

Roadmap for Costco Wholesale – MSDS 25.4: Voice of Market (VoM) Project

Phase 1: Planning and Setup (Weeks 1-2)

1. Kickoff and Alignment:

- Define project scope, deliverables, and timeline with the sponsor organization (Costco).
- Confirm team roles and responsibilities.
- Identify data sources (Reddit, Instagram, X) and set up Agile tools (Jira/Trello).

2. Technical Setup:

- Configure hardware and software environments.
- Finalize the tools and libraries for data collection, processing, and visualization.

Phase 2: Data Collection and Preprocessing (Weeks 3-5)

1. Data Collection:

- Extract data using APIs and web scraping tools like BeautifulSoup, Selenium, or Scrapy.
- Store raw data in a centralized database or data lake.

2. Data Preprocessing:

- Clean data to remove duplicates, noise, and irrelevant content.
- Segment data by product, service mentions, and geography.
- Conduct quality assurance checks to ensure reliability.

Phase 3: Sentiment Analysis and Knowledge Graph Development (Weeks 6-12)

1. Exploratory Data Analysis (EDA):

- Analyze patterns, trends, and anomalies in the dataset.

2. Sentiment Analysis:

- Use NLP models (e.g., VADER, TextBlob, or BERT) to classify sentiments as positive, negative, or neutral.
- Evaluate and optimize models for accuracy.

3. Knowledge Graph Development:

- Create a graph linking products, trends, and sentiment attribution tags.

- **Integrate propensity scoring to rank product/service relevance.**

Phase 4: Dashboard and Visualization Development (Weeks 13-15)

1. Dashboard Design:

- **Develop an interactive dashboard using Python (Plotly, Dash) or Tableau.**
- **Include visualizations such as knowledge graphs, sentiment trends, and product scores.**

2. User Testing:

- **Conduct testing sessions with stakeholders to gather feedback.**
- **Refine the dashboard based on usability and functionality.**

Phase 5: Summarization Chatbot Development (Weeks 16-17)

1. Chatbot Development:

- **Build a summarization chatbot prototype using Rasa, Dialogflow, or OpenAI APIs.**
- **Train the chatbot to summarize key insights dynamically.**

2. Integration and Testing:

- **Integrate the chatbot with the dashboard for seamless interaction.**
- **Conduct user testing and refine responses based on feedback.**

Phase 6: Final Deliverables and Presentation (Week 18)

1. Final Deliverables:

- **Cleaned and segmented dataset.**
- **Sentiment analysis report and recommendations.**
- **Knowledge graph with propensity scoring.**
- **Interactive dashboard and chatbot prototype.**
- **Reusable NLP models and framework in a notebook format.**

2. Stakeholder Presentation:

- **Present insights, deliverables, and technical documentation to Costco stakeholders.**
- **Collect feedback and discuss next steps.**

Stretch Goals (Optional, if Time Permits)

- 1. Real-Time Sentiment Monitoring:**
 - **Develop a real-time sentiment analysis and monitoring dashboard.**
 - 2. Trend Analysis:**
 - **Integrate advanced topic modeling for identifying emerging trends.**
 - 3. Enhanced Chatbot Functionality:**
 - **Enable live social media integration for dynamic summarization.**
 - 4. Scalable Web Application:**
 - **Convert the dashboard into a full-scale web application for broader usability.**
-

Deliverable Milestones

- **Week 5: Completion of data collection and preprocessing.**
- **Week 12: Delivery of sentiment analysis and knowledge graph.**
- **Week 15: Completion of dashboard development and user testing.**
- **Week 17: Summarization chatbot prototype ready.**
- **Week 18: Final deliverables and presentation to Costco.**