

Econometria I: Aplicações

Laboratório 3 – Análise estatística básica no Microsoft Excel

Profº Lindomar Pegorini Daniel¹

¹ Professor Adjunto da Universidade do Estado de Mato Grosso (UNEMAT) – Campus de Sinop.

LABORATÓRIO 3 – ANÁLISE ESTATÍSTICA BÁSICA NO MICROSOFT EXCEL

Visão geral

Nos laboratórios anteriores, você explorou um conjunto de dados contendo detalhes das vendas de limonada. Neste laboratório, você aplicará algumas técnicas de análise estatística para obter um melhor entendimento dos dados.

Do que você vai precisar

Para completar este laboratório você irá precisar:

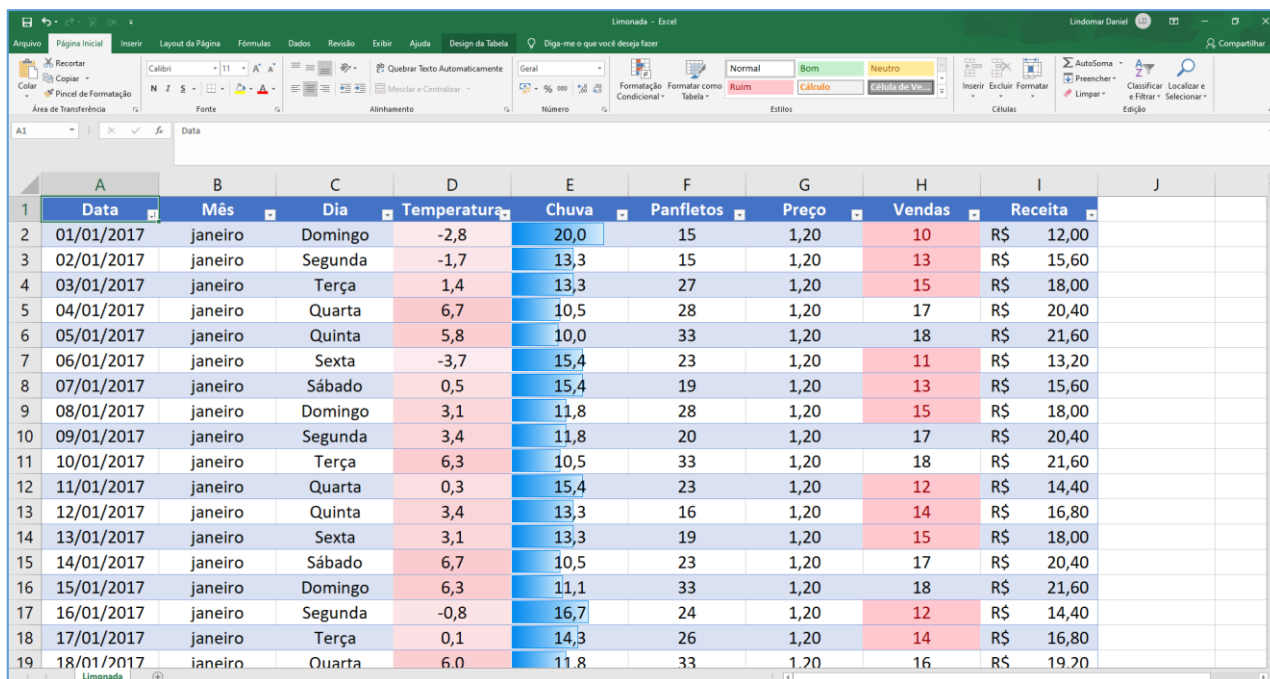
- Do arquivo de dados **Limonada.xlsx** dos laboratórios anteriores;
- De um computador com um aplicativo gerenciador de planilhas compatível com a extensão **xlsx**, ou, de um computador com acesso à internet e uma conta Microsoft (hotmail.com, live.com ou outlook.com) para acessar o Excel Online de forma gratuita.

EXERCÍCIO 1: Usando estatísticas descritivas

Estatísticas descritivas ajudam você a entender a “forma” ou a distribuição dos dados; por exemplo, encontrando medidas de tendência central (os valores “típicos” mais comuns) e medidas de variância (quanta diferença existe entre os valores mais comuns e outros valores que são maiores ou menores).

