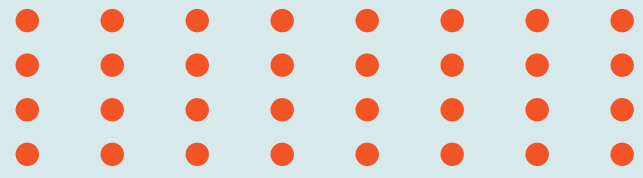Data Analyst Project
2025

SALES FORCE TRAINING
ANALYSIS

By: Muhammad Fakhri Azhar

Introduction

Hi! I'm Muhammad Fakhri Azhar, a physics graduate with a strong passion for data analysis. This project is part of my learning journey in turning data into insights.

Course License:

- Data Science Bootcamp @Kelas Work by Kelas.com
- Data Analyst Mini Course @RevoU
- Ms.Excel Short Class @MySkill
- Computer Training @FMIPA UNNES

Contact Info

Email : mfkriazh57@gmail.com

Phone : 0857-2454-9367

LinkedIn : Muhammad Fakhri Azhar

Portfolio : Click here

GitHub : mfakhriazhar

**Project Code Details on Github :**

https://github.com/mfakhriazhar/sales-force-training/blob/main/Case_01_Final_Project.ipynb

Overview

Company X wants to increase sales after conducting training for sales workers. However, it needs to analyze whether the training is really effective in improving sales performance. Previously, the average sales per transaction was $100, and to evaluate the effectiveness of the training, the company collected recent data from 25 sales workers.

| Transactions Amount	100	150	50	100	130	120	100	85	70	150
	150	120	50	100	100	140	90	150	50	90
	120	100	110	75	65					
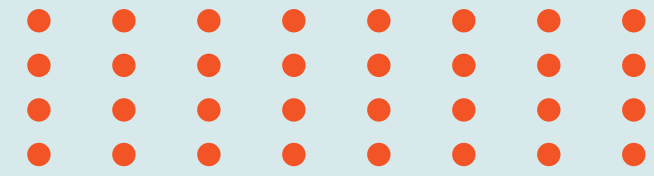
Project Goals

The main objective of this project was to measure the impact of training on sales performance and provide recommendations based on the analysis. This process included statistical analysis of recent sales data, evaluation of data variability, and hypothesis testing to determine if there was a significant change.
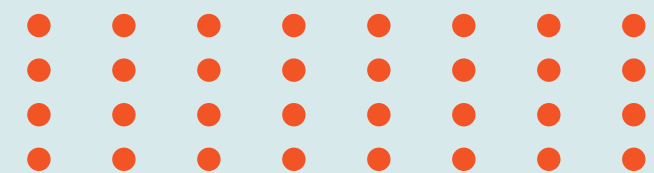
The steps in this project mechanism will be further explained through flowcharts and the analysis methods used, before finally presenting the findings, insights, and recommendations based on the results obtained.

