**DER23:** Distribution of drug era start dates, stratified by drug

This query is used to summary statistics of the drug era start dates (drug\_era\_start\_date) across all drug era records, stratified by drug (drug\_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 input to the query is a value (or a comma-separated list of values) of a drug\_concept\_id. If the input is omitted, all possible values are summarized.

**Sample query:**

with tt as (

SELECT

(t.drug\_era\_start\_date - MIN(t.drug\_era\_start\_date) OVER(partition by t.drug\_concept\_id)) AS start\_date\_num,

t.drug\_era\_start\_date AS start\_date, MIN(t.drug\_era\_start\_date) OVER(partition by t.drug\_concept\_id) min\_date,

t.drug\_concept\_id

FROM drug\_era t

where t.drug\_concept\_id in (1300978, 1304643, 1549080)

)

SELECT

tt.drug\_concept\_id,

min(tt.start\_date\_num) AS min\_value,

max(tt.start\_date\_num) AS max\_value,

tt.min\_date+avg(tt.start\_date\_num) AS avg\_value,

(round(stdDev(tt.start\_date\_num)) ) AS stdDev\_value ,

tt.min\_date+(select distinct PERCENTILE\_DISC(0.25) WITHIN GROUP(ORDER BY tt.start\_date\_num) OVER() from tt) AS percentile\_25,

tt.min\_date+(select distinct PERCENTILE\_DISC(0.5) WITHIN GROUP (ORDER BY tt.start\_date\_num) OVER() from tt) AS median\_value,

tt.min\_date+(select distinct PERCENTILE\_DISC(0.75) WITHIN GROUP (ORDER BY tt.start\_date\_num) OVER() from tt) AS percential\_75

FROM tt

group by

drug\_concept\_id,

tt.min\_date

order by

drug\_concept\_id

;

**Input:**

|  |
| --- |
|  |
| **Parameter** | **Example** | **Mandatory** | **Notes** |
| drug\_concept\_id | 1300978, 1304643, 1549080 | Yes |  |

**Output:**

|  |  |
| --- | --- |
| **Field** | **Description** |
| drug\_concept\_id | A foreign key that refers to a standard concept identifier in the vocabulary for the drug concept. |
| min\_value |  |
| max\_value |  |
| avg\_value | The start date for the drug era constructed from the individual instances of drug exposures. It is the start date of the very first chronologically recorded instance of utilization of a drug. |
| stddev\_value |  |
| percentile\_25 |  |
| median\_value |  |
| percentile\_75 |  |

**Sample output record:**

|  |  |
| --- | --- |
| **Field** | **Description** |
| drug\_concept\_id | 1300978 |
| min\_value | 0 |
| max\_value | 7156 |
| avg\_value | 2006-04-13 00:00:00 |
| stddev\_value | 1808 |
| percentile\_25 | 2000-03-21 00:00:00 |
| median\_value | 2002-07-29 00:00:00 |
| percentile\_75 | 2005-01-15 00:00:00 |