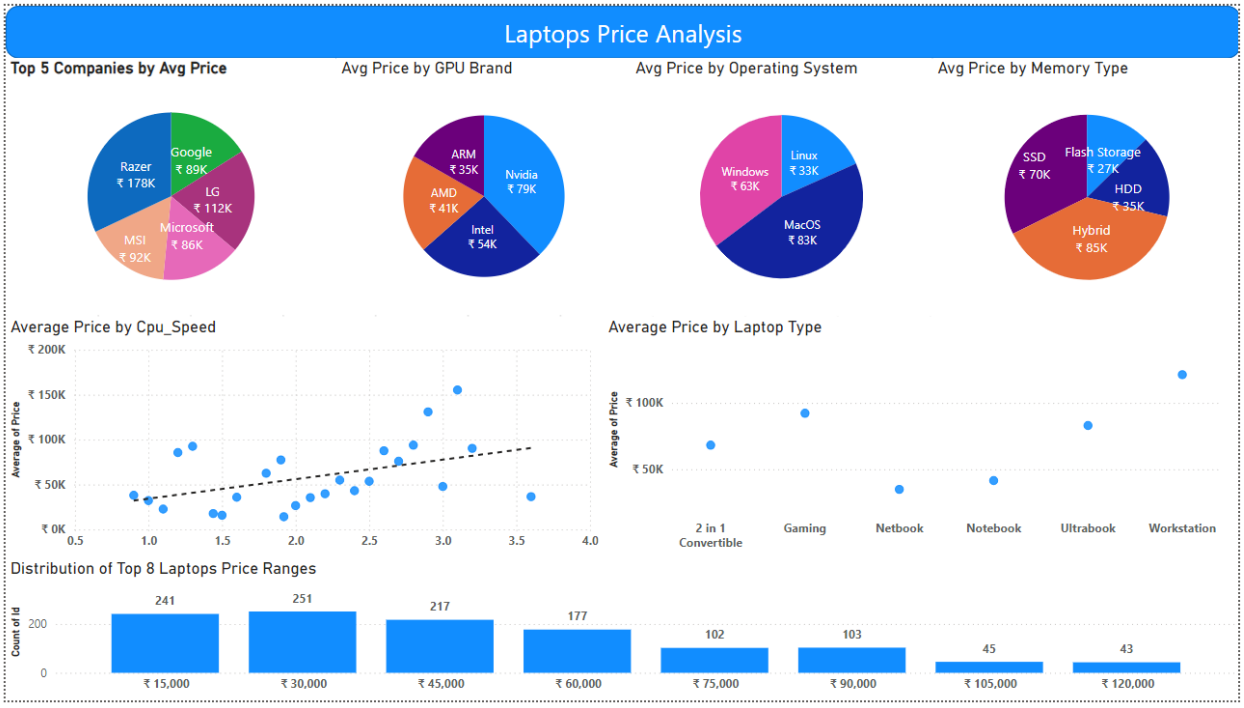



Dataset was obtained from Kaggle https://www.kaggle.com/datasets/ehtishamsadiq/uncleaned-laptop-price-dataset

Data was imported into SQL server from a .csv file

Table Schema -

- dbo.laptopdata
  - Columns
    - Unnamed_0 (smallint, null)
    - Company (nvarchar(50), null)
    - TypeName (nvarchar(50), null)
    - Inches (float, null)
    - ScreenResolution (nvarchar(50), null)
    - Cpu (nvarchar(50), null)
    - Ram (nvarchar(50), null)
    - Memory (nvarchar(50), null)
    - Gpu (nvarchar(50), null)
    - OpSys (nvarchar(50), null)
    - Weight (nvarchar(50), null)
    - Price (float, null)
  - Keys

Results Messages

| | Unnamed_0 | Company | TypeName | Inches | ScreenResolution | Cpu | Ram | Memory | Gpu | OpSys | Weight | Price |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0 | Apple | Ultrabook | 13.3000001907349 | IPS Panel Retina Display 2560x1600 | Intel Core i5 2.3GHz | 8GB | 128GB SSD | Intel Iris Plus Graphics 640 | macOS | 1.37kg | 71378.6796875 |
| 2 | 1 | Apple | Ultrabook | 13.3000001907349 | 1440x900 | Intel Core i5 1.8GHz | 8GB | 128GB Flash Storage | Intel HD Graphics 6000 | macOS | 1.34kg | 47895.5234375 |
| 3 | 2 | HP | Notebook | 15.6000003814697 | Full HD 1920x1080 | Intel Core i5 7200U 2.5GHz | 8GB | 256GB SSD | Intel HD Graphics 620 | No OS | 1.86kg | 30636 |
| 4 | 3 | Apple | Ultrabook | 15.3999996185303 | IPS Panel Retina Display 2880x1800 | Intel Core i7 2.7GHz | 16GB | 512GB SSD | AMD Radeon Pro 455 | macOS | 1.83kg | 135195.34375 |
| 5 | 4 | Apple | Ultrabook | 13.3000001907349 | IPS Panel Retina Display 2560x1600 | Intel Core i5 3.1GHz | 8GB | 256GB SSD | Intel Iris Plus Graphics 650 | macOS | 1.37kg | 96095.8046875 |
| 6 | 5 | Acer | Notebook | 15.6000003814697 | 1366x768 | AMD A9-Series 9420 3GHz | 4GB | 500GB HDD | AMD Radeon R5 | Windows 10 | 2.1kg | 21312 |
| 7 | 6 | Apple | Ultrabook | 15.3999996185303 | IPS Panel Retina Display 2880x1800 | Intel Core i7 2.2GHz | 16GB | 256GB Flash Storage | Intel Iris Pro Graphics | Mac OS X | 2.04kg | 114017.6015625 |
| 8 | 7 | Apple | Ultrabook | 13.3000001907349 | 1440x900 | Intel Core i5 1.8GHz | 8GB | 256GB Flash Storage | Intel HD Graphics 6000 | macOS | 1.34kg | 61735.53515625 |
| 9 | 8 | Asus | Ultrabook | 14 | Full HD 1920x1080 | Intel Core i7 8550U 1.8GHz | 16GB | 512GB SSD | Nvidia GeForce MX150 | Windows 10 | 1.3kg | 79653.6015625 |
| 10 | 9 | Acer | Ultrabook | 14 | IPS Panel Full HD 1920x1080 | Intel Core i5 8250U 1.6GHz | 8GB | 256GB SSD | Intel UHD Graphics 620 | Windows 10 | 1.6kg | 41025.6015625 |
| 11 | 10 | HP | Notebook | 15.6000003814697 | 1366x768 | Intel Core i5 7200U 2.5GHz | 4GB | 500GB HDD | Intel HD Graphics 620 | No OS | 1.86kg | 20986.9921875 |
| 12 | 11 | HP | Notebook | 15.6000003814697 | Full HD 1920x1080 | Intel Core i3 6006U 2GHz | 4GB | 500GB HDD | Intel HD Graphics 520 | No OS | 1.86kg | 18381.06640625 |

1. Rename column – Rename Unnamed_0 to Id.

```
--rename column Unnamed_0 to ID
EXEC sp_rename 'laptopdata.Unnamed_0', 'Id', 'COLUMN';

SELECT * FROM laptopdata
```

