|  |  |
| --- | --- |
| Business Analytics Skills for the Future-proofs Supply Chains - | **CASE STUDY**  **Optimizing logistics operation - Logisticsx**  Authors:  Kristijan Brglez |

Table of content

[DESCRIPTION OF THE COMPANY 2](#_Toc178609577)

[DECISION PROBLEM 3](#_Toc178609578)

[TASK 1 4](#_Toc178609579)

[TASK 1 Results 5](#_Toc178609580)

[TASK 2 6](#_Toc178609581)

[TASK 2 Results 7](#_Toc178609582)

[TASK 3 8](#_Toc178609583)

[TASK 3 Results 9](#_Toc178609584)

[TASK 4 10](#_Toc178609585)

[TASK 4 Results 11](#_Toc178609586)

[TASK 5 12](#_Toc178609587)

[TASK 5 Results 13](#_Toc178609588)

# DESCRIPTION OF THE COMPANY

Company LogisticsX is a third-party logistics (3PL) company, specialized in end-to-end logistics solutions for businesses across various industries. Headquartered in a central region, the company has build a reputation around efficient deliveries and reliable services related to transportation, warehousing, distribution and supply chain management. Leveraging its own fleet of vehicles and a robust network of rail and air transportation, it provides services to wide range of clients across multiple regions.

In recent years, LogisticsX has expanded its operation, driven by a mission to enhance efficiency through data-driven decision making. With the company’s expansion to new regions (A, B, C, D) the companies faced new logistical challenges related to terrain, customer expectation and utilizing transport modes. To stay in front of its competition, LogisticsX started to actively analyze its performance across critical metrics, such as delivery times, transport costs and customer satisfaction.

# DECISION PROBLEM

You are a data analyst working for LogisticsX. Due to the recent expansion of the company into new regions, the management believes that there is still room for further optimizing existing operations of the company. For this end they ordered you to analyze the collected data from the past year, to identify possible improvements that could improve the overall efficiency of operations.

You were provided with a data collections which includes the following variables:

* Sales, which are represented as monthly sales revenue (€) across the observed regions into which the company expanded.
* Delivery time, which was measured in days needed to finish a delivery.
* Transport mode, which indicated what transport mode was used for the delivery; this includes road (own vehicles), rail and air transport.
* Transport cost, which included the costs which incurred during the transportation phase presented in Euros (€).
* Delivery success, which indicated if the delivery was successful (1 = Yes) or the delivery was not successful (0 = No).
* Customer satisfaction, which indicated how satisfied the customers were with the delivery, presented on a scale from 1 to 5, with one being the least and 5 the most.
* Region, which represented the new regions into which the company expanded its new operations.

Based on the provided variables and the values for each of the variables, you need to now successfully finish the tasks you were given and provide a report, together with the solutions (screenshots) of each individual task.

**Figure 1 Example of data in Excel File**



# TASK 1

Before diving into deeper analyses, you need to test whether key variables like sales and delivery time follow a normal distribution, which will help determine the appropriate statistical tests to use.

* **Q-Q Plots**: Visualize the distribution of both sales and delivery times to see if they align with the expected normal distribution.
* **Shapiro-Wilk Test**: Perform this test on the same variables to statistically check for normality.

**Task**:

* Create Q-Q plots for both sales and delivery times.
* Perform the Shapiro-Wilk test to determine whether sales and delivery time data are normally distributed.

**Hypothesis**:

* **Null Hypothesis (H0)**: Data are normally distributed.
* **Alternative Hypothesis (H1)**: Data are not normally distributed.

Slika, ki vsebuje besede sličica, risanka, rumena, simbol

Opis je samodejno ustvarjen If you want you can do similar test for other variables as well.

