PACE Strategy Document: TikTok Claims Classification Data Preparation

---

1. Project Overview

Project Title: TikTok Claims Classification Data Preparation
Project Goal: Build a structured dataset for claims classification, perform initial data inspection, and summarize key descriptive statistics to support model development.

Key Deliverables:

- A structured DataFrame containing TikTok claims classification data
- Data type examination for each column
- Descriptive statistical summary of key variables
- Jupyter Notebook for data preparation and initial exploration
- Executive summary highlighting key data insights

---

2. PACE Workflow Breakdown

P - Plan

- **Identify stakeholders:**
    1. *Data Team:* Willow Jaffey (Lead), Rosie Mae Bradshaw (Manager), Orion Rainier (Scientist)
    2. *Cross-functional Team:* Mary Joanna Rodgers (Project Management), Margery Adebowale (Finance), Maika Abadi (Operations)
- **Define project milestones:**
    1. Build a DataFrame for TikTok claims classification data
    2. Examine data types of each column
    3. Compute and analyze descriptive statistics
    4. Prepare the Jupyter Notebook for data inspection and coding prep
    5. Communicate insights through an executive summary
- **Resources required:** Jupyter Notebook, Python (Pandas, NumPy), dataset access

A - Analyze

- **Data Inspection:** Identify missing values, inconsistencies, and data quality issues
- **Data Types Review:** Ensure correct data types for numerical, categorical, and text-based features
- **Descriptive Statistics:** Compute summary statistics, distribution analysis, and data variability
- **Initial Data Insights:** Identify key trends and anomalies in the dataset

C - Construct

- **Data Cleaning:** Handle missing values, remove duplicates, and standardize formats
- **Feature Engineering:** If necessary, create new variables to improve data usability
- **Data Visualizations:** Generate histograms, boxplots, and correlation heatmaps
- **Jupyter Notebook Documentation:** Ensure clear explanations and structured workflow

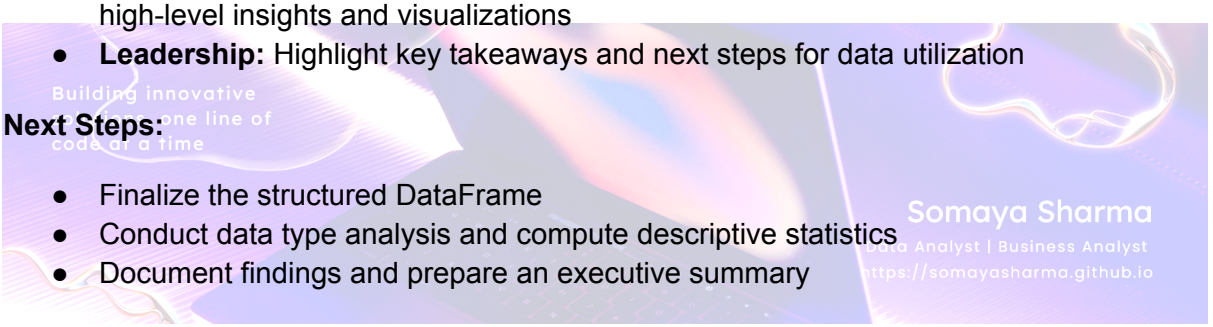
E - Execute

- **Implementation:** Finalize cleaned DataFrame for analysis
- **Validation:** Cross-check data integrity and correctness
- **Presentation:** Prepare an executive summary with key findings
- **Next Steps:** Outline how insights will contribute to the predictive model

---

3. Communication Strategy

- **Data Team (Technical):** Provide concise summaries in Jupyter Notebook with statistical outputs
- **Cross-functional Team (Non-Technical):** Share an executive summary with high-level insights and visualizations
- **Leadership:** Highlight key takeaways and next steps for data utilization



Next Steps:

- Finalize the structured DataFrame
- Conduct data type analysis and compute descriptive statistics
- Document findings and prepare an executive summary

This document serves as a structured guide for executing the TikTok Claims Classification Data Preparation project, ensuring clarity and alignment with data analysis objectives.