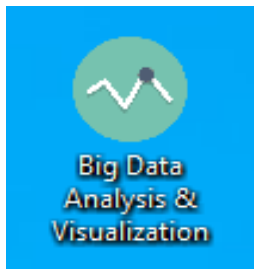




## Big Data Analysis & Visualization

Big Data Analysis has major impact on businesses in the era of the (Industry 4.0) where an ongoing huge hype for big data has been gained from academics and professionals. Big data analytics leads to valuable knowledge and promotion of innovative activity of enterprises and organizations, transforming economies in local, national and international level. In that context, “**Big Data Analysis & Visualization**” application can help to analyze critical data to support organizations in understanding their environment and in taking better decisions on time by collecting, storing and analyzing data with facilitates the creation of custom graph Visualization.

- Reduce the amount of time you spend managing data
- Analyse data in a variety of ways
- Quickly detect and address operational issues
- It helps companies gather reliable information
- It's an efficient, cost-effective solution compared to other data applications
- It helps businesses make profitable production and operational adjustments
- Reduce the amount of time you spend managing data
- Analyse data in a variety of ways
- Promote a disciplined approach to data management
- Turn disparate information into a valuable resource
- Improve the quality and consistency of information
- Increase business agility and optimization



- SQL Server Connection
- DataBase Information
- Data analysis functions
- Correlation coefficients
- Regression Model
- Execute SQL Command
- Data Store
- Messages

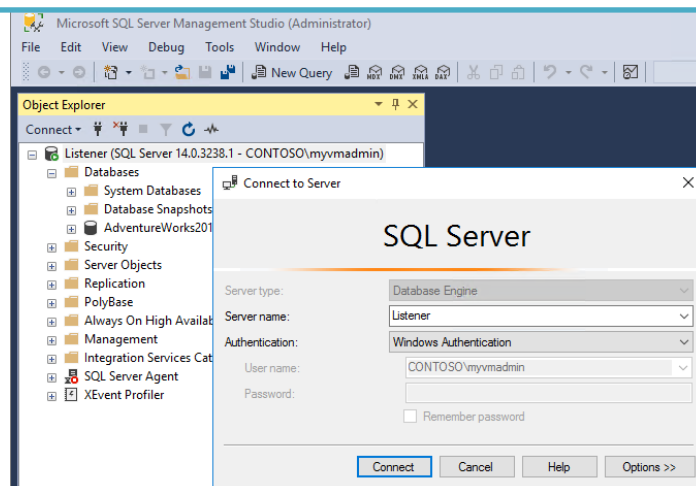
The screenshot shows the 'Big Data Analysis & Visualization' application window. It contains several sections: 'SQL Server Connection Setting' with fields for Server Name, Authentication (Windows Authentication), User Id, and Password; 'DataBase Information' with dropdowns for Database, 1st Table, and 1st Column; 'Data Analysis' with a MODE dropdown (Selected Table), 2nd Table dropdown, Visualization Method (Table View), 2nd Column dropdown, Model (Linear), Pol. Degree input, and buttons for 'Table Correlation' and 'Columns Regression'; 'Data Visualization' with Kind (Line), Marker (s), Color (blue), Line Style (-), and a 'Graph visualization' button; and 'Data Analysis Result Store' with a 'Folder Browser' button and a 'Save' button. A 'Command Line' section shows a sample SQL command: 'expl : CREATE TABLE table (column type)'. The bottom of the window indicates 'Trial Test Version' and 'version 1.0'.

1

## SQL Server Connection

This is a close-up of the 'SQL Server Connection Setting' dialog box. It features a 'Server Name' text field, an 'Authentication' dropdown menu set to 'Windows Authentication', 'User Id' and 'Password' text fields, and 'Connect' and 'Cancel' buttons.

windows authentication or sql server authentication



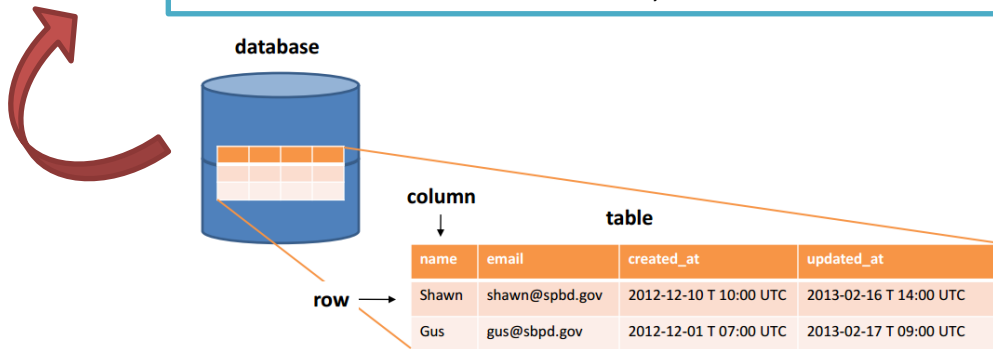
## 2

### DataBase Information

**DataBase Information**

Database  1st Table  1st Column

Show database information, include Tables and columns.



