

# DATA ANALYTICS

Shivam Thakur

19BME0237

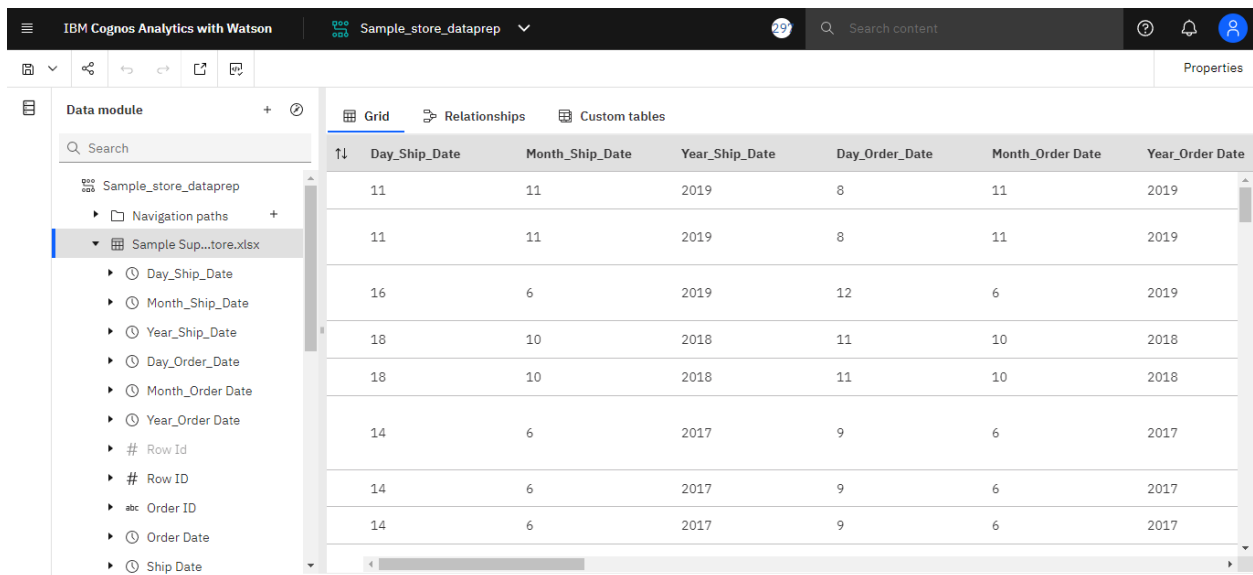
VIT Vellore

## Assignment 1:

### Cognos Data Analytics: Superstore

Sample Superstore data prep:

[https://eu2.ca.analytics.ibm.com/bi/?perspective=ca-modeller&pathRef=.my\\_folders%2FSample\\_store\\_dataprep](https://eu2.ca.analytics.ibm.com/bi/?perspective=ca-modeller&pathRef=.my_folders%2FSample_store_dataprep)

