PRO2514 - Pesquisa Quantitativa em Gestão de Operações

Análise de Clusters (Agrupamentos ou Conglomerados)

Prof. Dr. Renato de Oliveira Moraes

Sumário

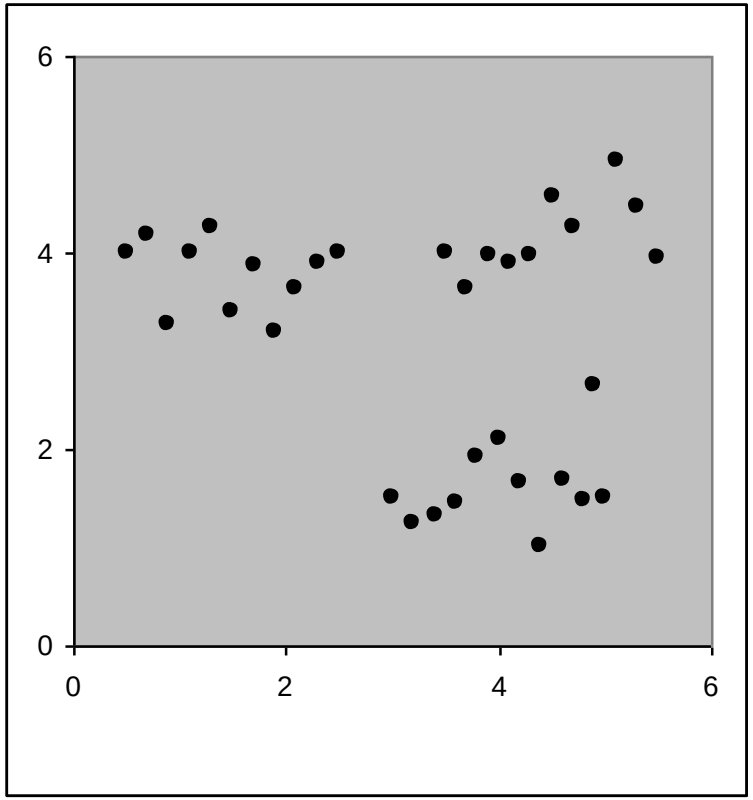
- Conceito geral (homogeneidade interna e heterogeneidade externa)
- Métodos de agrupamento: hierárquico x não hierárquico
- Medidas de (dis)similaridade
- Efeito da escala e padronização das variáveis
- Seleção das variáveis e sua influência no resultado final
- Quantidade de grupos formados
- Análise de Variância
- Significado dos grupos
- Validação dos resultados

Conceito geral

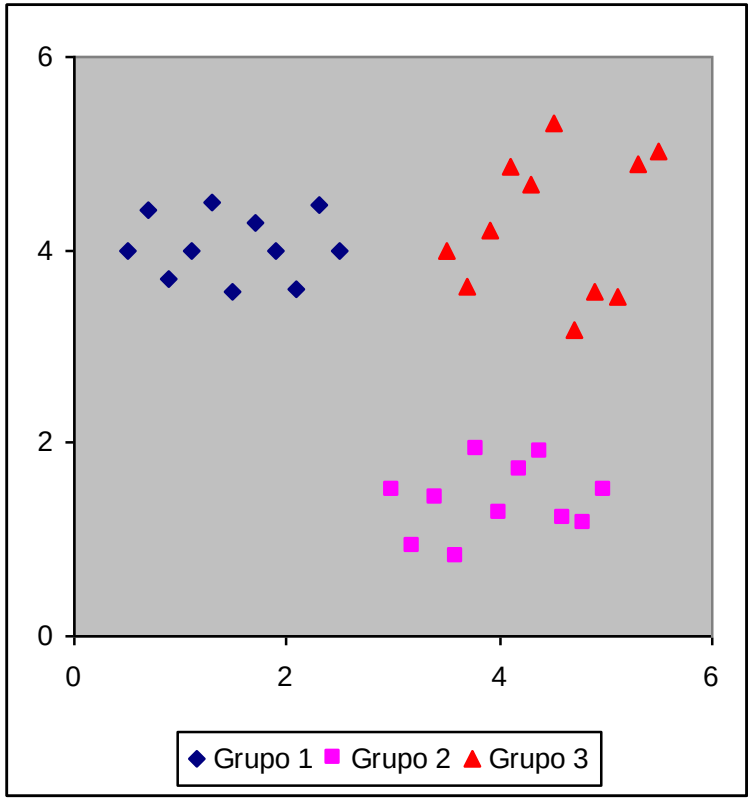
Divide a população em sub-populações que possuem características homogêneas dentro dos clusters e heterogêneas entre clusters, ou seja:

- Dentro do grupo (cluster) a variância é mínima;
- Entre grupos (clusters) a variância é máxima.

Análise de Conglomerados (*Clusters*)



Análise de Conglomerados (*Clusters*)



Métodos de agrupamento

- Hierárquico Aglomerativo:
  - Inicia-se com N grupos, cada grupo com um elemento;
  - A cada etapa, os dois grupos mais semelhantes são unidos. Ou seja, a cada etapa a quantidade de grupos se reduz em 1;
  - Em nenhum momento, os elementos de um grupo são separados.
- Hierárquico Divisivo:
  - Inicia-se com um único grupo com todos elementos;
  - A cada etapa, o grupo com menor homogeneidade interna é divido em dois grupos.
  - Ou seja, a cada etapa a quantidade de grupos aumenta em 1.
- Não hierárquico
  - Há uma definição prévia da quantidade (k) de grupos a serem formados;
  - Inicialmente os N elementos são divididos nos K grupos;
  - A cada etapa alguns elementos trocam de grupo de forma a maximizar a heterogeneidade entre os grupos.

Métodos Hierárquicos Aglomerativos

- Single linkage (SL) ou (vizinhos mais próximos): agrupa-se as pessoas com a distância mínima
- Complete linkage (CL) ou (vizinhos mais distantes): agrupa-se as pessoas com a distância máxima
- Average linkage (vizinhos comuns): é intermediária às duas anteriores (single e complete linkage), trabalha com valores ponderados das distâncias.
- Método da Centróide: distância entre dois clusters é a distância entre os centróides dos grupos
- Método Ward (mais usado): combina os indivíduos dentro dos clusters de acordo com o critério do menor incremento de soma total da distância euclidiana ao quadrado dentro do cluster

Sugestão de procedimento exploratório

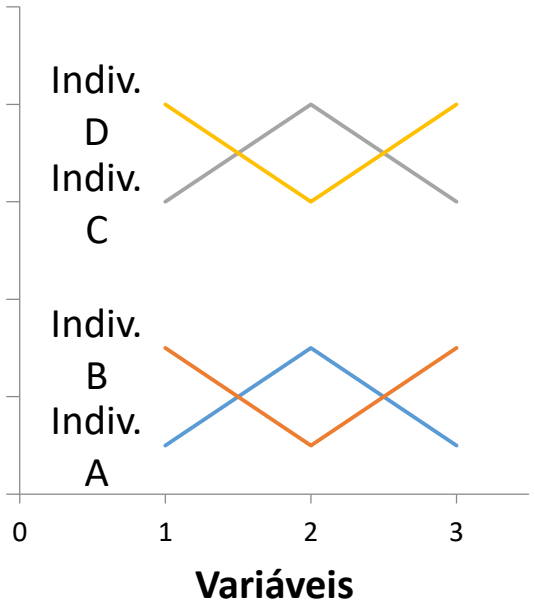
Para estudos exploratórios, onde deseja-se criar uma taxonomia (que não foi definida a priori)

1. Usar método hierárquico para determinar a quantidade K de grupos a serem formados
2. Utilizar um método não hierárquico (K-means no SPSS) para a formação dos grupos

Medidas de (dis)similaridade

É possível usar dois conceitos distintos para medida de similaridade:

- Distância
  - Grupo 1: A e B
  - Grupo 2: C e D
- Correlações.
  - Grupo 1: A e C
  - Grupo 2: B e D



Medidas de Distância

- distância euclidiana
- distância euclidiana ao quadrado
- distância euclidiana de Mahalanobis

Obs:

- Conforme o critério dois respondentes podem estar no mesmo grupo ou em grupos diferentes
- Pressupõe variáveis métricas
- Em princípio, a distância de Mahalanobis é o melhor critério, mas depende do caso, do número de variáveis e sobretudo do número de respondentes

Efeito da escala e padronização das variáveis

- É fortemente recomendável padronizar as variáveis quando diferentes escalas são utilizadas. Nestas condições, os dados padronizados se tornam adimensionais e podem comparados

- Uma opção é usar a Curva Normal (Z scores).
- Caso esteja trabalhando com fatores extraídos na análise fatorial, os dados estarão provavelmente padronizados.
- Cuidado com os outliers (valores absurdos, extremos, fora do padrão)

Efeito da escala e padronização das variáveis

Indivíduo	Peso (Kg)	Altura (m)
1	90	1,8
2	82,5	1,75
3	78	1,85
4	81	1,77

Medidas de similaridade (Distância Euclidiana)

