**OP14:** Distribution of age, stratified by gender

This query is used to provide summary statistics for the age across all observation 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 age value is defined by the earliest observation date. Age is summarized for all existing gender\_concept\_id values.

**Sample query:**

WITH t AS /\* person, gender, age \*/

( SELECT person\_id, NVL( concept\_name, 'MISSING' ) AS gender

, extract( year FROM first\_observation\_date ) - year\_of\_birth AS age

FROM -- person, first observation period date

( SELECT person\_id

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

FROM observation\_period

GROUP BY person\_id

)

JOIN person USING( person\_id )

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

WHERE year\_of\_birth IS NOT NULL

)

SELECT gender

, count(\*) AS num\_people

, min( age ) AS min\_age

, max( age ) AS max\_age

, round( avg( age ), 2 ) AS avg\_age

, round( stdDev( age ), 1 ) AS stdDev\_age

, (SELECT DISTINCT PERCENTILE\_DISC(0.25) WITHIN GROUP( ORDER BY age ) over ()

AS percentile\_25 FROM t)

, (SELECT DISTINCT PERCENTILE\_DISC(0.5) WITHIN GROUP (ORDER BY age ) over ()

AS median FROM t)

, (SELECT DISTINCT PERCENTILE\_DISC(0.75) WITHIN GROUP (ORDER BY age ) over ()

AS percential\_75 FROM t)

FROM t

GROUP BY gender

**Input:**

None

**Output:**

|  |  |
| --- | --- |
| **Field** | **Description** |
| gender | Gender concept name |
| num\_people | Number of people with specific gender |
| min\_age | Minimum age across observation of people with specific gender |
| max\_age | Maximum age across observation of people with specific gender |
| avg\_age | Average age across observation of people with specific gender |
| stdDev\_age | Standard deviation of age across observation within specific gender |
| percentile\_25 | 25th percentile age across observation within specific gender |
| median | Median age across observation within specific gender |
| percentile\_75 | 75th percentile age across observation within specific gender |

**Sample output record:**

|  |  |
| --- | --- |
| **Field** | **Description** |
| gender | MALE |
| num\_people | 1607472 |
| min\_age | 0 |
| max\_age | 103 |
| avg\_age | 40.78 |
| stdDev\_age | 18.60 |
| percentile\_25 | 29 |
| median | 45 |
| percentile\_75 | 55 |