SAMSUNG

Together for Tomorrow!
Enabling People
Education for Future Generations

Samsung Innovation Campus

Artificial Intelligence Course

Supervised By: Dr Doaa Mahmoud

PREDICTION OF LAPTOP PRICE

By Abdel-rahman Sabry & Moaz Atef

AGENDA

- Description of data
- Dataset
- Data visualization
- Preprocessing
- Modeling

QUESTIONS THAT OUR PRESENTATION WILL ANSWER:

- What is the range of the prices of the laptops?
- What is the prices of the laptops among companies?
- What is the types of laptops and their prices in euros?
- What is the processors types and their prices in euros?
- What is the Rams types and their prices in euros?

DESCRIPTION OF DATA

The data talks about several types of laptop computers and their manufacturers, components and prices. The following is a description ~~of each~~ column in the data:-

1 Company- String -Laptop Manufacturer
2 Product -String -Brand and Model
3 TypeName -String -Type (Notebook, Ultrabook, Gaming, etc.)
4 Inches -Numeric- Screen Size
5 ScreenResolution -String- Screen Resolution
6 CPU- String -Central Processing Unit (CPU)
7 Ram -String- Laptop RAM
8 Memory -String- Hard Disk / SSD Memory
9 GPU -String- Graphics Processing Units (GPU)
10 OpSys -String- Operating System
11 Weight -String- Laptop Weight
12 Price-euros -Numeric- Price (Euro)

DATASET

- Data has 1 column(int64), 2column (float64),10 (object)
- Shape of data **(1303, 13)**
- **Not found null data**

```
 #   Column           Non-Null Count  Dtype
---  ------           --------------  -----
 0   laptop_ID        1303 non-null   int64
 1   Company          1303 non-null   object
 2   Product          1303 non-null   object
 3   TypeName         1303 non-null   object
 4   Inches           1303 non-null   float64
 5   ScreenResolution 1303 non-null   object
 6   Cpu              1303 non-null   object
 7   Ram              1303 non-null   object
 8   Memory           1303 non-null   object
 9   Gpu              1303 non-null   object
 10  OpSys            1303 non-null   object
 11  Weight           1303 non-null   object
 12  Price_euros      1303 non-null   float64
dtypes: float64(2), int64(1), object(10)
```

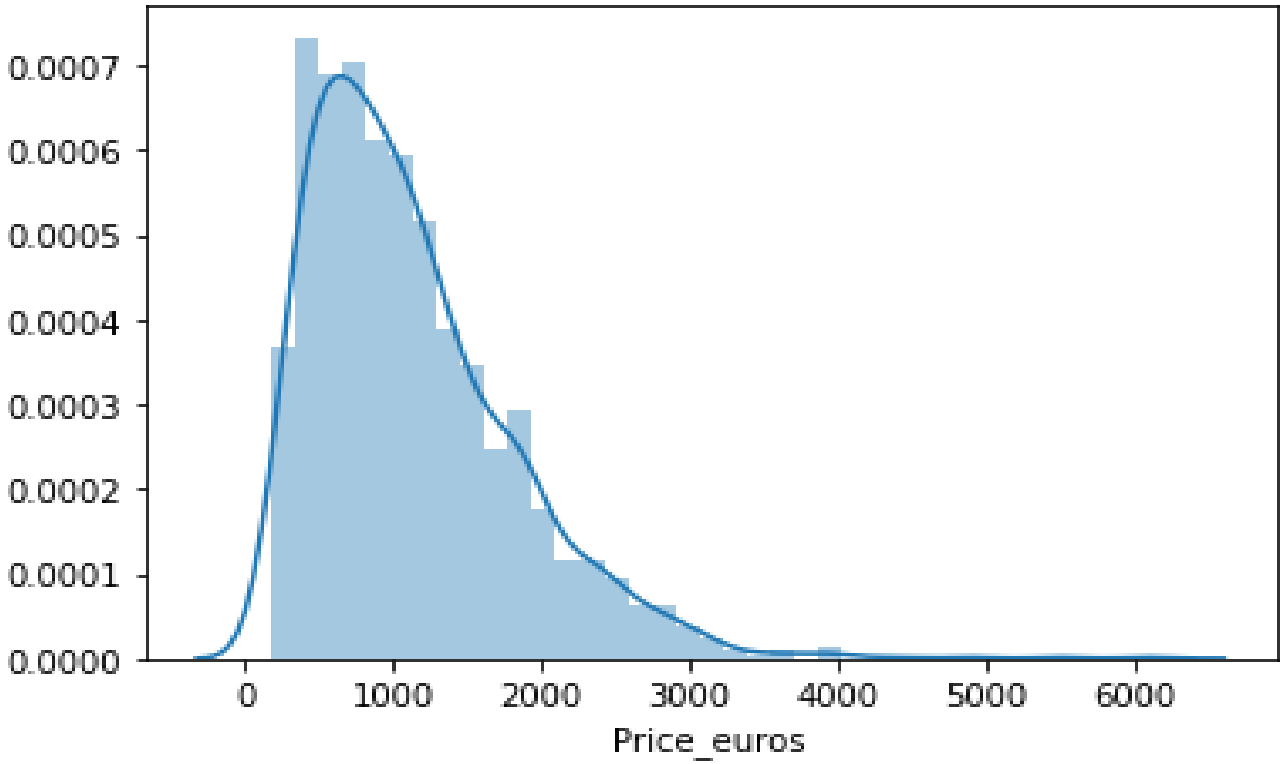
```
#Miss values
df.isnull().sum()
```

```
laptop_ID           0
Company             0
Product             0
TypeName            0
Inches              0
ScreenResolution    0
Cpu                 0
Ram                 0
Memory              0
Gpu                 0
OpSys               0
Weight              0
Price_euros         0
dtype: int64
```

DATA VISUALIZATION PRICE-EUROS

WHAT ARE THE PREVAILING PRICES?

