5-Day Practical Plan for Advanced Data Visualization

This document provides a detailed 5-day plan for practical sessions on data visualization, tailored for students of Master's in Computational Engineering Data Analytics.
Each session includes tasks that incorporate datasets, visualization techniques, and analysis.
Each session is designed to last approximately 3 hours.

Environment

Option 1. Python Interpreter
Option 2. Jupyter Servers [python -m notebook]

Day 1: Core Skills for Visualizing Algorithms and Complex Systems

Objective: Understand and demonstrate the core skills required to visualize algorithms, datasets, information, complex systems, and processes.

Datasets:
- None required (focus is on algorithm visualization).

Tasks:
1. **Visualize Sorting Algorithms (1 hour):**
   - Implement **Bubble Sort** and **visualize each step** using **Matplotlib**.
   - Create an animated visualization showing the array being sorted in real-time.

2. **Visualize Search Algorithms (1 hour):**
   - Implement and visualize a **Binary Search Tree traversal** (in-order, pre-order, post-order).
   - Create a graphical representation of the tree using **NetworkX**.

3. **Practical Assignment (1 hour):**
   - Choose another algorithm (e.g., Dijkstra's shortest path) and create a step-by-step visualization using Matplotlib or NetworkX.

 - For reference:
https://matplotlib.org/stable/gallery/animation/

https://interactivetextbooks.tudelft.nl/open-textbooks-demonstration/content/Basic_animation_demo.html

Day 2: Systematic Visualization of Procedures and Emergent Systems

Objective: Explore procedures, data, and emergent systems as the subject of visualizations in a systematic way.

Datasets:
- Emergent Systems: Simulation-based data (e.g., flocking behavior or Conway's Game of Life).

Tasks:
1. **Simulate and Visualize Emergent Behavior (1.5 hours):**
   - Implement **Conway's Game of Life** and visualize its evolution using Matplotlib.
   - Use animation to represent the states of the cells over time.

2. **Analyze and Visualize Cellular Automata (1.5 hours):**
   - Implement a cellular automaton based on a rule set (e.g., Rule 30).
   - Visualize the state transitions over time using **Seaborn** or **Matplotlib**.

- Reference: [Conway's Game of Life in Python](https://www.geeksforgeeks.org/conways-game-life-python-implementation/).

Day 3: Selecting and Applying Visualization Techniques

Objective: Assess visualization needs and select and apply an appropriate range of techniques to datasets relevant to a specific domain.

Datasets:
- [Flights Dataset](https://raw.githubusercontent.com/mwaskom/seaborn-data/master/flights.csv).

Tasks:
1. **Visualization Assessment and Application (1 hour):**
   - **Analyze** the dataset and determine which visualization techniques are most appropriate (e.g., heatmaps for month-to-month passenger trends, line charts for yearly trends).
   - Use **Seaborn** and **Matplotlib** to create these visualizations.

2. **Interactive Visualization (1 hour):**
   - Create interactive visualizations (e.g., scatter plots and line graphs) using **Plotly**.
   - Highlight key trends and annotate important points.

3. **Practical Assignment (1 hour):**
   - Design a dashboard using Plotly Dash to visualize flight trends interactively.

Day 4: Applying Data Analytics Techniques for Visualization

Objective: Apply data analytics techniques for visualization implementation and interpret the visualization results.

Datasets:
- [Iris Dataset](https://archive.ics.uci.edu/ml/machine-learning-databases/iris/iris.data).

Tasks:
1. **Data Preprocessing and Analysis (1 hour):**
   - Perform preprocessing on the Iris dataset (e.g., scaling, encoding).
   - Use **clustering** (e.g., K-Means) to group data and **visualize** clusters using scatter plots.

2. **Correlation and Feature Relationships (1 hour):**
   - Generate a **heatmap** of correlations between the features using **Seaborn**.
   - Use **pair plots** to visualize relationships between features grouped by species.

3. **Practical Assignment (1 hour):**
   - Perform **PCA** (Principal Component Analysis) and create a 2D visualization of the reduced dimensions, annotated with cluster information.

Day 5: Evaluating Visualizations for Usability and Human Factors

Objective: Evaluate visualizations in terms of different users and tasks with human factors and critically appraise visualization approaches.

Datasets:
- [Sentiment Analysis Dataset](https://github.com/vineetdhanawat/twitter-sentiment-analysis/tree/master/datasets).

Tasks:
1. **User-Centric Design (1 hour):**
   - Create visualizations tailored to specific personas (e.g., analysts vs. executives).
   - Design two different dashboards for the sentiment dataset, focusing on user goals.

2. **Evaluation and Critique (1 hour):**
   - Evaluate the usability of existing visualizations (provided by the instructor).
   - Write a critique on strengths, weaknesses, and accessibility.

3. **Practical Assignment (1 hour):**
   - Redesign a visualization to improve usability and align with human factors principles.