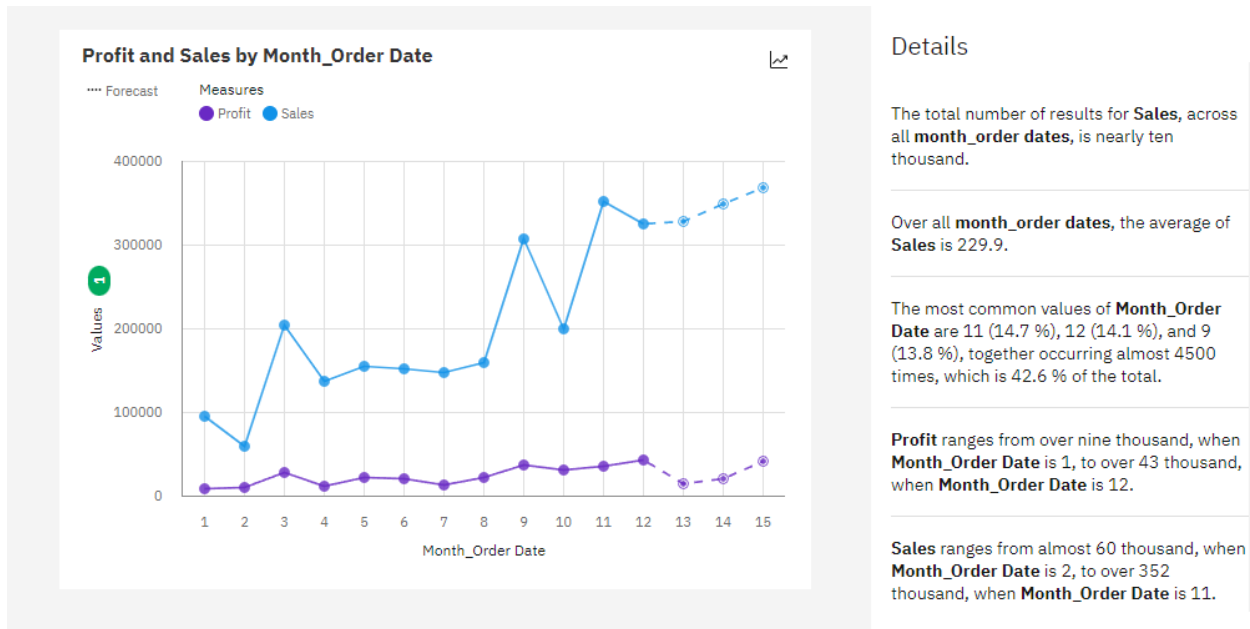


The screenshot displays the IBM Cognos Analytics interface. The top navigation bar includes the title 'IBM Cognos Analytics with Watson', the current workspace 'Sample\_store\_dataprep', and a search bar. The left sidebar shows the 'Data module' with a search bar and a tree view containing 'Sample\_store\_dataprep' and its sub-items: 'Navigation paths', 'Sample Sup...tore.xlsx', and a list of date fields (Day\_Ship\_Date, Month\_Ship\_Date, Year\_Ship\_Date, Day\_Order\_Date, Month\_Order\_Date, Year\_Order\_Date), 'Row ID', and 'Ship Date'. The main area shows a 'Grid' view of the data. The table has 7 columns: Day\_Ship\_Date, Month\_Ship\_Date, Year\_Ship\_Date, Day\_Order\_Date, Month\_Order\_Date, Year\_Order\_Date, and an unlabeled column. The data is sorted by Day\_Ship\_Date in descending order.

| Day_Ship_Date | Month_Ship_Date | Year_Ship_Date | Day_Order_Date | Month_Order_Date | Year_Order_Date |  |
|---------------|-----------------|----------------|----------------|------------------|-----------------|--|
| 11            | 11              | 2019           | 8              | 11               | 2019            |  |
| 11            | 11              | 2019           | 8              | 11               | 2019            |  |
| 16            | 6               | 2019           | 12             | 6                | 2019            |  |
| 18            | 10              | 2018           | 11             | 10               | 2018            |  |
| 18            | 10              | 2018           | 11             | 10               | 2018            |  |
| 14            | 6               | 2017           | 9              | 6                | 2017            |  |
| 14            | 6               | 2017           | 9              | 6                | 2017            |  |
| 14            | 6               | 2017           | 9              | 6                | 2017            |  |

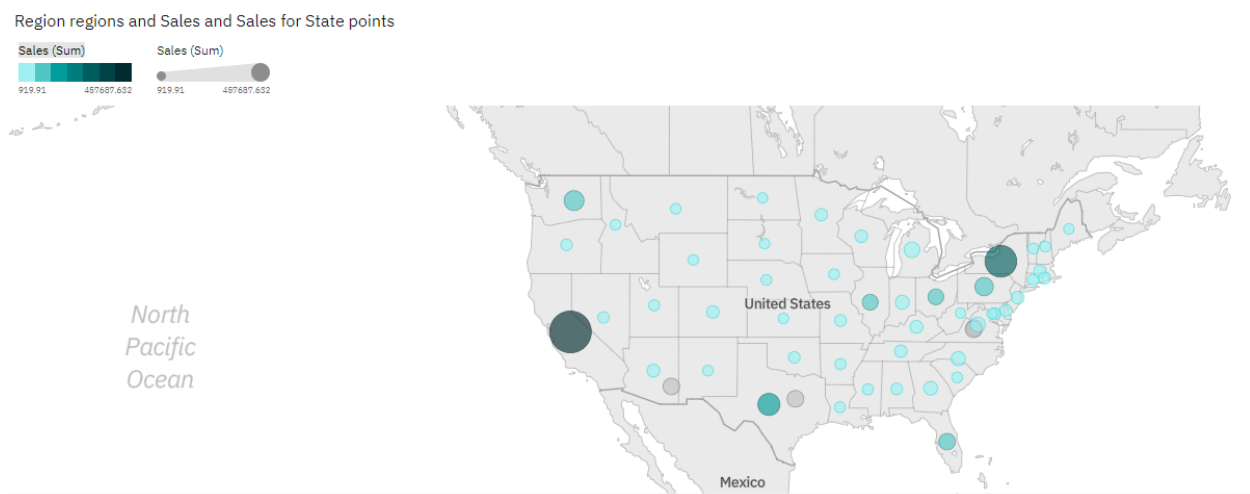
## Sample Superstore data Exploration:

