STATISTICAL ANALYSIS OF THE IMPACT OF EDUCATION ON INCOME

Introduction

The goal of this analysis is to determine whether there is a significant impact of education level on income using the **Customer Personality Analysis** dataset obtained from Kaggle. This project involves descriptive analysis, ANOVA testing, and T-Testing to evaluate hypotheses regarding the effect of education level on income.

Data Used

The dataset used in this analysis includes demographic and shopping behavior information for 2,240 customers. The two main variables analyzed are:

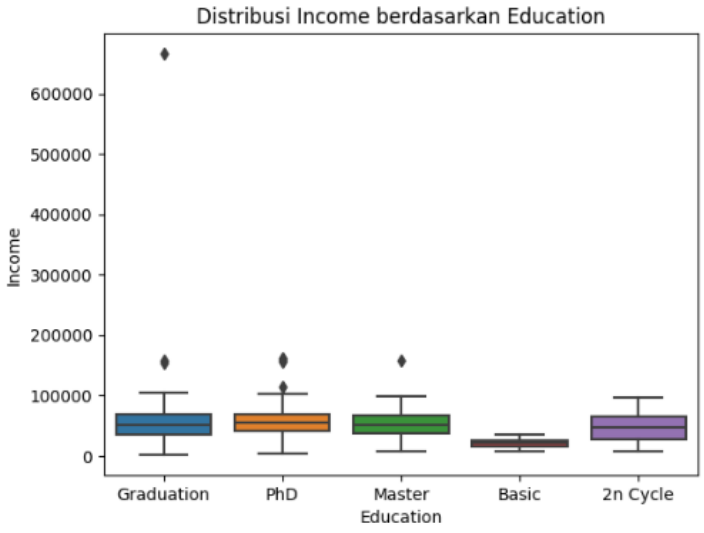
- **Education**: The education level of the customers.
- **Income**: The annual income of the customers.

Objectives

1. Perform **descriptive statistics** on the Education and Income variables.
2. Conduct **ANOVA** testing to determine if education level has a significant impact on income.
3. Perform a **T-Test** to compare income between two specific education groups (e.g., Graduation vs PhD).
4. Interpret the results and draw conclusions from the hypothesis testing.

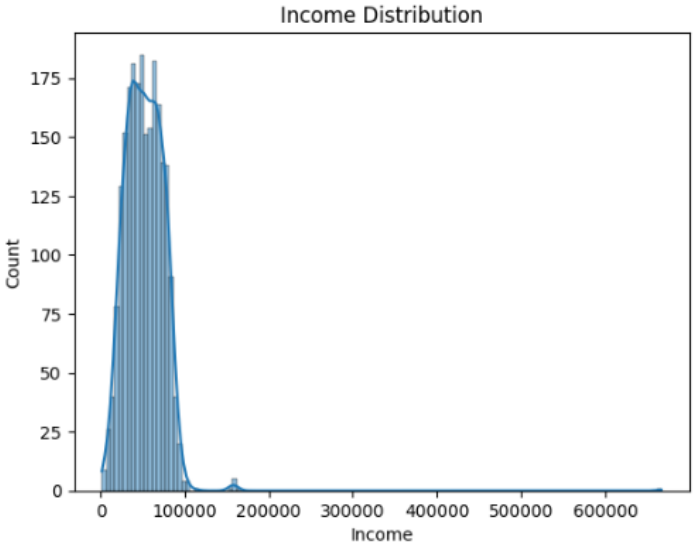
Descriptive Statistics

The descriptive statistics of the analyzed variables include:



Distribusi Income berdasarkan Education

- **Education**: Categories include Graduation, PhD, Bachelor, Master, and 2nd Cycle. Terdapat oulier pada Education dan perlu diatangani.



Income Distribution

- **Income**: Income distribution among customers shows considerable variability. Distribusi data income termasuk positive skew.

Hypotheses

ANOVA Test:

- **H0 (Null Hypothesis):** There is no significant difference in income across different education levels**.**
- **H1 (Alternative Hypothesis):** There is a significant difference in income across different education levels.

T-Test (Graduation vs PhD):

- **H0 (Null Hypothesis):** There is no significant difference in income between customers with Graduation and PhD education**.**
- **H1 (Alternative Hypothesis):** There is a significant difference in income between customers with Graduation and PhD education.

Analysis Results

ANOVA Test Results

- **F-statistic**: 39.6211
- **P-value**: 4.26e-32

Since the **P-value < 0.05**, we **reject the null hypothesis (H0)**. This indicates that there is a **significant impact** of education level on income. In other words, average income differs significantly across different education categories.

T-Test Results (Graduation vs PhD)

- **T-statistic**: -2.9811
- **P-value**: 0.0029

The T-Test results show a significant difference in income between customers with **Graduation** and **PhD** education. Since the **P-value < 0.05**, we **reject the null hypothesis (H0)**. This indicates that individuals with a **PhD** have significantly different average income compared to those with a **Graduation** level..

Conclusion

Based on the statistical tests:

1. The **ANOVA test** indicates that education level significantly affects income.
2. The **T-Test** between Graduation and PhD education shows a significant difference in income.

Overall, these results support that higher education levels are associated with higher income, and there is a significant difference between the education levels analyzed in this dataset.

Recommendations

Based on this analysis, companies may consider education level as an important indicator for customer segmentation in marketing and product strategies, as education is closely related to income

Referensi

- **Dataset: Customer Personality Analysis**