

# Assignment 1: EDA Documentation

Assignment 1: Exploratory Data Analysis on Laptop Prices

Generated documentation for the provided script 'assignment1\_eda.py'.

Author: (not specified)

# Dataset Info & Missing Values

Dataset information (df.info() output):

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1303 entries, 0 to 1302
Data columns (total 12 columns):
#   Column              Non-Null Count  Dtype
---  -
0   Unnamed: 0          1303 non-null   int64
1   Company             1303 non-null   object
2   TypeName            1303 non-null   object
3   Inches              1303 non-null   float64
4   ScreenResolution    1303 non-null   object
5   Cpu                 1303 non-null   object
6   Ram                 1303 non-null   object
7   Memory              1303 non-null   object
8   Gpu                 1303 non-null   object
9   OpSys               1303 non-null   object
10  Weight              1303 non-null   object
11  Price               1303 non-null   float64
dtypes: float64(2), int64(1), object(9)
memory usage: 122.3+ KB
```

Missing values summary (count and percent):

	missing_count	missing_percent
Unnamed: 0	0	0.0
Company	0	0.0
TypeName	0	0.0
Inches	0	0.0
ScreenResolution	0	0.0
Cpu	0	0.0
Ram	0	0.0
Memory	0	0.0
Gpu	0	0.0
OpSys	0	0.0
Weight	0	0.0
Price	0	0.0

# Dataset Preview (first 20 rows)

Cpu	Unnamed: 0	Ram	Company	Type	Name	Inches	Gpu	Screen	Resolution
					Memory			OpSys	Weight
0		0	Apple	Ultrabook	13.3		IPS Panel Retina Display	2560x1600	macOS
Core i5 2.3GHz		8GB		128GB SSD		Intel Iris Plus Graphics 640			
71378.6832									
1		1	Apple	Ultrabook	13.3				1440x900
Core i5 1.8GHz		8GB		128GB Flash Storage		Intel HD Graphics 6000			macOS
47895.5232									
2		2	HP	Notebook	15.6			Full HD	1920x1080
7200U 2.5GHz		8GB		256GB SSD		Intel HD Graphics 620			No OS
30636.0000									1
3		3	Apple	Ultrabook	15.4		IPS Panel Retina Display	2880x1800	macOS
Core i7 2.7GHz		16GB		512GB SSD		AMD Radeon Pro 455			
135195.3360									
4		4	Apple	Ultrabook	13.3		IPS Panel Retina Display	2560x1600	macOS
Core i5 3.1GHz		8GB		256GB SSD		Intel Iris Plus Graphics 650			
96095.8080									
