**Assignment: Week 7 – Data Visualization & Scientific Libraries**

**Theoretical Questions (Answer in 3–5 lines each)**

1. **Explain the difference between Seaborn and Matplotlib. When should you use Seaborn over Matplotlib?**

* Matplotlib is a low-level, powerful library for creating static, animated, and interactive visualizations in Python. Seaborn is built on top of Matplotlib and provides a high-level interface for drawing attractive and informative statistical graphics. Seaborn comes with built-in themes, color palettes, and complex plot types like heatmaps and violin plots.
* Use Seaborn when you want quick, clean, and beautiful statistical plots with less code. For more control and customization, especially for complex layouts or non-statistical plots, use Matplotlib.

1. **What is the purpose of a box plot? How does it help in detecting outliers and comparing distributions?**

* A box plot (or box-and-whisker plot) is used to display the distribution, central value, and variability of a dataset. It shows the median, quartiles, and potential outliers in the data.
* Outliers appear as individual points outside the “whiskers” (typically 1.5×IQR from the quartiles). By plotting multiple box plots side by side, you can easily compare distributions, detect skewness, and identify differences in spread or central tendency between groups.

1. **How do heatmaps help in understanding relationships between variables in a dataset? Give one real-life example.**

* Heatmaps use color gradients to represent the magnitude of values in a matrix format, making it easy to spot patterns, correlations, or anomalies between variables. They are especially useful for visualizing correlation matrices or categorical data frequencies.
* In a sales company, a heatmap of monthly sales across different regions can help identify peak sales periods and underperforming areas, enabling better strategic planning.

1. **What are color palettes in Seaborn, and why is customizing them important in data storytelling?**

* Color palettes in Seaborn are predefined sets of colors used to style plots consistently and aesthetically. Examples include deep, pastel, muted, and dark.
* **Importance in Data Storytelling:**  
  Customizing color palettes helps highlight key insights, maintain visual clarity, and align with branding or audience needs. Thoughtful color use improves readability, emphasizes comparisons, and makes the story behind the data more engaging and memorable.

1. **What is the role of SciPy in data science? Mention at least two use cases where SciPy is preferred over NumPy.**

* SciPy is a scientific computing library that builds on NumPy, providing additional functionality for advanced mathematical, statistical, and scientific computations.

**Use Cases Where SciPy is Preferred Over NumPy:**

* **Optimization:** For solving linear programming, curve fitting, or minimization problems using scipy.optimize.
* **Signal or Image Processing:** For filtering, transformation, or analysis using scipy.signal or scipy.ndimage, which go beyond NumPy's basic capabilities.