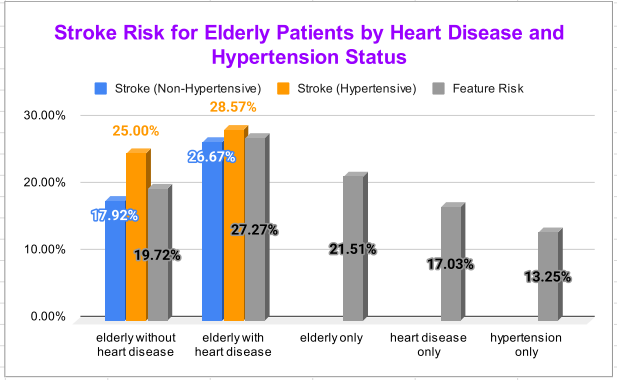
Multivariate Analysis - 3 Features vs Stroke Examples

Stroke Risk for Elderly Patients by Heart Disease and Hypertension Status

Feature Combination	Stroke (Non-Hypertensive)	Stroke (Hypertensive)	Feature Risk
elderly without heart disease	17.92%	25.00%	19.72%
elderly with heart disease	26.67%	28.57%	27.27%
elderly only			21.51%
heart disease only			17.03%
hypertension only			13.25%



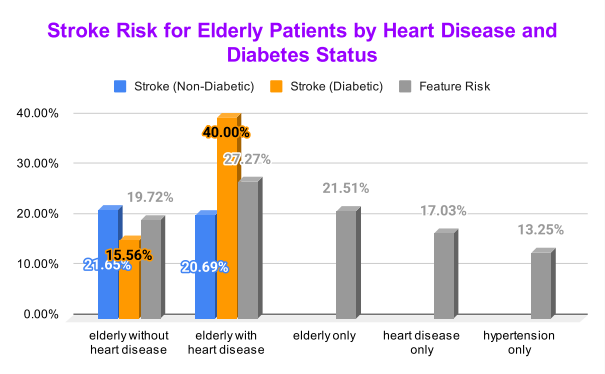
Insight and Comment

This visualization compares stroke risk for elderly patients, segmented by heart disease and hypertension status. It highlights how stroke risk increases as more risk factors are present, peaking at 28.57% for elderly patients with both heart disease and hypertension.

While this example explores only 4 of the 20 possible category combinations for age, heart disease, and hypertension, it illustrates how multivariate patterns quickly become complex. A full exploration of 3-feature combinations would be better handled through statistical modeling or machine learning to account for all interactions and subgroup variations.

Stroke Risk for Elderly Patients by Heart Disease and Diabetes Status

Feature Combination	Stroke (Non-Diabetic)	Stroke (Diabetic)	Feature Risk
elderly without heart disease	21.65%	15.56%	19.72%
elderly with heart disease	20.69%	40.00%	27.27%
elderly only			21.51%
heart disease only			17.03%
hypertension only			13.25%



Insight and Comment

Among elderly patients, those who have both heart disease and diabetes face a 40.00% risk of stroke — the highest observed stroke rate across all subgroup combinations in this analysis. This is nearly 8 times the overall population stroke rate of 4.87%, and around double the risk of the compounding risk factors in this comparison.

Interestingly, stroke risk does not increase uniformly for all combinations. For example, elderly patients without heart disease show a slightly lower stroke risk when diabetic (15.56%) compared to non-diabetic (21.65%), which may reflect smaller sample sizes or other confounding variables.

This example highlights how compounding risk factors can lead to extremely elevated stroke risk, but also reinforces the challenge of manually analyzing higher-order combinations. This 3-feature comparison covers only 4 of 20 possible combinations between age, heart disease, and a binary glucose level (diabetic vs non-diabetic).

Exploring the full set of glucose categories would expand the analysis to 40 combinations, further increasing complexity and reinforcing the need for statistical modeling or machine learning to handle the full scope of interactions.

To fully explore these relationships and control for interactions between multiple variables, statistical modeling or machine learning techniques would be required.