

Estatística Aplicada
Prof. Simões

Testes com um fator

ANOVA

$$\mu_1 = \mu_2 = \mu_3 = \mu_4 ?$$

Conceituação

- O teste ANOVA permite avaliar se as diferenças observadas entre as médias de três ou mais grupos são estatisticamente significantes.



MÉDIA DA NOTA
DOS ALUNOS

784,5



MÉDIA DA NOTA
DOS ALUNOS

832,4



MÉDIA DA NOTA
DOS ALUNOS

804,2

Conceituação

- As condições para a aplicação do teste são
 - As amostras devem ser aleatórias e independentes
 - A distribuição deve ser normal ou aproximadamente normal
 - A variância de cada população deve ser a mesma
- É baseada em duas avaliações:
 - A variabilidade “**entre**”, s_b , “b” de “between”: é a avaliação matemática da diferença entre as amostras avaliadas
 - A variabilidade “**dentro**”, s_w , “w” de “within”: é a avaliação matemática das diferenças dentro de cada amostra avaliada

Procedimento: teste de hipótese

1. Estabelecer as hipóteses

$$H_0: \mu_1 = \mu_2 = \mu_3, etc$$

H_1 : No mínimo uma amostra é diferente

2. Estabelecer o valor crítico

- É obtido usando a tabela F, que utiliza 2 graus de liberdade e o valor da significância:

$$Gl_{num.} = k - 1$$

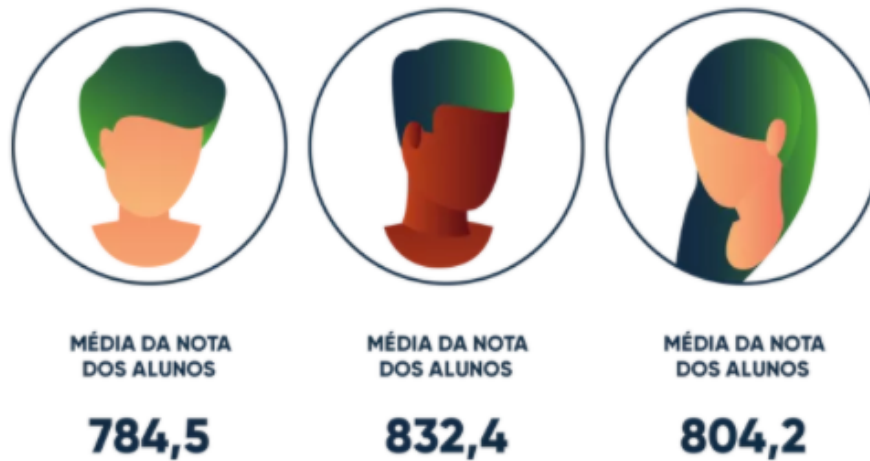
$$Gl_{den.} = N - k$$

k = número de amostras

$N = n_1 + n_2 + n_3 + \dots$ (soma dos tamanhos das amostras)

Procedimento: teste de hipótese

- Estabelecer o valor crítico (cont.)
 - Por exemplo, suponha que as turmas avaliadas tenham 14, 13 e 16 alunos respectivamente, e a significância desejada seja de 5%.



g.l.D: Graus de liberdade, denominador	$Gl_{num.}; \alpha = 0,05$		
	1	2	3
1	161,4	199,5	215,7
2	18,51	19,00	19,16
3	10,13	9,55	9,28
4	7,71	6,94	6,59
5	6,61	5,79	5,41
6	5,99	5,14	4,76
29	4,18	3,33	2,93
30	4,17	3,32	2,92
40	4,06	3,23	2,84
60	4,00	3,15	2,76