Results Messages

| | Id | Company | TypeName | Inches | ScreenResolution | Cpu | Ram | Memory | Gpu | OpSys | Weight | Price |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0 | Apple | Ultrabook | 13.3000001907349 | IPS Panel Retina Display 2560x1600 | Intel Core i5 2.3GHz | 8GB | 128GB SSD | Intel Iris Plus Graphics 640 | macOS | 1.37kg | 71378.6796875 |
| 2 | 1 | Apple | Ultrabook | 13.3000001907349 | 1440x900 | Intel Core i5 1.8GHz | 8GB | 128GB Flash Storage | Intel HD Graphics 6000 | macOS | 1.34kg | 47895.5234375 |
| 3 | 2 | HP | Notebook | 15.6000003814697 | Full HD 1920x1080 | Intel Core i5 7200U 2.5GHz | 8GB | 256GB SSD | Intel HD Graphics 620 | No OS | 1.86kg | 30636 |
| 4 | 3 | Apple | Ultrabook | 15.3999996185303 | IPS Panel Retina Display 2880x1800 | Intel Core i7 2.7GHz | 16GB | 512GB SSD | AMD Radeon Pro 455 | macOS | 1.83kg | 135195.34375 |
| 5 | 4 | Apple | Ultrabook | 13.3000001907349 | IPS Panel Retina Display 2560x1600 | Intel Core i5 3.1GHz | 8GB | 256GB SSD | Intel Iris Plus Graphics 650 | macOS | 1.37kg | 96095.8046875 |
| 6 | 5 | Acer | Notebook | 15.6000003814697 | 1366x768 | AMD A9-Series 9420 3GHz | 4GB | 500GB HDD | AMD Radeon R5 | Windows 10 | 2.1kg | 21312 |

2. Find and delete those records that have NULL values in all columns

```
--find rows that have null values in all columns
SELECT * FROM laptopdata
WHERE Company IS NULL AND TypeName IS NULL
AND ScreenResolution IS NULL AND Cpu IS NULL
AND Ram IS NULL AND Memory IS NULL AND Gpu IS NULL
AND OpSys IS NULL AND Weight IS NULL AND Price IS NULL;
--30 rows
```

Results Messages

| | Id | Company | TypeName | Inches | ScreenResolution | Cpu | Ram | Memory | Gpu | OpSys | Weight | Price |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | NULL | NULL | NULL | NULL | NULL | NULL | NULL | NULL | NULL | NULL | NULL | NULL |
| 2 | NULL | NULL | NULL | NULL | NULL | NULL | NULL | NULL | NULL | NULL | NULL | NULL |
| 3 | NULL | NULL | NULL | NULL | NULL | NULL | NULL | NULL | NULL | NULL | NULL | NULL |
| 4 | NULL | NULL | NULL | NULL | NULL | NULL | NULL | NULL | NULL | NULL | NULL | NULL |
| 5 | NULL | NULL | NULL | NULL | NULL | NULL | NULL | NULL | NULL | NULL | NULL | NULL |
| 6 | NULL | NULL | NULL | NULL | NULL | NULL | NULL | NULL | NULL | NULL | NULL | NULL |
| 7 | NULL | NULL | NULL | NULL | NULL | NULL | NULL | NULL | NULL | NULL | NULL | NULL |
| 8 | NULL | NULL | NULL | NULL | NULL | NULL | NULL | NULL | NULL | NULL | NULL | NULL |
| 9 | NULL | NULL | NULL | NULL | NULL | NULL | NULL | NULL | NULL | NULL | NULL | NULL |
| 10 | NULL | NULL | NULL | NULL | NULL | NULL | NULL | NULL | NULL | NULL | NULL | NULL |

```
DELETE FROM laptopdata WHERE Id IS NULL;
--total rows now = 1273
```

3. Convert Inches column from float to decimal(10,2). Also remove GB in Ram column.

```sql
-- convert Inches column from float to decimal(10,1)
ALTER TABLE laptopdata
ALTER COLUMN Inches decimal(10,1);

--remove GB from Ram column
UPDATE laptopdata
SET Ram = REPLACE(Ram, 'GB', '');

SELECT * FROM laptopdata

-- make Ram column a integer column from nvarchar(50)
ALTER TABLE laptopdata
ALTER COLUMN Ram int;

SELECT * FROM laptopdata
```

Before

Results Messages

| | Id | Company | TypeName | Inches | ScreenResolution | Cpu | Ram | Memory | Gpu | OpSys | Weight | Price |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0 | Apple | Ultrabook | 13.3000001907349 | IPS Panel Retina Display 2560x1600 | Intel Core i5 2.3GHz | 8GB | 128GB SSD | Intel Iris Plus Graphics 640 | macOS | 1.37kg | 71378.6796875 |
| 2 | 1 | Apple | Ultrabook | 13.3000001907349 | 1440x900 | Intel Core i5 1.8GHz | 8GB | 128GB Flash Storage | Intel HD Graphics 6000 | macOS | 1.34kg | 47895.5234375 |
| 3 | 2 | HP | Notebook | 15.6000003814697 | Full HD 1920x1080 | Intel Core i5 7200U 2.5GHz | 8GB | 256GB SSD | Intel HD Graphics 620 | No OS | 1.86kg | 30636 |
| 4 | 3 | Apple | Ultrabook | 15.3999996185303 | IPS Panel Retina Display 2880x1800 | Intel Core i7 2.7GHz | 16GB | 512GB SSD | AMD Radeon Pro 455 | macOS | 1.83kg | 135195.34375 |
| 5 | 4 | Apple | Ultrabook | 13.3000001907349 | IPS Panel Retina Display 2560x1600 | Intel Core i5 3.1GHz | 8GB | 256GB SSD | Intel Iris Plus Graphics 650 | macOS | 1.37kg | 96095.8046875 |
| 6 | 5 | Acer | Notebook | 15.6000003814697 | 1366x768 | AMD A9-Series 9420 3GHz | 4GB | 500GB HDD | AMD Radeon R5 | Windows 10 | 2.1kg | 21312 |

After

Results Messages

| | Id | Company | TypeName | Inches | ScreenResolution | Cpu | Ram | Memory | Gpu | OpSys | Weight | Price |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0 | Apple | Ultrabook | 13.3 | IPS Panel Retina Display 2560x1600 | Intel Core i5 2.3GHz | 8 | 128GB SSD | Intel Iris Plus Graphics 640 | macOS | 1.37kg | 71378.6796875 |
| 2 | 1 | Apple | Ultrabook | 13.3 | 1440x900 | Intel Core i5 1.8GHz | 8 | 128GB Flash Storage | Intel HD Graphics 6000 | macOS | 1.34kg | 47895.5234375 |
| 3 | 2 | HP | Notebook | 15.6 | Full HD 1920x1080 | Intel Core i5 7200U 2.5GHz | 8 | 256GB SSD | Intel HD Graphics 620 | No OS | 1.86kg | 30636 |
| 4 | 3 | Apple | Ultrabook | 15.4 | IPS Panel Retina Display 2880x1800 | Intel Core i7 2.7GHz | 16 | 512GB SSD | AMD Radeon Pro 455 | macOS | 1.83kg | 135195.34375 |
| 5 | 4 | Apple | Ultrabook | 13.3 | IPS Panel Retina Display 2560x1600 | Intel Core i5 3.1GHz | 8 | 256GB SSD | Intel Iris Plus Graphics 650 | macOS | 1.37kg | 96095.8046875 |
| 6 | 5 | Acer | Notebook | 15.6 | 1366x768 | AMD A9-Series 9420 3GHz | 4 | 500GB HDD | AMD Radeon R5 | Windows 10 | 2.1kg | 21312 |
| 7 | 6 | Apple | Ultrabook | 15.4 | IPS Panel Retina Display 2880x1800 | Intel Core i7 2.2GHz | 16 | 256GB Flash Storage | Intel Iris Pro Graphics | Mac OS X | 2.04kg | 114017.6015625 |
| 8 | 7 | Apple | Ultrabook | 13.3 | 1440x900 | Intel Core i5 1.8GHz | 8 | 256GB Flash Storage | Intel HD Graphics 6000 | macOS | 1.34kg | 61735.53515625 |
| 9 | 8 | Asus | Ultrabook | 14.0 | Full HD 1920x1080 | Intel Core i7 8550U 1.8GHz | 16 | 512GB SSD | Nvidia GeForce MX150 | Windows 10 | 1.3kg | 79653.6015625 |
| 10 | 9 | Acer | Ultrabook | 14.0 | IPS Panel Full HD 1920x1080 | Intel Core i5 8250U 1.6GHz | 8 | 256GB SSD | Intel UHD Graphics 620 | Windows 10 | 1.6kg | 41025.6015625 |