[https://eu2.ca.analytics.ibm.com/bi/?perspective=explore&pathRef=.my\\_folders%2Fsales\\_exploration](https://eu2.ca.analytics.ibm.com/bi/?perspective=explore&pathRef=.my_folders%2Fsales_exploration)



## Sample Superstore data Dashboard:

[https://eu2.ca.analytics.ibm.com/bi/?perspective=dashboard&pathRef=.my\\_folders%2Fsales\\_dashboard&action=view&mode=dashboard](https://eu2.ca.analytics.ibm.com/bi/?perspective=dashboard&pathRef=.my_folders%2Fsales_dashboard&action=view&mode=dashboard)



## Assignment 2:

### Watson studio: Challengers.csv space shuttle

Data refining:

[https://eu-gb.dataplatform.cloud.ibm.com/shaper?project\\_id=ef3daea2-9d65-4e55-b22d-813bababf489&flow\\_id=b7f98b83-8b09-4455-8abb-25302e43a05a&context=cpdaas](https://eu-gb.dataplatform.cloud.ibm.com/shaper?project_id=ef3daea2-9d65-4e55-b22d-813bababf489&flow_id=b7f98b83-8b09-4455-8abb-25302e43a05a&context=cpdaas)

My projects / Challenger\_LinearRegression / challengers.csv

Preview Visualization

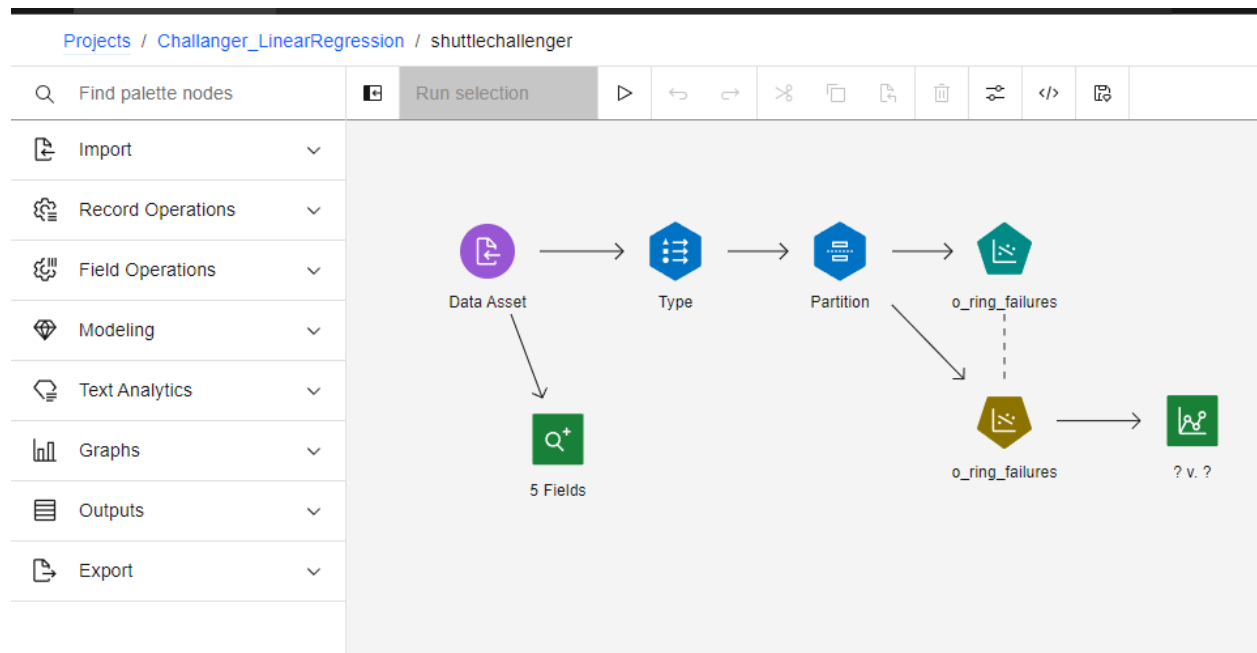
Schema: 5 Columns  
The preview includes only a limited set of columns and rows. ⓘ

