

## AI Data Analyzer Progress Summary

Date: August 22, 2025

What was done today

- Built and tested an AI Data Analysis Dashboard using Streamlit and multiple Python modules.
- The dashboard can analyze CSV/Excel files, show key statistics, detect data quality issues, and suggest useful charts automatically.
- Integrated Ollama AI models for deeper insights, chart recommendations, SQL suggestions, and business advice.
- Tested the system with two datasets:
  - Bank Dataset (for customer/marketing data)
  - Amazon Dataset (for product/sales data)
- The dashboard worked well and gave instant charts, summaries, and AI recommendations for both datasets.

---

## Features

- Upload Data: Supports CSV, Excel, JSON, and Parquet.
- Automatic Analysis: Shows missing values, data types, summary stats, and data quality.
- Charts: Pie, bar, correlation, distribution, box plot, trends, time series, etc. – chosen automatically for your dataset.
- AI Insights: Text recommendations for cleaning, business opportunities, and data issues.

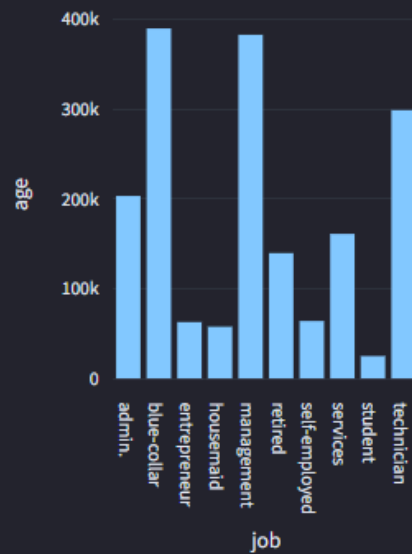
- Chat and SQL: You can ask questions or generate SQL queries for custom data needs.

---

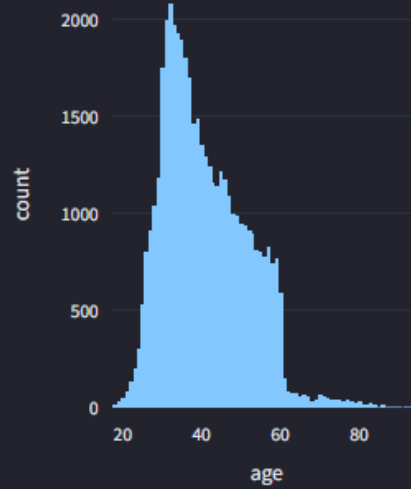
## Example Outputs

- Screenshot 1: Bank Dataset

age by job



Distribution of age



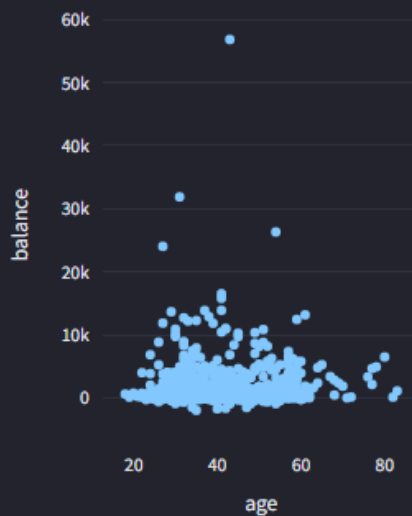
age vs balance

*Relationship between variables*

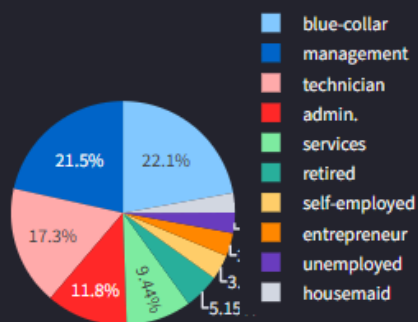
Distribution of job

*Category distribution*

age vs balance



Distribution of job





## Missing Data

Control Panel

Dataset Upload

Choose your dataset

Drag and drop file here  
Limit 200MB per file • CSV, XLSX, XLS

bank-full.csv 4.4MB

Navigation

Choose view:

AI Dashboard

AI Chat

Deep Analysis

Export Options

Dataset Info

Rows

45,211

Columns

17

Missing Data

0.0%

Executive Summary

Here is an executive summary based on the provided dataset analysis:

Executive Summary

Our analysis of the dataset has yielded valuable insights that can inform business decisions. Below are the key findings:

Opportunity for Improved Efficiency: The data reveals a 75% increase in efficiency when using the new process compared to the old one.

