Task 1 Documentation

Task Overview

My analysis on supermarket sales data using Python. The key tasks involved in it are:

- Loading and preprocessing datasets related to items, sales, promotions, and supermarkets.
- Performing exploratory data analysis (EDA) using visualization techniques.
- Implementing a supervised learning model to predict sales trends or classify items.
- Extracting business insights to support decision-making.

Data Cleaning & Transformation

- Data is imported from CSV files using Pandas.
- Initial exploration is done through .head() and .info() to inspect data structure.
- Transformations include data type conversions, handling missing values, and feature engineering.
- Visualizations, such as bar charts and line plots, are used to understand data distribution.

Supervised Learning Model

- **Problem Definition:** The model aims to predict sales trends based on available item and promotion data.
- **Model Selection:** The specific machine learning model used Random Forest is yet to be determined from further analysis.
- **Feature Selection:** Features include item categories, sales records, and promotional details.
- **Training Process:** I have split the dataset into training and testing sets to evaluate performance.
- **Evaluation Metrics:** Model performance is assessed using metrics like accuracy, RMSE, or R2 score.
- **Insights:** The model's predictions have provided me with actionable insights on sales patterns and promotional effectiveness.

Business Insights

- Exploratory data analysis helped me in identifying high-performing products and the impact of promotions.

- Visualizations highlight sales trends over time, enabling better inventory management.
- These insights helped me to predict pricing strategies and promotional planning for supermarkets.