

# Understanding Data Processing: Ingestion and Summarization

This table of contents outlines key information regarding the implementation of a Data Ingestion API and a Text Summarization Tool, focusing on their functionalities and setup procedures.

## 1. Data Ingestion API Documentation

This section provides a comprehensive overview of the Data Ingestion API, including its purpose, endpoints, implementation details, setup instructions, and testing procedures.

### 1.1. Overview

Briefly introduces the Data Ingestion API and the Text Summarization Tool.

The Data Ingestion API handles CSV file uploads, data storage, and provides basic summary statistics.

The Text Summarization tool uses NLP to condense news articles.

### 1.2. Assignment 1: Data Ingestion API

Outlines the objective of building a RESTful API for managing CSV data and retrieving summary statistics.

### 1.3. Endpoints

**1.3.1. Upload CSV File:** Details the endpoint for uploading CSV files, including the URL, method, required headers, and possible responses.

**1.3.2. Retrieve Summary Statistics:** Describes the endpoint for retrieving summary statistics, including the URL, method, required headers, and response format.

### 1.4. Implementation Details

**1.4.1. Model Definition:** Explains the data model used to store uploaded CSV data in JSON format.

**1.4.2. View for Data Upload:** Outlines the process of validating file types and content, converting CSV data to JSON, and storing it in the database.

**1.4.3. View for Summary Statistics:** Describes how the system retrieves datasets, calculates summary statistics (mean, median, mode) for numerical columns, and returns the results in a structured format.

**1.4.4. Decorator for Payload Validation:** Explains the mechanism used to ensure that API requests include valid headers and file data.

### 1.5. Setup Instructions

Provides step-by-step instructions for setting up the Data Ingestion API, including cloning the repository, creating a virtual environment, installing dependencies, applying migrations, and running the server.

### 1.6. Testing Endpoints

Provides commands for testing both the CSV file upload and summary statistics retrieval endpoints.

## 2. Contact

Provides contact information for further inquiries.

Ankit Anand

[ankitnd@outlook.com](mailto:ankitnd@outlook.com)

8982541015