A. Checking for Outliers

				1
	Scores	var	var	var
1	145.00			
2	123.00			
3	187.00			
4	166.00			
5	155.00			
6	190.00			
7	122.00			
8	144.00			
9	165.00			
10	175.00			
11	200.00			
12	133.00			
13	126.00			
14	153.00			
15	183.00			
16	127.00			
17	126.00			
18	189.00			
19	144.00			
20	113.00			
21	178.00			
22	156.00			
23	135.00			
24	185.00			
25	163.00			
26	145.00			
27	138.00			
28	174.00			
29	250.00			
30	60.00			

You may have a data set which looks similar to the one above







The following will appear in the output window. Cases 29 and 30 in our fictitious data set are outliers.



B. Checking for Normality

; 🔜 📖			-
	score1	var	var
1	145.00		
2	90.00		
3	187.00		
4	166.00		
5	155.00		
6	190.00		
7	122.00		
8	40.00		
9	20.00		
10	133.00		
11	50.00		
12	133.00		
13	90.00		
14	153.00		
15	183.00		
16	80.00		
17	100.00		
18	189.00		
19	192.00		
20	113.00		
21	178.00		
22	110.00		
23	135.00		
24	185.00		
25	163.00		
26	205.00		
27	138.00		
28	174.00		
29	10.00		
30	70.00		
31			
30			

You may have a data set which looks similar to the one above

	Click on Graphs, Leg F	gacy Di blot	ialogs and Box-	
lata.sav (Databet1) - P	ASW Statutics Data Edito	or		
iew <u>D</u> ata <u>T</u> ransform	Analyze Direct Marketing	<u>G</u> raphs	Unities Add-ons	<u>Window H</u> elp
	Reports Descriptive Statistics Tables	•	Frequencies	
score1 var	Compare Means General Linear Model	, , ,	🔤 Descriptives 4 Explore	var
90.00	Generali <u>z</u> ed Linear Mode Mixed Models	els 🕨	<mark>₩</mark> <u>C</u> rosstabs <mark>102</mark> <u>R</u> atio	
166.00	_ <u>C</u> orrelate <u>R</u> egression	• •	<u>ዎ</u> P-P Plots 🛃 <u>Q</u> -Q Plots	
190.00	Loglinear Neural Net <u>w</u> orks	•		
40.00	Classi <u>f</u> y <u>D</u> imension Reduction	۲ ۲		
133.00	Sc <u>a</u> le <u>N</u> onparametric Tests	*		
133.00	Forecasting Survival	> >		
153.00	Missing Value Analysis			
80.00	Multiple imputation Complex Samples	•		
189.00	ROC Curve	P		



Under 'Central Tendency' check the box marked 'Mean' Under 'Dispersion' check 'Std deviation', 'minimum', 'maximum', and under 'Distribution' check the boxes marked 'Skewness' and 'Kurtosis' Then 'Continue'

	Frequencies	: Statistics]	
	ercentile Value Quartiles Cut points fo Percentile(s Change Remove	25 pr: 10 equal):	groups	-Central Tende Mean Median Mode Sum Values are	group midpoints		
	ispersion	n v Minimum v Ma <u>x</u> imum S. <u>E</u> . mean Continue	Cancel	Distribution	3		







The output file appears as below

Statistics		
score1		
N Valid	30	
Missing	0	
Mean	129.9667	
Std. Deviation	54.36054	
Skewness	669	
Std. Error of Skewness	.427	
Kurtosis	437	
Std. Error of Kurtosis	.833	
Minimum	10.00	
Maximum	205.00	



Running the K-S test



	Click on 'Run'
🔢 One-Sar	mple Nonparametric Tests
Objective	Fields Settings
Identifies follow th	s differences in single fields using one or more nonparametric tests. Nonparametric tests do not assume your data normal distribution.
−What is Each o desire	s your objective? bjective corresponds to a distinct default configuration on the Settings Tab that you can further customize, if d.
©	Automatically compare observed data to typothesized Test sequence for randomness
C) <u>C</u> ustomize analysis
Descri Autom Kolmod	ption atically compares observed data to ypothesized using the Binomial test, Chi-Square test, or porov-Smirnov. The test chosen valies based on your data.
	Run Paste Reset Cancel Help

The following output will appear.

ion	parametric	Tests				
Data	aSet1] N:\d.	Other work\b.	Writing\3.	Methods	Bk\Normal	data.sa
		Hypothesis	Test Summa	у оі		1
	Null H	Hypothesis ypothesis	Test Summa _{Test}	Ƴ Sig.	Decision]

Asymptotic significances are displayed. The significance level is .05.

Transforming your data

Positive Skew



New Variable = SQRT (old variable)

Negative Skew



New variable = SQRT (K – old variable) where K = largest possible value + 1

Click on 'Transform' and 'Compute Variable'							
				1			
ew <u>D</u> a	ıta	Transform	<u>Ans</u> yze	Direct <u>M</u> arketing	Graphs	Utilities	Add- <u>o</u> ns
	ſ	📑 <u>C</u> ompu	te Variable.			88	
		Count "	Values with	in Cases			
		Shi <u>f</u> t ∨	alues				
1002	re1	🔤 Recode	e into Same	Variables		var	va
1	45.1		– e into Differ	ent Variables			
	90.1						
1	87.1	Autom:	atic Recode	·			
1	66.1	📑 Visual	<u>B</u> inning				
1	55.1	🔀 Optima	I Binning				
1	90.1	Prepar	e Data for N	Aodeling	•		
1	22.1	🛃 Ran <u>k</u> C	ases				
	40.1	🗎 Date a	nd Time Wiz	ard			
	20.1	树 Create	Time Serie:	s			
1	33.1	🍕 🖁 Replac	e Missing <u>V</u>	alues			
	50.1	🍘 Randoi	m Number G	enerators			
1	33.1	Run Pe	nding Trans	sforms Ctrl	+G		
	00.	_					

