

Quando fazemos teste a normalidade?

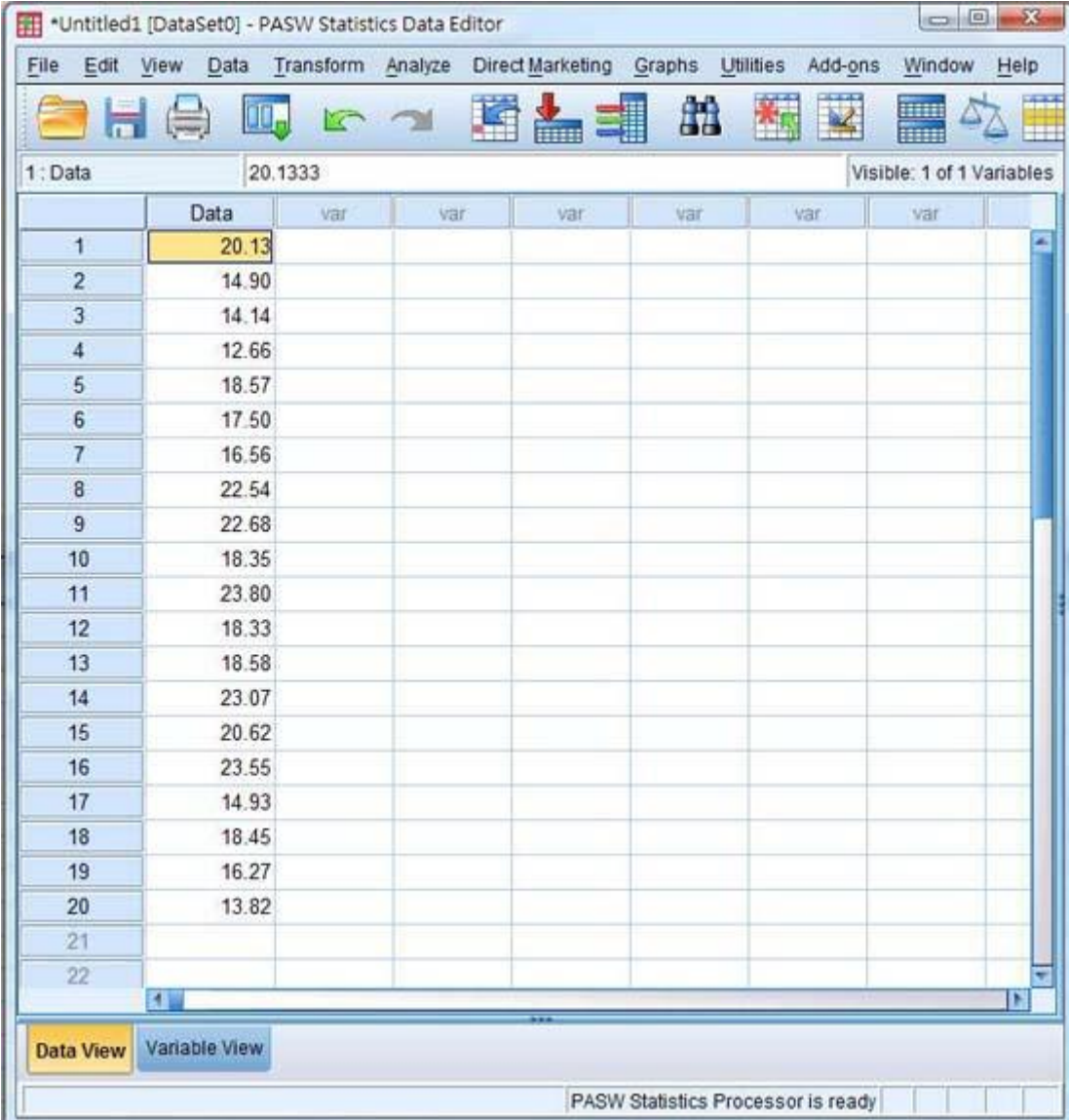
Bastantes testes estatísticos (por exemplo, teste t) exigem que os nossos dados sejam normalmente distribuídos e, portanto, deve sempre verificar se esta suposição é violada.

Exemplo Prático:

Dado um conjunto de dados, gostaríamos de verificar se sua distribuição é normal.

Neste exemplo, a hipótese nula é que os dados são normalmente distribuídos e a hipótese alternativa é que os dados não estão normalmente distribuídos.

Os dados a serem testados em armazenado na primeira coluna.



