use task\_37;

CREATE TABLE laptops\_backup LIKE laptopdata;

INSERT INTO laptops\_backup

SELECT \* FROM laptopdata;

SELECT \* FROM information\_schema.TABLES

WHERE TABLE\_SCHEMA = 'task\_37'

AND TABLE\_NAME ='laptopdata';

SELECT DATA\_LENGTH/1024 FROM information\_schema.TABLES

WHERE TABLE\_SCHEMA = 'task\_37'

AND TABLE\_NAME ='laptopdata';

ALTER TABLE laptopdata DROP COLUMN `Unnamed: 0`;

SELECT \* FROM task\_37.laptopdata;

SELECT \* FROM laptopdata

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;

DELETE FROM task\_37.laptopdata

WHERE `index` IN (

SELECT \* FROM (

SELECT `index` FROM task\_37.laptopdata

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

) AS t

);

SELECT COUNT(\*) FROM laptopdata;

ALTER TABLE laptopdata MODIFY COLUMN Inches DECIMAL(10,1);

SELECT Inches FROM laptopdata;

SELECT COUNT(\*) FROM laptopdata WHERE Inches = '';

SELECT COUNT(\*) FROM laptopdata WHERE Inches REGEXP '[^0-9.]';

SELECT COUNT(\*) FROM laptopdata WHERE Inches REGEXP '[^0-9.]' OR Inches LIKE '%.%.' OR Inches LIKE '%..%';

SELECT COUNT(\*) FROM laptopdata WHERE Inches = '' OR Inches IS NULL OR

LENGTH(TRIM(BOTH '0123456789.' FROM Inches)) > 0;

UPDATE laptopdata SET Inches = NULL WHERE Inches = '';

UPDATE laptopdata SET Inches = NULL WHERE Inches REGEXP '[^0-9.]';

SELECT \* FROM laptopdata;

UPDATE laptopdata l1

SET Ram = (SELECT REPLACE(l2.Ram, 'GB', '') FROM laptopdata l2 WHERE l2.index = l1.index);

CREATE TEMPORARY TABLE tmp\_laptopdata AS SELECT \* FROM laptops\_backup;

UPDATE tmp\_laptopdata SET Ram = REPLACE(Ram, 'GB', '');

UPDATE laptopdata SET Ram = (SELECT Ram FROM tmp\_laptopdata WHERE tmp\_laptopdata.index = laptopdata.index);

DROP TEMPORARY TABLE tmp\_laptopdata;

SELECT \* FROM laptopdata;

ALTER TABLE laptopdata MODIFY COLUMN Ram INTEGER;

SELECT DATA\_LENGTH/1024 FROM information\_schema.TABLES

WHERE TABLE\_SCHEMA = 'task\_37'

AND TABLE\_NAME = 'laptopdata';

UPDATE laptopdata l1

SET Weight = (SELECT REPLACE(Weight,'kg','')

FROM laptopdata l2 WHERE l2.index = l1.index);

CREATE TEMPORARY TABLE tmp\_laptopdata AS SELECT \* FROM laptops\_backup;

UPDATE tmp\_laptopdata SET Weight = REPLACE(Weight, 'kg', '');

UPDATE laptopdata SET Weight = (SELECT Weight FROM tmp\_laptopdata WHERE tmp\_laptopdata.index = laptopdata.index);

DROP TEMPORARY TABLE tmp\_laptopdata;

SELECT \* FROM laptopdata;

CREATE TEMPORARY TABLE temp\_laptopdata AS SELECT \* FROM laptopdata;

UPDATE temp\_laptopdata

SET Price = ROUND(Price);

UPDATE laptopdata

INNER JOIN temp\_laptopdata

ON laptopdata.index = temp\_laptopdata.index

SET laptopdata.Price = temp\_laptopdata.Price;

DROP TEMPORARY TABLE temp\_laptopdata;

ALTER TABLE laptopdata MODIFY COLUMN Price INTEGER;

SELECT \* FROM laptopdata;

SELECT DISTINCT OpSys FROM laptopdata;

-- 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'

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 \* FROM laptopdata;

ALTER TABLE laptopdata

ADD COLUMN gpu\_brand VARCHAR(255) AFTER Gpu,

ADD COLUMN gpu\_name VARCHAR(255) AFTER gpu\_brand;

UPDATE laptopdata l1

SET gpu\_brand = SUBSTRING\_INDEX(Gpu, ' ', 1);

UPDATE laptopdata l1

SET gpu\_name= TRIM(REPLACE(Gpu,gpu\_brand,''));

SELECT \* FROM laptopdata;

ALTER TABLE laptopdata DROP COLUMN Gpu;

ALTER TABLE laptopdata

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;

SELECT \* FROM laptopdata;

UPDATE laptopdata l1

JOIN (SELECT `index`, SUBSTRING\_INDEX(Cpu, ' ', 1) AS cpu\_brand FROM laptopdata) l2

ON l1.`index` = l2.`index`

SET l1.cpu\_brand = l2.cpu\_brand;

SELECT \* FROM laptopdata;

UPDATE laptopdata

SET cpu\_name = (SELECT

REPLACE(REPLACE(Cpu,cpu\_brand,''),SUBSTRING\_INDEX(REPLACE(Cpu,cpu\_brand,''),' ',-1),''));

SELECT \* FROM laptopdata;

ALTER TABLE laptopdata DROP COLUMN Cpu;

SELECT ScreenResolution,

SUBSTRING\_INDEX(SUBSTRING\_INDEX(ScreenResolution,' ',-1),'x',1),

SUBSTRING\_INDEX(SUBSTRING\_INDEX(ScreenResolution,' ',-1),'x',-1)

FROM laptopdata;

ALTER TABLE laptopdata

ADD COLUMN resolution\_width INTEGER AFTER ScreenResolution,

ADD COLUMN resolution\_height INTEGER AFTER resolution\_width;

SELECT \* FROM laptopdata;

UPDATE laptopdata

SET resolution\_width = SUBSTRING\_INDEX(SUBSTRING\_INDEX(ScreenResolution,' ',-1),'x',1),

resolution\_height = SUBSTRING\_INDEX(SUBSTRING\_INDEX(ScreenResolution,' ',-1),'x',-1);

SELECT \* FROM laptopdata;

ALTER TABLE laptopdata

ADD COLUMN touchscreen INTEGER AFTER resolution\_height;

SELECT ScreenResolution LIKE '%Touch%' FROM laptopdata;

UPDATE laptopdata

SET touchscreen = ScreenResolution LIKE '%Touch%';

SELECT \* FROM laptopdata;

ALTER TABLE laptopdata

DROP COLUMN ScreenResolution;

SELECT \* FROM laptopdata;

SELECT cpu\_name,

SUBSTRING\_INDEX(TRIM(cpu\_name),' ',2)

FROM laptopdata;

UPDATE laptopdata

SET cpu\_name = SUBSTRING\_INDEX(TRIM(cpu\_name),' ',2);

SELECT \* FROM laptopdata;

SELECT DISTINCT cpu\_name FROM laptopdata;

SELECT Memory FROM laptopdata;

ALTER TABLE laptopdata

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 FROM laptopdata;

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;

SELECT \* FROM laptopdata;

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 laptopdata;

UPDATE laptopdata

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 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;

SELECT \* FROM laptopdata;

ALTER TABLE laptopdata DROP COLUMN gpu\_name;