Last refresh: 9 minutes ago ↻ Refine

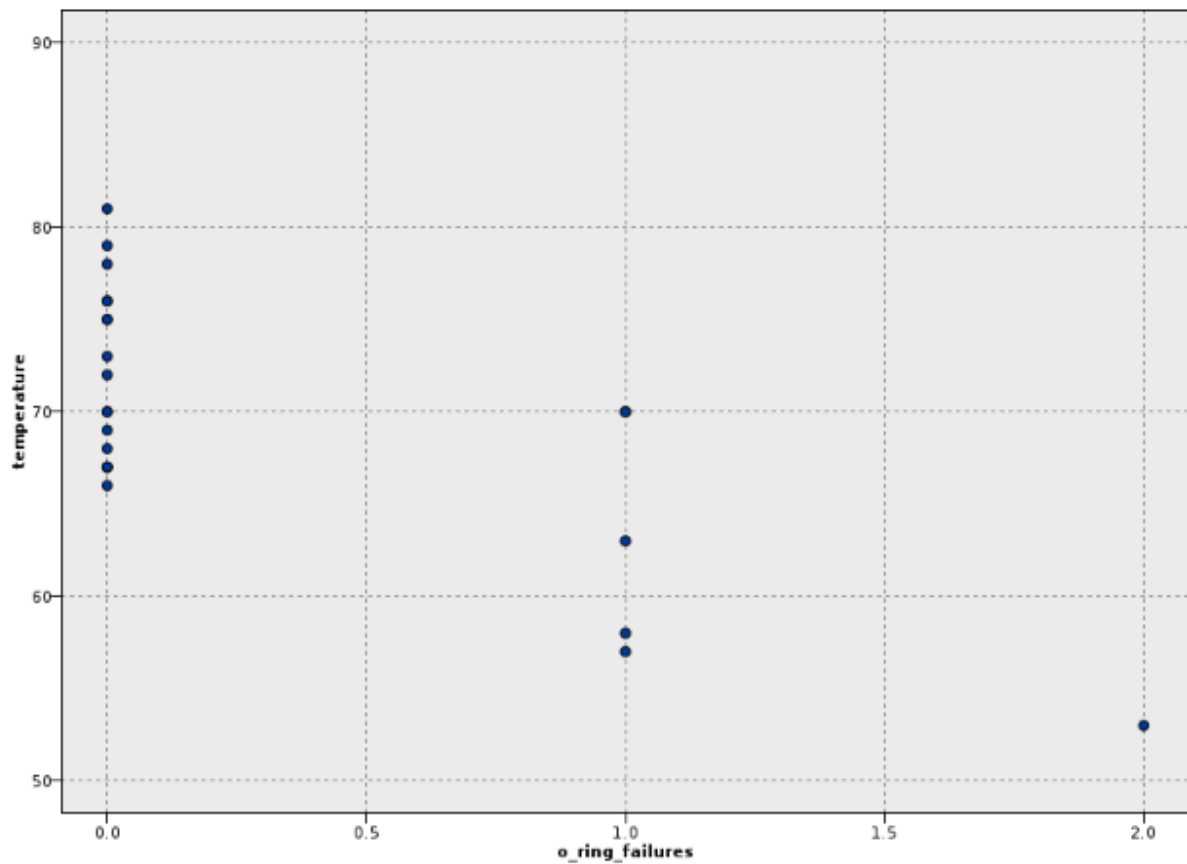
| o_ring_ct<br>String | o_ring_failures<br>String | temperature<br>String | pressure<br>String | launch_id<br>String |
|---------------------|---------------------------|-----------------------|--------------------|---------------------|
| 6                   | 0                         | 66                    | 50                 | 1                   |
| 6                   | 1                         | 70                    | 50                 | 2                   |
| 6                   | 0                         | 69                    | 50                 | 3                   |
| 6                   | 0                         | 68                    | 50                 | 4                   |
| 6                   | 0                         | 67                    | 50                 | 5                   |
| 6                   | 0                         | 72                    | 50                 | 6                   |
| 6                   | 0                         | 73                    | 100                | 7                   |
| 6                   | 0                         | 70                    | 100                | 8                   |
| 6                   | 1                         | 57                    | 200                | 9                   |
| 6                   | 1                         | 63                    | 200                | 10                  |
| 6                   | 1                         | 70                    | 200                | 11                  |
| 6                   | 0                         | 78                    | 200                | 12                  |

SPSS Modeler:

[https://eu-gb.dataplatform.cloud.ibm.com/canvas/flows/d08662ca-1eab-4e57-bd84-ae3d46ea43b8?context=cpdaas&project\\_id=ef3daea2-9d65-4e55-b22d-813bababf489](https://eu-gb.dataplatform.cloud.ibm.com/canvas/flows/d08662ca-1eab-4e57-bd84-ae3d46ea43b8?context=cpdaas&project_id=ef3daea2-9d65-4e55-b22d-813bababf489)



Output:



## Assignment 3:

### Watson studio: Diabetes.csv

Data refining:

[https://eu-gb.dataplatform.cloud.ibm.com/shaper?project\\_id=b8dcf6c2-d718-4b0d-be07-1d1fa42dea44&flow\\_id=18c3d526-545d-4792-bf8c-33f5994a6147&context=cpdaas](https://eu-gb.dataplatform.cloud.ibm.com/shaper?project_id=b8dcf6c2-d718-4b0d-be07-1d1fa42dea44&flow_id=18c3d526-545d-4792-bf8c-33f5994a6147&context=cpdaas)

My projects / Diabetes\_assingment / diabetes.csv

Preview Visualization

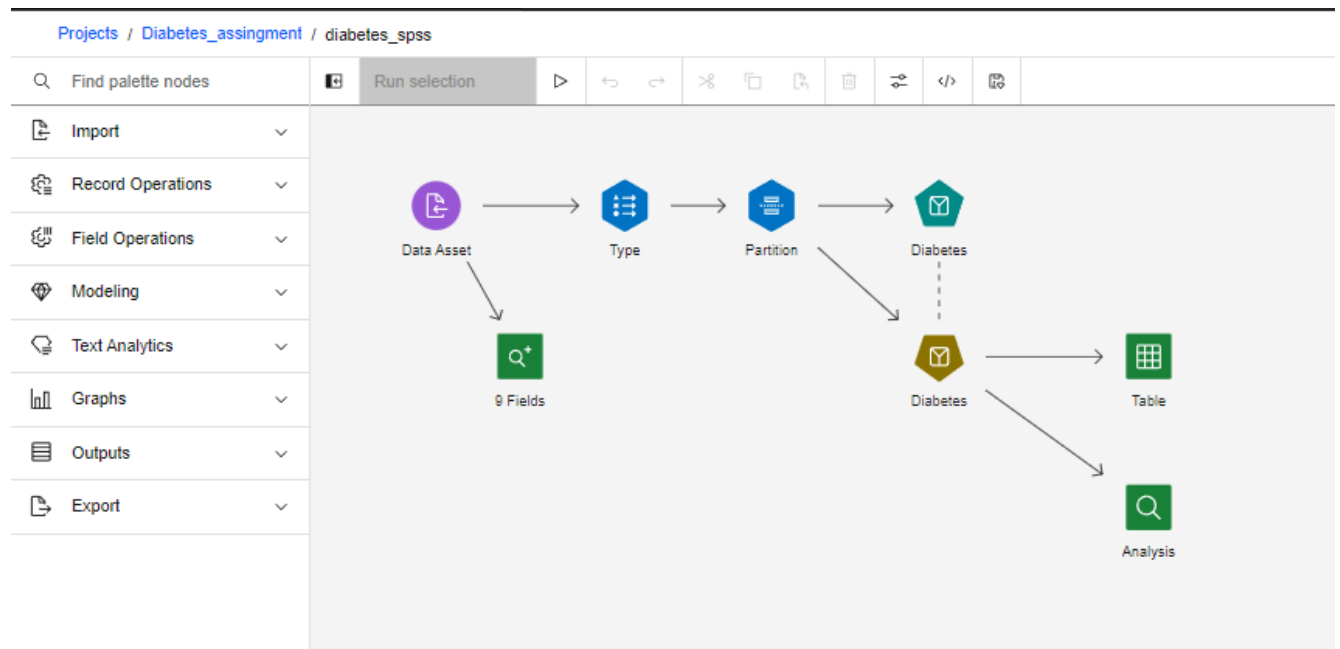
Schema: 9 Columns  
The preview includes only a limited set of columns and rows. ⓘ

