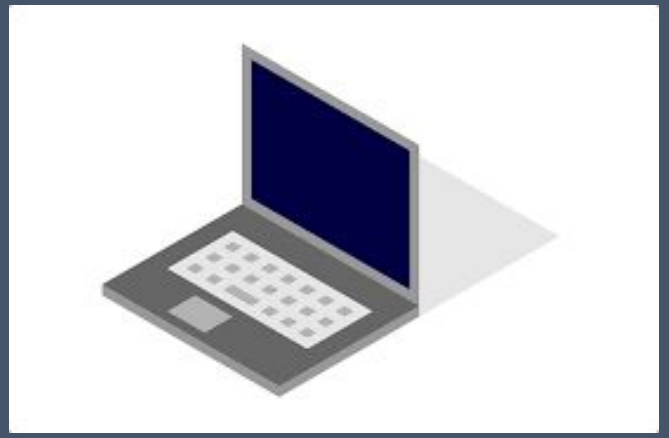
INNOMATICS RESEARCH LABS

Web Scraping
Laptop Price Analysis



Team Members :
Sai Charan
Nirosha Chennakesavula

INNOMATICS
RESEARCH LABS

CONTENTS :

Problem Statement :

- ✔ When a person want to buy a laptop under a required Budget .
- ✔ Change of prices in different e-commerce website by Brand .
- ✔ The specifications that affect the price .

Modules Used :

- ✔ Pandas
- ✔ requests
- ✔ BeautifulSoup (bs4)

Websites :

Reliance digital

Flipkart

DELL

Visualization Modules Used :

matplotlib

plotly

seaborn

INNOMATICS RESEARCH LABS

Steps Followed :

- ✔ **Step 1:** Collecting Contents from Web Pages
- ✔ **Step 2:** Extracting Data from the content
- ✔ **Step 3:** Storing the Data
- ✔ **Step 4:** Analyzing the Data

Code :

```
Brand1 = []
for i in soup.find_all("p",class_="sp__name"):
    a = i.text
    b = a.split()
    Brand1.append(b[0])
Model1 = []
for i in soup.find_all("p",class_="sp__name"):
    a = i.text
    b = a.split("(")
    Model1.append(b[0])
Processor1 = []
for i in soup.find_all("p",class_="sp__name"):
    a = i.text
    b = a.split("(")
    c = b[1].split("/")
    Processor1.append(c[0])
ram1 = []
for i in soup.find_all("p",class_="sp__name"):
    a = i.text
    b = re.findall("(\d+GB)/",a)
    if len(b)>=1:
        ram1.append(b[0])
    else:
        ram1.append(np.nan)
rom1 = []
for i in soup.find_all("p",class_="sp__name"):
    a = i.text
    b = a.split("/")[2]
    rom1.append(b)
Price = []
for i in soup.find_all("span",class_="sc-bxivhb dmBTBc"):
        a = i.text
        Price.append(a)
```

Code for scrapping Columns

Asus Vivobook 14 Laptop (10th Gen Intel i3-1005G1/8GB RAM/1TB...

₹28,999.00 ~~₹43,990.00~~ 34%(₹14,991)

OFFERS AVAILABLE

INNOMATICS RESEARCH LABS

Data Collection :

Scrapped Data / Dataset before Cleaning :

	Brand	Model	Processor	Ram	Rom	Price	Website
0	APPLE	APPLE MacBook Air M1	Apple M1 Processor	8GB	512GB SSD	₹1,06,990	FLIPCART
1	APPLE	APPLE 2020 Macbook Air M1	Apple M1 Processor	8GB	256GB SSD	₹92,900	FLIPCART
2	APPLE	APPLE 2020 Macbook Air M1	Apple M1 Processor	8GB	512GB SSD	₹1,11,490	FLIPCART
3	APPLE	APPLE 2020 Macbook Air M1	Apple M1 Processor	8GB	256GB SSD	₹92,900	FLIPCART
4	APPLE	APPLE 2022 MacBook Pro M2	Apple M2 Processor	8GB	256GB SSD	₹1,29,900	FLIPCART
...	...	...	...	...	...	...	...
208	DELL	DELL Vostro Core i5 11th Gen	Intel Core i5 Processor	8GB	1TB HDD	₹54,790	FLIPCART
209	DELL	DELL Inspiron Core i5 11th Gen	Intel Core i5 Processor	8GB	512GB SSD	₹72,490	FLIPCART
210	DELL	DELL Inspiron Core i3 10th Gen	Intel Core i3 Processor	8GB	256GB SSD	₹44,190	FLIPCART
211	DELL	DELL Inspiron Ryzen 7 Octa Core 5825U	AMD Ryzen 7 Octa Core Processor	8GB	512GB SSD	₹66,190	FLIPCART
212	DELL	DELL Inspiron Ryzen 5 Quad Core 3450U	AMD Ryzen 5 Quad Core Processor	8GB	256GB SSD	₹49,613	FLIPCART

213 rows × 7 columns

Info of data that is scraped from flipkart.

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 213 entries, 0 to 212
Data columns (total 7 columns):
 #   Column     Non-Null Count  Dtype
---  ------     --------------  -----
 0   Brand      213 non-null    object
 1   Model      213 non-null    object
 2   Processor  213 non-null    object
 3   Ram        213 non-null    object
 4   Rom        213 non-null    object
 5   Price      213 non-null    object
 6   Website    213 non-null    object
dtypes: object(7)
memory usage: 11.8+ KB
```

Dataset after Cleaning and Manipulation :

	Brand	Model	Ram	Rom	Processor	Price Category	Processor Type	Rom Type	Price (₹)	Website
0	APPLE	APPLE MacBook Air M1	8GB	512GB SSD	Apple M1 Processor	High End Laptop	Apple M1 Processor	Solid State Drive	106990	FLIPCART
1	APPLE	APPLE 2020 Macbook Air M1	8GB	256GB SSD	Apple M1 Processor	High End Laptop	Apple M1 Processor	Solid State Drive	92900	FLIPCART
2	APPLE	APPLE 2020 Macbook Air M1	8GB	512GB SSD	Apple M1 Processor	High End Laptop	Apple M1 Processor	Solid State Drive	111490	FLIPCART
3	APPLE	APPLE 2020 Macbook Air M1	8GB	256GB SSD	Apple M1 Processor	High End Laptop	Apple M1 Processor	Solid State Drive	92900	FLIPCART
4	APPLE	APPLE 2022 MacBook Pro M2	8GB	256GB SSD	Apple M2 Processor	High End Laptop	Apple M2 Processor	Solid State Drive	129900	FLIPCART
...	...	...	...	...	...	...	...	...	...	...
208	DELL	DELL Vostro Core i5 11th Gen	8GB	1TB HDD	Intel Core i5	Mid Range Laptop	Intel	Hard Disk Drive	54790	FLIPCART
209	DELL	DELL Inspiron Core i5 11th Gen	8GB	512GB SSD	Intel Core i5	High End Laptop	Intel	Solid State Drive	72490	FLIPCART
210	DELL	DELL Inspiron Core i3 10th Gen	8GB	256GB SSD	Intel Core i3	Budget Laptop	Intel	Solid State Drive	44190	FLIPCART
211	DELL	DELL Inspiron Ryzen 7 Octa Core 5825U	8GB	512GB SSD	AMD Ryzen 7	Mid Range Laptop	AMD	Solid State Drive	66190	FLIPCART
212	DELL	DELL Inspiron Ryzen 5 Quad Core 3450U	8GB	256GB SSD	AMD Ryzen 5	Budget Laptop	AMD	Solid State Drive	49613	FLIPCART

213 rows × 10 columns

Info after cleaning and manipulating the data

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 691 entries, 0 to 690
Data columns (total 13 columns):
 #   Column              Non-Null Count  Dtype
---  ------              --------------  -----
 0   Brand               691 non-null    object
 1   Model               691 non-null    object
 2   Ram                 691 non-null    object
 3   Rom                 691 non-null    object
 4   Processor           691 non-null    object
 5   Price Category      691 non-null    object
 6   Processor Type      691 non-null    object
 7   Rom Type            691 non-null    object
 8   Price  (₹)          691 non-null    int64
 9   Website             691 non-null    object
 10  Ram - b             691 non-null    int64
 11  Processor Type - b  691 non-null    int64
 12  Price Category - b  691 non-null    int64
dtypes: int64(4), object(9)
memory usage: 70.3+ KB
```