	2	3	4
1	7,50	12,00	9,00
2		4,50	1,50
3			3,00

Indivíduo	Peso (A)	Altura (cm)
1	6	180
2	5,5	175
3	5,2	185
4	5,4	177

Medidas de similaridade (Distância Euclidiana)

	2	3	4
1	5,02	5,06	3,06
2		10,00	2,00
3			8,00

Indivíduo	Peso	Altura
1	1,40	0,17
2	-0,07	-0,98
3	-0,95	1,32
4	-0,37	-0,52

Medidas de similaridade (Distância Euclidiana)

	2	3	4
1	1,87	2,62	1,89
2		2,46	0,55
3			1,93

Seleção das variáveis e sua influência no resultado final

Indivíduo	Peso 1 (Kg)	Peso 2 (Kg)	Altura (m)
1	95,00	32,00	1,75
2	92,00	30,00	1,70
3	85,00	28,00	1,72
4	82,00	25,00	1,66

	2	3	4
1	1,57	2,28	3,98
2		1,44	2,58
3			1,95

Indivíduo	Peso 1 (Kg)	Altura 1 (m)	Altura 2 (m)
1	95,00	1,75	0,98
2	92,00	1,70	0,95
3	85,00	1,72	0,97
4	82,00	1,66	0,93

	2	3	4
1	1,90	1,89	3,94
2		1,51	2,21
3			2,47

Procedimento no R Studio

1. Abrir o arquivo de dados
2. Visão geral o método hierárquico aglomerativo – Dendograma plot(hclust(dist(Fatores_Hatco)))
3. Calcular as distâncias (medidas de dissimilaridade)

Medidas de dissimilaridade

opções disponíveis para o parâmetro method da função dist

Exemplo: plot(hclust(dist(Fatores_Hatco, method = "manhattan")))

- euclidean
- maximum
- manhattan
- canberra
- binary
- minkowski

Métodos de aglomeração

opções disponíveis para o parâmetro method da função hclust

Exemplo: plot(hclust(dist(Fatores_Hatco, method = "manhattan"), method = "ward.D2"))

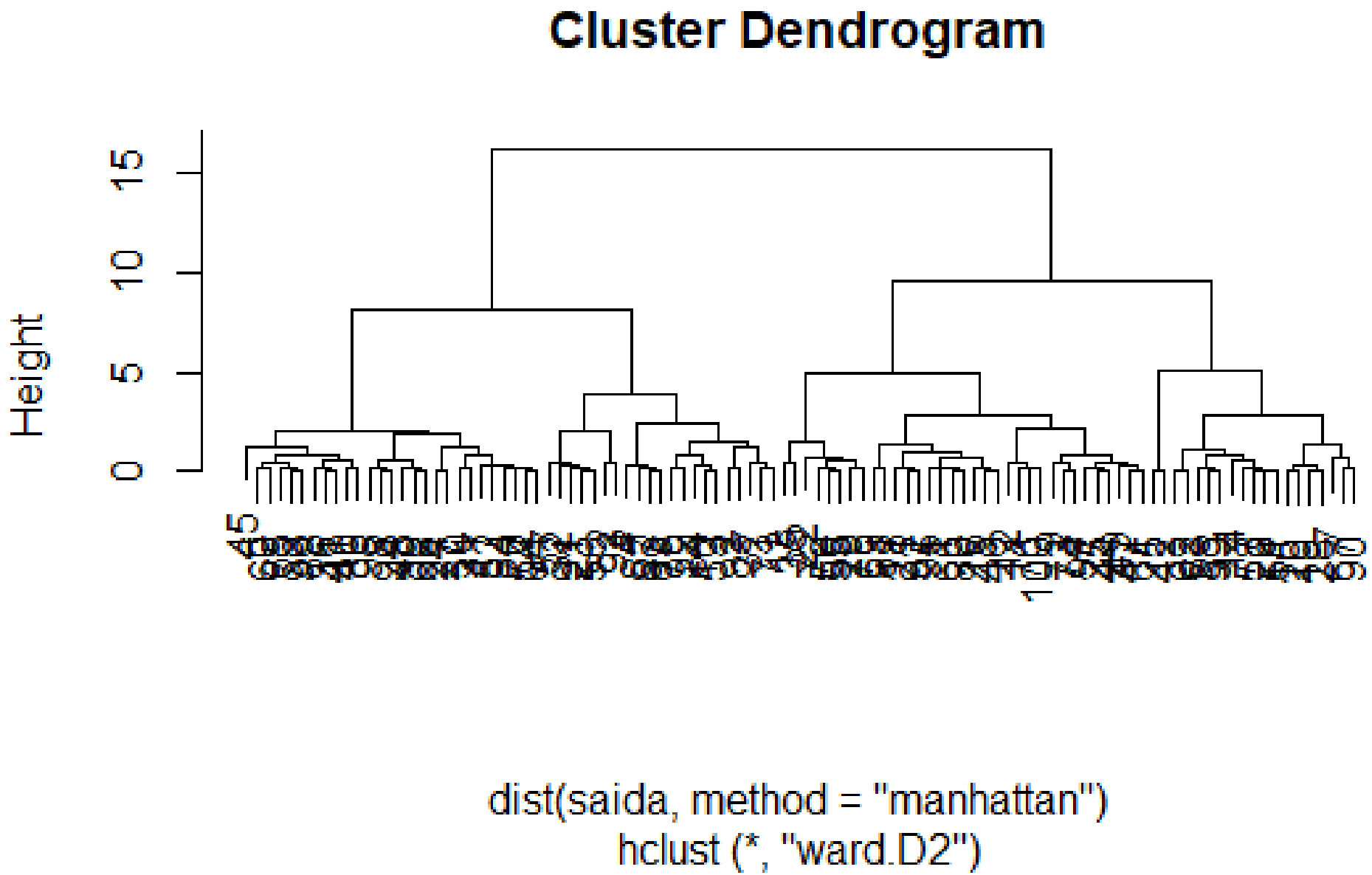
- ward.D
- ward.D2
- single
- complete
- average
- mcquitty
- median
- centroid

Criando os clusters (grupos) – cutree

cutree(tree, k = NULL, h = NULL)

- tree: dendograma
- k: número de grupos a serem formados
- h: número (ou vetor) com alturas onde a árvore deve ser cortada.

- Exemplo:
  grupos <- cutree(hclust(dist(Fatores_Hatco, method = "manhattan"), method = "ward.D"),4)


Cluster Dendrogram
Height
0
5
10
15
dist(saida, method = "manhattan")
hclust (*, "ward.D2")

Salvando os clusters (grupos) criados

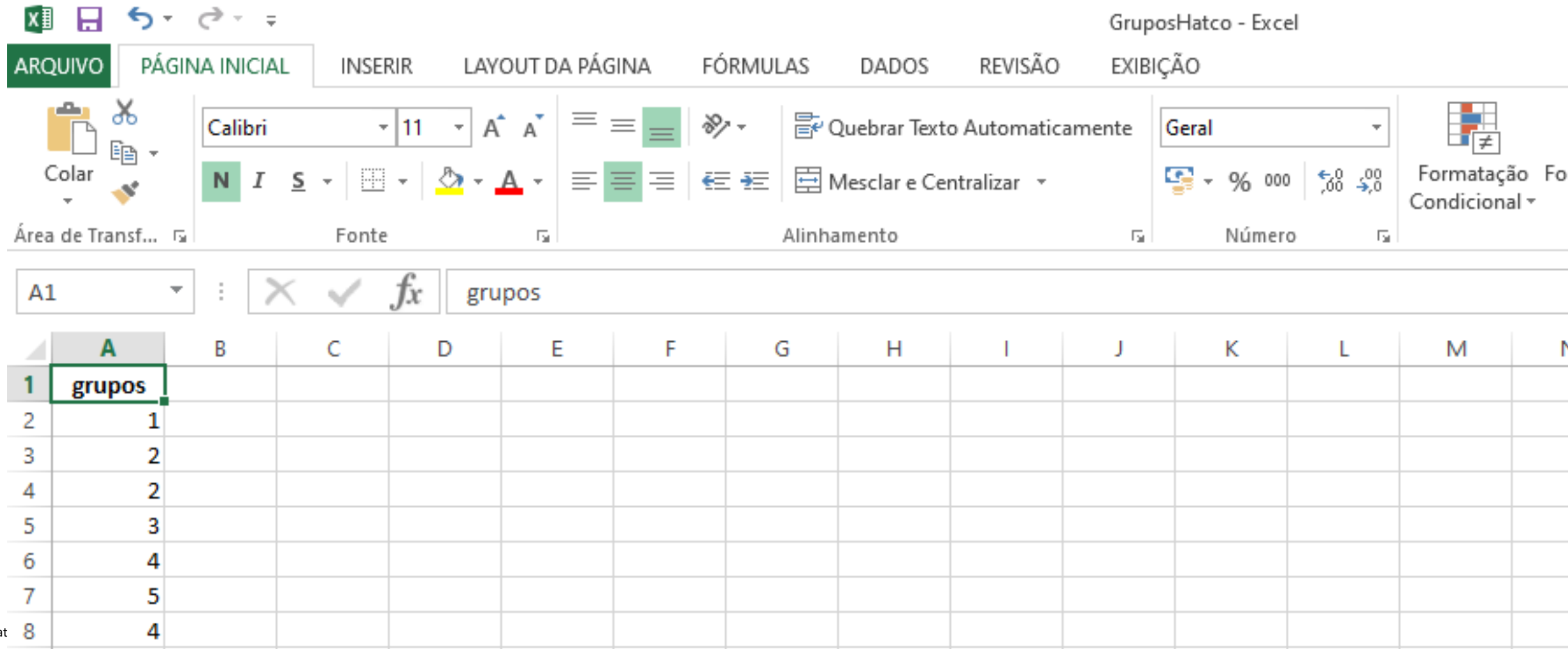
Convertendo os valores dos clusters em data frame:

- clusters <- as.data.frame(grupos)

Salvando em um arquivo Excel:

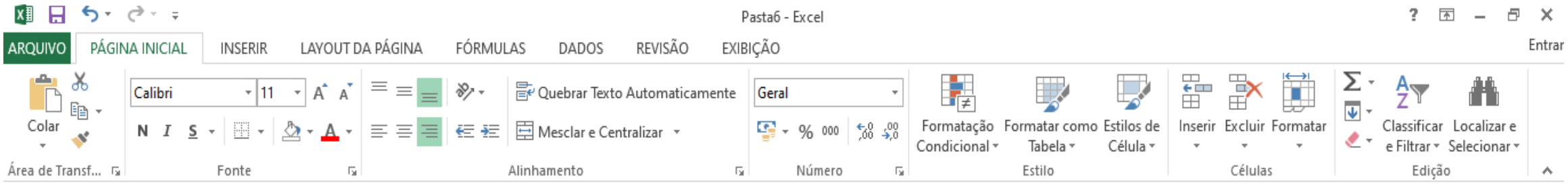
- library(writexl)
- write_xlsx(clusters, "Renato/2023/PRO2514 - Pesquisa Quantitativa em Gestão de Operações/GruposHatco.xlsx")

Arquivo gerado



	A
1	grupos
2	1
3	2
4	2
5	3
6	4
7	5
8	4

Dando significado aos grupos formados

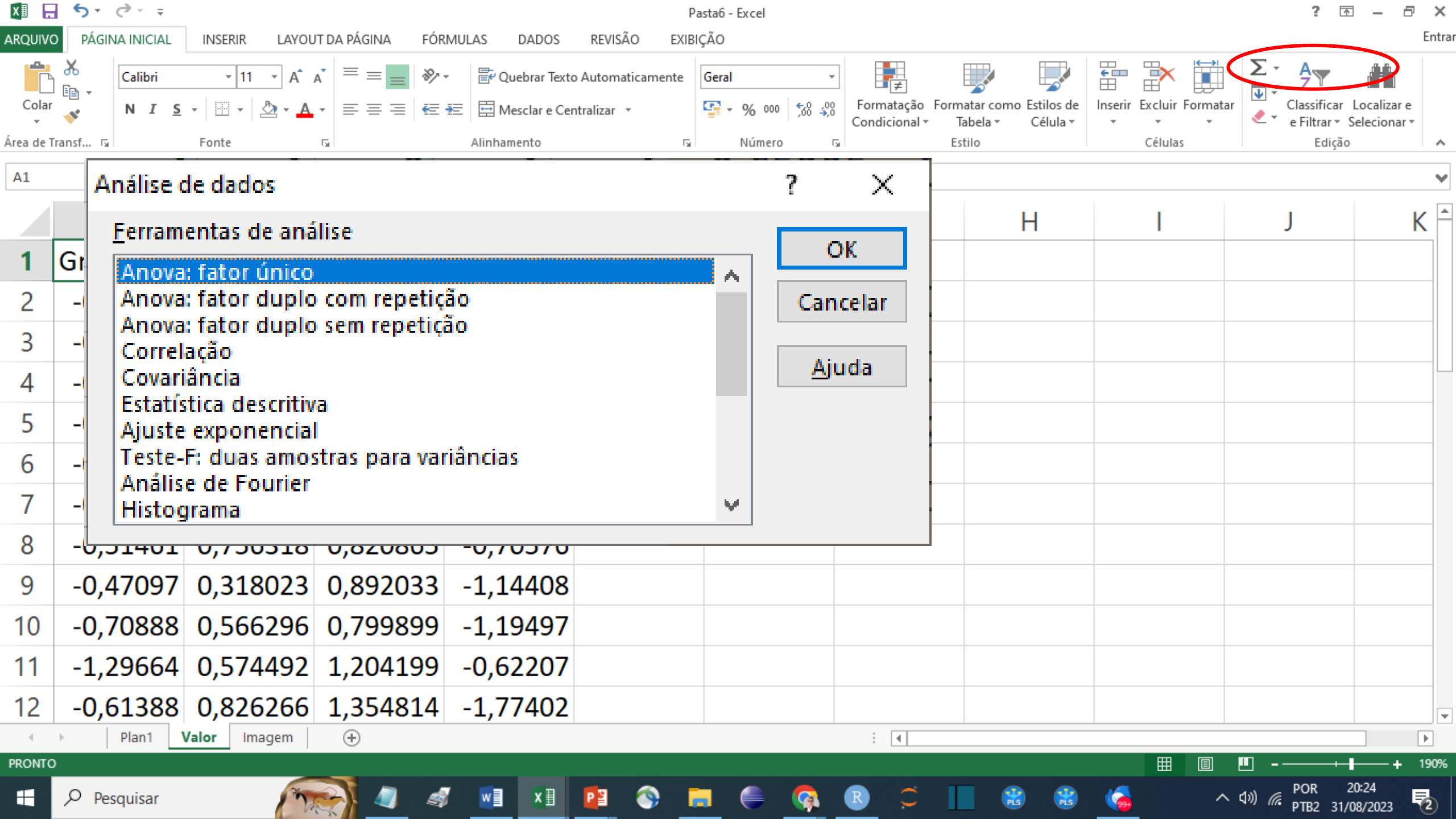


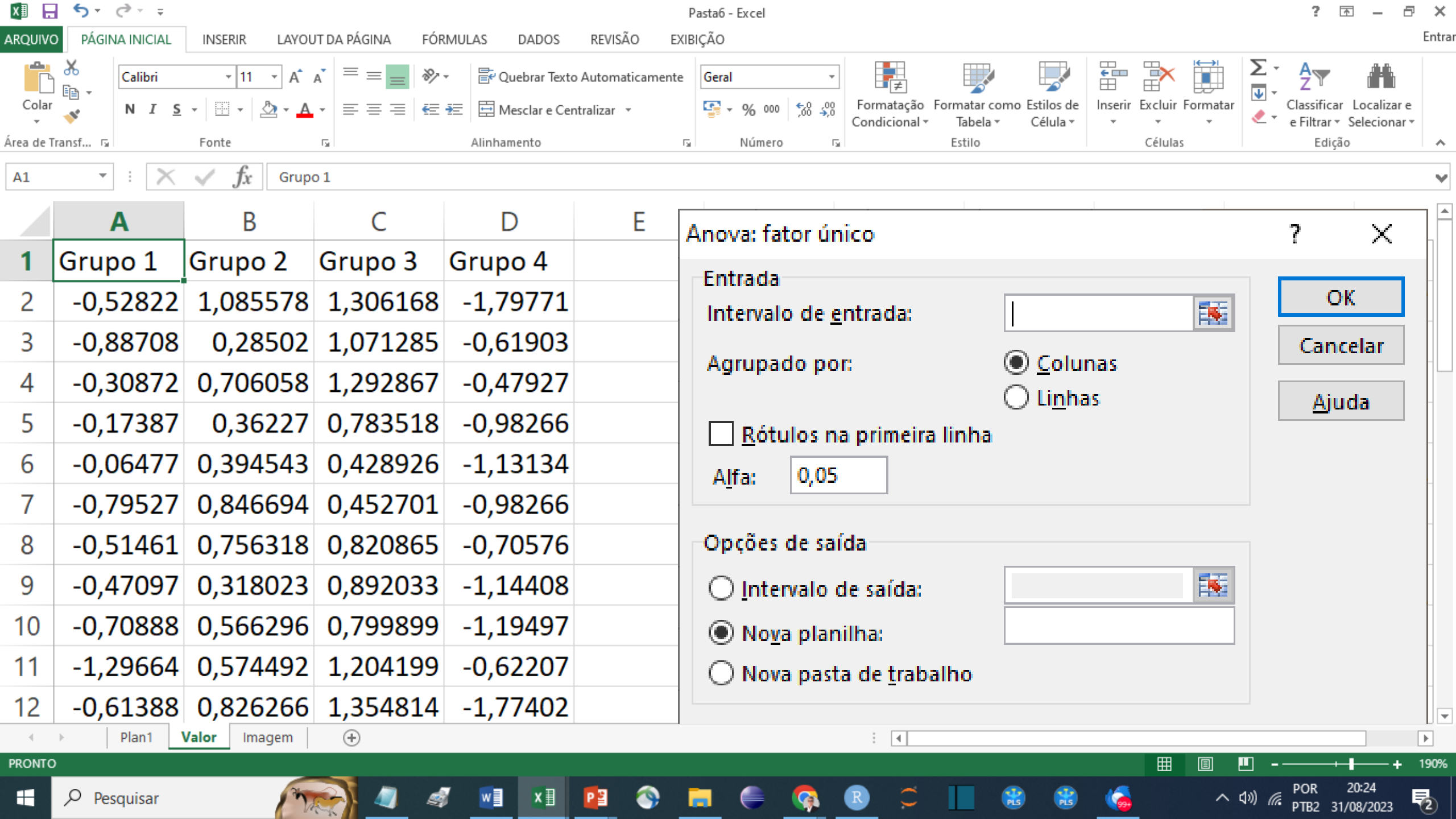
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1	id	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10	x11	x12	x13	x14	Valor	Imagem	grupos		
2	1	4	1	7	5	2	2	5	0	32	4	1	0	1	1	-0,52822	-0,47444	1		
3	2	2	3	6	7	3	4	8	1	43	4	0	1	0	1	1,08558	1,19907	2		
4	3	3	5	6	6	4	3	8	1	48	5	0	1	1	2	1,30617	0,45555	3		
5	4	3	1	7	6	2	2	8	1	32	4	0	1	1	1	0,28502	0,24076	2		
6	5	6	1	10	8	3	5	5	0	58	7	1	0	1	3	-1,79771	2,45369	4		
7	6	2	3	8	5	3	2	10	1	45	4	0	1	1	2	1,07129	-0,62269	3		
8	7	5	2	10	7	4	5	8	0	46	6	1	0	1	1	-0,61903	1,56408	4		
9	8	1	4	6	5	3	2	7	1	44	4	0	1	0	2	1,29287	-0,37783	3		
10	9	6	2	9	5	4	3	8	0	63	5	1	0	1	3	-0,88708	-0,13116	1		
11	10	4	4	7	6	4	3	9	1	54	5	0	1	0	2	0,70606	0,65252	2		
12	11	2	2	9	5	2	3	6	0	32	4	1	0	0	1	-0,30872	-0,27675	1		
13	12	4	2	9	5	3	3	8	0	47	5	1	0	1	2	-0,17387	-0,43877	1		