## 3

### Data analysis functions

**Data Analysis**

MODE  
Selected Table

Function  
Describe

Num. Rows

Result

2nd Table  Visualization Method  2nd Table  
Table Correlation

2nd Column  Model  1st VS 2nd Column  
Columns Regression

Pol. Degree

Command Line `expl : CREATE TABLE table (column type)`

SQL Execute

- **Head** : it shows the first number selected rows of the DataFrame.
- **Tail** : it shows the last number selected rows of the DataFrame.
- **Type** : it shows the data types of the variables of the DataFrame.
- **Describe** : it shows quick statistical summary for every numerical column of the DataFrame.
- **Null Values** : it shows missing values of the DataFrame.
- **Value Count** : it shows counts of unique values of the DataFrame.

	CODE	LOCAL	SAMPLE	TIMESTAMP	RADON	TEMPERATURE	HUMIDITY	PRESSURE	WIND
0	BARCELON	BAR_1	BAR_1	2010-03-06 11:30:00	365.6	19.7	55.0	999.3	1
1	BARCELON	BAR_1	BAR_1	2010-03-06 12:30:00	328.1	19.7	55.0	999.3	1
2	BARCELON	BAR_1	BAR_1	2010-03-06 13:30:00	302.2	19.4	55.6	999.0	1
3	BARCELON	BAR_1	BAR_1	2010-03-06 14:30:00	322.3	19.7	55.2	999.0	1
4	BARCELON	BAR_1	BAR_1	2010-03-06 15:30:00	477.0	20.0	54.0	999.2	1

	CODE	LOCAL	SAMPLE	TIMESTAMP	RADON	TEMPERATURE	HUMIDITY	PRESSURE	WIND
0	BARCELON	BAR_1	BAR_1	2010-03-06 11:30:00	365.6	19.7	55.0	999.3	1
1	BARCELON	BAR_1	BAR_1	2010-03-06 12:30:00	328.1	19.7	55.0	999.3	1
2	BARCELON	BAR_1	BAR_1	2010-03-06 13:30:00	302.2	19.4	55.6	999.0	1
3	BARCELON	BAR_1	BAR_1	2010-03-06 14:30:00	322.3	19.7	55.2	999.0	1
4	BARCELON	BAR_1	BAR_1	2010-03-06 15:30:00	477.0	20.0	54.0	999.2	1

	a	b	c
count	5.0	5.0	5.0
mean	0.43	0.68	0.4
std	0.32	0.29	0.26
min	0.13	0.37	0.12
25%	0.13	0.38	0.26
50%	0.45	0.84	0.34
75%	0.54	0.86	0.45
max	0.9	0.97	0.82

# 4

## Correlation coefficients

**Data Analysis**

MODE: Selected Table

Function: Describe

Num. Rows:

Result

Command Line: `expl : CREATE TABLE table (column type)`

SQL Execute

2nd Table:

Visualization Method: Table View

2nd Table: Table Correlation

1st VS 2nd Column: Columns Regression

Model: Linear

Pol. Degree:

Correlation coefficients are used to measure how strong a relationship is between two variables.

Heat Map



Table View

Table View showing correlation coefficients between variables A, B, C, D, and E.

	A	B	C	D	E
A	1.000000	0.830705	0.769591	-0.440535	-0.324389
B	0.830705	1.000000	0.972514	-0.007424	0.256854
C	0.769591	0.972514	1.000000	-0.092702	0.309316
D	-0.440535	-0.007424	-0.092702	1.000000	0.771163
E	-0.324389	0.256854	0.309316	0.771163	1.000000

# 5

## Regression Model

**Data Analysis**

MODE: Selected Table

Function: Describe

Num. Rows:

Result

Command Line: `expl : CREATE TABLE table (column type)`

SQL Execute

2nd Table:

Visualization Method: Table View

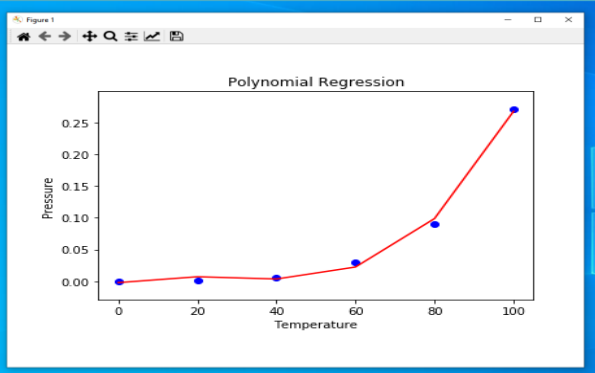
2nd Table: Table Correlation

1st VS 2nd Column: Columns Regression

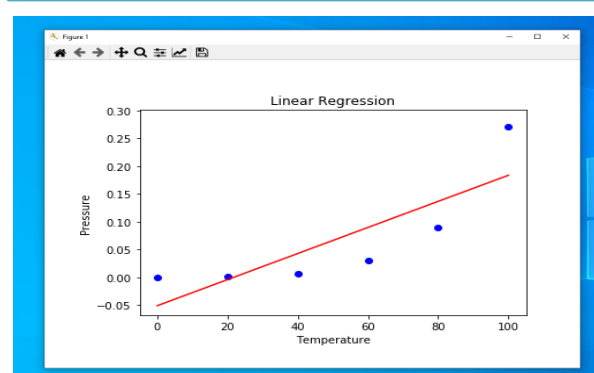
Model: Linear

Pol. Degree:

**Polynomial Regression** is a form of linear regression in which the relationship between the independent variable in  $n$ th degree polynomial.



