**Project Title**

**Unlocking Growth Potential: A Study on Increasing Fashion Sales at Daraz (Focused on Seller Segmentation for Improved Marketing Strategies)**

**Project Description**

This project applies machine learning clustering techniques to segment Daraz sellers based on their behavioral and performance metrics. The goal is to enable Daraz to enhance seller engagement, provide personalized support, and optimize marketing strategies in the fashion category.  
A segmentation model using **K-Means Clustering** was developed, validated, and integrated with an interactive **Power BI Dashboard** for easy stakeholder interpretation.

**Files Included**

* Aggregated and cleaned seller dataset.
* Python notebook containing data preprocessing, feature engineering, model training, and evaluation code.
* Power BI dashboard for visualization of seller segments and insights.
* Project overview, steps to run, and key insights.

**Tools Used**

* Python (Pandas, NumPy, Scikit-learn, Matplotlib, Seaborn)
* Power BI for dashboard development and business-friendly data visualization.
* Excel for data cleaning and preliminary exploration.

**Steps to Run the Project**

1. **Preprocessing and Feature Engineering**
   * Clean raw transactional data.
   * Aggregate seller-level KPIs (GMV, orders, buyers, return rates).
   * Engineer new features (GMV growth rate, retention rate).
2. **Model Training and Validation**
   * Apply **StandardScaler** for feature scaling.
   * Train **K-Means Clustering**.
   * Validate clusters using **Silhouette Score** and **PCA visualization**.
3. **Dashboard Creation**
   * Import labeled seller segments into Power BI.
   * Create visuals (Segment Distribution, GMV by Segment, Orders by Segment, Seller-level insights).
   * Publish the dashboard for stakeholder analysis.

**Results Summary**

* **High Growth Sellers:** 20.92%
* **Stable Performers:** 61.84%
* **New or Struggling Sellers:** 17.24%

The Power BI dashboard enables Daraz to monitor seller segments, identify sellers needing support, and improve strategic planning.

**Key Learnings**

* Applied machine learning for unsupervised seller segmentation.
* Gained expertise in data preprocessing, model validation, and business intelligence reporting.
* Improved understanding of seller engagement strategies using data-driven insights.

**Future Work**

* Integrate the segmentation model with real-time Daraz data pipelines.
* Extend segmentation analysis to other product categories and country markets.
* Enhance the dashboard with predictive analytics for seller churn and growth forecasting.