DS4200 Final Project - Design Choices
Gianni Diarbi, Sai Vishnu Chitra, Wali Qureshi

Link to our website:
https://saivishnuchitra.github.io/DS4200-Website

Visualization 1
This scatterplot with overlaid regression lines was chosen to examine how brain volume changes across age and whether this trajectory differs depending on dementia status. Age and nWBV are continuous variables, making scatterplots the most direct way to observe their relationship. Plotting both groups on the same axes facilitates a side-by-side comparison of decline patterns. The regression lines summarize the slope for each group, helping highlight differences in the rate of volume loss over aging.

Visualization 2
A split-panel scatterplot with regression lines was chosen because it allows a direct visual comparison between individuals with and without dementia while holding the visual structure constant. Scatterplots are ideal for showing relationships between two continuous variables—in this case, normalized whole brain volume (nWBV) and Mini-Mental State Exam (MMSE) scores. By separating the groups into two panels, the design prevents overplotting and makes the contrast in patterns more interpretable.

Visualization 3
This chart shows how age is spread out for people with and without dementia in the OASIS dataset. To make it, the ages were grouped into bins, and we counted how many people fell into each age range. We also created a simple label using the CDR score so we could split the sample into “No Dementia” (orange) and “Dementia” (blue).

Visualization 4
This chart compares brain volume between people with dementia and people without dementia. To make it, we used normalized whole brain volume (nWBV) and grouped participants based on whether their CDR score showed dementia or not. We plotted two boxplots (blue for dementia, orange for no dementia) and added light dots to show individual data points.

Visualization 5
To give a thorough statistical overview of variations in brain volume, a composite chart comprising a boxplot, jittered points, and mean markers was chosen. Standard boxplots can conceal sample size and density even when they show quartiles and medians. We display the actual distribution of the data by superimposing jittered points, revealing the

positions of individual individuals. The addition of a diamond-shaped mean marker highlights the clear decline in brain volume linked to dementia and enables a direct comparison of the arithmetic average between the groups.

Visualization 6
A faceted correlation heatmap was selected to simultaneously display the strength and direction of linear relationships between multiple continuous variables across different participant subgroups. Using "small multiples" allows for quick comparison of structural relationships without cluttering a single view. A diverging color scheme—Orange for positive correlations and Blue for negative correlations—was implemented to make strongly opposing relationships instantly distinguishable.