Analysis and Interpretation of Shipping and Flight Data

Course: DADS6001 Applied Modern Statistical Analysis

By: Miss Nattida Yotaprasert

Student ID: 6610422019

Advisor: Asst. Prof. Dr. Ramidha Srihera

Report Summary: This report aims to understand and analyze the marketing of shipping and air transportation. The analysis uses ship_data and flight_data obtained from transportation operations.

Date: May 23, 2567.

Title: Analysis and Comparison of Product Prices Transported by Sea and Air

Introduction:

The landscape of international trade is constantly evolving, with businesses navigating through various logistical challenges to maintain stability and optimize profits. One crucial aspect of this endeavor is selecting the most suitable transportation method. Analyzing and comparing the prices of goods transported by sea and air plays a pivotal role in facilitating strategic decision-making for businesses operating in the global market.

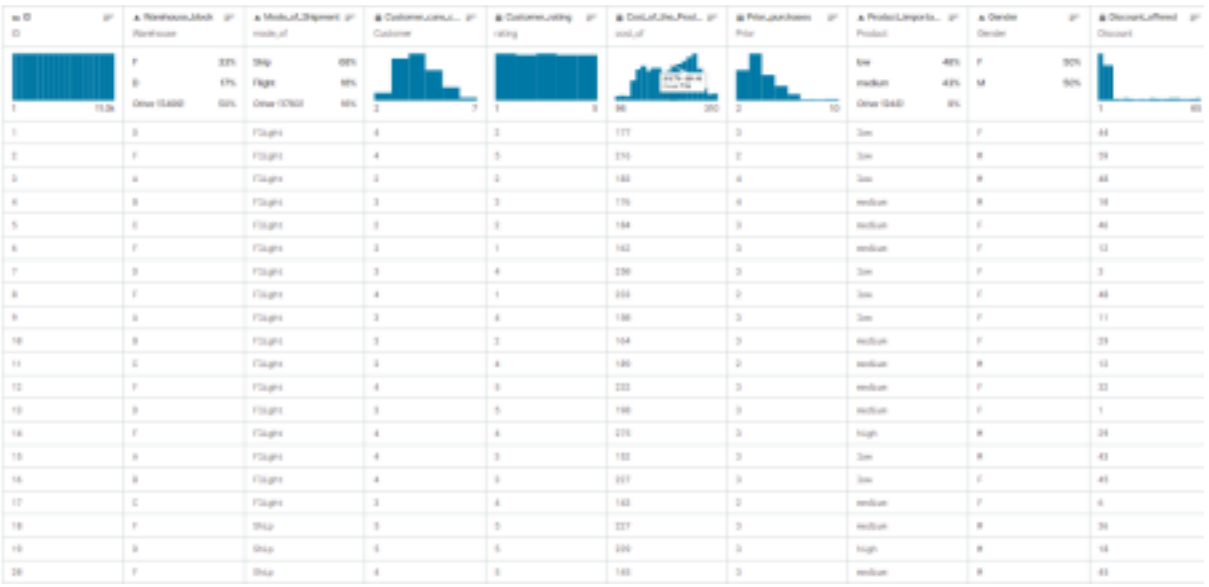
Objectives:

This research aims to conduct a comprehensive analysis and comparison of the prices of goods transported by sea and air, utilizing the ship_data and flight_data datasets. By focusing on the Cost of the Product as a primary determinant, this study seeks to provide valuable insights that can inform pricing strategies and profit analysis for businesses engaged in international trade.

Data Source:

The datasets used in this study were obtained from Kaggle, specifically the "On-Time Delivery" dataset provided by Willian Oliveira Gibin. The ship_data and flight_data subsets within this dataset contain information relevant to our analysis of transportation costs for goods shipped by sea and air.

Kaggle: https://www.kaggle.com/datasets/willianoliveiragibin/on-time-delivery



Step 1: Normality Test - Shapiro-Wilk Test:

Group 1 - Sea Transportation:

- H0: The data follows a normal distribution
- H1: The data does not follow a normal distribution
- Significance level: 0.05

The Shapiro-Wilk test conducted on the sea transportation data resulted in a p-value of 0.7935, indicating that it exceeds the significance level of 0.05. Thus, we fail to reject the null hypothesis (H0), concluding that the sea transportation data follows a normal distribution at the specified significance level.