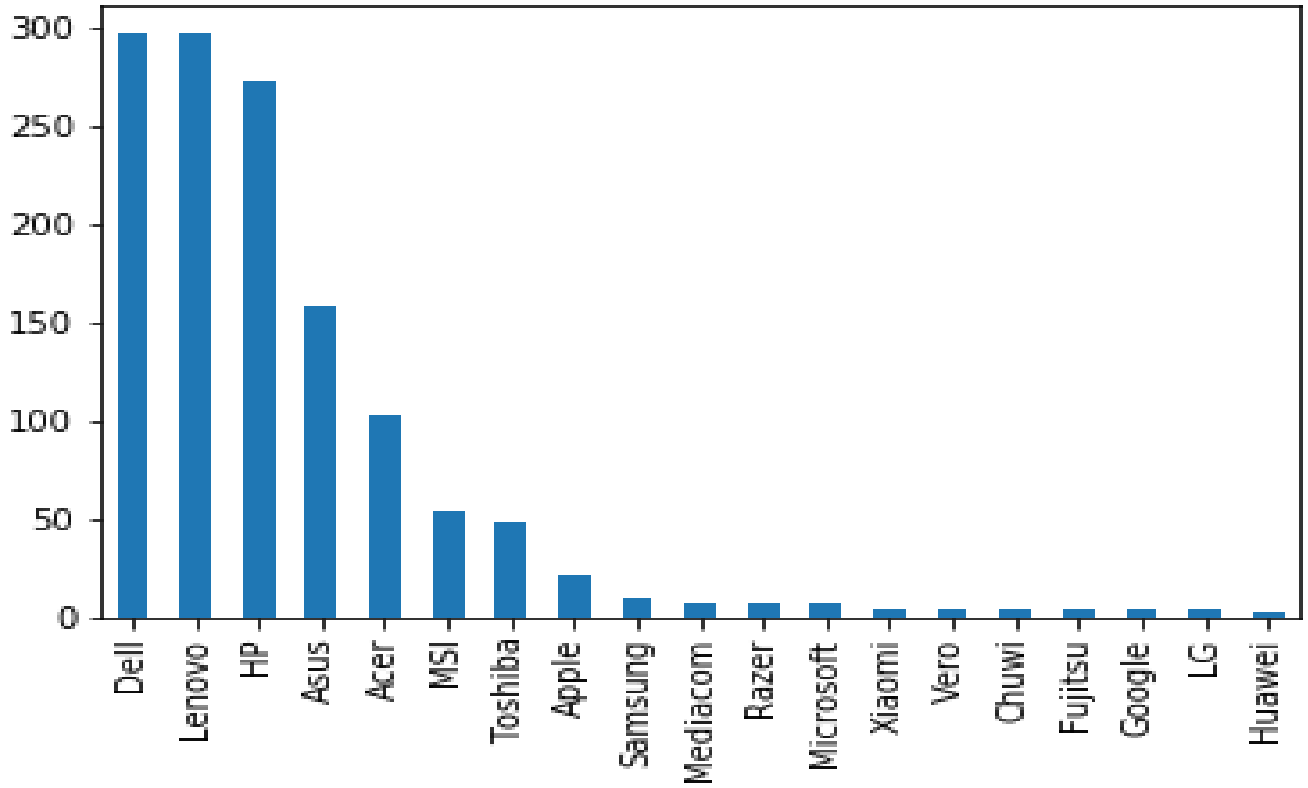
PRICES BETWEEN 500 AND 1500 ARE THE PREVAILING ONES, WHICH ARE THE LOW PRICES.



COMPANY

What are the best-selling companies in the market?

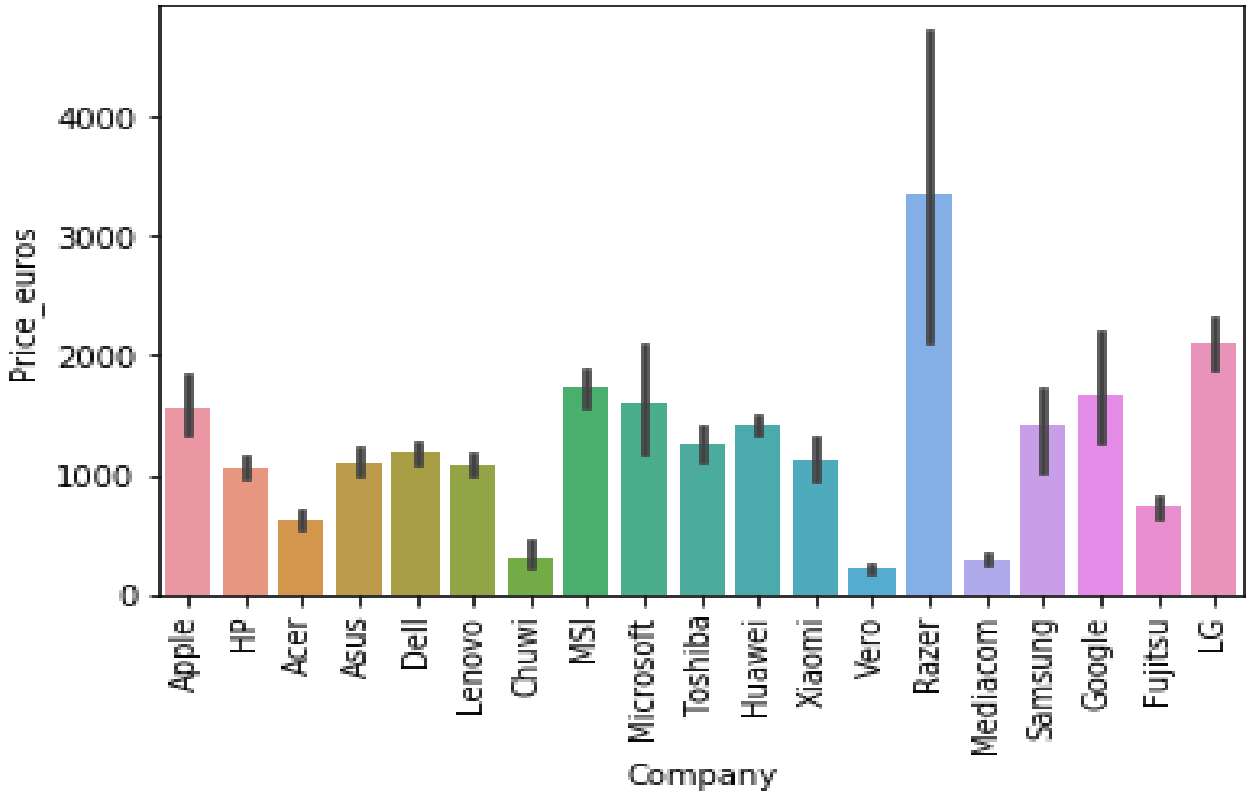
Dell, Lenovo and HP are the ones who dominate the market.



