

Cap.6 – Análise de dados em Excel

Estatística Aplicada

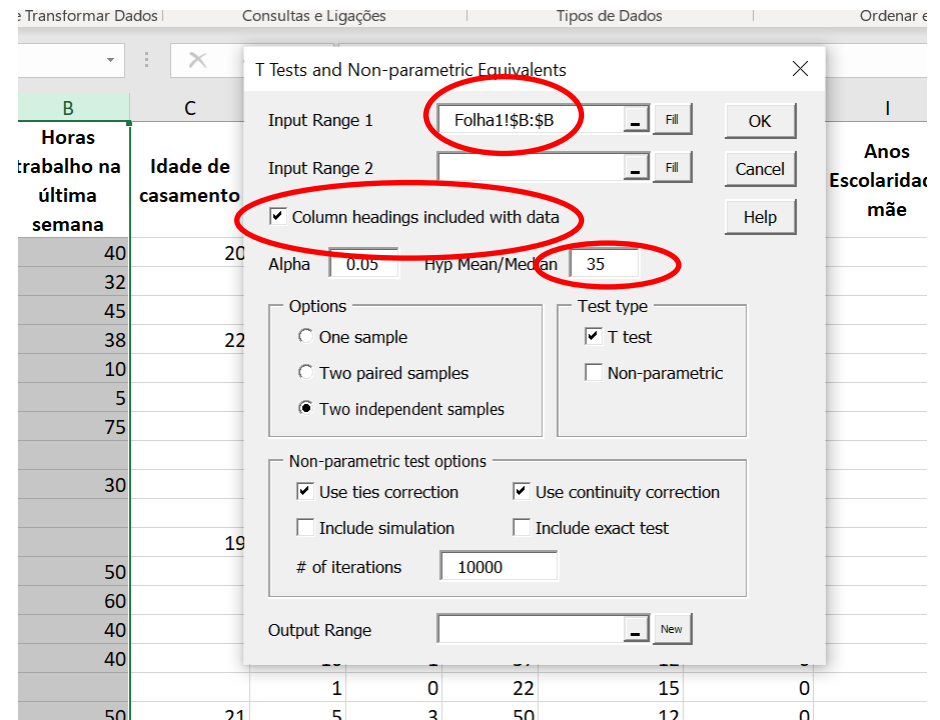
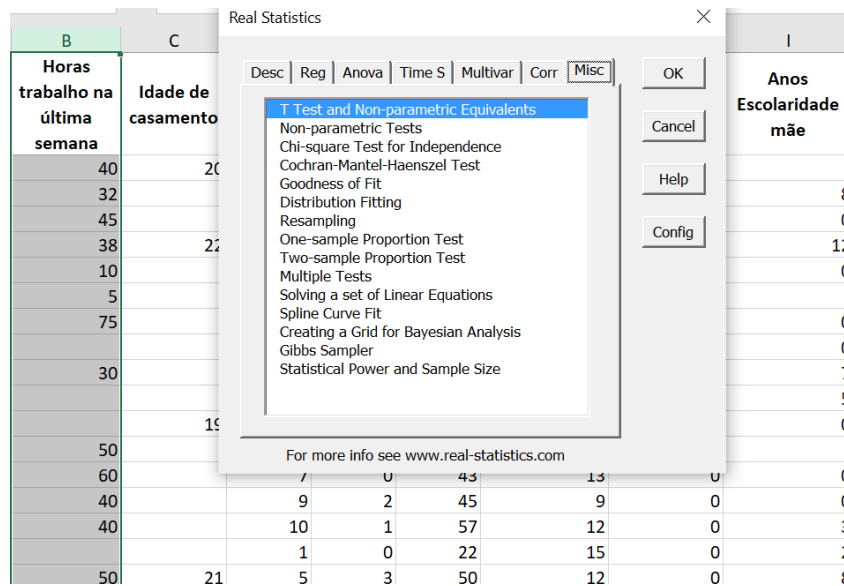
Curso Técnico Superior Profissional em Desenvolvimento
Web e Multimédia

Análise de Dados: Inferência Estatística

Análise de dados com Excel

Uma amostra

No ficheiro *Base de dados empresa XYZ Lda*, verificar se a empresa o número de horas de trabalho na última semana difere de 35 horas.



