

AULA 4 – Manipulação (preparação) de dados

A planilha abaixo contém as notas de duas turmas, dez para cada uma.

≽мп	NITAB - Un	titled - [W	orksheet 1	***]									Ð×
🏥 E	jile <u>E</u> dit Da	ata <u>⊂</u> alc g	<u>i</u> tat <u>G</u> raph	E <u>d</u> itor <u>T</u> o	ols <u>W</u> indow	<u>H</u> elp						E	٩×
🗃	86.	X 🖻 🛱	n ca	1	1 M A	🛇 ? 🗊	+🕻 🗖	h 🕞 🕕 💈	3 🗐 🔁	1 🗉 🖪			
	-2 🛶 🛱	20	0										
Ŧ	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	
	Nota	Turma											
1	2,6	1											
2	0,4	1											
3	9,9	1											
4	2,1	1											
5	3,0	1											
6	6,5	1											
7	1,9	1											
8	7,2	1											
9	1,9	1											
10	3,2	1											
11	3,8	2											
12	3,8	2											
13	6,4	2											
14	6,5	2											
15	6,4	2											
16	3,7	2											
17	5,5	2											
18	6,9	2											
_1q ∢	18	2											▼ ▶
Current	Worksheet: \	Worksheet 1										9:00	

Exemplo do uso do comando "Unstack"

O objetivo é separar as notas que estão armazenadas em apenas uma variável (Nota – coluna C1) em duas colunas, podendo ainda identificar a respectiva turma de cada nota em uma nova variável. As novas variáveis serão adicionadas na mesma planilha. O mesmo procedimento pode ser feito pondo as novas variáveis numa nova planilha.

≥мп	NITAB -	Unti	tled - [\	₩orks	heet 1	***]			
🏥 E	ile <u>E</u> dit	Dat	a <u>⊂</u> alc	<u>S</u> tat	<u>G</u> raph	Editor	<u>T</u> ools	<u>W</u> indow	<u>H</u> elp
🗃 🛛	86	4	Su <u>b</u> set	Works	heet		† ↓	m &	0 ? 1
	- <u>1</u> - <u>1</u> - <u>1</u>	a	Split Wo	orkshee	et	ŀ			
,	C1	*	Merge Worksheets				L I	C5	C6
	Nota		⊆ору			•			
1	2	→	Unstack	Colun	nns				
2	0		S <u>t</u> ack			•			
3	9	F	Tr <u>a</u> nspo	ose Col	umns				
4	2	≜ ↓	Sort						
5	3	312	Rank						
6	6	-	-			_			
7	1	¥.	Delete I	Rows	•				
8	7	Var	<u>E</u> rase V	ariable	s				
9	1		C <u>o</u> de						
10	3		C <u>h</u> ange	Data	Туре	•			
11	J 2		E <u>×</u> tract	from D)ate/Tim	e 🕨			
12	5	А+ В	Concate	enate.					
14	0 A	m	Display	Data					
15	6	4		2					

Você deverá ver a seguinte janela, onde serão escolhidas as colunas a serem desmembradas.



Σмп	NITAB - Un	titled - [Worksheet 1 ***]	
🏢 E	jile <u>E</u> dit Da	ata <u>C</u> alc <u>S</u> tat <u>G</u> raph E <u>d</u> itor	<u>I</u> ools <u>W</u> indow <u>H</u> elp
🚅		🖁 🖻 💼 🗠 🗠 🛄 🗠	t 🛛 🛤 🖧 🚫 💡 🗊 🛛 🔁 🖬 🐨 🖉 🐨 🗊 📰 🧏
	- I 🏎 🎜	1 A. 0	
Ŧ	C1	Unstack Columns	
	Nota	·	Unstack the data in: Nota
1	2,6		
2	0,4		
3	9,9		Using subscripts in: Turma
4	2,1		
5	3,0		Include missing as a subscript value
6	6,5		Store unstacked data:
7	1,9		C In new worksheet
8	7,2		Name: [Ontional]
9	1,9		(obuoun)
10	3,2		After last column in use
11	0,C 20		✓ Name the columns containing the unstacked data
12	0,C 6.4	,	
13	6,4	Palaat	
14	6,0	aerect	
15	3.7	Heln	
17	5,5		

Deverá aparecer a planilha original mais duas novas varáveis, como mostra a janela abaixo.

