Notes

Simpson Paradox

* It occurs when a trend observed in several groups reverses or disappears when the data is combined.
* This happens due to the influence of hidden variable affecting the analysis.
* In simpler words,
  + A trend exists within individual groups,
  + But flips or vanishes when we look at the overall data.
* It is caused by confounding variables that create misleading conclusions when data is aggregated.
* Confounders are hidden factors influencing the relationship between two variables.
* How to detect & solve it:
  + Always check group-level data before drawing conclusions from aggregated data.
  + Use statistical tests to control for confounding variables.
  + Analyze trends within each group separately before making decisions.

Problem Statement

* Analyzed Tesla’s stock returns over two years, separating them into Bull and Bear market conditions to see if performance trends where consistent when aggregated.
* Key Findings:
  + Stock performed better in both Bull & Bear markets in Year 2 compared to Year 1.
  + However, when aggregated, Year 1 appeared to have lower overall returns than Year 2.
  + This misleading trend happens because more quarters were spent in bear markets, causing Simpson’s Paradox.
* Why?
  + Investors might wrongly assume that TSLA underperformed based on aggregated yearly data.
  + The correct approach is to analyze stock performance separately in different market conditions.