Dando significado aos grupos formados

- Qual é comportamento das variáveis/fatores nos clusters formados?
- Espera-se que as variáveis/fatores usadas na formação dos clusters tenham médias não homogêneas nos clusters formados
- Análise de Variância

$$\begin{cases} H_o: \mu_1 = \mu_2 = \cdots \mu_k \\ H_1: \mu_i \neq \mu_y \end{cases}$$


Pasta6 - Excel
ARQUIVO
PÁGINA INICIAL
INSERIR
LAYOUT DA PÁGINA
FÓRMULAS
DADOS
REVISÃO
EXIBIÇÃO
Entrar
Colar
Calibri
11
Quebrar Texto Automaticamente
Mesclar e Centralizar
Geral
Formatação Condicional
Formatar como Tabela
Estilos de Célula
Inserir
Excluir
Formatar
Classificar e Filtrar
Localizar e Selecionar
Área de Transf...
Fonte
Alinhamento
Número
Estilo
Células
Edição
A1
Análise de dados
Ferramentas de análise
Anova: fator único
Anova: fator duplo com repetição
Anova: fator duplo sem repetição
Correlação
Covariância
Estatística descritiva
Ajuste exponencial
Teste-F: duas amostras para variâncias
Análise de Fourier
Histograma
OK
Cancelar
Ajuda
H
I
J
K
1
2
3
4
5
6
7
8
9
10
11
12
-0,47097 0,318023 0,892033 -1,14408
-0,70888 0,566296 0,799899 -1,19497
-1,29664 0,574492 1,204199 -0,62207
-0,61388 0,826266 1,354814 -1,77402
Plan1
Valor
Imagem
PRONTO
190%
Pesquisar
POR
PTB2
20:24
31/08/2023


Pasta6 - Excel
ARQUIVO
PÁGINA INICIAL
INSERIR
LAYOUT DA PÁGINA
FÓRMULAS
DADOS
REVISÃO
EXIBIÇÃO
Entrar
Colar
Calibri
11
Quebrar Texto Automaticamente
Mesclar e Centralizar
Geral
Formatação Condicional
Formatar como Tabela
Estilos de Célula
Inserir
Excluir
Formatar
Classificar e Filtrar
Localizar e Selecionar
Área de Transf...
Fonte
Alinhamento
Número
Estilo
Células
Edição
A1
Grupo 1
Grupo 2
Grupo 3
Grupo 4
-0,52822 1,085578 1,306168 -1,79771
-0,88708 0,28502 1,071285 -0,61903
-0,30872 0,706058 1,292867 -0,47927
-0,17387 0,36227 0,783518 -0,98266
-0,06477 0,394543 0,428926 -1,13134
-0,79527 0,846694 0,452701 -0,98266
-0,51461 0,756318 0,820865 -0,70576
-0,47097 0,318023 0,892033 -1,14408
-0,70888 0,566296 0,799899 -1,19497
-1,29664 0,574492 1,204199 -0,62207
-0,61388 0,826266 1,354814 -1,77402
Anova: fator único
Entrada
Intervalo de entrada:
Agrupado por:
Colunas
Linhas
Rótulos na primeira linha
Alfa:
0,05
OK
Cancelar
Ajuda
Opções de saída
Intervalo de saída:
Nova planilha:
Nova pasta de trabalho
Plan1
Valor
Imagem
PRONTO
190%
Pesquisar
POR
PTB2
20:24
31/08/2023

Pasta6 - Excel

ARQUIVO | PÁGINA INICIAL | INSERIR | LAYOUT DA PÁGINA | FÓRMULAS | DADOS | REVISÃO | EXIBIÇÃO | Entrar

Colar | Área de Transf... | Calibri | 11 | Fonte | Quebrar Texto Automaticamente | Mesclar e Centralizar | Alinhamento | Geral | Número | Formatação Condicional | Formatar como Tabela | Estilos de Célula | Estilo | Inserir | Excluir | Formatar | Células | Classificar e Filtrar | Localizar e Selecionar | Edição

A1 | Grupo 1

	A	B	C	D	E	F	G	H	I	J	K
1	Grupo 1	Grupo 2	Grupo 3	Grupo 4							
2	-0,52822	1,085578	1,306168	-1,79771							
3	-0,88708	0,28502	1,071285	-0,61903							
4	-0,30872	0,706058	1,292867	-0,47927							
5	-0,17387	0,36227	0,783518	-0,98266							
6	-0,06477	0,394543	0,428926	-1,13134							
7	-0,79527	0,846694	0,452701	-0,98266							
8	-0,51461	0,756318	0,820865	-0,70576							
9	-0,47097	0,318023	0,892033	-1,14408							
10	-0,70888	0,566296	0,799899	-1,19497							
11	-1,29664	0,574492	1,204199	-0,62207							
12	-0,61388	0,826266	1,354814	-1,77402							

Anova: fator único

Entrada

Intervalo de entrada: $A$1:$D$35

Agrupado por: Colunas / Linhas

Rótulos na primeira linha

Alfa: 0,05

Opções de saída

Intervalo de saída: $G$1

Nova planilha:

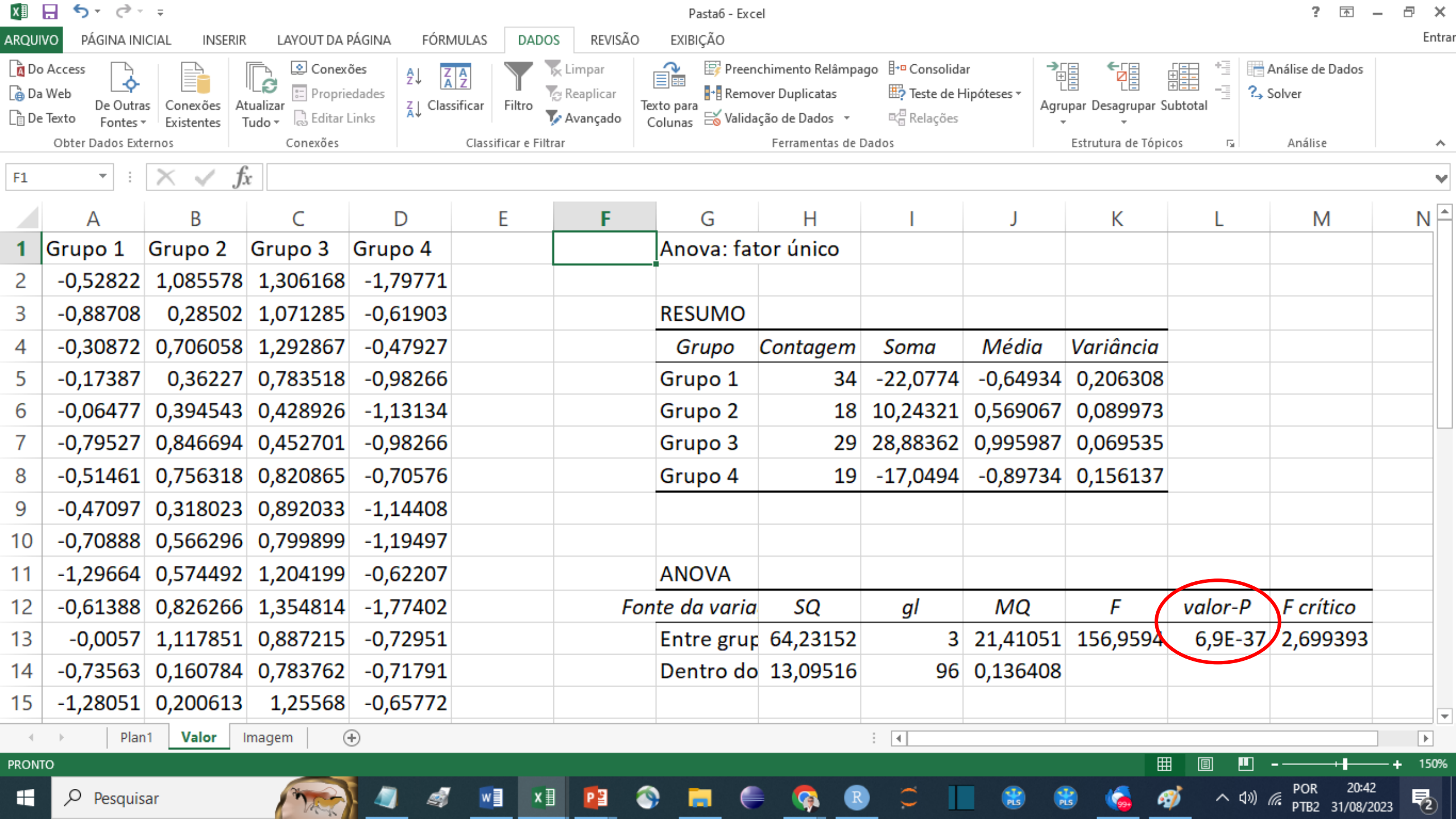
Nova pasta de trabalho

OK | Cancelar | Ajuda

Plan1 | Valor | Imagem

PRONTO | 190%

Pesquisar | POR PTB2 | 20:24 31/08/2023



Pasta6 - Excel

ARQUIVO | PÁGINA INICIAL | INSERIR | LAYOUT DA PÁGINA | FÓRMULAS | DADOS | REVISÃO | EXIBIÇÃO | Entrar

Do Access | Da Web | De Texto | De Outras Fontes | Conexões Existentes | Obter Dados Externos

Atualizar Tudo | Conexões | Propriedades | Editar Links | Conexões

Classificar | Filtro | Limpar | Reaplicar | Avançado | Classificar e Filtrar

Texto para Colunas | Preenchimento Relâmpago | Remover Duplicatas | Validação de Dados | Consolidar | Teste de Hipóteses | Relações | Ferramentas de Dados

Agrupar | Desagrupar | Subtotal | Estrutura de Tópicos

Análise de Dados | Solver | Análise

F1

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Grupo 1	Grupo 2	Grupo 3	Grupo 4			Anova: fator único							
2	-0,52822	1,085578	1,306168	-1,79771										
3	-0,88708	0,28502	1,071285	-0,61903			RESUMO							
4	-0,30872	0,706058	1,292867	-0,47927			*Grupo*	*Contagem*	*Soma*	*Média*	*Variância*			
5	-0,17387	0,36227	0,783518	-0,98266			Grupo 1	34	-22,0774	-0,64934	0,206308			
6	-0,06477	0,394543	0,428926	-1,13134			Grupo 2	18	10,24321	0,569067	0,089973			
7	-0,79527	0,846694	0,452701	-0,98266			Grupo 3	29	28,88362	0,995987	0,069535			
8	-0,51461	0,756318	0,820865	-0,70576			Grupo 4	19	-17,0494	-0,89734	0,156137			
9	-0,47097	0,318023	0,892033	-1,14408										
10	-0,70888	0,566296	0,799899	-1,19497										
11	-1,29664	0,574492	1,204199	-0,62207			ANOVA							
12	-0,61388	0,826266	1,354814	-1,77402			*Fonte da varia*	*SQ*	*gl*	*MQ*	*F*	*valor-P*	*F crítico*	
13	-0,0057	1,117851	0,887215	-0,72951			Entre grup	64,23152	3	21,41051	156,9594	6,9E-37	2,699393	
14	-0,73563	0,160784	0,783762	-0,71791			Dentro do	13,09516	96	0,136408				
15	-1,28051	0,200613	1,25568	-0,65772										

Plan1 | Valor | Imagem

PRONTO | 150%

Pesquisar | POR PTB2 | 20:42 31/08/2023

Comparação de médias dois a dois – teste t

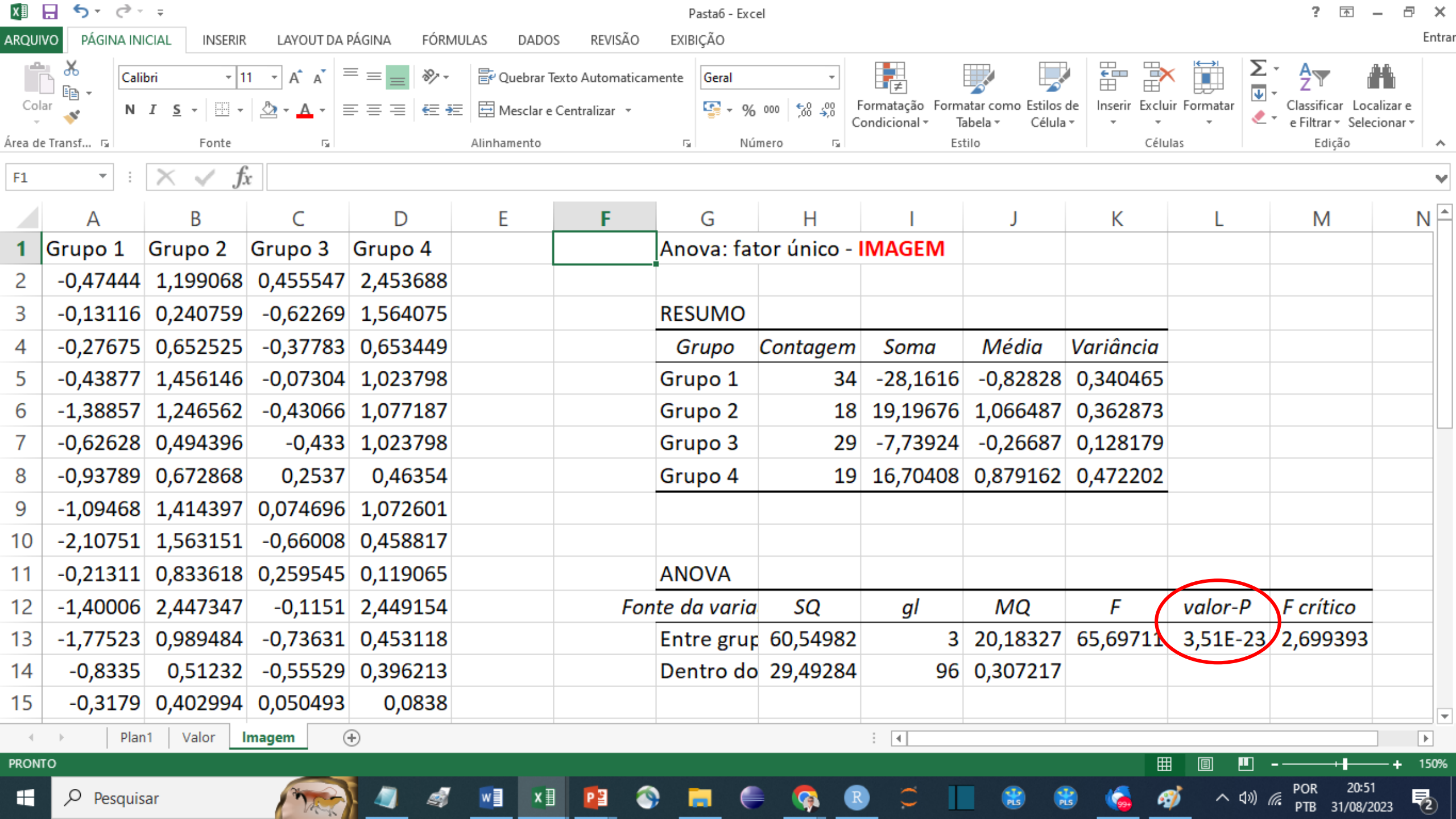
Valor: $\begin{cases} H_o: \mu_1 = \mu_2 \\ H_1: \mu_1 \neq \mu_2 \end{cases}$

	Grupo 1	Grupo 2	Grupo 3	Grupo 4
Grupo 1	1	0,000	0,000	0,052
Grupo 2	0,000	1	0,000	0,000
Grupo 3	0,000	0,000	1	0,000
Grupo 4	0,052	0,000	0,000	1