4. Remove Kg from Weight column. Round the Price column and change it to Integer. These Prices are in Rupees.

```
--remove kg from Weight column
UPDATE laptopdata
SET Weight = REPLACE(Weight, 'Kg', '');

--round Price to integer instead of float. First round it.
UPDATE laptopdata
SET Price = ROUND(Price, 0);


-- change Price data type
ALTER TABLE laptopdata
ALTER COLUMN Price INT;

select * from laptopdata
```

Before

| | OpSys | Weight | Price |
|---|---|---|---|
| 0 | macOS | 1.37kg | 71378.6796875 |
| | macOS | 1.34kg | 47895.5234375 |
| | No OS | 1.86kg | 30636 |
| | macOS | 1.83kg | 135195.34375 |
| 0 | macOS | 1.37kg | 96095.8046875 |
| | Windows 10 | 2.1kg | 21312 |
| | Mac OS X | 2.04kg | 114017.6015625 |
| | macOS | 1.34kg | 61735.53515625 |
| | Windows 10 | 1.3kg | 79653.6015625 |
| | Windows 10 | 1.6kg | 41025.6015625 |
| | No OS | 1.86kg | 20986.9921875 |
| | No OS | 1.86kg | 18381.06640625 |
| | macOS | 1.83kg | 130001.6015625 |

Results Messages

| | Id | Company | TypeName | Inches | ScreenResolution | Cpu | Ram | Memory | Gpu | OpSys | Weight | Price |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0 | Apple | Ultrabook | 13.3 | IPS Panel Retina Display 2560x1600 | Intel Core i5 2.3GHz | 8 | 128GB SSD | Intel Iris Plus Graphics 640 | macOS | 1.37 | 71379 |
| 2 | 1 | Apple | Ultrabook | 13.3 | 1440x900 | Intel Core i5 1.8GHz | 8 | 128GB Flash Storage | Intel HD Graphics 6000 | macOS | 1.34 | 47896 |
| 3 | 2 | HP | Notebook | 15.6 | Full HD 1920x1080 | Intel Core i5 7200U 2.5GHz | 8 | 256GB SSD | Intel HD Graphics 620 | No OS | 1.86 | 30636 |
| 4 | 3 | Apple | Ultrabook | 15.4 | IPS Panel Retina Display 2880x1800 | Intel Core i7 2.7GHz | 16 | 512GB SSD | AMD Radeon Pro 455 | macOS | 1.83 | 135195 |
| 5 | 4 | Apple | Ultrabook | 13.3 | IPS Panel Retina Display 2560x1600 | Intel Core i5 3.1GHz | 8 | 256GB SSD | Intel Iris Plus Graphics 650 | macOS | 1.37 | 96096 |
| 6 | 5 | Acer | Notebook | 15.6 | 1366x768 | AMD A9-Series 9420 3GHz | 4 | 500GB HDD | AMD Radeon R5 | Windows 10 | 2.1 | 21312 |
| 7 | 6 | Apple | Ultrabook | 15.4 | IPS Panel Retina Display 2880x1800 | Intel Core i7 2.2GHz | 16 | 256GB Flash Storage | Intel Iris Pro Graphics | Mac OS X | 2.04 | 114018 |
| 8 | 7 | Apple | Ultrabook | 13.3 | 1440x900 | Intel Core i5 1.8GHz | 8 | 256GB Flash Storage | Intel HD Graphics 6000 | macOS | 1.34 | 61736 |
| 9 | 8 | Asus | Ultrabook | 14.0 | Full HD 1920x1080 | Intel Core i7 8550U 1.8GHz | 16 | 512GB SSD | Nvidia GeForce MX150 | Windows 10 | 1.3 | 79654 |

5. Operating System – OpSys

```
--OpSys -
SELECT Opsys,COUNT(*) as count FROM laptopdata
GROUP BY OpSys;
```

Results Messages

| | Opsys | count |
|---|---|---|
| 1 | Windows 7 | 45 |
| 2 | No OS | 63 |
| 3 | Linux | 61 |
| 4 | Android | 1 |
| 5 | Chrome OS | 27 |
| 6 | Windows 10 | 1047 |
| 7 | macOS | 13 |
| 8 | Windows 10 S | 8 |
| 9 | Mac OS X | 8 |

***Group by Operating System Name instead of by Version number.

```
--Replace opsys with Windows, MacOS, Linux, Others
SELECT OpSys,
CASE
    WHEN opsys LIKE '%mac%' THEN 'MacOS'
    WHEN OpSys LIKE '%windows%' THEN 'Windows'
    WHEN OpSys LIKE '%Linux%' THEN 'Linux'
    WHEN OpSys = 'No OS' THEN 'N/A'
    ELSE 'Other'
END AS 'OS_Brand'
FROM laptopdata;

UPDATE laptopdata
SET OpSys =
CASE
    WHEN opsys LIKE '%mac%' THEN 'MacOS'
    WHEN OpSys LIKE '%windows%' THEN 'Windows'
    WHEN OpSys LIKE '%Linux%' THEN 'Linux'
    WHEN OpSys = 'No OS' THEN 'N/A'
    ELSE 'Other'
END;

--
SELECT Opsys,COUNT(*) as count FROM laptopdata
group by OpSys
```

Results Messages

| | Opsys | count |
|---|---|---|
| 1 | N/A | 63 |
| 2 | Linux | 61 |
| 3 | MacOS | 21 |
| 4 | Other | 28 |
| 5 | Windows | 1100 |

6. GPU column – break it into Gpu_Brand and Gpu_Name.

```sql
--Break Gpu into Gpu_Name and Gpu_Brand
ALTER TABLE laptopdata
ADD Gpu_Brand VARCHAR(50),
Gpu_Name VARCHAR(50);

--select substring(gpu,1,CHARINDEX(' ',Gpu)-1) from laptopdata
UPDATE t1
SET Gpu_Brand = (SELECT SUBSTRING(t2.gpu,1,CHARINDEX(' ',t2.Gpu)-1) FROM laptopdata t2
                    WHERE t1.Id = t2.Id) FROM laptopdata t1;

--replace brand name with nothing in gpu and that gives us the type of graphics card
--select gpu, replace(gpu, gpu_brand,'') from laptopdata

UPDATE t1
SET Gpu_Name = (SELECT REPLACE(Gpu, Gpu_Brand, '') FROM laptopdata t2
                    WHERE t1.Id = t2.Id) FROM laptopdata t1;


SELECT * FROM laptopdata

--drop Gpu column
ALTER TABLE laptopdata
DROP COLUMN Gpu;

SELECT * FROM laptopdata;
```

