

A STUDY ON TOURISM ACTIVITIES IN TANZANIA

BUSINESS UNDERSTANDING

Introduction

Tanzania's tourism sector generated USD 1 billion in revenues in 2020 versus 2.6 billion in 2019, being severely affected by the Covid-19 pandemic and its impact on international travel. The sector's contribution to GDP fell from 107% in 2019 to 5.3% in 2020. International tourist arrivals reached 620,900 in 2020, less than half the 1.5 million arrivals in 2019, the 1.4 million in 2018, and the 1.3 million in 2017.

Problem Statement

In Tanzania, the tourism sector is second in ranking in contributing to the national income. However, during the times of COVID, the industry took a huge blow due to travel restrictions. Now as countries are opening up, Tanzania should expect tourists to come back. The aim of this study is to give insight on the main tourism activity in Tanzania to invest more resources in.

Main Objective

To determine the main tourism activity in Tanzania

Objectives

1. To identify the countries or regions that contribute the most international tourists
2. To identify the age groups that contribute the most tourists
3. To identify the main tourism activity in Tanzania
4. To identify the most preferred tour arrangement by most tourists

Research Question

1. Which country or region of origin contributes the most international tourists?
2. Which age group contributes the most tourists and what attracts them the most to Tanzania?
3. Which is the main tourism activity in Tanzania?
4. What tour arrangement services do most tourists go for?

Hypothesis

H₀ Null -The average number of tourists who visit Tanzania for beach tourism are less or equal to the number of tourists who visit Tanzania to see wildlife.

H₁ Alternative(claim) - The average number of tourists who visit Tanzania for beach tourism is greater than the number of tourists who visit Tanzania to see wildlife.

Business Impact

Our analysis can be used by tours and travel operators and the Ministry of Tourism in Tanzania (Wizara ya Maliasili na Utalii Tanzania) to identify which activities to invest more resources in, to increase revenue from tourism.

DATA UNDERSTANDING

The data used for this research was provided by [Zindi](#). It is about tourist activities in Tanzania in the year 2018.

Below is a link to our github [link](#) to our analysis in Google Colab.

Data Description

The dataset is in CSV format and has **4809 rows** and **32 columns**. As per the research question, the columns likely to provide data to be used are:

- Country,
- Age-group,
- Travel with,
- Tour arrangements,
- Night in Zanzibar,
- Night in mainland,
- Main activity and purpose.

The dataset had a total of **1435 missing values** prompting the researcher to clean the data.

1. We chose to **replace the missing values** in the ‘**travel with**’ column with ‘**DNS**’, we assumed that respondents **chose not to give an answer**.
2. We chose to **replace the null values** in ‘**total male**’ and ‘**total female**’ with **0** because we assumed that they **were not filled** because they **were not there**.
3. There is presence of outliers in all numerical columns. However, we will not drop them as they are true values and not due to bad entry or anomalies.

DATA PREPARATION

Exploratory Data Analysis

The data was then subjected to **Exploratory Data Analysis: Univariate, Bivariate and Multivariate data Analysis** to help define the relationship between variables under study.

Univariate Analysis

- The frequency table from the age group column concluded that **most** of the **tourists** visiting were between the ages of **25 and 44 years**.
- **Most** of the **tourists** were **from the United States of America (302)**,
- A large percentage rarely used the tour packages
- **Wildlife tourism** is the **most popular activity**,
- **Leisure and holidays** came out as a common reason as to why most tourists visited the country.
- Our data is positively skewed
- The total male and total female columns have extreme values.

Bivariate Analysis

- From the pairplot visualization, it's evident that there is no relationship between the numerical columns
- The heat map concluded that the total male and total female are moderately correlated.

Multivariate Analysis

- The factors are dependent on each other
- The factors that determine most the main activity that tourists partake in Tanzania are age group, tour arrangement and package accommodation. Purpose, payment mode, package transport, package insurance, package food, and total cost are weakly determined. The rest don't.

Hypothesis testing

We first defined the null and alternate hypothesis.

Null Hypothesis - The average number of tourists who visit Tanzania for beach tourism are less or equal to the number of tourists who visit Tanzania to see wildlife.

Alternative hypothesis - The average number of tourists who visit Tanzania for beach tourism is greater than the number of tourists who visit Tanzania to see wildlife.

We will use a significance level of 0.05.

The sampling method will be stratified random sampling taking samples randomly from two populations, stratifying by age group.

From the testing, we found out that the p-value(0.245) is greater than the level of significance.

Therefore, we fail to reject the null hypothesis. The number of tourists who visit Tanzania for beach tourism are less or equal to the number of tourists who visit Tanzania for wildlife tourism.

CONCLUSION

The analysis above shows that the most popular tourism activity is Wildlife tourism. This is because Tanzania is known for its vast wilderness areas. They include the plains of Serengeti National Park, a safari mecca populated by the “big five” game (elephant, lion, leopard, buffalo, rhino), and Kilimanjaro National Park, home to Africa’s highest mountain.

RECOMMENDATIONS

- Introduce attractive offers and discounts
- Customize advertisements on social media
- Create travel technology
- Introduce a free travel visa
- Increase activities that attract people of ages 25 - 55
- Introduce packages that suit this age group travel needs
- The Zanzibar tourism agencies should strategize on increasing accommodation facilities for the tourists visiting during the day.

- The Tourism board should revisit its policies on Asian tourists. The number of Asian tourists is low.
- The country's tourism board should explore a range of alternative accommodations facilities to encourage more tourists to spend their leisure time in tourist sites.
- It is evident that a number of tourists are not aware of the travel packages which calls for a more aggressive marketing approach.