



Excel

Ferramentas de análise

Catarina
Oliveira

DCT DEPARTAMENTO DE CIÊNCIA
E TECNOLOGIA

CONTEÚDO

1. Ferramentas de análise
2. Instalação e utilização do suplemento Analysis Toolpak
3. Ferramentas de análise de:
 1. Variância
 2. Correlação
 3. Covariância
 4. Estatística descritiva
 5. Ajuste exponencial
 6. Teste F
 7. Análise de Fourier
 8. Histograma
 9. Média Móvel
 10. Geração de número Aleatório
 11. Ordem e Percentil
 12. Regressão
 13. Amostragem
 14. Teste T
 15. Teste Z

Ferramentas de análise

Suplemento *Analysis Toolpak*

Algumas ferramentas que poderão ser úteis na resolução de diversos tipos de problemas de

- Estatística descritiva
- Inferência estatística
- Correlação
- Regressão

Incluídas ferramentas de análise de:

- Variância
- Correlação
- Covariância
- Estatística descritiva
- Ajuste exponencial
- Teste F
- Análise de *Fourier*
- Histograma
- Média Móvel
- Geração de número Aleatório
- Ordem e Percentil
- Regressão
- Amostragem
- Teste T
- Teste Z

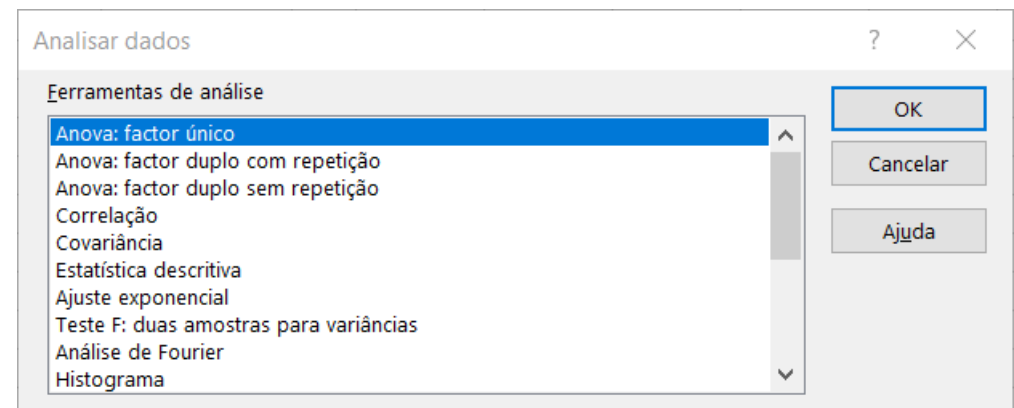
Instalação e utilização do suplemento *Analysis Toolpak*

Instalação:

1. Separador “Ficheiro” > Comando “Opções”
2. Na caixa de diálogo Opções do Excel
 1. Selecionar o separador Suplementos
 2. Na caixa listagem Gerir selecione a opção Suplementos do Excel e clique em Ir
3. Na caixa de diálogo Suplementos
 1. Na secção Suplementos disponíveis selecione a caixa do Suplemento *Analysis Toolpak*

Utilização:

1. Separador “Dados” > Grupo “Análise” > Comando “Análise de Dados”
2. Na caixa de diálogo Analisar Dados
 1. Selecionar a ferramenta pretendida
 2. Ok
 3. Preencher a caixa de diálogo da ferramenta



Análise de variância (ANOVA)

Objetivo: comparar a média (ou outra medida de localização) de diferentes amostras

Permite: concluir acerca da igualdade das mesmas

H0: as médias das distribuições são iguais

Três ferramentas (consoante o nº de amostras e o nº de fatores):

- **Anova – Fator Único:** executa uma análise simples de variância referente aos dados de uma ou mais amostras
- **Anova – Fator Duplo com Repetição:** é utilizado quando os dados podem ser classificados ou possuem duas dimensões diferentes
- **Anova – Fator Duplo sem Repetição:** semelhante à anterior, mas com a diferença de que existe apenas uma observação para cada par de dimensões

Análise de variância (ANOVA)