Before

| Gpu |
|---|
| Intel Iris Plus Graphics 640 |
| Intel HD Graphics 6000 |
| Intel HD Graphics 620 |
| AMD Radeon Pro 455 |
| Intel Iris Plus Graphics 650 |
| AMD Radeon R5 |
| Intel Iris Pro Graphics |
| Intel HD Graphics 6000 |
| Nvidia GeForce MX150 |
| Intel UHD Graphics 620 |
| Intel HD Graphics 620 |
| Intel HD Graphics 520 |
| AMD Radeon Pro 555 |

After

| Gpu_Brand | Gpu_Name |
|---|---|
| Intel | Iris Plus Graphics 640 |
| Intel | HD Graphics 6000 |
| Intel | HD Graphics 620 |
| AMD | Radeon Pro 455 |
| Intel | Iris Plus Graphics 650 |
| AMD | Radeon R5 |
| Intel | Iris Pro Graphics |
| Intel | HD Graphics 6000 |
| Nvidia | GeForce MX150 |
| Intel | UHD Graphics 620 |
| Intel | HD Graphics 620 |
| Intel | HD Graphics 520 |
| AMD | Radeon Pro 555 |

After

Results Messages

| | Id | Company | TypeName | Inches | ScreenResolution | Cpu | Ram | Memory | OpSys | Weight | Price | Gpu_Brand | Gpu_Name |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0 | Apple | Ultrabook | 13.3 | IPS Panel Retina Display 2560x1600 | Intel Core i5 2.3GHz | 8 | 128GB SSD | MacOS | 1.37 | 71379 | Intel | Iris Plus Graphics 640 |
| 2 | 1 | Apple | Ultrabook | 13.3 | 1440x900 | Intel Core i5 1.8GHz | 8 | 128GB Flash Storage | MacOS | 1.34 | 47896 | Intel | HD Graphics 6000 |
| 3 | 2 | HP | Notebook | 15.6 | Full HD 1920x1080 | Intel Core i5 7200U 2.5GHz | 8 | 256GB SSD | N/A | 1.86 | 30636 | Intel | HD Graphics 620 |
| 4 | 3 | Apple | Ultrabook | 15.4 | IPS Panel Retina Display 2880x1800 | Intel Core i7 2.7GHz | 16 | 512GB SSD | MacOS | 1.83 | 135195 | AMD | Radeon Pro 455 |
| 5 | 4 | Apple | Ultrabook | 13.3 | IPS Panel Retina Display 2560x1600 | Intel Core i5 3.1GHz | 8 | 256GB SSD | MacOS | 1.37 | 96096 | Intel | Iris Plus Graphics 650 |
| 6 | 5 | Acer | Notebook | 15.6 | 1366x768 | AMD A9-Series 9420 3GHz | 4 | 500GB HDD | Windows | 2.1 | 21312 | AMD | Radeon R5 |
| 7 | 6 | Apple | Ultrabook | 15.4 | IPS Panel Retina Display 2880x1800 | Intel Core i7 2.2GHz | 16 | 256GB Flash Storage | MacOS | 2.04 | 114018 | Intel | Iris Pro Graphics |
| 8 | 7 | Apple | Ultrabook | 13.3 | 1440x900 | Intel Core i5 1.8GHz | 8 | 256GB Flash Storage | MacOS | 1.34 | 61736 | Intel | HD Graphics 6000 |
| 9 | 8 | Asus | Ultrabook | 14.0 | Full HD 1920x1080 | Intel Core i7 8550U 1.8GHz | 16 | 512GB SSD | Windows | 1.3 | 79654 | Nvidia | GeForce MX150 |

7. Cpu column – separate Cpu column into Cpu_Brand, Cpu_Speed, Cpu_Name

```sql
--CPU - break Cpu into 3 cols - Cpu_Brand, Cpu_Speed, Cpu_Name
ALTER TABLE laptopdata
ADD
Cpu_Brand VARCHAR(50),
Cpu_Name VARCHAR(50),
Cpu_Speed DECIMAL(10,1);

--select CPU, substring(cpu,1,CHARINDEX(' ',Cpu)-1) from laptopdata
UPDATE t1
SET Cpu_Brand = (SELECT SUBSTRING(Cpu,1,CHARINDEX(' ',Cpu)-1) FROM laptopdata t2
                    WHERE t1.Id = t2.Id) FROM laptopdata t1;

--set speed as decimal and also remove GHz
UPDATE t1
SET Cpu_Speed = (SELECT CAST(REPLACE(TRIM(SUBSTRING(Cpu, LEN(Cpu) - CHARINDEX(' ', REVERSE(Cpu)) + 1, LEN(Cpu))),
                    'GHz', '') AS DECIMAL(10,2))
                    FROM laptopdata t2 WHERE t1.Id = t2.Id)
                    FROM laptopdata t1;

--updata Cpu_Name
WITH cpucte AS
(
    SELECT Id, TRIM(REPLACE(Cpu, Cpu_Brand, '')) AS temp FROM laptopdata
)
UPDATE t1
SET Cpu_Name = SUBSTRING(temp , 1, LEN(temp) - CHARINDEX(' ', REVERSE(temp))+1)
                FROM laptopdata t1
                INNER JOIN cpucte ON t1.Id = cpucte.Id;


SELECT * FROM laptopdata
```

Before

| Cpu | Ram |
|---|---|
| Intel Core i5 2.3GHz | 8 |
| Intel Core i5 1.8GHz | 8 |
| Intel Core i5 7200U 2.5GHz | 8 |
| Intel Core i7 2.7GHz | 16 |
| Intel Core i5 3.1GHz | 8 |
| AMD A9-Series 9420 3GHz | 4 |
| Intel Core i7 2.2GHz | 16 |
| Intel Core i5 1.8GHz | 8 |
| Intel Core i7 8550U 1.8GHz | 16 |
| Intel Core i5 8250U 1.6GHz | 8 |
| Intel Core i5 7200U 2.5GHz | 4 |
| Intel Core i3 6006U 2GHz | 4 |
| Intel Core i7 2.8GHz | 16 |

After

| Cpu_Brand | Cpu_Name | Cpu_Speed |
|---|---|---|
| Intel | Core i5 | 2.3 |
| Intel | Core i5 | 1.8 |
| Intel | Core i5 7200U | 2.5 |
| Intel | Core i7 | 2.7 |
| Intel | Core i5 | 3.1 |
| AMD | A9-Series 9420 | 3.0 |
| Intel | Core i7 | 2.2 |
| Intel | Core i5 | 1.8 |
| Intel | Core i7 8550U | 1.8 |
| Intel | Core i5 8250U | 1.6 |
| Intel | Core i5 7200U | 2.5 |
| Intel | Core i3 6006U | 2.0 |
| Intel | Core i7 | 2.8 |

