

VALIDATION OF STATEL RESULTS

NIST (National Institute of Standards and Technology) and its laboratory ITL (Information Technology Laboratory), located in USA, designed reference datasets with certified computational results that enable the objective evaluation of statistical software.

In order to certify users of StatEL about accuracy of its calculations, we confronted its results to certified results of ITL.

You can visit website “Statistical Reference Datasets“ of NIST on :

<http://www.itl.nist.gov/div898/strd/>

We used only datasets that are applicable with available tests of StatEL software, that is :

- Univariate statistics :
 - [PiDigits](#) (Lower level of difficulty)
 - [Lottery](#) (Lower level of difficulty)
 - [NumAcc2](#) (Average level of difficulty)
 - [NumAcc3](#) (Average level of difficulty)
 - [NumAcc4](#) (Higher level of difficulty)

- One-Way Anova :
 - [SiRstv](#) (Lower level of difficulty)
 - [SmLs01](#) (Lower level of difficulty)
 - [AtmWtAg](#) (Average level of difficulty)
 - [SmLs06](#) (Average level of difficulty)
 - [SmLs07](#) (Higher level of difficulty)
 - [SmLs08](#) (Higher level of difficulty)

- Simple and multiple linear regression
 - [Norris](#)
 - [Longley](#)

Following results are displayed through screenshots of NIST results and those supplied by StatEL on Excel spreadsheet.

UNIVARIATE STATISTICS

- Set [PiDigits](#) (Lower level of difficulty)

Certified Results :

Certified Values									
Dataset Name:	PiDigits								
Procedure:	Univariate Summary Statistics Certification Method & Definitions								
Data:	1 Response Variable (y) 5000 Observations Lower Level of Difficulty Observed Data								
Model:	3 Parameters (μ , σ , ρ_1) $y_i = \mu + \epsilon_i$								
	<table border="1"> <thead> <tr> <th>Parameter</th> <th>Certified Estimate</th> </tr> </thead> <tbody> <tr> <td>μ</td> <td>4.534800000000000</td> </tr> <tr> <td>σ</td> <td>2.86733906028871</td> </tr> <tr> <td>ρ_1</td> <td>-0.00355099287237972</td> </tr> </tbody> </table>	Parameter	Certified Estimate	μ	4.534800000000000	σ	2.86733906028871	ρ_1	-0.00355099287237972
Parameter	Certified Estimate								
μ	4.534800000000000								
σ	2.86733906028871								
ρ_1	-0.00355099287237972								

Results of StatEL :

Results	
Nb of values	5000
Minimum	0
Maximum	9
Range	9
Mean	4,535
Variation Co	0,632
Standard Er	0,0406
CI 95% (Con	4,455 < m < 4,614
Standard De	2,867
Variance	8,222
Median	5
25th Percent	2
75th Percent	7
Inter Percent	5
Skewness (F	0,122
Kurtosis (Fis	-1,453

- Set **Lottery** (Lower level of difficulty)

Certified Results :

Certified Values	
Dataset Name:	Lottery
Procedure:	Univariate Summary Statistics Certification Method & Definitions
Data:	1 Response Variable (y) 218 Observations Lower Level of Difficulty Observed Data
Model:	3 Parameters (μ , σ , ρ_1) $y_i = \mu + \varepsilon_i$

Parameter	Certified Estimate
μ	518.958715596330
σ	291.699727470969
ρ_1	-0.120948622967393

Results of StatEL :

Results	
Nb of values	218
Minimum	4
Maximum	999
Range	995
Mean	519,0
Variation Co	0,562
Standard Er	19,76
CI 95% (Com	480,2 < m < 557,7
Standard De	291,7
Variance	85089
Median	522,5
25th Percent	272,8
75th Percent	779,3
Inter Percent	506,5
Skewness (F	-0,0891
Kurtosis (Fis	-1,257

- Set **NumAcc2** (Average level of difficulty)

Certified Results :

Certified Values	
Dataset Name:	NumAcc2
Procedure:	Univariate Summary Statistics Certification Method & Definitions
Data:	1 Response Variable (y) 1001 Observations Average Level of Difficulty Generated Data
Model:	3 Parameters (μ , σ , ρ_1) $y_i = \mu + \epsilon_i$