-													
1 🖽 E	jile <u>E</u> dit D _é	<u>a</u> ta <u>⊂</u> alc g	<u>i</u> tat <u>G</u> raph	E <u>d</u> itor <u>T</u> ool	s <u>W</u> indow	Help						E	Ð
🗃		አ 🖻 🖻	S CH	📴 î i	M &	0 ? 1] +C 🦷	G 🕕 🖻) 🗐 🍋 I	5 E			
_ 	-2	20 0	0										
,	C1	[] []		64		6		6	P)	C10	C11	C12	_
	Nota	Turma	Nota 1	Nota 2			0.			010	011	UIL	
1	2.6	1	2.6	3.8									
2	0,4	1	0,4	3.8									
3	9,9	1	9,9	6,4									-
4	2,1	1	2,1	6,5									
5	3,0	1	3,0	6,4									Т
6	6,5	1	6,5	3,7									
7	1,9	1	1,9	5,5									
8	7,2	1	7,2	6,9									
9	1,9	1	1,9	1,8									
10	3,2	1	3,2	8,5									
11	3,8	2											
12	3,8	2											
13	6,4	2											
14	6,5	2											
15	6,4	2											_
16	3,7	2											_
17	5,5	2											+
18	6,9	2											
•	18												Þ
Connect	And and a data sets 2											0.10	_

Observação: As notas na coluna C2 não precisam estar ordenadas, como no exemplo acima, para usar o procedimento "*Unstack*"



Exemplo do uso do comando "Stack".

O próximo passo é fazer o procedimento inverso, isto é, a partir de duas ou mais colunas criar (compor) uma única coluna de dados, podendo ainda identificar a procedência (turma) de cada nota.



Na janela abaixo, foram selecionadas as duas colunas que comporão a nova variável. Além disso, essa nova variável será adicionada na planilha corrente juntamente com uma variável para identificar procedência (turma). As duas novas variáveis serão armazenadas nas colunas C5 e C6 como mostrado abaixo.

🚬 МП	NITAB - Un	titled - [Worksheet 1 ***]	
🏥 E	jile <u>E</u> dit D <u>a</u>	ita <u>C</u> alc <u>S</u> tat <u>G</u> raph E <u>d</u> itor	Iools Window Help
😅		X 🖻 💼 🗠 🗠 🛄	† ↓ 👫 糸 〇 ? 🗗 七 📾 🗟 🛈 🗵 🗟 📬 独 🗷 📰 国 🚿
-2 •	-2 🛶 🕅	1 d. 0	
Ŧ	C1	Stack Columns	X
	Nota	Lot N.	Stack the following columns:
1	2,6	C1 Nota C2 Turma	Nota 1' Nota 2'
2	0,4	C3 Nota_1	
3	9,9	C4 NOta_2	
4	2,1		-
5	3,0		
6	6,5		Store stacked data in:
7	1,9		
8	7,2		N <u>a</u> me: [Optional]
9	1,9		Column of current worksheet:
10	3,2		
11	3,8		Store <u>s</u> ubscripts in: c6 (Optional)
12	3,8		🔽 Lles verieble names in subscript column
13	6,4		I ose variable names in subscript column
14	6,5	Select	
15	6,4		
16	3,7	Help	<u>O</u> K Cancel
17	5,5 .		
18	6,9	2	
19	18	2	
للعا			



O resultado é mostrado abaixo:

Se MI	NITAR . He	titled - Dute	urk sheet_1	xxx]			
	ile Edit Di	ata Calc S	tat Graph	Editor Too	le Window	Help	
		ala <u>C</u> ail 2	uau <u>G</u> raph				
] 🖻		እ 🖻 🛱	l to ca		レぬる	0 8 🛛	.
	-2 📩 🖻	1 × 4	0				
+	C1	C2	C3	C4	C5	C6-T	
	Nota	Turma	Nota_1	Nota_2			
1	2,6	1	2,6	3,8	2,6	Nota_1	
2	0,4	1	0,4	3,8	0,4	Nota_1	
3	9,9	1	9,9	6,4	9,9	Nota_1	
4	2,1	1	2,1	6,5	2,1	Nota_1	
5	3,0	1	3,0	6,4	3,0	Nota_1	
6	6,5	1	6,5	3,7	6,5	Nota_1	
7	1,9	1	1,9	5,5	1,9	Nota_1	
8	7,2	1	7,2	6,9	7,2	Nota_1	
9	1,9	1	1,9	1,8	1,9	Nota_1	
10	3,2	1	3,2	8,5	3,2	Nota_1	
11	3,8	2			3,8	Nota_2	
12	3,8	2			3,8	Nota_2	
13	6,4	2			6,4	Nota_2	
14	6,5	2			6,5	Nota_2	
15	6,4	2			6,4	Nota_2	
16	3,7	2			3,7	Nota_2	
17	5,5	2			5,5	Nota_2	
18	6,9	2			6,9	Nota_2	
19	1.8	2			18	Nota 2	
Current	Worksheet	Worksheet 1					
		worksheet i	-				



Exemplo do uso do diagrama de caixa (boxplot) para fazer um acompanhamento ao longo do tempo. Os dados abaixo se referem às temperaturas médias diárias para todos os meses. Para este exemplo, todos os meses têm 30 dias. Para tanto, vamos fazer um boxplot para cada mês (mostrando a média do respectivo mês) e vamos conectar as médias para ver a evolução da temperatura média mensal ao longo do ano.