Fator Valor

Grupo	Média
Grupo 4	-0,897
Grupo 1	-0,649
Grupo 2	0,569
Grupo 3	0,996

Grupo 4		
Grupo 1		
	Grupo 2	
		Grupo 3



Pasta6 - Excel

ARQUIVO | PÁGINA INICIAL | INSERIR | LAYOUT DA PÁGINA | FÓRMULAS | DADOS | REVISÃO | EXIBIÇÃO | Entrar

Colar | Área de Transf... | Calibri | 11 | Fonte | Quebrar Texto Automaticamente | Mesclar e Centralizar | Alinhamento | Geral | Número | Formatação Condicional | Formatar como Tabela | Estilos de Célula | Estilo | Inserir | Excluir | Formatar | Células | Classificar e Filtrar | Localizar e Selecionar | Edição

F1

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Grupo 1	Grupo 2	Grupo 3	Grupo 4			Anova: fator único - IMAGEM							
2	-0,47444	1,199068	0,455547	2,453688										
3	-0,13116	0,240759	-0,62269	1,564075			RESUMO							
4	-0,27675	0,652525	-0,37783	0,653449			*Grupo*	*Contagem*	*Soma*	*Média*	*Variância*			
5	-0,43877	1,456146	-0,07304	1,023798			Grupo 1	34	-28,1616	-0,82828	0,340465			
6	-1,38857	1,246562	-0,43066	1,077187			Grupo 2	18	19,19676	1,066487	0,362873			
7	-0,62628	0,494396	-0,433	1,023798			Grupo 3	29	-7,73924	-0,26687	0,128179			
8	-0,93789	0,672868	0,2537	0,46354			Grupo 4	19	16,70408	0,879162	0,472202			
9	-1,09468	1,414397	0,074696	1,072601										
10	-2,10751	1,563151	-0,66008	0,458817										
11	-0,21311	0,833618	0,259545	0,119065			ANOVA							
12	-1,40006	2,447347	-0,1151	2,449154			*Fonte da varia*	*SQ*	*gl*	*MQ*	*F*	*valor-P*	*F crítico*	
13	-1,77523	0,989484	-0,73631	0,453118			Entre grup	60,54982	3	20,18327	65,69711	3,51E-23	2,699393	
14	-0,8335	0,51232	-0,55529	0,396213			Dentro do	29,49284	96	0,307217				
15	-0,3179	0,402994	0,050493	0,0838										

Plan1 | Valor | Imagem

PRONTO | 150%

Pesquisar | POR PTB | 20:51 31/08/2023

Comparação de médias dois a dois – teste t

Imagem: $\begin{cases} H_o: \mu_1 = \mu_2 \\ H_1: \mu_1 \neq \mu_2 \end{cases}$

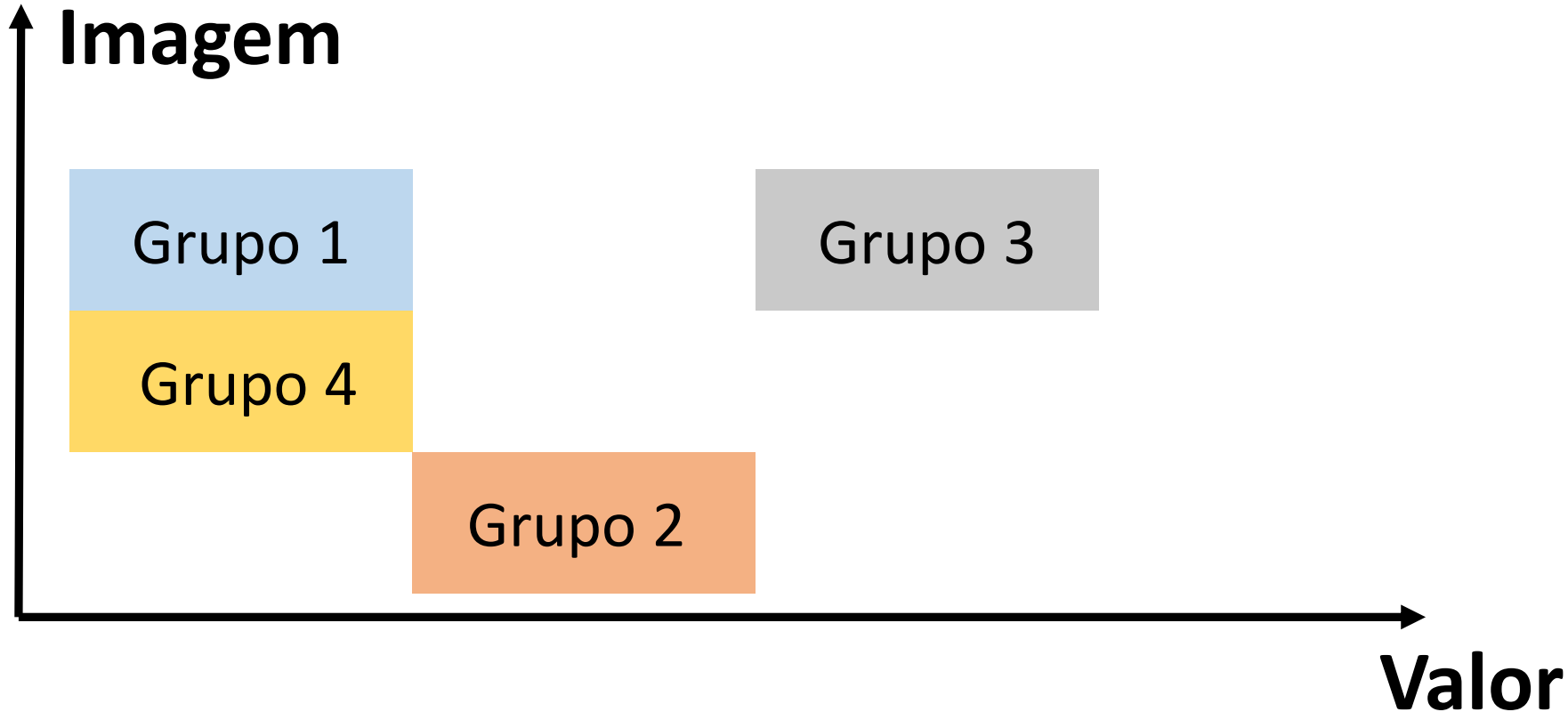
	Grupo 1	Grupo 2	Grupo 3	Grupo 4
Grupo 1	1	0,000	0,000	0,000
Grupo 2	0,000	1	0,000	0,385
Grupo 3	0,000	0,000	1	0,000
Grupo 4	0,000	0,385	0,000	1