*Untitled1 [DataSet0] - PASW Statistics Data Editor

File Edit View Data Transform Analyze Direct Marketing Graphs Utilities Add-ons Window Help

1: Data 20.1333 Visible: 1 of 1 Variables

	Data	var	var	var	var	var	var
1	20.13						
2	14.90						
3	14.14						
4	12.66						
5	18.57						
6	17.50						
7	16.56						
8	22.54						
9	22.68						
10	18.35						
11	23.80						
12	18.33						
13	18.58						
14	23.07						
15	20.62						
16	23.55						
17	14.93						
18	18.45						
19	16.27						
20	13.82						
21							
22							

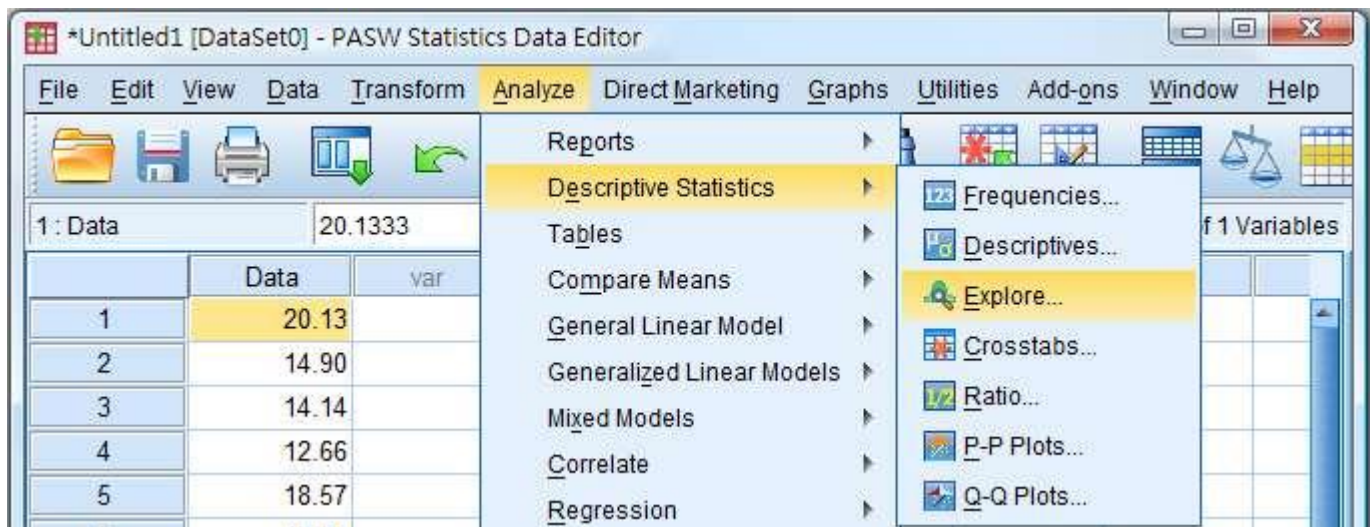
Data View Variable View

PASW Statistics Processor is ready

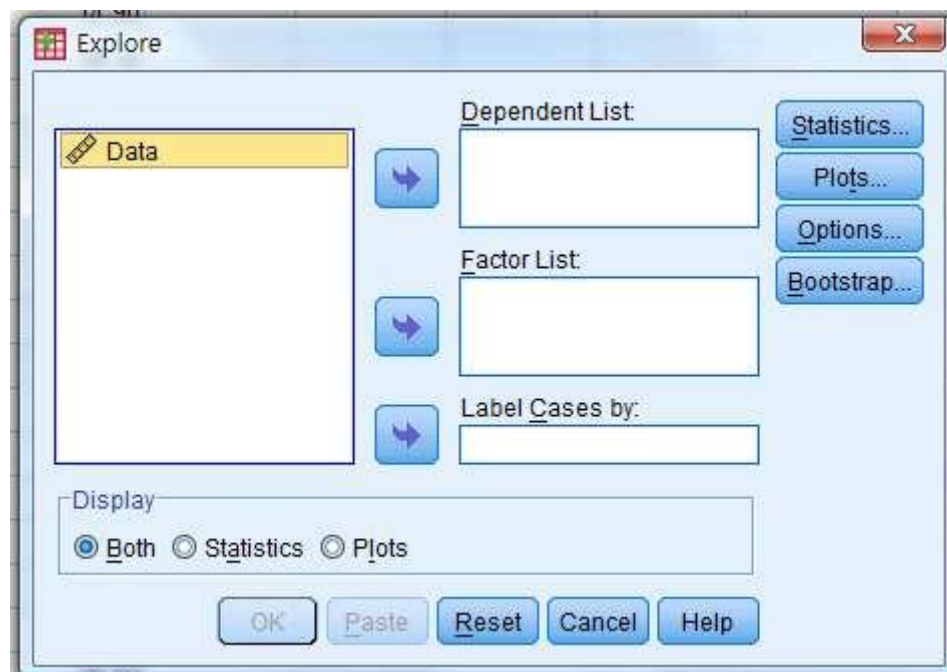
Passo

1

Selecione "Analisar -> Estatística Descritiva -> Explorar".



