

1. INTRODUCTION

Tourism plays an important role in economic development by creating employment opportunities, generating income, and contributing to national growth. Many countries depend on tourism-related activities such as accommodation services, international tourist arrivals, and outbound travel to support their economies. As a result, understanding the factors that influence tourism contribution is important for effective planning and policy decisions.

This study examines the factors that affect tourism contribution using a provided dataset. Tourism contribution is measured using **tourism_percent**, which serves as the dependent variable in the analysis. The independent variables considered in this study include **inbound arrivals**, **outbound departures**, **employment**, and a dummy variable (**high inbound**) created to represent countries with high levels of inbound tourism.

The analysis was carried out using SPSS. Descriptive statistics were first used to summarize the data, followed by correlation analysis to examine relationships among the variables. Multiple linear regression was then applied as the main method to assess how the selected independent variables influence tourism contribution.

2. DATA AND METHODOLOGY

2.1 Dataset Description

The dataset used in this study contains tourism-related indicators collected across several observations. The variables include measures of tourism contribution, international tourist arrivals, outbound departures, employment, and other related indicators.

During the initial data inspection, missing values were identified in some variables, including domestic accommodation, inbound arrivals, outbound departures, employment, tourism percentage, and SDG indicators. To address this issue, missing values were treated using the mean substitution method in SPSS.

2.2 Variables Used in the Study

Dependent Variable:

- Tourism Percentage (**tourism_percent**)

Independent Variables:

- Inbound Arrivals
- Outbound Departures
- Employment
- High Inbound (Dummy Variable: 1 = High inbound arrivals, 0 = Otherwise)

The dummy variable was created to satisfy the requirement of including a categorical variable in the regression model.

2.3 Analytical Methods

The study employed three main statistical techniques using SPSS. Descriptive statistics were first used to summarize the basic characteristics of the data. Pearson correlation analysis was then conducted to examine the relationships among the variables. Finally, multiple linear regression analysis was applied as the main method to assess the impact of the independent variables on tourism contribution.

3. DESCRIPTIVE STATISTICS

Descriptive analysis was conducted to understand the basic characteristics of the data. The main statistics examined include mean, median, standard deviation, minimum, and maximum values.

3.1 Tourism Percentage

The mean tourism percentage is approximately **2.17**, indicating a relatively low average contribution of tourism across the observations. The median value is close to the mean, suggesting a fairly balanced distribution. The standard deviation of **0.97** shows moderate variation. The minimum and maximum values indicate noticeable differences across countries.

3.2 Inbound Arrivals

Inbound arrivals have a high mean value, reflecting large differences in the number of tourists received across observations. The large standard deviation suggests significant variation, meaning some countries receive far more tourists than others.

3.3 Outbound Departures

Outbound departures also show wide dispersion. The difference between minimum and maximum values indicates unequal outbound travel behaviour across observations.

3.4 Employment

Employment figures show substantial variation, with very high maximum values compared to the minimum. This suggests that tourism employment levels differ greatly across regions.

Overall, the descriptive statistics show that tourism indicators vary widely, which justifies the use of regression analysis to examine their influence on tourism contribution.

4. CORRELATION ANALYSIS

Pearson correlation coefficients were used to examine the relationships among variables.

- Tourism percentage shows a **weak positive correlation** with inbound arrivals and outbound departures.
- A **weak negative correlation** exists between tourism percentage and employment.
- Inbound arrivals have a **strong positive correlation** with the high_inbound dummy variable, which is expected since the dummy variable was created from inbound arrivals.
- Most correlations are weak and not statistically significant, indicating limited linear relationships among the variables.

These results suggest that correlation alone does not strongly explain tourism contribution, making regression analysis necessary.

5. MULTIPLE LINEAR REGRESSION ANALYSIS

Multiple linear regression was conducted with tourism percentage as the dependent variable and four independent variables.

5.1 Model Summary

The R-square value of **0.046** indicates that approximately **4.6%** of the variation in tourism percentage is explained by the independent variables. The adjusted R-square is negative, suggesting that the model has weak explanatory power.

5.2 ANOVA Test

The ANOVA result shows that the model is **not statistically significant** (p-value = 0.475). This means that the independent variables do not jointly explain tourism percentage in a statistically reliable way.

5.3 Regression Coefficients

- **Inbound Arrivals:** Positive coefficient, but not statistically significant.
- **Outbound Departures:** Negative coefficient, indicating an inverse relationship, but not significant.
- **Employment:** Negative and insignificant effect on tourism percentage.
- **High Inbound (Dummy Variable):** Negative and not statistically significant.

None of the independent variables show a statistically significant impact on tourism percentage at the 5% significance level.

6. EXPLANATION OF CALCULATIONS AND RESULTS

The regression coefficients were calculated using the Ordinary Least Squares (OLS) method in SPSS. Each coefficient measures the expected change in tourism percentage when the independent variable increases by one unit, holding other variables constant.

The low R-square value indicates that other factors not included in the model may have stronger effects on tourism contribution. The insignificant p-values suggest that the selected variables do not individually or jointly explain tourism percentage within the dataset.

Collinearity diagnostics show acceptable tolerance and VIF values, indicating no serious multicollinearity problem among the independent variables.

7. CONCLUSION

This study examined the factors influencing tourism contribution using multiple linear regression. Although inbound arrivals, outbound departures, employment, and a dummy variable were included, the results show that these variables do not have a statistically significant impact on tourism percentage in the dataset used.

The findings suggest that tourism contribution may be influenced by other factors such as government policy, infrastructure quality, security, and investment levels, which were not captured in this analysis. Future studies may include additional variables and larger datasets to improve explanatory power. Despite the weak results, this study demonstrates the proper application of descriptive statistics, correlation analysis, and multiple linear regression using SPSS.