Parameter	Certified Estimate
μ	1.2 (exact)
σ	0.1 (exact)
ρ_1	-0.999 (exact)

Results of StatEL :

Results	
Nb of values	1001
Minimum	1,1
Maximum	1,3
Range	0,2
Mean	1,200
Variation Co	0,0833
Standard Er	3,16E-03
CI 95% (Com	1,194 < m < 1,206
Standard De	0,1
Variance	0,0100
Median	1,2
25th Percent	1,1
75th Percent	1,3
Inter Percent	0,2
Skewness (F	2,93E-13
Kurtosis (Fis	-2,242

- Set [NumAcc3](#) (Average level of difficulty)

Certified Results :

Certified Values	
Dataset Name:	NumAcc3
Procedure:	Univariate Summary Statistics Certification Method & Definitions
Data:	1 Response Variable (y) 1001 Observations Average Level of Difficulty Generated Data
Model:	3 Parameters (μ , σ , ρ_1) $y_i = \mu + \epsilon_i$
Parameter	Certified Estimate
μ	1000000.2 (exact)
σ	0.1 (exact)
ρ_1	-0.999 (exact)

Results of StatEL :

Results			
Nb of values	1001		
Minimum	1000000,1		
Maximum	1000000,3		
Range	0,200		
Mean	1000000,2		
Variation Co	1,00E-08		
Standard Er	3,16E-03		
CI 95% (Con	1000000,19 < m < 1000000,21		
Standard De	0,100		
Variance	0,0100		
Median	1000000,2		
25th Percent	1000000,1		
75th Percent	1000000,3		
Inter Percent	0,200		
Skewness (F	-2,55E-06		
Kurtosis (Fis	-2,242		

- Set [NumAcc4](#) (Higher level of difficulty)

Certified Results :

Certified Values	
Dataset Name:	NumAcc4
Procedure:	Univariate Summary Statistics Certification Method & Definitions
Data:	1 Response Variable (y) 1001 Observations Higher Level of Difficulty Generated Data
Model:	3 Parameters (μ , σ , ρ_1) $y_i = \mu + \epsilon_i$
Parameter	Certified Estimate
μ	10000000,2 (exact)
σ	0,1 (exact)
ρ_1	-0,999 (exact)

Results of StatEL :

Results	
Nb of values	1001
Minimum	10000000,1
Maximum	10000000,3
Range	0,200
Mean	10000000,2
Variation Co	1,00E-08
Standard Er	3,16E-03
CI 95% (Con	10000000,19 < m < 10000000,21
Standard De	0,100
Variance	0,0100
Median	10000000,2
25th Percent	10000000,1
75th Percent	10000000,3
Inter Percent	0,200
Skewness (F	-2,55E-06
Kurtosis (Fis	-2,242

ONE-WAY ANOVA

- Set [SiRstv](#) (Lower level of difficulty)

Certified Results :

Certified Values				
Dataset				
Name:	SiRstv			
Procedure: Analysis of Variance Certification Method & Definitions				
Data:				
	1 Factor			
	5 Treatments			
	5 Replicates/Cell			
	25 Observations			
	3 Constant Leading Digits			
	Lower Level of Difficulty			
	Observed Data			
Model:				
	6 Parameters ($\mu, \tau_1, \dots, \tau_5$)			
	$y_{ij} = \mu + \tau_i + \epsilon_{ij}$			
<hr/>				
Source of Variation	Certified Degrees of Freedom	Certified Sums of Squares	Certified Mean Squares	Certified F Statistic
Between Instrument	4	5.114626160000000E-02	1.278656540000000E-02	1.18046237440255E+00
Within Instrument	20	2.166365600000000E-01	1.083182800000000E-02	
Certified R-Squared			1.90999039051129E-01	
Certified Residual Standard Deviation			1.04076068334656E-01	

Results of StatEL :

Results of One-way Anova	
Between-group	0,0128
Within-group	0,0108
F :	1,18
F lim :	2,866
p <	0,35

- Set [SmLs01](#) (Lower level of difficulty)

Certified Results :