**Shortening the Cpu_Name column. For eg., Core i5 7200U becomes Core i5. This is better for categorical column analysis. The number of distinct Cpu_Names reduce from 92 to 39. After that Cpu column is dropped.

```sql
--deleting 7200 U in Core i5 7200U, keeping only first 2 for categorical analysis
UPDATE t1
SET Cpu_name = (SELECT
                -- Extract the first two words from cpu_name
                CASE
                    WHEN CHARINDEX(' ', cpu_name) = 0 THEN cpu_name
                    WHEN CHARINDEX(' ', cpu_name, CHARINDEX(' ', cpu_name) + 1) = 0 THEN cpu_name
                    ELSE
                        SUBSTRING(
                            cpu_name,
                            1,
                            CHARINDEX(' ', cpu_name, CHARINDEX(' ', cpu_name) + 1) - 1
                        )
                END AS FirstTwoWords
            FROM laptopdata t2
            WHERE t1.Id = t2.Id)
            FROM laptopdata t1;

SELECT  DISTINCT(cpu_name) FROM laptopdata;--39 distinct cpu_names now


--drop Cpu column
ALTER TABLE laptopdata
DROP COLUMN Cpu;

SELECT * FROM laptopdata;
```

Before

| Cpu_Brand | Cpu_Name | Cpu_Speed |
|---|---|---|
| Intel | Core i5 | 2.3 |
| Intel | Core i5 | 1.8 |
| Intel | Core i5 7200U | 2.5 |
| Intel | Core i7 | 2.7 |
| Intel | Core i5 | 3.1 |
| AMD | A9-Series 9420 | 3.0 |
| Intel | Core i7 | 2.2 |
| Intel | Core i5 | 1.8 |
| Intel | Core i7 8550U | 1.8 |
| Intel | Core i5 8250U | 1.6 |
| Intel | Core i5 7200U | 2.5 |
| Intel | Core i3 6006U | 2.0 |
| Intel | Core i7 | 2.8 |

After

| Cpu_Name |
|---|
| Core i5 |
| Core i5 |
| Core i5 |
| Core i7 |
| Core i5 |
| A9-Series 9420 |
| Core i7 |
| Core i5 |
| Core i7 |
| Core i5 |
| Core i5 |
| Core i3 |
| Core i7 |

Results Messages

| | Id | Company | TypeName | Inches | ScreenResolution | Ram | Memory | OpSys | Weight | Price | Gpu_Brand | Gpu_Name | Cpu_Brand | Cpu_Name | Cpu_Speed |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0 | Apple | Ultrabook | 13.3 | IPS Panel Retina Display 2560x1600 | 8 | 128GB SSD | MacOS | 1.37 | 71379 | Intel | Iris Plus Graphics 640 | Intel | Core i5 | 2.3 |
| 2 | 1 | Apple | Ultrabook | 13.3 | 1440x900 | 8 | 128GB Flash Storage | MacOS | 1.34 | 47896 | Intel | HD Graphics 6000 | Intel | Core i5 | 1.8 |
| 3 | 2 | HP | Notebook | 15.6 | Full HD 1920x1080 | 8 | 256GB SSD | N/A | 1.86 | 30636 | Intel | HD Graphics 620 | Intel | Core i5 | 2.5 |
| 4 | 3 | Apple | Ultrabook | 15.4 | IPS Panel Retina Display 2880x1800 | 16 | 512GB SSD | MacOS | 1.83 | 135195 | AMD | Radeon Pro 455 | Intel | Core i7 | 2.7 |
| 5 | 4 | Apple | Ultrabook | 13.3 | IPS Panel Retina Display 2560x1600 | 8 | 256GB SSD | MacOS | 1.37 | 96096 | Intel | Iris Plus Graphics 650 | Intel | Core i5 | 3.1 |
| 6 | 5 | Acer | Notebook | 15.6 | 1366x768 | 4 | 500GB HDD | Windows | 2.1 | 21312 | AMD | Radeon R5 | AMD | A9-Series 9420 | 3.0 |
| 7 | 6 | Apple | Ultrabook | 15.4 | IPS Panel Retina Display 2880x1800 | 16 | 256GB Flash Storage | MacOS | 2.04 | 114018 | Intel | Iris Pro Graphics | Intel | Core i7 | 2.2 |
| 8 | 7 | Apple | Ultrabook | 13.3 | 1440x900 | 8 | 256GB Flash Storage | MacOS | 1.34 | 61736 | Intel | HD Graphics 6000 | Intel | Core i5 | 1.8 |

8. Break Screen resolution column into width and height.

```sql
--ScreenResolution - add 2 columns to hold the width and height
ALTER TABLE laptopdata
ADD
Resolution_Width INT,
Resolution_Height INT;

--Get width and height
--width
UPDATE t1
SET Resolution_Width =
                    (SELECT
                    CASE
                        WHEN
                        (LEN(ScreenResolution) - CHARINDEX('x',REVERSE(ScreenResolution))+2)-
                        (LEN(ScreenResolution) - CHARINDEX(' ',REVERSE(ScreenResolution))+3) < 0
                        THEN
                        SUBSTRING(Screenresolution, 1,CHARINDEX('x', screenresolution)-1)
                    ELSE
                        SUBSTRING(Screenresolution,
                        (LEN(ScreenResolution) - CHARINDEX(' ',REVERSE(ScreenResolution))+2),
                        ABS((LEN(ScreenResolution) - CHARINDEX('x',REVERSE(ScreenResolution))+2)-
                        (LEN(ScreenResolution) - CHARINDEX(' ',REVERSE(ScreenResolution))+3))
                        )
                    END
                FROM laptopdata t2
                WHERE t1.Id = t2.Id
                )
        FROM laptopdata t1;
```

```sql
--height
UPDATE t1
SET Resolution_Height = (
                        SELECT
                        TRIM(SUBSTRING(ScreenResolution, CHARINDEX('x', screenresolution)+1, LEN(screenresolution)))
                        FROM laptopdata t2
                        WHERE t1.Id = t2.Id
                        )
        FROM laptopdata t1;


SELECT * FROM laptopdata;
```

Before

| ScreenResolution | Ram |
|---|---|
| IPS Panel Retina Display 2560x1600 | 8 |
| 1440x900 | 8 |
| Full HD 1920x1080 | 8 |
| IPS Panel Retina Display 2880x1800 | 16 |
| IPS Panel Retina Display 2560x1600 | 8 |
| 1366x768 | 4 |
| IPS Panel Retina Display 2880x1800 | 16 |
| 1440x900 | 8 |
| Full HD 1920x1080 | 16 |
| IPS Panel Full HD 1920x1080 | 8 |
| 1366x768 | 4 |
| Full HD 1920x1080 | 4 |
| IPS Panel Retina Display 2880x1800 | 16 |

After

| Resolution_Width | Resolution_Height |
|---|---|
| 2560 | 1600 |
| 1440 | 900 |
| 1920 | 1080 |
| 2880 | 1800 |
| 2560 | 1600 |
| 1366 | 768 |
| 2880 | 1800 |
| 1440 | 900 |
| 1920 | 1080 |
| 1920 | 1080 |
| 1366 | 768 |
| 1920 | 1080 |
| 2880 | 1800 |

