**OP20:** Distribution of observation period length, stratified by gender.

This query is used to provide summary statistics for the observation period length across all observation period records stratified by gender (gender\_concept\_id): the mean, the standard deviation, the minimum, the 25th percentile, the median, the 75th percentile, the maximum and the number of missing values. The length of an is defined as the difference between the start date and the end date. All existing gender\_concept\_id values are summarized.

**Sample query:**

SELECT

gender,

count(\*) AS observation\_periods\_cnt ,

min( period\_length ) AS min\_period ,

max( period\_length ) AS max\_period ,

round( avg( period\_length ), 2 ) AS avg\_period ,

round( stdDev( period\_length ), 1 ) AS stdDev\_period ,

percentile\_25,

median,

percentile\_75

FROM (

SELECT

person\_id,

gender,

period\_length,

PERCENTILE\_DISC(0.25) WITHIN GROUP( ORDER BY period\_length ) over(partition by gender) AS percentile\_25 ,

PERCENTILE\_DISC(0.5) WITHIN GROUP (ORDER BY period\_length ) over(partition by gender) AS median ,

PERCENTILE\_DISC(0.75) WITHIN GROUP (ORDER BY period\_length ) over(partition by gender) AS percentile\_75

FROM /\* person, age \*/ (

SELECT

person\_id ,

concept\_name as gender

FROM (

SELECT

person\_id ,

min( observation\_period\_start\_date ) AS first\_observation\_date

FROM observation\_period

GROUP BY person\_id

)

JOIN person USING( person\_id )

JOIN concept ON concept\_id = gender\_concept\_id

WHERE year\_of\_birth IS NOT NULL

)

JOIN (

SELECT

person\_id ,

observation\_period\_end\_date - observation\_period\_start\_date + 1 AS period\_length

FROM observation\_period

) USING( person\_id )

)

GROUP BY

gender,

percentile\_25 ,

median ,

percentile\_75;

**Input:**

None

**Output:**

|  |  |
| --- | --- |
| **Field** | **Description** |
| gender | Gender concept name |
| observations\_period\_cnt | Number of observation periods for specific gender |
| min\_period | Minimum duration of observation period in days |
| max\_period | Maximum duration of observation period in days |
| avg\_period | Average duration of observation in days |
| stdDev\_period | Standard deviation of observation |
| percentile\_25 | 25th percentile of observation periods in days |
| median | Median of observation periods in days |
| percentile\_75 | 75th percentile of observation periods in days |

**Sample output record:**

|  |  |
| --- | --- |
| **Field** | **Description** |
| gender | MALE |
| observations\_period\_cnt | 1812743 |
| min\_period | 1 |
| max\_period | 2372 |
| avg\_period | 653.77 |
| stdDev\_period | 502.40 |
| percentile\_25 | 365 |
| median | 457 |
| percentile\_75 | 731 |