# TASK 1 Results

**Shapiro-Wilk Test:**

* Sales (€): The Shapiro-Wilk test statistic is 0.996 with a significance value (p-value) of 0.829. Since the p-value is greater than 0.05, we fail to reject the null hypothesis, meaning that the sales data follow a normal distribution.
* Delivery Time (days): The Shapiro-Wilk test statistic is 0.991 with a significance value (p-value) of 0.251. Similarly, since the p-value is greater than 0.05, we fail to reject the null hypothesis, indicating that the delivery time data follow a normal distribution as well.

**Q-Q Plots:**

* The Q-Q plots for both sales and delivery times show that the data points mostly follow the straight line, confirming that the distributions are approximately normal. Some minor deviations in the tails are present, but they are not significant enough to suggest non-normality.

**Conclusion:**

For both sales and delivery times, the data appear to be normally distributed based on both the Q-Q plots and the Shapiro-Wilk test. Thus, we fail to reject the null hypothesis, indicating that the data are normally distributed.

# TASK 2

The management is particularly interested in whether the average delivery time meets their goal of 5 days and whether sales in Region A exceed €10,000. You will run two one-sample t-tests.

* **T-test 1**: Test if the average delivery time is significantly less than 5 days.
* **T-test 2**: Test if the average sales in Region A exceed €10,000.

**Task**:

* Perform a one-sample t-test to check if the average delivery time is less than 5 days.
* Perform a one-sample t-test to check if sales in Region A exceed €10,000.

**Hypothesis**:

* **Null Hypothesis (H0)**: Average delivery time equals 5 days; Sales in Region A are equal to €10,000.
* **Alternative Hypothesis (H1)**: Average delivery time is less than 5 days; Sales in Region A exceed €10,000.

Slika, ki vsebuje besede sličica, risanka, rumena, simbol

Opis je samodejno ustvarjen If you want you can do similar test for other variables as well.

# TASK 2 Results

**T-test 1: Average Delivery Time**

* Mean delivery time: 5.1716 days
* Test value: 5 days
* t-value: 37.046, df: 199, p-value: < 0.001.

The mean delivery time is slightly higher than 5 days, and the t-test is highly significant. Since the goal was to check if the average delivery time is less than 5 days, we fail to reject the null hypothesis. The results show that the average delivery time is not significantly less than 5 days.

**T-test 2: Sales in Region A**

* Mean sales: €11,877.6873
* Test value: €10,000
* t-value: 60.141, df: 199, p-value: < 0.001.

The mean sales in Region A are significantly higher than €10,000. Since the t-test is highly significant, we reject the null hypothesis that sales are equal to or less than €10,000. The results show that sales significantly exceed €10,000.

# TASK 3

Your supervisor suspects that higher transport costs may be linked to longer delivery times. Your job is to explore this potential relationship using correlation analysis.

* **Correlation**: Use Pearson correlation to test for a linear relationship between **transport costs** and **delivery times**.

**Task**:

* Calculate the Pearson correlation between transport costs and delivery times.

**Hypothesis**:

* **Null Hypothesis (H0)**: There is no significant correlation between transport costs and delivery times.
* **Alternative Hypothesis (H1)**: There is a significant correlation between transport costs and delivery times.

Slika, ki vsebuje besede sličica, risanka, rumena, simbol

Opis je samodejno ustvarjen If you want you can do similar test for other variables as well.

# TASK 3 Results

**Results:**

* Pearson correlation coefficient (r) between transport costs and delivery time: -0.053.
* p-value: 0.454.
* N (sample size): 200.

Interpretation:

The Pearson correlation coefficient of -0.053 suggests a very weak negative correlation between transport costs and delivery times. However, the p-value (0.454) is much greater than 0.05, meaning the result is not statistically significant.

**Conclusion:**

* We fail to reject the null hypothesis. There is no significant correlation between transport costs and delivery times based on the data provided.

# TASK 4

Next, management wants to know if the success of deliveries is influenced by the mode of transport used. You will perform a chi-square test of independence to investigate whether the two categorical variables (transport mode and delivery success) are related.