Análise de dados com Excel

T Test: One Sample							
SUMMARY			Alpha	0,05			
Count	Mean	Std Dev	Std Err	t	df	Cohen d	Effect r
937	41,49839915	14,30648	0,467372	13,90411	936	0,454228	0,413746
T TEST			Hyp Mean	35			
	p-value	t-crit	lower	upper	sig		
One Tail	2,15854E-40	1,646483			yes		
Two Tail	4,31707E-40	1,962502	40,58118	42,41562	yes		

média amostral →

Valor prova →

Intervalo de confiança →

$H_0: \mu = 35$

$H_1: \mu \neq 35$

$p \approx 0$

$\alpha = 0,05$

Como $p < \alpha$, rejeita-se H_0 e conclui-se que o número médio de horas semanais da empresa difere significativamente de 35.

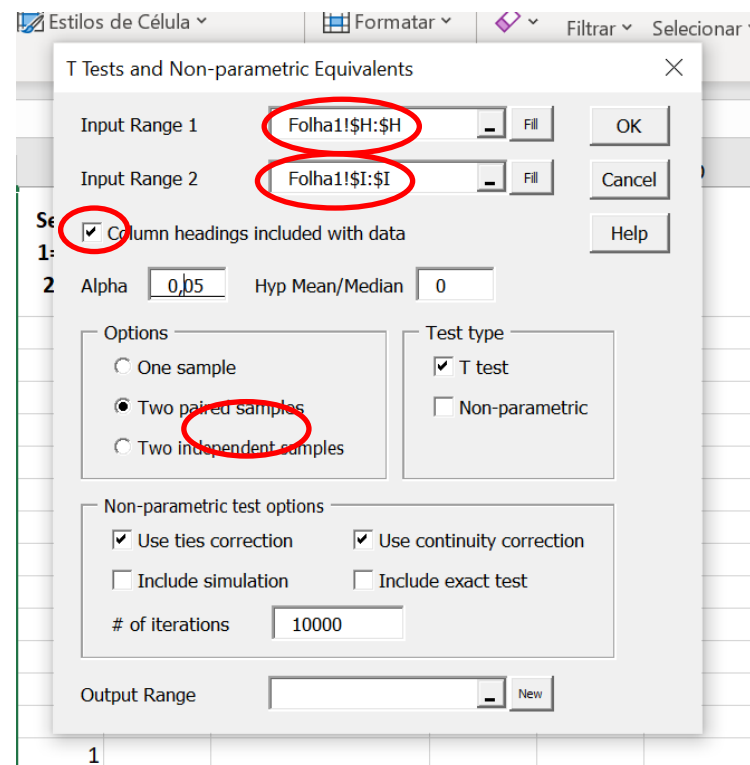
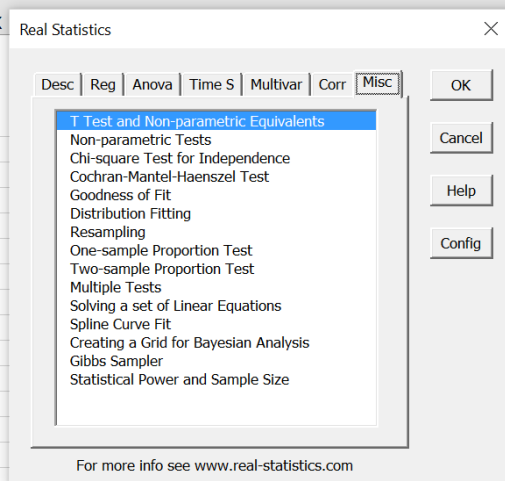
Pela análise da média amostral (> 35), tem-se que o número médio de horas semanais da empresa é significativamente superior a 35 horas.

