Quantitative Structure-Activity Relationships (QSARs) Tutorial

Single and Multi Linear Regression

SAPIENZA
UNIVERSITÀ DI ROMA

This tutorial will show you how to perform a QSAR with MLR and PLS statistical tools

In the Gnumerics file enclosed (Tutorial Data) you will find 29 sugars with different sucrose relative power and 8 different descriptors as reported in the publication.

The objective of this tutorial is to make you use of simple informatics tools to establish QSAR models.

RESEARCH ARTICLE

R.K. Singh, M.A. Khan and P.P. Singh,
S. Afr. J. Chem., 2014, **67**, 12–20,
<http://journals.sabinet.co.za/sajchem/>.

12

Rating of Sweetness by Molar Refractivity and Ionization Potential: QSAR Study of Sucrose and Guanidine Derivatives

Rajesh K. Singh*, Mohd. A. Khan and Pashupati P. Singh

Department of Chemistry, M.L.K. P.G. College, Balrampur, U.P., India.

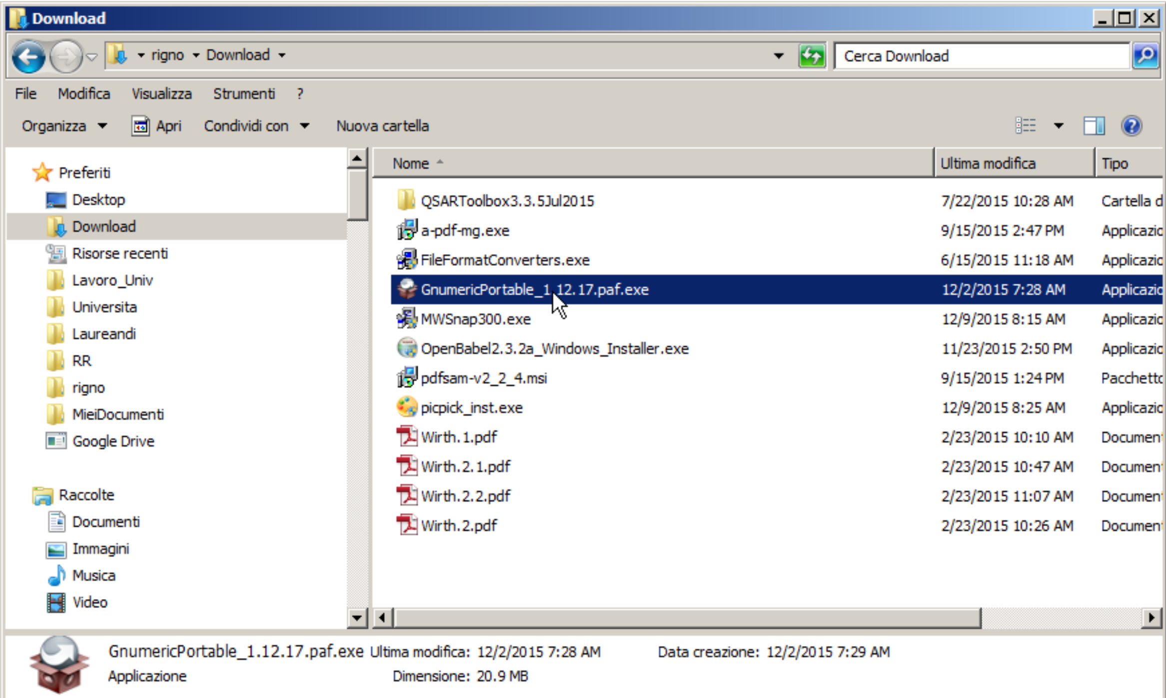
Received 31 December 2012, revised 29 August 2013, accepted 19 December 2013.

ABSTRACT

A quantitative structure activity relationship study of 31 sucrose derivatives and 30 guanidine derivatives has been undertaken. Their sweetness values, relative to sucrose (RS), have been taken from literature. The study has been made with the help of CAChe Pro software by using eight descriptors, *viz.* electron affinity, ionization potential, electrophilicity index, total energy, heat of formation, steric energy, molar refractivity and solvent accessible surface area. Multi-linear regression (MLR) analysis has been performed with different combinations of descriptors and the quality of regression has been adjudged by the correlation coefficient, cross-validation coefficient and other statistical parameters like the standard error, standard error of the estimate, degrees of freedom, etc. The study indicates that ionization potential appears an important descriptor for sucrose derivatives, whereas molar refractivity appears an important descriptor for guanidine derivatives. The ionization potential alone and in combination with the electrophilicity index, molar refractivity and solvent accessibility surface area provide dependable QSAR models for sucrose derivatives. Molar refractivity alone and in combination with solvent accessibility surface area, ionization potential and heat of formation provide dependable QSAR models for guanidine derivatives. The predicted sweetness values obtained by these QSAR models are close to observed sweetness.

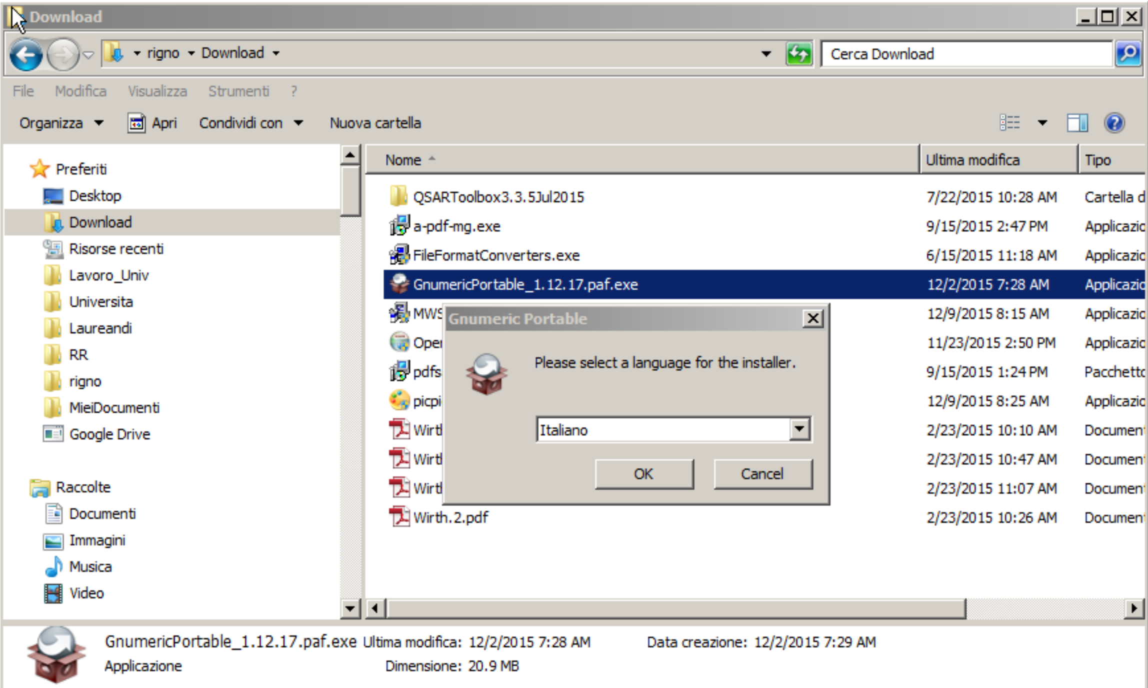
The tutorial rely on the data available from Singh's paper. You can fine the publication in elearning website

Gnumeric



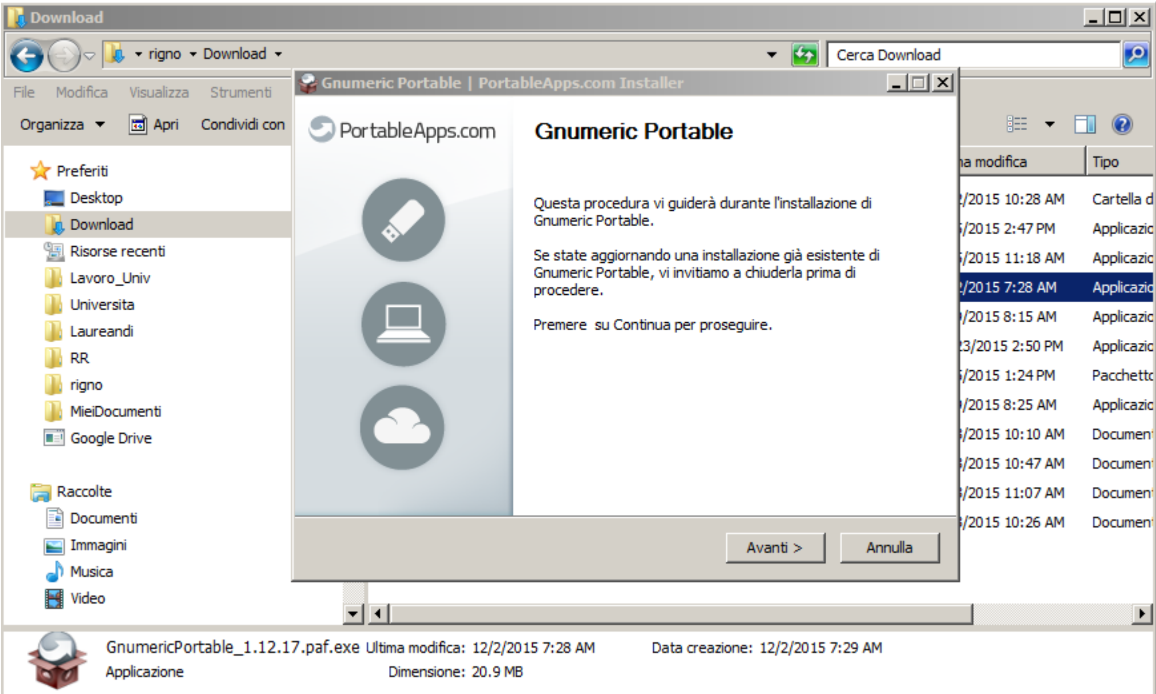
The gnumeric program will be used.
The program can be searched in the net as it is freely available.
Anyway it is also available from elarning.
Download it and install by double click on the file.

Gnumeric



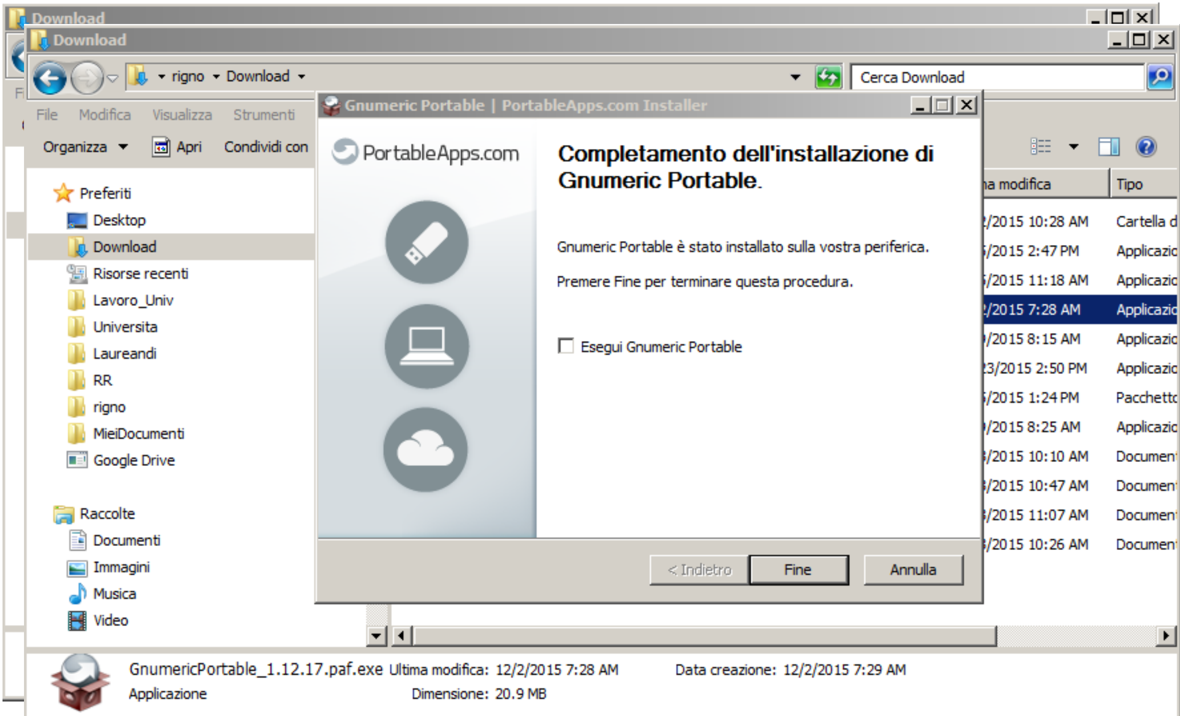
Answer to the question during the installation as shown in the slides

Gnumeric

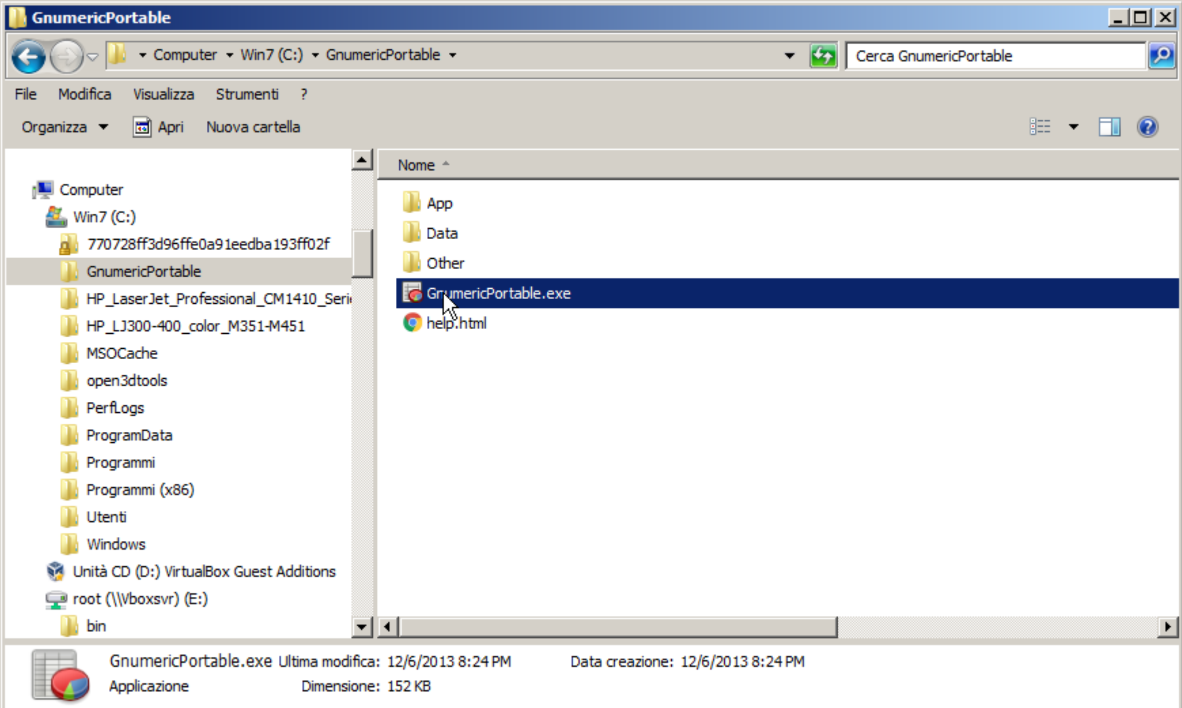


Download
rigno Download
Cerca Download
File Modifica Visualizza Strumenti
Organizza Apri Condividi con
Preferiti
Desktop
Download
Risorse recenti
Lavoro_Univ
Universita
Laureandi
RR
rigno
MieiDocumenti
Google Drive
Raccolte
Documenti
Immagini
Musica
Video
Gnumeric Portable | PortableApps.com Installer
PortableApps.com
Gnumeric Portable
Questa procedura vi guiderà durante l'installazione di Gnumeric Portable.
Se state aggiornando una installazione già esistente di Gnumeric Portable, vi invitiamo a chiuderla prima di procedere.
Premere su Continua per proseguire.
Avanti >
Annulla
Tipo
GnumericPortable_1.12.17.paf.exe Ultima modifica: 12/2/2015 7:28 AM Data creazione: 12/2/2015 7:29 AM
Applicazione Dimensione: 20.9 MB

Gnumeric

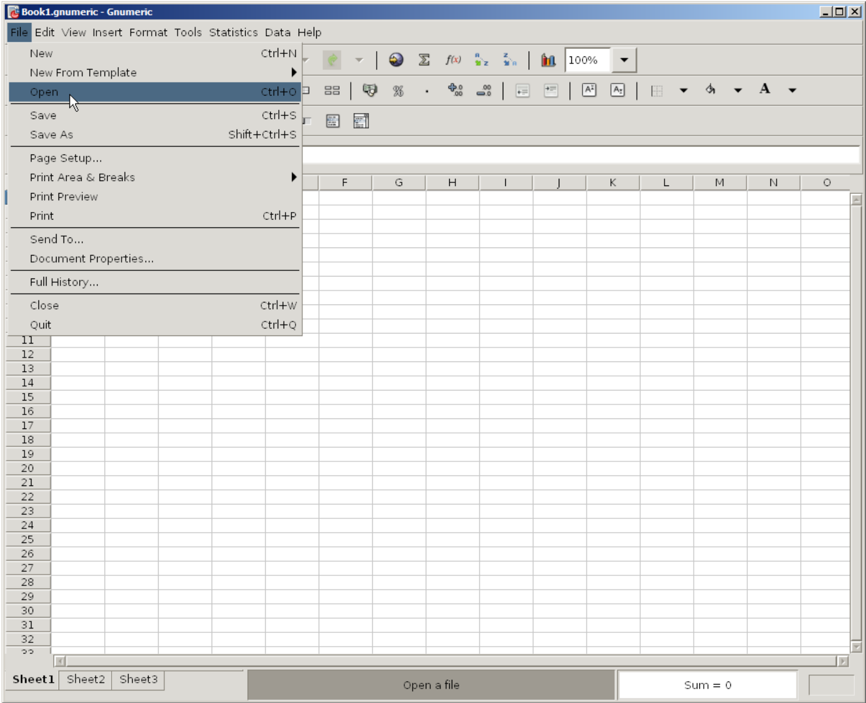

Download
Gnumeric Portable | PortableApps.com Installer
PortableApps.com
Completamento dell'installazione di Gnumeric Portable.
Gnumeric Portable è stato installato sulla vostra periferica.
Premere Fine per terminare questa procedura.
Esegui Gnumeric Portable
< Indietro
Fine
Annulla
Cerca Download
File Modifica Visualizza Strumenti
Organizza Apri Condividi con
Preferiti
Desktop
Download
Risorse recenti
Lavoro_Univ
Universita
Laureandi
RR
rigno
MieiDocumenti
Google Drive
Raccolte
Documenti
Immagini
Musica
Video
GnumericPortable_1.12.17.paf.exe Ultima modifica: 12/2/2015 7:28 AM
Data creazione: 12/2/2015 7:29 AM
Applicazione
Dimensione: 20.9 MB

Gnumeric



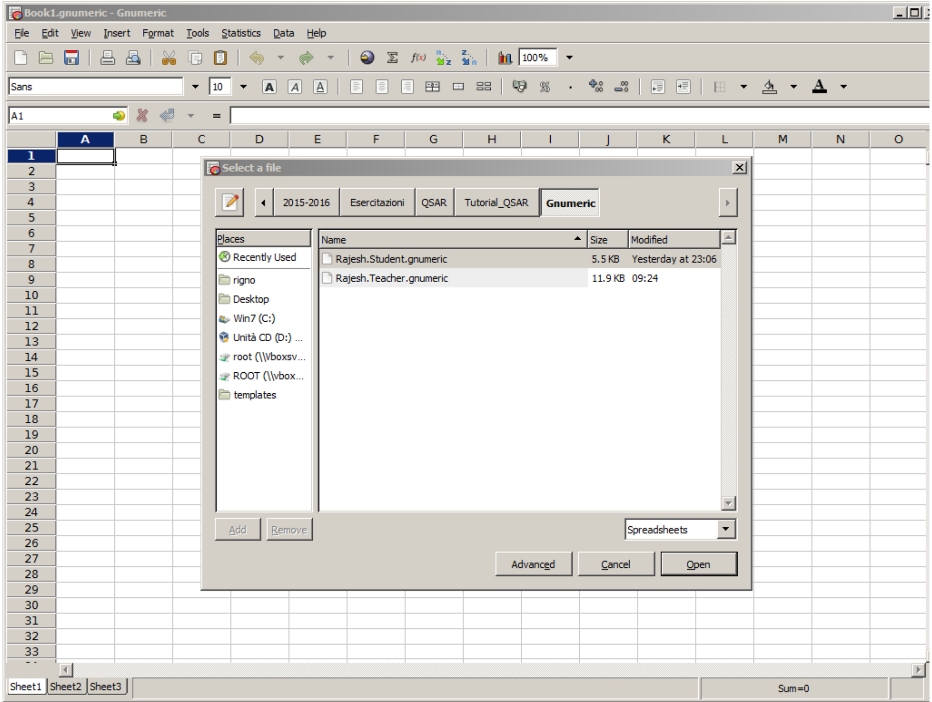
Look for the folder where gnumeric has been installed and open up it

Open the file



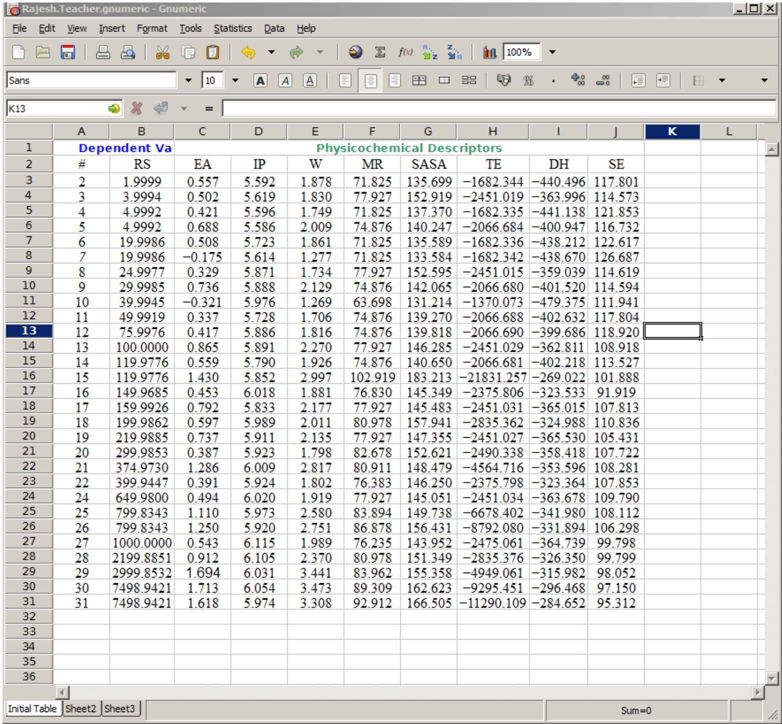
By mean of the «file» menu …

Open the file



From elearning download the file Rajesh.Student.gnumeric and open it with gnumeric

Initial Table



	A	B	C	D	E	F	G	H	I	J
1	Dependent Va				Physicochemical Descriptors					
2	#	RS	EA	IP	W	MR	SASA	TE	DH	SE
3	2	1.9999	0.557	5.592	1.878	71.825	135.699	−1682.344	−440.496	117.801
4	3	3.9994	0.502	5.619	1.830	77.927	152.919	−2451.019	−363.996	114.573
5	4	4.9992	0.421	5.596	1.749	71.825	137.370	−1682.335	−441.138	121.853
6	5	4.9992	0.688	5.586	2.009	74.876	140.247	−2066.684	−400.947	116.732
7	6	19.9986	0.508	5.723	1.861	71.825	135.589	−1682.336	−438.212	122.617
8	7	19.9986	−0.175	5.614	1.277	71.825	133.584	−1682.342	−438.670	126.687
9	8	24.9977	0.329	5.871	1.734	77.927	152.595	−2451.015	−359.039	114.619
10	9	29.9985	0.736	5.888	2.129	74.876	142.065	−2066.680	−401.520	114.594
11	10	39.9945	−0.321	5.976	1.269	63.698	131.214	−1370.073	−479.375	111.941
12	11	49.9919	0.337	5.728	1.706	74.876	139.270	−2066.688	−402.632	117.804
13	12	75.9976	0.417	5.886	1.816	74.876	139.818	−2066.690	−399.686	118.920
14	13	100.0000	0.865	5.891	2.270	77.927	146.285	−2451.029	−362.811	108.918
15	14	119.9776	0.559	5.790	1.926	74.876	140.650	−2066.681	−402.218	113.527
16	15	119.9776	1.430	5.852	2.997	102.919	183.213	−21831.257	−269.022	101.888
17	16	149.9685	0.453	6.018	1.881	76.830	145.349	−2375.806	−323.533	91.919
18	17	159.9926	0.792	5.833	2.177	77.927	145.483	−2451.031	−365.015	107.813
19	18	199.9862	0.597	5.989	2.011	80.978	157.941	−2835.362	−324.988	110.836
20	19	219.9885	0.737	5.911	2.135	77.927	147.355	−2451.027	−365.530	105.431
21	20	299.9853	0.387	5.923	1.798	82.678	152.621	−2490.338	−358.418	107.722
22	21	374.9730	1.286	6.009	2.817	80.911	148.479	−4564.716	−353.596	108.281
23	22	399.9447	0.391	5.924	1.802	76.383	146.250	−2375.798	−323.364	107.853
24	24	649.9800	0.494	6.020	1.919	77.927	145.051	−2451.034	−363.678	109.790
25	25	799.8343	1.110	5.973	2.580	83.894	149.738	−6678.402	−341.980	108.112
26	26	799.8343	1.250	5.920	2.751	86.878	156.431	−8792.080	−331.894	106.298
27	27	1000.0000	0.543	6.115	1.989	76.235	143.952	−2475.061	−364.739	99.798
28	28	2199.8851	0.912	6.105	2.370	80.978	151.349	−2835.376	−326.350	99.799
29	29	2999.8532	1.694	6.031	3.441	83.962	155.358	−4949.061	−315.982	98.052
30	30	7498.9421	1.713	6.054	3.473	89.309	162.623	−9295.451	−296.468	97.150
31	31	7498.9421	1.618	5.974	3.308	92.912	166.505	−11290.109	−284.652	95.312