Fator Imagem

Grupo	Média
Grupo 1	-0,828
Grupo 3	-0,267
Grupo 4	0,879
Grupo 2	1,066

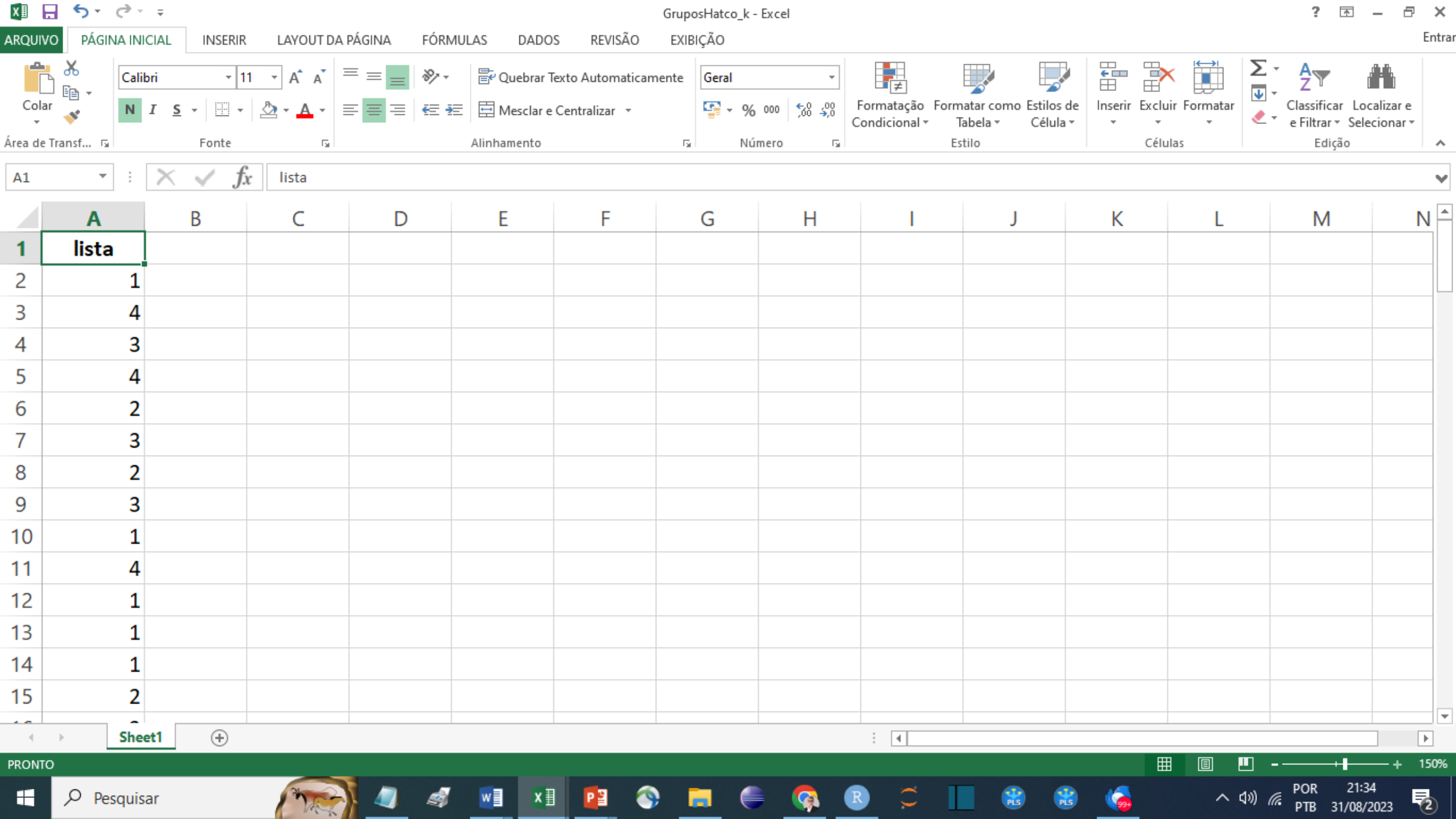
Grupo 1		
	Grupo 3	
		Grupo 4
		Grupo 2

Dando significado aos grupos



Método k-means no RStudio:

```
kmeans(saida, 4)
grp_k <- kmeans(saida, 4)
grp_k$cluster
lista <- grp_k$cluster
lista <- grp_k$cluster
clusters_k <- as.data.frame(lista)
write_xlsx(clusters_k, "Renato/2023/PRO2514 - Pesquisa Quantitativa em Gestão de Operações/GruposHatco_k.xlsx")
```



	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	lista													
2	1													
3	4													
4	3													
5	4													
6	2													
7	3													
8	2													
9	3													
10	1													
11	4													
12	1													
13	1													
14	1													
15	2													

ARQUIVO | PÁGINA INICIAL | INSERIR | LAYOUT DA PÁGINA | FÓRMULAS | DADOS | REVISÃO | EXIBIÇÃO | Entrar

Colar | Calibri | 11 | Quebrar Texto Automaticamente | Geral | Formatação Condicional | Formatar como Tabela | Estilos de Célula | Inserir | Excluir | Formatar | Classificar e Filtrar | Localizar e Selecionar

Área de Transf... | Fonte | Alinhamento | Número | Estilo | Células | Edição

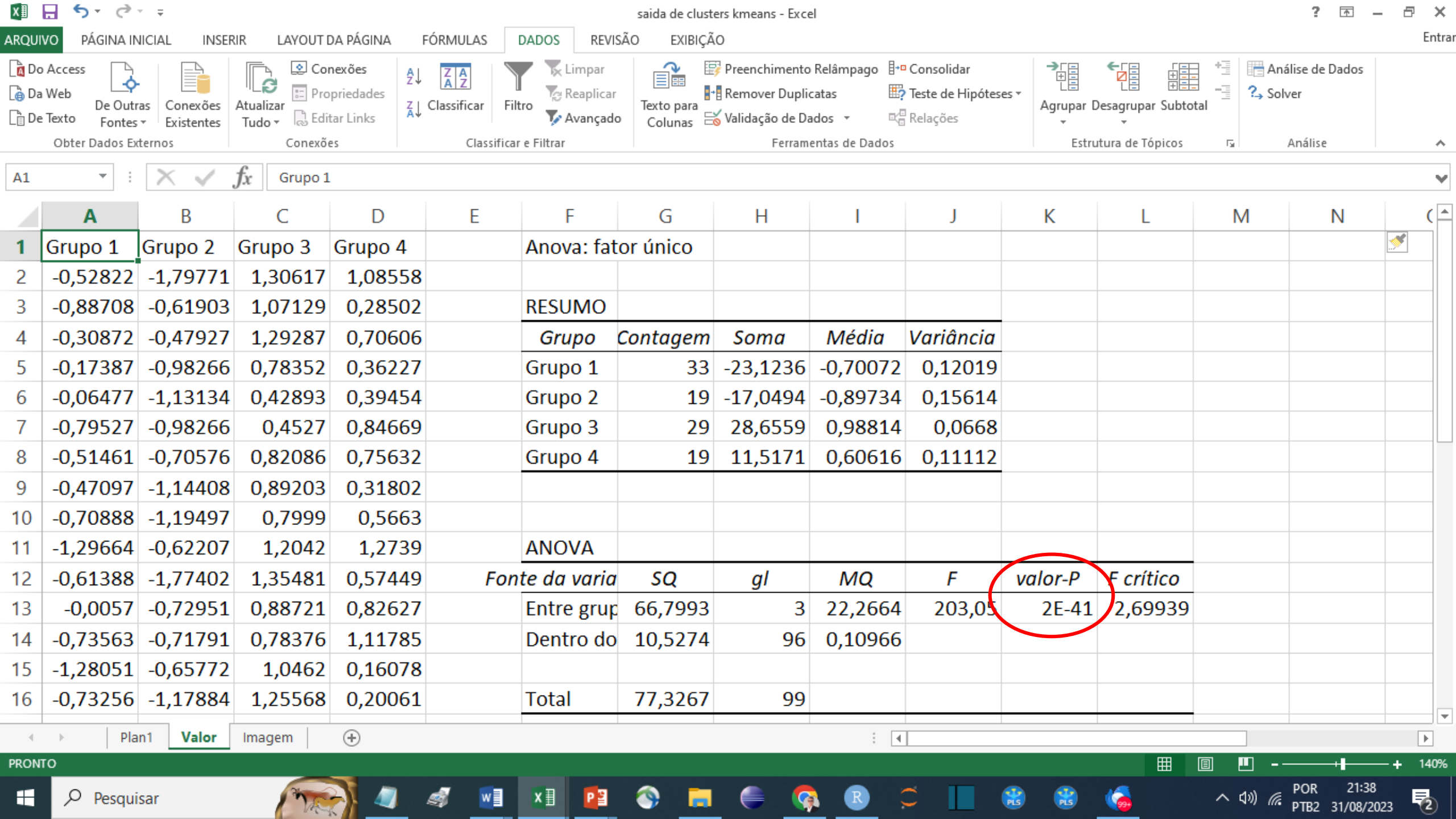
A1 | id

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1	id	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10	x11	x12	x13	x14	Valor	Imagem	grupos		
2	1	4	1	7	5	2	2	5	0	32	4	1	0	1	1	-0,52822	-0,47444	1		
3	2	2	3	6	7	3	4	8	1	43	4	0	1	0	1	1,08558	1,19907	4		
4	3	3	5	6	6	4	3	8	1	48	5	0	1	1	2	1,30617	0,45555	3		
5	4	3	1	7	6	2	2	8	1	32	4	0	1	1	1	0,28502	0,24076	4		
6	5	6	1	10	8	3	5	5	0	58	7	1	0	1	3	-1,79771	2,45369	2		
7	6	2	3	8	5	3	2	10	1	45	4	0	1	1	2	1,07129	-0,62269	3		
8	7	5	2	10	7	4	5	8	0	46	6	1	0	1	1	-0,61903	1,56408	2		
9	8	1	4	6	5	3	2	7	1	44	4	0	1	0	2	1,29287	-0,37783	3		
10	9	6	2	9	5	4	3	8	0	63	5	1	0	1	3	-0,88708	-0,13116	1		
11	10	4	4	7	6	4	3	9	1	54	5	0	1	0	2	0,70606	0,65252	4		
12	11	2	2	9	5	2	3	6	0	32	4	1	0	0	1	-0,30872	-0,27675	1		
13	12	4	2	9	5	3	3	8	0	47	5	1	0	1	2	-0,17387	-0,43877	1		
14	13	3	1	8	4	2	1	7	1	39	4	0	1	0	1	-0,06477	-1,38857	1		
15	14	4	2	9	6	3	4	7	0	38	5	1	0	1	1	-0,47927	0,65345	2		
16	15	5	1	10	7	3	3	7	0	54	6	1	0	0	3	-0,98266	1,0238	2		

Plan1 | Valor | Imagem

PRONTO | 140%

Pesquisar | POR PTB2 | 21:37 31/08/2023



saida de clusters kmeans - Excel

ARQUIVO | PÁGINA INICIAL | INSERIR | LAYOUT DA PÁGINA | FÓRMULAS | DADOS | REVISÃO | EXIBIÇÃO | Entrar

Do Access | Da Web | De Texto | De Outras Fontes | Conexões Existentes | Obter Dados Externos
Atualizar Tudo | Conexões | Propriedades | Editar Links | Conexões
Classificar | Filtro | Limpar | Reaplicar | Avançado | Classificar e Filtrar
Texto para Colunas | Preenchimento Relâmpago | Remover Duplicatas | Validação de Dados | Consolidar | Teste de Hipóteses | Relações | Ferramentas de Dados
Agrupar | Desagrupar | Subtotal | Estrutura de Tópicos
Análise de Dados | Solver | Análise