Results Messages

| | Id | Company | TypeName | Inches | ScreenResolution | Ram | Memory | OpSys | Weight | Price | Gpu_Brand | Gpu_Name | Cpu_Brand | Cpu_Name | Cpu_Speed | Resolution_Width | Resolution_Height |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0 | Apple | Ultrabook | 13.3 | IPS Panel Retina Display 256... | 8 | 128GB SSD | MacOS | 1.37 | 71379 | Intel | Iris Plus Graphics... | Intel | Core i5 | 2.3 | 2560 | 1600 |
| 2 | 1 | Apple | Ultrabook | 13.3 | 1440x900 | 8 | 128GB Flash Storage | MacOS | 1.34 | 47896 | Intel | HD Graphics 6000 | Intel | Core i5 | 1.8 | 1440 | 900 |
| 3 | 2 | HP | Notebook | 15.6 | Full HD 1920x1080 | 8 | 256GB SSD | N/A | 1.86 | 30636 | Intel | HD Graphics 620 | Intel | Core i5 | 2.5 | 1920 | 1080 |
| 4 | 3 | Apple | Ultrabook | 15.4 | IPS Panel Retina Display 288... | 16 | 512GB SSD | MacOS | 1.83 | 135195 | AMD | Radeon Pro 455 | Intel | Core i7 | 2.7 | 2880 | 1800 |
| 5 | 4 | Apple | Ultrabook | 13.3 | IPS Panel Retina Display 256... | 8 | 256GB SSD | MacOS | 1.37 | 96096 | Intel | Iris Plus Graphics... | Intel | Core i5 | 3.1 | 2560 | 1600 |
| 6 | 5 | Acer | Notebook | 15.6 | 1366x768 | 4 | 500GB HDD | Windows | 2.1 | 21312 | AMD | Radeon R5 | AMD | A9-Series 9420 | 3.0 | 1366 | 768 |
| 7 | 6 | Apple | Ultrabook | 15.4 | IPS Panel Retina Display 288... | 16 | 256GB Flash Storage | MacOS | 2.04 | 114018 | Intel | Iris Pro Graphics | Intel | Core i7 | 2.2 | 2880 | 1800 |
| 8 | 7 | Apple | Ultrabook | 13.3 | 1440x900 | 8 | 256GB Flash Storage | MacOS | 1.34 | 61736 | Intel | HD Graphics 6000 | Intel | Core i5 | 1.8 | 1440 | 900 |
| 9 | 8 | Asus | Ultrabook | 14.0 | Full HD 1920x1080 | 16 | 512GB SSD | Windows | 1.3 | 79654 | Nvidia | GeForce MX150 | Intel | Core i7 | 1.8 | 1920 | 1080 |
| 10 | 9 | Acer | Ultrabook | 14.0 | IPS Panel Full HD 1920x1080 | 8 | 256GB SSD | Windows | 1.6 | 41026 | Intel | UHD Graphics 620 | Intel | Core i5 | 1.6 | 1920 | 1080 |
| 11 | 10 | HP | Notebook | 15.6 | 1366x768 | 4 | 500GB HDD | N/A | 1.86 | 20987 | Intel | HD Graphics 620 | Intel | Core i5 | 2.5 | 1366 | 768 |

***Create a new column to hold if the laptop has touchscreen which is indicated in the ScreenResolution column. And then delete the ScreenResolution column**

```
1  --Adding if it is touchscreen or not
2  ALTER TABLE laptopdata
3  ADD Touchscreen INTEGER;
4
5  SELECT * FROM laptopdata WHERE ScreenResolution LIKE '%Touch%';
6
7  UPDATE laptopdata
8  SET Touchscreen = CASE
9                       WHEN ScreenResolution LIKE '%Touch%' THEN 1
10                      ELSE 0
11                    END;
12
13
14 SELECT * FROM laptopdata;
15
16 ALTER TABLE laptopdata
17 DROP COLUMN ScreenResolution;
18
19 SELECT * FROM laptopdata;
```

Results Messages

| | Id | Company | TypeName | Inches | Ram | Memory | OpSys | Weight | Price | Gpu_Brand | Gpu_Name | Cpu_Brand | Cpu_Name | Cpu_Speed | Resolution_Width | Resolution_Height | Touchscreen |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0 | Apple | Ultrabook | 13.3 | 8 | 128GB SSD | MacOS | 1.37 | 71379 | Intel | Iris Plus Graphics 640 | Intel | Core i5 | 2.3 | 2560 | 1600 | 0 |
| 2 | 1 | Apple | Ultrabook | 13.3 | 8 | 128GB Flash Storage | MacOS | 1.34 | 47896 | Intel | HD Graphics 6000 | Intel | Core i5 | 1.8 | 1440 | 900 | 0 |
| 3 | 2 | HP | Notebook | 15.6 | 8 | 256GB SSD | N/A | 1.86 | 30636 | Intel | HD Graphics 620 | Intel | Core i5 | 2.5 | 1920 | 1080 | 0 |
| 4 | 3 | Apple | Ultrabook | 15.4 | 16 | 512GB SSD | MacOS | 1.83 | 135195 | AMD | Radeon Pro 455 | Intel | Core i7 | 2.7 | 2880 | 1800 | 0 |
| 5 | 4 | Apple | Ultrabook | 13.3 | 8 | 256GB SSD | MacOS | 1.37 | 96096 | Intel | Iris Plus Graphics 650 | Intel | Core i5 | 3.1 | 2560 | 1600 | 0 |
| 6 | 5 | Acer | Notebook | 15.6 | 4 | 500GB HDD | Windows | 2.1 | 21312 | AMD | Radeon R5 | AMD | A9-Series 9420 | 3.0 | 1366 | 768 | 0 |
| 7 | 6 | Apple | Ultrabook | 15.4 | 16 | 256GB Flash Storage | MacOS | 2.04 | 114018 | Intel | Iris Pro Graphics | Intel | Core i7 | 2.2 | 2880 | 1800 | 0 |
| 8 | 7 | Apple | Ultrabook | 13.3 | 8 | 256GB Flash Storage | MacOS | 1.34 | 61736 | Intel | HD Graphics 6000 | Intel | Core i5 | 1.8 | 1440 | 900 | 0 |
| 9 | 8 | Asus | Ultrabook | 14.0 | 16 | 512GB SSD | Windows | 1.3 | 79654 | Nvidia | GeForce MX150 | Intel | Core i7 | 1.8 | 1920 | 1080 | 0 |