The opened file should appear as in the slide.
There is a dependent variable (RS = relative sweetness) and 8 descriptors

Descriptors

Molar refractivity is calculated by the Lorenz-Lorentz formula

$$MR = \frac{n^2 - 1}{n^2 + 1} \times \frac{M}{\rho},$$

Parr *et al.* introduced the electrophilicity index (ω) in terms of the chemical potential and hardness.[30] The operational definition of the electrophilicity index may be written as,

$\omega = \mu^2/2\eta$. W in the table

The total energy (TE) of a molecular system is the sum of the total electronic energy (Eee) and the energy of internuclear repulsion (Enr).[31]

$$TE = Eee + Enr .$$

The solvent accessibility surface area (SASA) is the surface area of a biomolecule that is accessible to a solvent and is usually quoted in square angstroms. Lee and Richards first described the solvent accessible surface area (SASA) of a molecular surface.[34]

The heat of formation is defined as:

$$\Delta H_f = E_{elect} + E_{nuc} - E_{isol} + E_{atom} ,$$

DH in the table

According to the Koopman's theorem, the ionization potential is simply the eigenvalue of the highest occupied molecular orbital (HOMO) with change of sign and the electron affinity is the eigenvalue of the lowest unoccupied molecular orbital (LUMO) with change of sign.[29]

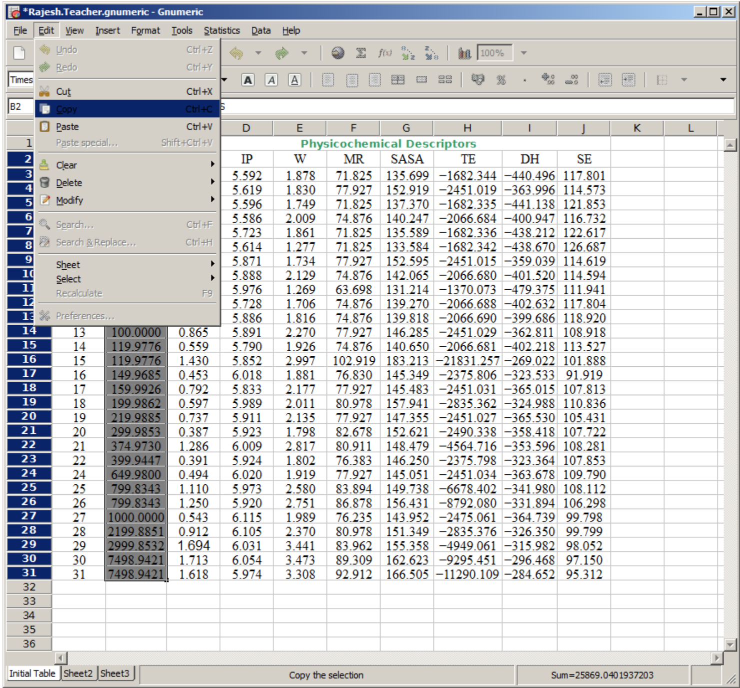
EA and IP in the table

The steric energy of a molecule is the sum of the molecular mechanics potential energies calculated for the bonds, bond angles, dihedral angles, nonbonded atoms and so forth. It is specific to mechanics and depends upon the force-field used.[33]

SE in the table

Here are described the 8 used descriptors

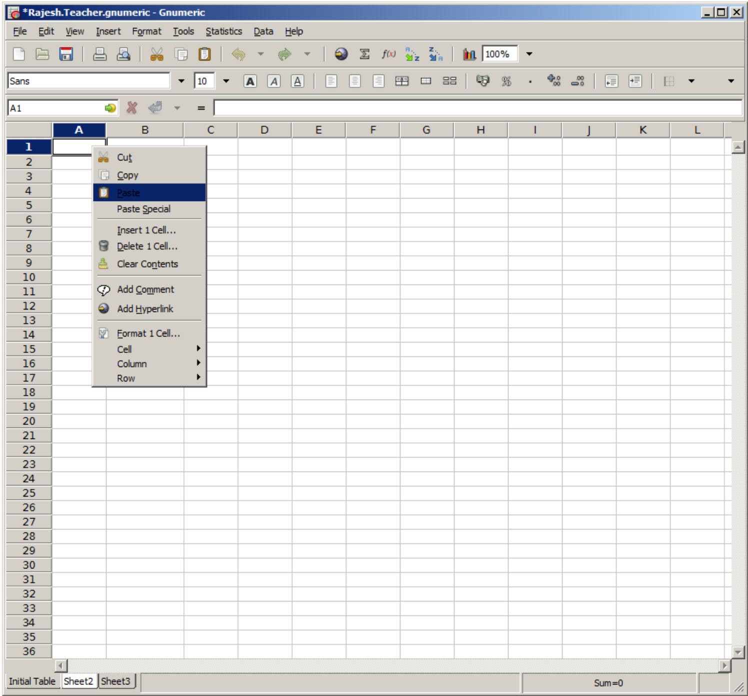
Check dependent variable (RS)



	A	B	C	D	E	F	G	H	I	J
1				Physicochemical Descriptors						
2				IP	W	MR	SASA	TE	DH	SE
3				5.592	1.878	71.825	135.699	−1682.344	−440.496	117.801
4				5.619	1.830	77.927	152.919	−2451.019	−363.996	114.573
5				5.596	1.749	71.825	137.370	−1682.335	−441.138	121.853
6				5.586	2.009	74.876	140.247	−2066.684	−400.947	116.732
7				5.723	1.861	71.825	135.589	−1682.336	−438.212	122.617
8				5.614	1.277	71.825	133.584	−1682.342	−438.670	126.687
9				5.871	1.734	77.927	152.595	−2451.015	−359.039	114.619
10				5.888	2.129	74.876	142.065	−2066.680	−401.520	114.594
11				5.976	1.269	63.698	131.214	−1370.073	−479.375	111.941
12				5.728	1.706	74.876	139.270	−2066.688	−402.632	117.804
13				5.886	1.816	74.876	139.818	−2066.690	−399.686	118.920
14	13	100.0000	0.865	5.891	2.270	77.927	146.285	−2451.029	−362.811	108.918
15	14	119.9776	0.559	5.790	1.926	74.876	140.650	−2066.681	−402.218	113.527
16	15	119.9776	1.430	5.852	2.997	102.919	183.213	−21831.257	−269.022	101.888
17	16	149.9685	0.453	6.018	1.881	76.830	145.349	−2375.806	−323.533	91.919
18	17	159.9926	0.792	5.833	2.177	77.927	145.483	−2451.031	−365.015	107.813
19	18	199.9862	0.597	5.989	2.011	80.978	157.941	−2835.362	−324.988	110.836
20	19	219.9885	0.737	5.911	2.135	77.927	147.355	−2451.027	−365.530	105.431
21	20	299.9853	0.387	5.923	1.798	82.678	152.621	−2490.338	−358.418	107.722
22	21	374.9730	1.286	6.009	2.817	80.911	148.479	−4564.716	−353.596	108.281
23	22	399.9447	0.391	5.924	1.802	76.383	146.250	−2375.798	−323.364	107.853
24	24	649.9800	0.494	6.020	1.919	77.927	145.051	−2451.034	−363.678	109.790
25	25	799.8343	1.110	5.973	2.580	83.894	149.738	−6678.402	−341.980	108.112
26	26	799.8343	1.250	5.920	2.751	86.878	156.431	−8792.080	−331.894	106.298
27	27	1000.0000	0.543	6.115	1.989	76.235	143.952	−2475.061	−364.739	99.798
28	28	2199.8851	0.912	6.105	2.370	80.978	151.349	−2835.376	−326.350	99.799
29	29	2999.8532	1.694	6.031	3.441	83.962	155.358	−4949.061	−315.982	98.052
30	30	7498.9421	1.713	6.054	3.473	89.309	162.623	−9295.451	−296.468	97.150
31	31	7498.9421	1.618	5.974	3.308	92.912	166.505	−11290.109	−284.652	95.312

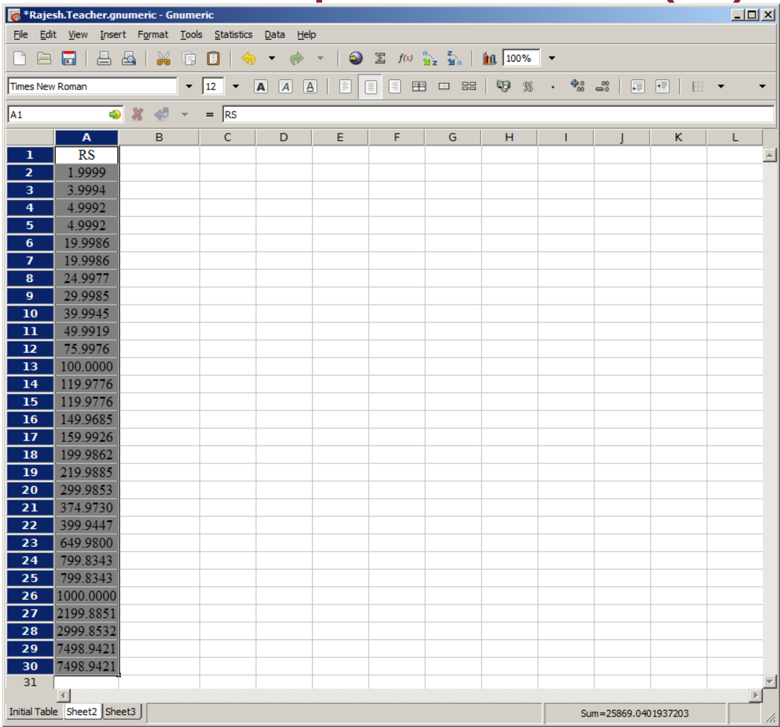
The first thing to do is to check the linear distribution of the dependent variable RS.
So copy the column of data belonging to RS it in the memory.

Check dependent variable (RS)



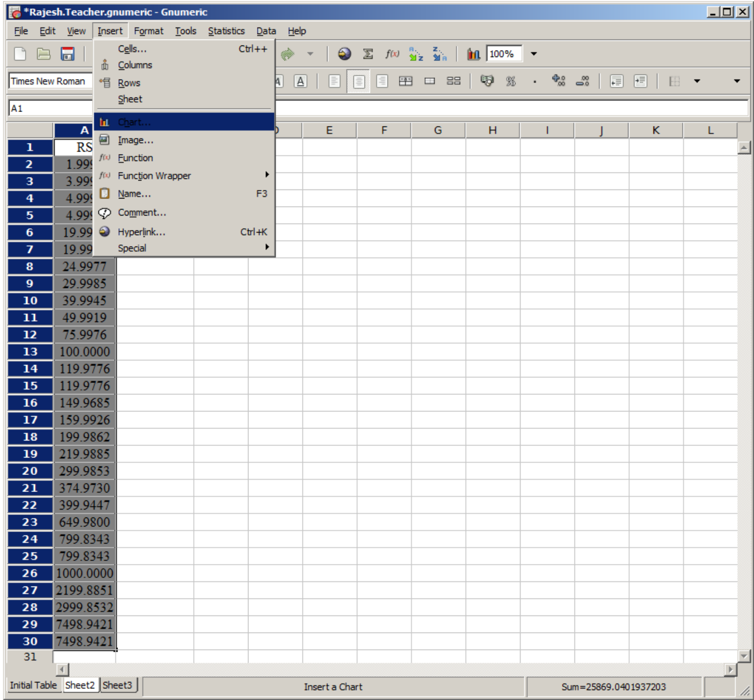
…. and paste it the into a new sheet

Check dependent variable (RS)



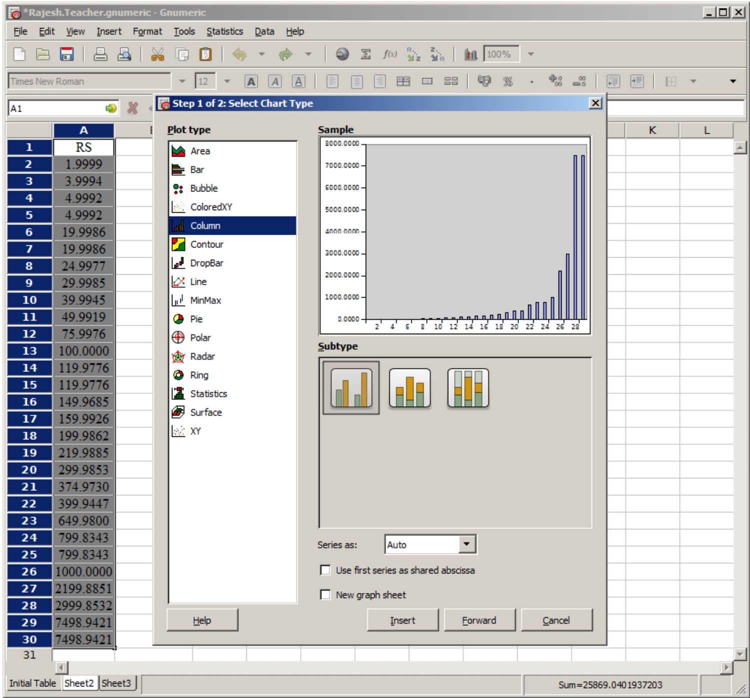
	A
1	RS
2	1.9999
3	3.9994
4	4.9992
5	4.9992
6	19.9986
7	19.9986
8	24.9977
9	29.9985
10	39.9945
11	49.9919
12	75.9976
13	100.0000
14	119.9776
15	119.9776
16	149.9685
17	159.9926
18	199.9862
19	219.9885
20	299.9853
21	374.9730
22	399.9447
23	649.9800
24	799.8343
25	799.8343
26	1000.0000
27	2199.8851
28	2999.8532
29	7498.9421
30	7498.9421

Check dependent variable (RS)



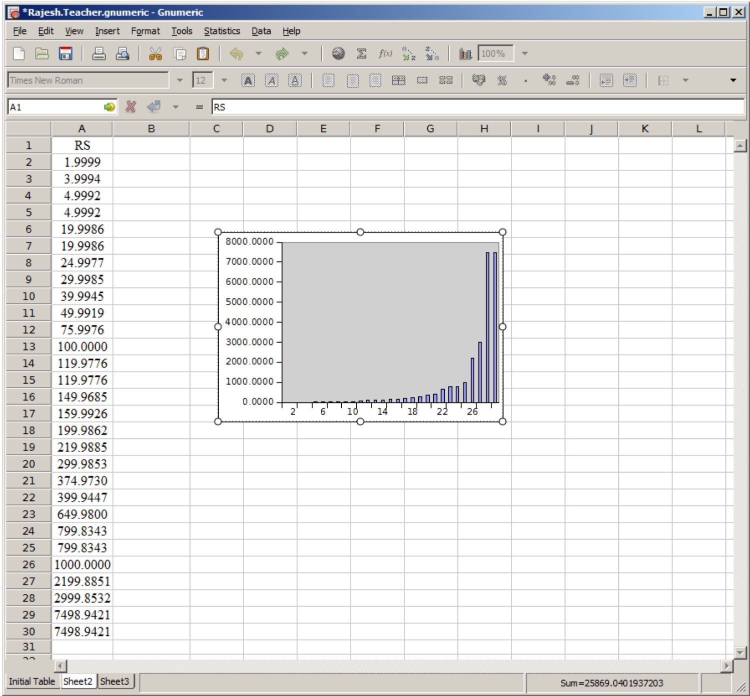
To check for linearity, just insert a column type chart

Check dependent variable (RS)



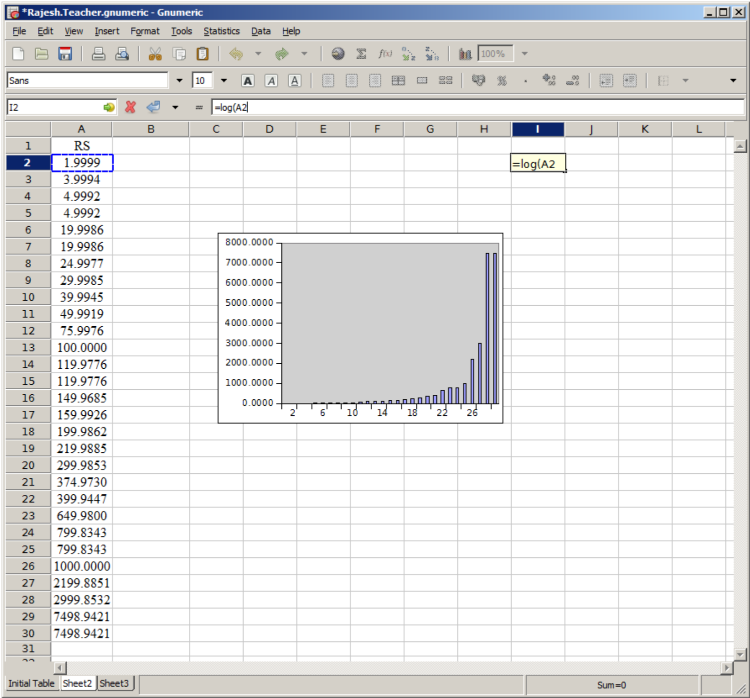
As you can see, the preview is showing the plot

Check dependent variable (RS)



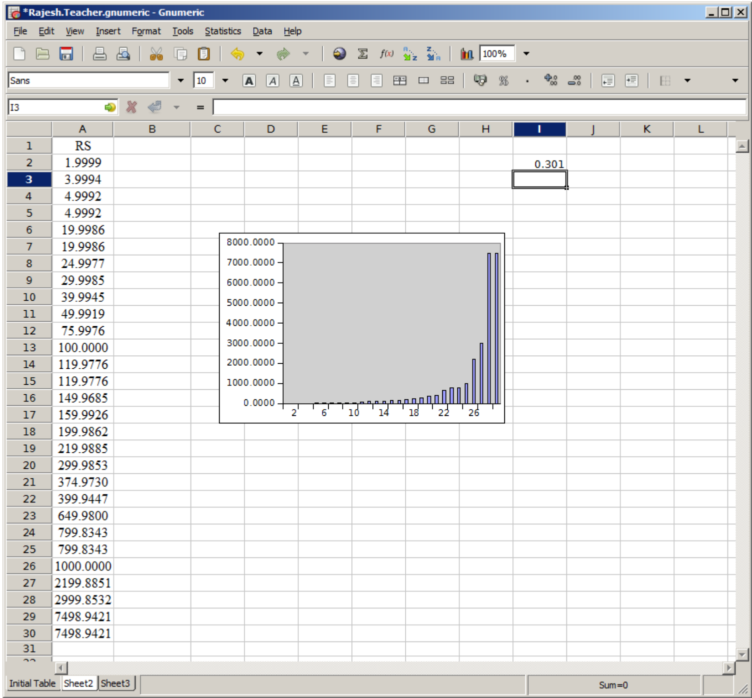
Then click in the area of the sheet grid and the plot will appear.
Clearly there is no linear distribution of the dependent variable

Check dependent variable (RS)

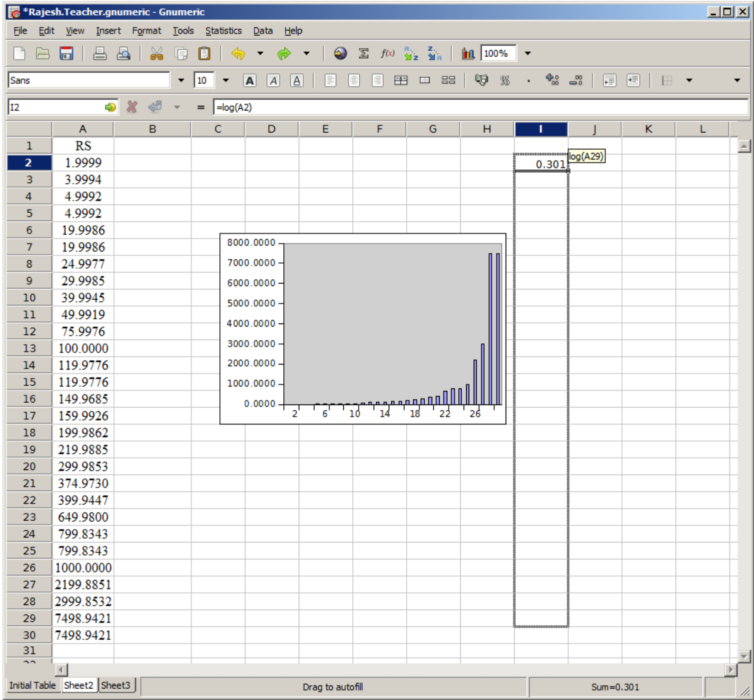


Let’s transform it in the logarithm scale as shown in the slide

Check dependent variable (RS)

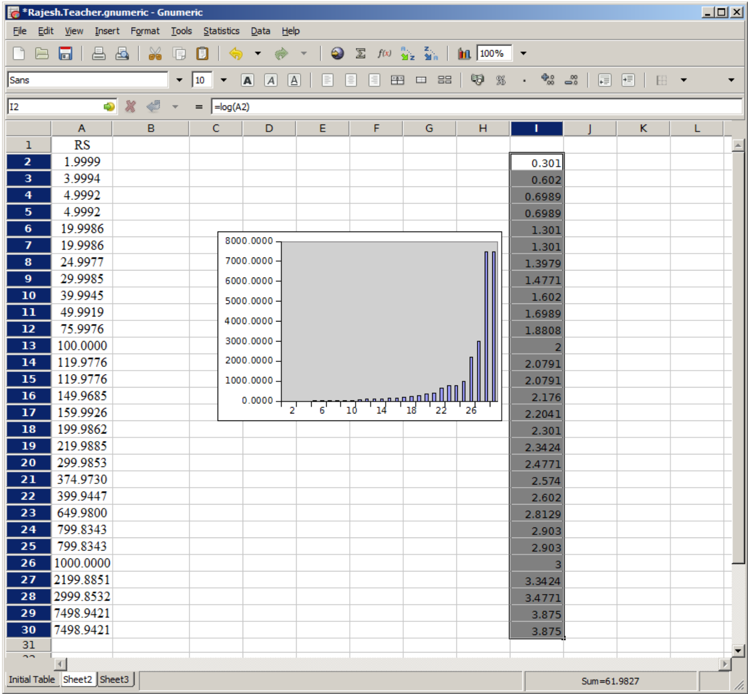


Check dependent variable (RS)



Then apply the formula for the cell «I2» to the other cells just grabbing the small square in the cell right bottom corner

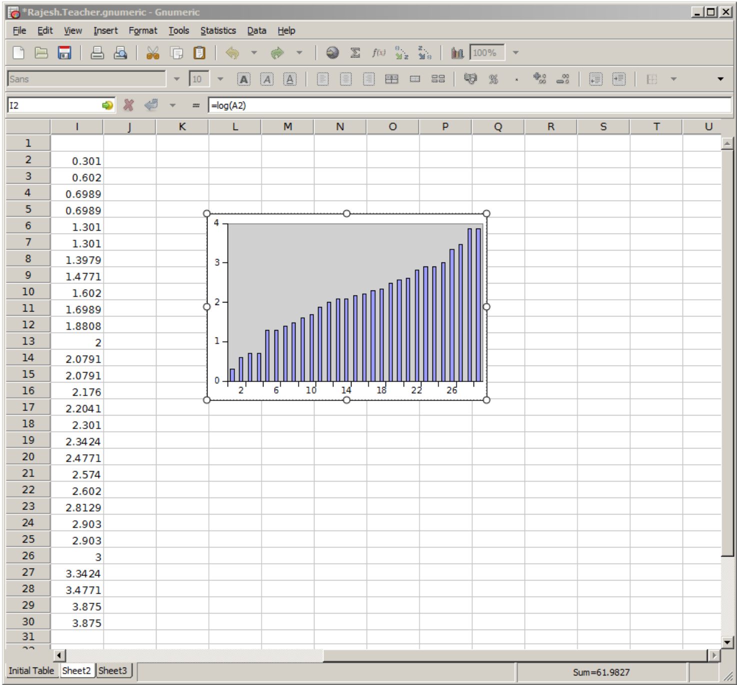
Check dependent variable (RS)



