Project : Food Claim Analysis

```sql
SELECT claim_id,time_to_close,claim_amount,amount_paid,
       location,individuals_on_claim,linked_cases,cause,
       COUNT(*)
  FROM food_claims_2212
 GROUP BY claim_id,time_to_close,claim_amount,amount_paid,
       location,individuals_on_claim,linked_cases,cause
HAVING COUNT(*) > 1; -- Checked for duplicate entries.
```

Data Output | Messages | Notifications

	claim_id [PK] integer	time_to_close integer	claim_amount numeric	amount_paid numeric	location character var	individuals_on_claim integer	linked_cases character varyir	cause character var	count bigint

```sql
--Category of the variable location which has the most observations
  SELECT location,
         COUNT(claim_id) AS number_of_claims -- No of locations
    FROM food_claims_2212
GROUP BY location -- Grouping the results by location
ORDER BY number_of_claims DESC; -- Ordering the results in descending order
```

Data Output | Messages | Notifications

	location character varying (20)	number_of_claims bigint
1	RECIFE	885
2	SAO LUIS	517
3	FORTALEZA	311
4	NATAL	287

Project : Food Claim Analysis

```sql
--Distribution of time to close for all claims
  SELECT location,
         MIN(time_to_close) AS min_time_to_close,
         MAX(time_to_close) AS max_time_to_close,
         PERCENTILE_DISC(0.5) WITHIN GROUP
                (ORDER BY time_to_close) AS median_time_to_close,-- Calculates the median value
         ROUND(AVG(time_to_close),2) AS mean_time_to_close,--Calculates the average & rounded to two decimal
         COUNT(time_to_close)        AS no_of_claims -- Counts the number of claims
    FROM food_claims_2212
GROUP BY location;
```

Data Output | Messages | Notifications

	location character varying (20)	min_time_to_close integer	max_time_to_close integer	median_time_to_close integer	mean_time_to_close numeric	no_of_claims bigint
1	FORTALEZA	76	453	180	185.31	311
2	NATAL	93	361	179	185.93	287
3	RECIFE	82	427	178	184.61	885
4	SAO LUIS	84	518	179	187.17	517

Project : Food Claim Analysis

```
/*Calculate various percentile statistics for the time_to_close column grouped by
the location column in the table.*/
SELECT location,
       PERCENTILE_DISC(0.25) WITHIN GROUP(ORDER BY time_to_close)-
            1.5*(PERCENTILE_DISC(0.75) WITHIN GROUP(ORDER BY time_to_close)-
       PERCENTILE_DISC(0.25) WITHIN GROUP(ORDER BY time_to_close)) AS bottom_outlier,
       PERCENTILE_DISC(0.25) WITHIN GROUP(ORDER BY time_to_close) AS "25th percentile",
       PERCENTILE_DISC(0.5)  WITHIN GROUP(ORDER BY time_to_close) AS median,
       PERCENTILE_DISC(0.75) WITHIN GROUP(ORDER BY time_to_close) AS "75th percentile",
       (PERCENTILE_DISC(0.75) WITHIN GROUP(ORDER BY time_to_close)-
            PERCENTILE_DISC(0.25) WITHIN GROUP(ORDER BY time_to_close)) AS IQR,
        PERCENTILE_DISC(0.75) WITHIN GROUP(ORDER BY time_to_close)+
            1.5*(PERCENTILE_DISC(0.75) WITHIN GROUP(ORDER BY time_to_close)-
       PERCENTILE_DISC(0.25) WITHIN GROUP(ORDER BY time_to_close)) AS upper_outlier
 FROM food_claims_2212
GROUP BY location;
```