PRICE-EUROS & COMPANY

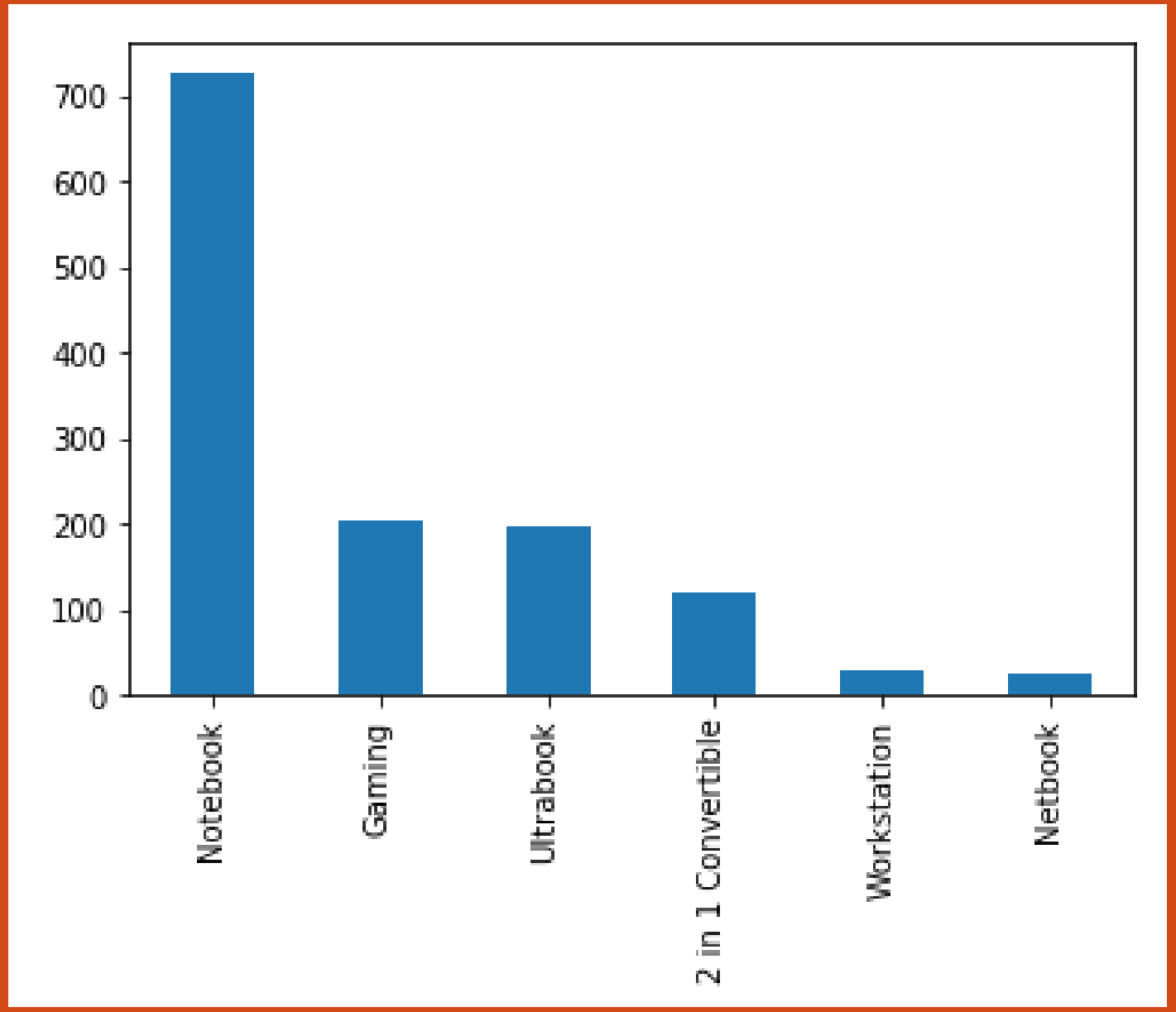
What is the company that owns the most expensive equipment?

Razer laptops are the most expensive laptops



TYPE-NAME

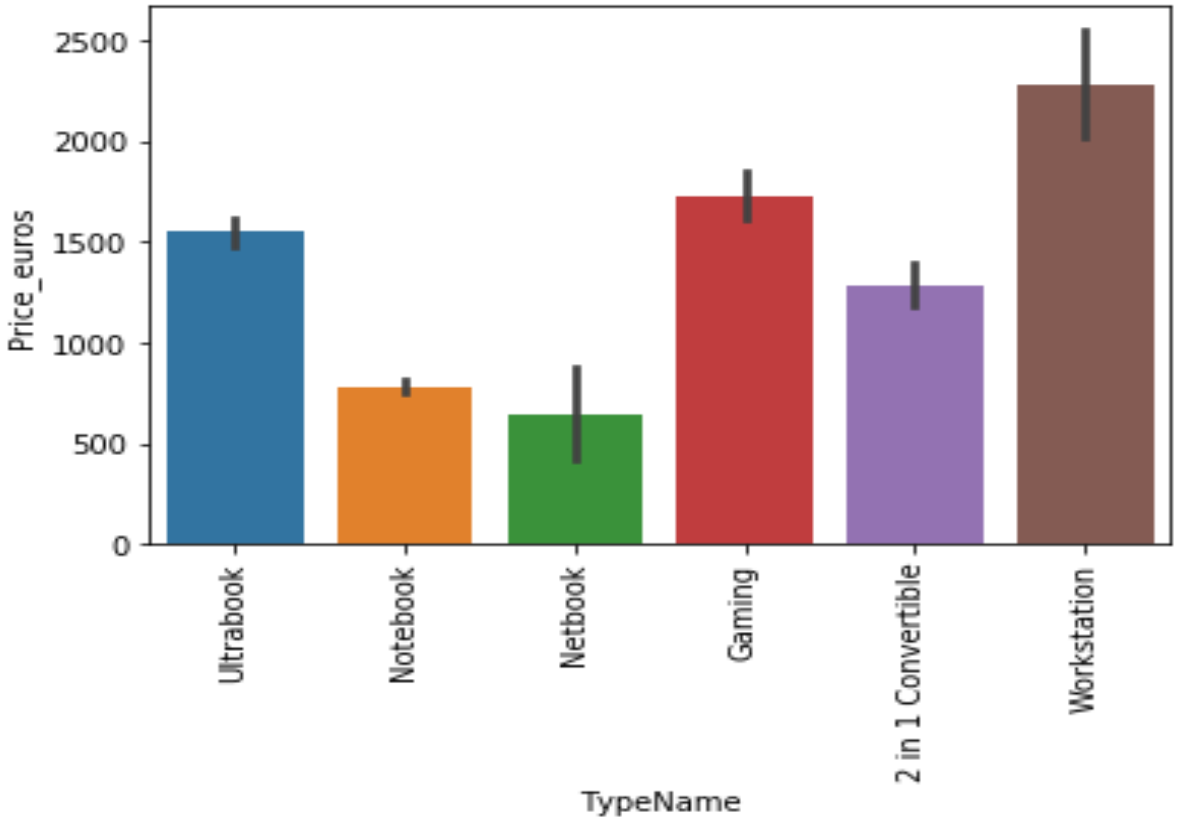
- From this figure, we conclude that students are the most buying laptops.



