Mango

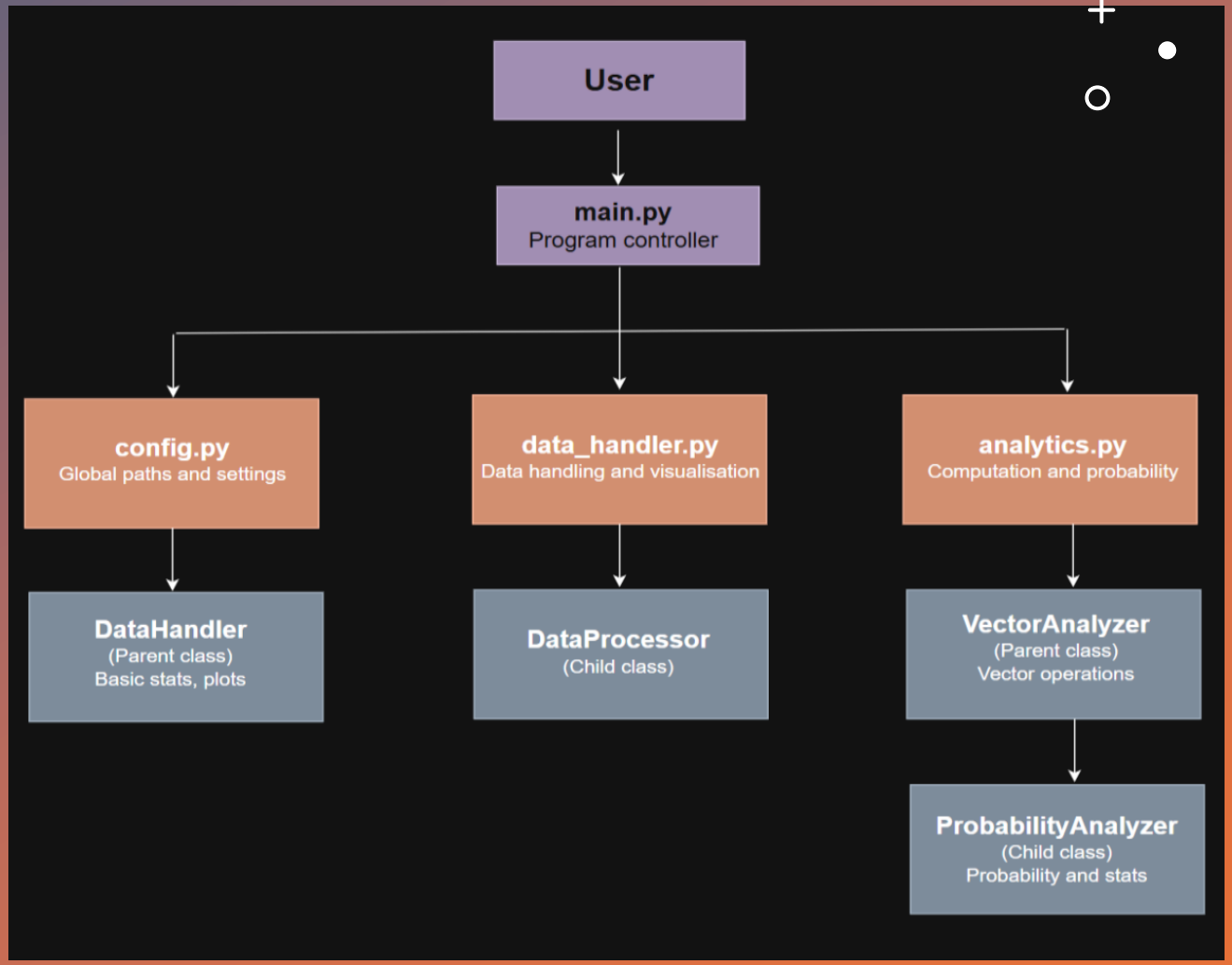
Team members-

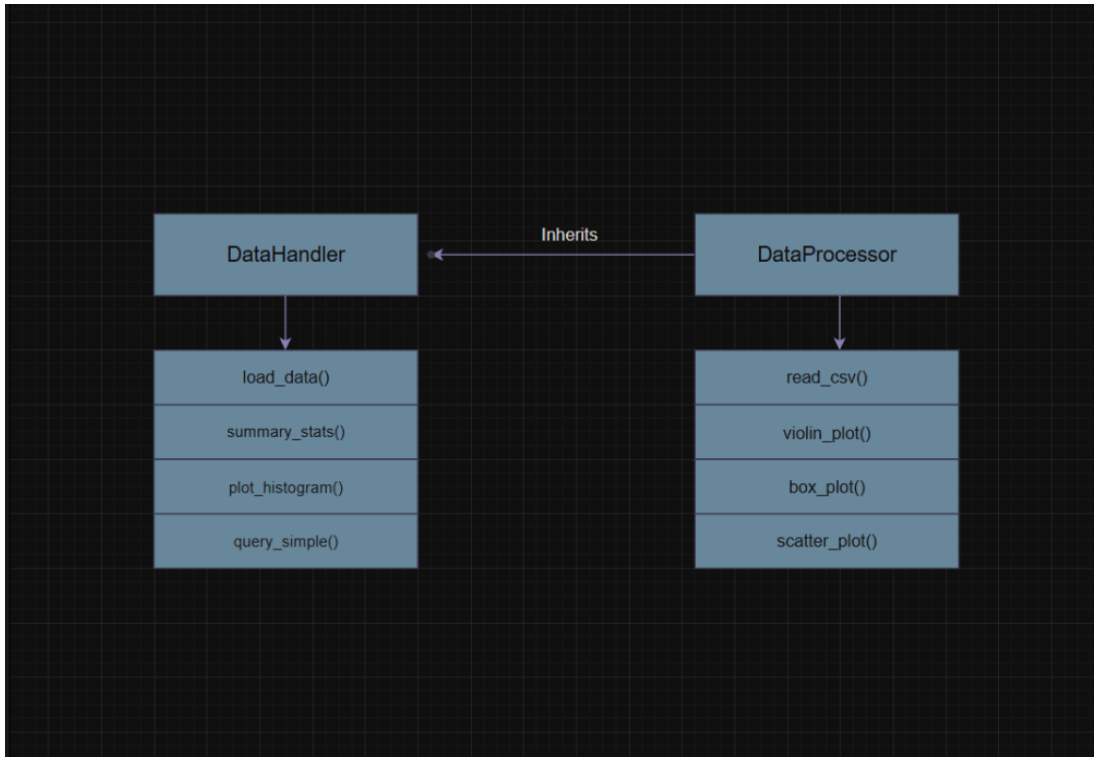
- Samiksha Gnawali
- Suyog Karki

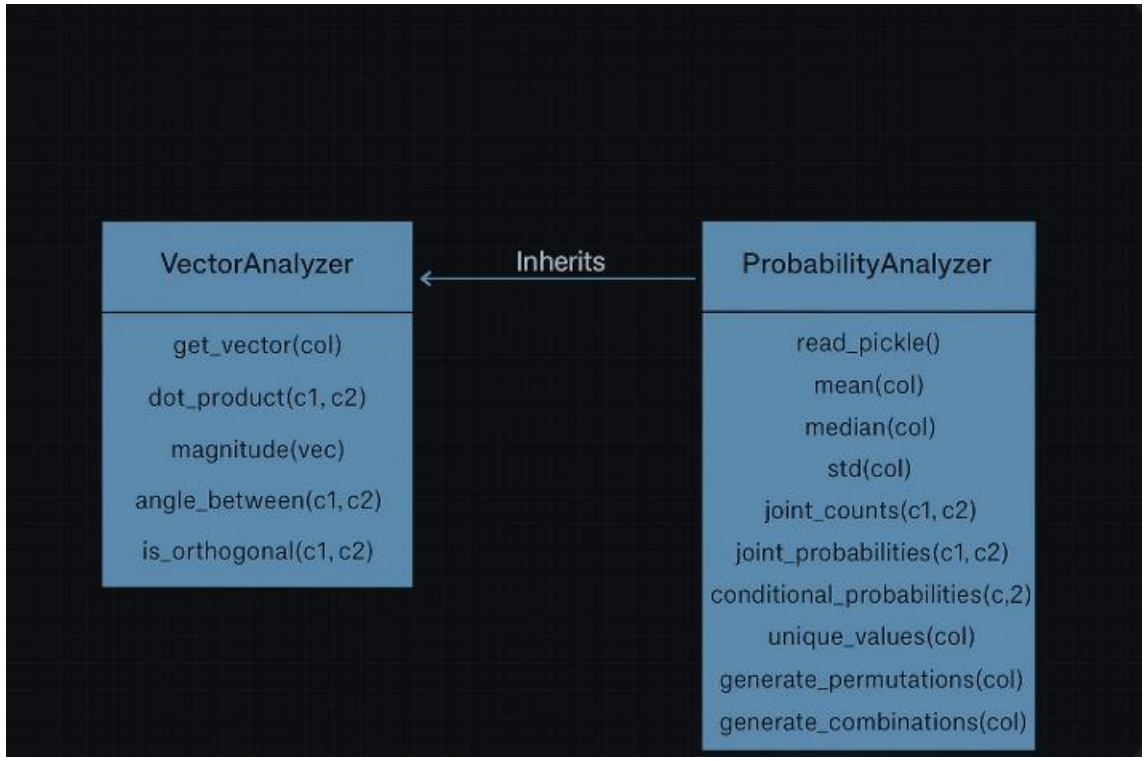
Goal of the project

- Our project is a Python-based data analysis tool that analyses patterns and relationships in a used car dataset. It focuses on understanding how different attributes such as year, selling price, fuel type, transmission, and kilometers driven affect a car’s resale value. The program will read and clean data, visualizing distributions and trends, and calculating key statistics. It can also compute probabilities and vector-based relationships between variables, helping to identify insights such as which car types tend to retain more value.