$$Gl_{num.} = k - 1 = 3 - 1 = 2$$

$$Gl_{den.} = N - k = (14 + 13 + 16) - 3 = 40$$

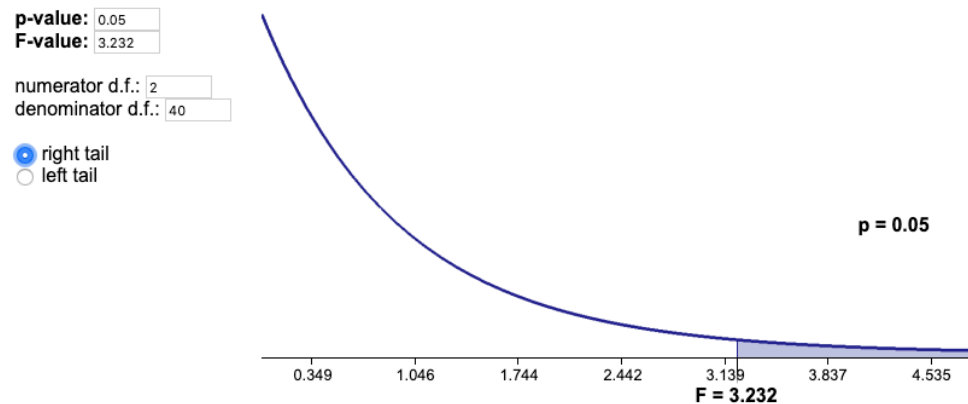
$$\alpha = 0,05$$

$$F_{\text{crítico}} = 3,23$$

Tabela F

- O uso de sites ou de planilhas pode ser conveniente para a determinação de valores de F

- Exemplo de site



<http://www.statdistributions.com/f?p=0.05&df1=2&df2=40>

- Exemplo no Excel

	B	C
34	Prob=>	0,05
35	Gl1=>	2
36	Gl2=>	40
37	F	=INV.F.CD(C34;C35;C36)

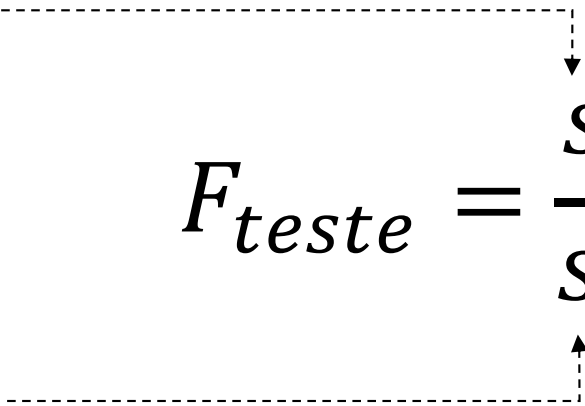
Comando

	B	C
34	Prob=>	0,05
35	Gl1=>	2
36	Gl2=>	40
37	F	3,232

Resultado

Procedimento: teste de hipótese

3. Estabelecer o valor de teste

$$s_b^2 = \frac{\sum n_i \cdot (\bar{x}_i - \bar{\bar{x}})^2}{k - 1}$$
$$F_{teste} = \frac{s_b^2}{s_w^2}$$
$$s_w^2 = \frac{\sum (n_i - 1) \cdot s_i^2}{N - k}$$