Group 2 - Air Transportation:

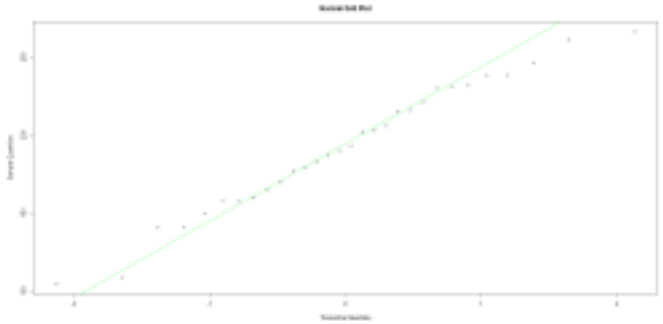
- H0: The data follows a normal distribution
- H1: The data does not follow a normal distribution
- Significance level: 0.05

Similarly, the Shapiro-Wilk test conducted on the air transportation data yielded a p-value of 0.5904, which exceeds the significance level of 0.05. Consequently, we fail to reject H0, indicating that the air transportation data follows a normal distribution at the specified significance level.

```
> shapiro.test(ship_data)

        Shapiro-Wilk normality test

data:  ship_data
W = 0.97882, p-value = 0.7935
```



```
> shapiro.test(flight_data)

        Shapiro-Wilk normality test

data:  flight_data
W = 0.97183, p-value = 0.5904
```

Step 2: F-Test for Equal Variances:

● H0: The variance of prices for sea transportation and the variance of prices for air transportation are not different

● H1: The variance of prices for sea transportation and the variance of prices for air transportation are different

● Significance level: 0.05

The F-test for equal variances produced a p-value of 0.7734, surpassing the significance level of 0.05. Therefore, we fail to reject H0, concluding that the variances of prices for sea transportation and air transportation do not significantly differ at the specified significance level.

```
var_ship <- var(ship_data)
var_flight <- var(flight_data)

var_ship
var_flight
```

```
1751.27471264368
1572.18965517241
```

```
var_test_result <- var.test(ship_data, flight_data)
var_test_result
```

```
        F test to compare two variances

data:  ship_data and flight_data
F = 1.1139, num df = 29, denom df = 29, p-value = 0.7734
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.530181 2.340316
sample estimates:
ratio of variances
          1.113908
```

```
var_test_result$p.value
```

```
0.773423774282557
```

Step 3: T-test:

● H0: The mean value of goods transported by air is greater than or equal to the mean value of goods transported by sea

● H1: The mean value of goods transported by air is less than the mean value of goods transported by sea

Results:

t = -1.7925, df = 58, p-value = 0.039

The mean price of goods transported by sea (192.3667) exceeds the mean price of goods transported by air (173.5000). With a p-value of 0.039, which is less than 0.05, we reject the null hypothesis (H0), providing sufficient evidence to conclude that the mean value of goods transported by air is lower than the mean value of goods transported by sea at the specified significance level of 0.05.

```
t.test(flight_data, ship_data, var.equal=TRUE,alternative='less')
```

```
        Two Sample t-test

data:  flight_data and ship_data
t = -1.7925, df = 58, p-value = 0.03913
alternative hypothesis: true difference in means is less than 0
95 percent confidence interval:
      -Inf -1.273056
sample estimates:
mean of x mean of y
 173.5000  192.3667
```

Conclusion:

Through rigorous statistical analysis, this study has revealed significant insights into the pricing dynamics of goods transported by sea and air. The findings suggest that, on average, the cost of transporting goods by sea is higher than that of air transportation. These results hold important implications for businesses engaged in international trade, informing their decisions regarding transportation methods and pricing strategies. However, it's imperative to acknowledge potential limitations of this analysis, such as data availability and other unaccounted variables. Future research could explore additional factors influencing transportation costs and delve deeper into the nuances of pricing strategies in the global marketplace.