Exploratory Data Analysis

Univariate Analysis

Top 5 Available Brands :

- HP - 175
- Lenovo - 167
- Dell - 121
- Asus - 86
- Apple - 35

Bar Plot

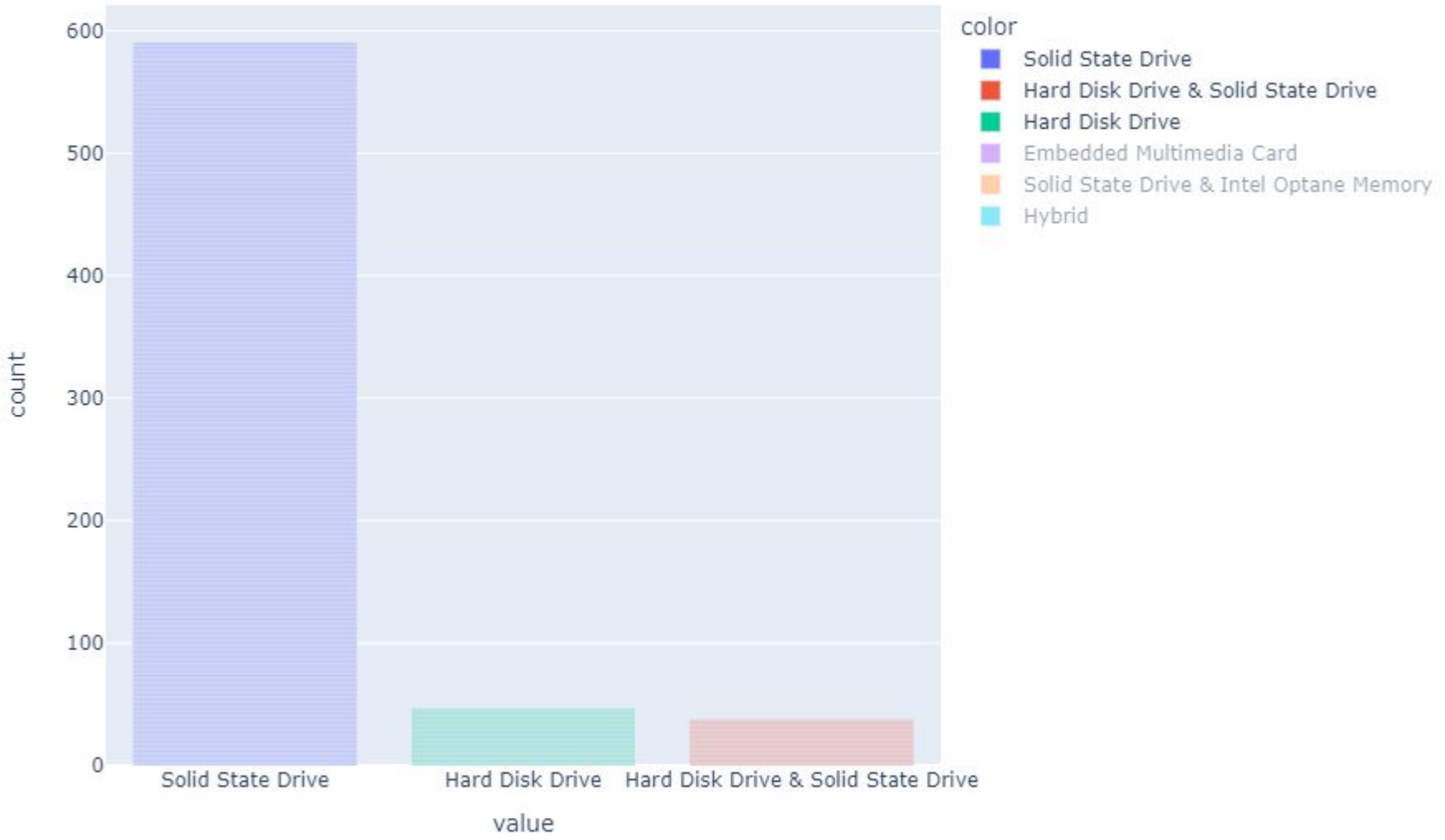
INNOMATİCS RESEARCH LABS

Bar Plot

Most Available Ram :

- 8GB - 426
- 16GB - 211
- 4GB - 50

INNOMATICS RESEARCH LABS

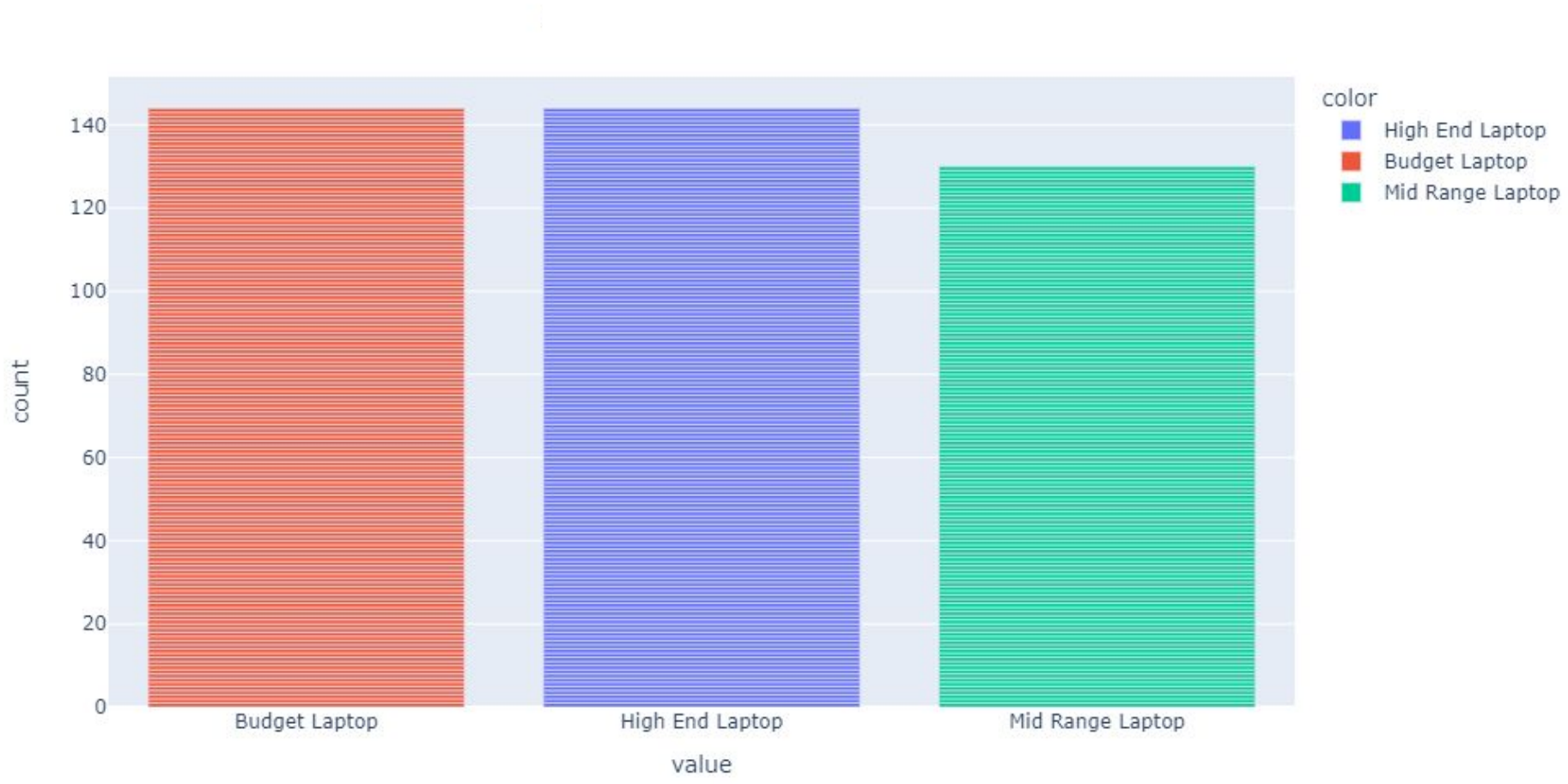


Most Available Rom type :