PRICE-EUROS & TYPE-NAME

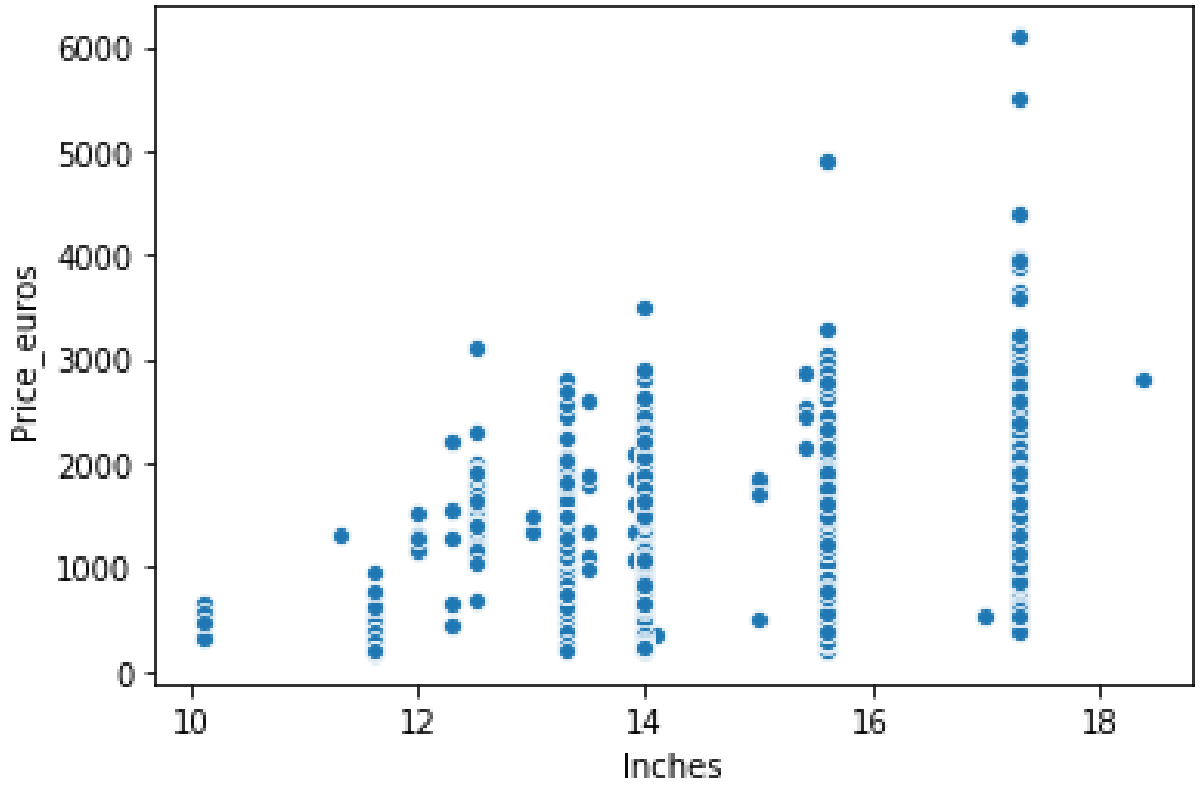
What is the Type that owns the most expensive equipment?

