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;