

Project architecture

Project architecture



### Why This Low Level Design

A low-level design document typically provides a detailed, in-depth description of the implementation of a system, focusing on specific aspects such as modules, classes, algorithms, data structures, and their interactions

### Scope

The scope of the customer segmentation project encompasses several key areas that define the boundaries and extent of the work to be undertaken. It outlines the specific objectives, deliverables, and limitations of the project, providing a clear framework for what will be achieved.

The primary objective of this project is to develop a machine learning-based system for customer segmentation, enabling businesses to identify and categorize distinct customer groups based on their behaviors and characteristics. This involves collecting and preprocessing customer data, implementing various clustering algorithms, evaluating the performance of these algorithms, and interpreting the results to derive actionable insights. The ultimate goal is to provide businesses with a tool that helps in crafting targeted marketing strategies and improving customer satisfaction.

### Constraints

This project on customer segmentation using machine learning faces several constraints that may impact its execution and outcomes. Identifying and understanding these constraints is essential for setting realistic expectations and mitigating potential challenges. The quality and availability of data are significant constraints for this project. The effectiveness of customer segmentation relies heavily on the accuracy, completeness, and relevance of the data collected. Issues such as missing values, inconsistent data formats, and insufficient data can hinder the preprocessing and analysis stages. The project requires substantial computational resources for data processing, model training, and evaluation. The complexity and volume of customer data, combined with the need to run multiple clustering algorithms, can demand significant processing power and memory.

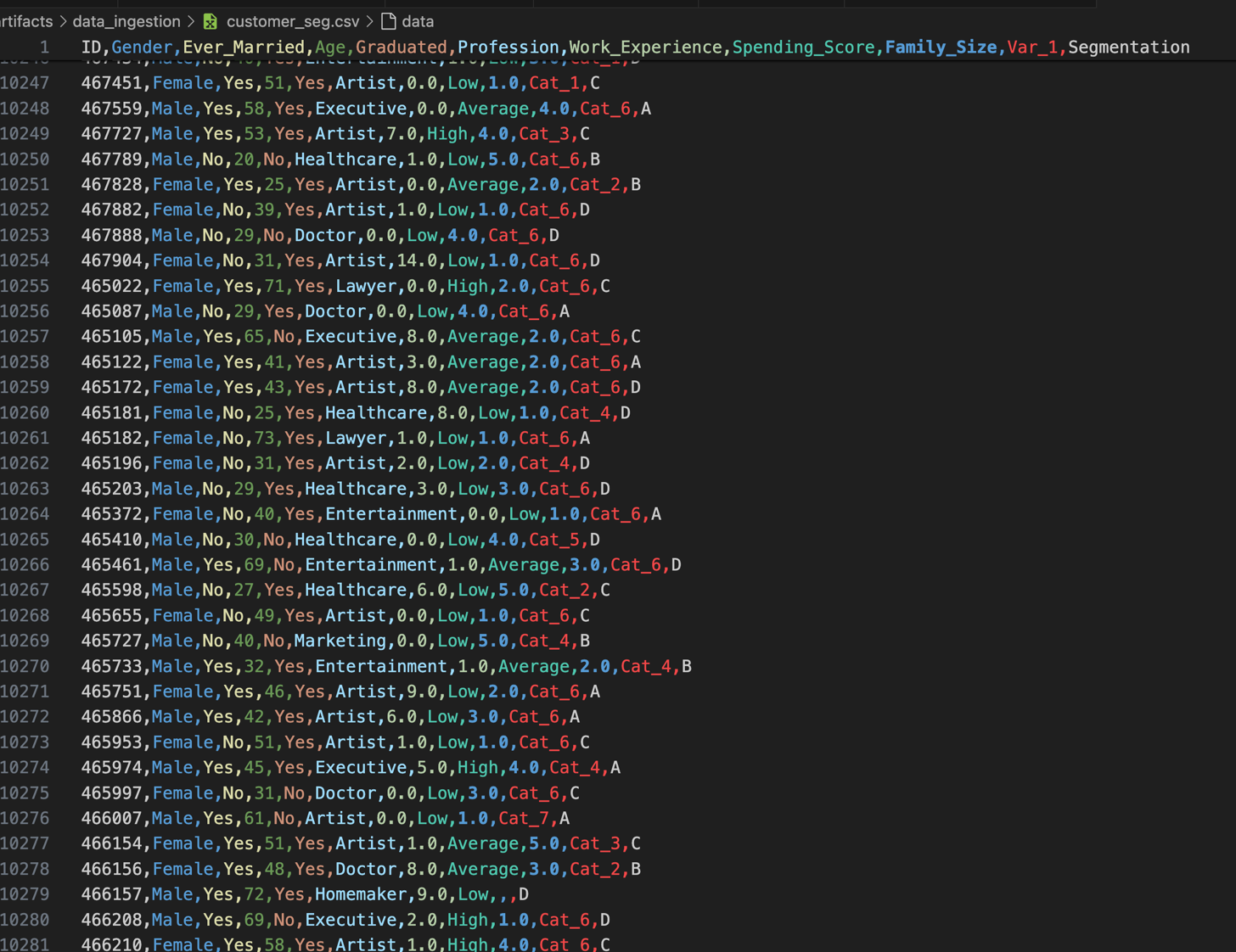
### Risks

The customer segmentation project using machine learning involves several risks that can potentially impact its success and effectiveness. Identifying these risks early on allows for the development of mitigation strategies to address them. Handling sensitive customer data poses significant privacy and security risks. Unauthorized access, data breaches, or misuse of customer information can lead to legal repercussions and damage the organization’s reputation. Ensuring robust data protection measures, such as encryption and secure access controls, is crucial to mitigate these risks.

### Out of Scope

To clearly define the boundaries of the customer segmentation project, it is essential to identify aspects that are beyond its scope. This helps manage expectations and focus resources on the primary objectives. The project will not address the implementation of real-time data processing or streaming analytics. While customer segmentation could benefit from real-time data updates, this project is limited to batch processing of historical data. Real-time processing requires additional infrastructure and technologies that are not covered within the current scope.

### Customer Segmentation Datasets Overview



### Dependent and Independent Schema

|  |  |
| --- | --- |
| Features Name | DataTypes |
| Gender | object |
| Ever married | object |
| Age | int |
| Graduated | object |
| Profession | object |
| Work Experiences | object |
| Spending score | object |
| Family size | int |
|  |  |
|  |  |
|  |  |

### Data Ingestion WorkFlow

### Data Validation WorkFlow

### Data Transformation WorkFlow

### ModelTraining WorkFlow