

Traceis™ Data Exploration Studio Summary table

1 Select tables and graphs step

This example uses the baseball hitters dataset. To generate a summary table, click on the tables and graphs step.

2 Select summary table tab

3 Select categorical variable to group on

A summary table will be generated where each row is a group of observations sharing a particular value for the 'group on' variable. In this example, Player team ('86) was selected.

Traceis Data Exploration Studio 2008

File Edit Help

Open Save Print Copy X Delete Search

1. Definition
2. Preparation
3. Implementation
4. Deployment

Tables and graphs
Statistics
Grouping
Prediction
Apply model

Contingency table Summary table Graphs

Group on: Player team ('86)

Show count

Number of columns: 7

Column (4)
Statistic: Mean
Variable: Runs ('86)

Column (5)
Statistic: Mean
Variable: Runs batted ('86)

Column (6)
Statistic: Mean
Variable: Walks ('86)

Column (7)
Statistic: Mean
Variable: Salary ('87)

Display

Selected items Histograms

Hitter's name | Bat times ('86) | Hits ('86) | Home runs ('86) | Runs ('86) | Runs batted ('86) | Walks ('86) | Player team ('86) | Player's position ('86) | Put outs ('86) | Assists ('86)

Opening file baseball hitters.txt

MAKING SENSE OF DATA

1 Optionally view count

The number of observations for each row is optionally displayed.

2 Select number of summary columns

Additional columns that summarize other variables can be displayed.

3 Define summary columns

Each summarized column requires the selection of another variable to summarize along with a summary statistic (e.g. maximum value or mean). In this example, the mean of 7 additional columns (such as Salary ('87)) are displayed.

4 Display summary table

Click on the display button to view the summary table.

The screenshot shows the 'Tracels Data Exploration Studio 2007 beta 1' interface. On the left, a navigation pane has four main sections: '1. Definition', '2. Preparation', '3. Implementation', and '4. Deployment'. The '3. Implementation' section is active, with sub-options for 'Tables and graphs', 'Statistics', 'Grouping', 'Prediction', and 'Apply model'. The 'Summary table' tab is selected in the top toolbar. The configuration panel for the summary table is visible, showing the following settings:

- Group on: Player team ('86')
- Show count:
- Number of columns: 7
- Column (4): Statistic: Mean, Variable: Runs ('86')
- Column (5): Statistic: Mean, Variable: Runs batted ('86')
- Column (6): Statistic: Mean, Variable: Walks ('86')
- Column (7): Statistic: Mean, Variable: Salary ('87')

A 'Display' button is located at the bottom of the configuration panel. The main workspace is currently empty. The bottom status bar indicates 'Opening file baseball hitters.txt'.

MAKING SENSE OF DATA

1 Common value

Each row represents a set of observations, all sharing the value in the first column.

2 Number of observations

The number of observations in each row is shown in the next columns.

3 Summarized columns

Each additional column defined is shown in the next columns. Any column may be sorted from low to high by clicking on the column header.

4 Selected row

Clicking on a row will display the selected observation.

The screenshot shows the Trace3 Data Exploration Studio 2007 beta 1 interface. The main window displays a data table with columns for Player team ('96), Count, Mean (Bat times ('96)), Mean (Hits ('96)), Mean (Home runs ('96)), Mean (Runs ('96)), Mean (Runs batted ('96)), Mean (Walks ('96)), and Mean (Salary ('07)). The table lists various teams and their corresponding statistics. The interface includes a sidebar with navigation options (1. Definition, 2. Preparation, 3. Implementation, 4. Deployment) and a 'Selected items' section at the bottom. The 'Selected items' section shows a list of players and their statistics, including Hitler's name, Bat times ('96), Hits ('96), Home runs ('96), Runs ('96), Runs batted ('96), Walks ('96), Player team ('96), Player's position ('96), Put outs ('96), and Assists.

Player team ('96)	Count	Mean (Bat times ('96))	Mean (Hits ('96))	Mean (Home runs ('96))	Mean (Runs ('96))	Mean (Runs batted ('96))	Mean (Walks ('96))	Mean (Salary ('07))
S.F.	13	364	96.54	7.85	47.23	44.31	37.54	276.92
Sea.	11	423.64	110.91	12.45	59.36	53.82	46.55	282.5
Tex.	11	395.09	104.36	13.36	54.36	50.27	32.64	311.23
Pt.	9	377.99	99.76	10.44	46.44	49.67	40.33	337.17
Min.	9	352.44	93.22	9	44.67	42.67	33.22	359.61
ML.	10	327.3	85.4	8.9	43.1	36.3	31.7	436.25
L.A.	10	334.8	88.8	8.2	41.4	37.5	33.7	476
Min.	11	434.18	115.36	16.73	59.18	57.27	41.55	479.27
Cal.	10	402.7	105.2	13	59.7	59.4	50.2	486.42
Dat.	11	397	105	16.55	61.55	59.27	45.36	497.62
S.D.	12	350.33	95.17	9.67	43.67	42.83	34.42	506.18
Hou.	10	410.8	113.2	11.1	63.9	61.7	39.2	519
Cle.	11	471.36	136.18	13.82	68.09	65.18	37.91	527.73
Ch.	10	404.9	107.8	12	56.9	51.3	43.3	533.82
K.C.	9	394.89	100.33	11.11	45.89	46	34	538.11
Chi.	22	412.73	106.32	11.18	50	49.18	37.14	538.26
Oak.	8	449.25	118.38	12.88	63.12	56.25	45.88	539.38
Phi.	10	368.1	95.8	9.3	52	47.9	38.4	590.58
S.L.	11	400.73	99	4.27	46.73	40.64	43.73	621.97
Bal.	10	373.3	101.6	11.8	51.3	47	40.2	653.5
Tor.	9	491.56	136.22	15.11	71.67	66.78	40.89	713.61
Atl.	8	386.88	96.62	9.88	45.12	41.88	44.62	716.88
N.Y.	20	443.45	125.85	14.8	68.35	62.85	53.5	843.73
Bos.	8	513.88	143.38	13.88	78	72.38	59.12	961.56