MAJOR PROJECT SYNOPSIS

ON

Supermarkets sales analysis and forecasting using Python

Master's of Computer Applications

In

Department of Computer Applications

By

Sparsh Bhardwaj

(22001602060)

Guided by

Dr. Manjeet Singh Tomar

JC Bose University of Science and Technology, YMCA

Faridabad - 121006

Synopsis: Sales Forecasting for Supermarkets using Python

Introduction:

In today's dynamic retail landscape, supermarkets face the challenge of managing inventory, optimizing resources, and meeting customer demand effectively. To address these challenges, this synopsis outlines a data-driven approach to supermarket sales forecasting using Python. By leveraging tools like NumPy, Pandas, Seaborn, Matplotlib, and ARIMA, we can extract valuable insights from historical sales data, enabling better decision-making and resource allocation.

Problem Statement:

The problem at hand is the need to accurately forecast supermarket sales to facilitate inventory management and resource allocation. This requires the analysis of historical sales data to identify trends, seasonality, and other patterns, and then applying a forecasting model that can predict future sales with precision. The key challenges include handling large datasets, managing time-series data effectively, and optimizing model parameters for accurate predictions.

Objectives

The primary objectives of this project are as follows:

Data Preparation: Collect and preprocess historical supermarket sales data, ensuring it is clean and structured for analysis.

Exploratory Data Analysis (EDA): Use Pandas and Seaborn to explore the data, identify trends, seasonality, and other patterns, and gain insights into the sales dataset.

Data Splitting: Split the data into training and testing sets for model development and evaluation.

Time-Series Forecasting: Utilize ARIMA, a powerful time-series forecasting model, to capture historical patterns and predict future supermarket sales.

Visualization: Create visualizations using Matplotlib to compare model forecasts with actual sales data, enhancing model evaluation.

Final Forecasting: Employ the tuned ARIMA model for future sales forecasting.

Reporting and Decision Making: Communicate findings and forecasts to stakeholders for data-driven business decisions, driving sales strategies and resource allocation.

Tools and Platform Required

For this project, the following tools and platforms are necessary:

Python: The primary programming language for data analysis and modeling.

NumPy: A fundamental library for numerical operations and data manipulation.

Pandas: Essential for data manipulation and analysis, especially in handling structured data.

Seaborn: A data visualization library that enhances the visual exploration of data.

Matplotlib: A versatile plotting library for creating visualizations.

ARIMA: The AutoRegressive Integrated Moving Average model for time-series forecasting, which is available through libraries like Statsmodels in Python.

Jupyter Notebook: A convenient environment for interactive data analysis and code development.

Conclusion:

This project aims to provide supermarkets with a data-driven solution for sales forecasting, facilitating efficient inventory management and resource allocation. By harnessing the power of Python and the specified libraries, businesses can make informed decisions based on historical sales data and accurate sales predictions.