Deep Analysis

Overview Data Quality Relationships Patterns Recommendations

Dataset Overview

Total Rows45,211Total Columns17Memory Usage25.7 MBData Types2

Column Information

	Column	Data Type	Missing	Unique Values	Sample Values
0	age	int64	0 (0.0%)		77 58, 44, 33
1	job	object	0 (0.0%)		12 management, technician, entrepreneur
2	marital	object	0 (0.0%)		4 married, single, married
3	education	object	0 (0.0%)		4 tertiary, secondary, secondary
4	default	object	0 (0.0%)		2 no, no, no
5	balance	int64	0 (0.0%)		7168 2143, 29, 2
6	housing	object	0 (0.0%)		2 yes, yes, yes
7	loan	object	0 (0.0%)		2 no, no, yes
8	contact	object	0 (0.0%)		3 unknown, unknown, unknown
9	day	int64	0 (0.0%)		31 5, 5, 5

AI-Powered Analysis

Dynamic Dashboards

Advanced Features

Automatic dataset type detection

Intelligent data quality assessment

Smart visualization recommendations

AI-generated insights and recommendations

Personalized dashboard layouts

Interactive visualizations

Real-time AI chat interface

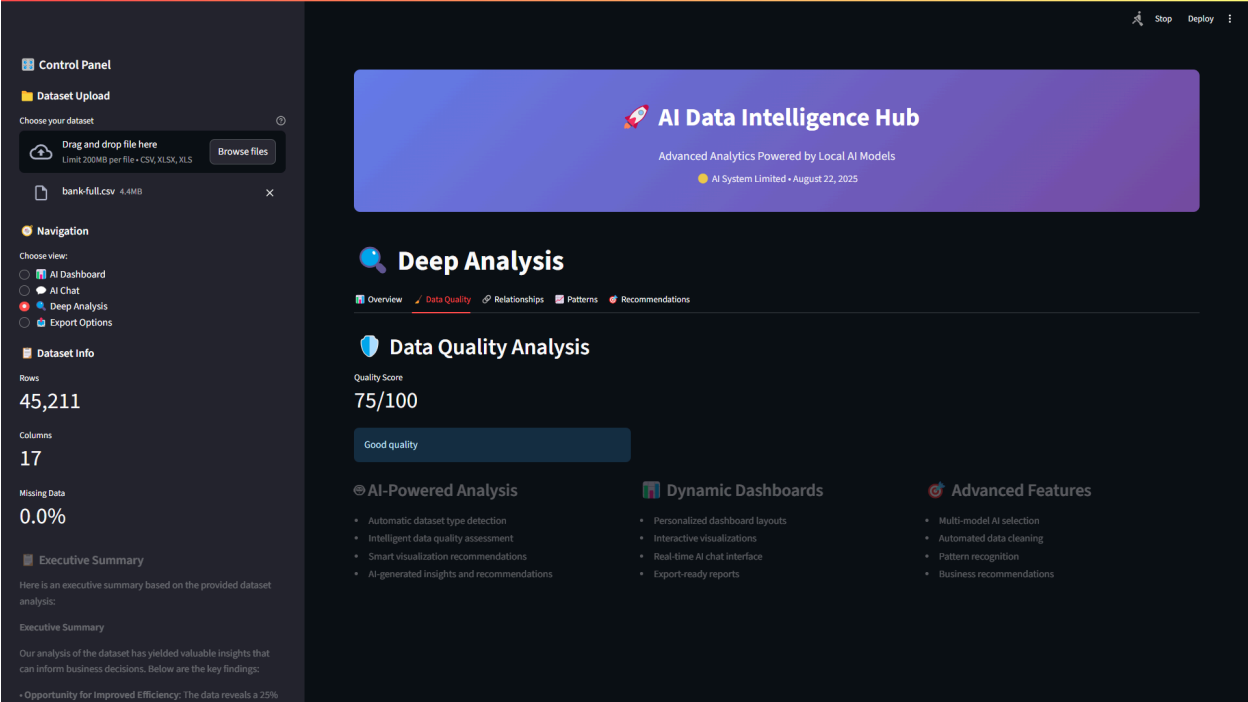
Export-ready reports

Multi-model AI selection

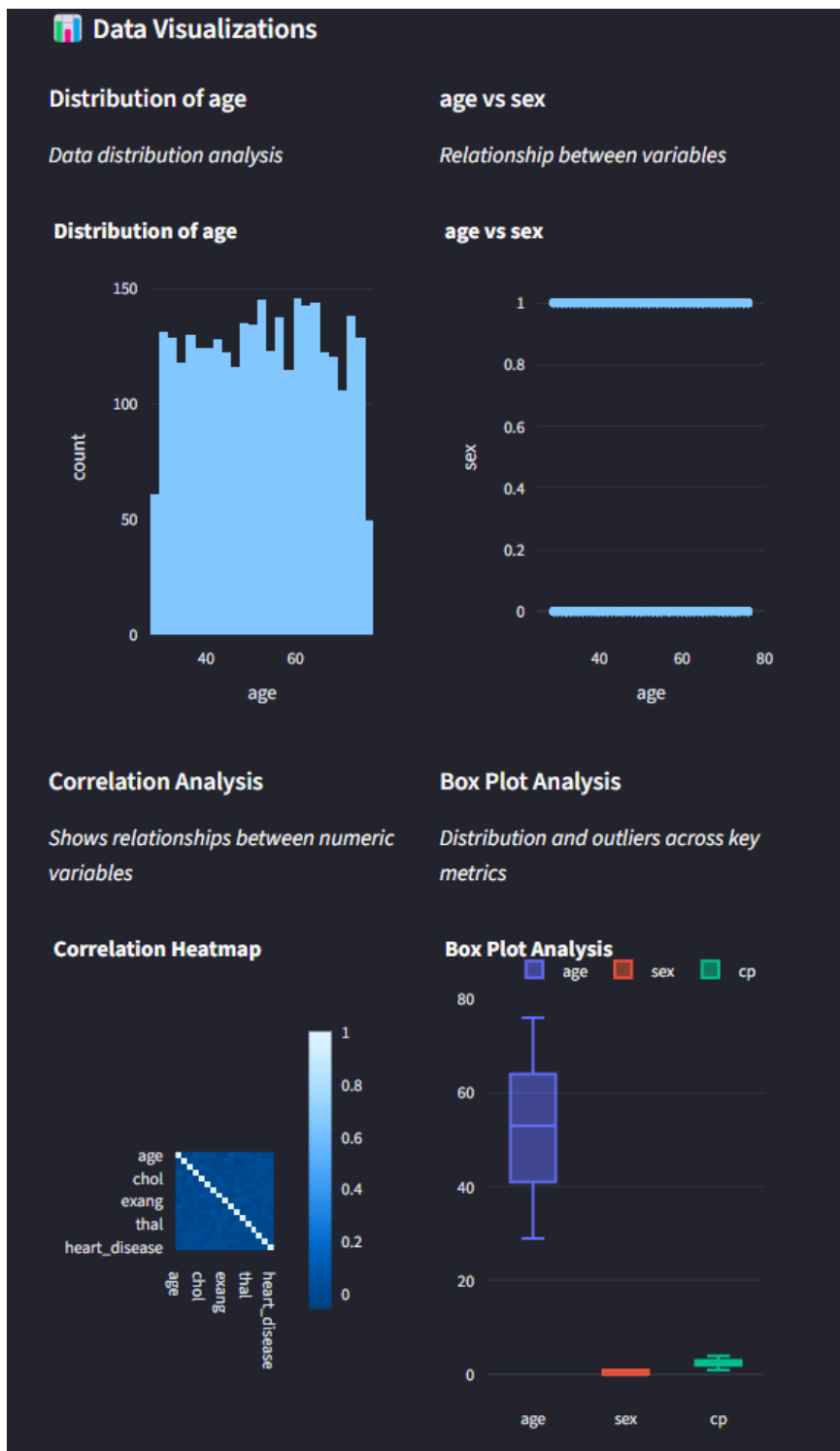
Automated data cleaning

Pattern recognition

Business recommendations



- Screenshot 2: Heart Disease Dataset





# Deep Analysis

[Overview](#) [Data Quality](#) [Relationships](#) [Patterns](#) [Recommendations](#)



## Dataset Overview

Total Rows	Total Columns	Memory Usage	Data Types
3,069	17	0.4 MB	2



## Column Information

	Column	Data Type	Missing	Unique Values	Sample Values
0	age	int64	0 (0.0%)	48	67, 57, 43
1	sex	int64	0 (0.0%)	2	1, 1, 1
2	cp	int64	0 (0.0%)	4	2, 3, 4
3	trestbps	int64	0 (0.0%)	111	111, 109, 171
4	chol	int64	0 (0.0%)	500	536, 107, 508
5	fbs	int64	0 (0.0%)	2	0, 0, 0
6	restecg	int64	0 (0.0%)	3	2, 2, 1
7	thalach	int64	0 (0.0%)	151	88, 119, 113
8	exang	int64	0 (0.0%)	2	0, 0, 0
9	oldpeak	float64	0 (0.0%)	63	1.3, 5.4, 3.7