# PACE Strategy Document

**1. Plan Stage**

**Data Understanding**:

* **Relevant Variables**: order\_id, customer\_age, purchase\_amount, product\_category, region
* **Units**: purchase\_amount in USD, customer\_age in years.
* **Data Format**: Mixed formats (dates as strings; need standardization).

**Data Quality**:

* **Missing Data**: 15% missing values in customer\_age; will impute with median.
* **Initial Presumptions**: Higher sales in Q4 due to holidays; will verify with EDA.

**EDA Prep**:

* Practices: Summary statistics, histograms, correlation analysis.

**2. Analyze Stage**

**EDA Approach**:

* **Steps**: Clean data → Calculate summary stats → Plot distributions.
* **Joining Data**: Merge with customer\_demographics table for richer insights.
* **Structuring**: Filter out test orders (purchase\_amount < $0).

**Audience Considerations**:

* **Executives**: High-level trend dashboards (sales by region).
* **Marketing**: Demographic breakdowns (age vs. product category).

**3. Construct Stage**

**Outputs**:

* **Visualizations**: Bar chart, box plot, histogram.
* **Key Variables**: product\_category (for segmentation), purchase\_amount (for KPIs).

**Processes**:

* Handle missing customer\_age with median imputation (from Plan stage).

**4. Execute Stage**

**Key Insights**:

* **Finding**: 45% of revenue comes from 18–35 age group; Electronics is top category.
* **New Question**: Why do Midwest sales underperform? (Further research needed).

**Recommendations**:

* **Business**: Target 18–35 age group with Electronics promotions.
* **Sharing**:
  + **Executives**: Slide deck with key metrics.
  + **Marketing Team**: Interactive Tableau dashboard.