- SSD - 590
- HDD - 46
- HDD & SSD - 37

Bar Plot

INNOMATICS RESEARCH LABS



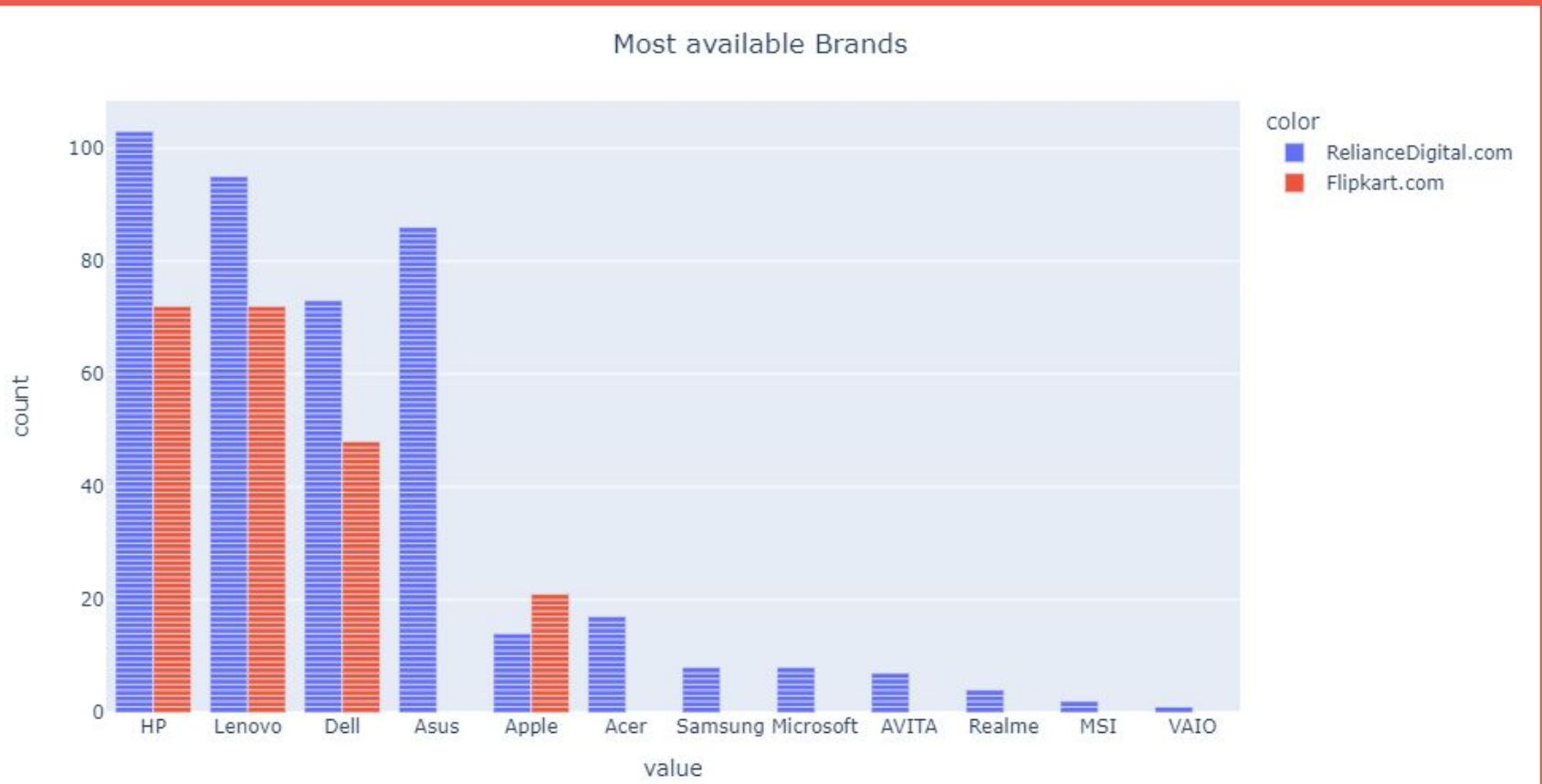
Bar Plot

Most Laptop type based on price :

- High End Laptop - 144
- Budget Laptop - 144
- Mid Range Laptop - 130

INNOMATICS RESEARCH LABS

Bivariate Analysis

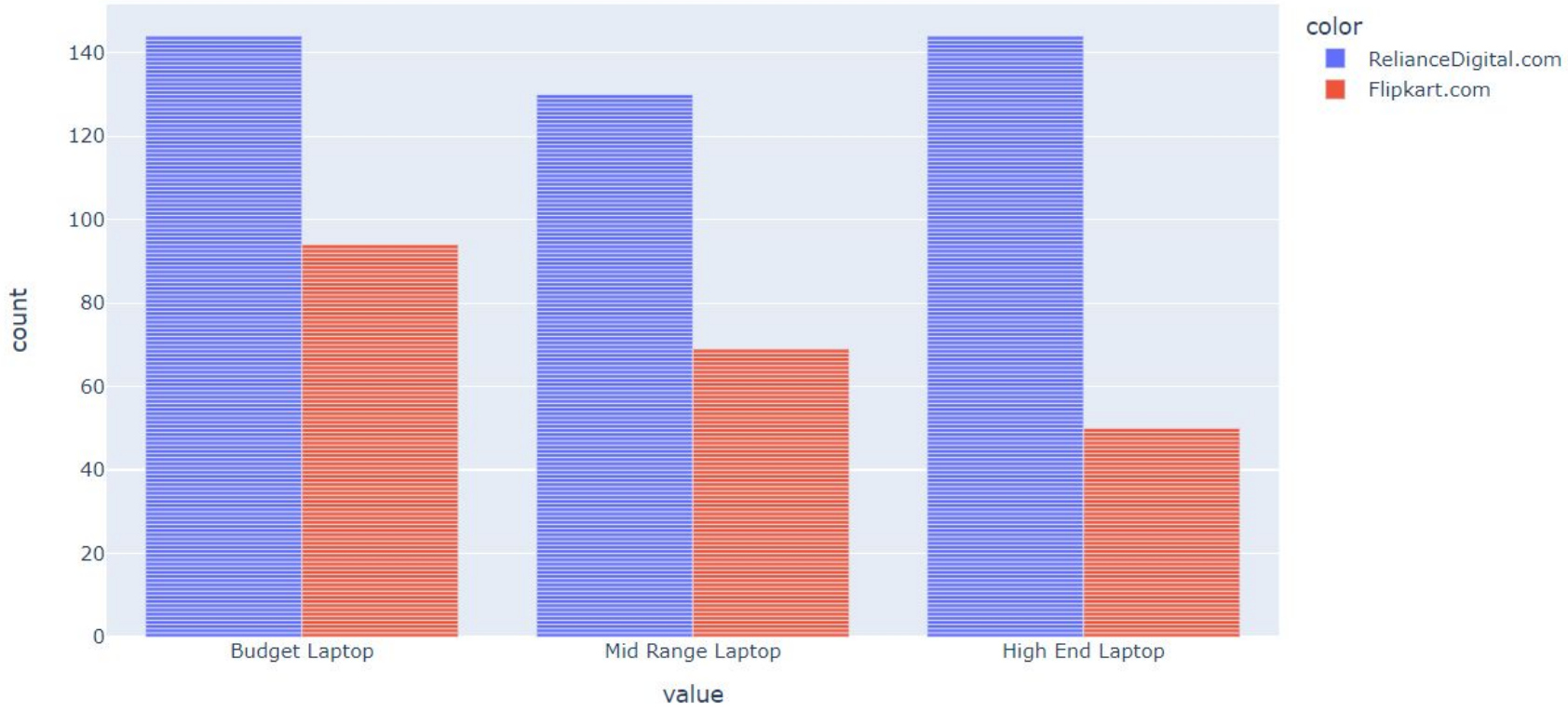


Bar Plot

Brand Count on different websites :

- **HP** RelianceDigital.com - 103 Flipkart.com - 72
- **Lenovo** RelianceDigital.com - 95 Flipkart.com - 72
- **Dell** RelianceDigital.com - 73 Flipkart.com - 48

