**CO17:** Distribution of condition occurrence records per person

This query is used to provide summary statistics for the number of condition occurrence records (condition\_occurrence\_id) for all persons: the mean, the standard deviation, the minimum, the 25th percentile, the median, the 75th percentile, the maximum and the number of missing values. There is no input required for this query.

**Sample query:**

with ranked as

(

SELECT num\_of\_conditions, sum(1) over (partition by 1

ORDER BY num\_of\_conditions asc rows BETWEEN unbounded preceding AND current row) AS rownumasc

FROM (

SELECT count(\*) as num\_of\_conditions

FROM condition\_occurrence

WHERE person\_id!=0

GROUP BY person\_id

)

),

other\_stat AS

(

SELECT count(num\_of\_conditions) AS condition\_num\_count,

min(num\_of\_conditions) AS condition\_num\_min,

max(num\_of\_conditions) AS condition\_num\_max,

avg(num\_of\_conditions) AS condition\_num\_averege,

stddev(num\_of\_conditions) as condition\_num\_stddev

FROM (

SELECT count(\*) AS num\_of\_conditions, person\_id

FROM condition\_occurrence

WHERE person\_id!=0

GROUP BY person\_id

)

)

SELECT

(

SELECT count(distinct person\_id)

FROM condition\_occurrence

WHERE person\_id!=0 AND condition\_occurrence\_id IS NULL

) AS condition\_null\_count,

\* FROM other\_stat,

(

SELECT num\_of\_conditions AS condition\_num\_25percentile

FROM (SELECT \*,(SELECT count(\*) FROM ranked) as rowno FROM ranked)

WHERE (rownumasc=cast (rowno\*0.25 as int) AND mod(rowno\*25,100)=0) OR

(rownumasc=cast (rowno\*0.25 as int) AND mod(rowno\*25,100)>0) OR

(rownumasc=cast (rowno\*0.25 as int)+1 AND mod(rowno\*25,100)>0)

) AS condition\_num\_25percentile,

(

SELECT num\_of\_conditions AS condition\_num\_median

FROM (SELECT \*,(SELECT count(\*) FROM ranked) AS rowno FROM ranked)

WHERE (rownumasc=cast (rowno\*0.50 AS int) AND mod(rowno\*50,100)=0) OR

(rownumasc=cast (rowno\*0.50 AS int) AND mod(rowno\*50,100)>0) OR

(rownumasc=cast (rowno\*0.50 AS int)+1 AND mod(rowno\*50,100)>0)

) AS condition\_num\_median,

(

SELECT num\_of\_conditions AS condition\_num\_75percentile

FROM (SELECT \*,(SELECT count(\*) FROM ranked) as rowno FROM ranked)

WHERE (rownumasc=cast (rowno\*0.75 as int) AND mod(rowno\*75,100)=0) OR

(rownumasc=cast (rowno\*0.75 as int) AND mod(rowno\*75,100)>0) OR

(rownumasc=cast (rowno\*0.75 as int)+1 AND mod(rowno\*75,100)>0)

) AS condition\_num\_75percentile

**Input:**

None

**Output:**

|  |  |
| --- | --- |
| **Field** | **Description** |
| condition\_null\_count | Number of persons with at least one null condition\_occurrence\_id |
| condition\_num\_count | Number of distinct persons with conditions |
| condition\_num\_min | The lowest number of condition occurences |
| condition\_num\_max | The highest number of condition occurences |
| condition\_num\_averege | The average number of condition occurences |
| condition\_num\_stddev | The standard deviation of condition occurence numbers |
| condition\_num\_25percentile | A condition occurence number where 25 percent of the other numbers are lower |
| condition\_num\_median | A condition occurence number where half of the other numbers are lower and half are higher |
| condition\_num\_75percentile | A condition occurence number where 75 percent of the other numbers are lower |

**Sample output record:**

|  |  |
| --- | --- |
| **Field** | **Description** |
| condition\_null\_count | 4395019 |
| condition\_num\_count |  |
| condition\_num\_min | 1 |
| condition\_num\_max | 7144 |
| condition\_num\_averege | 51 |
| condition\_num\_stddev | 86.63 |
| condition\_num\_25percentile | 11 |
| condition\_num\_median | 26 |
| condition\_num\_75percentile | 58 |