We notice that Workstation Type is the most expensive laptop



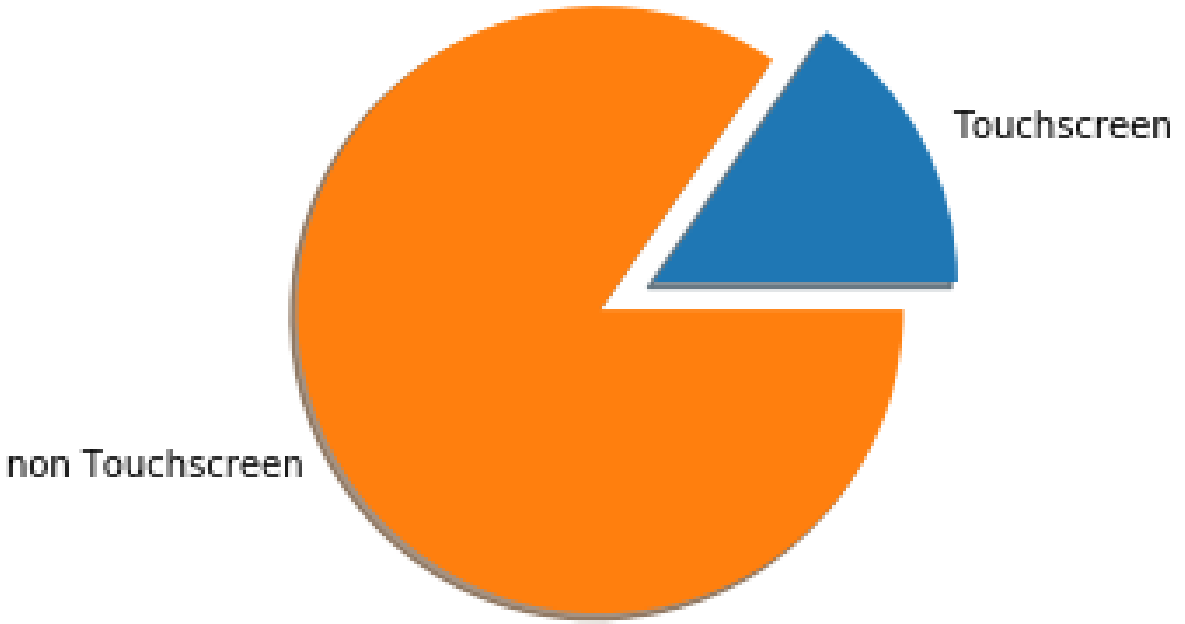
PRICE-EUROS & INCHES

There is a direct relationship between the inches and the price, as the laptop gets more bigger, it gets more expensive



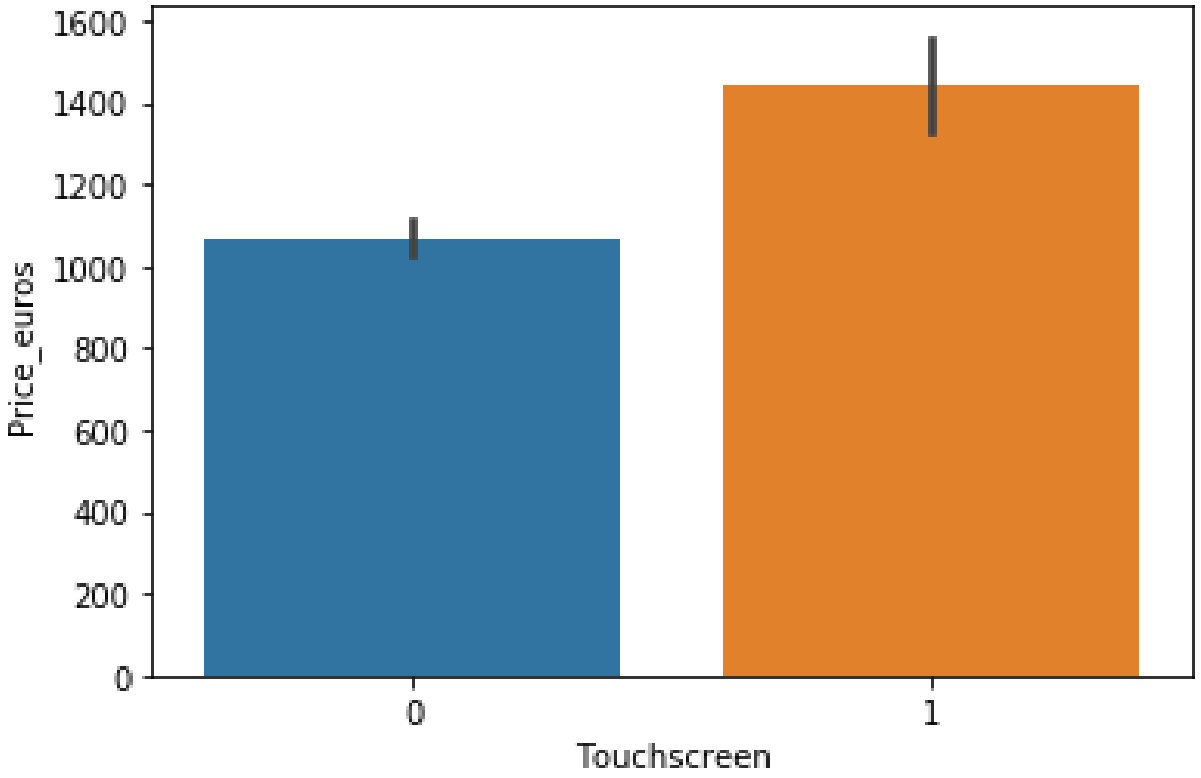
TYPE OF SCREEN

Touch screen ratio to non touch screens.



