



CSV Upload & Automated Data Summary Using Flask

A Mini Project Report Submitted By

K. Saketh Varma - 2403A510B6

A. Nihal - 2403A510B7

N. Akshaya - 2403A510B8

In partial fulfillment of the course “AI Assisted Coding”

Under the guidance of

Dr. Rishabh Mittal

Assistant Professor - CSE



ACKNOWLEDGEMENT

We would like to express our sincere gratitude to **Dr. Rishabh Mittal**, Assistant Professor, Department of CSE, SR University, for his continuous guidance, encouragement, and valuable support throughout the duration of this mini project. His insights and constructive feedback greatly contributed to the successful completion of this work.

INTRODUCTION

Data analysis plays a crucial role in modern software applications, enabling users to extract meaningful insights from raw datasets. CSV files are widely used for storing such data, but manually analyzing them can be slow and inefficient. This project provides a simple web-based solution where users can upload a CSV file and instantly receive an automatic summary that includes statistics, missing values, correlations, and sample records. Built using Flask and Pandas, the system makes data exploration faster, easier, and accessible through any web browser.

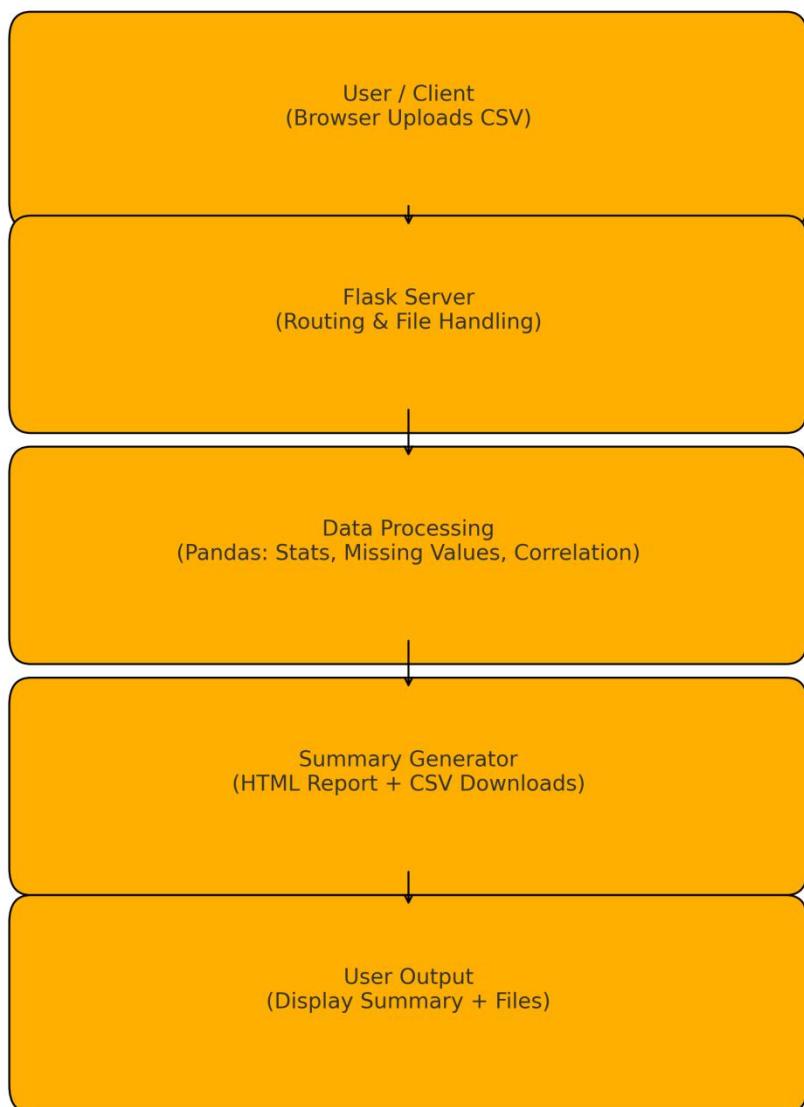
Key Features of the System:

- Upload any CSV file through a web interface
- Automatic data summary generation
- Detection of missing values
- Numerical statistics (mean, std, min, max, quartiles)
- Categorical column analysis
- Correlation matrix for numeric fields
- Display of first and last 5 rows
- Option to download summary as CSV
- Lightweight and user-friendly

TECHNOLOGIES USED

Component	Technology
Backend Framework	Flask (Python)
Data Processing	Pandas, NumPy
Frontend	HTML, CSS (via Flask templates)
File Handling	Flask File Upload
Optional Libraries	Matplotlib (safe backend), OS, IO
Programming Language	Python 3

System Architecture Diagram



SAMPLE OUTPUTS

127.0.0.1:5000

Gmail YouTube Maps

Upload CSV to get full summary

Drop a CSV file and get a complete human-friendly summary — statistics, missing values, sample rows and downloadable summaries.

height_students.csv

Or use curl to download CSV summary: curl -o summary.csv -F "file=@sample.csv" "http://127.0.0.1:5000/upload?format=csv"

Max upload size: 16 MB

127.0.0.1:5000/upload

Gmail YouTube Maps

CSV Summary Report

A quick, human-friendly overview of your uploaded CSV

Overall metrics

Total Rows	Total Columns	Duplicate Rows	Memory Usage (bytes)
19	2	0	1301

Missing values per column

column	missing_count	missing_pct
Name	0	0.0
Height	0	0.0

Column types & unique counts

column	dtype	unique_values
Name	object	19
Height	float64	15

Numeric descriptive statistics

column	count	mean	std	min	25%	50%	75%	max	skew	kurtosis
Height	19.0	6.757895	4.66051	2.3	5.1	5.8	6.15	23.2	2.883399	9.175463

Categorical summary (top values)

column	rank	value	count	pct
Name	1	anil	1	5.26
Name	2	iran	1	5.26
Name	3	karthika	1	5.26
Name	4	rita	1	5.26
Name	5	sta	1	5.26

CONCLUSION

This project successfully demonstrates how Flask and Pandas can be integrated to automate data exploration tasks. By allowing users to upload any CSV file and instantly receive detailed insights—such as summary statistics, missing values, correlations, and sample records—the system provides an efficient and user-friendly approach to initial data analysis. It eliminates the need for manual coding or complex tools and offers a quick, accessible solution that can be used by students, analysts, and beginners. Overall, the project enhances understanding of backend development, data processing, and web integration while delivering a practical and functional application.