Assignment: Demand Forecasting and Exploratory Data Analysis

Introduction:
Welcome to the demand forecasting and exploratory data analysis assignment! In this exercise, you will have the opportunity to demonstrate your skills in analysing and forecasting demand for a retail company's products. The dataset provided contains information about sales transactions, including the country of sale, department (brand), stock keeping unit (SKU), date of sale, and the number of items sold.

Objective:
Your task is to perform exploratory data analysis (EDA) and build a demand forecasting model. The ultimate goal is to help the company make informed decisions regarding inventory management, stock replenishment, and sales strategy. You are free to choose your approach and models for demand forecasting, but you must explain your reasons for selecting them.

Dataset Description:
The dataset you will be working with contains the following columns:

- COUNTRY: The country in which the items were sold.
- DEPARTMENT: The brand or department for which the items are being sold.
- SKU: The unique ID of a stock-keeping unit.
- DATE: The date of the sales transaction.
- NUM_ITEMS_SOLD: The number of items sold on a daily level.

Tasks:

- **Exploratory Data Analysis (EDA):**
  - Begin with an exploratory data analysis to gain insights into the dataset.
  - Visualise the distribution of NUM_ITEMS_SOLD over time (e.g., daily, monthly trends).
  - Explore any seasonality or trends in the data.
  - Analyse the impact of different departments and countries on sales.
  - Identify any missing data and propose a strategy to handle it.

- **Demand Forecasting:**
  - Build a demand forecasting model(s) to predict future sales for each SKU at a country level for daily granularity.
  - Experiment with at least 2-3 different models (e.g., time series models, regression models, machine learning models).
  - Clearly explain the rationale behind your choice of models.
  - Split the dataset into training and testing sets for model evaluation.
  - Evaluate and compare the performance of the models using appropriate metrics.

- **Performance Metrics:**
    - Since the predictions are to be made for the last quarter of 2022, choose appropriate performance metrics to assess the models.
    - State your reasons for selecting these metrics and explain why they are relevant for this forecasting task.

- **Model Selection:**
    - Select the best-performing model based on the chosen performance metrics.
    - Justify your choice and explain why this model is suitable for making predictions in the last quarter of 2022.

- **Model Deployment (Optional):**
    - If you wish, you can propose a strategy for deploying the selected demand forecasting model in a real-world scenario.

- **Conclusions and Recommendations:**
    - Summarise your findings from the EDA and demand forecasting.
    - Provide actionable recommendations for the company based on your analysis.

Submission Guidelines:

- Please provide a well-documented Jupyter Notebook or code file.
- Include clear explanations and comments throughout your code.
- Present your findings, insights, and recommendations in a clear and concise manner.
- Include visualisations where necessary to support your analysis.

Note:

- You are encouraged to use any programming language or tools you are comfortable with.
- Feel free to explore additional features or data transformations if you think they would enhance the forecasting accuracy.
- Remember to explain your thought process and reasoning behind every step of your analysis.

Good luck with your assignment! We look forward to seeing your insights and recommendations, especially in the context of predicting sales for the last quarter of 2022.