Certified Values																													
Dataset																													
Name:	SmLs01																												
Procedure: Analysis of Variance Certification Method & Definitions																													
Data:																													
1 Factor																													
9 Treatments																													
21 Replicates/Cell																													
189 Observations																													
1 Constant Leading Digit																													
Lower Level of Difficulty																													
Generated Data																													
Model: 10 Parameters ($\mu, \tau_1, \dots, \tau_9$) $y_{ij} = \mu + \tau_i + \epsilon_{ij}$																													
<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="width: 15%;">Source of Variation</th> <th style="width: 10%;">Certified Degrees of Freedom</th> <th style="width: 15%;">Certified Sums of Squares</th> <th style="width: 15%;">Certified Mean Squares</th> <th style="width: 15%;">Certified F Statistic</th> </tr> </thead> <tbody> <tr> <td>Between Treatment</td> <td style="text-align: center;">8</td> <td style="text-align: right;">1.68000000000000E+00</td> <td style="text-align: right;">2.10000000000000E-01</td> <td style="text-align: right;">2.10000000000000E+01</td> </tr> <tr> <td>Within Treatment</td> <td style="text-align: center;">180</td> <td style="text-align: right;">1.80000000000000E+00</td> <td style="text-align: right;">1.00000000000000E-02</td> <td></td> </tr> <tr> <td colspan="3">Certified R-Squared</td> <td style="text-align: right;">4.82758620689655E-01</td> <td></td> </tr> <tr> <td colspan="3">Certified Residual Standard Deviation</td> <td style="text-align: right;">1.00000000000000E-01</td> <td></td> </tr> </tbody> </table>					Source of Variation	Certified Degrees of Freedom	Certified Sums of Squares	Certified Mean Squares	Certified F Statistic	Between Treatment	8	1.68000000000000E+00	2.10000000000000E-01	2.10000000000000E+01	Within Treatment	180	1.80000000000000E+00	1.00000000000000E-02		Certified R-Squared			4.82758620689655E-01		Certified Residual Standard Deviation			1.00000000000000E-01	
Source of Variation	Certified Degrees of Freedom	Certified Sums of Squares	Certified Mean Squares	Certified F Statistic																									
Between Treatment	8	1.68000000000000E+00	2.10000000000000E-01	2.10000000000000E+01																									
Within Treatment	180	1.80000000000000E+00	1.00000000000000E-02																										
Certified R-Squared			4.82758620689655E-01																										
Certified Residual Standard Deviation			1.00000000000000E-01																										

Results of StatEL :

Results of One-way Anova	
Between-gro	0,210
Within-grou	0,01
F :	21
F lim :	1,99
p <	0,00001

- [AtmWtAg](#) (Average level of difficulty)

Certified Results :

Certified Values				
Dataset Name:	AtmWtAg			
Procedure:	Analysis of Variance Certification Method & Definitions			
Data:	1 Factor 2 Treatments 24 Replicates/Cell 48 Observations 7 Constant Leading Digits Average Level of Difficulty Observed Data			
Model:	3 Parameters (μ, τ_1, τ_2) $y_{ij} = \mu + \tau_i + \epsilon_{ij}$			

Source of Variation	Certified Degrees of Freedom	Certified Sums of Squares	Certified Mean Squares	Certified F Statistic
Between Instrument	1	3.63834187500000E-09	3.63834187500000E-09	1.59467335677930E+01
Within Instrument	46	1.04951729166667E-08	2.28155932971014E-10	
Certified R-Squared			2.57426544538321E-01	
Certified Residual Standard Deviation			1.51048314446410E-05	

Results of StatEL :

Results of One-way Anova	
Between-group	3,64E-09
Within-group	2,28E-10
F :	15,95
F lim :	4,052
p <	0,00023

- Set [SmLs06](#) (Average level of difficulty)

Certified Results :

Certified Values				
Dataset				
Name:	SmLs06			
Procedure: Analysis of Variance Certification Method & Definitions				
Data:				
1 Factor				
9 Treatments				
2001 Replicates/Cell				
18009 Observations				
7 Constant Leading Digits				
Average Level of Difficulty				
Generated Data				
Model: 10 Parameters ($\mu, \tau_1, \dots, \tau_9$) $y_{ij} = \mu + \tau_i + \epsilon_{ij}$				
<hr/>				
Source of Variation	Certified Degrees of Freedom	Certified Sums of Squares	Certified Mean Squares	Certified F Statistic
Between Treatment	8	1.600800000000000E+02	2.001000000000000E+01	2.001000000000000E+03
Within Treatment	18000	1.800000000000000E+02	1.000000000000000E-02	
Certified R-Squared			4.70712773465067E-01	
Certified Residual Standard Deviation			1.000000000000000E-01	