Uma nova janela surge.



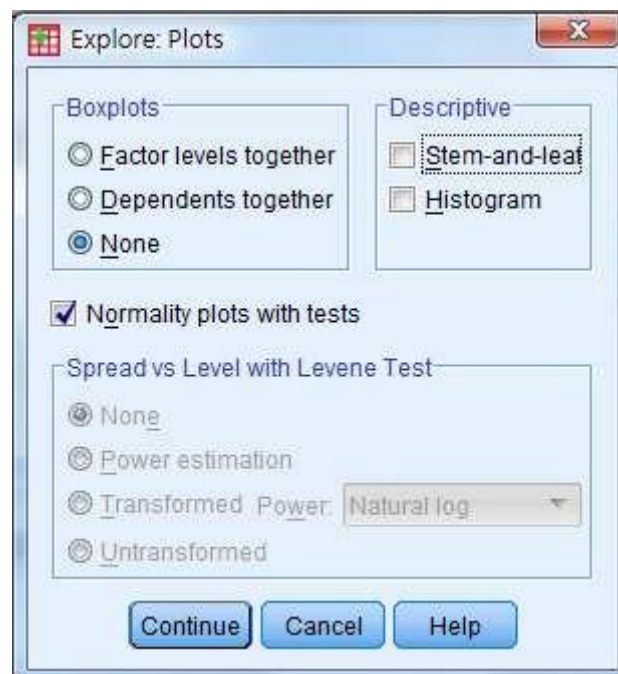
Passo

2

Na lista à esquerda, selecione a variável "dados" a "Lista Dependente".



Clique em "Lotes" à direita. Uma nova janela aparece fora. Marque a opção "None" para boxplot, desmarque tudo para descritivo e verifique se a caixa "parcelas de normalidade com testes" está marcada.



Passo

3

Os resultados agora aparecer na "saída" janela.

Explore
[DataSet0]

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Data	20	100.0%	0	.0%	20	100.0%

Descriptives

		Statistic	Std. Error
Data	Mean	18.4719	.77126
	95% Confidence Interval for Mean	Lower Bound: 16.8577	
		Upper Bound: 20.0862	
	5% Trimmed Mean	18.4986	
	Median	18.3962	
	Variance	11.897	
	Std. Deviation	3.44918	
	Minimum	12.66	
	Maximum	23.80	
	Range	11.14	
	Interquartile Range	6.80	
	Skewness	.070	.512
	Kurtosis	-1.060	.992

Tests of Normality

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Data	.137	20	.200 ^a	.946	20	.316

a. Lilliefors Significance Correction
*. This is a lower bound of the true significance.

PASW Statistics Processor is ready

Passo

Podemos agora interpretar o resultado.

4

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Data	20	100.0%	0	.0%	20	100.0%

Descriptives

				Statistic	Std. Error
Data	Mean			18.4719	.77126
	95% Confidence Interval for Mean	Lower Bound		16.8577	
		Upper Bound		20.0862	
	5% Trimmed Mean			18.4986	
	Median			18.3962	
	Variance			11.897	
	Std. Deviation			3.44918	
	Minimum			12.66	
	Maximum			23.80	
	Range			11.14	
	Interquartile Range			6.80	
	Skewness			.070	.512
	Kurtosis			-1.060	.992

Tests of Normality

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Data	.137	20	.200*	.946	20	.316

a. Lilliefors Significance Correction

*. This is a lower bound of the true significance.

As estatísticas de teste são mostrados na terceira tabel. Aqui dois testes de normalidade são executados. Para conjunto de dados pequeno de 2000 elementos, usamos o teste de Shapiro-Wilk, caso contrário, o teste de Kolmogorov-Smirnov é usado. No nosso caso, uma vez que temos apenas 20 elementos, o teste de Shapiro-Wilk é usado. De A, o valor-p é 0,316. Podemos rejeitar a hipótese alternativa e concluem que os dados têm de uma distribuição normal.