**PE12:** Distribution of year of birth

This query is used to to provide summary statistics for the age across all patient records: the mean, the standard deviation, the minimum, the 25th percentile, the median, the 75th percentile, the maximum. No input is required for this query.

**Sample query:**

SELECT percentile\_25

, median

, percentile\_75

, MIN( year\_of\_birth ) AS minimum

, MAX( year\_of\_birth ) AS maximum

, CAST(AVG( year\_of\_birth ) AS INTEGER) AS mean

, STDDEV( year\_of\_birth ) AS stddev

FROM

(SELECT MAX( CASE WHEN( percentile = 1 ) THEN year\_of\_birth END ) AS percentile\_25

, MAX( CASE WHEN( percentile = 2 ) THEN year\_of\_birth END ) AS median

, MAX( CASE WHEN( percentile = 3 ) THEN year\_of\_birth END ) AS percentile\_75

FROM -- year of birth / percentile

( SELECT year\_of\_birth, births

/\* The first sum is the sum of all the values from the first year of birth

to the current year. The second sum is the total of all the years of birth.

The result is a cumulative percent of the total for each year. You want to

capture when the percentage goes from 24 to 25 as percentile\_25, from 49 to 50

as the median and from 74 to 75 as the percentile\_75. Multiplying by 4 then SA

adding 1 just makes so that instead of looking at percentage, you get the whole

number 1 if the percentage is less than 25, 2 when the percentage is between 25

and 50, and so on.

\*/

, FLOOR( CAST( SUM( births ) OVER( ORDER BY year\_of\_birth ROWS UNBOUNDED PRECEDING ) AS DECIMAL )

/ CAST( SUM( births ) OVER( ORDER BY year\_of\_birth ROWS BETWEEN UNBOUNDED PRECEDING

AND UNBOUNDED FOLLOWING ) AS DECIMAL )

\* 4

) + 1 percentile

FROM -- Year with number of birthsQ

( SELECT year\_of\_birth, count(\*) AS births

FROM person

GROUP BY year\_of\_birth

)

)where percentile <= 3

) percentile\_table, person

GROUP BY percentile\_25, median, percentile\_75

**Input:**

None

**Output:**

|  |  |
| --- | --- |
| **Field** | **Description** |
| percentile\_25 | 25th percentile of year of birth |
| median | Median patient year of birth |
| percentile\_75 | 75th percentile of year of birth |
| minimum | Minimum year of birth |
| maximum | Maximum year of birth |
| mean | Mean patient year of birth |
| stddev | Standard deviation of year of birth |

**Sample output record:**

|  |  |
| --- | --- |
| **Field** | **Value** |
| percentile\_25 | 1954 |
| median | 1965 |
| percentile\_75 | 1979 |
| minimum | 1893 |
| maximum | 2010 |
| mean | 1968 |
| stddev | 17.277 |