Produção	Defeitos
922	11
1189	10
1109	5
1152	8
893	11
846	12
802	9
1110	14
1044	6
1117	14
1000	5
1065	0

Anova: factor único						
SUMÁRIO						
Grupos	Contagem	Soma	Média	Variância		
Produção	12	12249	1020,75	16176,56818		
Defeitos	12	105	8,75	17,29545455		
ANOVA						
Fonte de variação	SQ	gl	MQ	F	valor P	F crítico
Entre grupos	6144864	1	6144864	758,9126521	1,51182E-18	4,300949502
Dentro de grupos	178132,5	22	8096,931818			
Total	6322996,5	23				

H0: as médias das distribuições são iguais

$F > F_{\text{crítico}} \rightarrow$ rejeitamos H0 \rightarrow As médias são diferentes

A	B
922	1065
1189	998
1109	1117
1152	1044
893	1110
846	500
802	846
1110	893
1044	1152
1117	1109
1000	1185
1065	922

Anova: factor único						
SUMÁRIO						
Grupos	Contagem	Soma	Média	Variância		
A	12	12249	1020,75	16176,56818		
B	12	11941	995,0833333	35647,53788		
ANOVA						
Fonte de variação	SQ	gl	MQ	F	valor P	F crítica
Entre grupos	3952,666667	1	3952,666667	0,152541625	0,699874513	4,300949502
Dentro de grupos	570065,1667	22	25912,05303			
Total	574017,8333	23				

H0: as médias das distribuições são iguais

$F < F_{\text{crítico}} \rightarrow$ não rejeitamos H0 \rightarrow As médias são iguais

Correlação

Permite:

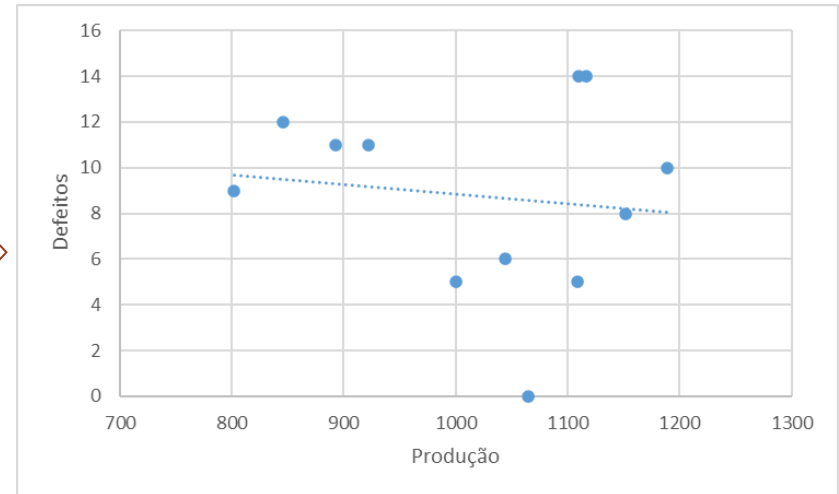
- Analisar e avaliar a magnitude e a direção da associação ou correlação entre duas variáveis sem qualquer tipo de assunção de funcionalidade (ou seja: nenhuma das variáveis é tida como dependente da outra variável)
- Determinar se as duas variáveis de medida tendem a deslocar-se em conjunto
 - **Correlação positiva:** se grandes valores de uma variável tendem a estar associados a grandes valores da outra
 - **Correlação negativa:** se pequenos valores de uma variável tendem a estar associados a grandes valores da outra
 - **Correlação próxima de zero:** se os valores de ambas as variáveis tendem a não estar relacionados

Correlação

Produção	Defeitos
922	11
1189	10
1109	5
1152	8
893	11
846	12
802	9
1110	14
1044	6
1117	14
1000	5
1065	0