Análise de dados com Excel

Teste t para duas amostras emparelhadas

No ficheiro *Base de dados empresa XYZ Lda*, comparar, através de um teste de hipóteses adequado, os *anos de escolaridade do pai e da mãe dos trabalhadores*.

	H	I	J	K
	Anos Escolaridade pai	Anos Escolaridade mãe	Sexo 1=M 2=F	
12			1	
8	0	8	1	
12	0	0	1	
16	0	12	1	
16	0	0	1	
16			1	
12	0	0	1	
16	0	0	1	
12	0	7	1	
13	0	5	1	
12	0	0	1	
16			1	
13	0	0	1	
9	0	0	1	
12	0	3	1	
15	0	2	1	
12	0	8	1	



Análise de dados com Excel

T Test: Two Paired Samples								
SUMMARY			Alpha	0,05	Hyp Mean		0	
Groups	Count	Mean	Std Dev	Std Err	t	df	Cohen d	Effect r
Anos Escolaridade pai	980	11,17347	4,249678					
Anos Escolaridade mãe	980	11,16837	3,505609					
Difference	980	0,005102	3,29127	0,105136	0,048528	979	0,00155	0,001551
T TEST								
	p-value	t-crit	lower	upper	sig			
One Tail	0,480653	1,646412			no			
Two Tail	0,961305	1,96239	-0,20122	0,211419	no			

médias amostrais

Valor prova

Intervalo de confiança

$$H_0: \mu_d = 0$$

$$H_1: \mu_d \neq 0$$

$$p=0,961 > \alpha=0,05$$

Não se rejeita H_0 pelo que não se pode concluir que existem diferenças significativas entre os anos de escolaridade do pai e da mãe destes trabalhadores.

Análise de dados com Excel

Exercício:

Comparar, através de um teste de hipóteses adequado, os o *número de filhos* e o *número de irmãos* dos trabalhadores.

Análise de dados com Excel

Teste t para duas amostras independentes

No caso do **teste t para duas amostras independentes** pretende-se verificar se existe uma diferença entre as médias de duas amostras.

Existe para este teste um pressuposto – homogeneidade das variâncias – ultrapassado no caso de as amostras terem dimensões aproximadamente iguais. No caso de não serem iguais, este pressuposto deve ser testado .

Análise de dados com Excel

Comparar as *horas de trabalho na última semana* entre as trabalhadoras e trabalhadores

Passo 1) Teste à homogeneidade das variâncias

H0: As variâncias são homogêneas

H1: As variâncias não são homogêneas

Análise de dados com Excel

Passo 1) Teste à homogeneidade das variâncias

H0: As variâncias são homogêneas

H1: As variâncias não são homogêneas

Ver Ajuda Power Pivot Diga-me o que pretende fazer Partilhar Comentários

Formatação Condicional Formatar como Tabela Estilos de Célula

Inserir Eliminar Formatar

Ordenar e Filtrar Selecionar

Ordenar do Mais Pequeno ao Maior

Ordenar do Maior ao Mais Pequeno

Ordenar do Mais Pequeno ao Maior

Do mais baixo para o mais elevado.

Mais informações

Reaplicar

I	J	K	L	M	N
Anos	Sexo				
Escolaridade	1=M				
mãe	2=F				
14	2				
	1				
8	1				
12	2				
12	2				
0	1				
12	2				
12	2				
12	2				
8	2				
12	1				
12	2				
0	1				
	1				
0	1				
0	1				

Aviso de ordenação

O Microsoft Excel localizou dados junto à seleção. Como não os selecionou, estes dados não serão ordenados.

O que pretende fazer?