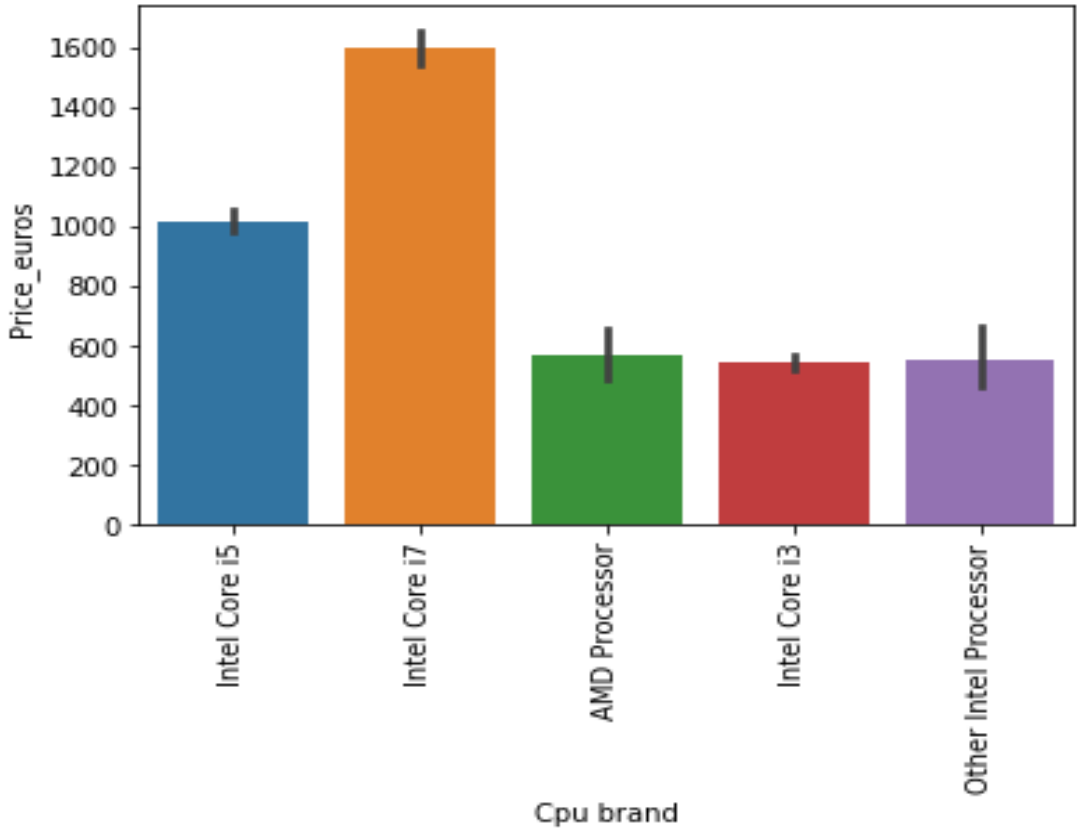
PRICE-EUROS & TOUCHSCREEN

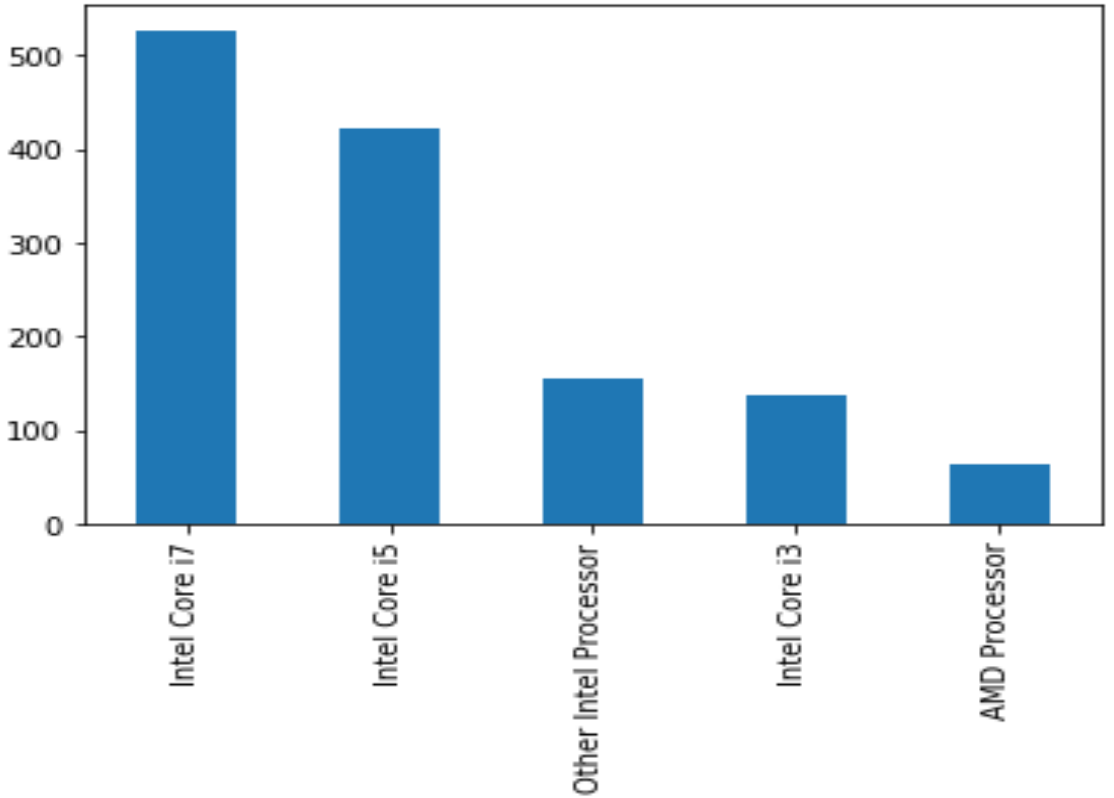
Touchscreen is More Expensive
Than other screens.



CPU brand

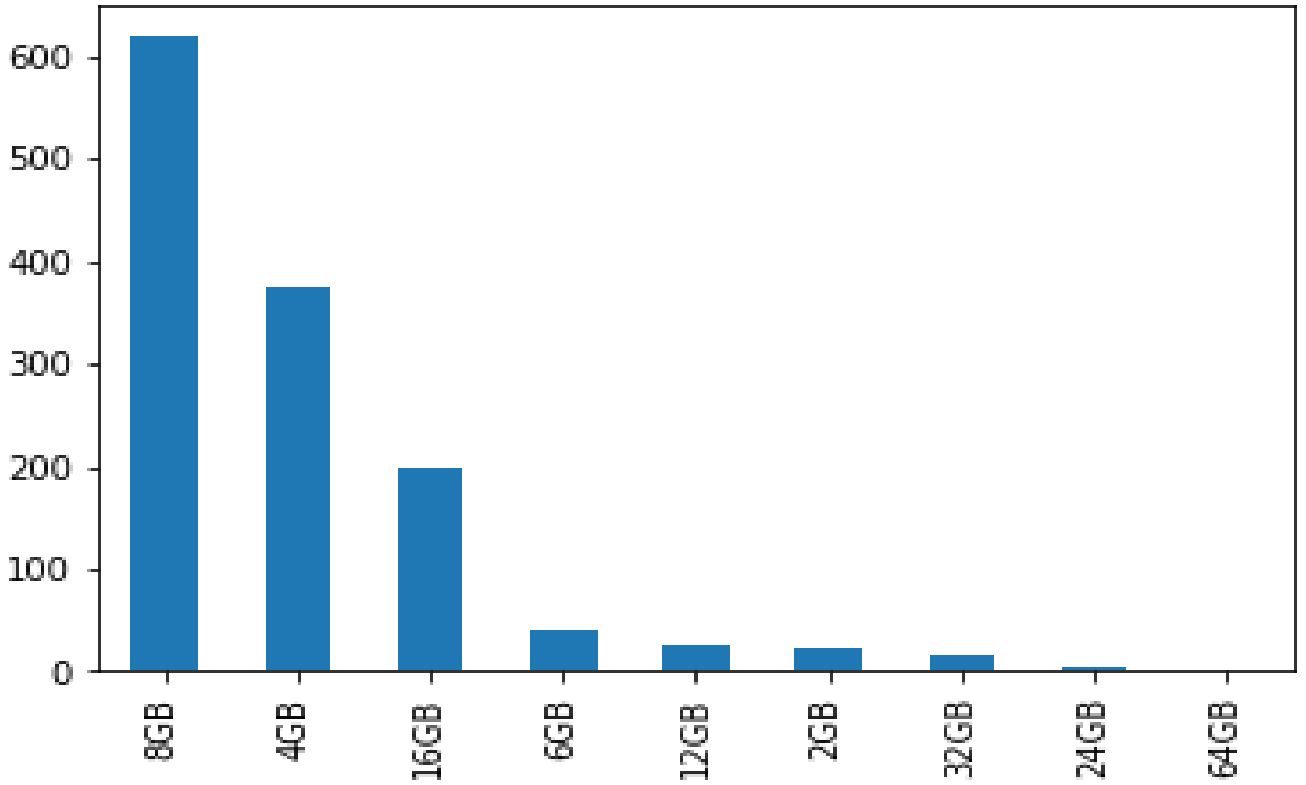
From slide No. 5 and from these two forms, we conclude that the CPU brand is the most expensive piece in a laptop.