Module Communication Flow







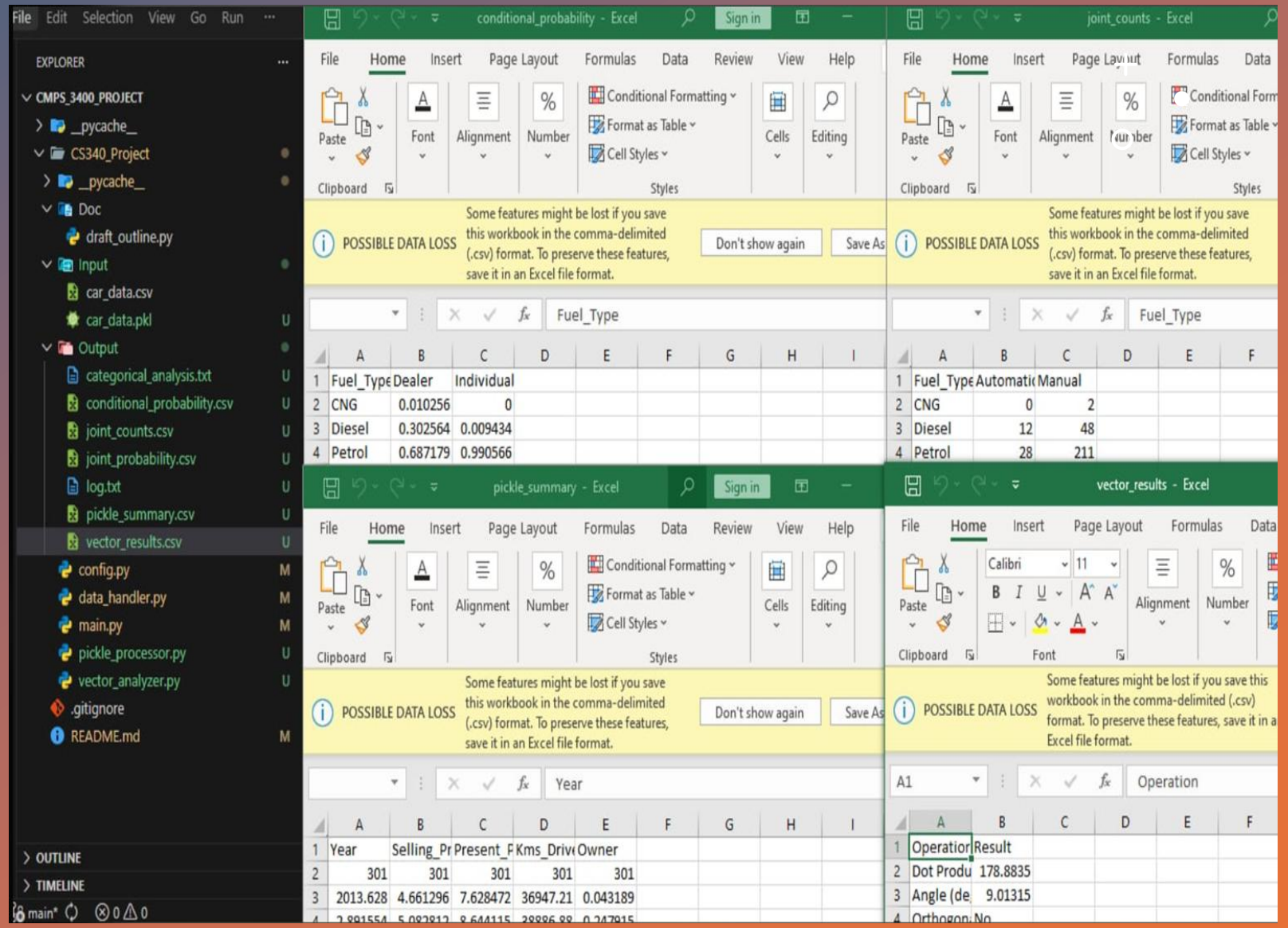
Class diagrams

Project manual

- Input data format
  - File type - CSV (car_data.csv)
  - Location – Input/ Folder
  - Data format

	Car_Name	Year	Selling_Price	Present_Price	Kms_Driven	Fuel_Type	Seller_Type	Transmission	Owner
	ritz	2014	3.35	5.59	27000	Petrol	Dealer	Manual	0
	sx4	2013	4.75	9.54	43000	Diesel	Dealer	Manual	0
	ciaz	2017	7.25	9.85	6900	Petrol	Dealer	Manual	0
	wagon r	2011	2.85	4.15	5200	Petrol	Dealer	Manual	0
	swift	2014	4.6	6.87	42450	Diesel	Dealer	Manual	0
	vitara brezza	2018	9.25	9.83	2071	Diesel	Dealer	Manual	0
	ciaz	2015	6.75	8.12	18796	Petrol	Dealer	Manual	0
	s cross	2015	6.5	8.61	33429	Diesel	Dealer	Manual	0
0	ciaz	2016	8.75	8.89	20273	Diesel	Dealer	Manual	0
1	ciaz	2015	7.45	8.92	42367	Diesel	Dealer	Manual	0
2	alto 800	2017	2.85	3.6	2135	Petrol	Dealer	Manual	0
3	ciaz	2015	6.85	10.38	51000	Diesel	Dealer	Manual	0
4	ciaz	2015	7.5	9.94	15000	Petrol	Dealer	Automatic	0
5	ertiga	2015	6.1	7.71	26000	Petrol	Dealer	Manual	0