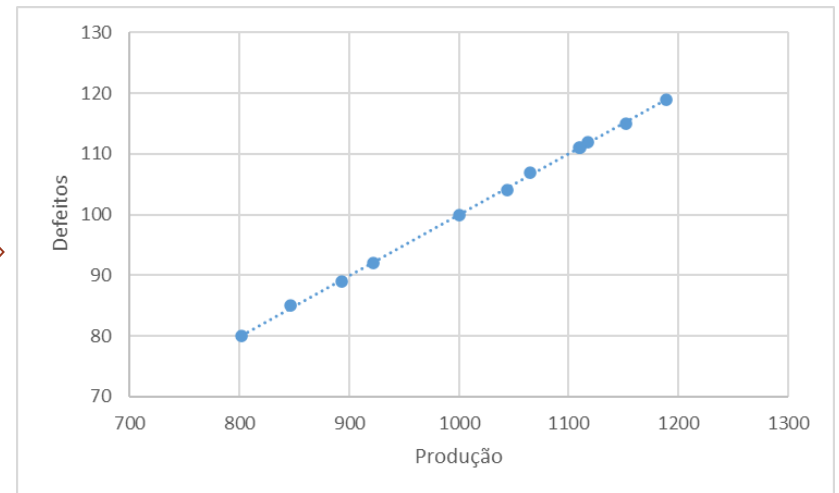
	Produção	Defeitos
Produção	1	
Defeitos	-0,129030824	1



Produção	Defeitos
922	92
1189	119
1109	111
1152	115
893	89
846	85
802	80
1110	111
1044	104
1117	112
1000	100
1065	107



	Produção	Defeitos
Produção	1	
Defeitos	0,999760629	1



Covariância

Permite:

- Analisar duas variáveis de modo a determinar se as duas variáveis tendem a deslocar-se em conjunto
 - **Covariância positiva:** se grandes valores de uma variável tendem a estar associados a grandes valores da outra
 - **Covariância negativa:** se pequenos valores de uma variável tendem a estar associados a grandes valores da outra
 - **Covariância próxima de zero:** se os valores de ambas as variáveis tendem a não estar relacionados

Estatística Descritiva

- Permite gerar automaticamente um relatório de estatística invariável para os dados definidos como intervalo de entrada, fornecendo um conjunto de informações sobre a tendência central e a variabilidade dos dados

Transporte A	Transporte B	Transporte C
144	128	132
139	123	134
137	112	131
131	129	111
109	110	126
138	128	128
126	146	107
132	134	105
128	120	126
128	109	117



Transporte A		Transporte B		Transporte C	
Média	131,2	Média	123,9	Média	121,7
Erro-padrão	3,065217049	Erro-padrão	3,686160303	Erro-padrão	3,419064199
Mediana	131,5	Mediana	125,5	Mediana	126
Moda	128	Moda	128	Moda	126
Desvio-padrão	9,693067397	Desvio-padrão	11,65666238	Desvio-padrão	10,81203034
Variância da amostra	93,95555556	Variância da amostra	135,8777778	Variância da amostra	116,9
Curtose	2,525244444	Curtose	-0,117946666	Curtose	-1,433295139
Assimetria	-1,230019145	Assimetria	0,381962526	Assimetria	-0,548977137
Intervalo	35	Intervalo	37	Intervalo	29
Mínimo	109	Mínimo	109	Mínimo	105
Máximo	144	Máximo	146	Máximo	134
Soma	1312	Soma	1239	Soma	1217
Contagem	10	Contagem	10	Contagem	10
Nível de confiança(95,0%)	6,934002702	Nível de confiança(95,0%)	8,338673932	Nível de confiança(95,0%)	7,734460569

Ajuste exponencial

- Ajuda a prever um valor baseado numa previsão do período anterior, ajustada para o erro nessa previsão anterior.
- Utiliza uma constante de ajuste, cuja magnitude determina o grau de suavização das previsões em relação aos erros na previsão anterior

Teste F: duas amostras para variâncias

- Executa um teste F de duas amostras para comparar as variâncias de duas populações.
- Utilizado para verificar a significância do modelo de regressão e analisar o ajuste originado pelo modelo

Análise de *Fourier*

- Permite resolver problemas nos sistemas lineares e analisa dados periódicos, usando o método de Transformação de *Fourier* Rápida (FFT) para transformar os dados.
- Oferece suporte a transformações inversas, nas quais o inverso dos dados transformados retorna os dados originais