Ram

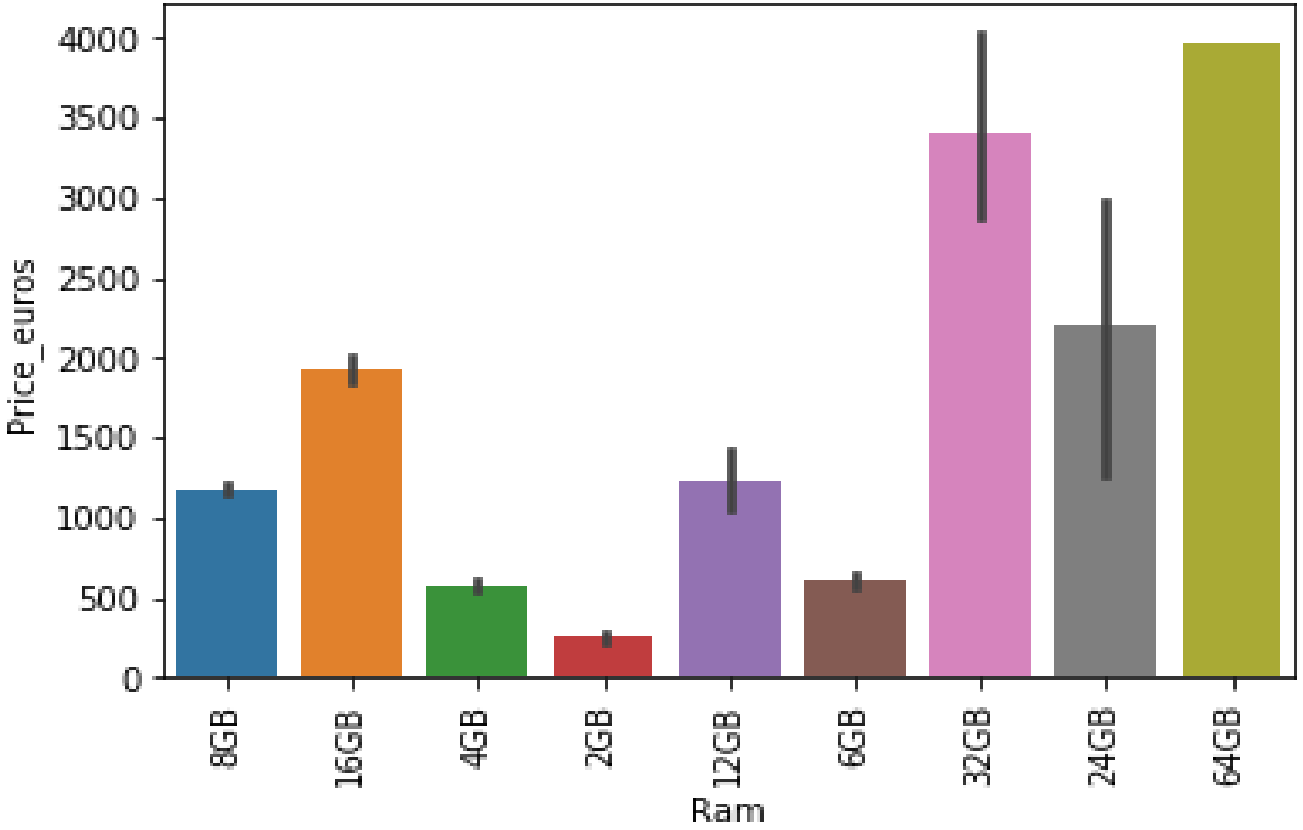
We conclude that 8 GB Ram is the most used among the laptops.