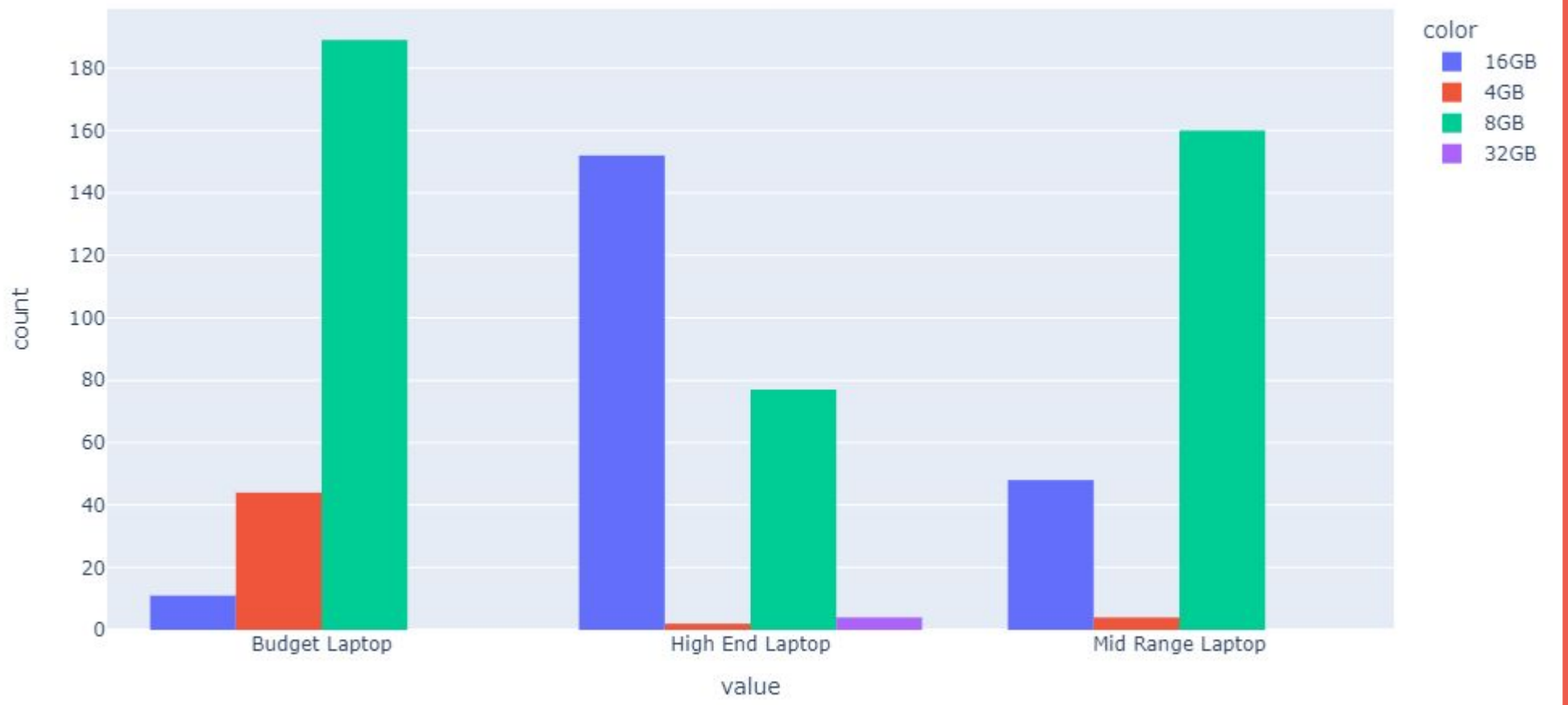
INNOMATICS RESEARCH LABS



Price category on different websites

Bar Plot

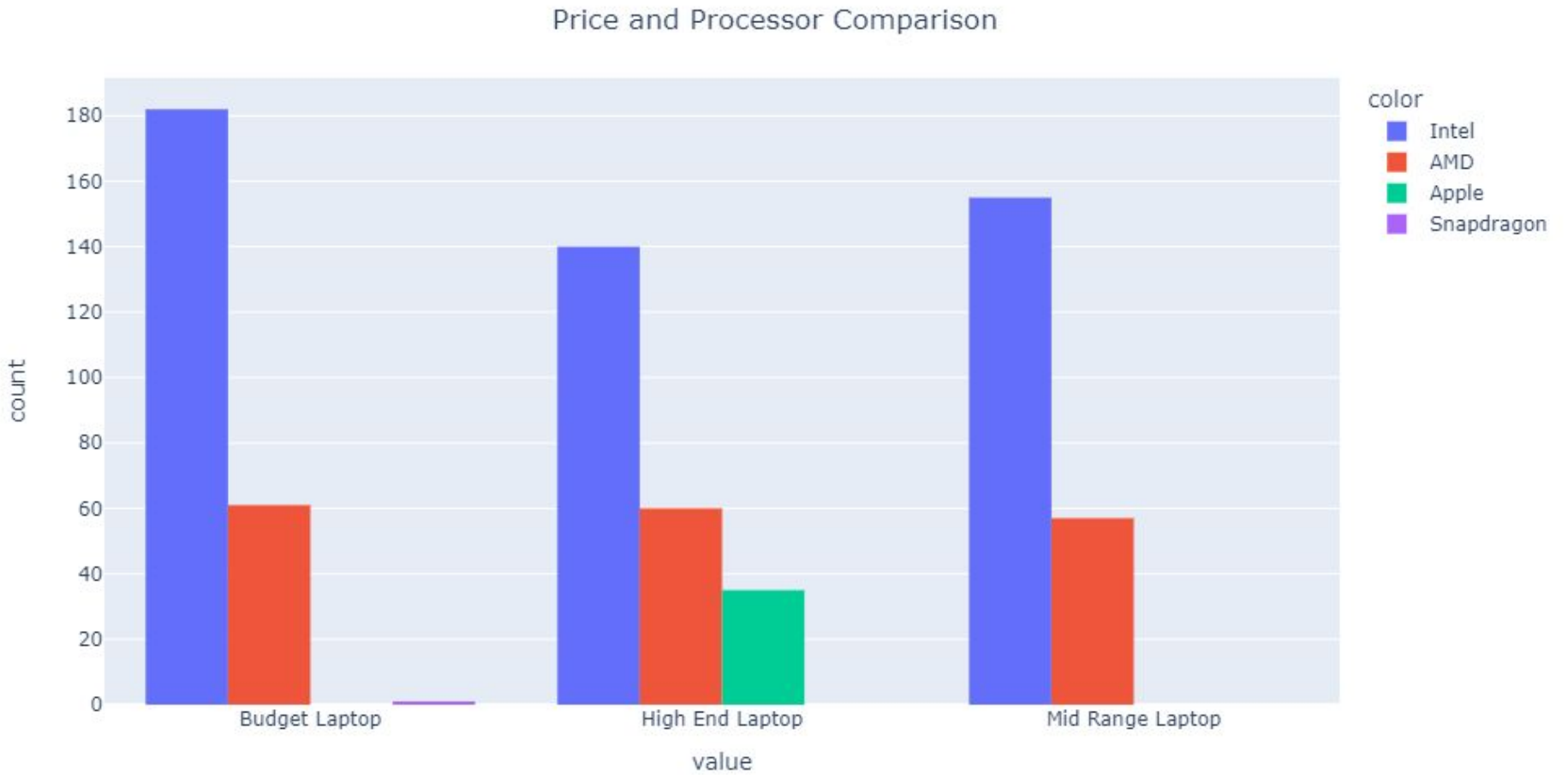
INNOMATICS RESEARCH LABS



Count based on Ram

Histogram

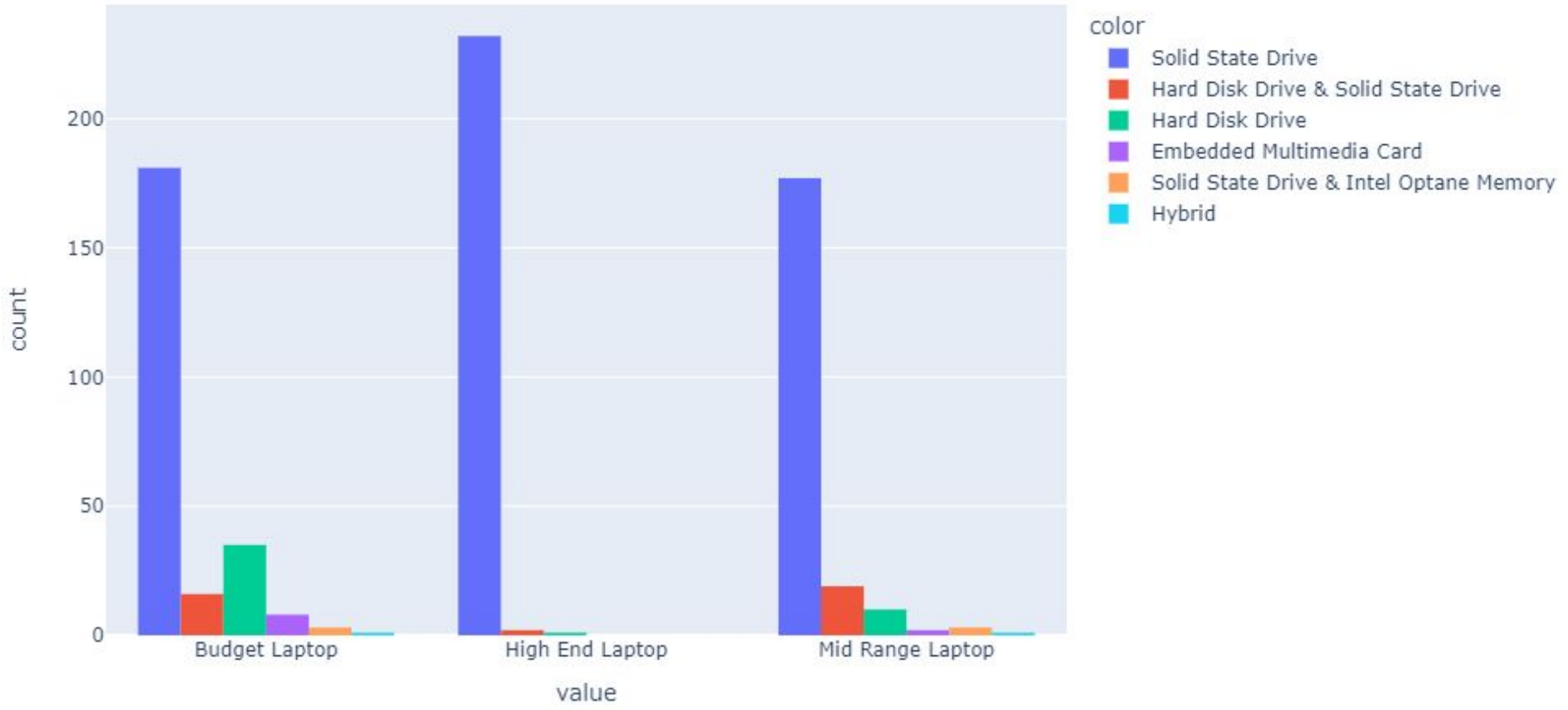
INNOMATICS RESEARCH LABS



Count based on Processor

Histogram

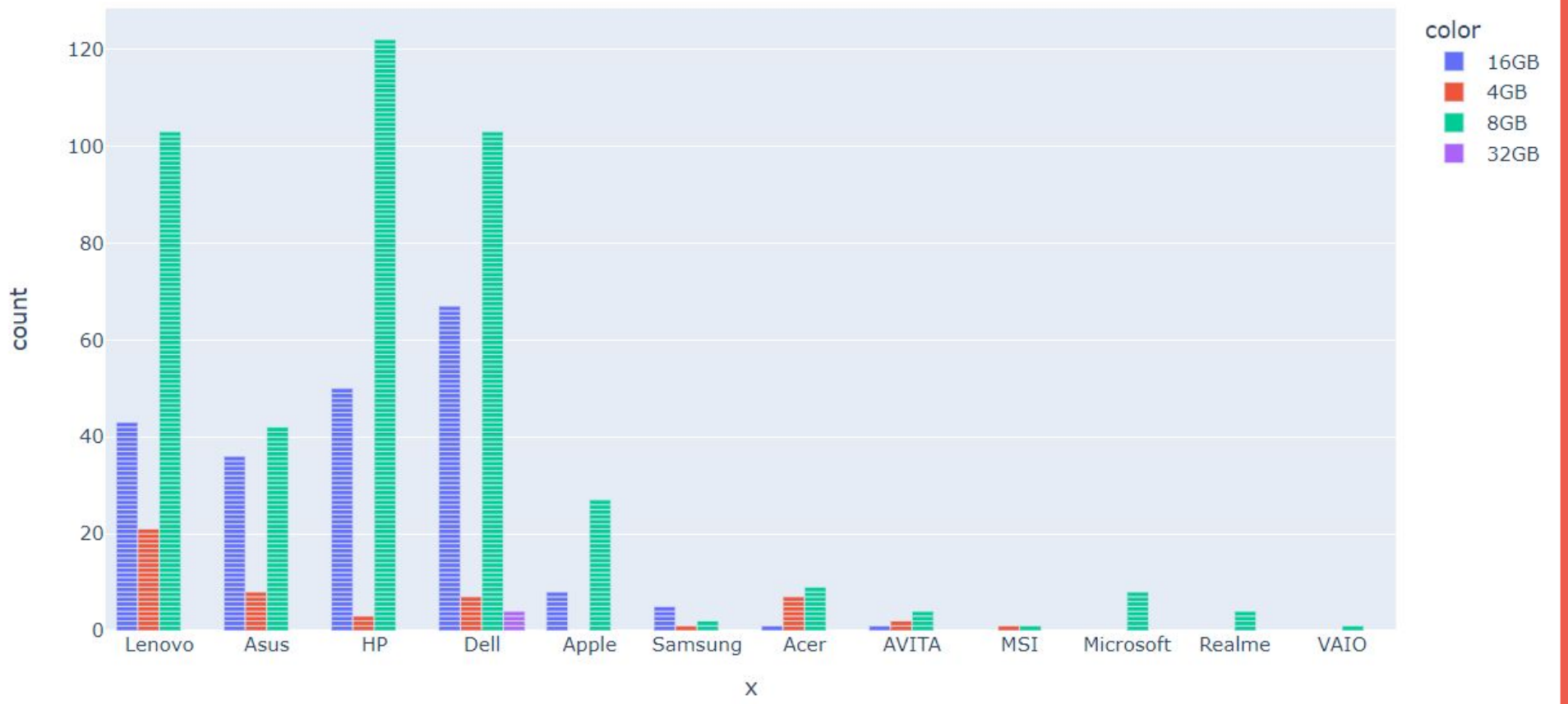