Results of StatEL :

Results of One-way Anova	
Between-gr	20,01
Within-group	0,0100
F :	2001
F lim :	1,939
p <	0,00001

- Set [SmLs07](#) (Higher level of difficulty)

Certified Results :

Certified Values				
Dataset Name:	SmLs07			
Procedure:	Analysis of Variance Certification Method & Definitions			
Data:	1 Factor 9 Treatments 21 Replicates/Cell 189 Observations 13 Constant Leading Digits Higher Level of Difficulty Generated Data			
Model:	10 Parameters ($\mu, \tau_1, \dots, \tau_9$) $y_{ij} = \mu + \tau_i + \epsilon_{ij}$			

Source of Variation	Certified Degrees of Freedom	Certified Sums of Squares	Certified Mean Squares	Certified F Statistic
Between Treatment	8	1.680000000000000E+00	2.100000000000000E-01	2.100000000000000E+01
Within Treatment	180	1.800000000000000E+00	1.000000000000000E-02	
Certified R-Squared			4.82758620689655E-01	
Certified Residual Standard Deviation			1.000000000000000E-01	

Results of StatEL :

Results of One way Anova	
Between-gr	0,209897513
Within-group	0,010000547
F :	20,99
F lim :	1,99
p <	0,00001

Note :

There is a slight difference between both results on this dataset whose data are 13 digits numbers.

- Set [SmLs08](#) (Higher level of difficulty)

Certified Results :

Certified Values				
Dataset				
Name:	SmLs08			
Procedure:	Analysis of Variance Certification Method & Definitions			
Data:	1 Factor 9 Treatments 201 Replicates/Cell 1809 Observations 13 Constant Leading Digits Higher Level of Difficulty Generated Data			
Model:	10 Parameters ($\mu, \tau_1, \dots, \tau_9$) $y_{ij} = \mu + \tau_i + \epsilon_{ij}$			

Source of Variation	Certified Degrees of Freedom	Certified Sums of Squares	Certified Mean Squares	Certified F Statistic
Between Treatment	8	1.608000000000000E+01	2.010000000000000E+00	2.010000000000000E+02
Within Treatment	1800	1.800000000000000E+01	1.000000000000000E-02	
Certified R-Squared			4.71830985915493E-01	
Certified Residual Standard Deviation			1.000000000000000E-01	

Results of StatEL :

Results of One-way Anova	
Between-group	2,009018674
Within-group	0,010000545
F :	200,9
F lim :	1,944
p <	0,00001

Note :

There is a slight difference between both results on this dataset whose data are 13 digits numbers.

- Set [SmLs09](#) (Higher level of difficulty)

Certified Results :

Certified Values				
Dataset				
Name:	SmLs09			
Procedure: Analysis of Variance Certification Method & Definitions				
Data:				
	1 Factor			
	9 Treatments			
	2001 Replicates/Cell			
	18009 Observations			
	13 Constant Leading Digits			
	Higher Level of Difficulty			
	Generated Data			
Model:				
	10 Parameters ($\mu, \tau_1, \dots, \tau_9$)			
	$y_{ij} = \mu + \tau_i + \epsilon_{ij}$			
<hr/>				
Source of Variation	Certified Degrees of Freedom	Certified Sums of Squares	Certified Mean Squares	Certified F Statistic
Between Treatment	8	1.600800000000000E+02	2.001000000000000E+01	2.001000000000000E+03
Within Treatment	18000	1.800000000000000E+02	1.000000000000000E-02	
		Certified R-Squared	4.70712773465067E-01	
		Certified Residual Standard Deviation	1.000000000000000E-01	

Results of StatEL :

Results of One-way Anova	
Between-group	20,02466590
Within-group	0,010000545
F :	2002
F lim :	1,939
p <	0,00001

Note :

There is a slight difference between both results on this dataset whose data are 13 digits numbers.