O primeiro passo seria agrupar todas as varáveis numa mesma coluna usando o comando "*Stack*" juntamente com a variável para identificar os respectivos meses. Ver o procedimento "*Stack*" descrito anteriormente.

≽мп	NITAB - Un	titled - [Wo	orkshe	et 1 ***]								_ 8
🏢 e	jie <u>E</u> dit Da	ta <u>⊂</u> alc <u>S</u>	tat G	raph E <u>d</u> itor]	ools <u>W</u> indow	Help						그리
B		X 🖻 🖷	× 2	<u>S</u> catterplot		0 ? 1	+🕻 📾	G 🛈 🖻) 🗒 🍖 9	i E		
,		26 6		🗄 Matrix Plot		F						
			_ 2	🛃 Margi <u>n</u> al Plot			67	<u></u>	<u> </u>	C40	C44	C42
+	ian	CZ fov	h .	h Histogram		iun	iul	200	C9 cot	Cit	UT1 nov	doz
	301 Jan	21 2710		J Dotplot		Jun 21.0227	34 CO07	ayu 24 2104	32 6040	10 NOE0	72 0011	16Z -
	22,9709	37,0000		2 Stem-and-Le	əf	21,9227	21,0507	24,2104	23,0343	20,0000	23,0311	23,1070
2	30,4467	27,0990		29 Deckability Di		21,0027	22,2394	21,0007	23,3340	27,1905	27,3005	22,2000
3	26,3157	27,9605	2			22,0604	22,7212	23,7546	25,2053	25,3049	29,2564	30,4012
4	24,9508	25,7958	_2t [2	Empirical CDr	·	20,4472	20,0455	22,8244	24,5335	29,0486	28,3284	31,7953
5	28,3830	25,9848		<mark>≬∮</mark> <u>B</u> oxplot		22,9913	24,9245	25,1553	22,1796	27,3665	27,3953	29,1571
6	29,8250	29,4209	[1]	III Interval Plot		21,2917	22,2317	26,1285	26,6646	26,9382	26,1732	27,6564
7	26,8281	27,2580	-26	Individual Value Plot		22,4805	20,4520	23,7318	23,4169	31,2184	27,0525	27,9371
8	25,7135	27,8277	- 28			24,0291	20,7969	23,1243	23,0669	27,2322	26,1381	27,5950
9	28,6929	26,6997	2; []	Bar Chart		23,0197	22,3672	23,5107	21,3909	26,8034	25,2633	27,4546
10	27,2754	22,2977	29	Pie Chart		22,2917	22,6393	24,1955	27,6111	27,3984	23,8277	27,1429
11	24,6301	25,2504	27	Time Series F	lot	23,0097	22,7365	22,6808	25,8932	26,3736	27,0539	32,7299
12	25,5496	26,4058	2!	🔺 Area Graph		21,7482	21,6178	21,5495	23,9645	26,1425	25,4191	29,8903
13	27,6376	27,8712	28			23,1581	22,2353	24,5625	25,3434	26,9492	27,5470	31,1014
14	25,1625	28,1956	2; 🗖	🧧 <u>C</u> ontour Plot		20,6992	20,9318	22,91348	25,5120	26,2819	26,6079	29,7815
15	26,5656	26,4328	2; 3	📫 <u>3</u> D Scatterpl	ot	23,0192	21,5926	22,7629	23,4518	24,1708	27,3657	28,2261
16	24,2955	27,6394	2; 💐	3D Syrface P	lot	22,2568	21,9491	25,7899	24,4981	27,1498	24,7931	24,9409
17	29,7571	26,9341	25,2	22,958	7 22,6896	22,1037	21,4975	23,1931	21,8605	26,8166	25,5909	27,3515
18	24,3948	26,0209	27,1	726 24,888	2 23,5604	22,1410	20,1390	23,0458	23,8866	25,4965	28,6429	27,4671
_19 ∢	27.7695	27 8145	26.7	726 24 216	6 23.9115	22 6467	23,5150	23.6408	22 8120	25 2017	25 9676	26 2112

O próximo passo é criar o boxplot, mostrado a seguir:

Selecionar a opção "With Groups", destacada em preto.