5	dzire	2009	2.25	7.21	77427	Petrol	Dealer	Manual	0
7	ertiga	2016	7.75	10.79	43000	Diesel	Dealer	Manual	0
8	ertiga	2015	7.25	10.79	41678	Diesel	Dealer	Manual	0
9	ertiga	2016	7.75	10.79	43000	Diesel	Dealer	Manual	0
0	wagon r	2015	3.25	5.09	35500	CNG	Dealer	Manual	0
1	sx4	2010	2.65	7.98	41442	Petrol	Dealer	Manual	0
2	alto k10	2016	2.85	3.95	25000	Petrol	Dealer	Manual	0
3	ignis	2017	4.9	5.71	2400	Petrol	Dealer	Manual	0
4	sx4	2011	4.4	8.01	50000	Petrol	Dealer	Automatic	0
5	alto k10	2014	2.5	3.46	45280	Petrol	Dealer	Manual	0
6	wagon r	2013	2.9	4.41	56879	Petrol	Dealer	Manual	0
7	swift	2011	3	4.99	20000	Petrol	Dealer	Manual	0
8	swift	2013	4.15	5.87	55138	Petrol	Dealer	Manual	0
9	swift	2017	6	6.49	16200	Petrol	Individual	Manual	0
0	alto k10	2010	1.95	3.95	44542	Petrol	Dealer	Manual	0
1	ciaz	2015	7.45	10.38	45000	Diesel	Dealer	Manual	0
2	ritz	2012	3.1	5.98	51439	Diesel	Dealer	Manual	0
3	ritz	2011	2.35	4.89	54200	Petrol	Dealer	Manual	0
4	swift	2014	4.95	7.49	39000	Diesel	Dealer	Manual	0
5	ertiga	2014	6	9.95	45000	Diesel	Dealer	Manual	0