SIMPLE LINEAR REGRESSION

- Set [Norris](#)

Certified Results :

Certified Values

Dataset
Name: Norris

Procedure: Linear Least Squares Regression
[Certification Method & Definitions](#)

Data: 1 Response Variable (y)
 1 Predictor Variable (x)
 36 Observations
 Lower Level of Difficulty
 Observed Data

Model: $y = \beta_0 + \beta_1 x + \epsilon$

Certified Regression Statistics

Parameter	Estimate	Standard Deviation of Estimate
β_0	-0.262323073774029	0.232818234301152
β_1	1.00211681802045	0.429796848199937E-03
Residual Standard Deviation	0.884796396144373	
R-Squared	0.999993745883712	

Results of StatEL :

Conclusions for analysis of Simple Linear Regression :	
Linear regression model : $Y = aX + b$	
$y = 1,002 * x - 0,2623$	
R² (Coefficient of Determination) :	99,999375%
This coefficient expresses percentage of variability that is expl:	
R²aj (Adjusted Coefficient of Determin	99,999356%
This coefficient allows to compare several linear models that d	

MULTIPLE LINEAR REGRESSION

- Set [Longley](#)

Certified Results :

Certified Values		
Dataset		
Name:	Longley	
Procedure:	Linear Least Squares Regression Certification Method & Definitions	
Data:	1 Response Variable (y) 6 Predictor Variable (x_1, \dots, x_6) 16 Observations Higher Level of Difficulty Observed Data	
Model:	$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \beta_6 x_6 + \epsilon$	
Certified Regression Statistics		
Parameter	Estimate	Standard Deviation of Estimate
β_0	-3482258.63459582	890420.383607373
β_1	15.0618722713733	84.9149257747669
β_2	-0.358191792925910E-01	0.334910077722432E-01
β_3	-2.02022980381683	0.488399681651699
β_4	-1.03322686717359	0.214274163161675
β_5	-0.511041056535807E-01	0.226073200069370
β_6	1829.15146461355	455.478499142212
Residual		
Standard Deviation	304.854073561965	
R-Squared	0.995479004577296	

Results of StatEL :

Model of linear multiple regression :				
$y = 15,06 * x1 - 0,03582 * x2 - 2,02 * x3 - 1,033 * x4 - 0,0511 * x5 + 1829 * x6 - 3482259$				
Coefficient of multiple correlation (expresses intensity of the linear relation between the				
r =	0,99774746			
Test of r versus 0 : p < 0,00001				
Coefficient of multiple determination (expresses percentage of variability that is explain				
R² =	99,55%			
Coefficient of adjusted multiple determination (allows to compare several linear models				
R²aj =	99,25%			
Tests on model parameters :				
Hypotheses :				
H0 = studied parameter is not significantly different from 0.				
H1 = studied parameter is significantly different from 0.				
Variable	Parameter	Standard-De t	p	Confidence Interval (95%)
Constant	-3482259	890420	-3,911	0,0036 [-5496529 ; -1467988]
x1	15,06	84,91	0,1774	0,86 [-177 ; 207,2]
x2	-0,03582	0,03349	-1,07	0,31 [-0,1116 ; 0,03994]
x3	-2,02	0,4884	-4,136	0,0025 [-3,125 ; -0,9154]
x4	-1,033	0,2143	-4,822	0,00094 [-1,518 ; -0,5485]
x5	-0,0511	0,2261	-0,2261	0,83 [-0,5625 ; 0,4603]
x6	1829	455,5	4,016	0,003 [798,8 ; 2860]

CONCLUSIONS

Except for data with more than 13 digits numbers, where Anova calculations differ very slightly, results of StatEL software are identical to certified results supplied by the NIST, on tested datasets.

Difference is due to the truncation error, the cancellation error, and/or the accumulation error. Any typical numerical algorithm will introduce computational inaccuracies, and will produce results which differ slightly from these certified values.

Consequently, if you need very precise and exact results for dataset whose values have more than 13 digits numbers, we suggest not to use StatEL, except if you are in a position to transform your data (see below).

One solution to avoid inaccuracies in calculation with such values is to subtract the leading constant from all the observations in that dataset before analyzing it.

Ex : if your data look all like 1000000000000.1, you just need to subtract 1000000000000 in order to launch calculations with 0.1...