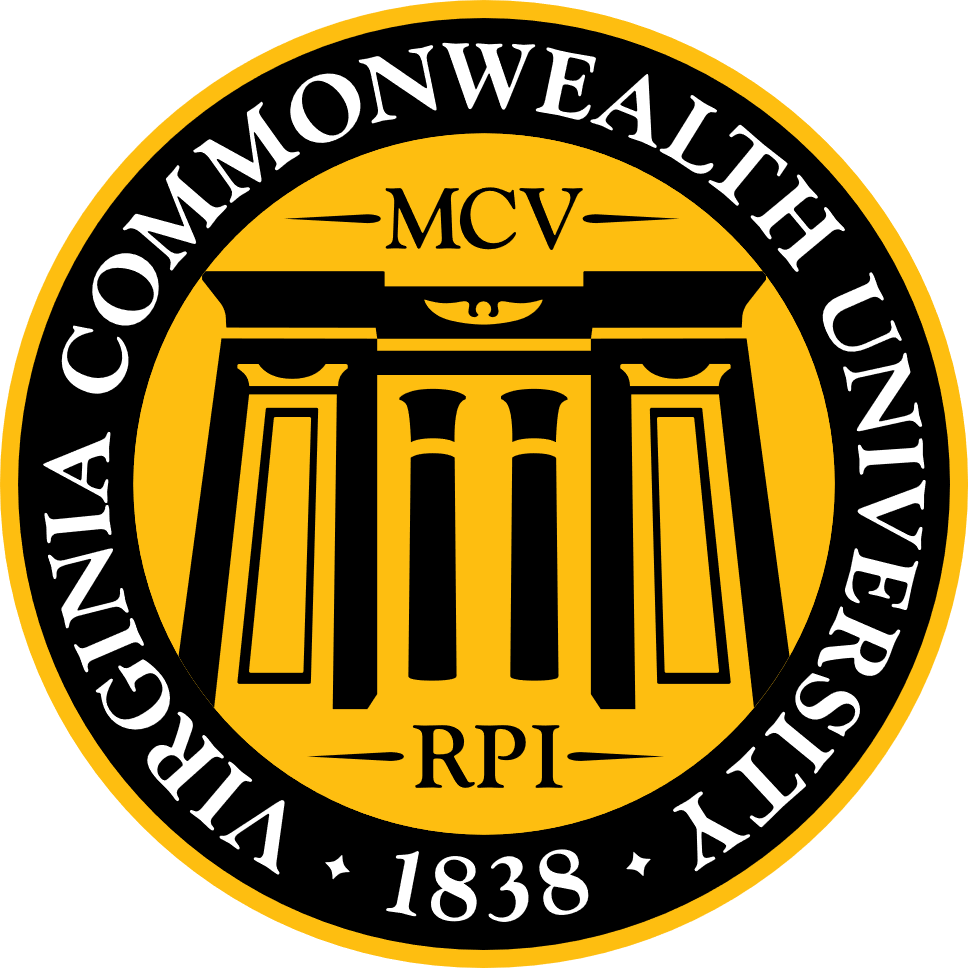
****

**VIRGINIA COMMONWEALTH UNIVERSITY**

**Statistical analysis and modelling (SCMA 632)**

**A2: Conducting a logistic regression, Confusion Matrix, ROC Curve and discussing Probit and Tobit regression**

**Arvind Srinivasan**

**V01107251**

**Date of Submission: 01-07-2024**

**CONTENTS**

|  |  |  |
| --- | --- | --- |
| **Sl. No.** | **Title** | **Page No.** |
| **1.** | Introduction | **1** |
| **2.** | Results | **1** |
| **3.** | Interpretations | **2** |
| **4.** | Recommendations | **4** |

**Introduction**

This report analyses the determinants of non-vegetarian dietary habits using data from the National Sample Survey Office (NSSO) 68th round. The dependent variable in this study is the binary indicator of non-vegetarian status, which is derived from various food consumption data points. We employ a Tobit model to account for the censored nature of the dependent variable, where the outcome is limited to binary values (0 for vegetarian and 1 for non-vegetarian).

The primary predictors examined in this analysis are:

* **Age**: The age of the individual.
* **MPCE\_URP**: Monthly Per Capita Expenditure on the usual resident population, representing the economic status.
* **Education**: The education level of the individual.

The NSSO dataset is comprehensive, containing information on demographic characteristics, socio-economic status, and detailed food consumption patterns. This allows for an in-depth examination of factors influencing dietary habits.

Our analysis aims to identify key socio-economic and demographic factors that influence the likelihood of individuals being non-vegetarian. Understanding these factors can help in policy formulation and targeted nutritional interventions. The results provide insights into how age, economic status, and education levels impact dietary choices, particularly in the context of vegetarianism and non-vegetarianism in India.

**Results**

The results from the Tobit model analysis provide insights into the determinants of non-vegetarian dietary habits among the surveyed population. The key predictors analyzed are Age, MPCE\_URP (Monthly Per Capita Expenditure on the usual resident population), and Education.

