

```

select distinct * from hospital;

UPDATE hospital
SET AdmissionDate =
CASE
    WHEN AdmissionDate LIKE '%-%-%' THEN STR_TO_DATE(AdmissionDate, '%d-%b-%Y')
    WHEN AdmissionDate LIKE '%/%/%' THEN STR_TO_DATE(AdmissionDate, '%c/%e/%Y')
    ELSE NULL
END;
UPDATE hospital
SET dischargedate =
CASE
    WHEN dischargedate LIKE '%/%/%'
        THEN STR_TO_DATE(dischargedate, '%d/%m/%Y')
    ELSE NULL
END;
UPDATE hospital
SET dischargedate = STR_TO_DATE(dischargedate, '%m/%d/%Y');

ALTER TABLE hospital
MODIFY COLUMN AdmissionDate DATE;

UPDATE hospital
SET Status = LOWER(Status);

UPDATE hospital
SET Doctor = 'Unknown',
    Name = 'Unknown'
WHERE doctor= "" or name IS NULL;
-----
-- Compare the row counts before and after cleaning
SELECT
    (SELECT COUNT(*) FROM hospital) AS row_count1,
    (SELECT COUNT(*) FROM hospital_cleaned) AS row_count2;
-- compare missing values
SELECT
    (SELECT COUNT(*) FROM hospital WHERE Name = "") AS row_count1,
    (SELECT COUNT(*) FROM hospital_cleaned WHERE name = "") AS row_count2;
-----
-- average age of patients for each diagnosis?
select Diagnosis ,avg(age) from hospital_cleaned group by Diagnosis;

-- Which department has the highest number of admitted patients?
select department, count(status) from hospital_cleaned where Status = 'admitted' group by Department limit 1;

-- How many patients have been discharged per month?
select monthname(dischargedate) as month,count(DischargeDate) as num_of_patients from hospital_cleaned group by DischargeDate;

-- What is the most common diagnosis among patients?
select Diagnosis,count(diagnosis) as count from hospital_cleaned group by Diagnosis order by count desc;

-- Which doctor has treated the most patients?
select doctor, count(doctor) as count_patients from hospital_cleaned group by doctor;

```