```sql
/*

Medicare Nursing Facilities 2014 Data Cleaning and Exploration in SQL Queries

Skills Used: CTE, Aggregate Functions, Window Function, Conditional Function, Subquery

Platform: GCP BigQuery

*/

-------------------------------------------------------------------------------------------

-- DATA CLEANING IN SQL QUERIES

-- Preview the nursing_facilities_2014 table

SELECT
    *
FROM
    `bigquery-public-data.cms_medicare.nursing_facilities_2014` LIMIT 200;

-- Check if the provider_id column contains duplicate

SELECT
    COUNT(DISTINCT provider_id) as unique_provider_id,
    COUNT(*) as total_row_number
FROM
    `bigquery-public-data.cms_medicare.nursing_facilities_2014`;

-- Check the percentage of null in each column

SELECT
    COUNTIF(provider_id IS NULL)/COUNT(*)*100 as null_provider_id,
    COUNTIF(facility_name IS NULL)/COUNT(*)*100 as null_facility_name,
    COUNTIF(street_address IS NULL)/COUNT(*)*100 as null_street_address,
    COUNTIF(city IS NULL)/COUNT(*)*100 as null_city,
    COUNTIF(state IS NULL)/COUNT(*)*100 as null_state,
    COUNTIF(zip_code IS NULL)/COUNT(*)*100 as null_zip_code,
    COUNTIF(total_stays IS NULL)/COUNT(*)*100 as null_total_stays,
    COUNTIF(distinct_beneficiaries_per_provider IS NULL)/COUNT(*)*100 as null_distinct_beneficiaries_per_provider,
    COUNTIF(average_length_of_stays_days IS NULL)/COUNT(*)*100 as null_average_length_of_stays_days,
    COUNTIF(total_snf_charge_amount IS NULL)/COUNT(*)*100 as null_total_snf_charge_amount,
    COUNTIF(total_snf_medicare_allowed_amount IS NULL)/COUNT(*)*100 as null_total_snf_medicare_allowed_amount,
    COUNTIF(total_snf_medicare_payment_amount IS NULL)/COUNT(*)*100 as null_total_snf_medicare_payment_amount,
    COUNTIF(total_snf_medicare_standard_payment_amount IS NULL)/COUNT(*)*100 as null_total_snf_medicare_standard_payment_amount,
    COUNTIF(average_age IS NULL)/COUNT(*)*100 as null_average_age,
    COUNTIF(male_beneficiaries IS NULL)/COUNT(*)*100 as null_male_beneficiaries,
    COUNTIF(female_beneficiaries IS NULL)/COUNT(*)*100 as null_female_beneficiaries,
    COUNTIF(nondual_beneficiaries IS NULL)/COUNT(*)*100 as null_nondual_beneficiaries,
    COUNTIF(dual_beneficiaries IS NULL)/COUNT(*)*100 as null_dual_beneficiaries,
    COUNTIF(white_beneficiaries IS NULL)/COUNT(*)*100 as null_white_beneficiaries,
    COUNTIF(black_beneficiaries IS NULL)/COUNT(*)*100 as null_black_beneficiaries,
    COUNTIF(asian_pacific_islander_beneficiaries IS NULL)/COUNT(*)*100 as null_asian_pacific_islander_beneficiaries,
    COUNTIF(hispanic_beneficiaries IS NULL)/COUNT(*)*100 as null_hispanic_beneficiaries,
```

```
    COUNTIF(american_indian_or_alaska_native_beneficiaries IS NULL)/COUNT(*)*100 as null_american_indian_or_alaska_native_beneficiaries,
    COUNTIF(other_unknown_beneficiaries IS NULL)/COUNT(*)*100 as null_other_unknown_beneficiaries,
    COUNTIF(average_hcc_score IS NULL)/COUNT(*)*100 as null_average_hcc_score,
    COUNTIF(percent_of_beneficiaries_with_atrial_fibrillation IS NULL)/COUNT(*)*100 as null_percent_of_beneficiaries_with_atrial_fibrillation,
    COUNTIF(percent_of_beneficiaries_with_alzheimers IS NULL)/COUNT(*)*100 as null_percent_of_beneficiaries_with_alzheimers,
    COUNTIF(percent_of_beneficiaries_with_asthma IS NULL)/COUNT(*)*100 as null_percent_of_beneficiaries_with_asthma,
    COUNTIF(percent_of_beneficiaries_with_cancer IS NULL)/COUNT(*)*100 as null_percent_of_beneficiaries_with_cancer,
    COUNTIF(percent_of_beneficiaries_with_chf IS NULL)/COUNT(*)*100 as null_percent_of_beneficiaries_with_chf,
