**Project Approach Overview**

**1. SQL: Data Extraction, Cleaning & Transformation**

**Goal:** Prepare raw marketing, engagement, and customer review data for analysis.

* **Product Data:**
  + Removed redundant category column.
  + Created a price\_category column using a CASE statement (low, medium, high).
* **Customer & Geography Data:**
  + Combined customer and geography tables with a LEFT JOIN to create a single dimension (dim\_customer) with demographics and location.
* **Customer Reviews:**
  + Cleaned review text (replaced double spaces with single spaces).
* **Engagement Data:**
  + Standardized content type spelling/casing.
  + Split combined views/clicks into separate columns.
  + Converted engagement date to date format.
  + Filtered out newsletter rows.
* **Customer Journey:**
  + Filled missing durations with the average for each visit date (window function).
  + Removed duplicates using ROW\_NUMBER() and retained only the first occurrence per journey.

**Outcome:**  
All tables are cleaned, structured, and ready for enrichment and modeling in Power BI.56

**2. Python: Sentiment Analysis & Data Enrichment**

**Goal:** Enrich customer reviews with sentiment insights for deeper analysis.

* **Data Loading:**
  + Fetched cleaned review data from SQL into a pandas DataFrame.
* **Sentiment Analysis:**
  + Used NLTK’s VADER SentimentIntensityAnalyzer to calculate a sentiment score for each review text (range: -1 to +1).
* **Categorization:**
  + Created SentimentCategory: Combines sentiment score and numeric rating to capture nuances (e.g., Positive, Mixed Negative, etc.).
  + Created SentimentBucket: Buckets sentiment score into ranges (e.g., "strongly positive", "mildly negative").
* **Output:**
  + Saved the enriched DataFrame (with original data + sentiment columns) as a CSV for use in Power BI.

**Outcome:**  
A sentiment-enriched reviews dataset ready for integration and visualization.78

**3. Power BI: Data Modeling, DAX, and Visualization**

**Goal:** Build an interactive, actionable analytics dashboard suite.

* **Data Model:**
  + Imported all cleaned and enriched tables:
    - Dimension tables: dim\_product, dim\_customer, calendar
    - Fact tables: fact\_customer\_journey, fact\_engagement\_data, fact\_customer\_reviews\_with\_sentiment
  + Created relationships between dimensions and facts on IDs and dates.
* **DAX Measures:**
  + Developed KPIs and calculations for views, clicks, likes, campaigns, journeys, reviews, average rating, conversion rate, engagement rate, cart abandonment, micro-conversions, sentiment mix, net sentiment score, and more.
* **Dashboards:**
  + **Overview:** KPIs, engagement/conversion trends, product benchmarks, funnel, sentiment by product.
  + **Conversion Rate:** Conversion KPIs, journey funnel, YoY trends, cart abandonment by product/country/month.
  + **Social Media Details:** Content type performance, campaign metrics, country sentiment, monthly trends.
  + **Customer Review Details:** Sentiment KPIs, key influencers, sentiment trends, **Python-powered word cloud**, negative review drivers, top issues.
    - *Note: The word cloud visual was created using Power BI’s Python visual, allowing for advanced customization and direct use of sentiment-enriched data.*

**Outcome:**  
A suite of interactive dashboards delivering actionable insights for marketing, product, and customer experience teams, with advanced text analytics powered by Python visuals