Data Output | Messages | Notifications

	location character varying (20)	bottom_outlier numeric	25th percentile integer	median integer	75th percentile integer	iqr integer	upper_outlier numeric
1	FORTALEZA	83.5	157	180	206	49	279.5
2	NATAL	83.5	157	179	206	49	279.5
3	RECIFE	88.0	157	178	203	46	272.0
4	SAO LUIS	95.0	161	179	205	44	271.0

```
--Total claimed amount, total paid amount, and the balanced amount for each location
  SELECT COALESCE(location, 'Total Amount') AS location,
         SUM(claim_amount) AS total_claim_amount,
         SUM(amount_paid) AS total_paid_amount,
         SUM(claim_amount) - SUM(amount_paid) AS balanced_amount
    FROM food_claims_2212
GROUP BY ROLLUP(location)
ORDER BY location;-- Total claimed amount, total paid amount and balanced amount by location wise.
```

Data Output | Messages | Notifications

	location character varying	total_claim_amount numeric	total_paid_amount numeric	balanced_amount numeric
1	FORTALEZA	8294581.01	6589450.95	1705130.06
2	NATAL	7914729.79	6219389.00	1695340.79
3	RECIFE	24116530.58	19096671.44	5019859.14
4	SAO LUIS	13988089.09	11126733.07	2861356.02
5	Total Amount	54313930.47	43032244.46	11281686.01

Project : Food Claim Analysis

```sql
--Analysis of individuals claims using bins.
WITH bins AS (SELECT GENERATE_SERIES(0,18, 3)+1 AS lower,-- Lower bin series
                     GENERATE_SERIES(3,21,3)    AS upper)-- Upper bin series

    SELECT lower, upper,
           COUNT(individuals_on_claim) AS no_of_individuals_claim
      FROM food_claims_2212 AS f
INNER JOIN bins AS b
           ON f.individuals_on_claim >= lower     -- individuals on claim should be between upper and lower
              AND f.individuals_on_claim <= upper
  GROUP BY lower, upper
  ORDER BY lower;
```

Data Output | Messages | Notifications

	lower integer	upper integer	no_of_individuals_claim bigint
1	1	3	363
2	4	6	402
3	7	9	448
4	10	12	428
5	13	15	359

```sql
-- Calculates the number of cases and the percentage distribution of linked cases
   SELECT linked_cases,
          COUNT(claim_id) AS no_of_cases,
          CONCAT(
              ROUND(
                    (COUNT(claim_id)/
                                  (SELECT COUNT(claim_id)::NUMERIC
                                     FROM food_claims_2212))*100,
                    2),
              ' %') AS percetage_distribution
     FROM food_claims_2212
 GROUP BY linked_cases;
```

Data Output | Messages | Notifications

	linked_cases character varying (15)	no_of_cases bigint	percetage_distribution text
1	TRUE	481	24.05 %
2	NA	26	1.30 %
3	FALSE	1493	74.65 %

Project : Food Claim Analysis

```
/*Count of cases for each location, considering only the cases where the linked_cases column
has a value of 'FALSE'*/
   SELECT location,
          COUNT(claim_id) AS no_of_cases,
          CONCAT(
              ROUND(
                    (COUNT(claim_id)/
                                    (SELECT COUNT(claim_id)::NUMERIC
                                       FROM food_claims_2212))*100,
                    2),
               ' %') AS percetage_distribution
     FROM food_claims_2212
     WHERE linked_cases = 'FALSE'
 GROUP BY location
 ORDER BY no_of_cases DESC;
```

Data Output | Messages | Notifications

	location character varying (20)	no_of_cases bigint	percetage_distribution text
1	RECIFE	656	32.80 %
2	SAO LUIS	381	19.05 %
3	FORTALEZA	237	11.85 %
4	NATAL	219	10.95 %

```
-- Count the number of claims for each cause where linked_cases is 'TRUE'.
  SELECT cause,
         COUNT(claim_id) AS no_of_claims
    FROM food_claims_2212
   WHERE linked_cases = 'TRUE' -- Filter for TRUE cases
GROUP BY cause
ORDER BY no_of_claims DESC;
```

Data Output | Messages | Notifications

	cause character varying (20)	no_of_claims bigint
1	Unknown	197
2	Vegetable	166
3	Meat	118