    COUNTIF(percent_of_beneficiaries_with_chronic_kidney_disease IS NULL)/COUNT(*)*100 as null_percent_of_beneficiaries_with_chronic_kidney_disease,
    COUNTIF(percent_of_beneficiaries_with_copd IS NULL)/COUNT(*)*100 as null_percent_of_beneficiaries_with_copd,
    COUNTIF(percent_of_beneficiaries_with_depression IS NULL)/COUNT(*)*100 as null_percent_of_beneficiaries_with_depression,
    COUNTIF(percent_of_beneficiaries_with_diabetes IS NULL)/COUNT(*)*100 as null_percent_of_beneficiaries_with_diabetes,
    COUNTIF(percent_of_beneficiaries_with_hyperlipidemia IS NULL)/COUNT(*)*100 as null_percent_of_beneficiaries_with_hyperlipidemia,
    COUNTIF(percent_of_beneficiaries_with_hypertension IS NULL)/COUNT(*)*100 as null_percent_of_beneficiaries_with_hypertension,
    COUNTIF(percent_of_beneficiaries_with_ihd IS NULL)/COUNT(*)*100 as null_percent_of_beneficiaries_with_ihd,
    COUNTIF(percent_of_beneficiaries_with_osteoporosis IS NULL)/COUNT(*)*100 as null_percent_of_beneficiaries_with_osteoporosis,
    COUNTIF(percent_of_beneficiaries_with_ra_oa IS NULL)/COUNT(*)*100 as null_percent_of_beneficiaries_with_ra_oa,
    COUNTIF(percent_of_beneficiaries_with_schizophrenia IS NULL)/COUNT(*)*100 as null_percent_of_beneficiaries_with_schizophrenia,
    COUNTIF(percent_of_beneficiaries_with_stroke IS NULL)/COUNT(*)*100 as null_percent_of_beneficiaries_with_stroke,
FROM
    `bigquery-public-data.cms_medicare.nursing_facilities_2014`;

-- Check the percentage of null in each column grouped by state

SELECT
    state AS state,
    COUNTIF(provider_id IS NULL)/COUNT(*)*100 as null_provider_id,
    COUNTIF(facility_name IS NULL)/COUNT(*)*100 as null_facility_name,
    COUNTIF(street_address IS NULL)/COUNT(*)*100 as null_street_address,
    COUNTIF(city IS NULL)/COUNT(*)*100 as null_city,
    COUNTIF(state IS NULL)/COUNT(*)*100 as null_state,
    COUNTIF(zip_code IS NULL)/COUNT(*)*100 as null_zip_code,
    COUNTIF(total_stays IS NULL)/COUNT(*)*100 as null_total_stays,
    COUNTIF(distinct_beneficiaries_per_provider IS NULL)/COUNT(*)*100 as null_distinct_beneficiaries_per_provider,
    COUNTIF(average_length_of_stays_days IS NULL)/COUNT(*)*100 as null_average_length_of_stays_days,
    COUNTIF(total_snf_charge_amount IS NULL)/COUNT(*)*100 as null_total_snf_charge_amount,
    COUNTIF(total_snf_medicare_allowed_amount IS NULL)/COUNT(*)*100 as null_total_snf_medicare_allowed_amount,
    COUNTIF(total_snf_medicare_payment_amount IS NULL)/COUNT(*)*100 as null_total_snf_medicare_payment_amount,
    COUNTIF(total_snf_medicare_standard_payment_amount IS NULL)/COUNT(*)*100 as null_total_snf_medicare_standard_payment_amount,
    COUNTIF(average_age IS NULL)/COUNT(*)*100 as null_average_age,
    COUNTIF(male_beneficiaries IS NULL)/COUNT(*)*100 as null_male_beneficiaries,
    COUNTIF(female_beneficiaries IS NULL)/COUNT(*)*100 as null_female_beneficiaries,
    COUNTIF(nondual_beneficiaries IS NULL)/COUNT(*)*100 as null_nondual_beneficiaries,
    COUNTIF(dual_beneficiaries IS NULL)/COUNT(*)*100 as null_dual_beneficiaries,
    COUNTIF(white_beneficiaries IS NULL)/COUNT(*)*100 as null_white_beneficiaries,
    COUNTIF(black_beneficiaries IS NULL)/COUNT(*)*100 as null_black_beneficiaries,
    COUNTIF(asian_pacific_islander_beneficiaries IS NULL)/COUNT(*)*100 as null_asian_pacific_islander_beneficiaries,
    COUNTIF(hispanic_beneficiaries IS NULL)/COUNT(*)*100 as null_hispanic_beneficiaries,
    COUNTIF(american_indian_or_alaska_native_beneficiaries IS NULL)/COUNT(*)*100 as null_american_indian_or_alaska_native_beneficiaries,
```

```sql
