Detection of Data Manipulation

Can We Trust "Normal-Looking" Data?
When Statistical Significance Hides the Truth

r = 0.5, p <0.001

Project Focus
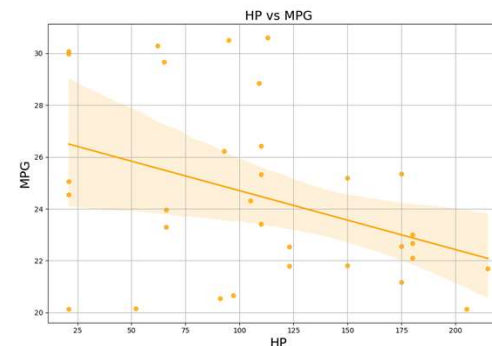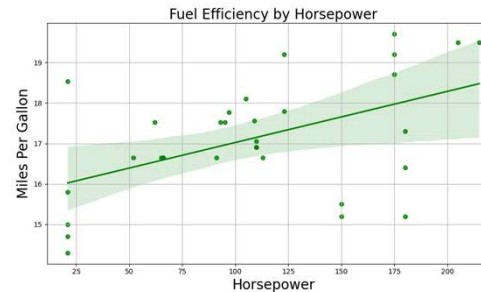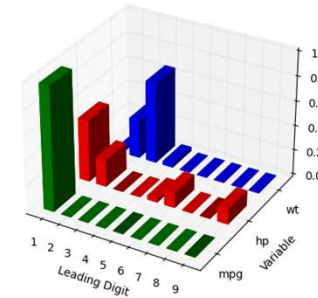Can seemingly valid data still be misleading?

Methods
- Benford's Law (statistical analysis)
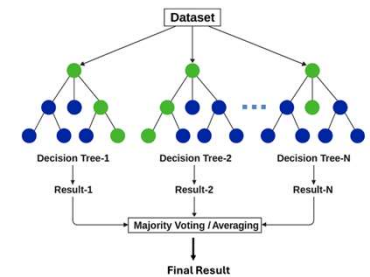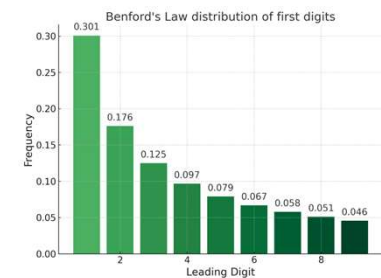- Random Forest (machine learning)

Tasks
- Classify manipulated vs. non-manipulated data
- Reconstruct original patterns via regression

Key Insight
Data manipulation can reverse the true correlation
—despite statistical significance.

Benford's Law
(Statistical Analysis)

Random Forest
(Machine Learning)