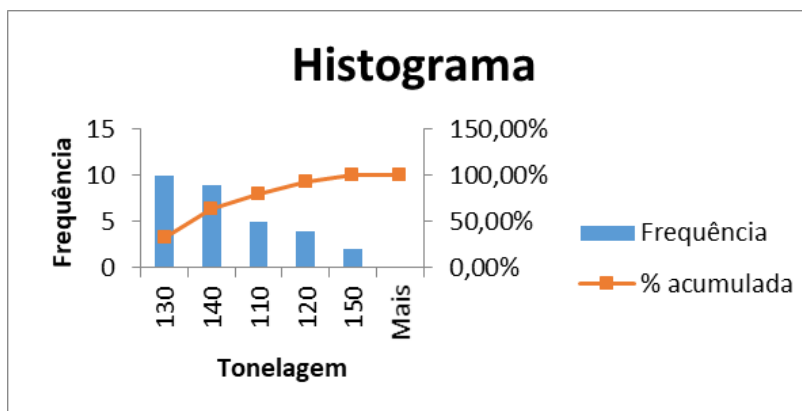
Histograma

- Calcula frequências individuais e cumulativas para um intervalo de células de dados e blocos de dados
- Gera dados para o número de ocorrências de um valor num conjunto de dados

Transporte A	Transporte B	Transporte C
144	128	132
139	123	134
137	112	131
131	129	111
109	110	126
138	128	128
126	146	107
132	134	105
128	120	126
128	109	117



Tonelagem	Frequência	% acumulada	Tonelagem	Frequência	% acumulada
110	5	16,67%	130	10	33,33%
120	4	30,00%	140	9	63,33%
130	10	63,33%	110	5	80,00%
140	9	93,33%	120	4	93,33%
150	2	100,00%	150	2	100,00%
Mais	0	100,00%	Mais	0	100,00%



Média móvel

- Projeta os valores do período de previsão, com base no valor médio da variável num número específico de períodos precedentes
- Fornece informações sobre a tendência que uma simples média de todos os dados históricos não mostra diretamente

Geração de número aleatório

- Preenche um intervalo de células com um conjunto de números aleatórios independentes obtidos numa das várias distribuições

Ordem e Percentil

- Produz uma tabela que contém a posição ordinal e de percentagem de cada valor num conjunto de dados
- É possível analisar a posição relativa de valores num conjunto de dados

Regressão

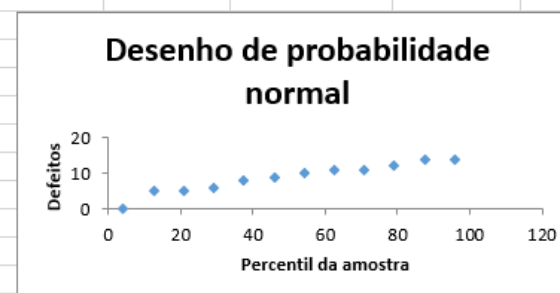
- Executa uma análise de regressão linear, utilizando o método dos quadrados mínimos para ajustar uma linha ao longo de um conjunto de observações
- Torna possível analisar como uma única variável dependente é afetada pelos valores de uma ou mais variáveis independentes

Regressão

Produção	Defeitos
922	11
1189	10
1109	5
1152	8
893	11
846	12
802	9
1110	14
1044	6
1117	14
1000	5
1065	0

SUMÁRIO DOS RESULTADOS								
Estatística de regressão								
R múltiplo	0,129030824							
Quadrado de R	0,016648953							
Quadrado de R ajustado	-0,081686151							
Erro-padrão	4,325303881							
Observações	12							
ANOVA								
	gl	SQ	MQ	F	F de significância			
Regressão	1	3,167463391	3,167463391	0,169308341	0,689407295			
Residual	10	187,0825366	18,70825366					
Total	11	190,25						
	Coefficientes	Erro-padrão	Stat t	valor P	95% inferior	95% superior	Inferior 95,0%	Superior 95,0%
Interceptar	13,05661106	10,54059388	1,238697858	0,243745278	-10,42929569	36,54251782	-10,42929569	36,54251782
Produção	-0,004219065	0,010253617	-0,411470948	0,689407295	-0,027065548	0,018627417	-0,027065548	0,018627417