    COUNTIF(other_unknown_beneficiaries IS NULL)/COUNT(*)*100 as null_other_unknown_beneficiaries,
    COUNTIF(average_hcc_score IS NULL)/COUNT(*)*100 as null_average_hcc_score,
    COUNTIF(percent_of_beneficiaries_with_atrial_fibrillation IS NULL)/COUNT(*)*100 as null_percent_of_beneficiaries_with_atrial_f
ibrillation,
    COUNTIF(percent_of_beneficiaries_with_alzheimers IS NULL)/COUNT(*)*100 as null_percent_of_beneficiaries_with_alzheimers,
    COUNTIF(percent_of_beneficiaries_with_asthma IS NULL)/COUNT(*)*100 as null_percent_of_beneficiaries_with_asthma,
    COUNTIF(percent_of_beneficiaries_with_cancer IS NULL)/COUNT(*)*100 as null_percent_of_beneficiaries_with_cancer,
    COUNTIF(percent_of_beneficiaries_with_chf IS NULL)/COUNT(*)*100 as null_percent_of_beneficiaries_with_chf,
    COUNTIF(percent_of_beneficiaries_with_chronic_kidney_disease IS NULL)/COUNT(*)*100 as null_percent_of_beneficiaries_with_chron
ic_kidney_disease,
    COUNTIF(percent_of_beneficiaries_with_copd IS NULL)/COUNT(*)*100 as null_percent_of_beneficiaries_with_copd,
    COUNTIF(percent_of_beneficiaries_with_depression IS NULL)/COUNT(*)*100 as null_percent_of_beneficiaries_with_depression,
    COUNTIF(percent_of_beneficiaries_with_diabetes IS NULL)/COUNT(*)*100 as null_percent_of_beneficiaries_with_diabetes,
    COUNTIF(percent_of_beneficiaries_with_hyperlipidemia IS NULL)/COUNT(*)*100 as null_percent_of_beneficiaries_with_hyperlipidemi
a,
    COUNTIF(percent_of_beneficiaries_with_hypertension IS NULL)/COUNT(*)*100 as null_percent_of_beneficiaries_with_hypertension,
    COUNTIF(percent_of_beneficiaries_with_ihd IS NULL)/COUNT(*)*100 as null_percent_of_beneficiaries_with_ihd,
    COUNTIF(percent_of_beneficiaries_with_osteoporosis IS NULL)/COUNT(*)*100 as null_percent_of_beneficiaries_with_osteoporosis,
    COUNTIF(percent_of_beneficiaries_with_ra_oa IS NULL)/COUNT(*)*100 as null_percent_of_beneficiaries_with_ra_oa,
    COUNTIF(percent_of_beneficiaries_with_schizophrenia IS NULL)/COUNT(*)*100 as null_percent_of_beneficiaries_with_schizophrenia,
    COUNTIF(percent_of_beneficiaries_with_stroke IS NULL)/COUNT(*)*100 as null_percent_of_beneficiaries_with_stroke,
FROM
    `bigquery-public-data.cms_medicare.nursing_facilities_2014`
GROUP BY state;

-------------------------------------------------------------------------------------------

-- EXPLORATORY DATA ANALYSIS WITH SQL QUERIES

-- Number of nursing facility per state

SELECT
    state AS state,
    COUNT(*) AS number_of_nursing_facility
FROM
    `bigquery-public-data.cms_medicare.nursing_facilities_2014`
GROUP BY state
ORDER BY number_of_nursing_facility DESC;

-- Number of nursing facility per city in top 5 states

SELECT
    state,
    city,
    COUNT(city) as city_count
FROM (
    SELECT
    state,
    city,
    COUNT(city) OVER(
    PARTITION BY state
    ROWS BETWEEN UNBOUNDED PRECEDING AND UNBOUNDED FOLLOWING)
    AS total_state_count
    FROM
```

```
    `bigquery-public-data.cms_medicare.nursing_facilities_2014`
    )
    WHERE state IN ('TX', 'CA', 'OH', 'IL', 'PA')
GROUP BY state, city,total_state_count
ORDER BY total_state_count DESC, city_count DESC;

-- Demographic information: National average number of male and female beneficiaries, dual and non-dual beneficiaries and
-- age of beneficiaries

SELECT
    ROUND(AVG(average_age), 2) AS national_avg_age,
    ROUND(AVG(male_beneficiaries), 0) as national_avg_male_beneficiaries,
    ROUND(AVG(female_beneficiaries), 0) as national_avg_female_beneficiaries,
    ROUND(AVG(nondual_beneficiaries), 0) as national_avg_nondual_beneficiaries,
    ROUND(AVG(dual_beneficiaries), 0) as national_avg_dual_beneficiariess,
