```
create database Lap;
use lap;
select * from laptop;

#upper lower###########33
select company,upper(company),lower(company) from laptop;
select company, concat( company,'  ', TypeName) from laptop;
select company, concat_ws('_', company,TypeName) from laptop;
select company, substr( company,1,5) from laptop;

#################################################3
#create copy data#####################3
create table laptops_backup like laptop;
insert into laptops_backup
select * from laptop;

#check number of rows###################
select count(*) from laptop;

#check memory consumption for references#################
select DATA_LENGTH/1024 from information_schema.Tables
where TABLE_SCHEMA = 'Lap' and TABLE_NAME ='laptop';

# DROP NON-IMPORTANT COLUMN#################
select * from laptop;
ALTER TABLE laptop DROP COLUMN `Unnamed: 0`; #backticks
select * from laptop;

#drop null values############################

SELECT * FROM laptop
WHERE Company IS NULL AND TypeName IS NULL AND Inches 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;
select count(*)from laptop;
#################################################################3
#DELETE FROM laptops
#WHERE `index` IN (SELECT `index` FROM laptops
#WHERE Company IS NULL AND TypeName IS NULL AND Inches 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);##

#######################################################
#drop duplicates #min first occurance distinct window also use###################
SELECT * FROM laptop
WHERE `index` NOT IN (
```

```
  SELECT MIN(`index`)
  FROM laptop
  GROUP BY Company, TypeName, Inches, ScreenResolution, Cpu, Ram, Memory, Gpu,
OpSys, Weight, Price
);
################################33#########################3
select * from laptop;
ALTER TABLE laptop ADD COLUMN `index` INT NOT NULL AUTO_INCREMENT UNIQUE;

#clean ram -> change col datatypes#######################
select distinct company from laptop;
select distinct TypeName from laptop;
alter table laptop modify column Inches decimal(10,1);

###cpu#################33
select replace(Ram,'GB','') from laptop;

UPDATE laptop
SET Ram = REPLACE(Ram, 'GB', '');

Alter table laptop modify column Ram integer;

###weight####################3
select replace(weight,'kg','') from laptop;
UPDATE laptop
SET weight = REPLACE(weight, 'kg', '');

select round(Price) from laptop;
UPDATE laptop
SET Price = round(Price);
Alter table laptop modify column Price integer;
select * from laptop;

#####################################
SELECT DISTINCT OpSys FROM laptops;

-- mac
-- windows
-- linux
-- no os
-- Android chrome(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' #column name
FROM laptop;

UPDATE laptop
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;

########gpu##################3
SELECT * FROM laptop;

ALTER TABLE laptop
ADD COLUMN gpu_brand VARCHAR(255) AFTER Gpu,
ADD COLUMN gpu_name VARCHAR(255) AFTER gpu_brand;
####new column
#SUBSTRING_INDEX(string, delimiter, count)
UPDATE laptop
SET gpu_brand = SUBSTRING_INDEX(Gpu, ' ', 1);

UPDATE laptop
SET gpu_name = TRIM(REPLACE(Gpu, gpu_brand, ''));


######drop##################
ALTER TABLE laptop DROP COLUMN Gpu;

#########cpu#####################
ALTER TABLE laptop
ADD COLUMN cpu_brand VARCHAR(255) AFTER Cpu,
ADD COLUMN cpu_name VARCHAR(255) AFTER cpu_brand,
ADD COLUMN cpu_speed DECIMAL(10,1) AFTER cpu_name;

####remove brandddd##################
UPDATE laptop
SET cpu_brand = SUBSTRING_INDEX(Cpu, ' ', 1);

##cast to convert decimal ################
SELECT Cpu
FROM laptop
WHERE Cpu NOT LIKE '%GHz';


UPDATE laptop
```

```
SET cpu_speed = CAST(REPLACE(SUBSTRING_INDEX(Cpu, ' ', -1), 'GHz', '') AS
DECIMAL(10,2))
WHERE Cpu LIKE '%GHz';


UPDATE laptop
SET cpu_name = TRIM(
   REPLACE(
     REPLACE(Cpu, cpu_brand, ''),
     SUBSTRING_INDEX(REPLACE(Cpu, cpu_brand, ''), ' ', -1),
     ''
   )
);
ALTER TABLE laptop DROP COLUMN Cpu;

####screen resolution#####
SELECT ScreenResolution,
SUBSTRING_INDEX(SUBSTRING_INDEX(ScreenResolution,' ',-1),'x',1),
SUBSTRING_INDEX(SUBSTRING_INDEX(ScreenResolution,' ',-1),'x',-1)
FROM laptop;


ALTER TABLE laptop
ADD COLUMN resolution_width INTEGER AFTER ScreenResolution,
ADD COLUMN resolution_height INTEGER AFTER resolution_width;

UPDATE laptop
SET
  resolution_width = SUBSTRING_INDEX(SUBSTRING_INDEX(ScreenResolution,'
',-1),'x',1),
  resolution_height = SUBSTRING_INDEX(SUBSTRING_INDEX(ScreenResolution,'
',-1),'x',-1);

###touch screen##############################33
ALTER TABLE laptop
ADD COLUMN touchscreen INTEGER AFTER resolution_height;

SELECT ScreenResolution LIKE '%Touch%' FROM laptop;

UPDATE laptop
SET touchscreen = ScreenResolution LIKE '%Touch%';

ALTER TABLE laptops
DROP COLUMN ScreenResolution;

####cpu###############
SELECT cpu_name,
SUBSTRING_INDEX(TRIM(cpu_name), ' ', 2)
```

```
FROM laptop;

UPDATE laptop
SET cpu_name = SUBSTRING_INDEX(TRIM(cpu_name), ' ', 2);

SELECT DISTINCT cpu_name FROM laptop;


################memory3333333#################################

ALTER TABLE laptop
ADD COLUMN memory_type VARCHAR(255) AFTER Memory,
ADD COLUMN primary_storage INTEGER AFTER memory_type,
ADD COLUMN secondary_storage INTEGER AFTER primary_storage;


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 laptop;

UPDATE laptop
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%' AND Memory LIKE '%HDD%' THEN 'Hybrid'
   WHEN Memory LIKE '%Flash Storage%' THEN 'Flash Storage'
   WHEN Memory LIKE '%Hybrid%' THEN 'Hybrid'
   ELSE NULL
END;

SELECT Memory,
REGEXP_SUBSTR(SUBSTRING_INDEX(Memory,'+',1),'[0-9]+'),
CASE WHEN Memory LIKE '%+%' THEN
REGEXP_SUBSTR(SUBSTRING_INDEX(Memory,'+',-1),'[0-9]+') ELSE 0 END
FROM laptop;

UPDATE laptop
SET primary_storage = REGEXP_SUBSTR(SUBSTRING_INDEX(Memory,'+',1),'[0-9]+'),
   secondary_storage = CASE
```

```
        WHEN Memory LIKE '%+%' THEN
REGEXP_SUBSTR(SUBSTRING_INDEX(Memory,'+',-1),'[0-9]+')
        ELSE 0
    END;


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 laptop;

UPDATE laptop
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;

ALTER TABLE laptop DROP COLUMN gpu_name;
```