**Business Report: Realtime Customer Insights Processor**

**1. Executive Summary**

* **Project Goal:** To develop and deploy a "Realtime Customer Insights Processor" application that enables businesses to automatically analyze customer feedback from Google Sheets, generate LLM-powered summary reports, and perform granular sentiment analysis to derive actionable insights quickly and efficiently.
* **Time Frame:**
  + **Start Date:** April 1, 2025
  + **Projected End Date** May 10, 2025
* **Key Highlights:**
  + Successfully integrated Google Sheets for real-time data fetching.
  + Implemented robust batch reporting functionality with configurable LLMs (Ollama-based) and PDF export.
  + Developed and integrated a custom TensorFlow/Keras sentiment analysis model for row-level feedback analysis.
  + Created an intuitive Streamlit-based user interface with three distinct functional tabs: Batch Reporter, Sentiment Analyzer, and Ad-hoc Sentiment Tester.
  + Established persistent state management for user configurations and processing progress.

**2. Project Objectives**

* **Short-Term Objectives (Achieved/In Progress):**
  + Develop a stable data ingestion pipeline from Google Sheets. (Achieved)
  + Implement LLM-driven batch reporting with configurable models and prompt templates. (Achieved)
  + Integrate a pre-trained sentiment analysis model for detailed feedback review. (Achieved)
  + Design and build a user-friendly Streamlit interface for all core features. (Achieved)
  + Enable PDF generation for batch reports. (Achieved)
  + Implement state persistence for application settings and progress. (Achieved)
* **Long-Term Objectives:**
  + Enhance scalability to support a larger volume of data and concurrent users.
  + Expand data source options beyond Google Sheets (e.g., direct database connections, CSV uploads).
  + Incorporate more advanced NLP features, such as topic modeling and trend detection within reports.
  + Offer user authentication and role-based access control for enterprise use.

**3. Progress Report**

* **Tasks Completed:**
  + **Core Application Structure:** app.py established with tab-based navigation.
  + **Configuration Management:** config.py created for LLM options, prompts, and styling.
  + **State Management:** state\_manager.py implemented for loading/saving application state.
  + **Google Sheet Data Fetching:** core\_logic.py functions for cached and fresh data retrieval.
  + **LLM Integration (Batch Reporter):** Implemented in core\_logic.py, including review formatting and report generation.
  + **PDF Generation:** pdf\_utils.py developed for Markdown to PDF conversion.
  + **Sentiment Model Loading & Preprocessing:** sentiment\_analyzer.py built with model loading, text cleaning, and encoding.
  + **Sentiment Prediction Logic:** Implemented in sentiment\_analyzer.py for both batch and ad-hoc analysis.
  + **UI Development (All Tabs):** Streamlit widgets and layouts for all three tabs completed.
* **Upcoming Work (Post-Initial Release/Next Phase):**
  + Deployment to a staging/production environment.
  + Monitoring of application performance and user feedback post-launch.
  + Begin planning for Phase 2 features (e.g., additional data sources, advanced NLP).

**4. Budget and Financials**

This project was successfully executed with **zero direct financial expenditure** for the initial development phase. This was achieved through the strategic utilization of freely available high-quality resources and tools.

* **Budget Allocation (Initial Phase):**
  + **Data Storage (Google Drive):** $0 (Utilized Google Drive's free 15GB tier for storing CSV data.)
  + **Model Training (Google Colab):** $0 (The TensorFlow/Keras sentiment analysis model was trained using Google Colab's free T4 GPU resources.)
  + **LLM Operations (Local Ollama):** $0 (Leveraged locally hosted Ollama with 4-bit quantized LLM models, enabling efficient operation on existing hardware without API costs.)
  + **Software/Tooling:** $0 (Core development tools like Python, Streamlit, TensorFlow, NLTK, and other libraries are open-source and free to use.)
* **Risks to Financials (Future Considerations):**
  + **Scalability of Free Tiers:** If data volume for Google Drive storage or processing needs for Google Colab significantly exceed free tier limits in future operational use, subscription costs could be incurred.
  + **Cloud Deployment & Operational Costs:** Future cloud deployment for wider accessibility or production-grade stability would introduce hosting, data transfer, and operational costs.
  + **Transition to Paid LLM APIs or Services:** If future requirements necessitate features, performance, or reliability levels only available through paid LLM APIs or other commercial services, this would require budget allocation.

**5. Challenges and Risks**

* **Challenges Encountered:**
  + **Initial LLM Prompt Engineering:** Iterative process to refine the SYSTEM\_PROMPT\_TEMPLATE in config.py to ensure consistent and high-quality report generation across different LLMs and review content.
  + **Sentiment Model Accuracy:** Ensuring the pre-trained sentiment model (saved\_model.h5) generalizes well to diverse customer feedback language. Initial tests showed good performance, but continuous monitoring will be needed.
* **Risk Mitigation:**
  + **LLM Prompting:** Established a clear, structured prompt template and tested with various data. Documentation includes guidance on prompt adaptation if needed.
  + **Sentiment Model:** The ad-hoc testing tab allows users to quickly validate sentiment on edge cases. Future work may involve fine-tuning or retraining the model with domain-specific data if performance dips.