FROM
    `bigquery-public-data.cms_medicare.nursing_facilities_2014`;

-- Demographic information: Average number of male and female beneficiaries, dual and non-dual beneficiaries and age of
-- beneficiaries by state

SELECT
    state,
    ROUND(AVG(average_age), 2) AS state_avg_age,
    ROUND(AVG(male_beneficiaries), 0) as state_avg_male_beneficiaries,
    ROUND(AVG(female_beneficiaries), 0) as state_avg_female_beneficiaries,
    ROUND(AVG(nondual_beneficiaries), 0) as state_avg_nondual_beneficiaries,
    ROUND(AVG(dual_beneficiaries), 0) as state_avg_dual_beneficiariess,
FROM
    `bigquery-public-data.cms_medicare.nursing_facilities_2014`
GROUP BY state
ORDER BY state_avg_age DESC;

-- National average, minimum and maximum average length of stay in days in nursing facilities

SELECT
    ROUND(AVG(average_length_of_stays_days), 2) AS national_average_length_of_stays_days,
    MIN(average_length_of_stays_days) AS national_min_average_length_of_stays_days,
    MAX(average_length_of_stays_days) AS national_max_length_of_stays_days
FROM
    `bigquery-public-data.cms_medicare.nursing_facilities_2014`;

-- Average, minimum and maximum average length of stay in days in nursing facilities for each state

SELECT
    state,
    ROUND(AVG(average_length_of_stays_days), 2) AS state_average_length_of_stays_days,
    MIN(average_length_of_stays_days) AS state_min_average_length_of_stays_days,
    MAX(average_length_of_stays_days) AS state_max_length_of_stays_days
FROM
    `bigquery-public-data.cms_medicare.nursing_facilities_2014`
GROUP BY state
ORDER BY state_average_length_of_stays_days DESC;
```

```sql
-- States with average length of stay in days above national average

SELECT
    state,
    state_average_length_of_stays_days
FROM (
    WITH new_nursing_facilities AS
    (SELECT
    state,
    average_length_of_stays_days,
    AVG(average_length_of_stays_days) OVER () AS national_average_length_of_stays_days,
    FROM
    `bigquery-public-data.cms_medicare.nursing_facilities_2014`)
    SELECT
    state,
    ROUND(AVG(average_length_of_stays_days), 2) AS state_average_length_of_stays_days,
    IF(AVG(average_length_of_stays_days) > national_average_length_of_stays_days, 'true', 'false') AS states_above_national_avg
    FROM new_nursing_facilities
    GROUP BY state, national_average_length_of_stays_days
    )
    WHERE states_above_national_avg = 'true'
    GROUP BY state, state_average_length_of_stays_days
    ORDER BY state_average_length_of_stays_days;

-- National total, average, minimum and maximum medicare standard payment amount for nursing facilities

SELECT
    SUM(total_snf_medicare_standard_payment_amount) AS national_total_medicare_standard_payment,
    ROUND(AVG(total_snf_medicare_standard_payment_amount), 2) AS national_avg_medicare_standard_payment,
    MIN(total_snf_medicare_standard_payment_amount) AS national_min_medicare_standard_payment,
    MAX(total_snf_medicare_standard_payment_amount) AS national_max_medicare_standard_payment
FROM
    `bigquery-public-data.cms_medicare.nursing_facilities_2014`;

-- State total, average, minimum and maximum medicare standard payment amount for nursing facilities

SELECT
    state,
    ROUND(AVG(total_snf_medicare_standard_payment_amount), 2) AS state_avg_medicare_standard_payment,
    MIN(total_snf_medicare_standard_payment_amount) AS state_min_medicare_standard_payment,
    MAX(total_snf_medicare_standard_payment_amount) AS state_max_medicare_standard_payment,
    SUM(total_snf_medicare_standard_payment_amount) AS state_total_medicare_standard_payment,