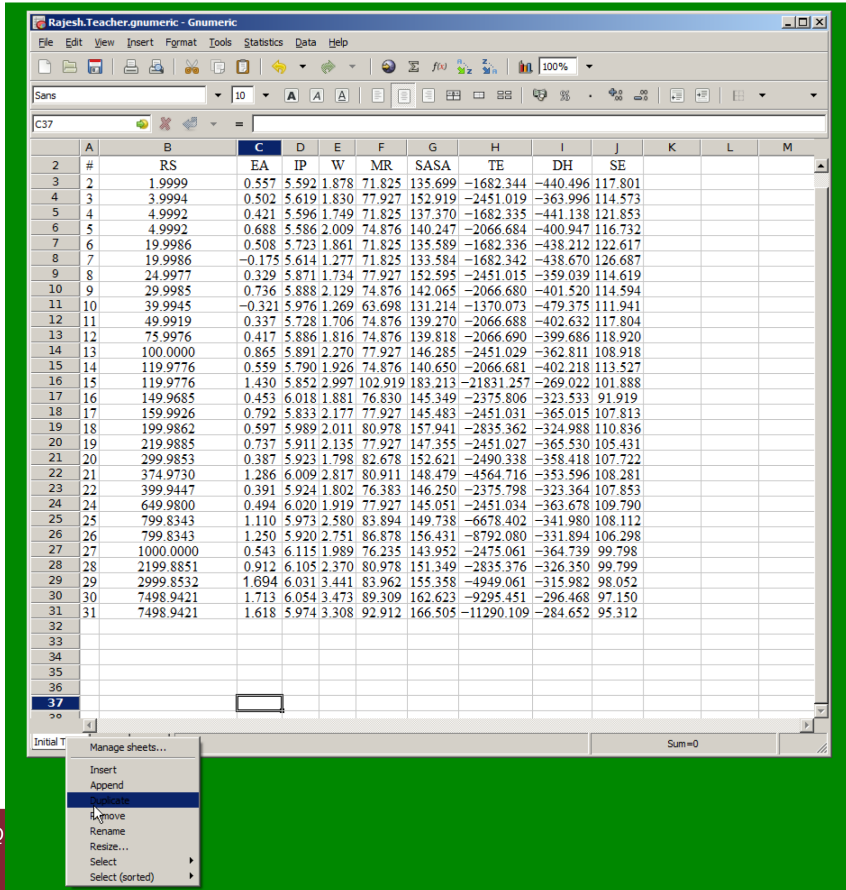
	A	I
1	RS	
2	1.9999	0.301
3	3.9994	0.602
4	4.9992	0.6989
5	4.9992	0.6989
6	19.9986	1.301
7	19.9986	1.301
8	24.9977	1.3979
9	29.9985	1.4771
10	39.9945	1.602
11	49.9919	1.6989
12	75.9976	1.8808
13	100.0000	2
14	119.9776	2.0791
15	119.9776	2.0791
16	149.9685	2.176
17	159.9926	2.2041
18	199.9862	2.301
19	219.9885	2.3424
20	299.9853	2.4771
21	374.9730	2.574
22	399.9447	2.602
23	649.9800	2.8129
24	799.8343	2.903
25	799.8343	2.903
26	1000.0000	3
27	2199.8851	3.3424
28	2999.8532	3.4771
29	7498.9421	3.875
30	7498.9421	3.875

And all the data will be transformed

Check dependent variable (RS)



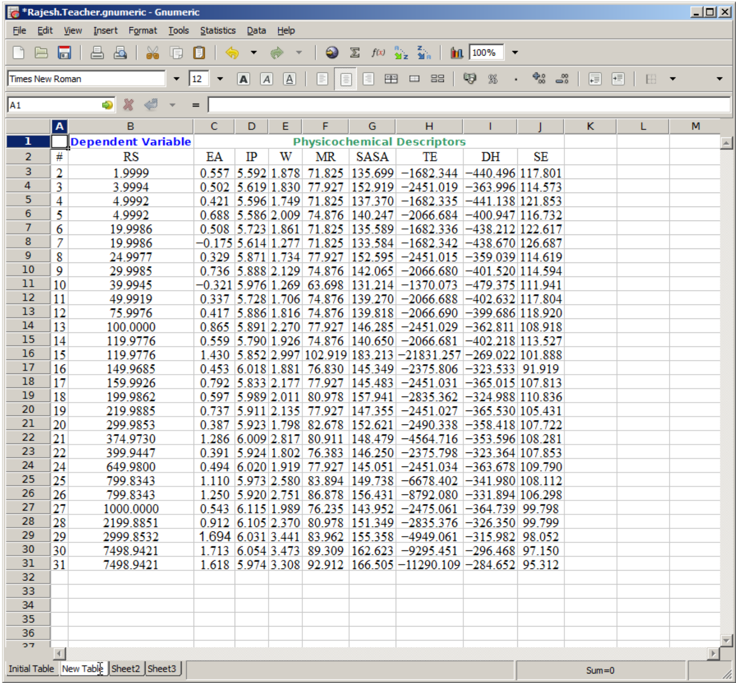
Make the plot as shown before and now the linearity is present.



#	RS	EA	IP	W	MR	SASA	TE	DH	SE
2	1.9999	0.557	5.592	1.878	71.825	135.699	−1682.344	−440.496	117.801
3	3.9994	0.502	5.619	1.830	77.927	152.919	−2451.019	−363.996	114.573
4	4.9992	0.421	5.596	1.749	71.825	137.370	−1682.335	−441.138	121.853
5	4.9992	0.688	5.586	2.009	74.876	140.247	−2066.684	−400.947	116.732
6	19.9986	0.508	5.723	1.861	71.825	135.589	−1682.336	−438.212	122.617
7	19.9986	−0.175	5.614	1.277	71.825	133.584	−1682.342	−438.670	126.687
8	24.9977	0.329	5.871	1.734	77.927	152.595	−2451.015	−359.039	114.619
9	29.9985	0.736	5.888	2.129	74.876	142.065	−2066.680	−401.520	114.594
10	39.9945	−0.321	5.976	1.269	63.698	131.214	−1370.073	−479.375	111.941
11	49.9919	0.337	5.728	1.706	74.876	139.270	−2066.688	−402.632	117.804
12	75.9976	0.417	5.886	1.816	74.876	139.818	−2066.690	−399.686	118.920
13	100.0000	0.865	5.891	2.270	77.927	146.285	−2451.029	−362.811	108.918
14	119.9776	0.559	5.790	1.926	74.876	140.650	−2066.681	−402.218	113.527
15	119.9776	1.430	5.852	2.997	102.919	183.213	−21831.257	−269.022	101.888
16	149.9685	0.453	6.018	1.881	76.830	145.349	−2375.806	−323.533	91.919
17	159.9926	0.792	5.833	2.177	77.927	145.483	−2451.031	−365.015	107.813
18	199.9862	0.597	5.989	2.011	80.978	157.941	−2835.362	−324.988	110.836
19	219.9885	0.737	5.911	2.135	77.927	147.355	−2451.027	−365.530	105.431
20	299.9853	0.387	5.923	1.798	82.678	152.621	−2490.338	−358.418	107.722
21	374.9730	1.286	6.009	2.817	80.911	148.479	−4564.716	−353.596	108.281
22	399.9447	0.391	5.924	1.802	76.383	146.250	−2375.798	−323.364	107.853
24	649.9800	0.494	6.020	1.919	77.927	145.051	−2451.034	−363.678	109.790
25	799.8343	1.110	5.973	2.580	83.894	149.738	−6678.402	−341.980	108.112
26	799.8343	1.250	5.920	2.751	86.878	156.431	−8792.080	−331.894	106.298
27	1000.0000	0.543	6.115	1.989	76.235	143.952	−2475.061	−364.739	99.798
28	2199.8851	0.912	6.105	2.370	80.978	151.349	−2835.376	−326.350	99.799
29	2999.8532	1.694	6.031	3.441	83.962	155.358	−4949.061	−315.982	98.052
30	7498.9421	1.713	6.054	3.473	89.309	162.623	−9295.451	−296.468	97.150
31	7498.9421	1.618	5.974	3.308	92.912	166.505	−11290.109	−284.652	95.312

Duplicate the sheet by right clicking on the sheet name.

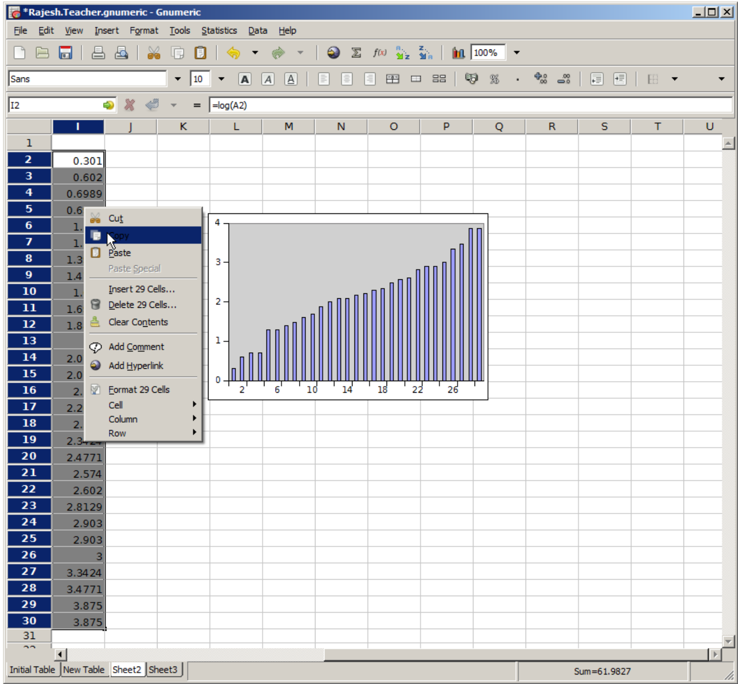
Check dependent variable (RS)



#	Dependent Variable RS	EA	IP	W	MR	SASA	TE	DH	SE
		Physicochemical Descriptors							
2	1.9999	0.557	5.592	1.878	71.825	135.699	−1682.344	−440.496	117.801
3	3.9994	0.502	5.619	1.830	77.927	152.919	−2451.019	−363.996	114.573
4	4.9992	0.421	5.596	1.749	71.825	137.370	−1682.335	−441.138	121.853
5	4.9992	0.688	5.586	2.009	74.876	140.247	−2066.684	−400.947	116.732
6	19.9986	0.508	5.723	1.861	71.825	135.589	−1682.336	−438.212	122.617
7	19.9986	−0.175	5.614	1.277	71.825	133.584	−1682.342	−438.670	126.687
8	24.9977	0.329	5.871	1.734	77.927	152.595	−2451.015	−359.039	114.619
9	29.9985	0.736	5.888	2.129	74.876	142.065	−2066.680	−401.520	114.594
10	39.9945	−0.321	5.976	1.269	63.698	131.214	−1370.073	−479.375	111.941
11	49.9919	0.337	5.728	1.706	74.876	139.270	−2066.688	−402.632	117.804
12	75.9976	0.417	5.886	1.816	74.876	139.818	−2066.690	−399.686	118.920
13	100.0000	0.865	5.891	2.270	77.927	146.285	−2451.029	−362.811	108.918
14	119.9776	0.559	5.790	1.926	74.876	140.650	−2066.681	−402.218	113.527
15	119.9776	1.430	5.852	2.997	102.919	183.213	−21831.257	−269.022	101.888
16	149.9685	0.453	6.018	1.881	76.830	145.349	−2375.806	−323.533	91.919
17	159.9926	0.792	5.833	2.177	77.927	145.483	−2451.031	−365.015	107.813
18	199.9862	0.597	5.989	2.011	80.978	157.941	−2835.362	−324.988	110.836
19	219.9885	0.737	5.911	2.135	77.927	147.355	−2451.027	−365.530	105.431
20	299.9853	0.387	5.923	1.798	82.678	152.621	−2490.338	−358.418	107.722
21	374.9730	1.286	6.009	2.817	80.911	148.479	−4564.716	−353.596	108.281
22	399.9447	0.391	5.924	1.802	76.383	146.250	−2375.798	−323.364	107.853
24	649.9800	0.494	6.020	1.919	77.927	145.051	−2451.034	−363.678	109.790
25	799.8343	1.110	5.973	2.580	83.894	149.738	−6678.402	−341.980	108.112
26	799.8343	1.250	5.920	2.751	86.878	156.431	−8792.080	−331.894	106.298
27	1000.0000	0.543	6.115	1.989	76.235	143.952	−2475.061	−364.739	99.798
28	2199.8851	0.912	6.105	2.370	80.978	151.349	−2835.376	−326.350	99.799
29	2999.8532	1.694	6.031	3.441	83.962	155.358	−4949.061	−315.982	98.052
30	7498.9421	1.713	6.054	3.473	89.309	162.623	−9295.451	−296.468	97.150
31	7498.9421	1.618	5.974	3.308	92.912	166.505	−11290.109	−284.652	95.312

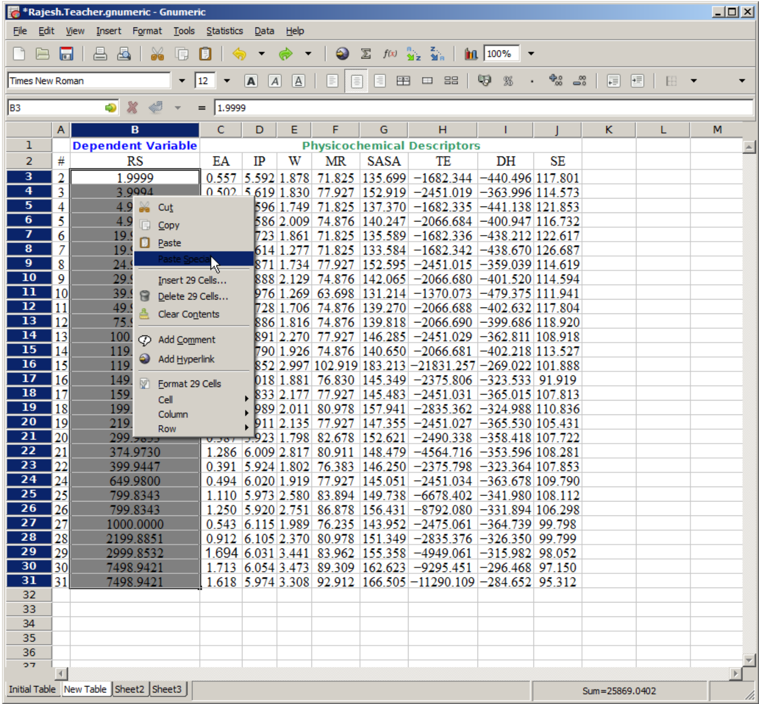
Rename the sheet into «New Table»

Check dependent variable (RS)



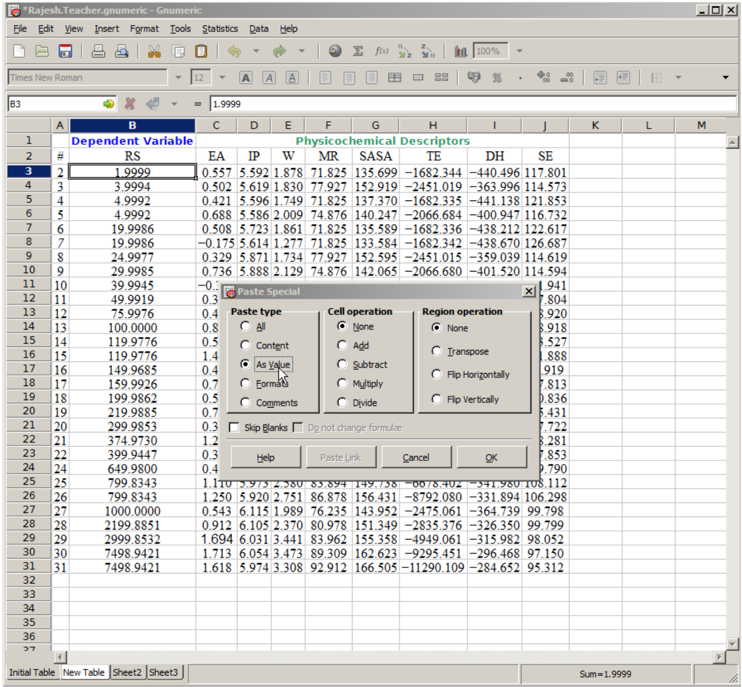
Highlight the transformed RS variable and copy into the memory with the right click

Check dependent variable (RS)



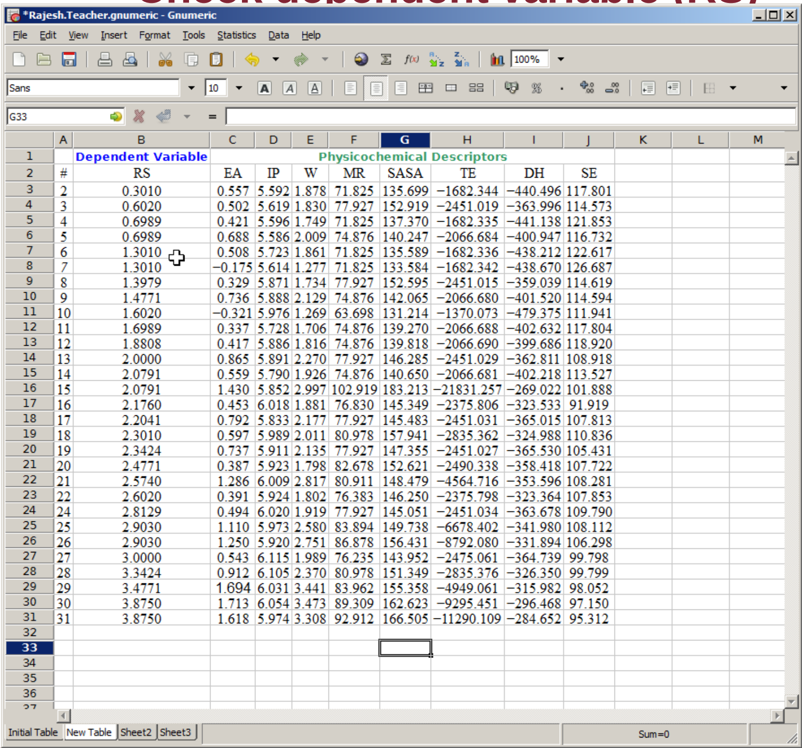
Then paste it to the «New Table» sheet. Use the «paste special» option

Check dependent variable (RS)



… and select «as value»

Check dependent variable (RS)



A	B	C	D	E	F	G	H	I	J
	Dependent Variable	Physicochemical Descriptors							
#	RS	EA	IP	W	MR	SASA	TE	DH	SE
2	0.3010	0.557	5.592	1.878	71.825	135.699	−1682.344	−440.496	117.801
3	0.6020	0.502	5.619	1.830	77.927	152.919	−2451.019	−363.996	114.573
4	0.6989	0.421	5.596	1.749	71.825	137.370	−1682.335	−441.138	121.853
5	0.6989	0.688	5.586	2.009	74.876	140.247	−2066.684	−400.947	116.732
6	1.3010	0.508	5.723	1.861	71.825	135.589	−1682.336	−438.212	122.617
7	1.3010	−0.175	5.614	1.277	71.825	133.584	−1682.342	−438.670	126.687
8	1.3979	0.329	5.871	1.734	77.927	152.595	−2451.015	−359.039	114.619
9	1.4771	0.736	5.888	2.129	74.876	142.065	−2066.680	−401.520	114.594
10	1.6020	−0.321	5.976	1.269	63.698	131.214	−1370.073	−479.375	111.941
11	1.6989	0.337	5.728	1.706	74.876	139.270	−2066.688	−402.632	117.804
12	1.8808	0.417	5.886	1.816	74.876	139.818	−2066.690	−399.686	118.920
13	2.0000	0.865	5.891	2.270	77.927	146.285	−2451.029	−362.811	108.918
14	2.0791	0.559	5.790	1.926	74.876	140.650	−2066.681	−402.218	113.527
15	2.0791	1.430	5.852	2.997	102.919	183.213	−21831.257	−269.022	101.888
16	2.1760	0.453	6.018	1.881	76.830	145.349	−2375.806	−323.533	91.919
17	2.2041	0.792	5.833	2.177	77.927	145.483	−2451.031	−365.015	107.813
18	2.3010	0.597	5.989	2.011	80.978	157.941	−2835.362	−324.988	110.836
19	2.3424	0.737	5.911	2.135	77.927	147.355	−2451.027	−365.530	105.431
20	2.4771	0.387	5.923	1.798	82.678	152.621	−2490.338	−358.418	107.722
21	2.5740	1.286	6.009	2.817	80.911	148.479	−4564.716	−353.596	108.281
22	2.6020	0.391	5.924	1.802	76.383	146.250	−2375.798	−323.364	107.853
24	2.8129	0.494	6.020	1.919	77.927	145.051	−2451.034	−363.678	109.790
25	2.9030	1.110	5.973	2.580	83.894	149.738	−6678.402	−341.980	108.112
26	2.9030	1.250	5.920	2.751	86.878	156.431	−8792.080	−331.894	106.298
27	3.0000	0.543	6.115	1.989	76.235	143.952	−2475.061	−364.739	99.798
28	3.3424	0.912	6.105	2.370	80.978	151.349	−2835.376	−326.350	99.799
29	3.4771	1.694	6.031	3.441	83.962	155.358	−4949.061	−315.982	98.052
30	3.8750	1.713	6.054	3.473	89.309	162.623	−9295.451	−296.468	97.150
31	3.8750	1.618	5.974	3.308	92.912	166.505	−11290.109	−284.652	95.312

Now you have the dependent variable fixed

Independent variables analysis

Notice that the 8 molecular descriptors have very different magnitudes. To guarantee the comparability of the MLR coefficients (see previous lesson on QSAR), proceed to the normalization of all the descriptors (for instance for the interval [0; 1]).

The first step is to calculate the MAX and MIN values of each descriptor with the functions “=MAX(range)” and “=MIN(range)” as described in the following slides

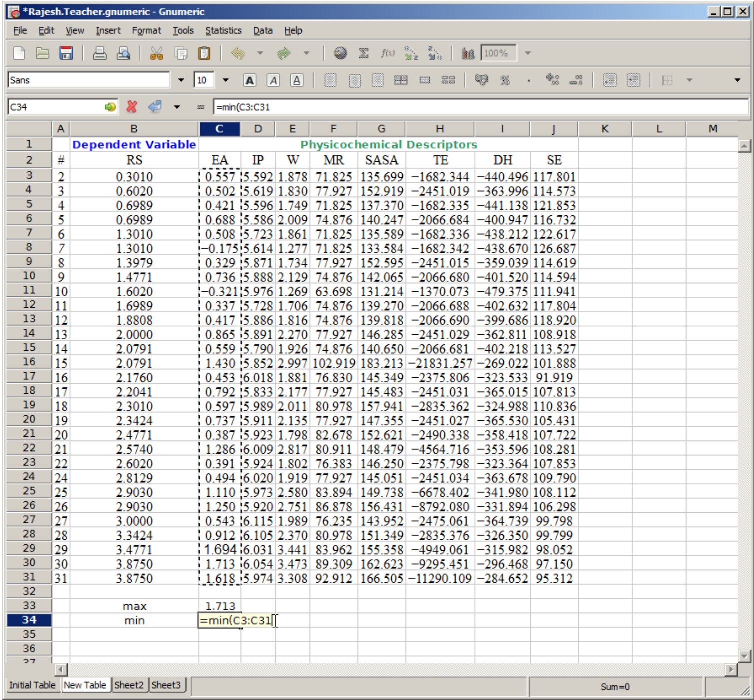
Once we have checked the dependent variable (RS)
we need to analyse the independent variables (descriptors).