5		5	Acer	Notebook	15.6				1366x768
A9-Series 9420		3GHz		4GB	500GB HDD			AMD Radeon R5	Windows
21312.0000									
6		6	Apple	Ultrabook	15.4		IPS Panel Retina Display	2880x1800	Mac OS X
Core i7 2.2GHz		16GB		256GB Flash Storage		Intel Iris Pro Graphics			
114017.6016									
7		7	Apple	Ultrabook	13.3				1440x900
Core i5 1.8GHz		8GB		256GB Flash Storage		Intel HD Graphics 6000			macOS
61735.5360									
8		8	Asus	Ultrabook	14.0			Full HD	1920x1080
8550U 1.8GHz		16GB		512GB SSD		Nvidia GeForce MX150		Windows 10	
79653.6000									
9		9	Acer	Ultrabook	14.0		IPS Panel Full HD	1920x1080	
8250U 1.6GHz		8GB		256GB SSD		Intel UHD Graphics 620		Windows 10	
41025.6000									
10		10	HP	Notebook	15.6				1366x768
7200U 2.5GHz		4GB		500GB HDD		Intel HD Graphics 620			No OS
20986.9920									1
11		11	HP	Notebook	15.6			Full HD	1920x1080
i3 6006U 2GHz		4GB		500GB HDD		Intel HD Graphics 520			No OS
18381.0672									
12		12	Apple	Ultrabook	15.4		IPS Panel Retina Display	2880x1800	macOS
Core i7 2.8GHz		16GB		256GB SSD		AMD Radeon Pro 555			
130001.6016									
13		13	Dell	Notebook	15.6			Full HD	1920x1080
i3 6006U 2GHz		4GB		256GB SSD		AMD Radeon R5 M430		Windows 10	
26581.3920									
14		14	Apple	Ultrabook	12.0		IPS Panel Retina Display	2304x1440	macOS
M m3 1.2GHz		8GB		256GB SSD		Intel HD Graphics 615			0.1
67260.6720									
15		15	Apple	Ultrabook	13.3		IPS Panel Retina Display	2560x1600	macOS
Core i5 2.3GHz		8GB		256GB SSD		Intel Iris Plus Graphics 640			
80908.3440									
16		16	Dell	Notebook	15.6			Full HD	1920x1080
7500U 2.7GHz		8GB		256GB SSD		AMD Radeon R5 M430		Windows 10	
39693.6000									
17		17	Apple	Ultrabook	15.4		IPS Panel Retina Display	2880x1800	macOS
Core i7 2.9GHz		16GB		512GB SSD		AMD Radeon Pro 560			
152274.2400									
18		18	Lenovo	Notebook	15.6			Full HD	1920x1080
7100U 2.4GHz		8GB		1TB HDD		Nvidia GeForce 940MX			No OS
26586.7200									
19		19	Dell	Ultrabook	13.3	IPS Panel Full HD / Touchscreen		1920x1080	
8250U 1.6GHz		8GB		128GB SSD		Intel UHD Graphics 620		Windows 10	1

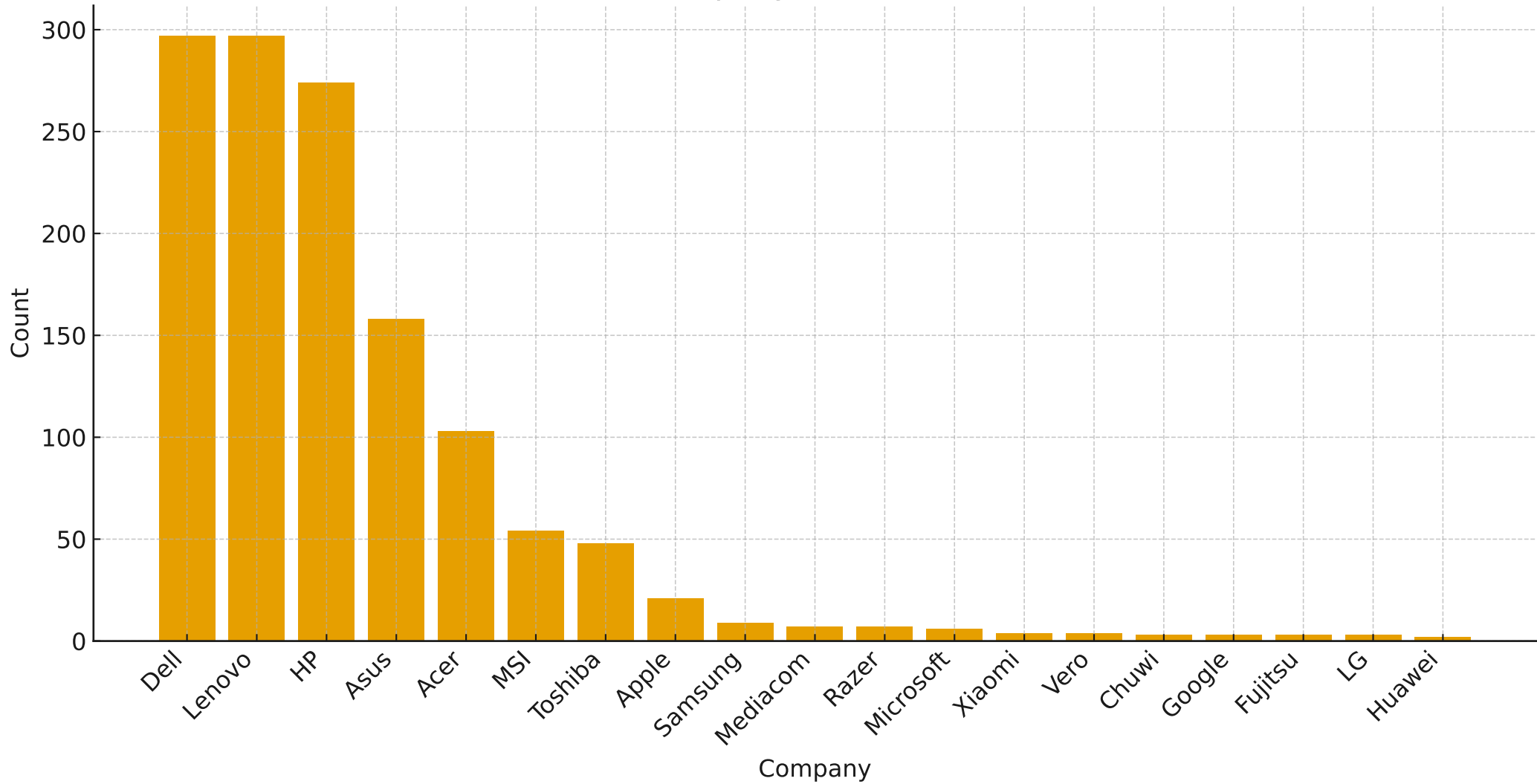
# Key Code Snippet (assignment1\_eda.py)