9. Memory column – First group Memory based on type (SSD, HDD, Flash Storage, Hybrid). Then for a memory of type '128GB SSD + 1TB HDD' break into two additional columns with Primary_Storage having 128 and Secondary_Storage having 1. Later convert all TBs to GBs and there should be no units.

```sql
--Memory
ALTER TABLE laptopdata
ADD
Memory_Type VARCHAR(50),
Primary_Storage INTEGER,
Secondary_Storage INTEGER;


SELECT Memory,
CASE
    WHEN Memory LIKE '%SSD%' AND Memory LIKE '%HDD%' THEN 'Hybrid'
    WHEN Memory LIKE '%SSD%' THEN 'SSD'
    WHEN Memory LIKE '%HDD%' THEN 'HDD'
    WHEN Memory LIKE '%Flash Storage%' THEN 'Flash Storage'
    WHEN Memory LIKE '%Hybrid%' THEN 'Hybrid'
    WHEN Memory LIKE '%Flash Storage%' AND Memory LIKE '%HDD%' THEN 'Hybrid'
    ELSE NULL
END AS 'memory_type'
FROM laptopdata;

UPDATE laptopdata
SET memory_type = CASE
    WHEN Memory LIKE '%SSD%' AND Memory LIKE '%HDD%' THEN 'Hybrid'
    WHEN Memory LIKE '%SSD%' THEN 'SSD'
    WHEN Memory LIKE '%HDD%' THEN 'HDD'
    WHEN Memory LIKE '%Flash Storage%' THEN 'Flash Storage'
    WHEN Memory LIKE '%Hybrid%' THEN 'Hybrid'
    WHEN Memory LIKE '%Flash Storage%' AND Memory LIKE '%HDD%' THEN 'Hybrid'
    ELSE NULL
END;
```

```sql
-- populate the columns with cte
WITH cte AS
(
    SELECT
        Id, Memory,
        -- Extract the number before the '+' sign
        CASE
            WHEN CHARINDEX('+', Memory) > 0 THEN
                -- Extract substring before '+', then remove non-numeric characters
                LEFT(SUBSTRING(Memory, 1, CHARINDEX('+', Memory) - 1),
                    PATINDEX('%[^0-9]%', SUBSTRING(Memory, 1, CHARINDEX('+', Memory) - 1)) - 1)
            ELSE
                -- Remove non-numeric characters
                LEFT(Memory, PATINDEX('%[^0-9]%', Memory) - 1)
        END AS Memory_Before_Plus,

        -- Extract the number after the '+' sign, if present
        CASE
            WHEN CHARINDEX('+', Memory) > 0 THEN
                CAST(
                LEFT(
                    LTRIM(RTRIM(SUBSTRING(Memory, CHARINDEX('+', Memory) + 1,
                        LEN(Memory) - CHARINDEX('+', Memory)))),
                    PATINDEX('%[^0-9]%', LTRIM(RTRIM(SUBSTRING(Memory, CHARINDEX('+', Memory) + 1,
                        LEN(Memory) - CHARINDEX('+', Memory)))) + 'X') - 1
                ) AS INT
            )
            ELSE
                '0'
        END AS Memory_After_Plus
    FROM laptopdata
)
```

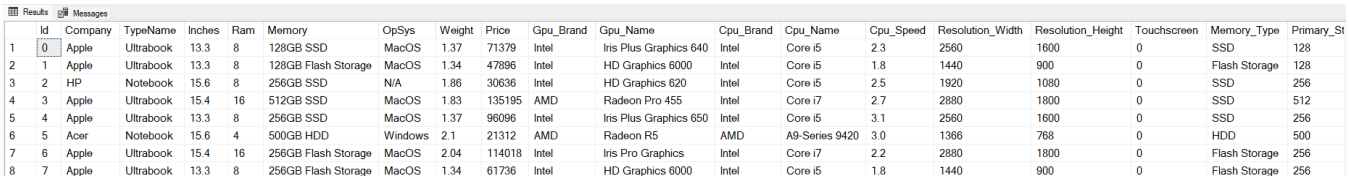
```sql
-- Update the laptopdata table using the cleaned values from the CTE
UPDATE t1
SET
    t1.primary_storage = TRY_CAST(cte.Memory_Before_Plus AS INT),
    t1.secondary_storage = TRY_CAST(cte.Memory_After_Plus AS INT)
FROM laptopdata t1
INNER JOIN cte ON t1.Id = cte.Id;

SELECT * FROM laptopdata
```

Before

| Memory | OpSys |
|---|---|
| 128GB SSD | MacOS |
| 128GB Flash Storage | MacOS |
| 256GB SSD | N/A |
| 512GB SSD | MacOS |
| 256GB SSD | MacOS |
| 500GB HDD | Window |
| 256GB Flash Storage | MacOS |
| 256GB Flash Storage | MacOS |
| 512GB SSD | Window |
| 256GB SSD | Window |
| 500GB HDD | N/A |
| 500GB HDD | N/A |
| 256GB SSD | MacOS |

After

| Memory_Type | Primary_Storage | Secondary_Storage |
|---|---|---|
| SSD | 128 | 0 |
| Flash Storage | 128 | 0 |
| SSD | 256 | 0 |
| SSD | 512 | 0 |
| SSD | 256 | 0 |
| HDD | 500 | 0 |
| Flash Storage | 256 | 0 |
| Flash Storage | 256 | 0 |
| SSD | 512 | 0 |
| SSD | 256 | 0 |
| HDD | 500 | 0 |
| HDD | 500 | 0 |
| SSD | 256 | 0 |



| | Id | Company | TypeName | Inches | Ram | Memory | OpSys | Weight | Price | Gpu_Brand | Gpu_Name | Cpu_Brand | Cpu_Name | Cpu_Speed | Resolution_Width | Resolution_Height | Touchscreen | Memory_Type | Primary_St |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0 | Apple | Ultrabook | 13.3 | 8 | 128GB SSD | MacOS | 1.37 | 71379 | Intel | Iris Plus Graphics 640 | Intel | Core i5 | 2.3 | 2560 | 1600 | 0 | SSD | 128 |
| 2 | 1 | Apple | Ultrabook | 13.3 | 8 | 128GB Flash Storage | MacOS | 1.34 | 47896 | Intel | HD Graphics 6000 | Intel | Core i5 | 1.8 | 1440 | 900 | 0 | Flash Storage | 128 |
| 3 | 2 | HP | Notebook | 15.6 | 8 | 256GB SSD | N/A | 1.86 | 30636 | Intel | HD Graphics 620 | Intel | Core i5 | 2.5 | 1920 | 1080 | 0 | SSD | 256 |
| 4 | 3 | Apple | Ultrabook | 15.4 | 16 | 512GB SSD | MacOS | 1.83 | 135195 | AMD | Radeon Pro 455 | Intel | Core i7 | 2.7 | 2880 | 1800 | 0 | SSD | 512 |
| 5 | 4 | Apple | Ultrabook | 13.3 | 8 | 256GB SSD | MacOS | 1.37 | 96096 | Intel | Iris Plus Graphics 650 | Intel | Core i5 | 3.1 | 2560 | 1600 | 0 | SSD | 256 |
| 6 | 5 | Acer | Notebook | 15.6 | 4 | 500GB HDD | Windows | 2.1 | 21312 | AMD | Radeon R5 | AMD | A9-Series 9420 | 3.0 | 1366 | 768 | 0 | HDD | 500 |
| 7 | 6 | Apple | Ultrabook | 15.4 | 16 | 256GB Flash Storage | MacOS | 2.04 | 114018 | Intel | Iris Pro Graphics | Intel | Core i7 | 2.2 | 2880 | 1800 | 0 | Flash Storage | 256 |
| 8 | 7 | Apple | Ultrabook | 13.3 | 8 | 256GB Flash Storage | MacOS | 1.34 | 61736 | Intel | HD Graphics 6000 | Intel | Core i5 | 1.8 | 1440 | 900 | 0 | Flash Storage | 256 |

***updating TB to GBs and deleting the memory column.****

```sql
--replace TB with GBs
SELECT primary_storage,
    CASE
        WHEN primary_storage <= 2 THEN primary_storage*1024 ELSE primary_storage END,
    secondary_storage,
    CASE
        WHEN secondary_storage <= 2 THEN secondary_storage*1024 ELSE secondary_storage END
FROM laptopdata;


UPDATE laptopdata
SET
    primary_storage =
        CASE
            WHEN primary_storage <= 2
            THEN primary_storage * 1024
            ELSE primary_storage
        END,
    secondary_storage =
        CASE
            WHEN secondary_storage <= 2
            THEN secondary_storage * 1024
            ELSE secondary_storage
        END;

--Drop memory column
ALTER TABLE laptopdata
DROP COLUMN Memory;

SELECT * FROM laptopdata
```