**Coefficients and Statistical Significance**

1. **Intercepts**:
   * **Intercept 1**: The coefficient is 25.71 with a highly significant p-value (< 2e-16), indicating a strong baseline influence on the likelihood of being non-vegetarian.
   * **Intercept 2**: The coefficient is 3.818 with a highly significant p-value (< 2e-16), further supporting the significant baseline effect.
2. **Age**:
   * The coefficient for Age is -0.01 with a p-value of 0.45959. This indicates that Age is not a statistically significant predictor of non-vegetarian status in this model. The negative sign suggests a slight tendency for older individuals to be less likely to be non-vegetarian, but this effect is not statistically supported.
3. **MPCE\_URP (Economic Status)**:
   * The coefficient for MPCE\_URP is -0.0001312 with a p-value of 0.00438. This negative and significant coefficient indicates that individuals with higher monthly per capita expenditure are less likely to be non-vegetarian. This suggests that economic status inversely affects non-vegetarian dietary habits, with wealthier individuals showing a preference for vegetarianism.
4. **Education**:
   * The coefficient for Education is -0.6782 with a highly significant p-value (< 2e-16). This negative coefficient suggests that higher levels of education are associated with a lower likelihood of being non-vegetarian. Educated individuals tend to prefer vegetarian diets more than their less educated counterparts.

**Model Fit and Diagnostics**

* **Log-likelihood**:
  + The log-likelihood value is -64831.43. While this value indicates the fit of the model, it is best interpreted comparatively with other models to assess relative fit.
* **Convergence Warning**:
  + The model generated a warning indicating that convergence was not obtained in 30 IRLS (Iteratively Reweighted Least Squares) iterations. This suggests potential issues with model convergence, which could be due to data characteristics or model specification.
* **Hauck-Donner Effect**:
  + The Hauck-Donner effect detected for Intercept 2 indicates potential anomalies in parameter estimation, possibly due to near-perfect prediction or other data-related issues.

**Interpretation and Implications**

The results highlight that economic status and education significantly influence dietary habits. Higher economic status and education levels are associated with a lower likelihood of being non-vegetarian. Age, however, does not have a significant impact on non-vegetarian status in this model.

These findings can inform policy makers and health practitioners about the socio-economic and educational factors influencing dietary choices. Interventions aimed at promoting healthy eating habits may need to consider these demographic factors to be more effective. For instance, promoting the benefits of vegetarian diets in economically well-off and educated communities may yield different results compared to other groups.

Overall, the analysis provides valuable insights into the socio-economic determinants of dietary habits, emphasizing the role of education and economic status in shaping non-vegetarian dietary choices.

**Objectives**

The primary objectives of this analysis are to:

1. **Identify Key Determinants**: Determine the socio-economic and demographic factors that significantly influence the likelihood of individuals being non-vegetarian. Specifically, we aim to assess the impact of Age, MPCE\_URP (Monthly Per Capita Expenditure on the usual resident population), and Education on non-vegetarian dietary habits.
2. **Model Dietary Habits**: Utilize a Tobit model to appropriately handle the censored nature of the dependent variable (binary outcome: vegetarian or non-vegetarian) and provide robust estimates of the influence of the key predictors.
3. **Inform Policy and Interventions**: Provide insights that can help policymakers and health practitioners develop targeted nutritional interventions and policies based on the socio-economic and educational profiles of the population.

**Business Insights**

The results of this analysis offer several valuable business insights:

1. **Targeted Marketing Strategies**:
   * **Economic Status**: Since higher MPCE\_URP is associated with a lower likelihood of being non-vegetarian, businesses in the food and beverage industry can tailor their marketing strategies based on economic segments. For example, vegetarian product lines might be more appealing to higher-income groups, while non-vegetarian products could be marketed more aggressively in lower-income segments.
2. **Educational Campaigns**:
   * **Education Level**: The significant negative relationship between education and non-vegetarian status suggests that more educated individuals prefer vegetarian diets. Companies can design educational campaigns highlighting the health and environmental benefits of vegetarianism, specifically targeting highly educated demographics.
3. **Product Development**:
   * **Product Innovation**: Understanding that wealthier and more educated individuals are more likely to be vegetarians, businesses can develop innovative vegetarian and vegan products catering to these segments. This could include high-end, premium vegetarian products that appeal to their lifestyle and preferences.
4. **Policy Formulation**:
   * **Public Health Policies**: Policymakers can use these insights to design public health campaigns promoting vegetarian diets in line with educational outreach programs. This can help address dietary-related health issues and promote sustainable eating habits.
5. **Market Segmentation**:
   * **Demographic Targeting**: Businesses can segment their market more effectively by considering the socio-economic and educational profiles of their target consumers. For instance, premium vegetarian brands might focus their distribution and marketing efforts in urban areas with higher education levels and incomes.
6. **Nutritional Programs**:
   * **Community Programs**: Organizations involved in community nutrition programs can use these findings to tailor their initiatives. Programs promoting balanced diets can be customized to address the specific needs and preferences of different socio-economic and educational groups.

**Recommendation**

1. **Develop and Promote Premium Vegetarian Products.**
2. Implement Educational Campaigns on Vegetarianism
3. Segment Market for Tailored Product Offerings