```
# Key parts of assignment1_eda.py (simplified)
import pandas as pd
import matplotlib.pyplot as plt

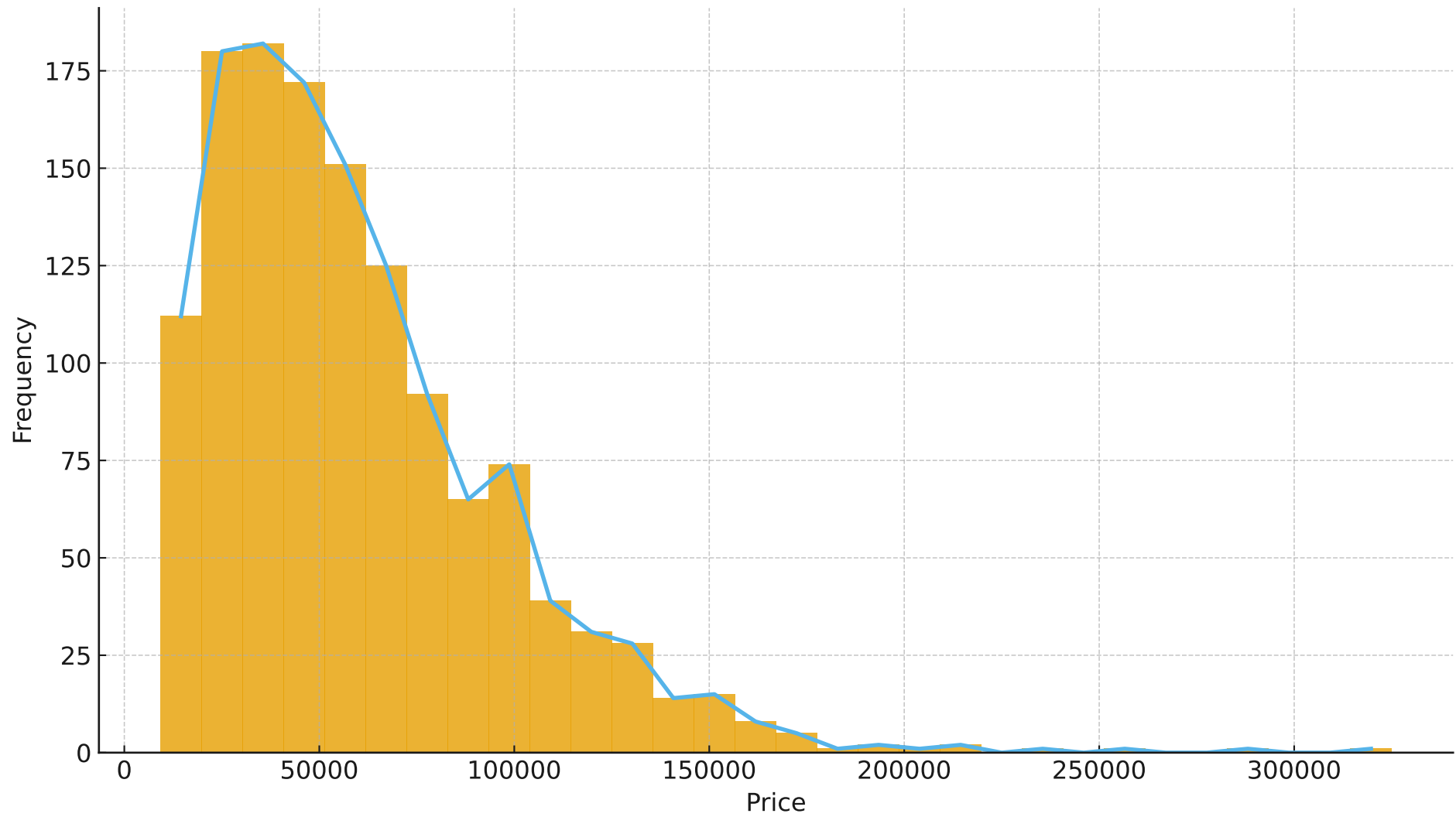
class LaptopPriceEDA:
    def __init__(self, csv_path):
        self.df = pd.read_csv(csv_path)