Output data format

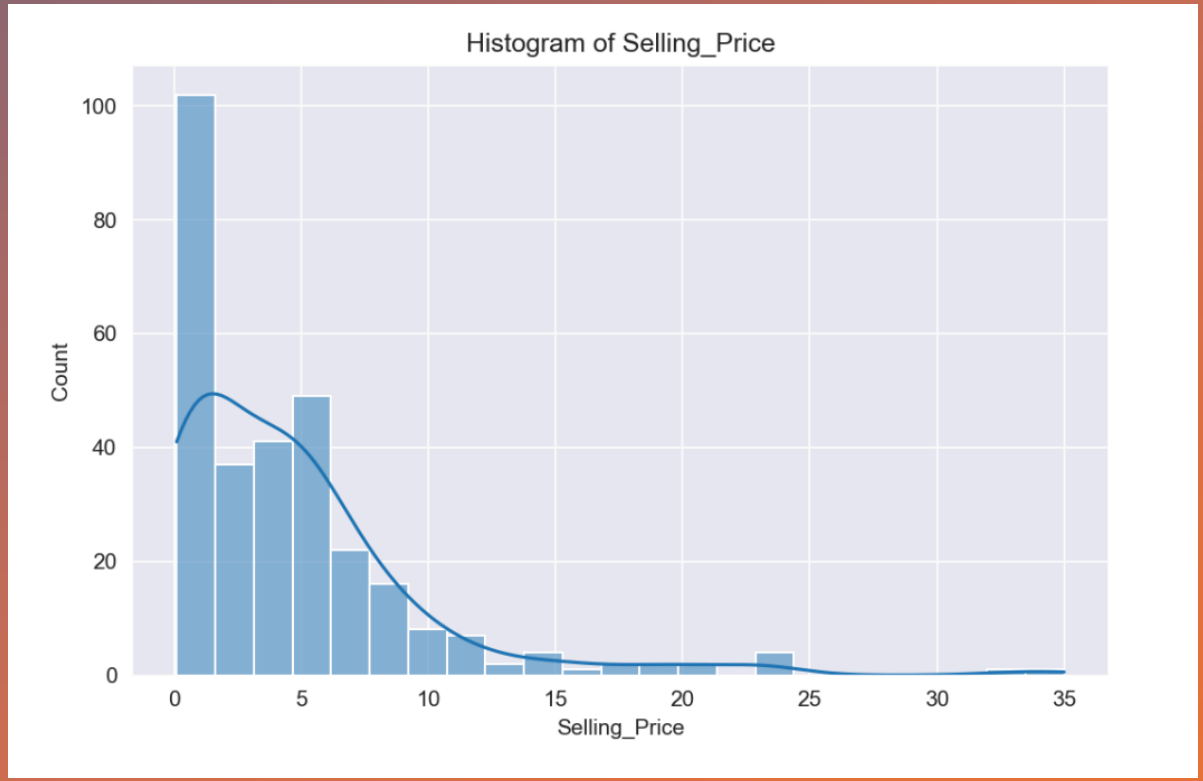


User manual

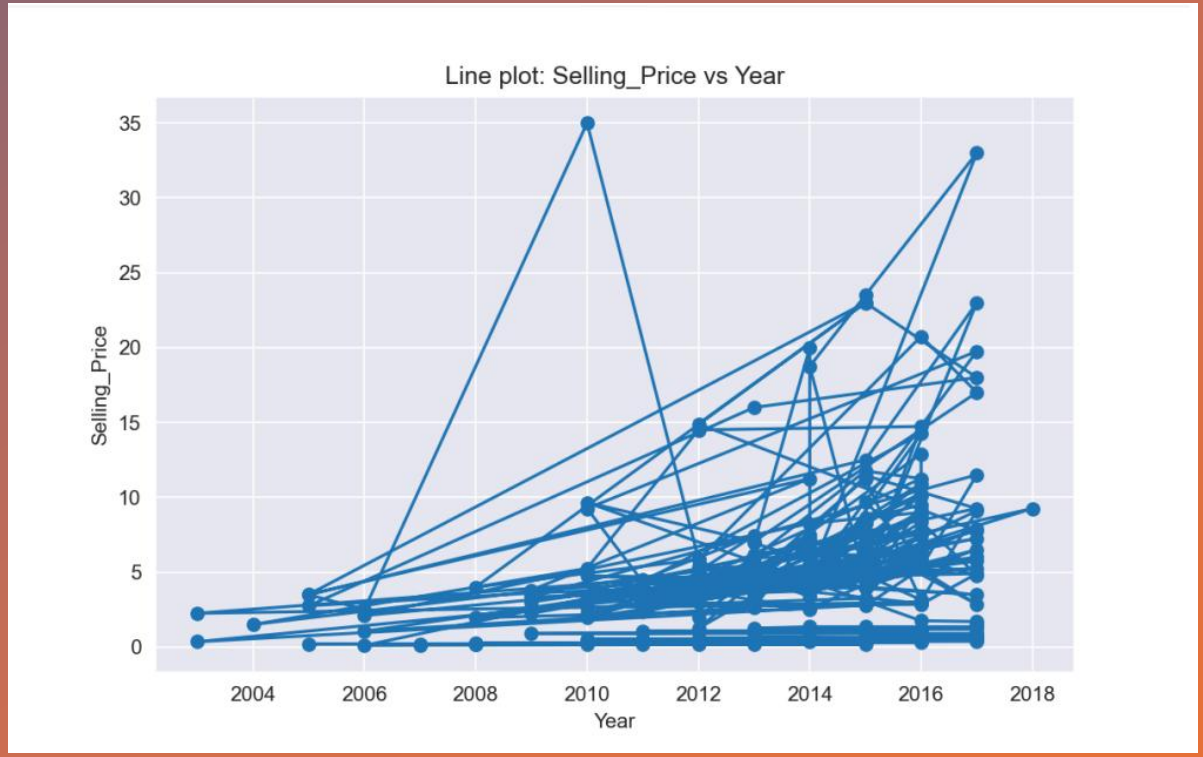
- **Environment Setup**
  - Python 3.12 or later
  - Required libraries: pandas, numpy, matplotlib, seaborn
  - Install using pip install pandas numpy matplotlib seaborn
- **How to Run**

  1. Open the project folder in VS Code or terminal.
  2. Run: python main.py
  3. Program will:
     - Load input data (CSV + Pickle)
     - Generate plots and analyses
     - Export results to /Output/ folder
- **Output Files**
  - pickle_summary.csv
  - joint_counts.csv
  - joint_probability.csv
  - conditional_probability.csv
  - vector_results.csv
  - categorical_analysis.txt
- **Tip**
  - Close the pop-up plots to continue execution