INNOMATICS RESEARCH LABS



Count Based on ROM

Histogram

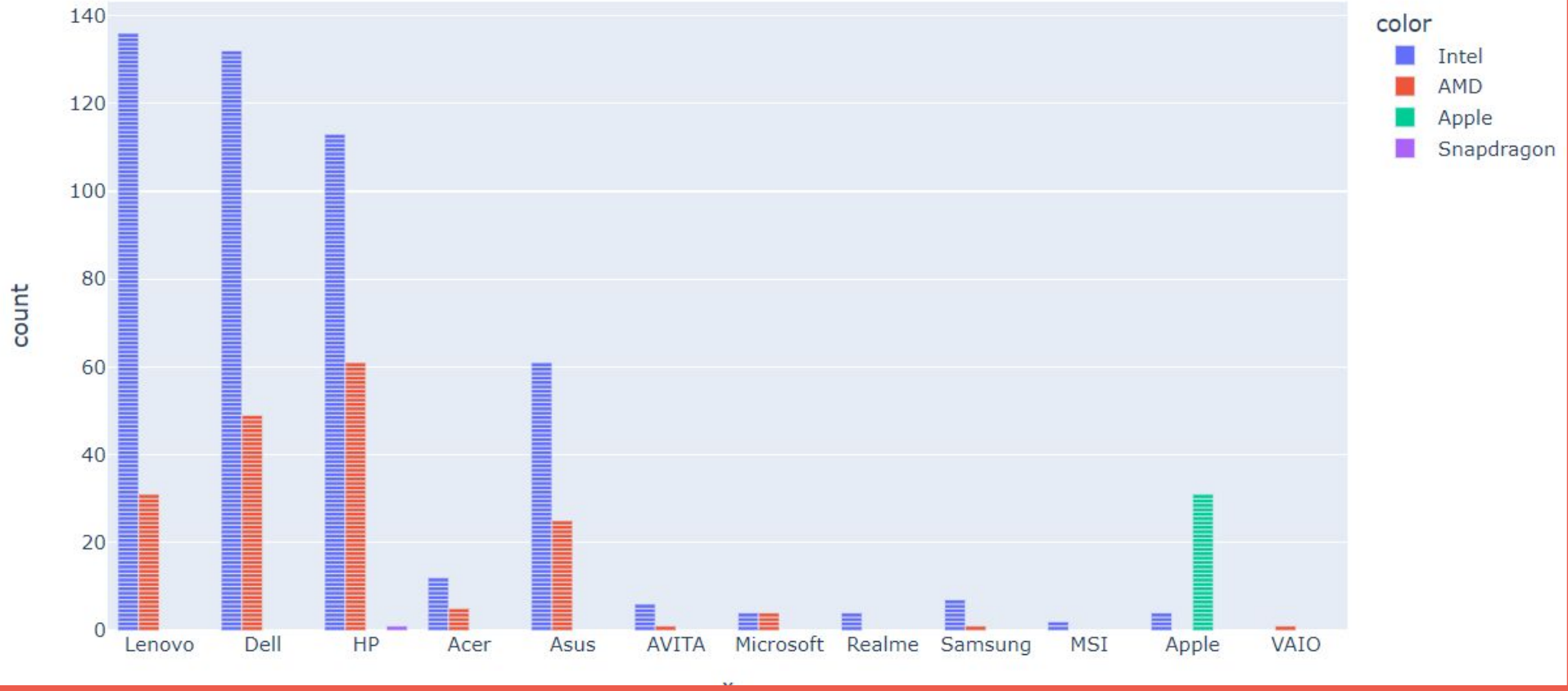
INNOMATICS RESEARCH LABS



Count Based on Ram

Bar Plot

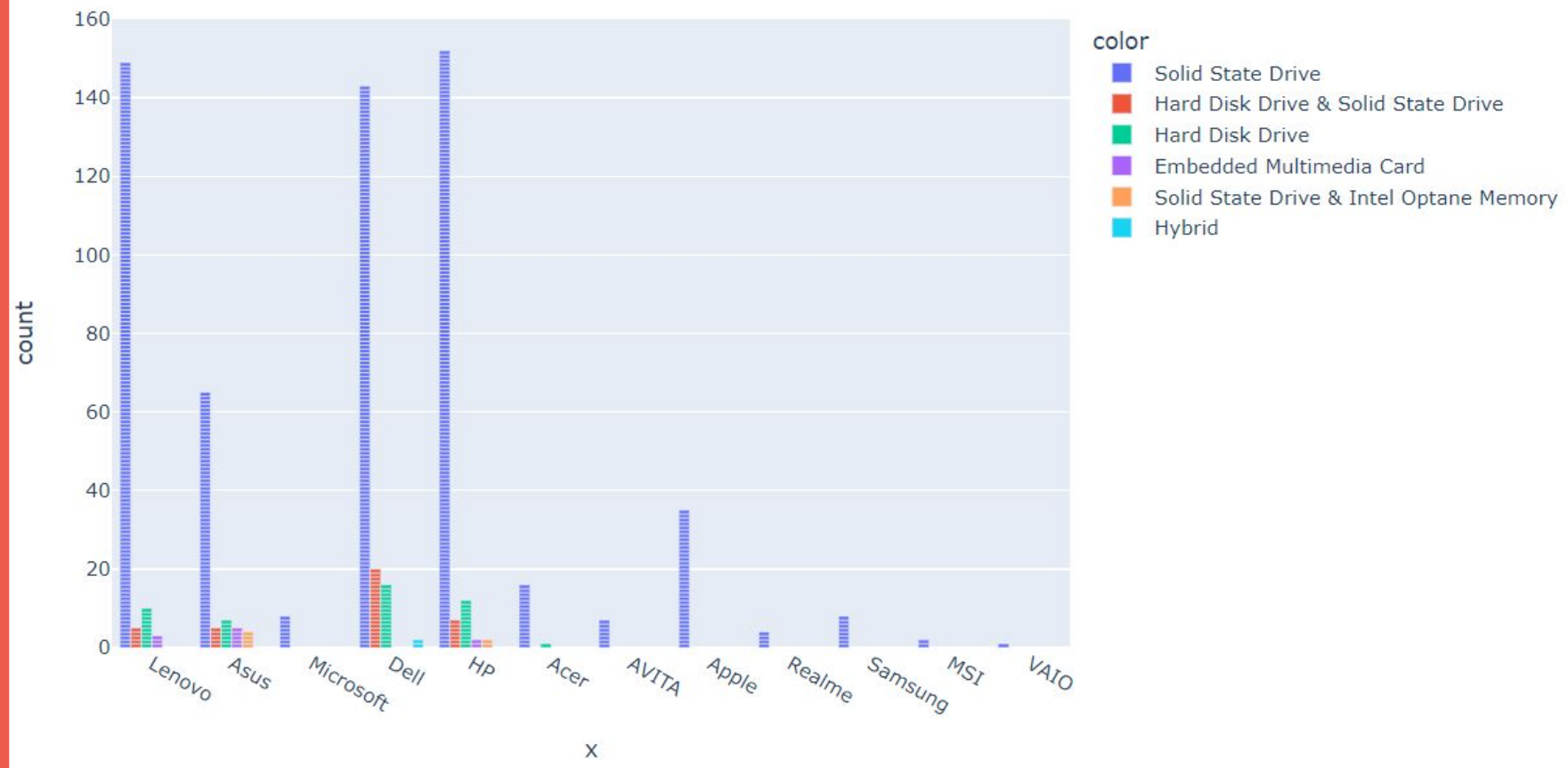
INNOMATICS RESEARCH LABS



Count Based on Processor

Bar Plot

INNOMATICS RESEARCH LABS

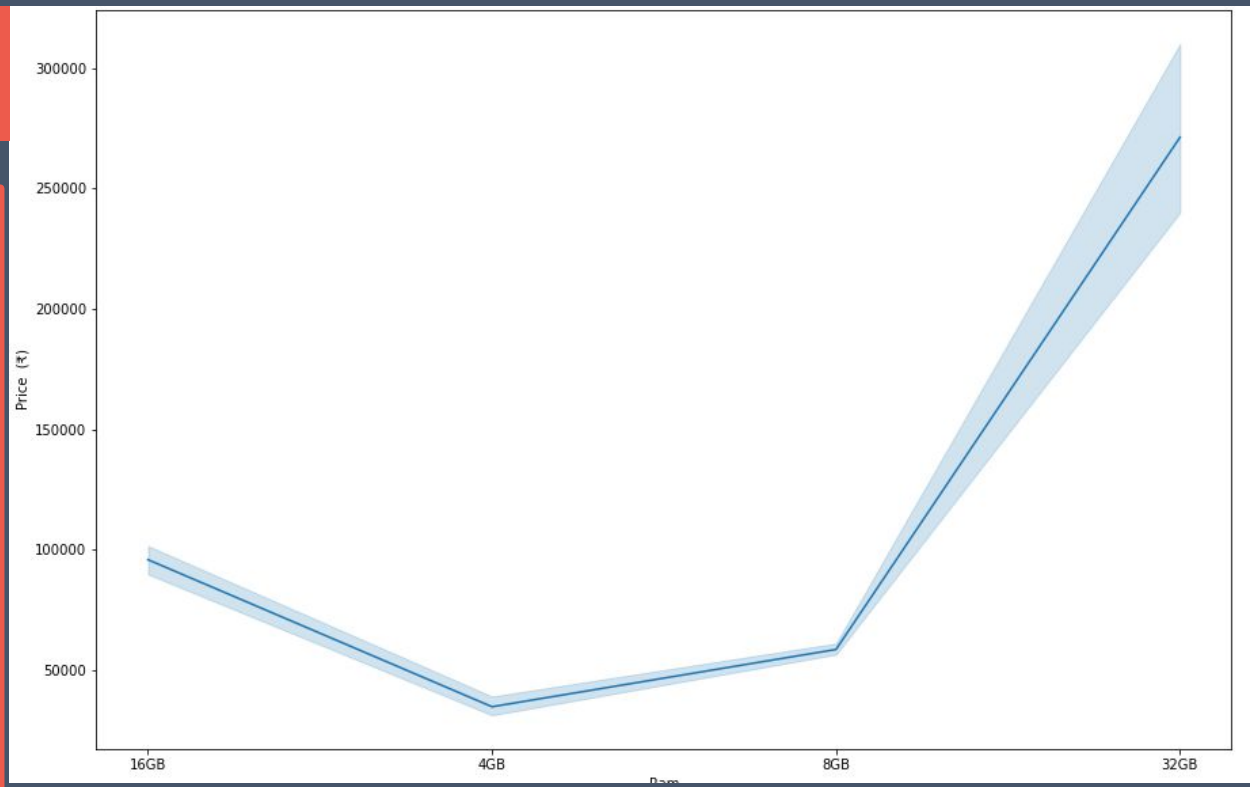


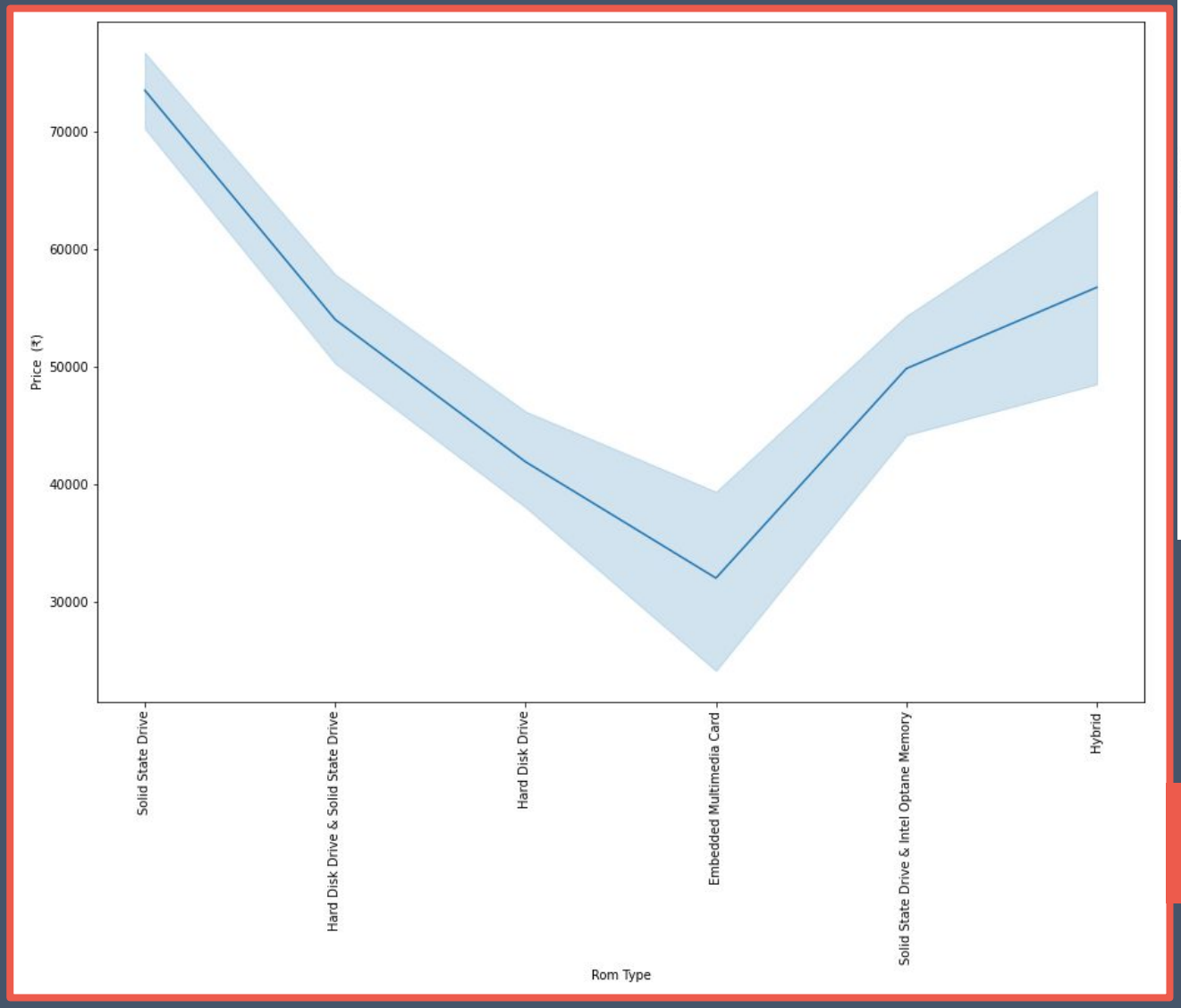
Count based on Rom type

