**CSE-564 - Visualization**

**Lab 2b – Assignment: Advanced Data Visualization with MDS and PCP on Kidney Disease Dataset**

**MDS and PCP Visualization Dashboard**

**Introduction**

This lab's project extends the visualization capabilities introduced in Lab 2a by incorporating Multidimensional Scaling (MDS) and Parallel Coordinates Plot (PCP) analyses. The objective is to further unravel the complex high-dimensional nature of the chronic kidney disease (CKD) dataset and to provide interactive visual explorations that could aid medical diagnostics.

**Source of the Dataset**

The dataset used in this project is obtained from Kaggle, which is a platform for predictive modeling and analytics competitions. The specific dataset for chronic kidney disease can be accessed through the following link:[**CK Disease Dataset**](https://www.kaggle.com/datasets/mansoordaku/ckdisease) on Kaggle. It comprises various medical attributes that are essential in diagnosing CKD in patients.

**Attributes Description**

The dataset contains multiple attributes, including both numerical and categorical data, such as age, blood pressure, specific gravity, albumin, sugar, red blood cells, and more. Each attribute provides critical information regarding the patient's health status, contributing to the diagnosis of CKD.

**MDS Plots**

The MDS plots translate the high-dimensional data into two-dimensional scatterplots using both Euclidean distances for the data MDS plot and (1-|correlation|) distances for the variables' MDS plot.

a) The data MDS plot colors points by cluster ID, reflecting the clusters identified through K-Means, offering an insightful visual representation of the dataset's inherent groupings.

b) The variables' MDS plot uses correlations to capture relationships between variables, providing an informative scatterplot that showcases the data's structure in terms of its variables.

**PCP Plot**

The PCP plot presents all data dimensions, both categorical and numerical, allowing users to interactively determine a meaningful axis ordering. The coloring of polylines by cluster ID establishes a clear connection between this plot and the patterns discovered in the MDS plots.

**Insights and Interactivity**

Lab 2b emphasizes user interaction, especially in the context of PCP axis ordering. Users click points in the variables' MDS plot to sequence axes in the PCP plot, an innovative approach that utilizes correlations to streamline the analysis process.

**Implementation Highlights**

Noteworthy Implementation Aspects

In Lab 2b, the dashboard's architecture and design have undergone significant refinement to enable more intricate analysis and user engagement:

* **Sophisticated Data Mapping:** The MDS plots, a central feature of this lab, employ sophisticated data mapping techniques to distill high-dimensional data into two dimensions, preserving as much of the data's variability as possible.
* **Enhanced Interactivity:** The interactive components, such as cluster selection and variable ordering in the PCP, are more than mere visual elements—they serve as intuitive controls for users to steer the analysis, reflecting an integration of visualization and data mining concepts.
* **Context-Aware Visual Elements:** The dashboard elements react to each other contextually. For instance, selecting a cluster in the K-Means MSE plot immediately updates the MDS plots to highlight the chosen cluster, fostering a cohesive analytical narrative.
* **Automatic Elbow Point Detection:** The K-Means MSE plot is equipped with an automatic elbow point detection feature, providing users with an objective method for selecting the number of clusters, which is crucial for subsequent MDS and PCP visualizations.
* **Responsive and Accessible Design:** The visualizations are engineered to be fully responsive, maintaining clarity and interaction quality across devices, thereby democratizing access to data analysis.
* **Algorithm-Driven User Interface:** The dashboard's user interface elegantly balances algorithmic complexity with user simplicity, allowing users to interact with advanced statistical models without needing deep technical knowledge.
* **Technical Integration:** The dashboard's back-end integration has been enhanced to process data more efficiently, translating into faster response times for user interactions and smoother transitions between visual states.

**Conclusion**

The enhancements introduced in Lab 2b represent a significant leap forward in the field of interactive data visualization for healthcare datasets. The dashboard now not only serves as an analytical tool but also acts as an educational platform where users can learn about data relationships through direct manipulation and exploration.

The addition of MDS and PCP plots has enriched the dashboard with advanced visualization techniques, allowing for more nuanced interpretations of the CKD dataset. These methods help uncover patterns that might be less apparent in traditional plots, offering potential for new insights in medical diagnostics. The improved interactivity empowers users to tailor the analysis process to their specific queries or hypotheses, making the dashboard not just a visualization tool but an interactive canvas for discovery.

Lab 2b's visualization dashboard exemplifies the synergy between data visualization, machine learning, and user interface design. It stands as a sophisticated, yet accessible, analysis tool—one that can adapt to different users' needs, from researchers and clinicians looking for detailed data insights to students learning the ropes of data science.

In summary, this lab's achievements underscore the transformative power of combining data science with user-centric design—a combination that can bring complex data to life and make sophisticated analytics accessible to a broader audience.

**Running the Code**

To initiate the dashboard, direct to the project folder and execute the command python app.py in the terminal. The Flask server will activate, making the dashboard available at http://localhost:5000. It is crucial that all dependencies are installed and that the dataset files are located in the designated directory.

**Dashboard UI Snapshots**

The enhanced dashboard showcases integrated MDS and PCP visualizations, presenting an evolved user interface for complex data analysis.

A close-up of a graph

Description automatically generatedA screenshot of a graph

Description automatically generated

A graph of different colored lines

Description automatically generated