

## Assignment Solution

### Question 1.

Answer: Given the scale of the business and the sales volume, it's crucial to select a tool that can handle large datasets efficiently, provide deep insights through interactive visualizations, and accommodate the iterative addition of new data. Let's evaluate Power BI and Streamlit based on these requirements.

#### Power BI

##### Pros:

- **Scalability:** Power BI can handle large datasets efficiently, especially when using Power BI Pro or Premium for enhanced data refresh rates and more storage.
- **Integration:** Seamless integration with other Microsoft products, such as Excel and Azure, which could be beneficial if your company already uses these tools.
- **Visualization:** Offers a wide range of visualization options that are easy to use for creating comprehensive dashboards.
- **Collaboration:** Power BI reports can be easily shared and published within the organization, including scheduled automatic updates.

##### Cons:

- **Cost:** While there is a free version, the Pro and Premium versions, which offer necessary capabilities for large-scale businesses, come with a subscription cost.
- **Complexity:** Some of the more advanced features have a steep learning curve.

#### Streamlit

##### Pros:

- **Flexibility:** Being Python-based, it offers vast flexibility in terms of data manipulation, analysis, and the ability to use the entire Python ecosystem.
- **Development Speed:** Quick to develop and deploy data applications, especially for data scientists and analysts familiar with Python.
- **Cost-Effective:** Streamlit itself is free, and you can host applications on various platforms, depending on your budget and requirements.

##### Cons:

- **Scalability:** While Streamlit is excellent for quick development, handling very large datasets efficiently may require additional engineering effort, such as optimizing the app or managing the server infrastructure.
- **Visualization and UI Polish:** While Streamlit supports many charting libraries, the out-of-the-box visualization options are not as extensive or as polished as those in Power BI.

#### Recommendation

Given the large scale of your business and the emphasis on weekly sales reporting across various stores and departments, **Power BI** would likely be the more suitable choice. Its robust data handling capabilities, integration with existing business systems, and extensive visualization options make it a powerful tool for creating comprehensive dashboards. Power BI's ability to manage large datasets efficiently and its native integration with Microsoft's ecosystem (potentially including your POS system) can provide the scalability and reliability needed for your business.

Moreover, the collaborative and sharing features of Power BI align well with the need to regularly update and distribute reports to the CEO and other stakeholders. While Streamlit offers great flexibility and development speed, the requirements for scalability, data integration, and polished visualization lean towards Power BI as the better tool for this specific scenario.

However, it's important to consider the subscription cost of Power BI Pro or Premium as part of your decision. If cost is a significant concern, starting with Power BI's free version to validate the approach before scaling up could be a prudent strategy.

Question 10.

Answer: When updating the dashboard with the new week's data, covering the period starting 26th October 2012, the process involves a few steps that directly impact the overall time needed for the refresh. This includes moving the new data from the "New week" folder to the "All week" folder, ensuring it's correctly formatted and aligned with the existing dataset structure, and then performing the actual data refresh in Power BI. The entire process, depending on the dataset size and complexity, could take anywhere from a few minutes to potentially longer if manual adjustments or data transformations are required. In this case it was only a few minutes

#### **Recommendations for Organizational Data Management:**

To maintain and even improve the efficiency of our weekly updates, I propose a few strategic recommendations for managing our organizational data:

1. **Automated Data Collection:** Implementing a more automated data collection process from our POS systems could significantly reduce the time needed to prepare and load the data each week. Automation ensures data consistency and accuracy, reducing manual errors.
2. **Centralized Data Storage:** Establishing a centralized data storage solution, such as a cloud-based data warehouse, would allow for more streamlined data management. This centralization supports better data governance, security, and accessibility across the organization.
3. **Scheduled Refreshes:** Utilizing Power BI's scheduled refresh feature can automate the dashboard update process. By setting up specific times for the data to be refreshed automatically, we ensure that the dashboard always displays the most current data without manual intervention.
4. **Data Quality Checks:** Implementing regular data quality checks can help identify and correct any issues before they impact the dashboard. This proactive approach keeps our data reliable and trustworthy.

**Making the Case:**

By adopting these recommendations, we're not just streamlining our weekly dashboard updates. We're setting a foundation for a more data-driven culture within our organization. Automated data collection and centralized storage reduce manual workload and improve data accuracy. Scheduled refreshes ensure our insights are always current, empowering us to make swift, informed decisions. Finally, regular data quality checks maintain the integrity of our insights, building trust and confidence in our data-driven strategies. These steps will further solidify our competitive edge, enabling us to respond more agilely to market changes and opportunities.