

Saleh AL-Frhan  
T5 Bootcamp  
6 November 2021

# Documentation

## Abstract

For most companies, sales is the largest source of revenue, sales does not represent profit. To determine actual profit from sales we need to find the most profitable sales in the sales market and the items with the lowest cost and the highest selling in the market and the items with the highest cost and lowest selling in the market.

## Design

The project will show the most profit of selling item by compare between the item cost and revenue, in addition this project is one of the T5 Data Science BootCamp requirements.

## Data

- The datasets are generated through random logic in VBA.
- This data set can be categorized under "Sales" category.
- The sample in this project are groups.
- The dataset is provided in CSV format. It contains 50,000 record, each record has 14 features which will be mentioned.

## Algorithms

- Cleaning the data.
- Visualized the data.
- Comparing between the item cost and revenue.
- Visualized the result from the cleaned data.

## Tools

- Python visualization libraries (matplotlib).
- Data analysis in pandas for discovering the data and train a model.