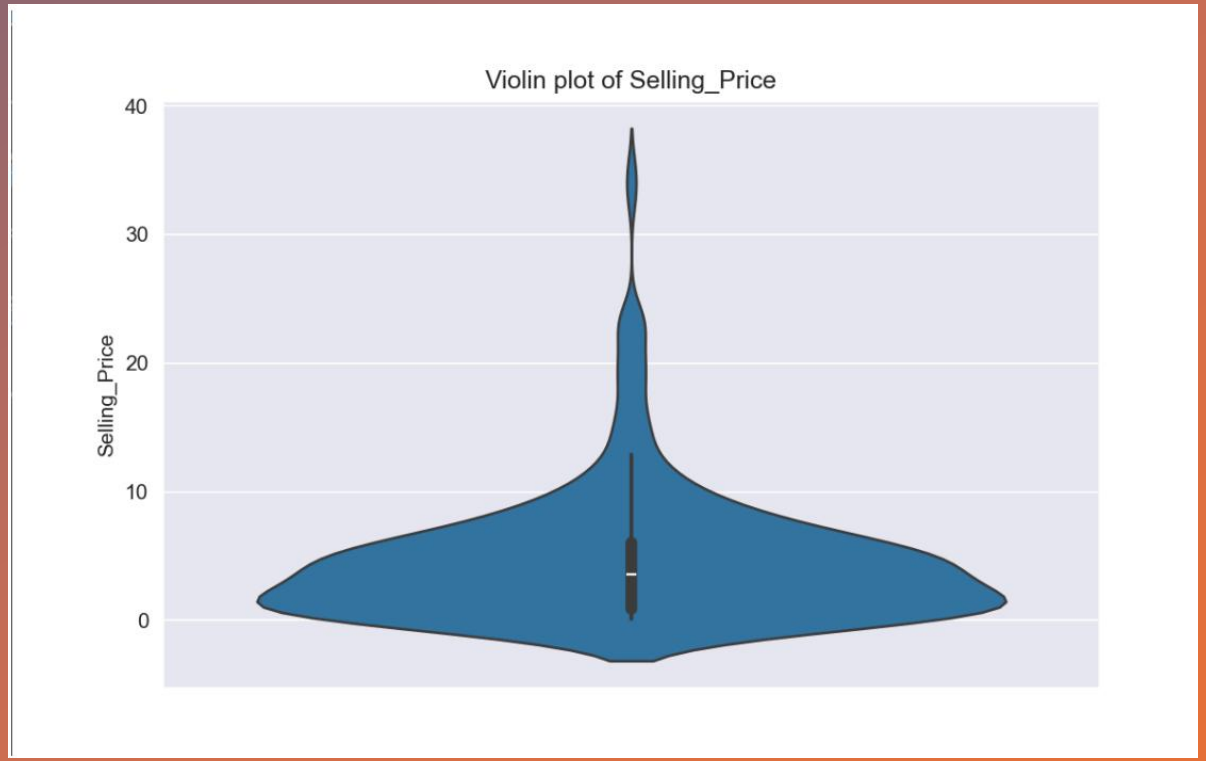
Histogram plot



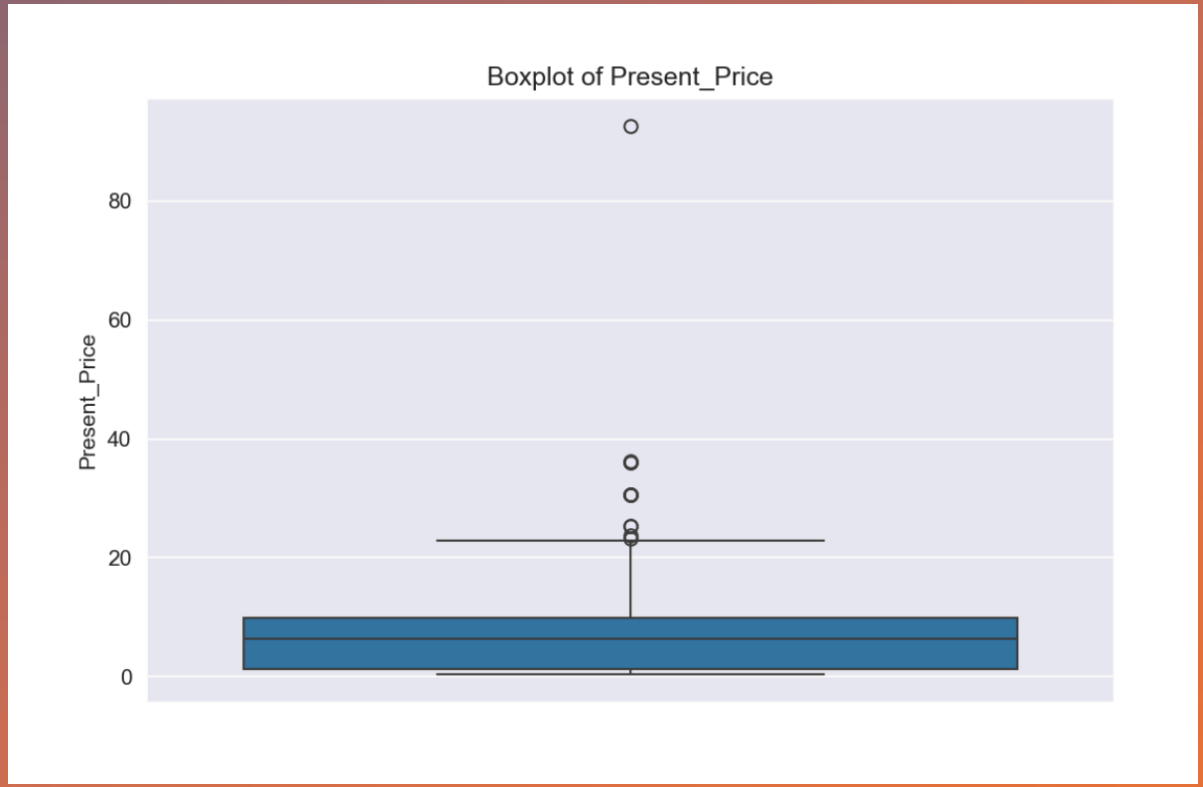
Line Plot


Line plot: Selling_Price vs Year
Selling_Price
Year
0
5
10
15
20
25
30
35
2004
2006
2008
2010
2012
2014
2016
2018

Violin plot



Box plot



Scatter plot