.sheet 1	***]						
: <u>G</u> raph	E <u>d</u> itor <u>T</u> ools	<u>W</u> indow	<u>H</u> elp				
ဂ လ	🖪 † 🖡	M #	0 ? 1] +(] 🖷	i 🕞 🕕 💈) 🗒 🍡	i I E
2							
СЗ	C4	C5	C6	C7	C8	C9	C10
mar	Boxplots					×	out
28,2802							26,08
26,5118	Une Y Circela	No. 201-1	C				27,19
29,9554	Simple	With	Groups				25,38
26,5073							29,04
26,7765			1 부 Y				27,36
25,8774		- B 1 2 A 1	1 2				25,93
26,4402							31,21
28,1925	- Multiple Y's Simple	No. 201- 1	C				27,23
27,6507			Groups				26,80
29,1008	日本古	L L L					27,39
27,8752							26,37:
25,1085	Ý1 Ý2		1 2 Y2				26,14:
28,2902							26,94
27,2785	Help	b	<u>0</u>	ĸ	Can	cel j	26,28
27,5745	26,7648	22,6641	23,0192	21,5926	22,7629	23,4518	24,17
27 6486	25,9842	24 3882	22 2568	21 9491	25 7899	24 4981	27 14



As varáveis criadas foram armazenas nas colunas C13 (temperaturas) e C14 (meses). Selecionar a opção "Data View", destacada em vermelho abaixo.

MINITAB - Untitled - [Worksheet 1 ***]													
E 📰 E	jile <u>E</u> dit Da	ata <u>⊂</u> alc g	itat <u>G</u> raph	E <u>d</u> itor <u>T</u> ool	ls <u>W</u> indow	Help							
- 		X 🖻 🖻	l no ce l		A &	0 ? 1	+🕄 🔚	🗟 🛈 🆻) 🗐 🍖 🤅	ð 🔳 📰			
	▝▋▖▎⋒	20 12	0										
+	C1	C2	C3	C4	យ	C6	C7	C8	C9	C10	C11		
	jan	fev Bo	oxplot - One	Y, With G	roups					>	nov		
1	22,9789	31,37	C1 ian		<u>G</u> raph ·	variables:					23,8911		
2	30,4467	27,09	2 fev		C13					A	27,3665		
3	26,3157	27,96	C4 abr								29,2564		
4	24,9508	25,79	C5 mai C6 iun							V	28,3284		
5	28,3830	25,98	C7 jul		Catego	rical variable:	s for grouping	g (1-4, outerm	iost first):		27,3953		
6	29,8250	29,42	28 ago C9 set		C14					<u> </u>	26,1732		
7	26,8281	27,25	C10 out							T	27,0525		
8	25,7135	27,82	C12 dez		1					_	26,1381		
9	28,6929	26,69	213 C14								25,2633		
10	27,2754	22,29			Sc	ale	Lat	oels	Data	View	23,8277		
11	24,6301	25,25									27,0539		
12	25,5496	26,40	Sel	ect	<u>M</u> ultiple	Graphs	D <u>a</u> ta C)ptions			25,4191		
13	27,6376	27 ,87									27,5470		
14	25,1625	28,19	Help				<u>(</u>	<u>o</u> k	Ca	ncel	26,6079		
15	26,5656	26,43									27,3657 [:		
16	24,2955	27,6394	27,6486	25,9842	24,3882	22,2568	21,9491	25,7899	24,4981	27,1498	24,7931		
17	29,7571	26,9341	25,2057	22,9587	22,6896	22,1037	21,4975	23,1931	21,8605	26,8166	25,5909		
18	24,3948	26,0209	27,1726	24,8882	23,5604	22,1410	20,1390	23,0458	23,8866	25,4965	28,6429		
19	27.7695	27 8145	26 7726	24 2166	23,9115	22 6467	23,5150	23 6408	22 8120	25 2017	25.9676		

Aparecerá a seguinte janela:

2	Boxplot -	Data View		<u>×</u> 10	1	Ţ
Box	Data Dis	splay			×	
7 19 16 19 18 12 15 12 19 19 15 10 17 19		Select	Data Display Median <u>c</u> onfidence interval box Interquartile range <u>b</u> ox Bange box Ogtlier symbols Individual symbols Median symbol Megian connect line Mgan symbot Megan connect line Lajegorical variables for attribute assignment:			
3 394	27н	elp	<u>O</u> K Cancel			₽7 24

Selecionar as duas últimas opções, destacadas acima em vermelho.



Após a seleção clicar em "OK".



Deverá aparecer a seguinte janela de gráfico:

