**Statistical Analysis Plan**

**Objective:**

1. Exploring FDA Recall Trends by State: Develop a Tableau dashboard showcasing states with the highest FDA recall occurrences

2. Categorizing Major Recall Types: Present a visualization highlighting primary categories subject to FDA recalls

3. Understanding Recall Reasons: Create a dashboard segmenting key reasons behind FDA recalls

**Data Source:** FDA Open Website

**Data Pre-processing:**

1. Text Data Cleansing: Thoroughly pre-process text data to ensure accurate NLP analysis

2. Stop Word Removal: Eliminate irrelevant words for meaningful insights

3. Tokenization: Break down text data into tokens for analysis

4. Vectorization with Term Frequency-Inverse Document Frequency

5. Clustering and Labelling: Group similar recall reasons using clustering, with appropriate labels

**Analysis Population:**

1. Full Dataset Inclusion: Analyse entire FDA recall dataset for comprehensive insights

2. Exclusion Criteria: Exclude incomplete or irrelevant data points

**Additional Considerations:**

**- Interactive Visualization:** Develop user-friendly interactive dashboard

**- Visual Storytelling:** Create a coherent narrative for user guidance