**Plan: (applying the PCIP system)**

Before beginning the analysis, we outlined the following research questions:

* Does the number of Walmart stores in each U.S. state correlate with state-level GDP?
* What patterns emerge in GDP per store across different regions?
* Are there outliers that challenge the assumption that higher GDP equals more Walmart stores?

We identified two datasets:

1. Walmart’s official open dataset of store locations.
2. State-level GDP data from the U.S. Bureau of Economic Analysis.

We planned to clean, join, and visualize these datasets using R and Quarto, with all steps documented in a reproducible file.

**Code**

Code was written in modular sections with clear comments. Each block focuses on a single task: importing data, cleaning it, merging datasets, generating plots, or calculating correlation statistics.

All code was structured with:

* Clear chunk labels for traceability.
* Reusability in mind (avoiding hardcoded paths and maintaining reproducible formats).

**Interpret**

We interpreted output using statistical reasoning (e.g., a Pearson correlation coefficient of 0.75) and visualization techniques (e.g., dual-axis bar chart, scatter plot with labeled outliers). Outliers such as New York and California were noted and explained in terms of regional and strategic retail differences.

Interpretation was directly embedded below each visualization, using EPT techniques such as comparison, emphasis, and spatial encoding to guide understanding.

**Present**

The final analysis was presented as a structured Quarto report with:

* Clean sectioning (Intro, Data, FAIR/CARE, Visuals, Conclusion, Appendix)
* Hidden code in the main body and full visibility in the appendix
* All visual outputs supported by text-based interpretation and external context