☒ Expandir a seleção

☐ Continuar com a seleção atual

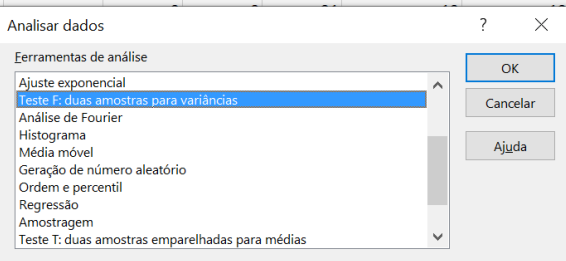
Ordenar Cancelar

40	20	4	3	59	12
32		6	3	59	8
20		3	0	21	13
35					
45					
40					
40					
65					
38					
10					
5					
75					
		1	0	53	12
		1	0	65	16

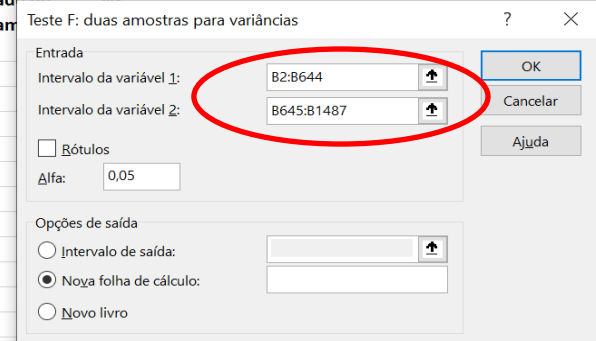
Análise de dados com Excel



Idade da mãe	casamento	irmãos	Idade da mãe	Idade da mãe	Escolaridade da mãe	Escolaridade da mãe	Escolaridade da mãe
40	20	4	2	33	12	12	12
32		4	3	59	8	0	0
20		6	3	59			



ID	Horas trabalho na última semana	Idade da mãe	Anos de escolaridade da mãe	Anos de escolaridade da mãe	Anos de escolaridade da mãe
1	40				
2	32				
3	45				
4	38				
5	10				
6	5				
7	75				
8	30				
9	50				
10	60				
11	40				
12					
13					
14					
15					



$$p=0,005 < \alpha=0,05$$

Rejeita-se H_0 pelo que se pode concluir que existe um grupo com variância significativamente superior à do outro. Logo as variâncias **não** são homogêneas.