Independent variables analysis

	A	B	C	D	E	F	G	H	I	J
1		Dependent Variable	Physicochemical Descriptors							
2	#	RS	EA	IP	W	MR	SASA	TE	DH	SE
3	2	0.3010	0.557	5.592	1.878	71.825	135.699	−1682.344	−440.496	117.801
4	3	0.6020	0.502	5.619	1.830	77.927	152.919	−2451.019	−363.996	114.573
5	4	0.6989	0.421	5.596	1.749	71.825	137.370	−1682.335	−441.138	121.853
6	5	0.6989	0.688	5.586	2.009	74.876	140.247	−2066.684	−400.947	116.732
7	6	1.3010	0.508	5.723	1.861	71.825	135.589	−1682.336	−438.212	122.617
8	7	1.3010	−0.175	5.614	1.277	71.825	133.584	−1682.342	−438.670	126.687
9	8	1.3979	0.329	5.871	1.734	77.927	152.595	−2451.015	−359.039	114.619
10	9	1.4771	0.736	5.888	2.129	74.876	142.065	−2066.680	−401.520	114.594
11	10	1.6020	−0.321	5.976	1.269	63.698	131.214	−1370.073	−479.375	111.941
12	11	1.6989	0.337	5.728	1.706	74.876	139.270	−2066.688	−402.632	117.804
13	12	1.8808	0.417	5.886	1.816	74.876	139.818	−2066.690	−399.686	118.920
14	13	2.0000	0.865	5.891	2.270	77.927	146.285	−2451.029	−362.811	108.918
15	14	2.0791	0.559	5.790	1.926	74.876	140.650	−2066.681	−402.218	113.527
16	15	2.0791	1.430	5.852	2.997	102.919	183.213	−21831.257	−269.022	101.888
17	16	2.1760	0.453	6.018	1.881	76.830	145.349	−2375.806	−323.533	91.919
18	17	2.2041	0.792	5.833	2.177	77.927	145.483	−2451.031	−365.015	107.813
19	18	2.3010	0.597	5.989	2.011	80.978	157.941	−2835.362	−324.988	110.836
20	19	2.3424	0.737	5.911	2.135	77.927	147.355	−2451.027	−365.530	105.431
21	20	2.4771	0.387	5.923	1.798	82.678	152.621	−2490.338	−358.418	107.722
22	21	2.5740	1.286	6.009	2.817	80.911	148.479	−4564.716	−353.596	108.281
23	22	2.6020	0.391	5.924	1.802	76.383	146.250	−2375.798	−323.364	107.853
24	24	2.8129	0.494	6.020	1.919	77.927	145.051	−2451.034	−363.678	109.790
25	25	2.9030	1.110	5.973	2.580	83.894	149.738	−6678.402	−341.980	108.112
26	26	2.9030	1.250	5.920	2.751	86.878	156.431	−8792.080	−331.894	106.298
27	27	3.0000	0.543	6.115	1.989	76.235	143.952	−2475.061	−364.739	99.798
28	28	3.3424	0.912	6.105	2.370	80.978	151.349	−2835.376	−326.350	99.799
29	29	3.4771	1.694	6.031	3.441	83.962	155.358	−4949.061	−315.982	98.052
30	30	3.8750	1.713	6.054	3.473	89.309	162.623	−9295.451	−296.468	97.150
31	31	3.8750	1.618	5.974	3.308	92.912	166.505	−11290.109	−284.652	95.312
32										
33		max	=max(C3:C31							
34		min								

First we calculate the maximum and the minimum values for each colum

Independent variables analysis



	A	B	C	D	E	F	G	H	I	J
1		Dependent Variable	Physicochemical Descriptors							
2	#	RS	EA	IP	W	MR	SASA	TE	DH	SE
3	2	0.3010	0.557	5.592	1.878	71.825	135.699	−1682.344	−440.496	117.801
4	3	0.6020	0.502	5.619	1.830	77.927	152.919	−2451.019	−363.996	114.573
5	4	0.6989	0.421	5.596	1.749	71.825	137.370	−1682.335	−441.138	121.853
6	5	0.6989	0.688	5.586	2.009	74.876	140.247	−2066.684	−400.947	116.732
7	6	1.3010	0.508	5.723	1.861	71.825	135.589	−1682.336	−438.212	122.617
8	7	1.3010	−0.175	5.614	1.277	71.825	133.584	−1682.342	−438.670	126.687
9	8	1.3979	0.329	5.871	1.734	77.927	152.595	−2451.015	−359.039	114.619
10	9	1.4771	0.736	5.888	2.129	74.876	142.065	−2066.680	−401.520	114.594
11	10	1.6020	−0.321	5.976	1.269	63.698	131.214	−1370.073	−479.375	111.941
12	11	1.6989	0.337	5.728	1.706	74.876	139.270	−2066.688	−402.632	117.804
13	12	1.8808	0.417	5.886	1.816	74.876	139.818	−2066.690	−399.686	118.920
14	13	2.0000	0.865	5.891	2.270	77.927	146.285	−2451.029	−362.811	108.918
15	14	2.0791	0.559	5.790	1.926	74.876	140.650	−2066.681	−402.218	113.527
16	15	2.0791	1.430	5.852	2.997	102.919	183.213	−21831.257	−269.022	101.888
17	16	2.1760	0.453	6.018	1.881	76.830	145.349	−2375.806	−323.533	91.919
18	17	2.2041	0.792	5.833	2.177	77.927	145.483	−2451.031	−365.015	107.813
19	18	2.3010	0.597	5.989	2.011	80.978	157.941	−2835.362	−324.988	110.836
20	19	2.3424	0.737	5.911	2.135	77.927	147.355	−2451.027	−365.530	105.431
21	20	2.4771	0.387	5.923	1.798	82.678	152.621	−2490.338	−358.418	107.722
22	21	2.5740	1.286	6.009	2.817	80.911	148.479	−4564.716	−353.596	108.281
23	22	2.6020	0.391	5.924	1.802	76.383	146.250	−2375.798	−323.364	107.853
24	24	2.8129	0.494	6.020	1.919	77.927	145.051	−2451.034	−363.678	109.790
25	25	2.9030	1.110	5.973	2.580	83.894	149.738	−6678.402	−341.980	108.112
26	26	2.9030	1.250	5.920	2.751	86.878	156.431	−8792.080	−331.894	106.298
27	27	3.0000	0.543	6.115	1.989	76.235	143.952	−2475.061	−364.739	99.798
28	28	3.3424	0.912	6.105	2.370	80.978	151.349	−2835.376	−326.350	99.799
29	29	3.4771	1.694	6.031	3.441	83.962	155.358	−4949.061	−315.982	98.052
30	30	3.8750	1.713	6.054	3.473	89.309	162.623	−9295.451	−296.468	97.150
31	31	3.8750	1.618	5.974	3.308	92.912	166.505	−11290.109	−284.652	95.312
32										
33		max	1.713							
34		min	=min(C3:C31							

The method is similar to that already shown.
Just do the same as displayed in the slide.

Independent variables analysis

*Rajesh.Teacher.gnumeric - Gnumeric

C33 =max(C3:C31)

	A	B	C	D	E	F	G	H	I	J
1		Dependent Variable	Physicochemical Descriptors							
2	#	RS	EA	IP	W	MR	SASA	TE	DH	SE
3	2	0.3010	0.557	5.592	1.878	71.825	135.699	−1682.344	−440.496	117.801
4	3	0.6020	0.502	5.619	1.830	77.927	152.919	−2451.019	−363.996	114.573
5	4	0.6989	0.421	5.596	1.749	71.825	137.370	−1682.335	−441.138	121.853
6	5	0.6989	0.688	5.586	2.009	74.876	140.247	−2066.684	−400.947	116.732
7	6	1.3010	0.508	5.723	1.861	71.825	135.589	−1682.336	−438.212	122.617
8	7	1.3010	−0.175	5.614	1.277	71.825	133.584	−1682.342	−438.670	126.687
9	8	1.3979	0.329	5.871	1.734	77.927	152.595	−2451.015	−359.039	114.619
10	9	1.4771	0.736	5.888	2.129	74.876	142.065	−2066.680	−401.520	114.594
11	10	1.6020	−0.321	5.976	1.269	63.698	131.214	−1370.073	−479.375	111.941
12	11	1.6989	0.337	5.728	1.706	74.876	139.270	−2066.688	−402.632	117.804
13	12	1.8808	0.417	5.886	1.816	74.876	139.818	−2066.690	−399.686	118.920
14	13	2.0000	0.865	5.891	2.270	77.927	146.285	−2451.029	−362.811	108.918
15	14	2.0791	0.559	5.790	1.926	74.876	140.650	−2066.681	−402.218	113.527
16	15	2.0791	1.430	5.852	2.997	102.919	183.213	−21831.257	−269.022	101.888
17	16	2.1760	0.453	6.018	1.881	76.830	145.349	−2375.806	−323.533	91.919
18	17	2.2041	0.792	5.833	2.177	77.927	145.483	−2451.031	−365.015	107.813
19	18	2.3010	0.597	5.989	2.011	80.978	157.941	−2835.362	−324.988	110.836
20	19	2.3424	0.737	5.911	2.135	77.927	147.355	−2451.027	−365.530	105.431
21	20	2.4771	0.387	5.923	1.798	82.678	152.621	−2490.338	−358.418	107.722
22	21	2.5740	1.286	6.009	2.817	80.911	148.479	−4564.716	−353.596	108.281
23	22	2.6020	0.391	5.924	1.802	76.383	146.250	−2375.798	−323.364	107.853
24	24	2.8129	0.494	6.020	1.919	77.927	145.051	−2451.034	−363.678	109.790
25	25	2.9030	1.110	5.973	2.580	83.894	149.738	−6678.402	−341.980	108.112
26	26	2.9030	1.250	5.920	2.751	86.878	156.431	−8792.080	−331.894	106.298
27	27	3.0000	0.543	6.115	1.989	76.235	143.952	−2475.061	−364.739	99.798
28	28	3.3424	0.912	6.105	2.370	80.978	151.349	−2835.376	−326.350	99.799
29	29	3.4771	1.694	6.031	3.441	83.962	155.358	−4949.061	−315.982	98.052
30	30	3.8750	1.713	6.054	3.473	89.309	162.623	−9295.451	−296.468	97.150
31	31	3.8750	1.618	5.974	3.308	92.912	166.505	−11290.109	−284.652	95.312
32										
33		max	1.713							
34		min	0.321							

Initial Table | New Table | Sheet2 | Sheet3

Sum=1.392

Independent variables analysis

*Rajesh.Teacher.gnumeric - Gnumeric

C33 =max(C3:C31)

#	Dependent Variable RS	Physicochemical Descriptors EA	IP	W	MR	SASA	TE	DH	SE
2	0.3010	0.557	5.592	1.878	71.825	135.699	−1682.344	−440.496	117.801
3	0.6020	0.502	5.619	1.830	77.927	152.919	−2451.019	−363.996	114.573
4	0.6989	0.421	5.596	1.749	71.825	137.370	−1682.335	−441.138	121.853
5	0.6989	0.688	5.586	2.009	74.876	140.247	−2066.684	−400.947	116.732
6	1.3010	0.508	5.723	1.861	71.825	135.589	−1682.336	−438.212	122.617
7	1.3010	−0.175	5.614	1.277	71.825	133.584	−1682.342	−438.670	126.687
8	1.3979	0.329	5.871	1.734	77.927	152.595	−2451.015	−359.039	114.619
9	1.4771	0.736	5.888	2.129	74.876	142.065	−2066.680	−401.520	114.594
10	1.6020	−0.321	5.976	1.269	63.698	131.214	−1370.073	−479.375	111.941
11	1.6989	0.337	5.728	1.706	74.876	139.270	−2066.688	−402.632	117.804
12	1.8808	0.417	5.886	1.816	74.876	139.818	−2066.690	−399.686	118.920
13	2.0000	0.865	5.891	2.270	77.927	146.285	−2451.029	−362.811	108.918
14	2.0791	0.559	5.790	1.926	74.876	140.650	−2066.681	−402.218	113.527
15	2.0791	1.430	5.852	2.997	102.919	183.213	−21831.257	−269.022	101.888
16	2.1760	0.453	6.018	1.881	76.830	145.349	−2375.806	−323.533	91.919
17	2.2041	0.792	5.833	2.177	77.927	145.483	−2451.031	−365.015	107.813
18	2.3010	0.597	5.989	2.011	80.978	157.941	−2835.362	−324.988	110.836
19	2.3424	0.737	5.911	2.135	77.927	147.355	−2451.027	−365.530	105.431
20	2.4771	0.387	5.923	1.798	82.678	152.621	−2490.338	−358.418	107.722
21	2.5740	1.286	6.009	2.817	80.911	148.479	−4564.716	−353.596	108.281
22	2.6020	0.391	5.924	1.802	76.383	146.250	−2375.798	−323.364	107.853
24	2.8129	0.494	6.020	1.919	77.927	145.051	−2451.034	−363.678	109.790
25	2.9030	1.110	5.973	2.580	83.894	149.738	−6678.402	−341.980	108.112
26	2.9030	1.250	5.920	2.751	86.878	156.431	−8792.080	−331.894	106.298
27	3.0000	0.543	6.115	1.989	76.235	143.952	−2475.061	−364.739	99.798
28	3.3424	0.912	6.105	2.370	80.978	151.349	−2835.376	−326.350	99.799
29	3.4771	1.694	6.031	3.441	83.962	155.358	−4949.061	−315.982	98.052
30	3.8750	1.713	6.054	3.473	89.309	162.623	−9295.451	−296.468	97.150
31	3.8750	1.618	5.974	3.308	92.912	166.505	−11290.109	−284.652	95.312
	max	1.713	max(J3:J31)						
	min	0.321	min(J3:J31)						

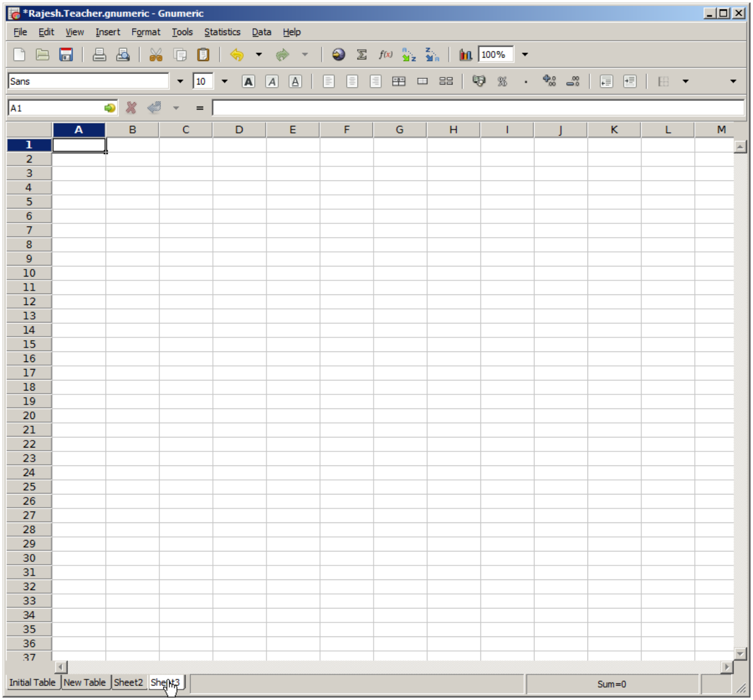
Initial Table | New Table | Sheet2 | Sheet3 — Drag to autofill — Sum=1.392

Independent variables analysis

	Dependent Variable	Physicochemical Descriptors							
#	RS	EA	IP	W	MR	SASA	TE	DH	SE
2	0.3010	0.557	5.592	1.878	71.825	135.699	−1682.344	−440.496	117.801
3	0.6020	0.502	5.619	1.830	77.927	152.919	−2451.019	−363.996	114.573
4	0.6989	0.421	5.596	1.749	71.825	137.370	−1682.335	−441.138	121.853
5	0.6989	0.688	5.586	2.009	74.876	140.247	−2066.684	−400.947	116.732
6	1.3010	0.508	5.723	1.861	71.825	135.589	−1682.336	−438.212	122.617
7	1.3010	−0.175	5.614	1.277	71.825	133.584	−1682.342	−438.670	126.687
8	1.3979	0.329	5.871	1.734	77.927	152.595	−2451.015	−359.039	114.619
9	1.4771	0.736	5.888	2.129	74.876	142.065	−2066.680	−401.520	114.594
10	1.6020	−0.321	5.976	1.269	63.698	131.214	−1370.073	−479.375	111.941
11	1.6989	0.337	5.728	1.706	74.876	139.270	−2066.688	−402.632	117.804
12	1.8808	0.417	5.886	1.816	74.876	139.818	−2066.690	−399.686	118.920
13	2.0000	0.865	5.891	2.270	77.927	146.285	−2451.029	−362.811	108.918
14	2.0791	0.559	5.790	1.926	74.876	140.650	−2066.681	−402.218	113.527
15	2.0791	1.430	5.852	2.997	102.919	183.213	−21831.257	−269.022	101.888
16	2.1760	0.453	6.018	1.881	76.830	145.349	−2375.806	−323.533	91.919
17	2.2041	0.792	5.833	2.177	77.927	145.483	−2451.031	−365.015	107.813
18	2.3010	0.597	5.989	2.011	80.978	157.941	−2835.362	−324.988	110.836
19	2.3424	0.737	5.911	2.135	77.927	147.355	−2451.027	−365.530	105.431
20	2.4771	0.387	5.923	1.798	82.678	152.621	−2490.338	−358.418	107.722
21	2.5740	1.286	6.009	2.817	80.911	148.479	−4564.716	−353.596	108.281
22	2.6020	0.391	5.924	1.802	76.383	146.250	−2375.798	−323.364	107.853
24	2.8129	0.494	6.020	1.919	77.927	145.051	−2451.034	−363.678	109.790
25	2.9030	1.110	5.973	2.580	83.894	149.738	−6678.402	−341.980	108.112
26	2.9030	1.250	5.920	2.751	86.878	156.431	−8792.080	−331.894	106.298
27	3.0000	0.543	6.115	1.989	76.235	143.952	−2475.061	−364.739	99.798
28	3.3424	0.912	6.105	2.370	80.978	151.349	−2835.376	−326.350	99.799
29	3.4771	1.694	6.031	3.441	83.962	155.358	−4949.061	−315.982	98.052
30	3.8750	1.713	6.054	3.473	89.309	162.623	−9295.451	−296.468	97.150
31	3.8750	1.618	5.974	3.308	92.912	166.505	−11290.109	−284.652	95.312
	max	1.713	6.115	3.473	102.919	183.213	1370.073	269.022	126.687
	min	0.321	5.586	1.269	63.698	131.214	21831.257	479.375	91.919

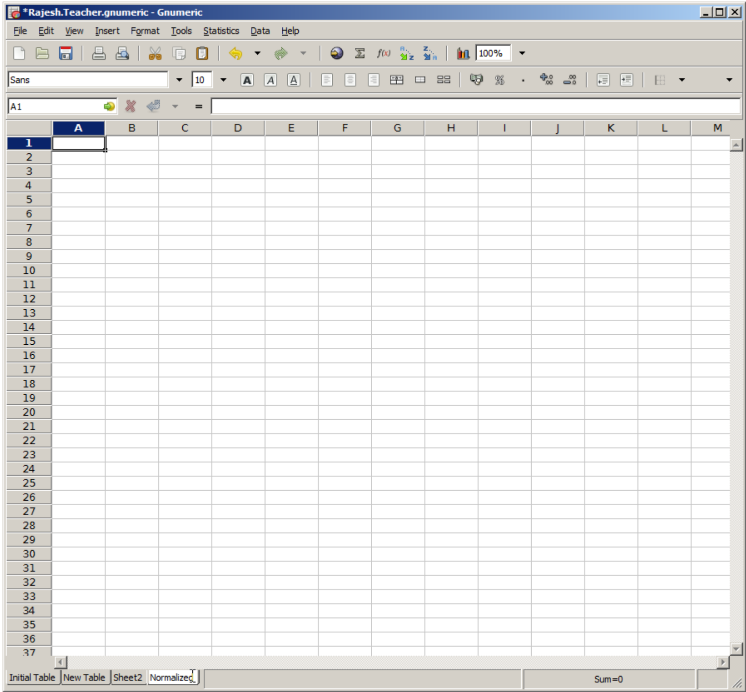
And all the min and max values are promptly calculated

Independent variables analysis



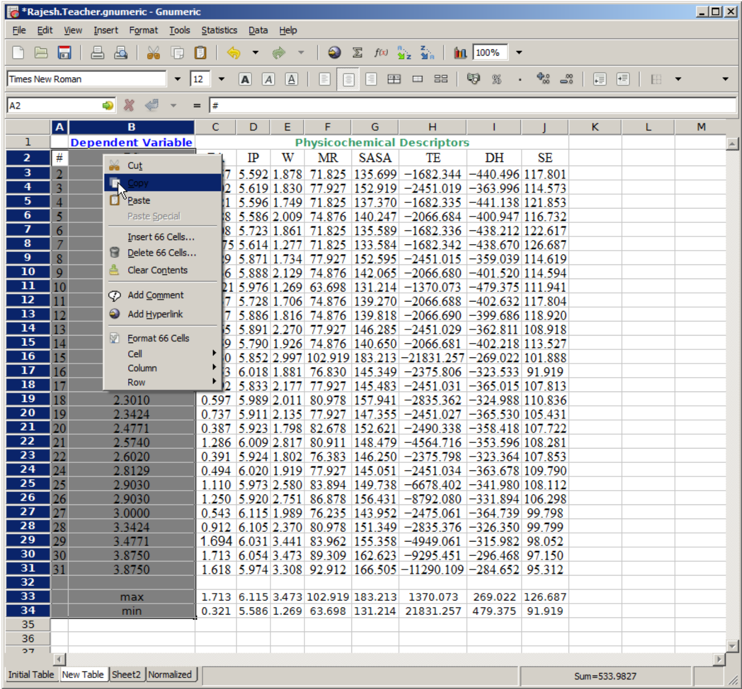
In a new sheet ....

Independent variables analysis



… renamed «Normalized» …

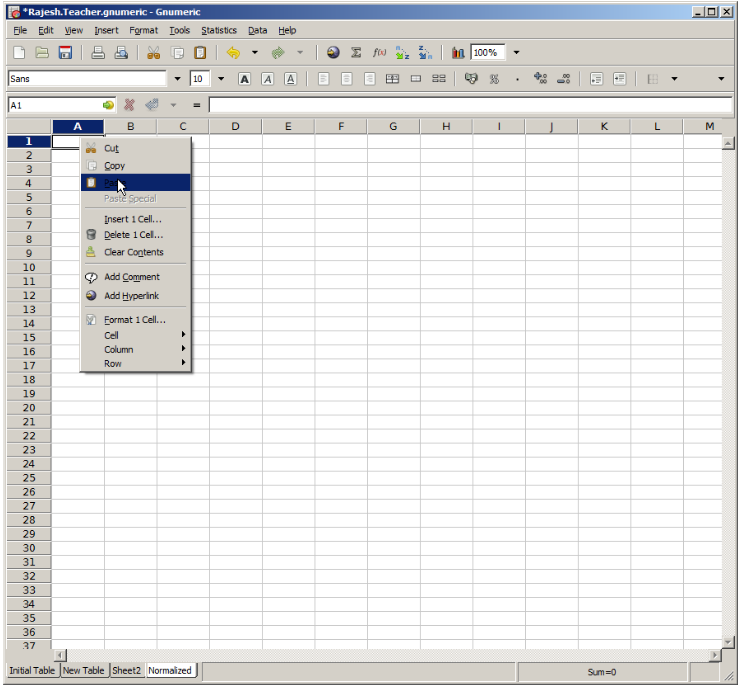
Independent variables analysis



	A	B	C	D	E	F	G	H	I	J
1		Dependent Variable	Physicochemical Descriptors							
2	#	[illegible]	[illegible]	IP	W	MR	SASA	TE	DH	SE
3	2	[illegible]	[illegible]	5.592	1.878	71.825	135.699	−1682.344	−440.496	117.801
4	3	[illegible]	[illegible]	5.619	1.830	77.927	152.919	−2451.019	−363.996	114.573
5	4	[illegible]	[illegible]	5.596	1.749	71.825	137.370	−1682.335	−441.138	121.853
6	5	[illegible]	[illegible]	5.586	2.009	74.876	140.247	−2066.684	−400.947	116.732
7	6	[illegible]	[illegible]	5.723	1.861	71.825	135.589	−1682.336	−438.212	122.617
8	7	[illegible]	[illegible]	5.614	1.277	71.825	133.584	−1682.342	−438.670	126.687
9	8	[illegible]	[illegible]	5.871	1.734	77.927	152.595	−2451.015	−359.039	114.619
10	9	[illegible]	[illegible]	5.888	2.129	74.876	142.065	−2066.680	−401.520	114.594
11	10	[illegible]	[illegible]	5.976	1.269	63.698	131.214	−1370.073	−479.375	111.941
12	11	[illegible]	[illegible]	5.728	1.706	74.876	139.270	−2066.688	−402.632	117.804
13	12	[illegible]	[illegible]	5.886	1.816	74.876	139.818	−2066.690	−399.686	118.920
14	13	[illegible]	[illegible]	5.891	2.270	77.927	146.285	−2451.029	−362.811	108.918
15	14	[illegible]	[illegible]	5.790	1.926	74.876	140.650	−2066.681	−402.218	113.527
16	15	[illegible]	[illegible]	5.852	2.997	102.919	183.213	−21831.257	−269.022	101.888
17	16	[illegible]	[illegible]	6.018	1.881	76.830	145.349	−2375.806	−323.533	91.919
18	17	[illegible]	[illegible]	5.833	2.177	77.927	145.483	−2451.031	−365.015	107.813
19	18	2.3010	0.597	5.989	2.011	80.978	157.941	−2835.362	−324.988	110.836
20	19	2.3424	0.737	5.911	2.135	77.927	147.355	−2451.027	−365.530	105.431
21	20	2.4771	0.387	5.923	1.798	82.678	152.621	−2490.338	−358.418	107.722
22	21	2.5740	1.286	6.009	2.817	80.911	148.479	−4564.716	−353.596	108.281
23	22	2.6020	0.391	5.924	1.802	76.383	146.250	−2375.798	−323.364	107.853
24	24	2.8129	0.494	6.020	1.919	77.927	145.051	−2451.034	−363.678	109.790
25	25	2.9030	1.110	5.973	2.580	83.894	149.738	−6678.402	−341.980	108.112
26	26	2.9030	1.250	5.920	2.751	86.878	156.431	−8792.080	−331.894	106.298
27	27	3.0000	0.543	6.115	1.989	76.235	143.952	−2475.061	−364.739	99.798
28	28	3.3424	0.912	6.105	2.370	80.978	151.349	−2835.376	−326.350	99.799
29	29	3.4771	1.694	6.031	3.441	83.962	155.358	−4949.061	−315.982	98.052
30	30	3.8750	1.713	6.054	3.473	89.309	162.623	−9295.451	−296.468	97.150
31	31	3.8750	1.618	5.974	3.308	92.912	166.505	−11290.109	−284.652	95.312
32										
33		max	1.713	6.115	3.473	102.919	183.213	1370.073	269.022	126.687
34		min	0.321	5.586	1.269	63.698	131.214	21831.257	479.375	91.919

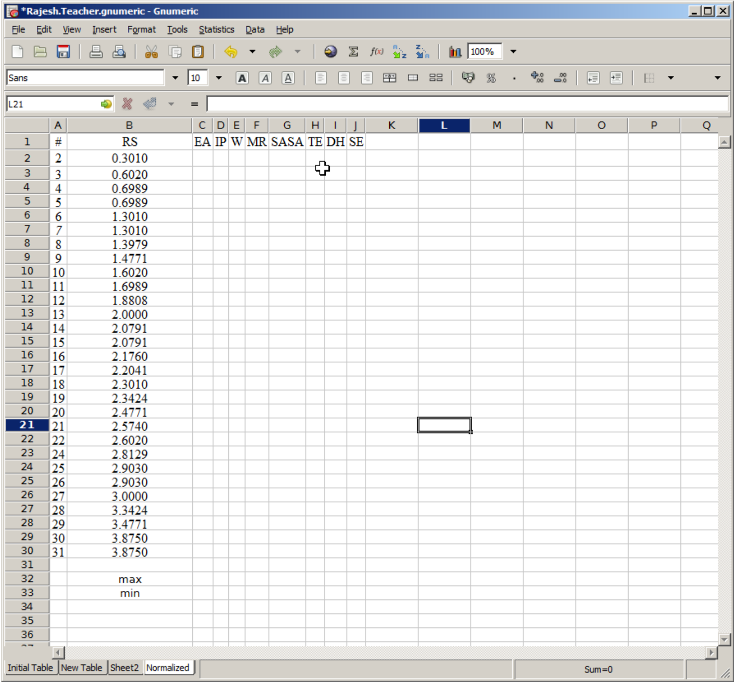
Copy the ID and labels from «New Table» and also the descriptors labels