RESULTADOS DE PROBABILIDADE	
Percentil	Defeitos
4,166666667	0
12,5	5
20,83333333	5
29,16666667	6
37,5	8
45,83333333	9
54,16666667	10
62,5	11
70,83333333	11
79,16666667	12
87,5	14
95,83333333	14



Apenas 1% da variação na
variável dependente

Defeitos é explicada pela
variação na variável
independente **Produção**

Regressão

Produção	Defeitos
922	92
1189	119
1109	111
1152	115
893	89
846	85
802	80
1110	111
1044	104
1117	112
1000	100
1065	107

SUMÁRIO DOS RESULTADOS

Estatística de regressão	
R múltiplo	0,999760629
Quadrado de R	0,999521316
Quadrado de R ajustado	0,999473447
Erro-padrão	0,293120834
Observações	12

ANOVA

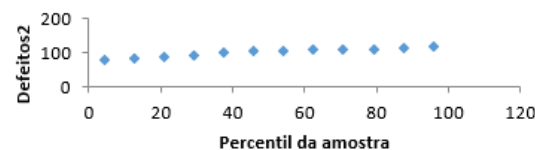
	gl	SQ	MQ	F	F de significância
Regressão	1	1794,057468	1794,057468	20880,60009	6,18632E-18
Residual	10	0,859198232	0,085919823		
Total	11	1794,916667			

	Coeficientes	Erro-padrão	Stat t	valor P	95% inferior	95% superior	Inferior 95,0%	Superior 95,0%
Interceptar	-0,410568129	0,714323838	-0,574764704	0,57815115	-2,002180825	1,181044567	-2,002180825	1,181044567
Produção	0,100410386	0,000694876	144,5012114	6,18632E-18	0,098862106	0,101958666	0,098862106	0,101958666

RESULTADOS DE PROBABILIDADE

Percentil	Defeitos2
4,166666667	80
12,5	85
20,83333333	89
29,16666667	92
37,5	100
45,83333333	104
54,16666667	107
62,5	111
70,83333333	111
79,16666667	112
87,5	115
95,83333333	119

Desenho de probabilidade normal



99,9% da variação na
variável dependente
Defeitos é explicada pela
variação na variável
independente **Produção**

Amostragem

- Permite criar uma amostra de uma população, tratando o intervalo de entrada como a população em estudo
- A amostragem é bastante útil quando uma população em estudo é demasiado grande para ser processada e analisada
- Permite criar amostras aleatórias e periódicas

Teste T

Duas Amostras Emparelhadas para Médias

- Utilizado quando existe um emparelhamento natural de observações nas amostras, tal como quando um grupo de amostras é testado duas vezes, antes e após uma experiência
- Executa um teste T de *Student* para determinar se é provável que as observações efetuadas antes e depois de um tratamento resultem de distribuições com médias de população iguais.
- Não parte do princípio de que as variâncias de ambas as populações são iguais

Duas Amostras com Variâncias Iguais

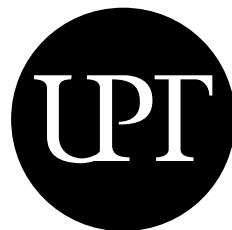
- Executa um teste T de *Student* de duas amostras
- Parte do princípio que os dois conjuntos de dados provêm de distribuições com as mesmas variâncias
- Pode ser utilizado para determinar se é provável que as duas amostras tenham resultado de distribuições com médias de população iguais

Duas Amostras com Variâncias Desiguais

- Executa um teste T de *Student* de duas amostras
- Parte do princípio que os dois conjuntos de dados provêm de distribuições com variâncias desiguais
- Pode ser utilizado para determinar se é provável que as duas amostras tenham resultado de distribuições com médias de população iguais

Teste Z

- Executa um teste Z de duas amostras para médias com variâncias conhecidas
- Utilizada para testar a hipótese nula de que não existe diferença entre as duas médias de população, contra as hipóteses alternativas unilaterais ou bilaterais
- Se as variâncias não forem conhecidas, a função TESTEZ deverá ser usada em alternativa



UNIVERSIDADE
PORTUCALENSE

Do conhecimento à prática.