Bar Plot

INNOMATİCS RESEARCH LABS

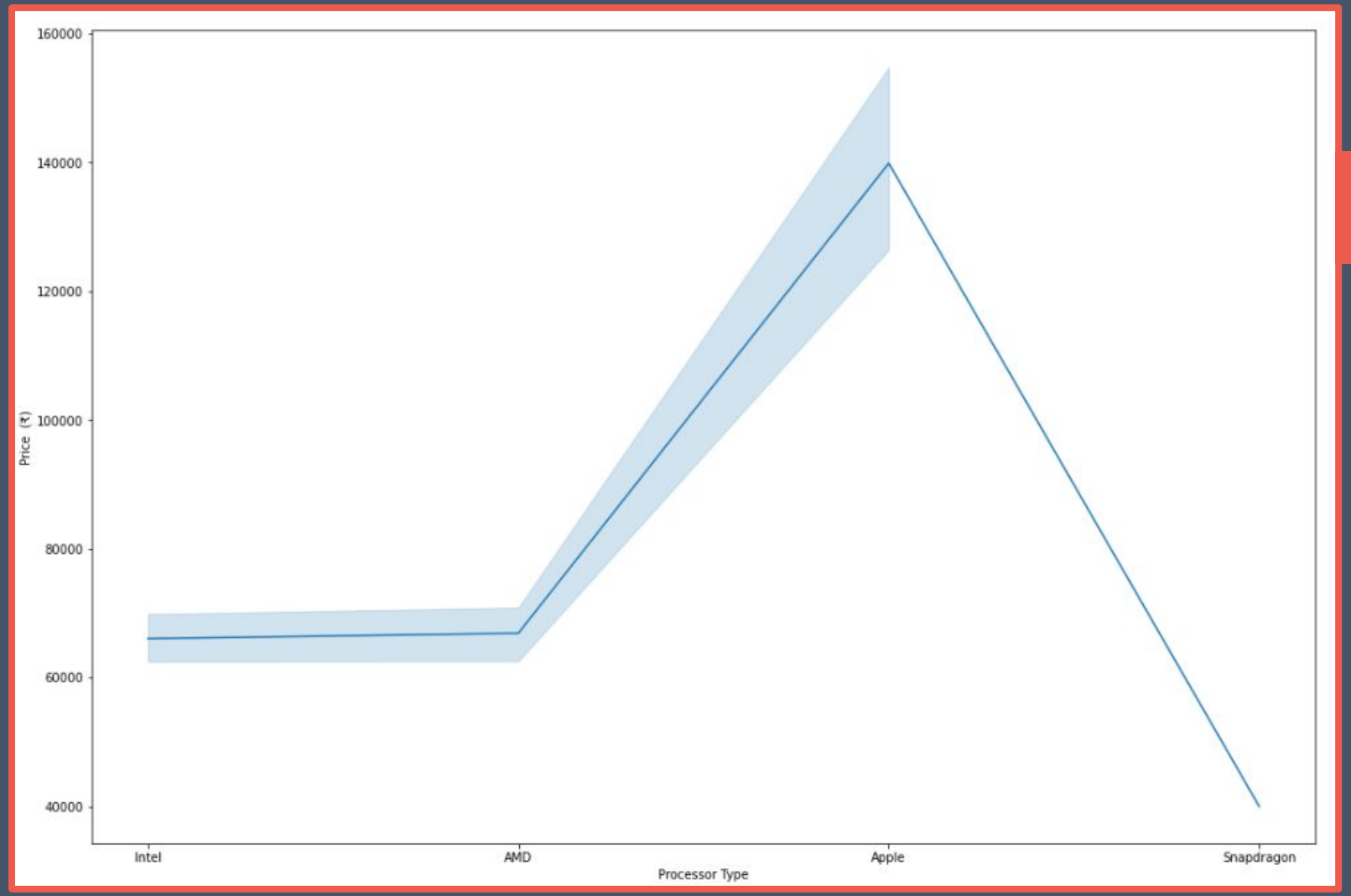
Ram





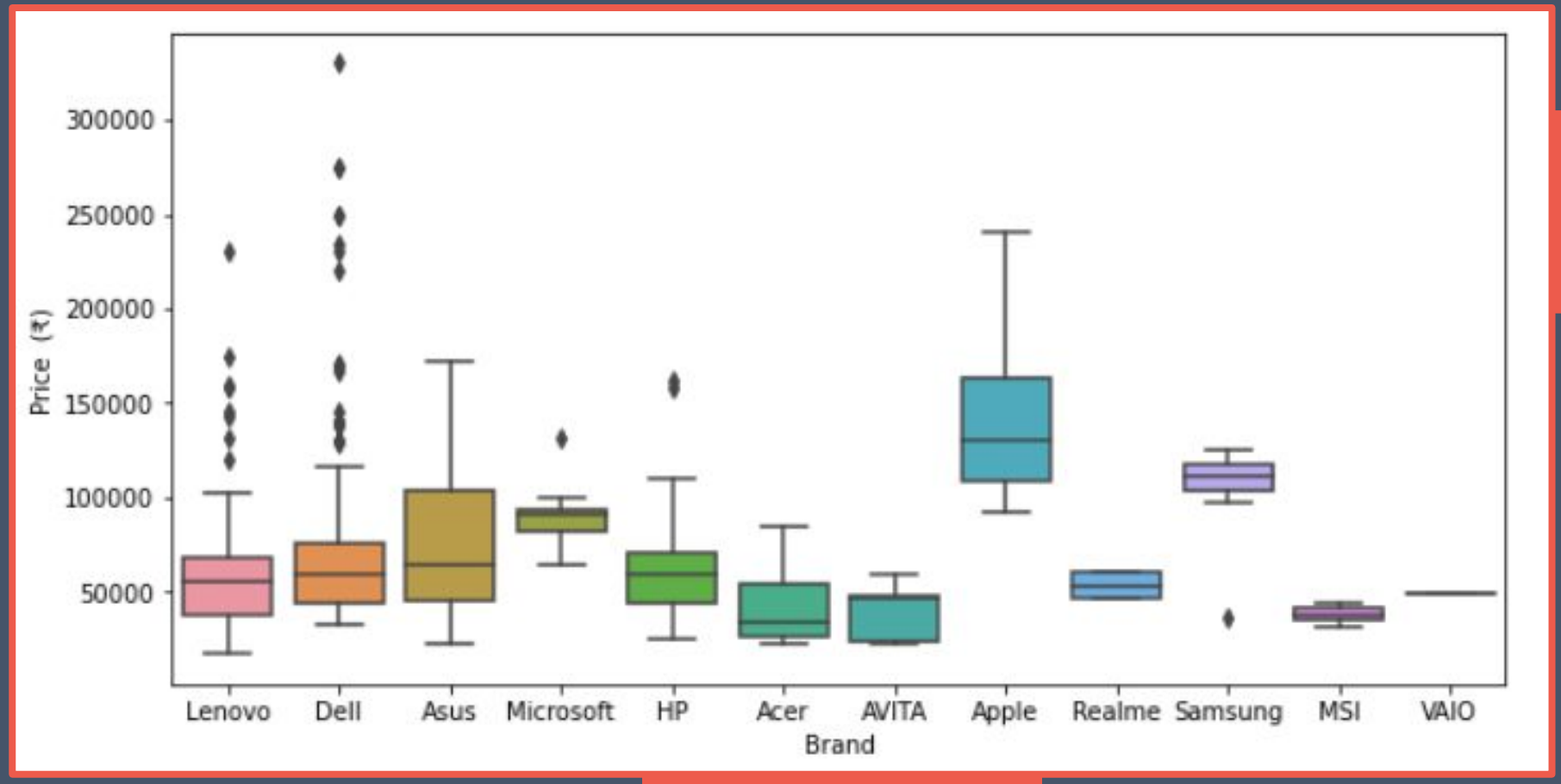
Rom

INNOMATICS RESEARCH LABS



Processor

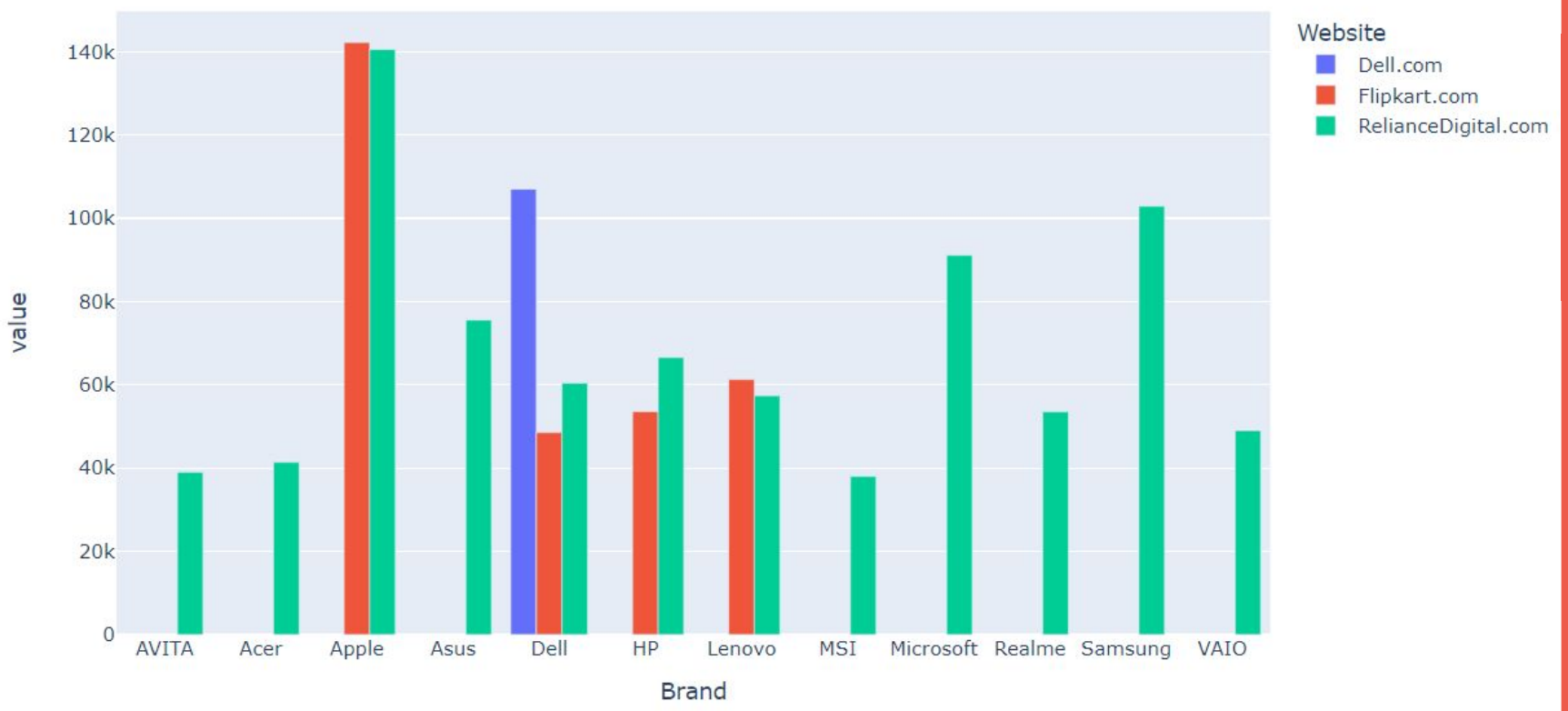
INNOMATICS RESEARCH LABS



Box plot of Brand with Price

Box Plot

Multivariate Analysis



Price of laptops in different websites

Bar Plot

Correlation of Numeric columns

HeatMap

INNOMATICS RESEARCH LABS

Conclusion :

- ✔ ***The price of a laptop is dependent on the Type of Ram, Processor Type and Storage type.***
- ✔ Top 3 Brands that are available are HP , Lenovo and Dell .
- ✔ Most available Ram is 8GB in most of the laptops .
- ✔ Most used storage type is SSD .
- ✔ There are equal number of Budget and High end laptops
- ✔ HP is the most available laptop in Flipkart and Reliance digital website .
- ✔ Budget Laptops are having 8GB the most and having SSD as storage type .

Challenges faced :

- ✔ Getting data from three websites .
- ✔ More number of unique values in a column even they are same .
- ✔ Selecting a suitable plot to visualize the data.

THANK YOU

INNOMATICS
RESEARCH LABS