

QUESTIONS:

- 1- What are the most profitable sales in the sales market in the timeline given ?
- 2- What are the items with the lowest cost and the highest selling in the market ?
- 3- What are the items with the highest cost and lowest selling in the market ?
- 4- What are the most stable item regarding the profit in the market ?

DATA DESCRIPTION:

- Data provided by (E for Excel) has been used in this project.
- 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.

FEATURES:

- | | |
|--------------------------------------|---------------------------------|
| • Region | • Ship date |
| • Country | • Unit soled |
| • Item type | • Unit price (in dollars \$) |
| • Cells channel (offline / online) | • Unit cost (in dollars \$) |
| • Order priority | • Total revenue |
| • Order date | • Total cost (in dollars \$) |
| • Order ID | • Total profit (in dollars \$) |

TOOLS:

Tools that will be used to achieve the goal of this study are:

- Python visualization libraries (matplotlib)
- Data analysis in pandas for discovering the data and train a model
- Jupyter notebook
- Explore the data and come up with EDA phases then use a model to fit the data

THE MVP GOALS:

- Explore and Analyze the data.
- To prioritize the suitable field of investment in the developing market for the future.