Linear regression is a **linear model**, e.g. a model that assumes a linear relationship between the input variables.



## 6

### Execute SQL Command

**Data Analysis**

MODE: Selected Table

2nd Table: [ ]

Visualization Method: Table View

2nd Table: [ ]

Function: Describe

2nd Column: [ ]

Model: Linear

1st VS 2nd Column: [ ]

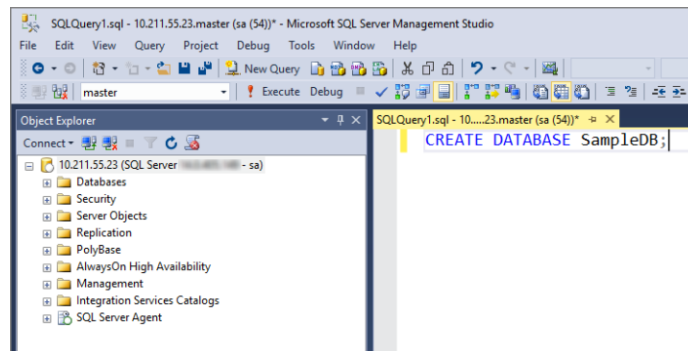
Num. Rows: [ ]

Pol. Degree: [ ]

**SQL Execute**

Command Line: `expl : CREATE TABLE table (column type)`

SQL stands for Structured Query Language which is basically a language used by databases.



## 7

### Graphic Visualization

**Data Visualization**

Kind: Line

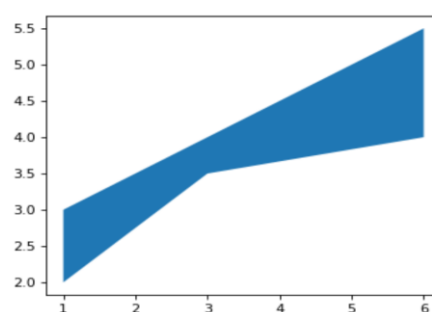
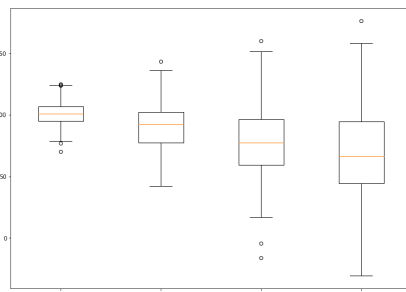
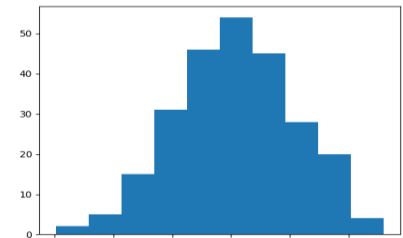
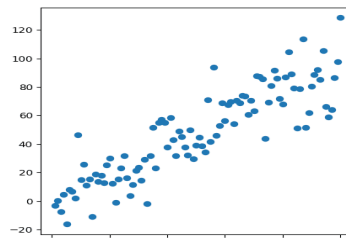
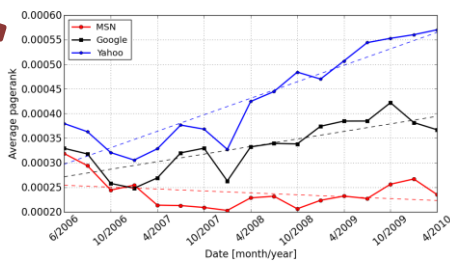
Marker: s

Color: blue

Line Style: -

MODE: Selected Table

**Graph visualization**



8

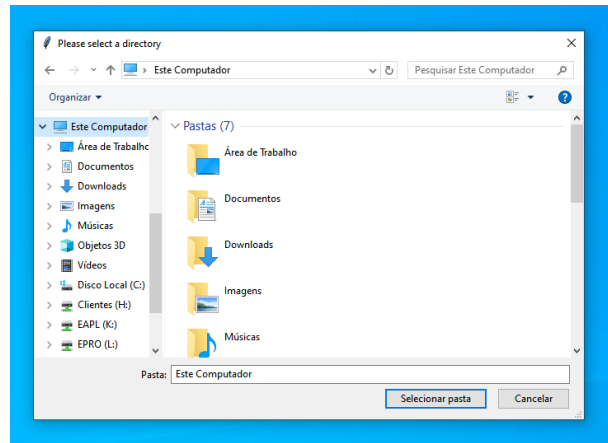
## Data Store

**Data Analysis Result Store**

expl : .xml.format

Folder Browser File Name  Save

Save data analysis result in selected folder



9

## Messages