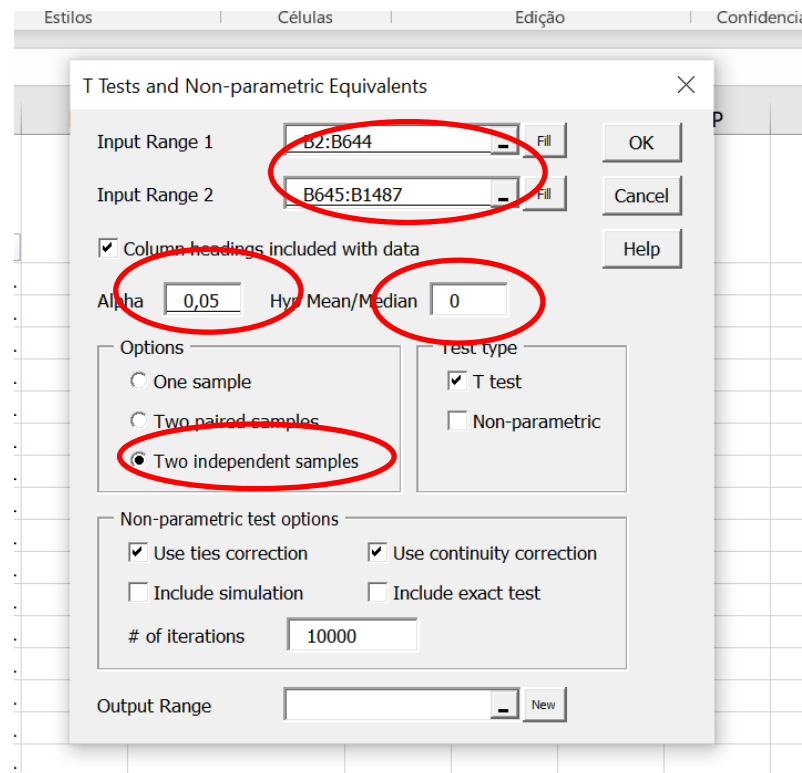
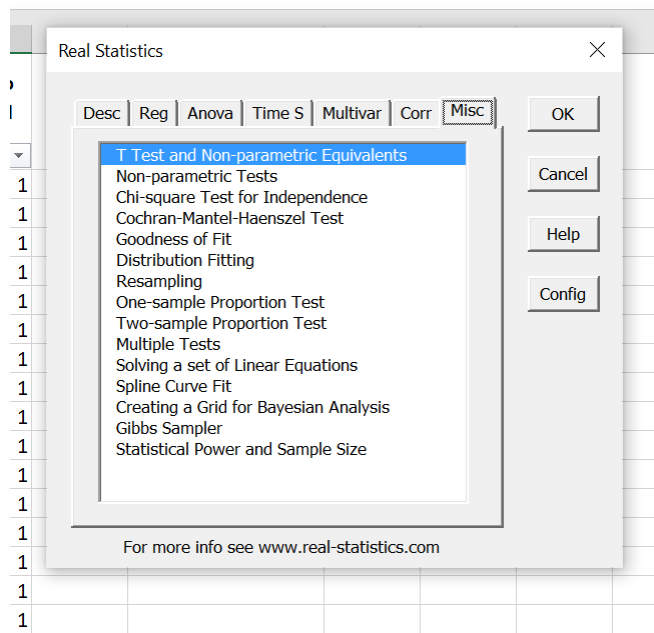
Teste F: duas amostras para variâncias		
	Variável 1	Variável 2
Média	43,09251	40
Variância	227,1614	179,3237
Observações	454	483
gl	453	482
F	1,266768	
P(F<=f) uni-caudal	0,00531	
F crítico uni-caudal	1,164356	

Análise de dados com Excel

Passo 2) Teste t para amostras independentes

$$H_0: \mu_1 = \mu_2$$

$$H_1: \mu_1 \neq \mu_2$$



Análise de dados com Excel

médias amostrais

Grupo 1 (M)

Grupo 2 (F)

T Test: Two Independent Samples									
SUMMARY				Hyp Mean	0				
Groups	Count	Mean	Variance	Cohen d					
40	453	43,09934	227,6428						
0	483	40	179,3237						
Pooled			202,7072	0,217688					
T TEST: Equal Variances				Alpha	0,05				
	std err	t-stat	df	p-value	t-crit	lower	upper	sig	effect r
One Tail	0,931215	3,328274	934	0,000454	1,646487			yes	0,108264
Two Tail	0,931215	3,328274	934	0,000908	1,962507	1,271822	4,926853	yes	0,108264
T TEST: Unequal Variances				Alpha	0,05				
	std err	t-stat	df	p-value	t-crit	lower	upper	sig	effect r
One Tail	0,934769	3,315619	903,919	0,000475	1,646541			yes	0,109616
Two Tail	0,934769	3,315619	903,919	0,000951	1,962592	1,264768	4,933908	yes	0,109616

Valor prova

Intervalo de confiança

$$p=0,00096 < \alpha=0,05$$

Rejeita-se H_0 pelo que se pode concluir que os grupos apresentam valores médios significativamente diferentes.

Pela análise das médias amostrais, conclui-se que o 1º grupo (masculino) apresenta média de horas de trabalho na última semana significativamente superior ao 2º grupo (feminino).

Análise de dados com Excel

Exercício:

Comparar, através de um teste de hipóteses adequado, a idade entre trabalhadoras e trabalhadores.