    def perform_eda(self):
        print(self.df.info())
        print(self.df.isnull().sum())
        # plots: Company distribution, Price distribution, RAM vs Price
```

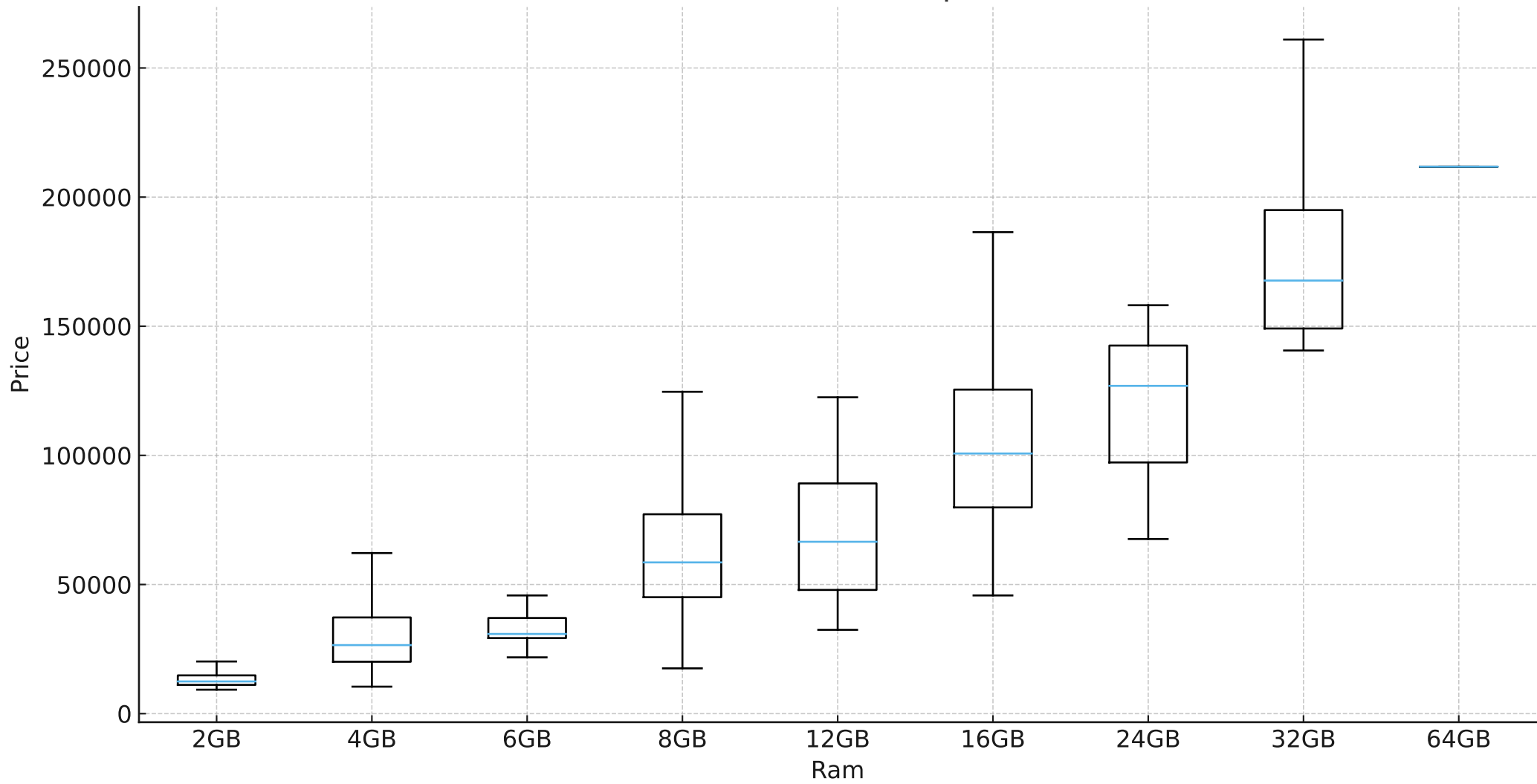
# Company Distribution



# Price Distribution



RAM vs Price (boxplot)



# Conclusion & Next Steps

Notes and suggestions:

- Consider converting string columns like 'Ram' and 'Weight' to numeric types for modeling.
- Review missing value patterns and decide on imputation or row dropping based on domain knowledge.
- Use log transformation on Price if distribution is heavily right-skewed when modeling.