$s_b^2 \Rightarrow$ variância entre

$s_w^2 \Rightarrow$ variância dentro

$n_i \Rightarrow$ número de elementos de cada amostra

$\bar{x}_i \Rightarrow$ média de cada amostra

$\bar{\bar{x}} \Rightarrow$ média de todos os elementos da amostra

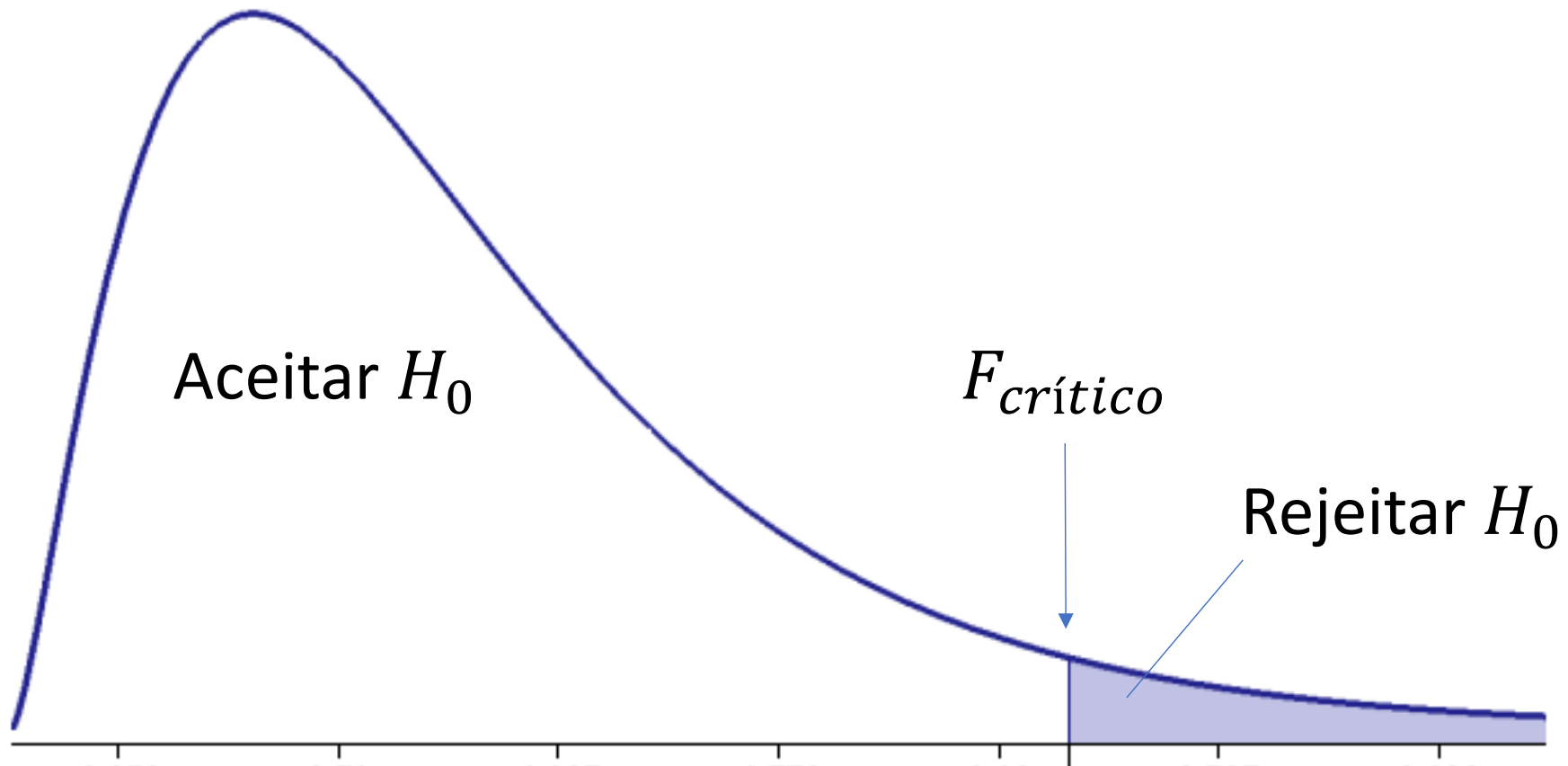
$k \Rightarrow$ número de amostras

$N \Rightarrow$ número total de elementos

Procedimento: teste de hipótese

4. Fazer a comparação

- Assim com nos demais testes, o valor do F_{teste} será comparado com o valor do $F_{crítico}$, o que definirá se a H_0 deve ou não ser aceita, ou seja, se as médias entre os grupos são ou não estatisticamente iguais



