Project Instructions
1.1.1 Comprehensive Data Analytics Workflow for a Retail Business

Summary

This project integrates various lessons from the program so far into a coherent assignment designed to give learners hands-on experience with the end-to-end data analytics process. The focal point of the project is a hypothetical retail business. You will use various tools like Excel, Python, and SQL, to collect, clean, analyze, and visualize data to help the business make informed decisions..

Project Timeline:

Day 1-2: Data Analytics Overview and Data Collection

Day 2-5: Data Cleaning and Transformation in Excel

Day 5-8: Exploratory Data Analysis in Excel

Day 8-10: Automate Data Manipulation in Excel

Day 10-18: Data Analysis in Python

Day 18-24: Database Management and SQL

Day 24-27: Data-Driven Decision Making and Presentation

Objectives:

- Understand the importance and applications of data analytics in a retail business context.

- Utilize various tools like Excel, Python, and SQL for data-related tasks.
- Go through the entire data analytics pipeline from collecting data to making business decisions.

Materials Needed:

Software

- Microsoft Excel: For initial data collection, cleaning, and transformation tasks, as well as some basic exploratory data analysis.
- Python Environment: A Python environment like Jupyter Notebook or PyCharm will be necessary for more advanced data manipulation, analysis, and visualization.
    - Libraries: Pandas, Matplotlib, Seaborn, and others for data analysis and visualization.
- SQL Database Software: Software like MySQL, PostgreSQL, or SQLite to implement SQL queries for data investigation.
- Text Editor: A simple text editor like Notepad or a more advanced one like Visual Studio Code for script writing, note-taking, and coding.
- Presentation Software: Microsoft PowerPoint, Google Slides, or similar software for creating the final presentation.

Data

- sales_data.csv: Contains historical sales data.
- customer_data.csv: Contains information about customers.
- product_data.csv: Contains information about the products.

Tasks:

Part 1: Data Analytics Overview and Planning

- Define what data analytics means in the context of a retail business.
- Identify the types of decisions that could be informed by data analytics.
- Import the CSV files into Excel.
- Identify the types and quality of data.

Part 2: Data Cleaning and Transformation in Excel

- Clean the data for any inconsistencies or missing values.
- Use Excel functions to transform the data into a more usable format.

Part 3: Exploratory Data Analysis in Excel

- Use pivot tables to summarize key metrics.
- Create graphs to visualize sales trends.

Part 4: Automate Data Manipulation in Excel

- Write macros to automate repetitive tasks.

Part 5: Data Analysis in Python

- Import the cleaned Excel files into Python.
- Use libraries like Pandas for data manipulation and Matplotlib for data visualization.

Part 6: Database Management and SQL

- Import the cleaned and transformed data into an SQL database.
- Use SQL queries to extract specific insights, such as identifying the best-selling products or understanding customer behavior.

Part 7: Data-Driven Decision Making

- Analyze the data to answer business questions such as:
    - What products are underperforming?
    - What are the peak sales times?
- Present the data and insights via a presentation deck.

Deliverables:

- Excel spreadsheets with cleaned and analyzed data.
- Python code for data analysis and visualization.
- SQL queries used for data extraction and investigation.

- A final report summarizing the entire data analytics workflow, insights, and business recommendations.
- A presentation deck.

You will find the submission course upload area within the learning management system to submit your deliverable documents and files.

Evaluation Criteria:

- Quality and accuracy of data cleaning and transformation.
- Relevance and depth of exploratory data analysis.
- Efficiency and correctness of SQL queries.
- Clarity and impact of the final report and presentation.