FROM
    `bigquery-public-data.cms_medicare.nursing_facilities_2014`
GROUP BY state
ORDER BY state_avg_medicare_standard_payment DESC;

-- National average, minimum and maximum total stays in nursing facilities and medicare payment amount per stay

SELECT
    ROUND(AVG(total_stays), 2) AS national_average_total_stays,
    MIN(total_stays) AS national_min_total_stays,
    MAX(total_stays) AS national_max_total_stays,
    SUM(total_stays) AS national_year_total_stays,
```

```sql
    SUM(total_snf_medicare_standard_payment_amount) AS national_year_medicare_standard_payment,
    ROUND(SUM(total_snf_medicare_standard_payment_amount)/SUM(total_stays), 2) AS national_amount_per_stay
FROM
    `bigquery-public-data.cms_medicare.nursing_facilities_2014`;

-- Average, minimum and maximum total stays in nursing facilities and medicare payment amount per stay for each state

SELECT
    state,
    ROUND(AVG(total_stays), 2) AS state_average_total_stays,
    MIN(total_stays) AS state_min_total_stays,
    MAX(total_stays) AS state_max_total_stays,
    SUM(total_stays) AS state_year_total_stays,
    SUM(total_snf_medicare_standard_payment_amount) AS state_year_medicare_standard_payment,
    ROUND(SUM(total_snf_medicare_standard_payment_amount)/SUM(total_stays), 2) AS state_amount_per_stay
FROM
    `bigquery-public-data.cms_medicare.nursing_facilities_2014`
GROUP BY state
ORDER BY state_amount_per_stay DESC;

-- States with medicare payment amount per stay above national average

SELECT
    state,
    state_amount_per_stay
FROM (
    WITH nursing_facilities_total_stays AS
    (SELECT
    state,
    total_stays,
    total_snf_medicare_standard_payment_amount,
    SUM(total_snf_medicare_standard_payment_amount) OVER () AS national_standard_payment_amount,
    SUM(total_stays) OVER () AS national_total_stays
    FROM
    `bigquery-public-data.cms_medicare.nursing_facilities_2014`)
    SELECT
    state,
    ROUND(SUM(total_snf_medicare_standard_payment_amount)/SUM(total_stays), 2) AS state_amount_per_stay,
    ROUND(national_standard_payment_amount/national_total_stays, 2) AS national_amount_per_stay,
    FROM nursing_facilities_total_stays
    GROUP BY state, national_amount_per_stay
    )
    WHERE state_amount_per_stay > national_amount_per_stay
    GROUP BY state, state_amount_per_stay
    ORDER BY state_amount_per_stay DESC;

-- National total, average, minimum and maximum average HCC score for nursing facilities

SELECT
    ROUND(AVG(average_hcc_score), 2) AS national_avg_average_hcc_score,
    MIN(average_hcc_score) AS national_min_average_hcc_score,
    MAX(average_hcc_score) AS national_max_average_hcc_score
FROM
    `bigquery-public-data.cms_medicare.nursing_facilities_2014`;
```

```sql
-- State average, minimum, maximum HCC score and disparity between highest and lowest HCC score for nursing facilities

SELECT
    state,
    ROUND(AVG(average_hcc_score), 2) AS state_avg_average_hcc_score,
    MIN(average_hcc_score) AS state_min_average_hcc_score,
    MAX(average_hcc_score) AS state_max_average_hcc_score,
    ROUND(MAX(average_hcc_score) - MIN(average_hcc_score), 2) AS diff_max_min_average_hcc_score
FROM
    `bigquery-public-data.cms_medicare.nursing_facilities_2014`
GROUP BY state
ORDER BY diff_max_min_average_hcc_score DESC;

-- States with HCC score above national average

SELECT
    state,
    state_average_hcc_score
FROM (
    WITH nursing_facilities_hcc_score AS
    (SELECT
    state,
    average_hcc_score,
    AVG(average_hcc_score) OVER () AS national_average_hcc_score,
    FROM
    `bigquery-public-data.cms_medicare.nursing_facilities_2014`)
