Data Science Project 2023 - Documentation

# Introduction

This project is the final deliverable for a Data Science course in 2023. It involves data cleaning, processing, and analysis using Python. The main goal is to demonstrate the application of data science techniques on a provided dataset.

# Project Structure

- `main.py`: Main script for executing the data analysis pipeline.  
- `Dag.py`: DAG (Directed Acyclic Graph) setup for the pipeline.  
- `cleaned\_dataset.csv`: Preprocessed dataset used for analysis.  
- `requirements.txt`: List of dependencies required to run the project.

# Installation

Clone the repository and install the dependencies:  
  
```bash  
git clone https://github.com/Alvaroinfantee/DataSciencePrject.git  
cd DataSciencePrject  
pip install -r requirements.txt  
```

# Usage

Run the main script to execute the analysis pipeline:  
  
```bash  
python main.py  
```

# Requirements

- Python 3.x  
- Pandas  
- Other libraries listed in `requirements.txt`

# Contributing

Contributions are welcome. Feel free to submit issues or pull requests.

# License

This project is licensed under the MIT License.