Last refresh: just now ↻ Refine

| Pregnancies String | Glucose String | BloodPressure String | SkinThickness String | Insulin String | BMI String | DiabetesPedigreeFunction String | Age String | Diabetes String |
|--------------------|----------------|----------------------|----------------------|----------------|------------|---------------------------------|------------|-----------------|
| 6                  | 148            | 72                   | 35                   | 0              | 33.6       | 0.627                           | 50         | pos             |
| 1                  | 85             | 66                   | 29                   | 0              | 26.6       | 0.351                           | 31         | neg             |
| 8                  | 183            | 64                   | 0                    | 0              | 23.3       | 0.672                           | 32         | pos             |
| 1                  | 89             | 66                   | 23                   | 94             | 28.1       | 0.167                           | 21         | neg             |
| 0                  | 137            | 40                   | 35                   | 168            | 43.1       | 2.288                           | 33         | pos             |
| 5                  | 116            | 74                   | 0                    | 0              | 25.6       | 0.201                           | 30         | neg             |
| 3                  | 78             | 50                   | 32                   | 88             | 31         | 0.248                           | 26         | pos             |
| 10                 | 115            | 0                    | 0                    | 0              | 35.3       | 0.134                           | 29         | neg             |
| 2                  | 197            | 70                   | 45                   | 543            | 30.5       | 0.158                           | 53         | pos             |
| 8                  | 125            | 96                   | 0                    | 0              | 0          | 0.232                           | 54         | pos             |
| 4                  | 110            | 92                   | 0                    | 0              | 37.6       | 0.191                           | 30         | neg             |

SPSS Modeler:

[https://eu-gb.dataplatform.cloud.ibm.com/canvas/flows/78e5e044-860f-4252-a430-7529dacabf49?context=cpdaas&project\\_id=b8dcf6c2-d718-4b0d-be07-1d1fa42dea44](https://eu-gb.dataplatform.cloud.ibm.com/canvas/flows/78e5e044-860f-4252-a430-7529dacabf49?context=cpdaas&project_id=b8dcf6c2-d718-4b0d-be07-1d1fa42dea44)



Output:

### Results for output field Diabetes

#### Comparing \$D-Diabetes with Diabetes

| 'Partition' | 1_Training |       | 2_Testing |        |
|-------------|------------|-------|-----------|--------|
| Correct     | 394        | 73.1% | 171       | 74.67% |
| Wrong       | 145        | 26.9% | 58        | 25.33% |
| Total       | 539        |       | 229       |        |

## Assignment 4:

### Watson studio: Mall customers

Data refining:

[https://eu-gb.dataplatform.cloud.ibm.com/projects/7eeba2ab-b82c-4dfd-95a0-5d66a8cc77c0/data-assets/e0b54419-2f23-4e33-b296-f02841eafe94?context=cpdaas&walkme\\_guided\\_tutorial=false](https://eu-gb.dataplatform.cloud.ibm.com/projects/7eeba2ab-b82c-4dfd-95a0-5d66a8cc77c0/data-assets/e0b54419-2f23-4e33-b296-f02841eafe94?context=cpdaas&walkme_guided_tutorial=false)

My projects / MallCustomer\_clusteranalysis / Mall\_Customers.csv

Preview Visualization

Schema: 5 Columns  
The preview includes only a limited set of columns and rows. ⓘ

Last refresh: 2 minutes ago ↻ Refine

| CustomerID<br>String | Gender<br>String | Age<br>String | Annual Income (k\$)<br>String | Spending Score (1-100)<br>String |
|----------------------|------------------|---------------|-------------------------------|----------------------------------|
| 1                    | Male             | 19            | 15                            | 39                               |
| 2                    | Male             | 21            | 15                            | 81                               |
| 3                    | Female           | 20            | 16                            | 6                                |
| 4                    | Female           | 23            | 16                            | 77                               |
| 5                    | Female           | 31            | 17                            | 40                               |
| 6                    | Female           | 22            | 17                            | 76                               |
| 7                    | Female           | 35            | 18                            | 6                                |
| 8                    | Female           | 23            | 18                            | 94                               |
| 9                    | Male             | 64            | 19                            | 3                                |
| 10                   | Female           | 30            | 19                            | 72                               |
| 11                   | Male             | 67            | 19                            | 14                               |
| 12                   | Female           | 35            | 19                            | 99                               |

SPSS Modeler:

[https://eu-gb.dataplatform.cloud.ibm.com/canvas/flows/a7b64b82-0a8d-4cb3-86b1-9e48d8ec13d2?context=cpdaas&project\\_id=7eeba2ab-b82c-4dfd-95a0-5d66a8cc77c0](https://eu-gb.dataplatform.cloud.ibm.com/canvas/flows/a7b64b82-0a8d-4cb3-86b1-9e48d8ec13d2?context=cpdaas&project_id=7eeba2ab-b82c-4dfd-95a0-5d66a8cc77c0)

Output:

