Project 8 – Visual Sommelier (Wine project)

Github website: <https://jessyyy0299.github.io/DS4200_Project/>

Github repo: <https://github.com/Jessyyy0299/DS4200_Project>

Visualization 1- pair analysis

The visualization uses a hexbin heatmap with rectangular bins arranged in a 30x30 grid pattern, where each bin represents a specific range of x and y values in the data space. The primary visual channels include position and color saturation. Position on both x and y axes encodes the values of two selected wine properties, while color saturation uses a sequential blue scale to encode the density of observations, with darker blues indicating higher concentration of data points and lighter blues or white showing fewer observations. The visualization maintains consistent visual representation through fixed-size bins. Interactive elements enhance the exploration through dropdown menus for selecting x and y variables, accompanied by dynamic updates to axis labels and plot titles.

Visualization 2 - cost-reduction analysis

The visualization utilizes density plot structures with smooth curves representing the distribution of a selected chemical property across various wine quality levels. The visual encoding relies on several key channels: position, shape, and color. The x-axis position represents the chemical property values, while the y-axis position encodes the density, indicating the relative frequency of these values. The shape of each curve illustrates how the property is distributed within each quality level. The color scheme employs a gradient palette, transitioning from purple for lower qualities to yellow for higher qualities, creating a clear distinction between quality levels. The visualization's interactivity is enhanced through a dropdown menu for selecting different chemical properties, allowing for a focused analysis of specific attributes.

Visualization 3 – regression analysis

The marks are the bars that represent the regression line slopes for each wine category. The channels used include the height of the bars (representing the magnitude of the slope). The labels along the x-axis (representing different wine categories). The tasks for this visualization are to enable users to quickly compare the wine quality across different combinations of alcohol, pH, and volatile acidity. By analyzing the heights of the bars, users can determine which wine categories have stronger or weaker relationships between different combinations and quality, providing insights into how different combinations influence wine quality.

Visualization 4 – feature analysis

The marks in this visualization are the stacked bars, which represent different categories of feature levels (e.g. 'Low, Medium, High') for each type of wine. The channels used include color, height, and position. Colors are used to distinguish between red and white wines, while the height of each segment within the bars conveys the proportion of each level (e.g., low, medium, high). The position along the x-axis shows the feature being compared, allowing users to easily discern differences in feature values between wine types. The tasks of the visualization are to enable users to compare the distribution of feature levels between red and white wines, identify which features are more prominent in higher-quality wines, and explore the overall impact of these features on wine quality. By interacting with the dropdown, slider, and toggle button, users can effectively filter and compare different aspects of the data, gaining a deeper understanding of the factors influencing wine quality.

Visualization 5 – sustainability analysis

The visualization facilitates comparative analysis of the trade-offs between achieving high wine quality and minimizing the use of non-environmentally friendly compounds such as alcohol and sulfates. The primary objective is to uncover actionable insights that balance sustainability with wine quality. To achieve this, the chart employs circles, where each represents an individual wine sample. The position of the circles is defined by the x-axis, representing a primary attribute such as alcohol or sulfates, and the y-axis, representing additional properties such as acidity or pH. This layout highlights the relationships between these attributes and wine quality. Color encodes wine quality, with hues for qualities 7 (pink), 8 (orange), and 9 (purple), enabling easy differentiation. Size emphasizes higher-quality wines, with larger circles making these samples more prominent. Additionally, brush and link function is added to make comparison easier, allowing user to select a particular region they are interested.

Visualization 6 – covariance analysis

This visualization uses points as marks, where each point represents an individual wine sample, encoding its relationship between a feature ratio and quality. Points are mapped using position on the x-axis, representing feature ratios, and on the y-axis, representing quality levels. This arrangement visualizes variability across different ratio ranges. Color differentiates between red and white wines, facilitating comparison. Density or clustering of points indicates consistency, with densely populated regions reflecting reliable quality at specific ratios and sparse areas highlighting variability. This design effectively illustrates the greater variability in red wines, particularly at low-to-medium ranges, and the relative predictability of white wines, particularly at medium-to-high ratio ranges. These insights can guide producers in refining chemical management strategies to achieve consistent quality in both wine types.

Visualization 7 – heatmap

The visualizations use heat maps as the primary mark type, where each cell represents the average quality score for a specific combination of volatile acidity and residual sugar levels. The two categorical axes determine the position of the cells: volatile acidity groups (y-axis) and residual sugar groups (x-axis). The color channel encodes the quality scores, with darker shades representing higher quality and lighter shades representing lower quality. Annotations within the cells provide additional clarity by displaying precise quality values. The task of the visualizations is to enable comparison and identification of trends between different ranges of volatile acidity and residual sugar. By visualizing the quality scores for each group, the heat maps allow producers to determine the optimal combination of these factors for achieving higher-quality wines. The inclusion of distinct color gradients for red and white wines further aids in differentiating their quality patterns.