Independent variables analysis



… paste everything in the «Normalized» sheet

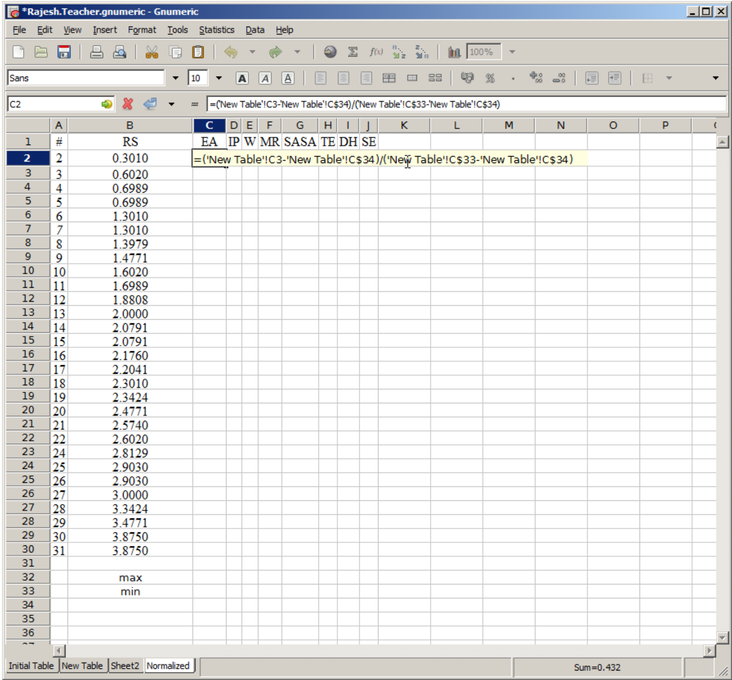
Independent variables analysis



#	RS	EA	IP	W	MR	SASA	TE	DH	SE
2	0.3010								
3	0.6020								
4	0.6989								
5	0.6989								
6	1.3010								
7	1.3010								
8	1.3979								
9	1.4771								
10	1.6020								
11	1.6989								
12	1.8808								
13	2.0000								
14	2.0791								
15	2.0791								
16	2.1760								
17	2.2041								
18	2.3010								
19	2.3424								
20	2.4771								
21	2.5740								
22	2.6020								
24	2.8129								
25	2.9030								
26	2.9030								
27	3.0000								
28	3.3424								
29	3.4771								
30	3.8750								
31	3.8750								
	max								
	min								

Make sure you have the same values for RS

Independent variables analysis



*Rajesh.Teacher.gnumeric - Gnumeric

C2 =('New Table'!C3-'New Table'!C$34)/('New Table'!C$33-'New Table'!C$34)

	A	B	C	D	E	F	G	H	I	J
1	#	RS	EA	IP	W	MR	SASA	TE	DH	SE
2	2	0.3010	=('New Table'!C3-'New Table'!C$34)/('New Table'!C$33-'New Table'!C$34)							
3	3	0.6020								
4	4	0.6989								
5	5	0.6989								
6	6	1.3010								
7	7	1.3010								
8	8	1.3979								
9	9	1.4771								
10	10	1.6020								
11	11	1.6989								
12	12	1.8808								
13	13	2.0000								
14	14	2.0791								
15	15	2.0791								
16	16	2.1760								
17	17	2.2041								
18	18	2.3010								
19	19	2.3424								
20	20	2.4771								
21	21	2.5740								
22	22	2.6020								
23	24	2.8129								
24	25	2.9030								
25	26	2.9030								
26	27	3.0000								
27	28	3.3424								
28	29	3.4771								
29	30	3.8750								
30	31	3.8750								
31										
32		max								
33		min								

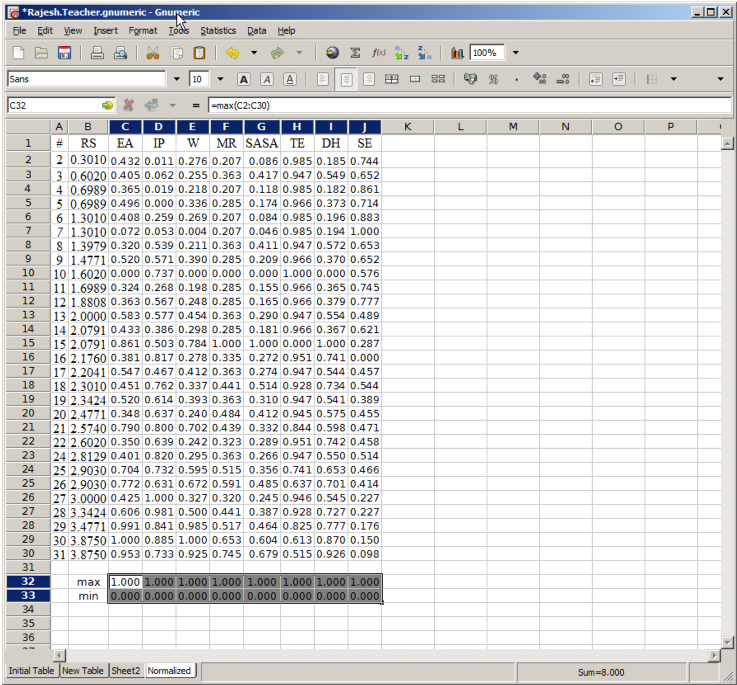
Initial Table | New Table | Sheet2 | Normalized   Sum=0.432

Now apply the normalization as described in the lesson. (see also the formula in the slide).

Here the MinMax normalization is being applied:

Normalized valued = (original value – min) / (max –min)

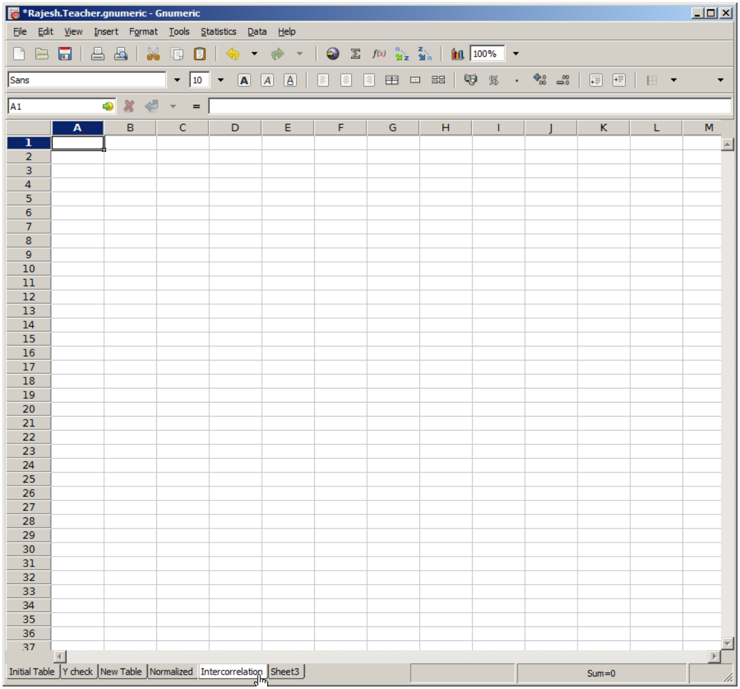
Independent variables analysis



#	RS	EA	IP	W	MR	SASA	TE	DH	SE
2	0.3010	0.432	0.011	0.276	0.207	0.086	0.985	0.185	0.744
3	0.6020	0.405	0.062	0.255	0.363	0.417	0.947	0.549	0.652
4	0.6989	0.365	0.019	0.218	0.207	0.118	0.985	0.182	0.861
5	0.6989	0.496	0.000	0.336	0.285	0.174	0.966	0.373	0.714
6	1.3010	0.408	0.259	0.269	0.207	0.084	0.985	0.196	0.883
7	1.3010	0.072	0.053	0.004	0.207	0.046	0.985	0.194	1.000
8	1.3979	0.320	0.539	0.211	0.363	0.411	0.947	0.572	0.653
9	1.4771	0.520	0.571	0.390	0.285	0.209	0.966	0.370	0.652
10	1.6020	0.000	0.737	0.000	0.000	0.000	1.000	0.000	0.576
11	1.6989	0.324	0.268	0.198	0.285	0.155	0.966	0.365	0.745
12	1.8808	0.363	0.567	0.248	0.285	0.165	0.966	0.379	0.777
13	2.0000	0.583	0.577	0.454	0.363	0.290	0.947	0.554	0.489
14	2.0791	0.433	0.386	0.298	0.285	0.181	0.966	0.367	0.621
15	2.0791	0.861	0.503	0.784	1.000	1.000	0.000	1.000	0.287
16	2.1760	0.381	0.817	0.278	0.335	0.272	0.951	0.741	0.000
17	2.2041	0.547	0.467	0.412	0.363	0.274	0.947	0.544	0.457
18	2.3010	0.451	0.762	0.337	0.441	0.514	0.928	0.734	0.544
19	2.3424	0.520	0.614	0.393	0.363	0.310	0.947	0.541	0.389
20	2.4771	0.348	0.637	0.240	0.484	0.412	0.945	0.575	0.455
21	2.5740	0.790	0.800	0.702	0.439	0.332	0.844	0.598	0.471
22	2.6020	0.350	0.639	0.242	0.323	0.289	0.951	0.742	0.458
24	2.8129	0.401	0.820	0.295	0.363	0.266	0.947	0.550	0.514
25	2.9030	0.704	0.732	0.595	0.515	0.356	0.741	0.653	0.466
26	2.9030	0.772	0.631	0.672	0.591	0.485	0.637	0.701	0.414
27	3.0000	0.425	1.000	0.327	0.320	0.245	0.946	0.545	0.227
28	3.3424	0.606	0.981	0.500	0.441	0.387	0.928	0.727	0.227
29	3.4771	0.991	0.841	0.985	0.517	0.464	0.825	0.777	0.176
30	3.8750	1.000	0.885	1.000	0.653	0.604	0.613	0.870	0.150
31	3.8750	0.953	0.733	0.925	0.745	0.679	0.515	0.926	0.098
	max	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	min	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

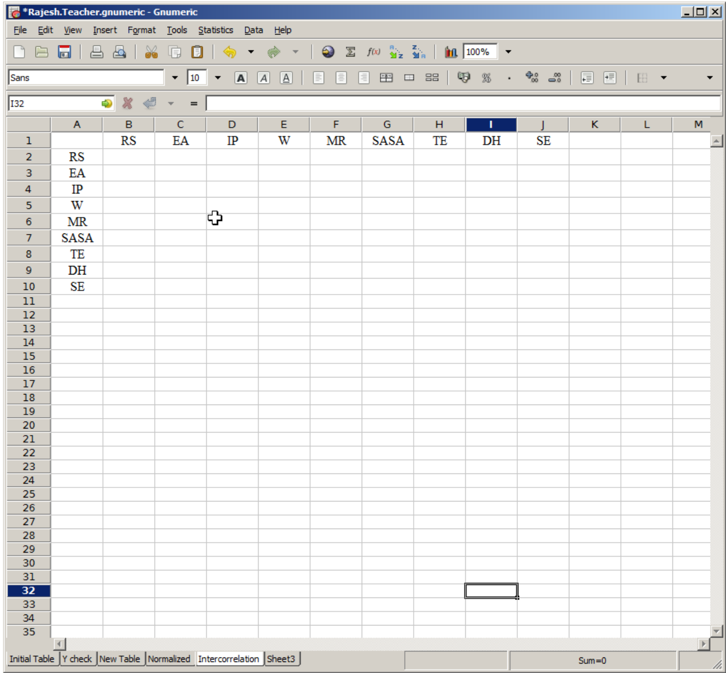
And let's obtain the normalized matrix, that can be checked by calculating min and max.
Their values should range all between 0 and 1.

Independent variables analysis



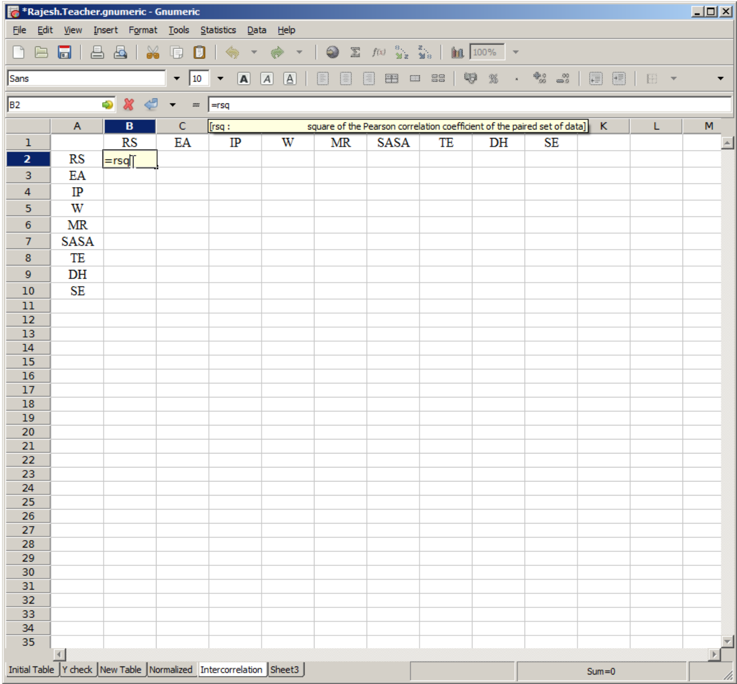
At this point we can continue analysing the data.

Independent variables analysis



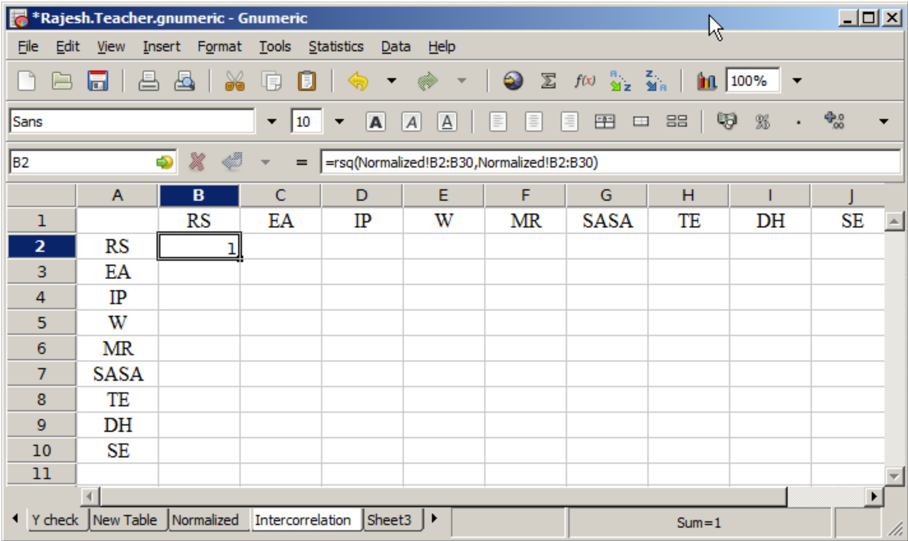
Prepare a new sheet «Intercorrelation» as in the slide

Independent variables analysis



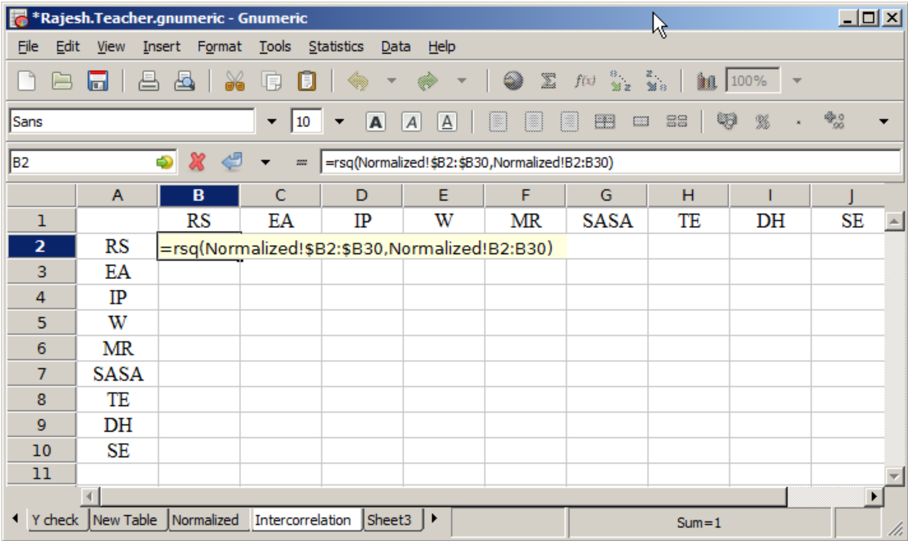
And by the proper function calculate the squared pearson coefficients (r2)

Independent variables analysis



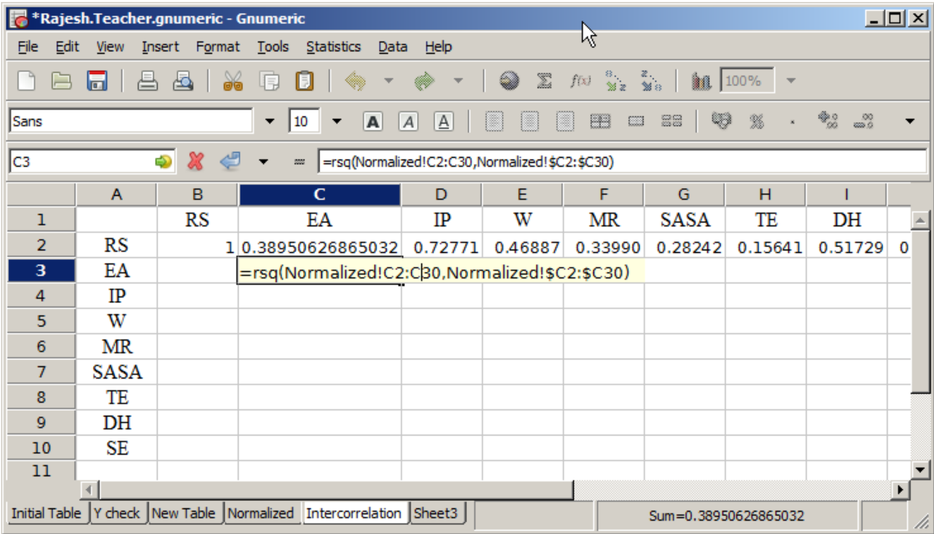
First RS vs RS. It should be = 1!

Independent variables analysis



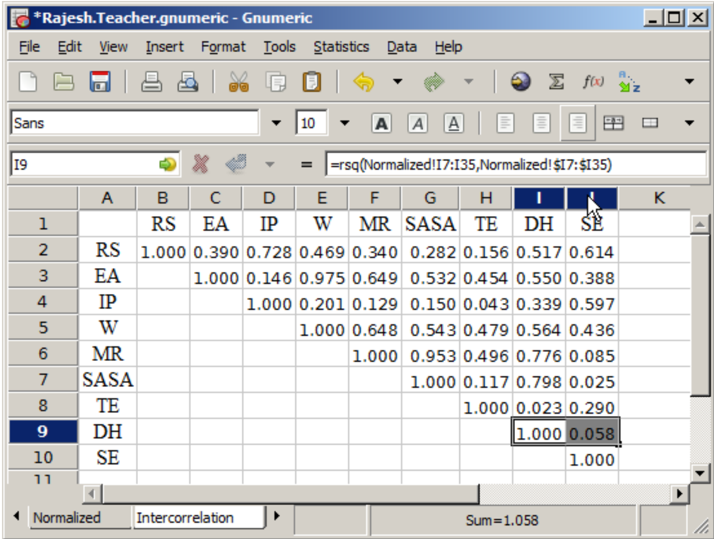
Lock at the formula and define fixed variable (insert $ signs as in the slide)

Independent variables analysis



Then just by drugging the r2 is calculated for all the values in the first row.
Make a similar operation in the others rows

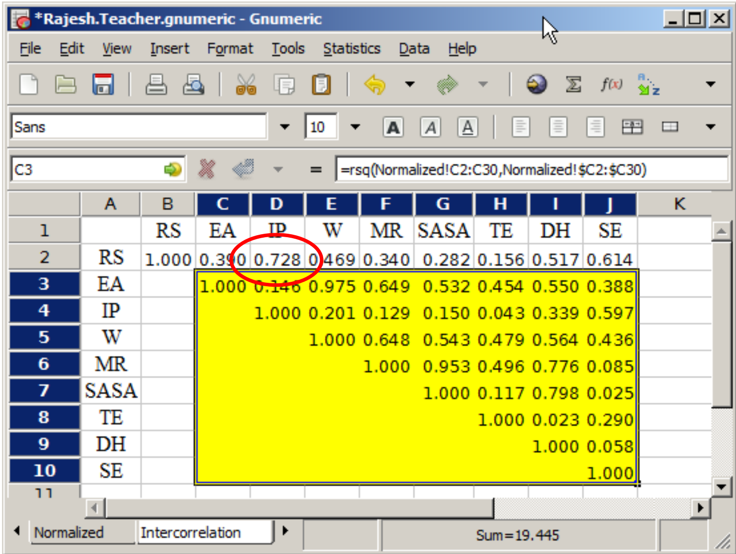
Independent variables analysis



	RS	EA	IP	W	MR	SASA	TE	DH	SE
RS	1.000	0.390	0.728	0.469	0.340	0.282	0.156	0.517	0.614
EA		1.000	0.146	0.975	0.649	0.532	0.454	0.550	0.388
IP			1.000	0.201	0.129	0.150	0.043	0.339	0.597
W				1.000	0.648	0.543	0.479	0.564	0.436
MR					1.000	0.953	0.496	0.776	0.085
SASA						1.000	0.117	0.798	0.025
TE							1.000	0.023	0.290
DH								1.000	0.058
SE									1.000

And the correlation matrix is obtained

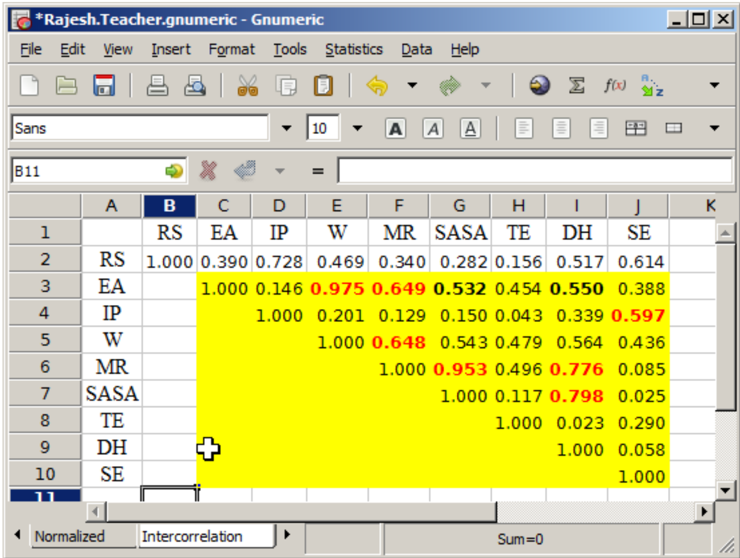
Independent variables analysis



	RS	EA	IP	W	MR	SASA	TE	DH	SE
RS	1.000	0.390	0.728	0.469	0.340	0.282	0.156	0.517	0.614
EA		1.000	0.146	0.975	0.649	0.532	0.454	0.550	0.388
IP			1.000	0.201	0.129	0.150	0.043	0.339	0.597
W				1.000	0.648	0.543	0.479	0.564	0.436
MR					1.000	0.953	0.496	0.776	0.085
SASA						1.000	0.117	0.798	0.025
TE							1.000	0.023	0.290
DH								1.000	0.058
SE									1.000

In the first row are reported the r2 values for each variable with the dependent variable, these are monoparametric regressions. The best one is IP that as the highest r^2 value.

Independent variables analysis



	RS	EA	IP	W	MR	SASA	TE	DH	SE
RS	1.000	0.390	0.728	0.469	0.340	0.282	0.156	0.517	0.614
EA		1.000	0.146	**0.975**	**0.649**	**0.532**	0.454	**0.550**	0.388
IP			1.000	0.201	0.129	0.150	0.043	0.339	**0.597**
W				1.000	**0.648**	0.543	0.479	0.564	0.436
MR					1.000	**0.953**	0.496	**0.776**	0.085
SASA						1.000	0.117	**0.798**	0.025
TE							1.000	0.023	0.290
DH								1.000	0.058
SE									1.000

In the yellow area are highlighted the intercorrelations among the descriptors to check for any collinearity.
As a limit an r2 greater than 0.5 means that two variable are correlated and should not used together.