Exemplo

- Três catalisadores estão sendo comparados para aplicação em um processo, quanto ao tempo de reação. Foram feitos ensaios, obtendo-se os valores indicados na tabela abaixo, em minutos. Determine, com uma significância de 5% se há ou não diferença estatisticamente significativa entre os três.
 - Resolução manual
 - Resolução usando Excel

Catalisador 1	Catalisador 2	Catalisador 3
10	6	14
8	9	13
5	8	10
12	13	17
14		16
11		



Resolução manual

1. Hipóteses

$$H_0: \mu_1 = \mu_2 = \mu_3$$

H_1 : No mínimo uma amostra é diferente

2. Valor crítico

$$Gl_{\text{num}} = 3 - 1 = 2$$

$$Gl_{\text{den}} = N - k = 15 - 3 = 12$$

$$\alpha = 5\% = 0,05$$

$$F_c = 3,89$$

$\alpha = 0,05$			
g.l.N: Graus de liberdade, numerador			
g.l.D: Graus de liberdade, denominador			
	1	2	3
1	161,4	199,5	215,7
2	18,51	19,00	19,16
3	10,13	9,55	9,28
4	7,71	6,94	6,59
5	6,61	5,79	5,41
6	5,99	5,14	4,76
7	5,59	4,74	4,35
8	5,32	4,46	4,07
9	5,12	4,26	3,86
10	4,96	4,10	3,71
11	4,84	3,98	3,59
12	4,75	3,89	3,49
13	4,67	3,81	3,41

Resolução manual

3. Valor de teste

$$F = \frac{s_b^2}{s_w^2}$$

	Catalisador 1	Catalisador 2	Catalisador 3	
	10	6	14	
	8	9	13	
	5	8	10	
	12	13	17	
	14		16	
	11			
n	6	4	5	
Média	10,00	9,00	14,00	
Variância	10,00	8,67	7,50	
Média geral				11,07

$$s_b^2 = \frac{\sum n_i \cdot (\bar{x}_i - \bar{\bar{x}})^2}{k - 1}$$

$$s_b^2 = \frac{6 \cdot (10 - 11,07)^2 + 4 \cdot (9 - 11,07)^2 + 5 \cdot (14 - 11,07)^2}{3 - 1}$$

$$s_b^2 = 33,47$$

Resolução manual

- Valor de teste (cont.)

$$s_w^2 = \frac{\sum (n_i - 1) \cdot s_i^2}{N - k}$$

	Catalisador 1	Catalisador 2	Catalisador 3	
	10	6	14	
	8	9	13	
	5	8	10	
	12	13	17	
	14		16	
	11			
n	6	4	5	
Média	10,00	9,00	14,00	
Variância	10,00	8,67	7,50	
Média geral				11,07

$$s_w^2 = \frac{(6 - 1) \cdot 10 + (4 - 1) \cdot 8,67 + (5 - 1) \cdot 7,5}{15 - 3}$$

$$s_w^2 = 8,83$$

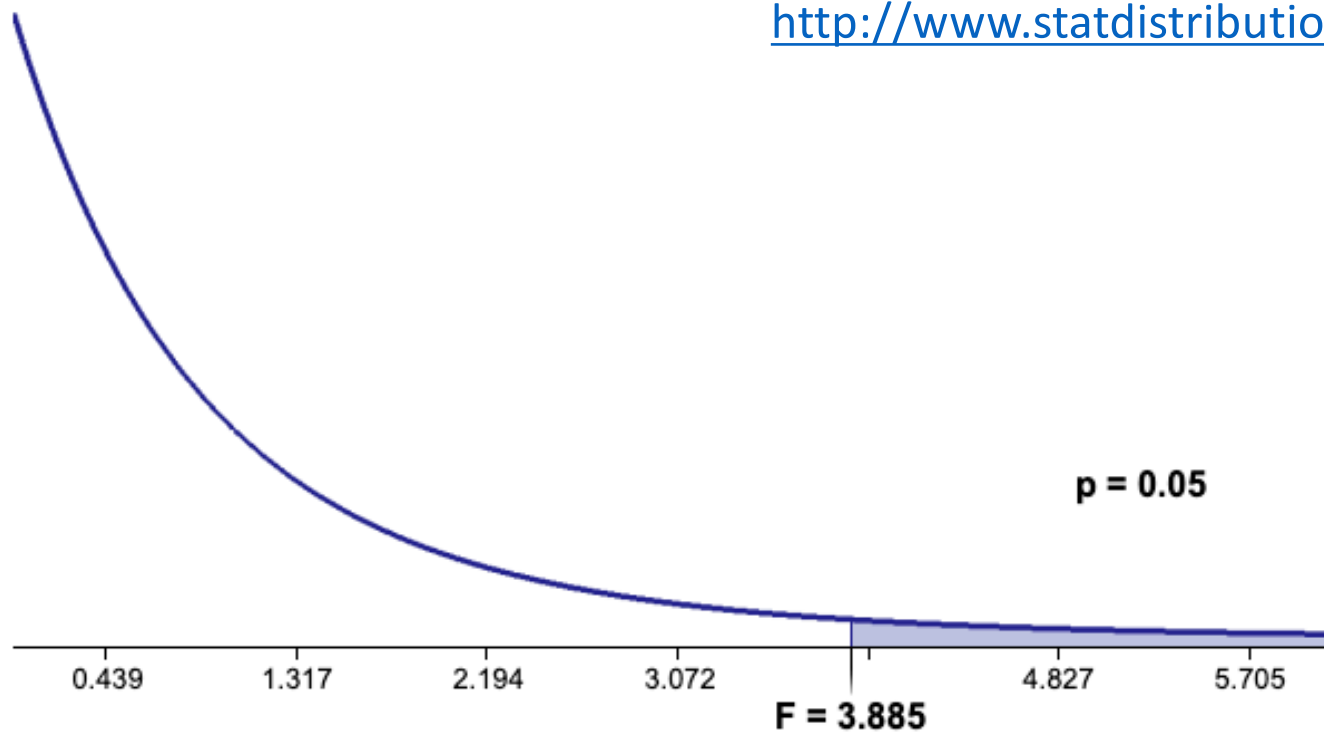
$$F = \frac{s_b^2}{s_w^2} = \frac{33,47}{8,83} = 3,79$$

Resolução manual

- Conclusão

F_{teste} está na região de aceitação de H_0 . Portanto, as médias podem ser consideradas estatisticamente iguais

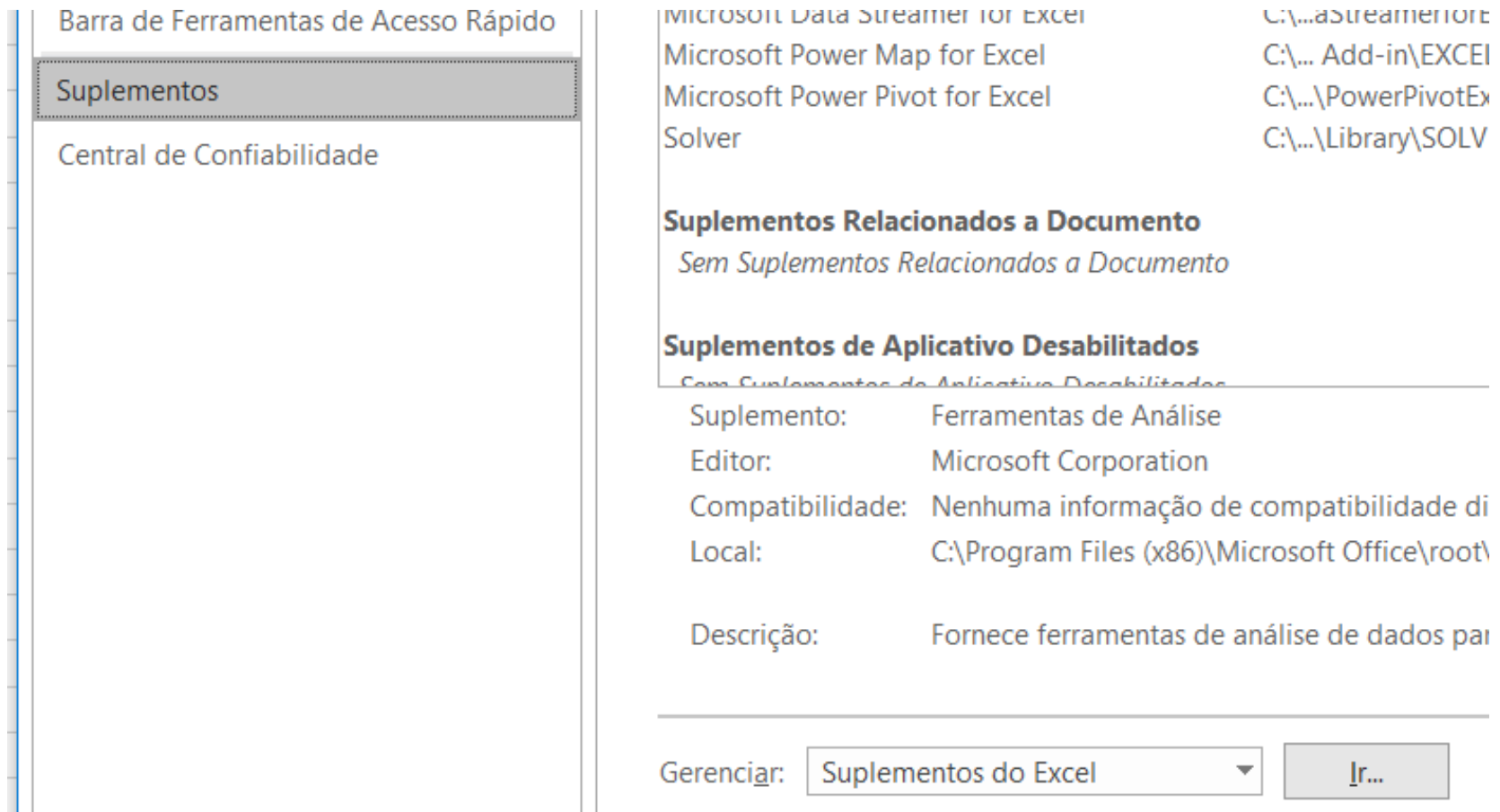
<http://www.statdistributions.com/f/>



Resolução usando Excel

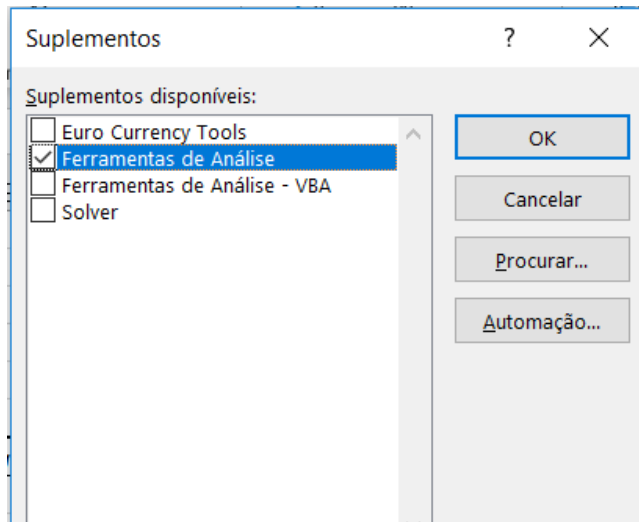
1. Habilitar as ferramentas estatísticas (se necessário)

