

```
SELECT *
FROM Healthcare_Dataset

SELECT *
FROM Healthcare_Dataset
WHERE [Name] LIKE '% %'
      OR [Name] IS NULL
      OR [Name] = ''
--Checked for Empty Values.

DELETE FROM Healthcare_Dataset
WHERE [Name] IS NULL

ALTER TABLE Healthcare_Dataset
DROP COLUMN [Room Number], [Test Results]
--Deleted some Irrelevant Columns.

UPDATE Healthcare_Dataset
SET [Name] = UPPER([Name])
--Formatted the patient Names into Uppercase; since most were not in a uniform format.

UPDATE Healthcare_Dataset
SET [Billing Amount] = ROUND([Billing Amount], 0)
--Removed the several decimals by Converting Billing Amount to the nearest Whole Number.

UPDATE Healthcare_Dataset
SET [Name] = LTRIM(RTRIM([Name])),
    [Doctor] = LTRIM(RTRIM([Doctor])),
    [Hospital] = LTRIM(RTRIM([Hospital]))
--Remove leading and trailing spaces from the columns.

ALTER TABLE Healthcare_Dataset
ADD AgeGroup VARCHAR(20)
UPDATE Healthcare_Dataset
SET AgeGroup = CASE
    WHEN Age < 18 THEN 'Child'
    WHEN Age BETWEEN 18 AND 35 THEN 'Young Adult'
    WHEN Age BETWEEN 36 AND 55 THEN 'Adult'
    ELSE 'Senior'
END;
--Added an extra column 'AgeGroup'; to help with further queries.

--QUERIES FOR THE DATASET:
--For the Gender and Medical Conditions:
SELECT [Gender], [Medical Condition], COUNT(*) AS ConditionCount
FROM Healthcare_Dataset
GROUP BY [Gender], [Medical Condition]
ORDER BY [Medical Condition] ASC

--For the Medical Conditions Across Age Groups:
SELECT [AgeGroup], [Medical Condition], COUNT(*) AS ConditionCount
FROM Healthcare_Dataset
```

---

```
GROUP BY [AgeGroup], [Medical Condition]
ORDER BY ConditionCount DESC
```

```
--For the Doctors Preferred for Specific Conditions:
```

```
SELECT [Doctor], [Medical Condition], COUNT(*) AS PatientCount
FROM Healthcare_Dataset
GROUP BY [Doctor], [Medical Condition]
HAVING COUNT(*) >= 5
```