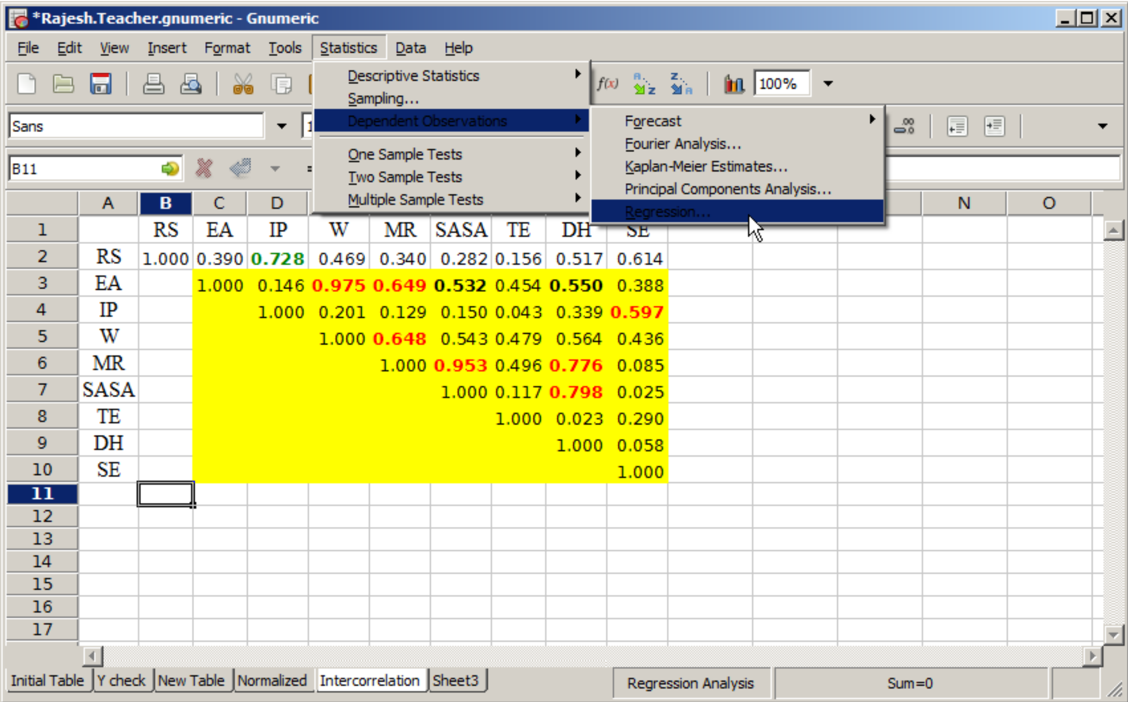
Choice of the best descriptors (stepwise method)

In order to get an overview of the uniparametric relationship between the various descriptors we use a linear regression analysis function built into Gnumeric

(Statistics > Dependent Observations > Regression)

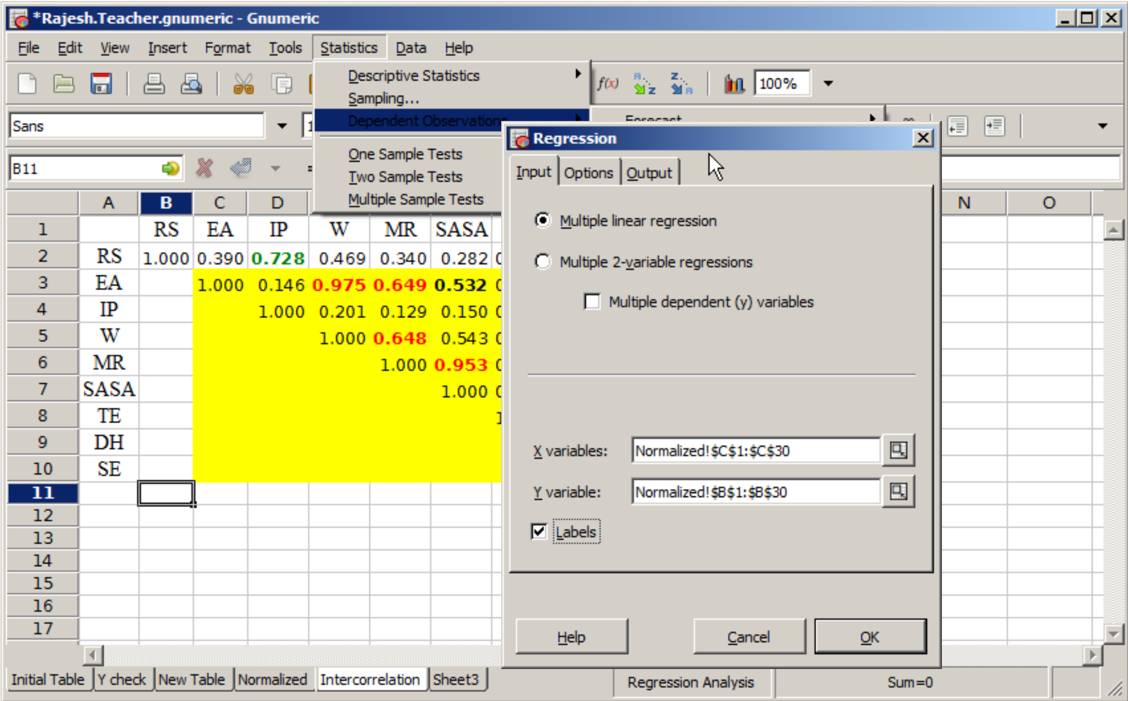
In the next slides we are going to make a stepwise building of the QSAR model

Choice of the best descriptors (stepwise method)



	RS	EA	IP	W	MR	SASA	TE	DH	SE
RS	1.000	0.390	0.728	0.469	0.340	0.282	0.156	0.517	0.614
EA		1.000	0.146	0.975	0.649	0.532	0.454	0.550	0.388
IP			1.000	0.201	0.129	0.150	0.043	0.339	0.597
W				1.000	0.648	0.543	0.479	0.564	0.436
MR					1.000	0.953	0.496	0.776	0.085
SASA						1.000	0.117	0.798	0.025
TE							1.000	0.023	0.290
DH								1.000	0.058
SE									1.000

Choice of the best descriptors (stepwise method)



We start with the uniparametric regressions using the «Statistic → Regression» menu. Then by clicking in the X and Y variable boxes we define the areas containing the values. While doing this operation select also the label of the variables. Then click on the OK button ….

Choice of the best descriptors (stepwise method)

*Rajesh.Teacher.gnumeric - Gnumeric

File Edit View Insert Format Tools Statistics Data Help

A1 = SUMMARY OUTPUT

	A	B	C	D	E	F
1	*SUMMARY OUTPUT*		*Response Variable*	*RS*		
2						
3	*Regression Statistics*					
4	*Multiple R*	0.62410437320237				
5	*R^2*	0.38950626865032				
6	*Standard Error*	0.75297888457587				
7	*Adjusted R^2*	0.36689538971144				
8	*Observations*	29				
9						
10	*ANOVA*					
11		*df*	*SS*	*MS*	*F*	*Significance of F*
12	*Regression*	1	9.76703180885487	9.76703180885487	17.2264983463603	0.00029690051514
13	*Residual*	27	15.3083844166624	0.56697720061712		
14	*Total*	28	25.0754162255172			
15						
16		*Coefficients*	*Standard Error*	*t-Statistics*	*p-Value*	*Lower 95%*
17	*Intercept*	0.90633024200569	0.32789999371423	2.76404470686136	0.01015913570582	0.23353502854898
18	*EA*	2.40923773665525	0.58047183688409	4.15048170052107	0.00029690051514	1.21820790777883
19						

Initial Table | Y check | New Table | Normalized | Intercorrelation | Sheet3 | Regression (1) — Sum=0

… and a new sheet «Regression (1)» will be created with all the statistical values

Choice of the best descriptors (stepwise method)

*Rajesh.Teacher.gnumeric - Gnumeric

	A	B	C	D	E	F	G	H	I
17	*Intercept*	0.90633024200569	0.32789999371423	2.76404470686136	0.01015913570582	0.23353502854898	1.57912545546239		
18	*EA*	2.40923773665525	0.58047183688409	4.15048170052107	0.00029690051514	1.21820790777883	3.60026756553168		
19									
20	*Constant*	*EA*	*Prediction*	*RS*	*Residual*	*Leverages*	*Internally studentized*	*Externally studentized*	*p-Value*
21	1	0.43166175024582	1.94630602016857	0.301	−1.64530602016857	0.03821898823821	−2.2281	−2.4098	2.33%
22	1	0.40462143559489	1.88115947370051	0.602	−1.27915947370051	0.04120185573742	−1.7349	−1.8015	8.32%
23	1	0.36479842674533	1.785216377993	0.6989	−1.086316377993	0.04717721504616	−1.4780	−1.5097	14.32%
24	1	0.4960668633235	2.10147324902887	0.6989	−1.40257324902887	0.0346144289285	−1.8958	−1.9928	5.69%
25	1	0.40757128810226	1.88826636967884	1.301	−0.58726636967884	0.04083421996198	−0.7964	−0.7904	43.64%
26	1	0.07177974434612	1.07926471081181	1.301	0.22173528918819	0.14910423874043	0.3192	0.3138	75.62%
27	1	0.31956735496559	1.67624397299188	1.3979	−0.27834397299188	0.05625031936562	−0.3805	−0.3743	71.12%
28	1	0.5196656833825	2.15832841685554	1.4771	−0.68122841685554	0.03452788443552	−0.9207	−0.9175	36.73%
29	1	0	0.90633024200569	1.602	0.69566975799431	0.18963444343223	1.0263	1.0234	31.55%
30	1	0.32350049164208	1.68571983429632	1.6989	0.01318016570368	0.05536482337662	0.0180	0.0177	98.60%
31	1	0.36283185840708	1.78047844734078	1.8808	0.10032155265922	0.0475211343019	0.1365	0.1340	89.44%
32	1	0.58308751229105	2.31112668038972	2	−0.31112668038972	0.03757517487862	−0.4212	−0.4146	68.18%
33	1	0.43264503441495	1.94867498549468	2.0791	0.13042501450532	0.03812689602123	0.1766	0.1734	86.37%
34	1	0.86086529008[illegible]	[illegible]	2.0[illegible]91	−0.9[illegible]01618	0.1072[illegible]3334[illegible]	−1.[illegible]668	−1.2775	21.27%
35	1	0.3805309734[illegible]	[illegible]	2.[illegible]76	0.352[illegible]78922	0.044[illegible]	0.[illegible]795	0.4724	64.06%
36	1	0.54719764011799	2.22[illegible]944598666	2.2041	−0.02055944598666	0.0352635126258	−0.0278	−0.0273	97.84%
37	1	0.45132743362832	1.9936853266908	2.301	0.3073146733092	0.03659548646297	0.4158	0.4093	68.57%
38	1	0.52015732546706	2.15951289951859	2.3424	0.18288710048141	0.03453312009996	0.2472	0.2428	81.00%
39	1	0.34808259587021	1.74494396744911	2.4771	0.73215603255089	0.05024704807097	0.9977	0.9967	32.81%
40	1	0.79006882989184	2.80979388153616	2.574	−0.23579388153616	0.08078153719068	−0.3266	−0.3211	75.07%
41	1	0.35004916420846	1.74968189810133	2.602	0.85231810189867	0.04986865367381	1.1613	1.1676	25.36%
42	1	0.40068829891839	1.87168361239606	2.8129	0.94121638760394	0.04170812517068	1.2769	1.2909	20.81%
43	1	0.70353982300885	2.60132493283836	2.903	0.30167506716164	0.05652498618567	0.4125	0.4060	68.81%
44	1	0.77236971484759	2.76715250566616	2.903	0.13584749433384	0.07509599196959	0.1876	0.1842	85.53%
45	1	0.42477876106195	1.92972326288579	3	1.07027673711421	0.03889581055506	1.4499	1.4792	15.11%
46	1	0.60619469026549	2.36679736555334	3.3424	0.97560263444666	0.03987367119381	1.3223	1.3399	19.19%
47	1	0.99065880039331	3.29306280806288	3.4771	0.18403719193712	0.17123955311578	0.2685	0.2638	79.41%
48	1	1	3.31556797866094	3.875	0.55943202133906	0.17661746218617	0.8188	0.8118	42.43%
49	1	0.95329400196657	3.20304212567065	3.875	0.67195787432935	0.15076504400564	0.9684	0.9646	34.36%
50									

New Table | Normalized | Intercorrelation | Regressions | Regression (1) | Regression (2) | Regression (3) | Regression (4) | Regression (5) — Sum=0

logRS = 0.90 + 2.40 * EA

By analysing the new sheet we can observe the intercept and constant values of the line and thus we can write its equation

Choice of the best descriptors (stepwise method)

*Rajesh.Teacher.gnumeric - Gnumeric

	A	B	C	D	E	F
1	*SUMMARY OUTPUT*		*Response Variable*	*RS*		
2						
3	*Regression Statistics*					
4	*Multiple R*	0.85305842334863				
5	*R^2*	0.72770867364606				
6	*Standard Error*	0.50287390153896				
7	*Adjusted R^2*	0.71762380970702				
8	*Observations*	29				
9						
10	*ANOVA*					
11		*df*	*SS*	*MS*	*F*	*Significance of F*
12	*Regression*	1	18.2475978825939	18.2475978825939	72.1585018940464	4.1569053282004
13	*Residual*	27	6.8278183429233	0.25288216084901		
14	*Total*	28	25.0754162255172			
15						
16		*Coefficients*	*Standard Error*	*t-Statistics*	*p-Value*	*Lower 95%*
17	*Intercept*	0.66259068100773	0.19713001110905	3.36118624089754	0.00232964121029	0.2581133
18	*IP*	2.6878988550149	0.31642381929091	8.49461605336265	4.15690532820046E−009	2.0386508
19						

Regression (2) | Regression (3) | Regression (4) | Regression (5) | Regression (6) — Sum=0

Reapeat for all the descriptors

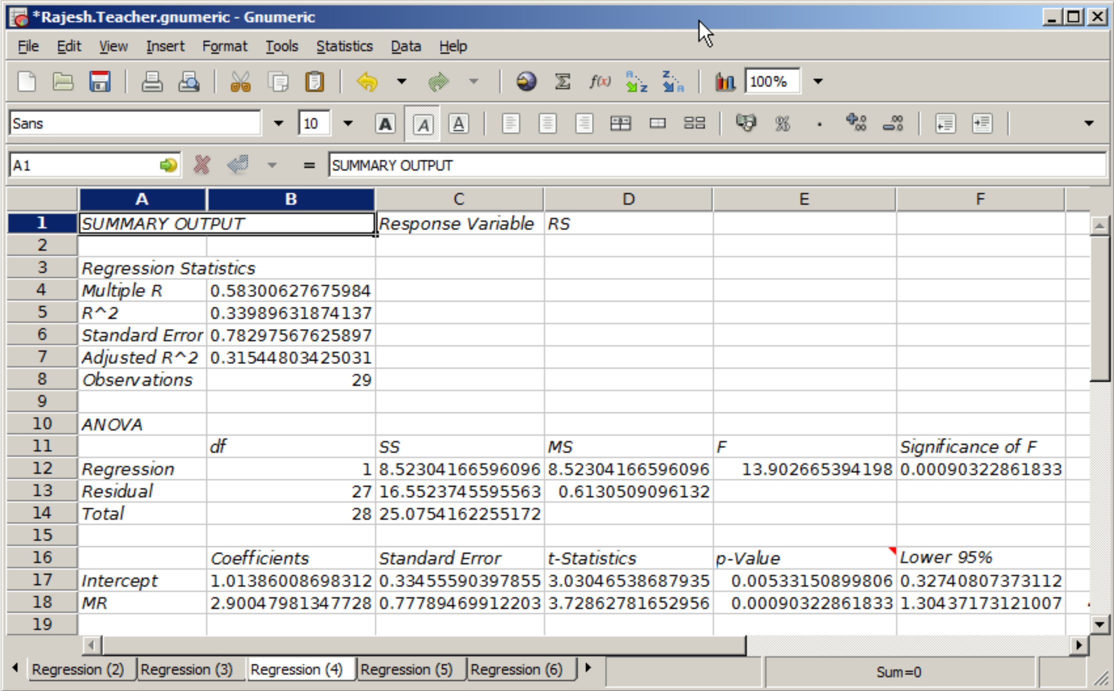
Choice of the best descriptors (stepwise method)

*Rajesh.Teacher.gnumeric - Gnumeric

	A	B	C	D	E	F
1	*SUMMARY OUTPUT*		*Response Variable*	*RS*		
2						
3	*Regression Statistics*					
4	*Multiple R*	0.68473965334663				
5	*R^2*	0.46886839286526				
6	*Standard Error*	0.70233331011526				
7	*Adjusted R^2*	0.44919685186027				
8	*Observations*	29				
9						
10	*ANOVA*					
11		*df*	*SS*	*MS*	*F*	*Significance of F*
12	*Regression*	1	11.7570701060858	11.7570701060858	23.8348583238253	4.181154129624
13	*Residual*	27	13.3183461194314	0.49327207849746		
14	*Total*	28	25.0754162255172			
15						
16		*Coefficients*	*Standard Error*	*t-Statistics*	*p-Value*	*Lower 95%*
17	*Intercept*	1.13322035863428	0.24353781773917	4.65315969878639	7.73582994698183E−005	0.6335220
18	*W*	2.45877550982804	0.50363115884698	4.88209568974486	4.181154129625E−005	1.4254097
19						

Regression (2) | Regression (3) | Regression (4) | Regression (5) | Regression (6) Sum=0

Choice of the best descriptors (stepwise method)



	A	B	C	D	E	F
1	SUMMARY OUTPUT		Response Variable	RS		
2						
3	Regression Statistics					
4	Multiple R	0.58300627675984				
5	R^2	0.33989631874137				
6	Standard Error	0.78297567625897				
7	Adjusted R^2	0.31544803425031				
8	Observations	29				
9						
10	ANOVA					
11		df	SS	MS	F	Significance of F
12	Regression	1	8.52304166596096	8.52304166596096	13.902665394198	0.00090322861833
13	Residual	27	16.5523745595563	0.6130509096132		
14	Total	28	25.0754162255172			
15						
16		Coefficients	Standard Error	t-Statistics	p-Value	Lower 95%
17	Intercept	1.01386008698312	0.33455590397855	3.03046538687935	0.00533150899806	0.32740807373112
18	MR	2.90047981347728	0.77789469912203	3.72862781652956	0.00090322861833	1.30437173121007
19						

Choice of the best descriptors (stepwise method)

*Rajesh.Teacher.gnumeric - Gnumeric

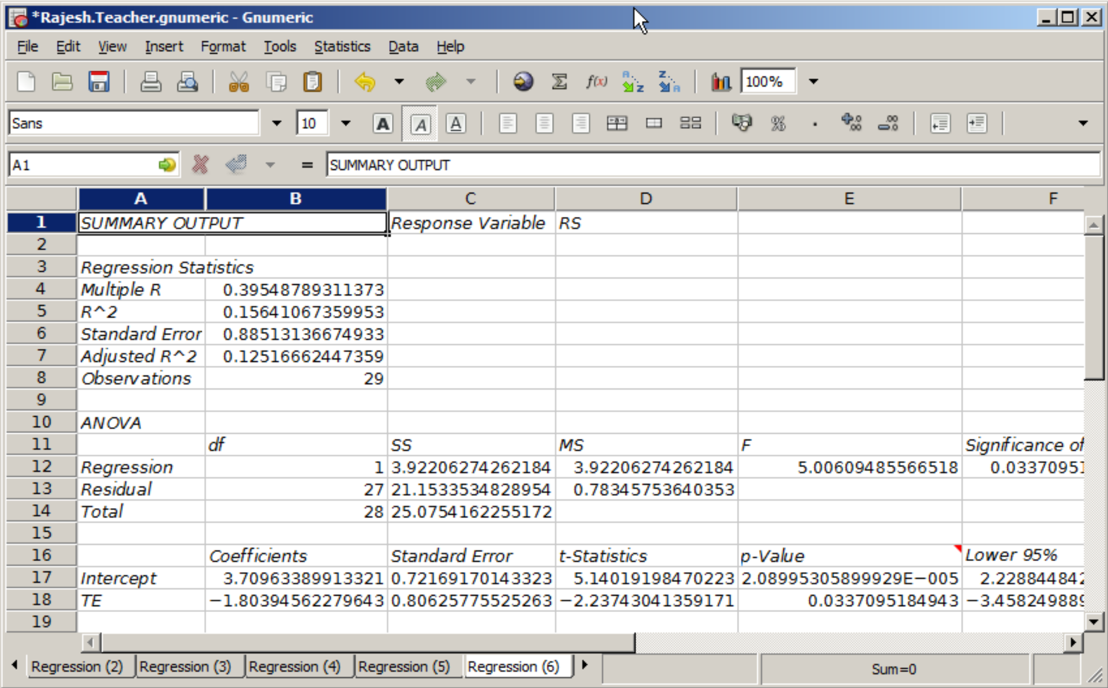
A1 = SUMMARY OUTPUT

	A	B	C	D	E	F
1	*SUMMARY OUTPUT*		*Response Variable*	*RS*		
2						
3	*Regression Statistics*					
4	*Multiple R*	0.5314367259195				
5	*R^2*	0.28242499365604				
6	*Standard Error*	0.81634896217295				
7	*Adjusted R^2*	0.25584814156922				
8	*Observations*	29				
9						
10	*ANOVA*					
11		*df*	*SS*	*MS*	*F*	*Significance of F*
12	*Regression*	1	7.08192426841418	7.08192426841418	10.6267285807022	0.0030108119412
13	*Residual*	27	17.9934919571031	0.66642562804085		
14	*Total*	28	25.0754162255172			
15						
16		*Coefficients*	*Standard Error*	*t-Statistics*	*p-Value*	*Lower 95%*
17	*Intercept*	1.37607010059851	0.27841463014938	4.94252080021867	3.55431958487801E−005	0.80481046622
18	*SASA*	2.39260481334891	0.73395794823225	3.2598663439936	0.00301081194124	0.88664749735
19						

Regression (2) | Regression (3) | Regression (4) | Regression (5) | Regression (6)

Sum=0

Choice of the best descriptors (stepwise method)



	A	B	C	D	E	F
1	SUMMARY OUTPUT		Response Variable	RS		
2						
3	Regression Statistics					
4	Multiple R	0.39548789311373				
5	R^2	0.15641067359953				
6	Standard Error	0.88513136674933				
7	Adjusted R^2	0.12516662447359				
8	Observations	29				
9						
10	ANOVA					
11		df	SS	MS	F	Significance of
12	Regression	1	3.92206274262184	3.92206274262184	5.00609485566518	0.03370951
13	Residual	27	21.1533534828954	0.78345753640353		
14	Total	28	25.0754162255172			
15						
16		Coefficients	Standard Error	t-Statistics	p-Value	Lower 95%
17	Intercept	3.70963389913321	0.72169170143323	5.14019198470223	2.08995305899929E−005	2.22884484
18	TE	−1.80394562279643	0.80625775525263	−2.23743041359171	0.0337095184943	−3.45824988
19						

Choice of the best descriptors (stepwise method)

*Rajesh.Teacher.gnumeric - Gnumeric

	A	B	C	D	E	F
1	SUMMARY OUTPUT		Response Variable	RS		
2						
3	Regression Statistics					
4	Multiple R	0.719230883066				
5	R^2	0.51729306315589				
6	Standard Error	0.66955146280539				
7	Adjusted R^2	0.49941502845796				
8	Observations	29				
9						
10	ANOVA					
11		df	SS	MS	F	Significance of F
12	Regression	1	12.9713388692068	12.9713388692068	28.9345597486613	1.1016391606563
13	Residual	27	12.1040773563104	0.44829916134483		
14	Total	28	25.0754162255172			
15						
16		Coefficients	Standard Error	t-Statistics	p-Value	Lower 95%
17	Intercept	0.63847706491587	0.30512596126205	2.09250324775717	0.0459264945109	0.0124103
18	DH	2.80306342879905	0.5211040948343	5.37908540075925	1.10163916065638E-005	1.7338461
19						

Regression (3) | Regression (4) | Regression (5) | Regression (6) | Regression (7) — Sum=0

Choice of the best descriptors (stepwise method)

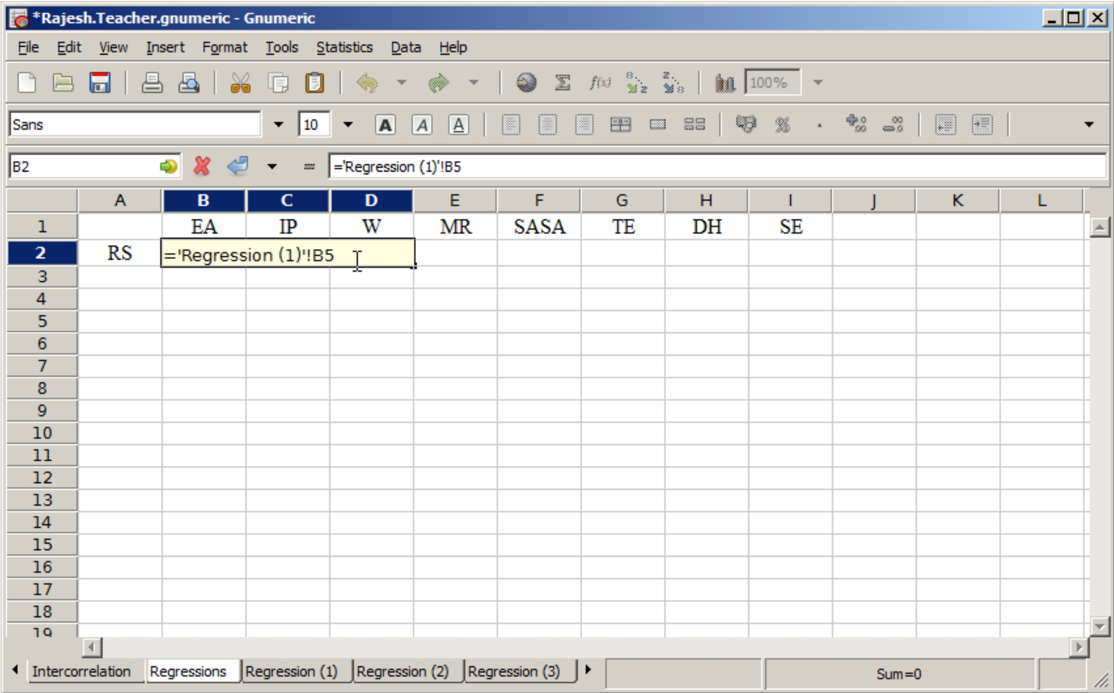
*Rajesh.Teacher.gnumeric - Gnumeric

	A	B	C	D	E	F
1	*SUMMARY OUTPUT*		*Response Variable*	*RS*		
2						
3	*Regression Statistics*					
4	*Multiple R*	0.78335040208227				
5	*R^2*	0.61363785244245				
6	*Standard Error*	0.59901745673537				
7	*Adjusted R^2*	0.59932814327365				
8	*Observations*	29				
9						
10	*ANOVA*					
11		*df*	*SS*	*MS*	*F*	*Significance of*
12	*Regression*	1	15.387224561727	15.387224561727	42.8826222255078	5.0549005851
13	*Residual*	27	9.68819166379022	0.35882191347371		
14	*Total*	28	25.0754162255172			
15						
16		*Coefficients*	*Standard Error*	*t-Statistics*	*p-Value*	*Lower 95%*
17	*Intercept*	3.65275893626293	0.2567617455734	14.2262583863715	4.60469169074284E−014	3.12592
18	*SE*	−2.99020144187781	0.45662509913122	−6.54848243683281	5.05490058511252E−007	−3.92711
19						