- Ir em **Arquivo > Opções > Suplementos**, e na janela **Gerenciar: Suplementos do Excel**, clicar em **Ir...**

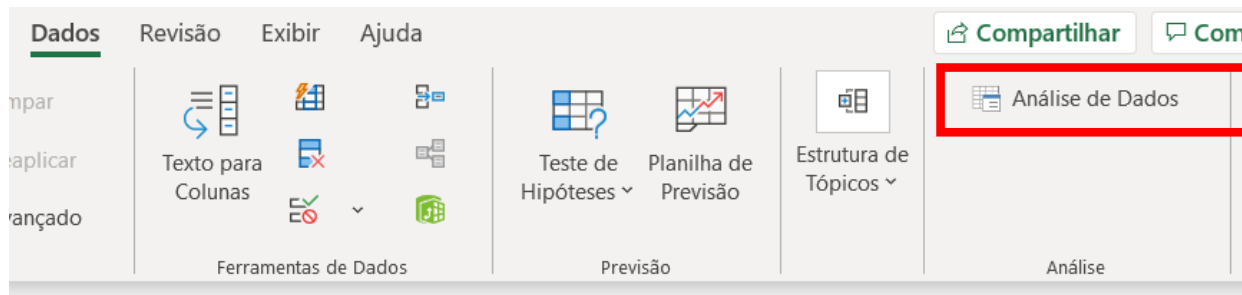


Resolução usando Excel

- Na tela que surgirá, escolha **Ferramentas de Análise**, e clique em **OK**

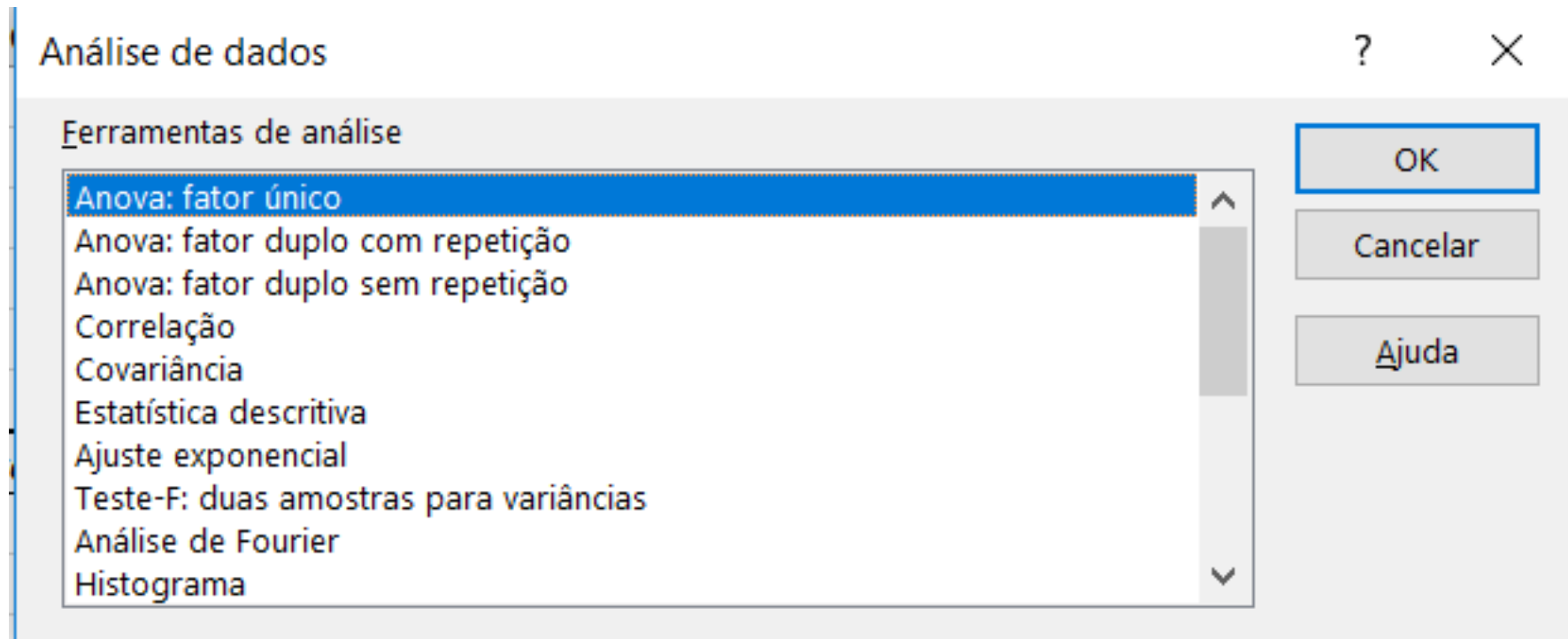


- De volta à planilha, no menu **Dados** da barra de menus, haverá agora o botão **Análise de Dados**. Clicar nele.



Resolução usando Excel

4. Na janela que surgirá, escolher a ferramenta de análise
Anova: fator único



Obs.: vale observar a grande variedade de funções estatísticas no Excel, muitas delas também presentes no Google Spreadsheets

Resolução usando Excel

5. Na janela que surgirá, informar as colunas com os dados, a área de saída e a significância. Clique em **OK**

	A	B	C	D
1				
2				
3				
4		10	6	14
5		8	9	13
6		5	8	10
7		12	13	17
8		14		16
9		11		

Anova: fator único

Entrada

Intervalo de entrada:

Agrupado por:

☒ Colunas

☐ Linhas

☐ Rótulos na primeira linha

Alfa:

Opções de saída

☒ Intervalo de saída:

☐ Nova planilha:

☐ Nova pasta de trabalho

OK

Cancelar

Ajuda

Resolução usando Excel

6. Os resultados serão informados

RESUMO

<i>Grupo</i>	<i>Contagem</i>	<i>Soma</i>	<i>Média</i>	<i>Variância</i>
Catalisador 1	6	60	10	10
Catalisador 2	4	36	9	8,67
Catalisador 3	5	70	14	7,5

ANOVA

<i>Fonte da variação</i>	<i>SQ</i>	<i>gl</i>	<i>MQ</i>	<i>F</i>	<i>valor-P</i>	<i>F crítico</i>
Entre grupos	66,9	2,0	33,5	3,79	0,053	3,885
Dentro dos grupos	106,0	12,0	8,83			
Total	172,9	14,0				

- Observe que é possível identificar as colunas e agregar várias outras funções como média, variância, etc, que podem ser úteis em um estudo/relatório mais completo.

Exercício

- Três equipes foram treinadas para a montagem de um produto utilizando métodos diferentes. Foram feitas montagens de teste para avaliar se há ou não diferença na eficiência entre os métodos, e os resultados estão abaixo. Determine se há ou não diferença significativa, usando uma significância de 1%

Método 1	Método 2	Método 3
12	16	14
15	14	17
17	21	20
12	15	15
	19	



Resultados

- Conclusão: não há evidência para concluir que a eficiência dos métodos são diferentes, já que o valor de teste está na região de aceitação da hipótese nula

RESUMO

<i>Grupo</i>	<i>Contagem</i>	<i>Soma</i>	<i>Média</i>	<i>Variância</i>
Método 1	4	56	14	6
Método 2	5	85	17	8,5
Método 3	4	66	16,5	7

ANOVA

<i>Fonte da variação</i>	<i>SQ</i>	<i>gl</i>	<i>MQ</i>	<i>F</i>	<i>valor-P</i>	<i>F crítico</i>
Entre grupos	21,92	2	10,96	1,50	0,269	7,56
Dentro dos grupos	73	10	7,3			
Total	94,9	12				

$$1,50 < 7,56$$