General Instructions for Capstone

Overview

Throughout your coursework, you have acquired a diverse set of skills essential for data analysis, enabling you to explore data-driven questions and derive insights effectively. These skills encompass:

- Microsoft Excel
- SQL (Structured Query Language)
- Power BI

The primary objective of this capstone project is to provide you with an opportunity to apply and enhance these competencies, thereby challenging yourself to deepen your learning experience.

As part of your project requirements, you are expected to utilize some of the aforementioned skills, with the inclusion of Power BI being obligatory. Your task will culminate in the creation of a dashboard within Power BI, which will serve as a visual representation of your analytical journey and findings.

To guide your approach, consider the following methodology:

- Employ **Microsoft Excel** for preliminary data inspection and cleaning. This stage is critical for gaining an initial understanding of your dataset and preparing it for more sophisticated analysis.
- Utilize **SQL** for advanced data querying capabilities. This tool is particularly useful for extracting specific insights from large datasets or conducting complex manipulations that are beyond the scope of Excel.
- Integrate your analytical efforts in **Power BI** to visualize and present your findings. This step involves creating a comprehensive dashboard that not only showcases your analytical work but also tells a compelling story through data.
- In the final step, you compile your analysis into a final presentation to share your findings through storytelling.

This structured approach ensures a well-rounded application of your analytical skills, from initial data exploration to presenting your insights in a coherent and visually engaging manner.

Instructions

A collection of datasets has been curated, spanning across four distinct domains: **Financial Data**, **Sales Data**, **Entertainment Data**, and **Economic Data**. These datasets will be made available to you in CSV (Comma Separated Values) file format, offering a versatile foundation for your analysis.

You are provided with the flexibility to engage with these datasets using several tools and methodologies:

- **Excel**: You can open the CSV/Excel files directly in Microsoft Excel. This approach is particularly beneficial for initial data inspection, cleaning, and basic analysis.
- **SQL Database**: For those seeking to perform more complex data manipulation and querying, you have the option to import the CSV files into an SQL Database. An example of such a database is `SQLite`, which can facilitate advanced data analysis tasks.
- **Power BI**: Alternatively, you can connect the CSV files directly to Power BI for the purpose of creating dynamic visualizations and comprehensive dashboards. This tool is especially useful for synthesizing insights from the data and presenting them in an accessible format.

Note

To ensure all participants have a clear understanding of how to effectively utilize these tools with the provided datasets, the instructor will conduct a detailed walkthrough. This session is designed to demonstrate the process of opening CSV files in Excel, loading them into an SQL Database, and connecting them to Power BI, thereby equipping you with the necessary skills to commence your analysis.

Expectations

For the successful completion of this capstone project, participants are expected to deliver the following components, which are critical in demonstrating the application of your analytical skills and the ability to convey your findings effectively:

1. **Power BI Dashboard**: Create a Power BI dashboard that encapsulates your analysis and narrative. The dashboard should contain at least one screen, thoughtfully designed to support your story with visual data representations. It is imperative that this dashboard not only showcases your analytical prowess but also serves as a pivotal tool in narrating your data-driven story.
2. **PowerPoint Presentation**: Develop a PowerPoint presentation to present your findings and tell your story. This presentation should be structured in a manner that guides the audience through your analytical journey, from the initial questions and objectives to the insights and conclusions derived from your data. It should be visually engaging and articulate, with a clear focus on conveying the significance of your findings in a coherent narrative form.
3. **Submission of Work**: All components of your work, including (if any) Excel files , SQL queries, and Power BI files , must be submitted to your instructor for review.

Data Sources

If you are looking for additional data source you can explore the following websites:

- Our World in Data
- U.S. Government's Open Data
- Kaggle Datasets
- National Center for Health Statistics
- World Bank Open Data