A1 | Grupo 1

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Grupo 1	Grupo 2	Grupo 3	Grupo 4		Anova: fator único								
2	-0,52822	-1,79771	1,30617	1,08558										
3	-0,88708	-0,61903	1,07129	0,28502		RESUMO								
4	-0,30872	-0,47927	1,29287	0,70606		*Grupo*	*Contagem*	*Soma*	*Média*	*Variância*				
5	-0,17387	-0,98266	0,78352	0,36227		Grupo 1	33	-23,1236	-0,70072	0,12019				
6	-0,06477	-1,13134	0,42893	0,39454		Grupo 2	19	-17,0494	-0,89734	0,15614				
7	-0,79527	-0,98266	0,4527	0,84669		Grupo 3	29	28,6559	0,98814	0,0668				
8	-0,51461	-0,70576	0,82086	0,75632		Grupo 4	19	11,5171	0,60616	0,11112				
9	-0,47097	-1,14408	0,89203	0,31802										
10	-0,70888	-1,19497	0,7999	0,5663										
11	-1,29664	-0,62207	1,2042	1,2739		ANOVA								
12	-0,61388	-1,77402	1,35481	0,57449		*Fonte da varia*	*SQ*	*gl*	*MQ*	*F*	*valor-P*	*F crítico*		
13	-0,0057	-0,72951	0,88721	0,82627		Entre grup	66,7993	3	22,2664	203,05	2E-41	2,69939		
14	-0,73563	-0,71791	0,78376	1,11785		Dentro do	10,5274	96	0,10966					
15	-1,28051	-0,65772	1,0462	0,16078										
16	-0,73256	-1,17884	1,25568	0,20061		Total	77,3267	99						

Plan1 | Valor | Imagem

PRONTO | 140%

Pesquisar | POR PTB2 | 21:38 31/08/2023

Comparação de médias dois a dois – teste t

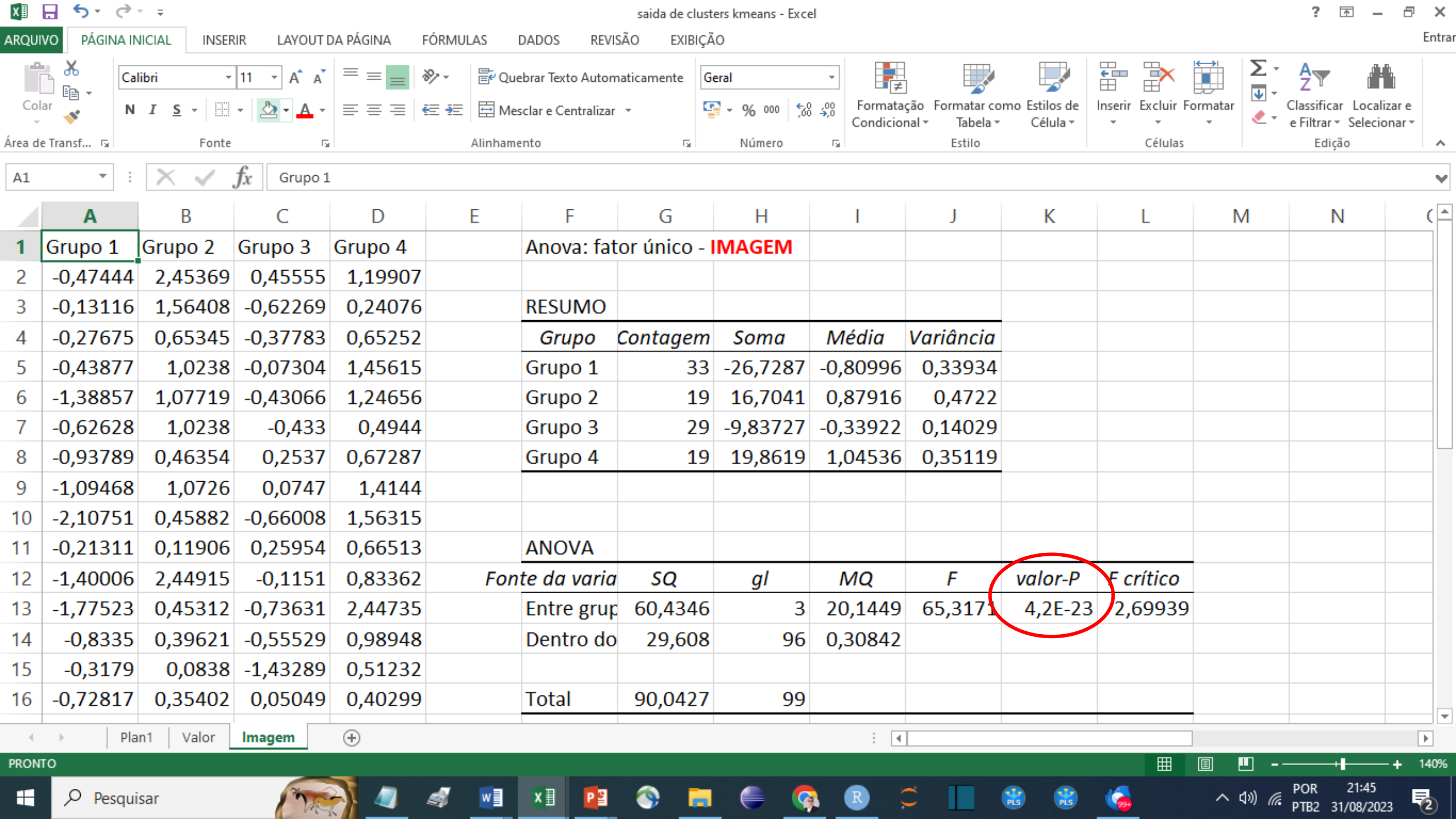
Valor: $\begin{cases} H_o: \mu_1 = \mu_2 \\ H_1: \mu_1 \neq \mu_2 \end{cases}$

	Grupo 1	Grupo 2	Grupo 3	Grupo 4
Grupo 1	1	0,067	0,000	0,000
Grupo 2	0,067	1	0,000	0,000
Grupo 3	0,000	0,000	1	0,000
Grupo 4	0,000	0,000	0,000	1

Fator Valor

Grupo	Média
Grupo 2	-0,90
Grupo 1	-0,70
Grupo 4	0,61
Grupo 3	0,99

Grupo 2		
Grupo 1		
	Grupo 4	
		Grupo 3



saida de clusters kmeans - Excel

ARQUIVO | PÁGINA INICIAL | INSERIR | LAYOUT DA PÁGINA | FÓRMULAS | DADOS | REVISÃO | EXIBIÇÃO | Entrar

Colar | Área de Transf... | Calibri | 11 | Fonte | Quebrar Texto Automaticamente | Mesclar e Centralizar | Alinhamento | Geral | Número | Formatação Condicional | Formatar como Tabela | Estilos de Célula | Estilo | Inserir | Excluir | Formatar | Células | Classificar e Filtrar | Localizar e Selecionar | Edição

A1 | Grupo 1

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Grupo 1	Grupo 2	Grupo 3	Grupo 4		Anova: fator único - IMAGEM								
2	-0,47444	2,45369	0,45555	1,19907										
3	-0,13116	1,56408	-0,62269	0,24076		RESUMO								
4	-0,27675	0,65345	-0,37783	0,65252		*Grupo*	*Contagem*	*Soma*	*Média*	*Variância*				
5	-0,43877	1,0238	-0,07304	1,45615		Grupo 1	33	-26,7287	-0,80996	0,33934				
6	-1,38857	1,07719	-0,43066	1,24656		Grupo 2	19	16,7041	0,87916	0,4722				
7	-0,62628	1,0238	-0,433	0,4944		Grupo 3	29	-9,83727	-0,33922	0,14029				
8	-0,93789	0,46354	0,2537	0,67287		Grupo 4	19	19,8619	1,04536	0,35119				
9	-1,09468	1,0726	0,0747	1,4144										
10	-2,10751	0,45882	-0,66008	1,56315										
11	-0,21311	0,11906	0,25954	0,66513		ANOVA								
12	-1,40006	2,44915	-0,1151	0,83362		*Fonte da varia*	*SQ*	*gl*	*MQ*	*F*	*valor-P*	*F crítico*		
13	-1,77523	0,45312	-0,73631	2,44735		Entre grup	60,4346	3	20,1449	65,3171	4,2E-23	2,69939		
14	-0,8335	0,39621	-0,55529	0,98948		Dentro do	29,608	96	0,30842					
15	-0,3179	0,0838	-1,43289	0,51232										
16	-0,72817	0,35402	0,05049	0,40299		Total	90,0427	99						

Plan1 | Valor | Imagem

PRONTO | 140%

Pesquisar | POR PTB2 | 21:45 31/08/2023

Comparação de médias dois a dois – teste t

Imagem: $\begin{cases} H_o: \mu_1 = \mu_2 \\ H_1: \mu_1 \neq \mu_2 \end{cases}$

	Grupo 1	Grupo 2	Grupo 3	Grupo 4
Grupo 1	1	0,000	0,000	0,000
Grupo 2	0,000	1	0,000	0,430
Grupo 3	0,000	0,000	1	0,000
Grupo 4	0,000	0,430	0,000	1

Fator Imagem

Grupo	Média
Grupo 1	-0,810
Grupo 3	-0,339
Grupo 2	0,879
Grupo 4	1,045

Grupo 1		
	Grupo 3	
		Grupo 2
		Grupo 4

Dando significado aos grupos