

## SQL Source

### Exercise 1

**How many facilities are represented in this table? How many insurance carriers?**

```
SELECT
    COUNT (DISTINCT facility_id) AS facility_count
    , COUNT (DISTINCT insurance_carrier_id) AS insurance_count
FROM imaging_coverage
;
```

**Which 5 states have the most facilities?**

```
SELECT
    state
    , COUNT (DISTINCT facility_id) AS facility_count
FROM imaging_coverage
GROUP BY state
ORDER BY facility_count DESC
LIMIT 5
;
```

**Which 10 cities have the most?**

```
SELECT
    city
    , COUNT (DISTINCT facility_id) AS facility_count
FROM imaging_coverage
GROUP BY city
ORDER BY facility_count DESC
LIMIT 10
;
```

**Which 5 states have the highest proportion of cost-efficient facilities?**

WITH cef\_state AS (

```
SELECT
    state
    , facility_id
    , COUNT(DISTINCT insurance_carrier_id) AS count_ins
    , COUNT(DISTINCT CASE WHEN cost_efficient_flag = 'Yes' THEN
insurance_carrier_id ELSE NULL END) AS
count_cost_efficient_ins
    , count_cost_efficient_ins * 1.0/count_ins AS
percent_cost_efficient_ins
```

```

        , CASE WHEN percent_cost_efficient_ins >= 0.75 THEN 1 ELSE 0
        END AS cef

FROM imaging_coverage
GROUP BY state, facility_id
)

SELECT

    state

    , COUNT(DISTINCT CASE WHEN cef = 1 THEN facility_id ELSE
    NULL END) AS count_cef

    , COUNT(DISTINCT facility_id) AS count_facilities
    , count_cef*1.0/count_facilities AS prop_cef

FROM cef_state
GROUP BY state
ORDER BY prop_cef DESC
LIMIT 5
;

```

**How many facilities have at least 10 insurance carrier IDs that are all labeled cost efficient?**

```

WITH ceic AS (
    SELECT
        facility_id
        , COUNT(DISTINCT CASE WHEN cost_efficient_flag = 'Yes' THEN
        insurance_carrier_id ELSE NULL END) AS count_cost_efficient_ins
    FROM imaging_coverage
    GROUP BY facility_id
)
SELECT COUNT(DISTINCT facility_id)
FROM ceic
WHERE count_cost_efficient_ins >= 10
;

```

## Exercise 2

How many facilities in the Houston region are considered “cost efficient” (75% of the insurance carrier IDs are considered cost efficient)?

```

WITH cef_houston AS (
    SELECT
        facility_id
        , COUNT(DISTINCT insurance_carrier_id) AS count_ins

```

```

        ,COUNT(DISTINCT CASE WHEN cost_efficient_flag = 'Yes' THEN
insurance_carrier_id ELSE NULL END) AS count_cost_efficient_ins
        ,count_cost_efficient_ins * 1.0/count_ins AS percent_cost_efficient_ins
        ,CASE WHEN percent_cost_efficient_ins >= 0.75 THEN 1 ELSE 0 END AS
cost_efficient_facility
FROM imaging_coverage
WHERE city = 'Houston'
GROUP BY facility_id
)
SELECT COUNT(DISTINCT facility_id)
FROM cef_houston
WHERE cost_efficient_facility = 1
;

```

**What proportion of Houston facilities are considered “cost efficient” by this definition?**

```

WITH cef_houston AS (
    SELECT
        facility_id
        ,COUNT(DISTINCT insurance_carrier_id) AS count_ins
        ,COUNT(DISTINCT CASE WHEN cost_efficient_flag = 'Yes' THEN
insurance_carrier_id ELSE NULL END) AS count_cost_efficient_ins
        ,count_cost_efficient_ins*1.0/count_ins AS percent_cost_efficient_ins
        ,CASE WHEN percent_cost_efficient_ins >= 0.75 THEN 1 ELSE 0 END AS
cef
FROM imaging_coverage
WHERE city = 'Houston'
GROUP BY facility_id
)
SELECT
    COUNT(DISTINCT CASE WHEN cef = 1 THEN facility_id ELSE NULL
END) AS count_cef
    ,COUNT(DISTINCT facility_id) AS count_facilities
    ,count_cef*1.0/count_facilities AS prop_cef_houston
FROM cef_houston
;

```

**Which Houston facilities would you recommend sending Hendrix employees to for imaging services based on cost efficiency and having a wide variety of insurance carriers?**

```

WITH cef_houston_rank AS (
    SELECT
        facility_id
        ,COUNT(DISTINCT insurance_carrier_id) AS count_ins
        ,COUNT(DISTINCT CASE WHEN cost_efficient_flag = 'Yes' THEN
insurance_carrier_id ELSE NULL END) AS count_cost_efficient_ins

```

```
        ,count_cost_efficient_ins*1.0/count_ins AS percent_cost_efficient_ins
        ,CASE WHEN percent_cost_efficient_ins >=0.75 THEN 1 ELSE 0 END AS
cef
FROM imaging_coverage
WHERE city = 'Houston'
GROUP BY facility_id
)

SELECT
    facility_id
    , count_cost_efficient_ins
    , count_ins
    , percent_cost_efficient_ins
FROM cef_houston_rank
ORDER BY count_cost_efficient_ins DESC
;
```