    SELECT
    state,
    ROUND(AVG(average_hcc_score), 2) AS state_average_hcc_score,
    ROUND(national_average_hcc_score, 2) AS national_average_hcc_score
    FROM nursing_facilities_hcc_score
    GROUP BY state, national_average_hcc_score
    )
    WHERE state_average_hcc_score > national_average_hcc_score
    GROUP BY state, state_average_hcc_score
    ORDER BY state_average_hcc_score DESC;

-- Nationwide average percent of beneficiaries having some specific chronic conditions

SELECT
    ROUND(AVG(percent_of_beneficiaries_with_atrial_fibrillation), 2) AS national_avg_percent_atrial_fibrillation,
    ROUND(AVG(percent_of_beneficiaries_with_alzheimers), 2) AS national_avg_percent_alzheimers,
    ROUND(AVG(percent_of_beneficiaries_with_asthma), 2) AS national_avg_percent_asthma,
    ROUND(AVG(percent_of_beneficiaries_with_cancer), 2) AS national_avg_percent_cancer,
    ROUND(AVG(percent_of_beneficiaries_with_chf), 2) AS national_avg_percent_chf,
    ROUND(AVG(percent_of_beneficiaries_with_chronic_kidney_disease), 2) AS national_avg_percent_chronic_kidney_disease,
    ROUND(AVG(percent_of_beneficiaries_with_copd), 2) AS national_avg_percent_copd,
    ROUND(AVG(percent_of_beneficiaries_with_depression), 2) AS national_avg_percent_depression,
    ROUND(AVG(percent_of_beneficiaries_with_diabetes), 2) AS national_avg_percent_diabetes,
    ROUND(AVG(percent_of_beneficiaries_with_hyperlipidemia), 2) AS national_avg_percent_hyperlipidemia,
    ROUND(AVG(percent_of_beneficiaries_with_ihd), 2) AS national_avg_percent_ihd,
    ROUND(AVG(percent_of_beneficiaries_with_osteoporosis), 2) AS national_avg_percent_osteoporosis,
    ROUND(AVG(percent_of_beneficiaries_with_ra_oa), 2) AS national_avg_percent_ra_oa,
```

```
    ROUND(AVG(percent_of_beneficiaries_with_schizophrenia), 2) AS national_avg_percent_schizophrenia,
    ROUND(AVG(percent_of_beneficiaries_with_stroke), 2) AS national_avg_percent_stroke
FROM
    `bigquery-public-data.cms_medicare.nursing_facilities_2014`;

-- Average percent of beneficiaries having some specific chronic conditions by state

SELECT
    state,
    ROUND(AVG(percent_of_beneficiaries_with_atrial_fibrillation), 2) AS state_avg_percent_atrial_fibrillation,
    ROUND(AVG(percent_of_beneficiaries_with_alzheimers), 2) AS state_avg_percent_alzheimers,
    ROUND(AVG(percent_of_beneficiaries_with_asthma), 2) AS state_avg_percent_asthma,
    ROUND(AVG(percent_of_beneficiaries_with_cancer), 2) AS state_avg_percent_cancer,
    ROUND(AVG(percent_of_beneficiaries_with_chf), 2) AS state_avg_percent_chf,
    ROUND(AVG(percent_of_beneficiaries_with_chronic_kidney_disease), 2) AS state_avg_percent_chronic_kidney_disease,
    ROUND(AVG(percent_of_beneficiaries_with_copd), 2) AS state_avg_percent_copd,
    ROUND(AVG(percent_of_beneficiaries_with_depression), 2) AS state_avg_percent_depression,
    ROUND(AVG(percent_of_beneficiaries_with_diabetes), 2) AS state_avg_percent_diabetes,
    ROUND(AVG(percent_of_beneficiaries_with_hyperlipidemia), 2) AS state_avg_percent_hyperlipidemia,
    ROUND(AVG(percent_of_beneficiaries_with_ihd), 2) AS state_avg_percent_ihd,
    ROUND(AVG(percent_of_beneficiaries_with_osteoporosis), 2) AS state_avg_percent_osteoporosis,
    ROUND(AVG(percent_of_beneficiaries_with_ra_oa), 2) AS state_avg_percent_ra_oa,
    ROUND(AVG(percent_of_beneficiaries_with_schizophrenia), 2) AS state_avg_percent_schizophrenia,
    ROUND(AVG(percent_of_beneficiaries_with_stroke), 2) AS state_avg_percent_stroke
FROM
    `bigquery-public-data.cms_medicare.nursing_facilities_2014`
GROUP BY state;
```