**6. Timeline and Milestones**

* **Completed Milestones:**
  + **Project Kick-off & Requirements Gathering:** April 1 - April 5, 2025 (On Time)
  + **Initial Data Collection & Preprocessing (for Sentiment Model):** April 6 - April 15, 2025 (Delayed - see "Delays" section below)
  + **Sentiment Model Training & Evaluation:** April 16 - April 22, 2025 (Adjusted due to data collection delay, completed within revised timeframe)
  + **Core Logic & Backend Development (Data Fetching, LLM, Sentiment Integration):** April 10 - April 28, 2025 (Partially overlapped with model tasks, progress maintained)
  + **UI Development & Integration (Streamlit Tabs):** April 20 - May 5, 2025
  + **PDF Feature Completion & Initial End-to-End Testing:** May 1 - May 7, 2025
  + **Internal Review, Final Bug Fixes & Documentation Finalization:** May 8 - May 9, 2025
  + **Project Completion & Sign-off (Initial Version):** May 10, 2025 (Achieved target end date despite initial delay)
* **Future Milestones (Post-Initial Version - Next Phase):**
  + **Real-world Deployment & Pilot Program:**
    - Deploy the application to a controlled staging/pilot environment (e.g., internal cloud server or dedicated machine).
  + **Integration of Advanced NLP Techniques:**
    - Research and implement topic modeling to automatically identify prevalent themes in batch reports.
    - Investigate trend analysis capabilities to track sentiment and themes over time
  + **Expansion of Data Sources:**
    - Develop connectors for direct database access (e.g., PostgreSQL, MySQL).
    - Implement functionality for users to upload CSV/Excel files directly.
    - Explore integration with survey platforms or social media APIs
* **Delays:**
  + **Data Collection for Sentiment Model:** A delay of approximately 5-7 days was encountered during the initial data collection phase (originally planned for April 6 - April 10, completed around April 15).
    - **Reason:** The process of sourcing, cleaning, and manually labeling a sufficiently diverse dataset for training the TensorFlow/Keras sentiment model took longer than initially estimated. Ensuring data quality and representativeness was prioritized.

**7. Next Steps**

With the successful completion and sign-off of the initial version of the Realtime Customer Insights Processor, the focus now shifts to validating its real-world utility and strategically planning for future enhancements outlined in the long-term objectives.  
The following steps are prioritized:

* **Deployment and Pilot Program (Immediate Priority):**
  + **Package and Deploy V1:** Prepare the application and deploy it to a controlled, stable, internal pilot environment.
  + **Launch Pilot Program:** Initiate testing with a targeted group of end-users to use the application with real, operational data sources (starting with Google Sheets).
  + **Monitor & Gather Feedback:** Actively monitor application performance and stability. Establish clear channels to gather detailed user feedback on functionality, usability, report quality, and sentiment analysis accuracy during the pilot.
* **Analysis and Phase 2 Planning (Near-Term Priorities):**
  + **Evaluate Pilot Results:** Systematically analyse the performance data and user feedback gathered from the pilot program. Use these insights to confirm the value of the tool, identify any critical V1 refinements, and accurately prioritise features for the next development phase.
  + **Scope Advanced NLP Features:** Based on pilot feedback and the long-term objectives, begin research, technical design, and effort-scoping for the priority NLP enhancements, focusing initially on Topic Modeling and Sentiment Trend Analysis.
  + **Scope Data Source Expansion:** Investigate the technical requirements and design options for implementing the most requested new data sources, likely beginning with direct CSV/Excel uploads and common Database connectors.
  + **Assess Resource & Infrastructure Needs:** Evaluate the scalability requirements and potential costs (for hosting, and/or paid services, if required) necessary for a wider rollout or for implementing Phase 2 features. Define the resource plan (developer time, budget, etc.) required for the next development phase, building on the success of the zero-cost V1.

**8. Conclusion and Recommendations**

The "Realtime Customer Insights Processor" project has successfully met its initial objectives, delivering a functional and valuable tool for automated customer feedback analysis. The application's ability to leverage LLMs for insightful reporting and provide granular sentiment analysis offers significant potential for businesses to quickly understand customer sentiment and make data-driven decisions.

* **Recommendations:**
  + **Approve Project for Internal Pilot:** Proceed with a controlled internal pilot program with select teams (e.g., Customer Service, Marketing) to gather real-world usage data and further refine the tool.
  + **Allocate Resources for Phase 2:** Based on pilot feedback, allocate budget and resources for Phase 2 development, focusing on enhanced scalability, broader data source compatibility, and advanced analytical features.
  + **Explore Cloud Deployment Options:** Investigate cost-effective and secure cloud deployment strategies to make the application more widely accessible within the organization.
  + **Invest in Ongoing Model Maintenance:** For the custom sentiment model, plan for periodic review and potential retraining/fine-tuning with new data to maintain accuracy.