Before

| Memory_Type | Primary_Storage | Secondary_Storage |
|---|---|---|
| Hybrid | 128 | 1 |
| HDD | 500 | 0 |
| SSD | 256 | 0 |
| SSD | 256 | 0 |
| HDD | 1 | 0 |
| Flash Storage | 128 | 0 |
| SSD | 256 | 0 |
| SSD | 256 | 256 |
| HDD | 1 | 0 |
| Flash Storage | 64 | 0 |
| Flash Storage | 32 | 0 |

After

| Memory_Type | Primary_Storage | Secondary_Storage |
|---|---|---|
| Hybrid | 128 | 1024 |
| HDD | 500 | 0 |
| SSD | 256 | 0 |
| SSD | 256 | 0 |
| HDD | 1024 | 0 |
| Flash Storage | 128 | 0 |
| SSD | 256 | 0 |
| SSD | 256 | 256 |
| HDD | 1024 | 0 |
| Flash Storage | 64 | 0 |
| Flash Storage | 32 | 0 |

Results Messages

| | Id | Company | TypeName | Inches | Ram | OpSys | Weight | Price | Gpu_Brand | Gpu_Name | Cpu_Brand | Cpu_Name | Cpu_Speed | Resolution_Width | Resolution_Height | Touchscreen | Memory_Type | Primary_Storage | Secondary_St |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0 | Apple | Ultrabook | 13.3 | 8 | MacOS | 1.37 | 71379 | Intel | Iris Plus Graphics 640 | Intel | Core i5 | 2.3 | 2560 | 1600 | 0 | SSD | 128 | 0 |
| 2 | 1 | Apple | Ultrabook | 13.3 | 8 | MacOS | 1.34 | 47896 | Intel | HD Graphics 6000 | Intel | Core i5 | 1.8 | 1440 | 900 | 0 | Flash Storage | 128 | 0 |
| 3 | 2 | HP | Notebook | 15.6 | 8 | N/A | 1.86 | 30636 | Intel | HD Graphics 620 | Intel | Core i5 | 2.5 | 1920 | 1080 | 0 | SSD | 256 | 0 |
| 4 | 3 | Apple | Ultrabook | 15.4 | 16 | MacOS | 1.83 | 135195 | AMD | Radeon Pro 455 | Intel | Core i7 | 2.7 | 2880 | 1800 | 0 | SSD | 512 | 0 |
| 5 | 4 | Apple | Ultrabook | 13.3 | 8 | MacOS | 1.37 | 96096 | Intel | Iris Plus Graphics 650 | Intel | Core i5 | 3.1 | 2560 | 1600 | 0 | SSD | 256 | 0 |
| 6 | 5 | Acer | Notebook | 15.6 | 4 | Windows | 2.1 | 21312 | AMD | Radeon R5 | AMD | A9-Series 9420 | 3.0 | 1366 | 768 | 0 | HDD | 500 | 0 |
| 7 | 6 | Apple | Ultrabook | 15.4 | 16 | MacOS | 2.04 | 114018 | Intel | Iris Pro Graphics | Intel | Core i7 | 2.2 | 2880 | 1800 | 0 | Flash Storage | 256 | 0 |
| 8 | 7 | Apple | Ultrabook | 13.3 | 8 | MacOS | 1.34 | 61736 | Intel | HD Graphics 6000 | Intel | Core i5 | 1.8 | 1440 | 900 | 0 | Flash Storage | 256 | 0 |
| 9 | 8 | Asus | Ultrabook | 14.0 | 16 | Windows | 1.3 | 79654 | Nvidia | GeForce MX150 | Intel | Core i7 | 1.8 | 1920 | 1080 | 0 | SSD | 512 | 0 |
| 10 | 9 | Acer | Ultrabook | 14.0 | 8 | Windows | 1.6 | 41026 | Intel | UHD Graphics 620 | Intel | Core i5 | 1.6 | 1920 | 1080 | 0 | SSD | 256 | 0 |

10. Revisited the Gpu_Name column. There are 106 distinct Gpu_Name values and it will not result in meaningful Categorical data analysis. So dropping this column.

```
1  --drop Gpu_Name column as there are too many categorical columns
2  -- for meaning ful ML prediction
3  ALTER TABLE laptopdata
4  DROP COLUMN gpu_name;
5
6  SELECT * FROM laptopdata
```

Before

| Gpu_Brand | Gpu_Name |
|---|---|
| Nvidia | GeForce GTX 1050 |
| AMD | Radeon R2 |
| Intel | UHD Graphics 620 |
| Intel | HD Graphics 620 |
| Intel | HD Graphics 520 |
| Intel | HD Graphics 6000 |
| AMD | Radeon 530 |
| Intel | UHD Graphics 620 |
| Nvidia | GeForce 930MX |
| Intel | HD Graphics |
| AMD | Radeon R2 |
| AMD | Radeon 530 |

Results Messages

| | Id | Company | TypeName | Inches | Ram | OpSys | Weight | Price | Gpu_Brand | Cpu_Brand | Cpu_Name | Cpu_Speed | Resolution_Width | Resolution_Height | Touchscreen | Memory_Type | Primary_Storage | Secondary_Storage |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0 | Apple | Ultrabook | 13.3 | 8 | MacOS | 1.37 | 71379 | Intel | Intel | Core i5 | 2.3 | 2560 | 1600 | 0 | SSD | 128 | 0 |
| 2 | 1 | Apple | Ultrabook | 13.3 | 8 | MacOS | 1.34 | 47896 | Intel | Intel | Core i5 | 1.8 | 1440 | 900 | 0 | Flash Storage | 128 | 0 |
| 3 | 2 | HP | Notebook | 15.6 | 8 | N/A | 1.86 | 30636 | Intel | Intel | Core i5 | 2.5 | 1920 | 1080 | 0 | SSD | 256 | 0 |
| 4 | 3 | Apple | Ultrabook | 15.4 | 16 | MacOS | 1.83 | 135195 | AMD | Intel | Core i7 | 2.7 | 2880 | 1800 | 0 | SSD | 512 | 0 |
| 5 | 4 | Apple | Ultrabook | 13.3 | 8 | MacOS | 1.37 | 96096 | Intel | Intel | Core i5 | 3.1 | 2560 | 1600 | 0 | SSD | 256 | 0 |
| 6 | 5 | Acer | Notebook | 15.6 | 4 | Windows | 2.1 | 21312 | AMD | AMD | A9-Series 9420 | 3.0 | 1366 | 768 | 0 | HDD | 500 | 0 |
| 7 | 6 | Apple | Ultrabook | 15.4 | 16 | MacOS | 2.04 | 114018 | Intel | Intel | Core i7 | 2.2 | 2880 | 1800 | 0 | Flash Storage | 256 | 0 |
| 8 | 7 | Apple | Ultrabook | 13.3 | 8 | MacOS | 1.34 | 61736 | Intel | Intel | Core i5 | 1.8 | 1440 | 900 | 0 | Flash Storage | 256 | 0 |

11. After the cleanup the table schema looks like this