Regression (4) | Regression (5) | Regression (6) | Regression (7) | Regression (8)

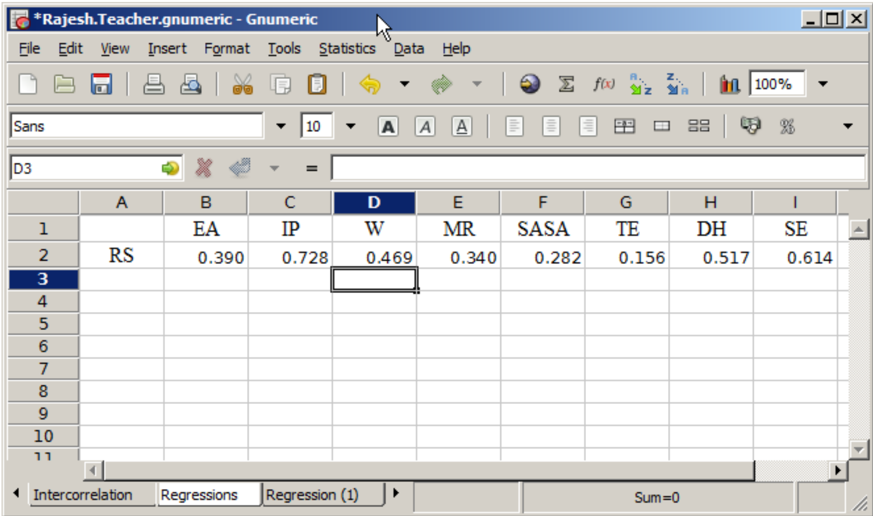
Sum=0

Choice of the best descriptors (stepwise method)



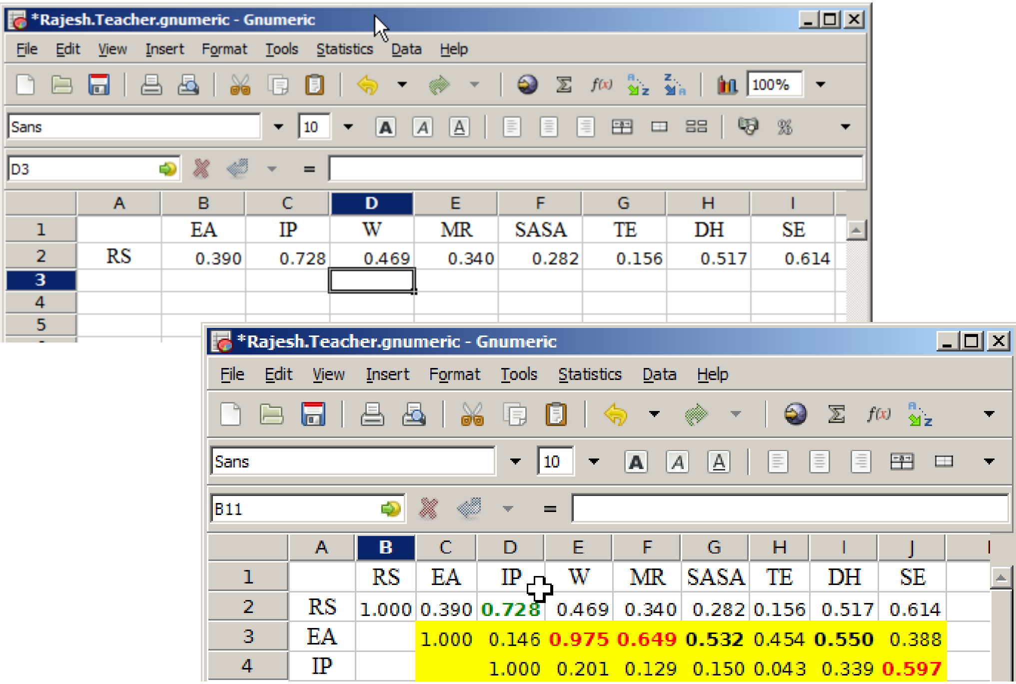
And record the r^2 values in a new sheet «Regressions»

Choice of the best descriptors (stepwise method)



	EA	IP	W	MR	SASA	TE	DH	SE
RS	0.390	0.728	0.469	0.340	0.282	0.156	0.517	0.614

Choice of the best descriptors (stepwise method)

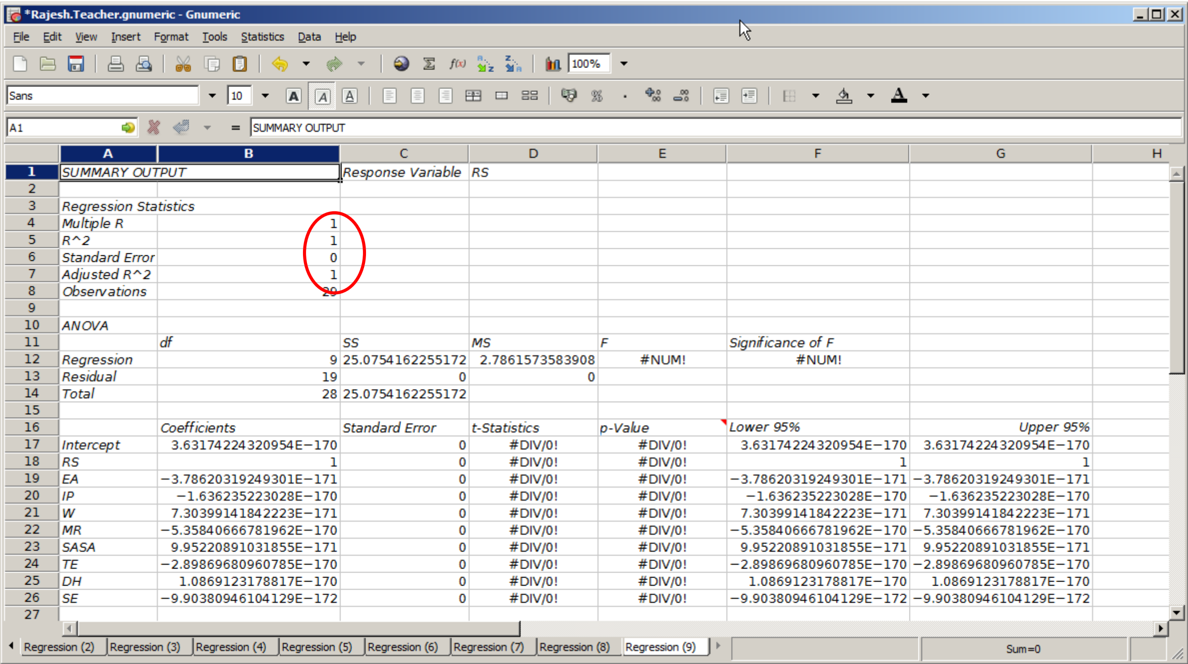


	EA	IP	W	MR	SASA	TE	DH	SE
RS	0.390	0.728	0.469	0.340	0.282	0.156	0.517	0.614

	RS	EA	IP	W	MR	SASA	TE	DH	SE
RS	1.000	0.390	**0.728**	0.469	0.340	0.282	0.156	0.517	0.614
EA		1.000	0.146	**0.975**	**0.649**	**0.532**	0.454	**0.550**	0.388
IP			1.000	0.201	0.129	0.150	0.043	0.339	**0.597**

Compare the Regressions with the initial correlations
and the they should be the same values.

MLR with all variables → Overfitting!



	A	B	C	D	E	F	G
1	SUMMARY OUTPUT		Response Variable	RS			
2							
3	Regression Statistics						
4	Multiple R	1					
5	R^2	1					
6	Standard Error	0					
7	Adjusted R^2	1					
8	Observations	29					
9							
10	ANOVA						
11		df	SS	MS	F	Significance of F	
12	Regression	9	25.0754162255172	2.7861573583908	#NUM!	#NUM!	
13	Residual	19	0	0			
14	Total	28	25.0754162255172				
15							
16		Coefficients	Standard Error	t-Statistics	p-Value	Lower 95%	Upper 95%
17	Intercept	3.63174224320954E−170	0	#DIV/0!	#DIV/0!	3.63174224320954E−170	3.63174224320954E−170
18	RS	1	0	#DIV/0!	#DIV/0!	1	1
19	EA	−3.78620319249301E−171	0	#DIV/0!	#DIV/0!	−3.78620319249301E−171	−3.78620319249301E−171
20	IP	−1.636235223028E−170	0	#DIV/0!	#DIV/0!	−1.636235223028E−170	−1.636235223028E−170
21	W	7.30399141842223E−171	0	#DIV/0!	#DIV/0!	7.30399141842223E−171	7.30399141842223E−171
22	MR	−5.35840666781962E−170	0	#DIV/0!	#DIV/0!	−5.35840666781962E−170	−5.35840666781962E−170
23	SASA	9.95220891031855E−171	0	#DIV/0!	#DIV/0!	9.95220891031855E−171	9.95220891031855E−171
24	TE	−2.89869680960785E−170	0	#DIV/0!	#DIV/0!	−2.89869680960785E−170	−2.89869680960785E−170
25	DH	1.0869123178817E−170	0	#DIV/0!	#DIV/0!	1.0869123178817E−170	1.0869123178817E−170
26	SE	−9.90380946104129E−172	0	#DIV/0!	#DIV/0!	−9.90380946104129E−172	−9.90380946104129E−172
27							

Just to try, instead of highlighting one parameter, select all the 8 descriptors while doung the regressions,

in the new sheet «Regression (9)» the values are clearlyindicating that 8 parameters give a perfect regression of r2 = 1!

This is overfitting!

We are using collinear parameters and the model is LYING!

No differences with unscaled data!

Rajesh.Teacher.NO-Normalized.gnumeric - Gnumeric

F21 = {=mmult(mmult(mmult('Regression (2)'!$A$21:$B$49,minverse(mmult(transpose('Regression (2)'!$A$21:$B$49),'Regression (2)'!$A$21:$B$49))),transpose('Regression (

	A	B	C	D	E	F	G
1	*SUMMARY OUTPUT*		*Response Variable*	*RS*			
2							
3	*Regression Statistics*						
4	*Multiple R*	0.85305842334863					
5	*R^2*	0.72770867364606					
6	*Standard Error*	0.50287390153896					
7	*Adjusted R^2*	0.71762380970702					
8	*Observations*	29					
9							
10	*ANOVA*						
11		*df*	*SS*	*MS*	*F*	*Significance of F*	
12	*Regression*	1	18.2475978825939	18.2475978825939	72.1585018940464	4.1569053282004E−009	
13	*Residual*	27	6.8278183429233	0.25288216084901			
14	*Total*	28	25.0754162255172				
15							
16		*Coefficients*	*Standard Error*	*t-Statistics*	*p-Value*	*Lower 95%*	*Upper 95%*
17	*Intercept*	−27.72040176533	3.51614143537824	−7.883756178411	1.78162061194726E−008	−34.9349280627027	−20.5058754679595
18	*IP*	5.081094243884 5	0.59815466784671	8.49461605336265	4.1569053282004 6E−009	3.85378224282119	6.30840624494782
19							

NO-Normalized | Intercorrelation | Regressions | Regression (1) | Regression (2) — Sum=0.14879254218158

IF we repeated with unscaled data (without normalizing) we can observe that the results are the same, the only differences are the intercept and the variable coefficient values.

Although it seems that scaling is not necessary, apply always a normalization in you matrix!

Step Forward Multi Linear Regressions

	EA	IP	W	MR	SASA	TE	DH	SE
EA	1.00	0.15	0.98	0.65	0.53	0.45	0.55	0.39
IP		1.00	0.20	0.13	0.15	0.04	0.34	0.60
W			1.00	0.65	0.54	0.48	0.56	0.44
MR				1.00	0.95	0.50	0.78	0.08
SASA					1.00	0.12	0.80	0.02
TE						1.00	0.02	0.29
DH							1.00	0.06
SE								1.00

	EA	IP	W	MR	SASA	TE	DH	SE
RS	0.39	0.73	0.47	0.34	0.28	0.16	0.52	0.61

Going back to the intercorrelation data it is possible to see that the first parameter to use is «IP» and that we can associate with it only «W», «MR», «SASA», «TE» and «DH» values.

The ones that display low correlation with «IP».

At first glance we could expect a godd correlation by using atriparametric model including «IP», «W» and «DH». But, this can be confirmed by a stepwise regression.

Step Forward Multi Linear Regressions

IP	W	IP	MR	IP	SASA	IP	TE	IP	DH
5.59	1.88	5.59	71.83	5.59	135.70	5.59	−1682.34	5.59	−440.50
5.62	1.83	5.62	77.93	5.62	152.92	5.62	−2451.02	5.62	−364.00
5.60	1.75	5.60	71.83	5.60	137.37	5.60	−1682.34	5.60	−441.14
5.59	2.01	5.59	74.88	5.59	140.25	5.59	−2066.68	5.59	−400.95
5.72	1.86	5.72	71.83	5.72	135.59	5.72	−1682.34	5.72	−438.21
5.61	1.28	5.61	71.83	5.61	133.58	5.61	−1682.34	5.61	−438.67
5.87	1.73	5.87	77.93	5.87	152.60	5.87	−2451.01	5.87	−359.04
5.89	2.13	5.89	74.88	5.89	142.07	5.89	−2066.68	5.89	−401.52
5.98	1.27	5.98	63.70	5.98	131.21	5.98	−1370.07	5.98	−479.38
5.73	1.71	5.73	74.88	5.73	139.27	5.73	−2066.69	5.73	−402.63
5.89	1.82	5.89	74.88	5.89	139.82	5.89	−2066.69	5.89	−399.69
5.89	2.27	5.89	77.93	5.89	146.29	5.89	−2451.03	5.89	−362.81
5.79	1.93	5.79	74.88	5.79	140.65	5.79	−2066.68	5.79	−402.22
5.85	3.00	5.85	102.92	5.85	183.21	5.85	−21831.26	5.85	−269.02
6.02	1.88	6.02	76.83	6.02	145.35	6.02	−2375.81	6.02	−323.53
5.83	2.18	5.83	77.93	5.83	145.48	5.83	−2451.03	5.83	−365.01
5.99	2.01	5.99	80.98	5.99	157.94	5.99	−2835.36	5.99	−324.99
5.91	2.13	5.91	77.93	5.91	147.35	5.91	−2451.03	5.91	−365.53
5.92	1.80	5.92	82.68	5.92	152.62	5.92	−2490.34	5.92	−358.42
6.01	2.82	6.01	80.91	6.01	148.48	6.01	−4564.72	6.01	−353.60
5.92	1.80	5.92	76.38	5.92	146.25	5.92	−2375.80	5.92	−323.36
6.02	1.92	6.02	77.93	6.02	145.05	6.02	−2451.03	6.02	−363.68
5.97	2.58	5.97	83.89	5.97	149.74	5.97	−6678.40	5.97	−341.98
5.92	2.75	5.92	86.88	5.92	156.43	5.92	−8792.08	5.92	−331.89
6.12	1.99	6.12	76.23	6.12	143.95	6.12	−2475.06	6.12	−364.74
6.11	2.37	6.11	80.98	6.11	151.35	6.11	−2835.38	6.11	−326.35
6.03	3.44	6.03	83.96	6.03	155.36	6.03	−4949.06	6.03	−315.98
6.05	3.47	6.05	89.31	6.05	162.62	6.05	−9295.45	6.05	−296.47
5.97	3.31	5.97	92.91	5.97	166.51	5.97	−11290.11	5.97	−284.65

Let’s prepare the biparametric combinations

Step Forward Multi Linear Regressions

Rajesh.Teacher.NO-Normalized.gnumeric - Gnumeric

	A	B	C	D	E	F	G
1	SUMMARY OUTPUT		Response Variable	RS			
2							
3	Regression Statistics						
4	Multiple R	0.91769207990091					
5	R^2	0.84215875351285					
6	Standard Error	0.39016441077004					
7	Adjusted R^2	0.83001711916768					
8	Observations	29					
9							
10	ANOVA						
11		df	SS	MS	F	Significance of F	
12	Regression	2	21.1174812722975	10.5587406361487	69.3612350340779	3.77456166633094E−011	
13	Residual	26	3.95793495321975	0.15222826743153			
14	Total	28	25.0754162255172				
15							
16		Coefficients	Standard Error	t-Statistics	p-Value	Lower 95%	Upper 95%
17	Intercept	−23.12151769879	2.92646506772578	−7.900835022356	2.23320906535491E−008	−29.1369527966614	−17.10608260092
18	IP	4.07088235004434	0.51914507766559	7.84151198803542	2.56892074143486E−008	3.00376435997592	5.13800034011275
19	W	0.61656170789146	0.14200122009646	4.34194655139328	0.00019078662105	0.32467401965991	0.908449396123
20							

Regression (4) | Regression (5) | Regression (6) | Regression (7) | Regression (8) | Regression (9) | Sheet1 | Sheet2 | Regression (10)

Sum=0

And by using the «Statistic → Regression» menu the five biparametric are easily build.

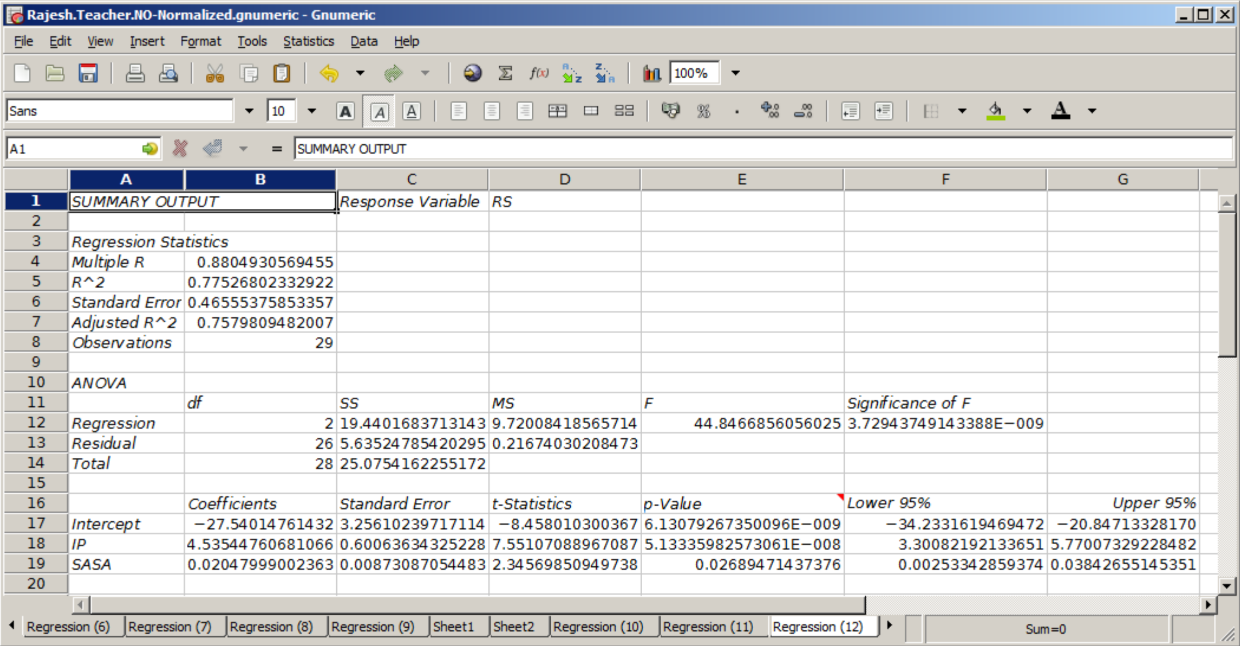
Step Forward Multi Linear Regressions

Rajesh.Teacher.NO-Normalized.gnumeric - Gnumeric

	A	B	C	D	E	F	G
1	*SUMMARY OUTPUT*		*Response Variable*	*RS*			
2							
3	*Regression Statistics*						
4	*Multiple R*	0.90317927777274					
5	*R^2*	0.81573280779809					
6	*Standard Error*	0.42156195553287					
7	*Adjusted R^2*	0.80155840839794					
8	*Observations*	29					
9							
10	*ANOVA*						
11		*df*	*SS*	*MS*	*F*	*Significance of F*	
12	*Regression*	2	20.454839684347	10.2274198421735	57.5497264307116	2.82373695431237E−010	
13	*Residual*	26	4.62057654117024	0.1777144823527			
14	*Total*	28	25.0754162255172				
15							
16		*Coefficients*	*Standard Error*	*t-Statistics*	*p-Value*	*Lower 95%*	*Upper 95%*
17	*Intercept*	−26.90950309644	2.956567694227	−9.101602222397	1.44960940692087E−009	−32.9868150292698	−20.83219116362
18	*IP*	4.40181436773323	0.5372051748974	8.19391653956787	1.12563128581626E−008	3.29757331614008	5.50605541932638
19	*MR*	0.04031839293135	0.01144035937905	3.52422433557411	0.00159489291772	0.01680239743906	0.06383438842364
20							

Regression (5) | Regression (6) | Regression (7) | Regression (8) | Regression (9) | Sheet1 | Sheet2 | Regression (10) | Regression (11) — Sum=0

Step Forward Multi Linear Regressions



	A	B	C	D	E	F	G
1	SUMMARY OUTPUT		Response Variable	RS			
2							
3	Regression Statistics						
4	Multiple R	0.8804930569455					
5	R^2	0.77526802332922					
6	Standard Error	0.46555375853357					
7	Adjusted R^2	0.7579809482007					
8	Observations	29					
9							
10	ANOVA						
11		df	SS	MS	F	Significance of F	
12	Regression	2	19.4401683713143	9.72008418565714	44.8466856056025	3.72943749143388E−009	
13	Residual	26	5.63524785420295	0.21674030208473			
14	Total	28	25.0754162255172				
15							
16		Coefficients	Standard Error	t-Statistics	p-Value	Lower 95%	Upper 95%
17	Intercept	−27.54014761432	3.25610239717114	−8.458010300367	6.13079267350096E−009	−34.2331619469472	−20.84713328170
18	IP	4.53544760681066	0.60063634325228	7.55107088967087	5.13335982573061E−008	3.30082192133651	5.77007329228482
19	SASA	0.02047999002363	0.00873087054483	2.34569850949738	0.02689471437376	0.00253342859374	0.03842655145351
20							

Step Forward Multi Linear Regressions

Rajesh.Teacher.NO-Normalized.gnumeric - Gnumeric

	A	B	C	D	E	F	G
1	*SUMMARY OUTPUT*		*Response Variable*	*RS*			
2							
3	*Regression Statistics*						
4	*Multiple R*	0.88170499144918					
5	*R^2*	0.77740369194639					
6	*Standard Error*	0.46333635708462					
7	*Adjusted R^2*	0.76028089901919					
8	*Observations*	29					
9							
10	*ANOVA*						
11		*df*	*SS*	*MS*	*F*	*Significance of F*	
12	*Regression*	2	19.4937211508096	9.7468605754048	45.4016874029613	3.29407467335866E−009	
13	*Residual*	26	5.58169507470764	0.21468057979645			
14	*Total*	28	25.0754162255172				
15							
16		*Coefficients*	*Standard Error*	*t-Statistics*	*p-Value*	*Lower 95%*	*Upper 95*
17	*Intercept*	−26.2642533135557	3.29558699475141	−7.96952207767067	1.9000194819616E−008	−33.0384293988774	−19.49007722823
18	*IP*	4.79872863573448	0.56344968096452	8.51669421042164	5.3632217820068E−009	3.6405412293177	5.9569160421512
19	*TE*	−5.08066165E−005	2.10880368E−005	−2.40926251133444	0.02336381837501	−9.415369698587E−005	−7.45953606E−00
20							

Regression (7) | Regression (8) | Regression (9) | Sheet1 | Sheet2 | Regression (10) | Regression (11) | Regression (12) | Regression (13)

Sum=0

Step Forward Multi Linear Regressions

Rajesh.Teacher.NO-Normalized.gnumeric - Gnumeric

	A	B	C	D	E	F	G
1	*SUMMARY OUTPUT*		*Response Variable*	*RS*			
2							
3	*Regression Statistics*						
4	*Multiple R*	0.89582250409996					
5	*R^2*	0.80249795885192					
6	*Standard Error*	0.43643864169934					
7	*Adjusted R^2*	0.78730549414822					
8	*Observations*	29					
9							
10	*ANOVA*						
11		*df*	*SS*	*MS*	*F*	*Significance of F*	
12	*Regression*	2	20.1229703383399	10.0614851691699	52.8221045434813	6.95713848758163E−010	
13	*Residual*	26	4.95244588717739	0.19047868796836			
14	*Total*	28	25.0754162255172				
15							
16		*Coefficients*	*Standard Error*	*t-Statistics*	*p-Value*	*Lower 95%*	*Upper 95%*
17	*Intercept*	−18.57278677660	4.22036925011418	−4.400749241574	0.00016342050294	−27.2478800121582	−9.897693541049
18	*IP*	3.9136043985866	0.63870094037851	6.12744424059761	1.77378147024894E−006	2.60073581314938	5.22647298402382
19	*DH*	0.00623383839961	0.00198671224603	3.13776613199662	0.00420172888299	0.00215009289178	0.01031758390743
20							

