Practice Queries - sqldf in R

1. Top 5 Hospitals by Visits

```
sqldf("SELECT facility_name, COUNT(*) as visits FROM hospital GROUP BY facility_name ORDER BY
visits DESC LIMIT 5")
```

2. Common Procedure Types

```
sqldf("SELECT    ccsr_procedure_description,    COUNT(*)    as    freq    FROM    hospital    GROUP    BY
ccsr_procedure_description ORDER BY freq DESC LIMIT 5")
```

3. Average Length of Stay by Diagnosis

```
sqldf("SELECT ccsr_diagnosis_description, AVG(length_of_stay) as avg_days FROM hospital GROUP BY
ccsr_diagnosis_description ORDER BY avg_days DESC LIMIT 5")
```

4. Total Charges by County

```
sqldf("SELECT c.hospital_county, SUM(h.total_charges) AS total_charges FROM hospital h INNER JOIN
counties c ON h.hospital_county = c.county_id GROUP BY c.hospital_county ORDER BY total_charges
DESC")
```

5. Hospital Discharges by Age Group

```
sqldf("SELECT    age_group,    COUNT(*)    as    discharges    FROM    hospital    GROUP    BY    age_group    ORDER    BY
discharges DESC")
```

6. Charges by Race and Age Group

```
sqldf("SELECT    race,    age_group,    AVG(total_charges)    as    avg_charge    FROM    hospital    GROUP    BY    race,
age_group ORDER BY avg_charge DESC")
```

7. Linking Hospital and County Info

```
sqldf("SELECT    h.facility_name,    c.hospital_county    FROM    hospital    h    LEFT    JOIN    counties    c    ON
h.hospital_county = c.county_id LIMIT 10")
```

8. Procedures in the Bronx

```
sqldf("SELECT ccsr_procedure_description, COUNT(*) as freq FROM hospital WHERE hospital_county =
'Bronx' GROUP BY ccsr_procedure_description ORDER BY freq DESC LIMIT 5")
```

9. Most Expensive Diagnoses

```
sqldf("SELECT ccsr_diagnosis_description, AVG(total_charges) as avg_charge FROM hospital GROUP BY
ccsr_diagnosis_description ORDER BY avg_charge DESC LIMIT 5")
```

10. Hospital Visits by Race in Queens

```
sqldf("SELECT race, COUNT(*) as visits FROM hospital WHERE hospital_county = 'Queens' GROUP BY
race ORDER BY visits DESC")
```

Practice Queries - sqldf in R

11. High Frequency Procedures (10,000+ cases)

```
sqldf("SELECT   ccsr_procedure_description,   COUNT(*)   as   n   FROM   hospital   GROUP   BY
ccsr_procedure_description HAVING n > 10000 ORDER BY n DESC")
```

12. Age Distribution by Facility

```
sqldf("SELECT  facility_name,  age_group,  COUNT(*)  as  n  FROM  hospital  GROUP  BY  facility_name,
age_group ORDER BY facility_name")
```

11. High Frequency Procedures (10,000+ cases)

```
sqldf("SELECT   ccsr_procedure_description,   COUNT(*)   as   n   FROM   hospital   GROUP   BY
ccsr_procedure_description HAVING n > 10000 ORDER BY n DESC")
```