Price-euros & ram

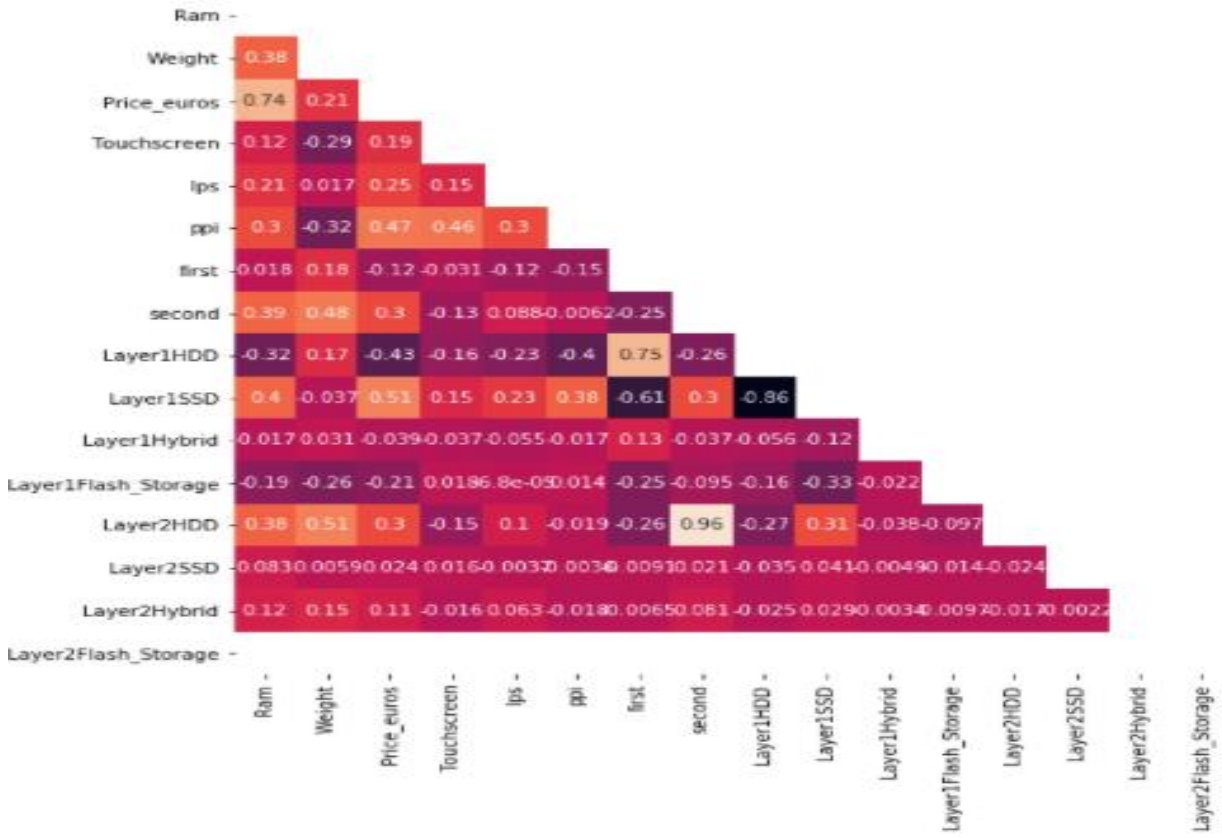
What is the relationship between RAM and prices?

There is a direct relationship between RAMs and prices.



The relationship between columns:

We conclude from the heatmap that the most influential columns on prices are RAM, SSD and CPU.



MODELING

LINEAR REGRESSION

DECISION TREE

RANDOM FOREST REGRESSION

LINEAR REGRESSION

We have found that accuracy of LINEAR REGRESSION IS 0.7

```
In [165]: # Let's see how accurate is our model.
          from sklearn import metrics
          accuracy_lin = metrics.r2_score(Y, y_pred_lr)
          print("Linear Regression ac: ",accuracy_lin)

          Linear Regression r2:  0.6567644514526807
```

RANDOM FOREST REGRESSION

We have found that accuracy of **RANDOM FOREST REGRESSION** IS 0.8

```
In [77]: accuracy_rf=metrics.r2_score(y_test,y_pred_rf)

         print("Random Forest Regression r2: ",accuracy_rf)

         Random Forest Regression r2:  0.7400989461569603
```

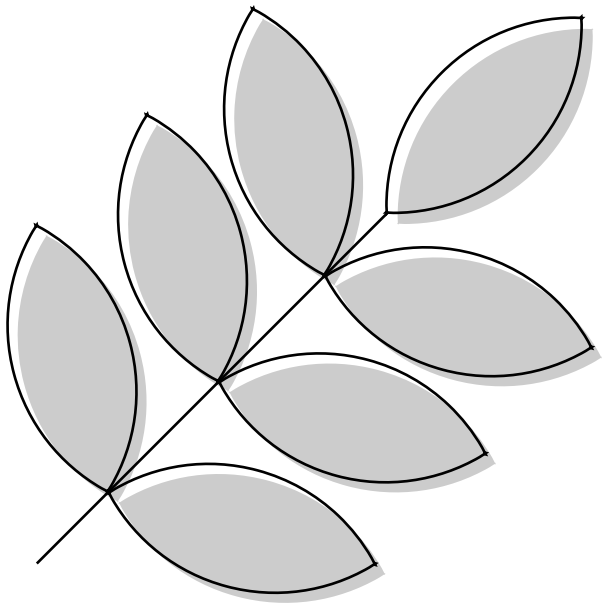
DECISION TREE

We have found that accuracy of **DECISION TREE** IS 0.97

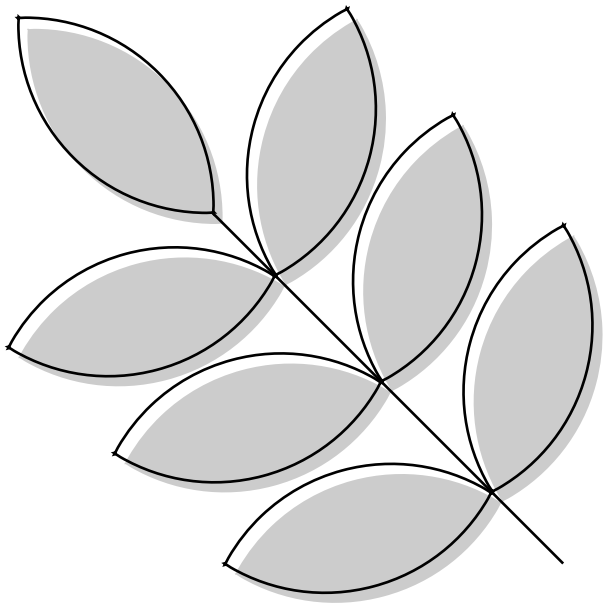
```
In [62]: accuracy_dt = metrics.r2_score(Y, y_pred_dt)
         print("Decision Tree Regression r2: ",accuracy_dt)

         Decision Tree Regression r2:  0.9683258162385 98
```

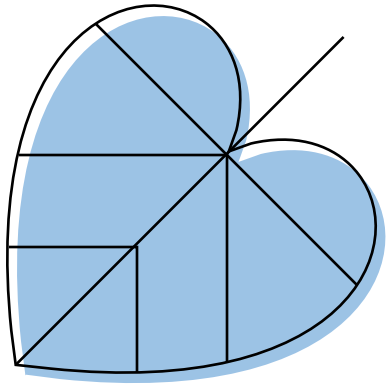
SO,

- We will use **decision tree**, Because it have the highest accuracy .



Thank you





Honored to have any questions!