Biparametric models

1		
2		R^2
3	IP + W	0.84215875351285
4	IP + MR	0.81573280779809
5	IP + SASA	0.77526802332922
6	IP + TE	0.77740369194639
7	IP + DH	0.80249795885192
8		
9		

The best combination resulted to be IP + W, but also IP + MR should be analyzed with other descriptor

Step Forward Multi Linear Regressions

AC	AD	AE	AF	AG	AH	AI	AJ	AK	AL	AM	AN	AO	AP	AQ	AR
IP	W	MR		IP	W	SASA		IP	W	TE		IP	W	DH	
5.59	1.88	71.83		5.59	1.88	135.70		5.59	1.88	−1682.34		5.59	1.88	−440.50	
5.62	1.83	77.93		5.62	1.83	152.92		5.62	1.83	−2451.02		5.62	1.83	−364.00	
5.60	1.75	71.83		5.60	1.75	137.37		5.60	1.75	−1682.34		5.60	1.75	−441.14	
5.59	2.01	74.88		5.59	2.01	140.25		5.59	2.01	−2066.68		5.59	2.01	−400.95	
5.72	1.86	71.83		5.72	1.86	135.59		5.72	1.86	−1682.34		5.72	1.86	−438.21	
5.61	1.28	71.83		5.61	1.28	133.58		5.61	1.28	−1682.34		5.61	1.28	−438.67	
5.87	1.73	77.93		5.87	1.73	152.60		5.87	1.73	−2451.01		5.87	1.73	−359.04	
5.89	2.13	74.88		5.89	2.13	142.07		5.89	2.13	−2066.68		5.89	2.13	−401.52	
5.98	1.27	63.70		5.98	1.27	131.21		5.98	1.27	−1370.07		5.98	1.27	−479.38	
5.73	1.71	74.88		5.73	1.71	139.27		5.73	1.71	−2066.69		5.73	1.71	−402.63	
5.89	1.82	74.88		5.89	1.82	139.82		5.89	1.82	−2066.69		5.89	1.82	−399.69	
5.89	2.27	77.93		5.89	2.27	146.29		5.89	2.27	−2451.03		5.89	2.27	−362.81	
5.79	1.93	74.88		5.79	1.93	140.65		5.79	1.93	−2066.68		5.79	1.93	−402.22	
5.85	3.00	102.92		5.85	3.00	183.21		5.85	3.00	−21831.26		5.85	3.00	−269.02	
6.02	1.88	76.83		6.02	1.88	145.35		6.02	1.88	−2375.81		6.02	1.88	−323.53	
5.83	2.18	77.93		5.83	2.18	145.48		5.83	2.18	−2451.03		5.83	2.18	−365.01	
5.99	2.01	80.98		5.99	2.01	157.94		5.99	2.01	−2835.36		5.99	2.01	−324.99	
5.91	2.13	77.93		5.91	2.13	147.35		5.91	2.13	−2451.03		5.91	2.13	−365.53	
5.92	1.80	82.68		5.92	1.80	152.62		5.92	1.80	−2490.34		5.92	1.80	−358.42	
6.01	2.82	80.91		6.01	2.82	148.48		6.01	2.82	−4564.72		6.01	2.82	−353.60	
5.92	1.80	76.38		5.92	1.80	146.25		5.92	1.80	−2375.80		5.92	1.80	−323.36	
6.02	1.92	77.93		6.02	1.92	145.05		6.02	1.92	−2451.03		6.02	1.92	−363.68	
5.97	2.58	83.89		5.97	2.58	149.74		5.97	2.58	−6678.40		5.97	2.58	−341.98	
5.92	2.75	86.88		5.92	2.75	156.43		5.92	2.75	−8792.08		5.92	2.75	−331.89	
6.12	1.99	76.23		6.12	1.99	143.95		6.12	1.99	−2475.06		6.12	1.99	−364.74	
6.11	2.37	80.98		6.11	2.37	151.35		6.11	2.37	−2835.38		6.11	2.37	−326.35	
6.03	3.44	83.96		6.03	3.44	155.36		6.03	3.44	−4949.06		6.03	3.44	−315.98	
6.05	3.47	89.31		6.05	3.47	162.62		6.05	3.47	−9295.45		6.05	3.47	−296.47	
5.97	3.31	92.91		5.97	3.31	166.51		5.97	3.31	−11290.11		5.97	3.31	−284.65	

So prepare the triparametric combination for IP + W

Step Forward Multi Linear Regressions

IP	MR	SASA	IP	MR	TE	IP	MR	DH
5.59	71.83	135.70	5.59	71.83	−1682.34	5.59	71.83	−440.50
5.62	77.93	152.92	5.62	77.93	−2451.02	5.62	77.93	−364.00
5.60	71.83	137.37	5.60	71.83	−1682.34	5.60	71.83	−441.14
5.59	74.88	140.25	5.59	74.88	−2066.68	5.59	74.88	−400.95
5.72	71.83	135.59	5.72	71.83	−1682.34	5.72	71.83	−438.21
5.61	71.83	133.58	5.61	71.83	−1682.34	5.61	71.83	−438.67
5.87	77.93	152.60	5.87	77.93	−2451.01	5.87	77.93	−359.04
5.89	74.88	142.07	5.89	74.88	−2066.68	5.89	74.88	−401.52
5.98	63.70	131.21	5.98	63.70	−1370.07	5.98	63.70	−479.38
5.73	74.88	139.27	5.73	74.88	−2066.69	5.73	74.88	−402.63
5.89	74.88	139.82	5.89	74.88	−2066.69	5.89	74.88	−399.69
5.89	77.93	146.29	5.89	77.93	−2451.03	5.89	77.93	−362.81
5.79	74.88	140.65	5.79	74.88	−2066.68	5.79	74.88	−402.22
5.85	102.92	183.21	5.85	102.92	−21831.26	5.85	102.92	−269.02
6.02	76.83	145.35	6.02	76.83	−2375.81	6.02	76.83	−323.53
5.83	77.93	145.48	5.83	77.93	−2451.03	5.83	77.93	−365.01
5.99	80.98	157.94	5.99	80.98	−2835.36	5.99	80.98	−324.99
5.91	77.93	147.35	5.91	77.93	−2451.03	5.91	77.93	−365.53
5.92	82.68	152.62	5.92	82.68	−2490.34	5.92	82.68	−358.42
6.01	80.91	148.48	6.01	80.91	−4564.72	6.01	80.91	−353.60
5.92	76.38	146.25	5.92	76.38	−2375.80	5.92	76.38	−323.36
6.02	77.93	145.05	6.02	77.93	−2451.03	6.02	77.93	−363.68
5.97	83.89	149.74	5.97	83.89	−6678.40	5.97	83.89	−341.98
5.92	86.88	156.43	5.92	86.88	−8792.08	5.92	86.88	−331.89
6.12	76.23	143.95	6.12	76.23	−2475.06	6.12	76.23	−364.74
6.11	80.98	151.35	6.11	80.98	−2835.38	6.11	80.98	−326.35
6.03	83.96	155.36	6.03	83.96	−4949.06	6.03	83.96	−315.98
6.05	89.31	162.62	6.05	89.31	−9295.45	6.05	89.31	−296.47
5.97	92.91	166.51	5.97	92.91	−11290.11	5.97	92.91	−284.65

And the other threes for IP + MR

Step Forward Multi Linear Regressions

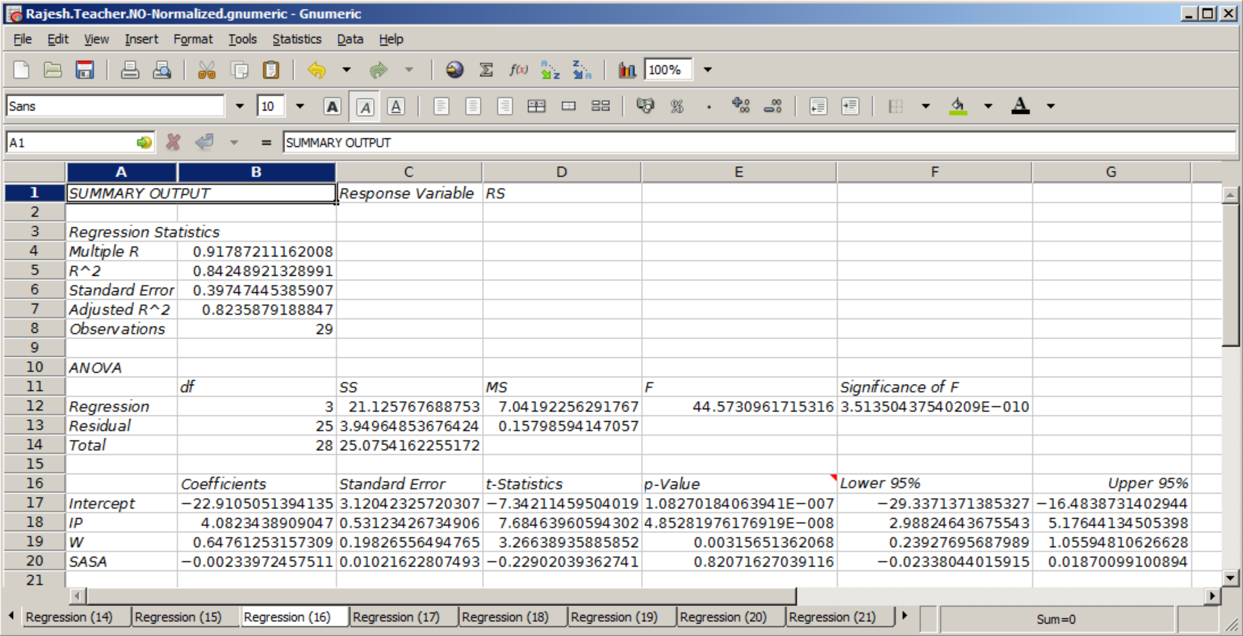
Rajesh.Teacher.NO-Normalized.gnumeric - Gnumeric

	A	B	C	D	E	F	G
1	*SUMMARY OUTPUT*		*Response Variable*	*RS*			
2							
3	*Regression Statistics*						
4	*Multiple R*	0.91937513026079					
5	*R^2*	0.84525063014204					
6	*Standard Error*	0.39397486517924					
7	*Adjusted R^2*	0.82668070575908					
8	*Observations*	29					
9							
10	*ANOVA*						
11		*df*	*SS*	*MS*	*F*	*Significance of F*	
12	*Regression*	3	21.1950113656923	7.06500378856409	45.5171821226071	2.82079971524628E−010	
13	*Residual*	25	3.88040485982496	0.155216194393			
14	*Total*	28	25.0754162255172				
15							
16		*Coefficients*	*Standard Error*	*t-Statistics*	*p-Value*	*Lower 95%*	*Upper 95%*
17	*Intercept*	−23.80208339730	3.10798434250768	−7.658366572753	5.15812132000212E−008	−30.2030969720504	−17.4010698225512
18	*IP*	4.07242818164555	0.52421975621695	7.76855151555969	3.99603520849363E−008	2.99277738360157	5.15207897968954
19	*W*	0.49322499897487	0.22586418121859	2.18372384817194	0.03857769116353	0.02804901006902	0.95840098788072
20	*MR*	0.01190276432794	0.01684151656373	0.70675133577771	0.48625672532005	−0.0227829883219	0.04658851697778
21							

Regression (8) | Regression (9) | Sheet1 | Sheet2 | Regression (10) | Regression (11) | Regression (12) | Regression (13) | Regression (14) | Regression (15) — Sum=0

And build the seven triparametric models

Step Forward Multi Linear Regressions



	A	B	C	D	E	F	G
1	*SUMMARY OUTPUT*		*Response Variable*	*RS*			
2							
3	*Regression Statistics*						
4	*Multiple R*	0.91787211162008					
5	*R^2*	0.84248921328991					
6	*Standard Error*	0.39747445385907					
7	*Adjusted R^2*	0.8235879188847					
8	*Observations*	29					
9							
10	*ANOVA*						
11		*df*	*SS*	*MS*	*F*	*Significance of F*	
12	*Regression*	3	21.125767688753	7.04192256291767	44.5730961715316	3.51350437540209E−010	
13	*Residual*	25	3.94964853676424	0.15798594147057			
14	*Total*	28	25.0754162255172				
15							
16		*Coefficients*	*Standard Error*	*t-Statistics*	*p-Value*	*Lower 95%*	*Upper 95%*
17	*Intercept*	−22.9105051394135	3.12042325720307	−7.34211459504019	1.08270184063941E−007	−29.3371371385327	−16.4838731402944
18	*IP*	4.08234389090 47	0.53123426734906	7.68463960594302	4.85281976176919E−008	2.98824643675543	5.17644134505398
19	*W*	0.64761253157309	0.19826556494765	3.26638935885852	0.00315651362068	0.23927695687989	1.05594810626628
20	*SASA*	−0.00233972457511	0.01021622807493	−0.22902039362741	0.82071627039116	−0.02338044015915	0.01870099100894
21							

Step Forward Multi Linear Regressions

	A	B	C	D	E	F	G
1	*SUMMARY OUTPUT*		*Response Variable*	*RS*			
2							
3	*Regression Statistics*						
4	*Multiple R*	0.91779372625134					
5	*R^2*	0.84234532394632					
6	*Standard Error*	0.3976559629602					
7	*Adjusted R^2*	0.82342676281988					
8	*Observations*	29					
9							
10	*ANOVA*						
11		*df*	*SS*	*MS*	*F*	*Significance of F*	
12	*Regression*	3	21.1221596035722	7.04071986785741	44.5248091710865	3.55355920184728E−010	
13	*Residual*	25	3.95325662194501	0.1581302648778			
14	*Total*	28	25.0754162255172				
15							
16		*Coefficients*	*Standard Error*	*t-Statistics*	*p-Value*	*Lower 95%*	*Upper 95%*
17	*Intercept*	−23.06835340559	2.99862854379441	−7.692967991427	4.75995517532941E−008	−29.2441444969226	−16.8925623142578
18	*IP*	4.05603702187155	0.53610620613174	7.56573413155911	6.40070834234719E−008	2.95190562197269	5.16016842177042
19	*W*	0.64017720163657	0.19949041255275	3.20906249801492	0.00363431772472	0.22931900607942	1.05103539719371
20	*TE*	4.29096886E−006	2.49469454E−005	0.17200377809545	0.86481936105994	−4.708822690858E−005	5.567016462E−005
21							

Step Forward Multi Linear Regressions

Rajesh.Teacher.NO-Normalized.gnumeric - Gnumeric

	A	B	C	D	E	F	G
1	*SUMMARY OUTPUT*		*Response Variable*	*RS*			
2							
3	*Regression Statistics*						
4	*Multiple R*	0.91997255702248					
5	*R^2*	0.84634950567447					
6	*Standard Error*	0.39257356500249					
7	*Adjusted R^2*	0.82791144635541					
8	*Observations*	29					
9							
10	*ANOVA*						
11		*df*	*SS*	*MS*	*F*	*Significance of F*	
12	*Regression*	3	21.2225661270482	7.07418870901607	45.9023095125543	2.58192448871272E−010	
13	*Residual*	25	3.85285009846905	0.15411400393876			
14	*Total*	28	25.0754162255172				
15							
16		*Coefficients*	*Standard Error*	*t-Statistics*	*p-Value*	*Lower 95%*	*Upper 95%*
17	*Intercept*	−21.00497066704	3.90386643144781	−5.380555671125	1.39678266422564E−005	−29.0451340874092	−12.9648072466733
18	*IP*	3.87297837817665	0.57470840268017	6.73903210761297	4.61167922794022E−007	2.68934426626531	5.05661249008799
19	*W*	0.51135155559572	0.19143621517046	2.67113281121017	0.01310353726964	0.11708129005888	0.90562182113255
20	*DH*	0.00197715383419	0.00239437102872	0.82575081742642	0.41675629700302	−0.00295414560906	0.00690845327743
21							

Regression (14) | Regression (15) | Regression (16) | Regression (17) | Regression (18) | Regression (19) | Regression (20) | Regression (21)

Sum=0

Step Forward Multi Linear Regressions

Rajesh.Teacher.NO-Normalized.gnumeric - Gnumeric

	A	B	C	D	E	F	G
1	*SUMMARY OUTPUT*		*Response Variable*	*RS*			
2							
3	*Regression Statistics*						
4	*Multiple R*	0.93029450784265					
5	*R^2*	0.86544787132221					
6	*Standard Error*	0.36736633654297					
7	*Adjusted R^2*	0.84930161588087					
8	*Observations*	29					
9							
10	*ANOVA*						
11		*df*	*SS*	*MS*	*F*	*Significance of F*	
12	*Regression*	3	21.7014655948922	7.23382186496407	53.6005313719128	4.96403320051664E−011	
13	*Residual*	25	3.37395063062503	0.134958025225			
14	*Total*	28	25.0754162255172				
15							
16		*Coefficients*	*Standard Error*	*t-Statistics*	*p-Value*	*Lower 95%*	*Upper 95%*
17	*Intercept*	−25.5541806822902	2.61478082168975	−9.77297235405616	5.08641309667451E−010	−30.9394225915613	−20.168938773019
18	*IP*	4.63976084281589	0.47464398225434	9.77524421731669	5.06281839926264E−010	3.66221326253052	5.61730842310127
19	*MR*	0.13889262202855	0.03393123830337	4.09335553234888	0.00038969353447	0.06900992860008	0.20877531545701
20	*SASA*	−0.07126526122011	0.02344818134093	−3.03926603875767	0.00549348056003	−0.11955769468371	−0.02297282775651
21							

Regression (14) | Regression (15) | Regression (16) | Regression (17) | Regression (18) | Regression (19) | Regression (20) | Regression (21)

Sum=0

Step Forward Multi Linear Regressions

Rajesh.Teacher.NO-Normalized.gnumeric - Gnumeric

	A	B	C	D	E	F	G
1	*SUMMARY OUTPUT*		*Response Variable*	*RS*			
2							
3	*Regression Statistics*						
4	*Multiple R*	0.90928297517243					
5	*R^2*	0.82679552893842					
6	*Standard Error*	0.41680567193787					
7	*Adjusted R^2*	0.80601099241103					
8	*Observations*	29					
9							
10	*ANOVA*						
11		*df*	*SS*	*MS*	*F*	*Significance of F*	
12	*Regression*	3	20.7322420215277	6.91074734050922	39.7793584595404	1.14188494827146E−009	
13	*Residual*	25	4.34317420398958	0.17372696815958			
14	*Total*	28	25.0754162255172				
15							
16		*Coefficients*	*Standard Error*	*t-Statistics*	*p-Value*	*Lower 95%*	*Up*
17	*Intercept*	−27.9060560150851	3.0277237249918	−9.21684359267643	1.61928437395792E−009	−34.1417697537929	−21.670342
18	*IP*	4.20517239877165	0.55347131795757	7.5978144889055	5.93895741308059E−008	3.06527688159461	5.3450679
19	*MR*	0.07043044183343	0.02637803064221	2.67004170207962	0.01313653540738	0.01610387078008	0.1247570
20	*TE*	5.59017392867349E−005	4.42388554946174E−005	1.26363439247511	0.21801810719595	−3.52098891341316E−005	0.0001470
21							

Regression (14) | Regression (15) | Regression (16) | Regression (17) | Regression (18) | Regression (19) | Regression (20) | Regression (21) — Sum=0

Step Forward Multi Linear Regressions

Rajesh.Teacher.NO-Normalized.gnumeric - Gnumeric

	A	B	C	D	E	F	G
1	*SUMMARY OUTPUT*		*Response Variable*	*RS*			
2							
3	*Regression Statistics*						
4	*Multiple R*	0.90353582275065					
5	*R^2*	0.8163769829937					
6	*Standard Error*	0.42915841271113					
7	*Adjusted R^2*	0.79434222095294					
8	*Observations*	29					
9							
10	*ANOVA*						
11		*df*	*SS*	*MS*	*F*	*Significance of F*	
12	*Regression*	3	20.4709926454989	6.82366421516631	37.0495030299709	2.35619727643141E−009	
13	*Residual*	25	4.60442358001831	0.18417694320073			
14	*Total*	28	25.0754162255172				
15							
16		*Coefficients*	*Standard Error*	*t-Statistics*	*p-Value*	*Lower 95%*	*Upper 95%*
17	*Intercept*	−25.2420591257467	6.38444099079169	−3.95368351937992	0.00055797571893	−38.3910614840635	−12.0930567674299
18	*IP*	4.2808836684156	0.68251744425095	6.27219671009844	1.45858235903516E−006	2.87521267905382	5.68655465777739
19	*MR*	0.03388456270511	0.02464994614108	1.37463029376118	0.18144331720092	−0.01688295169575	0.08465207710597
20	*DH*	0.00122449759115	0.00413475224001	0.29614775446424	0.76956340741969	−0.00729118405324	0.00974017923553
21							

Regression (14) | Regression (15) | Regression (16) | Regression (17) | Regression (18) | Regression (19) | Regression (20) | Regression (21) — Sum=0

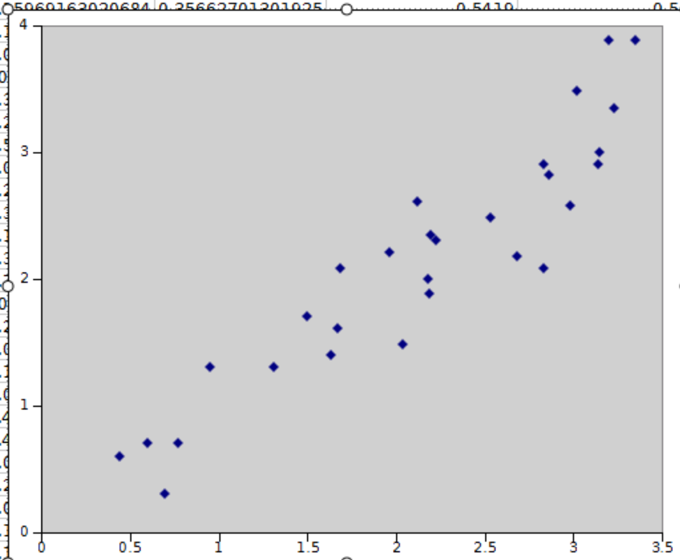
Triparametric models

		R^2
1		
2		R^2
3	IP + W + MR	0.84525063014204
4	IP + W + SASA	0.84248921328991
5	IP + W + TE	0.84234532394632
6	IP + W + DH	0.84634950567447
7	IP + MR + SASA	0.86544787132221
8	IP + MR + TE	0.82679552893842
9	IP + MR + DH	0.8163769829937
10		
11		

From the results the best triparametric model is that with «IP» + «MR» + «SASA»

Triparametric models

Prediction	RS	Residual	Leverages	Internally studentized	Externally studentized	p-V
0.6966998456292	0.301	−0.3956998456292	0.15990656393278	−1.1752	−1.1791	24.
0.44230836979316	0.602	0.15969163020684	0.35662701301925	0.5419	0.5[illegible]30	59.
0.59617463750167	0.6989	0.1			76	76.
0.76850826235235	0.6989	−0.0			25	84.
1.3123476947723	1.301	−0			18	97.
0.94950061165168	1.301	0.3			97	29.
1.63461804681808	1.3979	−0.2			78	48.
2.04015579198459	1.4771	−0.5			68	11.
1.66921236661669	1.602	−0.0			88	82.
1.49698046224425	1.6989	0.2			85	57.
2.19100931226055	1.8808	−0.3			11	39.
2.17709706197328	2	−0.1			33	63.
1.68629957401509	2.0791	0.3			40	28.
2.83566723250474	2.0791	−0.			12	0.
2.6806857671476	2.176	−0			75	15.
1.96514567258849	2.2041	0.2			59	51.
2.22488712959675	2.301	0.0			42	81.
2.19363844932408	2.3424	0.1			64	68.
2.53391156111041	2.4771	−0.0			52	87.
2.98268844244182	2.574	−0.4			90	25.
2.11825324551683	2.602	0.4			87	18.
2.86356754304214	2.8129	−0.0			14	88.
3.14025077973549	2.903	−0.2			00	49.
2.83182064585324	2.903	0.0			05	84.
3.14765902871825	3	−0.1			89	67.
3.23287998932635	3.3424	0.1			03	75.
3.01829083885974	3.4771	0.45880916114026	0.07286085508795	1.2971	1.3125	20.
3.34992206546706	3.875	0.52507793453294	0.11986408862865	1.5235	1.5577	13.
3.20251957115419	3.875	0.67248042884581	0.16466634031334	2.0029	2.1088	4.



The fitting plot can be made by the «Insert → Chart» menu

Compare result with those published

$$^{SUCROSE}PS1 = 5.08109 \times IP - 27.7204.$$

$r^2 = 0.727709$, $rCV^2 = 0.713409$, $SE = 0.1177$, $SEE = 0.5028$, t-value = 8.4963, P-value = 0, DOF = 0.7177, n = 29.

$$^{SUCROSE}PS2 = 4.07088 \times IP + 0.616562 \times \omega - 23.1215.$$

$r^2 = 0.842159$, $rCV^2 = 0.805992$, $SE = 0.0833$, $SEE = 0.3829$, t-value = 12.0028, P-value = 0, DOF = 0.8363, n = 29.

$$^{SUCROSE}PS3 = 4.63976 \times IP + 0.138893 \times MR - 0.0712653 \times SASA - 25.5542.$$

$r^2 = 0.865448$, $rCV^2 = 0.816901$, $SE = 0.0759$, $SEE = 0.3536$, t-value = 13.1759, P-value = 0, DOF = 0.8604, n = 29.

Are the same!

Comparing the results with those reported in the publication the same results were obtained using data not normalized.
The same results are obtained either with normalized or autoscaled data!