**📝 Project 2 – Customer Segmentation Report**

**Project Title:**

**Customer Segmentation using K-Means Clustering**

**Intern:**

**Suhani Gajre**  
Infotact Solutions – Internship Project 2  
Individual Submission

**📁 Dataset Description:**

* **Dataset Name**: Mall\_Customers\_5.csv
* **Source**: Kaggle – Customer Segmentation Dataset
* **Attributes**:
  + CustomerID
  + Gender
  + Age
  + Annual Income (in k$)
  + Spending Score (1–100)

**🧰 Tools & Libraries:**

* **Python** (Google Colab)
* pandas, matplotlib, seaborn
* scikit-learn for KMeans clustering
* GitHub for version control

**📊 Project Objectives:**

* Segment customers based on behavioral and demographic data
* Apply K-Means clustering algorithm
* Identify meaningful customer groups
* Provide marketing insights and suggestions

**📋 Project Steps:**

**✅ Week 1 – Data Cleaning & EDA:**

* Loaded and explored the dataset
* Checked for nulls, datatypes, and distribution
* Plotted distributions of Age, Gender, Income, and Spending Score
* Encoded categorical values (Gender)
* Scaled data for better clustering

**✅ Week 2 – K-Means Clustering:**

* Used **Elbow Method** to determine optimal clusters (k = 4)
* Applied **KMeans** from scikit-learn
* Created new column Cluster in dataset
* Visualized clusters using scatter plot

**✅ Week 3 – Cluster Profiling:**

* Grouped by cluster and calculated averages
* Identified distinct customer segments based on behavior

**✅ Week 4 – Summary & Insights:**

* Completed GitHub repo structure
* Added .ipynb notebook with code, plots, and notes
* Wrote final analysis and uploaded to GitHub

**🧠 Cluster Profiles:**

| **Cluster** | **Profile Name** | **Description** |
| --- | --- | --- |
| 0 | **Premium Shoppers** | High income, high spending (mostly male) |
| 1 | **Stable Buyers** | Older, moderate income, high spenders (mostly male) |
| 2 | **Cautious Rich Shoppers** | High income, low spending (all male) |
| 3 | **Low Engagement Females** | High income, low spending (all female) |

**💡 Recommendations:**

* 🎯 **Target Cluster 0** with luxury promotions and loyalty benefits
* 🛍 **Re-engage Cluster 3** (high-income females with low spending)
* 📦 **Upsell to Cluster 2** with bundles and seasonal offers
* 🤝 **Retain Cluster 1** with consistent, quality experiences

**✅ Outcome:**

The project successfully grouped customers into meaningful segments using KMeans and revealed actionable marketing insights. This lays a foundation for personalized strategies and data-driven decision-making in retail businesses.

**📎 GitHub Repo:**

📌https://github.com/suhanigajre/customer-segmentation-clustering.git