Flowchart
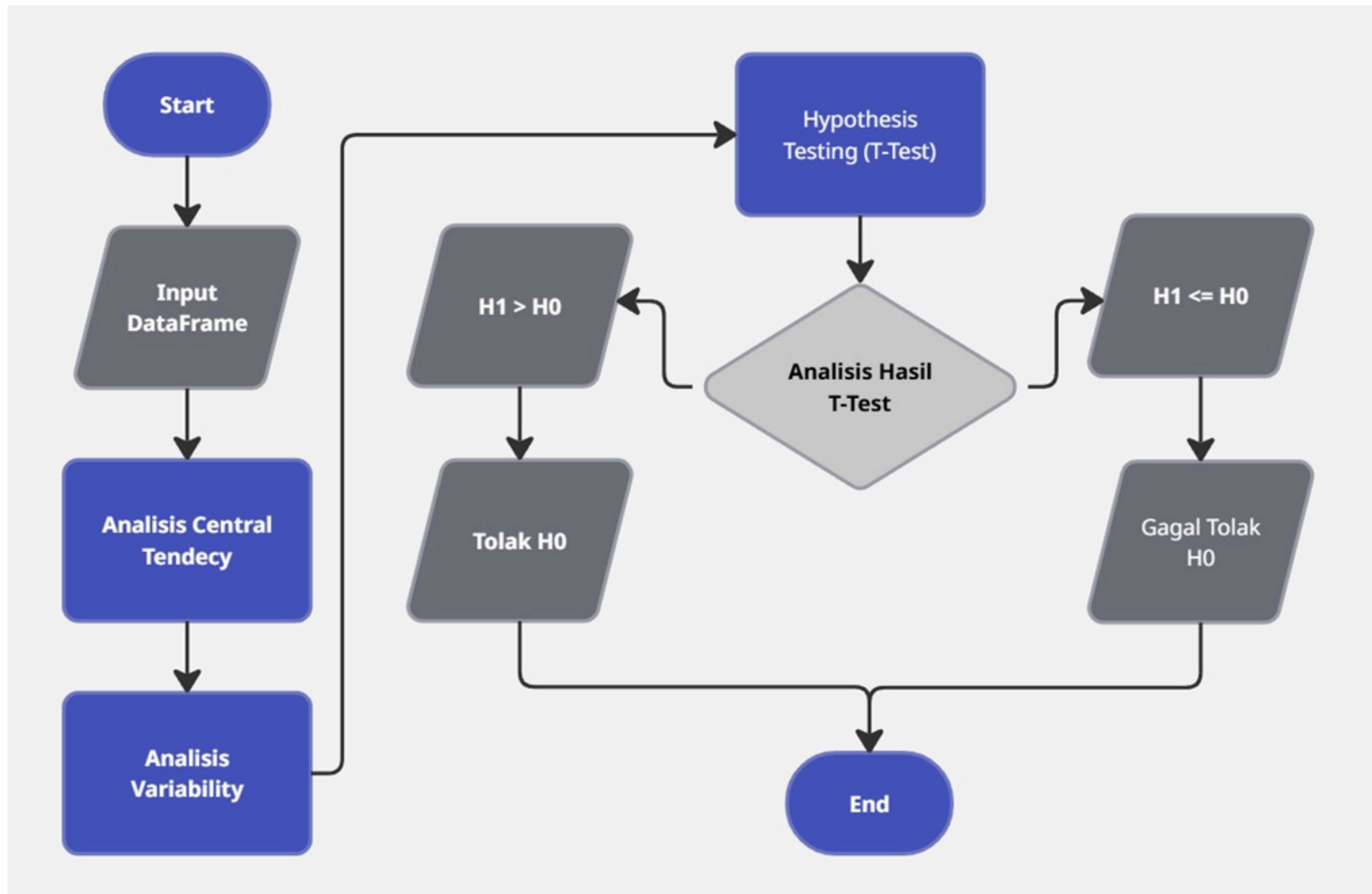
Methods

To analyze the effectiveness of sales worker training in increasing sales, a data-driven approach that includes several statistical methods is used. These methods help provide a comprehensive overview before drawing conclusions and providing recommendations for the company.

The process of working on the project is carried out through several stages as follows:

- Central Tendency Analysis
- Variability Analysis
- Hypothesis Testing (T-Test)

Central Tendecy Analysis

```python
TransactionAmount = TransactionAmount['TransactionAmount'].tolist()
print('Mean: ',statistics.mean(TransactionAmount))
print('Median: ',statistics.median(TransactionAmount))
print('Mode: ',statistics.mode(TransactionAmount))
```

```
Mean:   102.6
Median:   100
Mode:   100
```

Central Tendecy :

Calculate the mean, median, and mode to understand the trend of the data after training.

Insights:

- The mean sales increased to $102.6 from $100 previously, indicating an increase after training.

- The median remains $100, indicating half of the workers still have sales below or equal to this figure.
- The increase in the mean is likely influenced by a few workers with high sales (outliers).
- The mode remained $100, indicating most workers still recorded sales around this figure.

Variability Analysis

```python
print('Variance: ', np.var(TransactionAmount))
print('Std Deviasi: ', np.std(TransactionAmount))
print('Range: ', np.max(TransactionAmount) - np.min(TransactionAmount))
print('Quartile 1: ', np.quantile(TransactionAmount, 0.25))
print('Quartile 2: ', np.quantile(TransactionAmount, 0.5))
print('Quartile 3: ', np.quantile(TransactionAmount, 0.75))
```

```
Variance:  972.2399999999999
Std Deviasi:  31.18076330047101
Range:  100
Quartile 1:  85.0
Quartile 2:  100.0
Quartile 3:  120.0
```

Variablity:

Measuring variance, standard deviation, range, and quartile to see the distribution of data and variation in sales between sales.

Insights:

- The standard deviation of 31.18 and variance of 972.24 indicate a high difference in sales between workers.

- The range of 100 reflects the large difference between the highest and lowest transactions.
- The majority of workers have sales in the 85-120 range, but the presence of outliers causes large variations.
- The impact of training is uneven, so further evaluation is needed.

Hypothesis Testing: T-Test

- In the T-Test we will determine whether or not there is a significant difference after training by rejecting or accepting H0 and H1.
- One Sample T-Test is used to compare the new sales average with the previous average ($100).
- If the p-value < 0.05, then the training is considered to have a significant impact on increasing sales. Conversely, if the p-value > 0.05, then the training is considered not to have a significant impact on increasing sales. Otherwise, there is no significant change.
- The Critical Region is used to compare with the t-statistic value to determine whether the hypothesis test results are strong enough to reject H_0 (null hypothesis) or not.
- If the t-statistic falls within the critical region → Reject H_0, meaning there is a statistically significant difference.
- If the t-statistic does not fall within the critical region → Fail to reject H_0, meaning there is not enough evidence to declare a significant difference.

Insights:

After conducting the T-test, it was found that the **P-value (0.68) was greater than alpha (0.05)**, so the training **failed to reject H_0**, which means that there is **no significant difference between the average sales after the training ($102.6) and before the training ($100)**. This indicates that **the training has not had a significant impact on increasing average sales**. In addition, since the **t-statistic is smaller than the critical region**, it **fails to reject H_0**, which means there is **not enough evidence that the average sales after the training is significantly different from before**. In other words, the training conducted has not had a significant effect on increasing sales.

```
T-Statistic: 0.4085001556802841
P-Value: 0.6865284813438117
Degree of Freedom: 24
Gagal tolak H0, tidak ada perbedaan signifikan dengan $100.
```

```python
#H1: larger (right side)
crit_region = stats.t.ppf(1-alpha, 24)
print(f'Critical Region: {crit_region}')

Critical Region: 1.7108820799094275
```

Conlusion

- The analysis showed that the average sales increased to $102.6, but the t-test revealed that this change was not statistically significant (p-value = 0.686).
- This means that the training provided has not had a meaningful impact on increasing sales.
- If the company only relied on averages without statistical tests, the results could appear as if there was an increase.
- However, with more in-depth analysis, it is found that this increase is not significant, so the company needs to review the effectiveness of the training program.
- Other solutions to increase sales Company X can re-evaluate the training materials, add other strategies such as incentives, additional training, or personalized approach in sales to increase its effectiveness.

Data Analyst Project | 2025

Thank You

Muhammad Fakhri Azhar