Industry Standard Documentation

**1. Project Charter:**

* **Project Title:** Customer Segmentation for a Retail Store
* **Project Manager:** Sai Charan Narsina
* **Start Date:** 15 July
* **End Date:** 24 July
* **Objectives:** To segment customers into distinct groups based on their purchasing behavior.
* **Scope:** Data cleaning, EDA, customer segmentation using K-Means, visualization using Matplotlib and Power BI.
* **Deliverables:** Insights, conclusions, and recommendations.

**2. Business Requirements Document (BRD):**

* **Business Problem:** Lack of understanding of different customer profiles leading to untargeted marketing strategies.
* **Business Objectives:** To improve customer satisfaction and sales by understanding customer segments.
* **Functional Requirements:** Data analysis, clustering, and visualization.
* **Non-functional Requirements:** Performance, scalability, and usability.

**3. Technical Requirements Document (TRD):**

* **Data Sources:** Mall Customers dataset
* **Technologies:** Python, Jupyter Notebook, Matplotlib, Seaborn, Scikit-learn, Power BI
* **Architecture:** Data preprocessing, EDA, clustering, and visualization
* **Data Flow:** Import data → Clean data → Analyze data → Segment customers → Visualize results

**4. Project Plan:**

* **Tasks:** Data collection, data cleaning, EDA, clustering, visualization, documentation
* **Timeline:** One weak
* **Risks:** Data quality issues, algorithm performance, visualization limitations

5. Final Report:

* Executive Summary:
* This project aimed to segment customers of a retail store into distinct groups based on their purchasing behavior to enhance targeted marketing strategies. Using the Mall Customers dataset, we performed data cleaning, exploratory data analysis (EDA), and customer segmentation using the K-Means clustering algorithm. The results were visualized to provide actionable insights.
* Introduction:
* Understanding customer profiles is crucial for retail stores to design effective marketing strategies. This project focuses on segmenting customers based on their demographics and purchasing behavior using data analysis and clustering techniques.
* Methodology:
* Data Collection:
* The dataset 'Mall\_Customers.csv' was loaded and examined for completeness and consistency.
* Data Cleaning:
* Missing values were handled by imputing the mean for 'Age' and the mode for 'Gender'.
* Columns were renamed for better readability.
* Categorical variables were encoded numerically.
* Exploratory Data Analysis (EDA):
* Distribution of age, annual income, and spending score were visualized using histograms.
* Relationships between features were explored using scatter plots.
* Clustering:
* Selected features: Age, Annual Income, Spending Score.
* Features were standardized using StandardScaler.
* K-Means clustering algorithm was applied to segment customers into 5 clusters.
* Visualization:
* Distributions were visualized using histograms.
* Customer segments were visualized using scatter plots.
* Results:
* The dataset was successfully cleaned and prepared for analysis. The K-Means clustering algorithm identified five distinct customer segments based on age, annual income, and spending score. The scatter plots illustrated clear distinctions between the clusters, providing valuable insights into customer behavior.
* Conclusion:
* The customer segmentation analysis revealed distinct groups with unique purchasing behaviors. These insights can help the retail store tailor marketing strategies, improve customer satisfaction, and potentially increase sales.