* **Chi-Square Test**: Analyze whether the mode of transport (road, rail, air) affects whether a delivery is successful (Yes/No).

**Task**:

* Perform a chi-square test to check if there is a significant relationship between transport mode and delivery success.

**Hypothesis**:

* **Null Hypothesis (H0)**: Transport mode and delivery success are independent.
* **Alternative Hypothesis (H1)**: There is a relationship between transport mode and delivery success.

Slika, ki vsebuje besede sličica, risanka, rumena, simbol

Opis je samodejno ustvarjen If you want you can do similar test for other variables as well.

# TASK 4 Results

**Chi-Square Test Results:**

* Pearson Chi-Square value: 0.616
* Degrees of Freedom (df): 2
* p-value: 0.735

Interpretation:

The p-value (0.735) is much greater than 0.05, indicating that there is no statistically significant relationship between the mode of transport (air, rail, road) and delivery success (yes/no).

**Conclusion:**

* We fail to reject the null hypothesis. There is no significant relationship between the mode of transport and the success of deliveries. This suggests that the success of deliveries is independent of the transport mode used.

# TASK 5

The company also wants to explore whether delivery times differ significantly depending on the mode of transport. Use a one-way ANOVA to test for differences in delivery times across the three transport modes (road, rail, air).

* **One-Way ANOVA**: Test if there are significant differences in delivery times based on the mode of transport.

**Task**:

* Conduct a one-way ANOVA to check for differences in delivery times across the different transport modes.

**Hypothesis**:

* **Null Hypothesis (H0)**: There are no significant differences in delivery times across transport modes.
* **Alternative Hypothesis (H1)**: There are significant differences in delivery times across transport modes.

Slika, ki vsebuje besede sličica, risanka, rumena, simbol

Opis je samodejno ustvarjen If you want you can do similar test for other variables as well.

# TASK 5 Results

**PRE-TEST Settings**

As can be observed from the associated file for the teachers, there is a special Column Transport\_Mode\_Num, which is not included in the Excel file. This Column is actually based on the Column Transport Mode, which normally enables all the other analysis, but in ANOVA the students will be confronted with error. To avoid this, SPSS enables creating new Variables. To accomplish this, the students need to go “Transform” 🡪 “Recode into Different Variables”. Input Variable should be the variable we want to change “Transport Mode”. Then in the Output Variable 🡪 Name; here write the name of the new variable, in our case it is Transport\_Mode\_Num. Then click Old and New Values 🡪 A new window will open, where you input into Old Value 🡪 “Value:” the value you have (for e.g. “Road”) and then into New Value 🡪 “Value:” the value we want (for e.g. “1”). You then click on “Add” button near window “Old🡪New:”. After repeating for all the Values (for e.g. “2” for “Rail”), we can click Continue and the new Variable with values will be created in our current DataSet. Similar can be done for other variables, when their “Type” doesn’t allow analysis.

**One-Way ANOVA Test**

* Sum of Squares (Between Groups): 19.975
* Degrees of Freedom (Between Groups): 2
* Mean Square (Between Groups): 9.988
* F-value: 2.604
* p-value (Sig.): 0.077

Interpretation:

The p-value (0.077) is greater than the significance level of 0.05, which means that we fail to reject the null hypothesis. There is no statistically significant difference in delivery times across the three transport modes (road, rail, air).

**Post Hoc Tests:**

The post hoc tests (Tukey HSD and Bonferroni) confirm that there are no significant pairwise differences between any of the transport modes in terms of delivery times.

However, based on the results, we can look at the mean delivery times from the ANOVA results to get an idea of which transport mode tends to have shorter delivery times:

* Air (1) has a mean delivery time of 5.3932 days.
* Rail (2) has a mean delivery time of 4.7313 days (the shortest).
* Road (3) has a mean delivery time of 5.4036 days.

Business Analytics Skills for the Future-proofs Supply Chains -