Calculando estatísticas descritivas para as vendas

1. Ao abrir o arquivo Limonada.xlsx, ele deveria parecer com o seguinte:



	A	B	C	D	E	F	G	H	I	J
1	Data	Mês	Dia	Temperatura	Chuva	Panfletos	Preço	Vendas	Receita	
2	01/01/2017	janeiro	Domingo	-2,8	20,0	15	1,20	10	R\$ 12,00	
3	02/01/2017	janeiro	Segunda	-1,7	13,3	15	1,20	13	R\$ 15,60	
4	03/01/2017	janeiro	Terça	1,4	13,3	27	1,20	15	R\$ 18,00	
5	04/01/2017	janeiro	Quarta	6,7	10,5	28	1,20	17	R\$ 20,40	
6	05/01/2017	janeiro	Quinta	5,8	10,0	33	1,20	18	R\$ 21,60	
7	06/01/2017	janeiro	Sexta	-3,7	15,4	23	1,20	11	R\$ 13,20	
8	07/01/2017	janeiro	Sábado	0,5	15,4	19	1,20	13	R\$ 15,60	
9	08/01/2017	janeiro	Domingo	3,1	11,8	28	1,20	15	R\$ 18,00	
10	09/01/2017	janeiro	Segunda	3,4	11,8	20	1,20	17	R\$ 20,40	
11	10/01/2017	janeiro	Terça	6,3	10,5	33	1,20	18	R\$ 21,60	
12	11/01/2017	janeiro	Quarta	0,3	15,4	23	1,20	12	R\$ 14,40	
13	12/01/2017	janeiro	Quinta	3,4	13,3	16	1,20	14	R\$ 16,80	
14	13/01/2017	janeiro	Sexta	3,1	13,3	19	1,20	15	R\$ 18,00	
15	14/01/2017	janeiro	Sábado	6,7	10,5	23	1,20	17	R\$ 20,40	
16	15/01/2017	janeiro	Domingo	6,3	11,1	33	1,20	18	R\$ 21,60	
17	16/01/2017	janeiro	Segunda	-0,8	16,7	24	1,20	12	R\$ 14,40	
18	17/01/2017	janeiro	Terça	0,1	14,3	26	1,20	14	R\$ 16,80	
19	18/01/2017	janeiro	Quarta	6,0	11,8	33	1,20	16	R\$ 19,20	

2. Na célula **K1**, insira o texto **Estatísticas de vendas** e formate-o como negrito.
3. Na célula **K2**, insira o texto **Média** e, em seguida, na célula **L2**, insira a seguinte fórmula:
= MÉDIA (H2: H366)
Isso calcula a média aritmética das vendas, que deve ser ligeiramente superior a 25,3.
4. Na célula **K3**, insira o texto **Mediana** e, na célula **L3**, insira a seguinte fórmula:
= MED (H2: H366)
Isso calcula a mediana de vendas, que deve ser 25.
5. Na célula **K4**, digite o texto **Moda** e, na célula **L4**, insira a seguinte fórmula:
= MODO.ÚNICO (H2: H366)
Isso calcula a moda de vendas, que deve ser 25.
6. Na célula **K5**, insira o texto **Variância** e, na célula **L5**, insira a seguinte fórmula:
= VAR.P (H2: H366)
Isso calcula a variação das vendas, que deve ser um pouco acima de 47,39.
Observe que a fórmula para a **variância** neste caso aplica-se à população completa de dados, por isso a extensão **.P** no nome da função - você explorará amostras de dados posteriormente neste laboratório.
7. Na célula **K6**, insira o texto **Desv Pad** e, na célula **L6**, insira a seguinte fórmula:
= DESVPAD.P (H2: H366)
Isso calcula o desvio padrão das vendas, que deve ser ligeiramente superior a 6,88.
Observe mais uma vez que a fórmula para desvio padrão neste caso aplica-se ao total da população de dados.

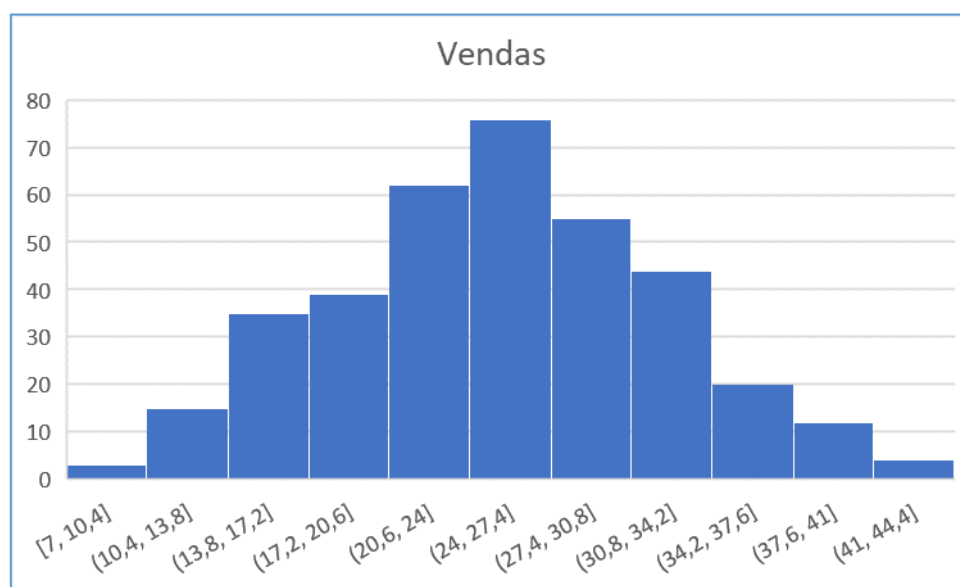
Sua planilha agora deve ficar assim:

	D	E	F	G	H	I	J	K	L	M
1	Temperatura	Chuva	Panfletos	Preço	Vendas	Receita		Estatísticas de vendas		
2	-2,8	20,0	15	1,20	10	R\$ 12,00		Média	25,32	
3	-1,7	13,3	15	1,20	13	R\$ 15,60		Mediana	25	
4	1,4	13,3	27	1,20	15	R\$ 18,00		Moda	25	
5	6,7	10,5	28	1,20	17	R\$ 20,40		Variância	47,39	
6	5,8	10,0	33	1,20	18	R\$ 21,60		Desv Pad	6,88	
7	-3,7	15,4	23	1,20	11	R\$ 13,20				
8	0,5	15,4	19	1,20	13	R\$ 15,60				
9	3,1	11,8	28	1,20	15	R\$ 18,00				
10	3,4	11,8	20	1,20	17	R\$ 20,40				
11	6,3	10,5	33	1,20	18	R\$ 21,60				
12	0,3	15,4	23	1,20	12	R\$ 14,40				
13	3,4	13,3	16	1,20	14	R\$ 16,80				
14	3,1	13,3	19	1,20	15	R\$ 18,00				
15	6,7	10,5	23	1,20	17	R\$ 20,40				
16	6,3	11,1	33	1,20	18	R\$ 21,60				
17	-0,8	16,7	24	1,20	12	R\$ 14,40				
18	0,1	14,3	26	1,20	14	R\$ 16,80				
19	6,0	11,8	33	1,20	16	R\$ 19,20				

As estatísticas que você calculou dizem algo sobre a distribuição dos valores de vendas, mas pode ser mais fácil visualizar os dados para ter uma ideia de como os dados são distribuídos.

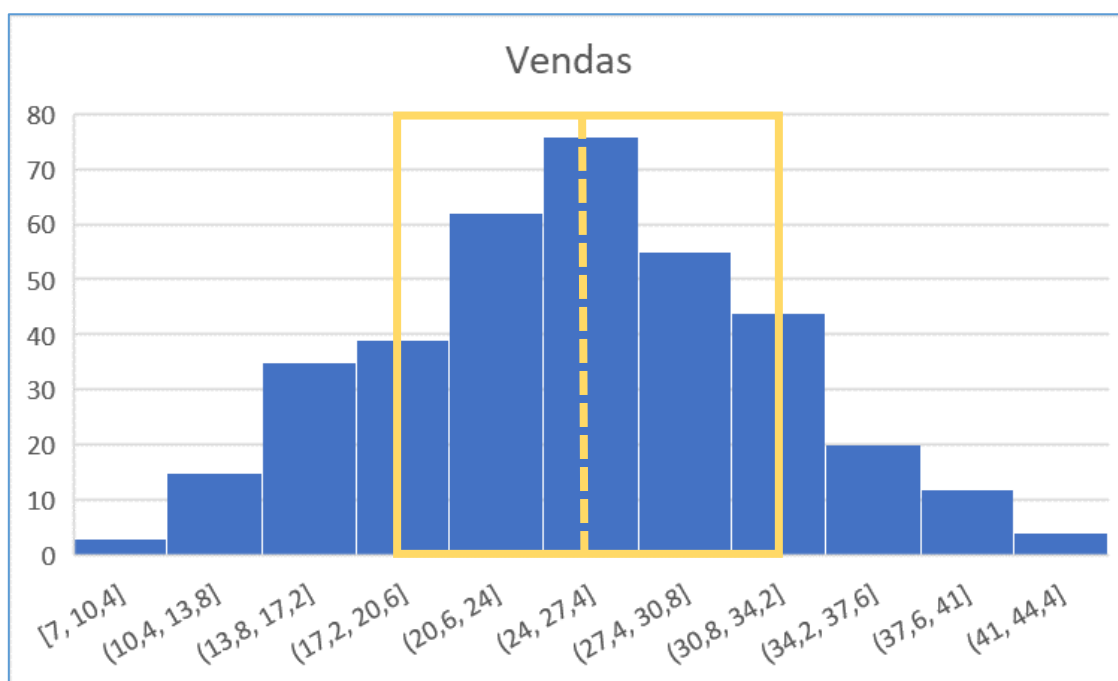
Visualize a distribuição dos valores de vendas

1. Selecione todos os dados na coluna **Vendas**, incluindo o cabeçalho. Em seguida, na guia **Inserir** da faixa de opções, na lista suspensa **Gráficos**, clique no gráfico **Histograma** (que é o primeiro na seção **Inserir Gráfico de Estatística**).
2. Selecione o gráfico que é produzido e edite o título do gráfico para alterá-lo para **Vendas**. Então mova o gráfico para que fique à direita das estatísticas que você calculou no exercício anterior. O gráfico deve ficar assim:



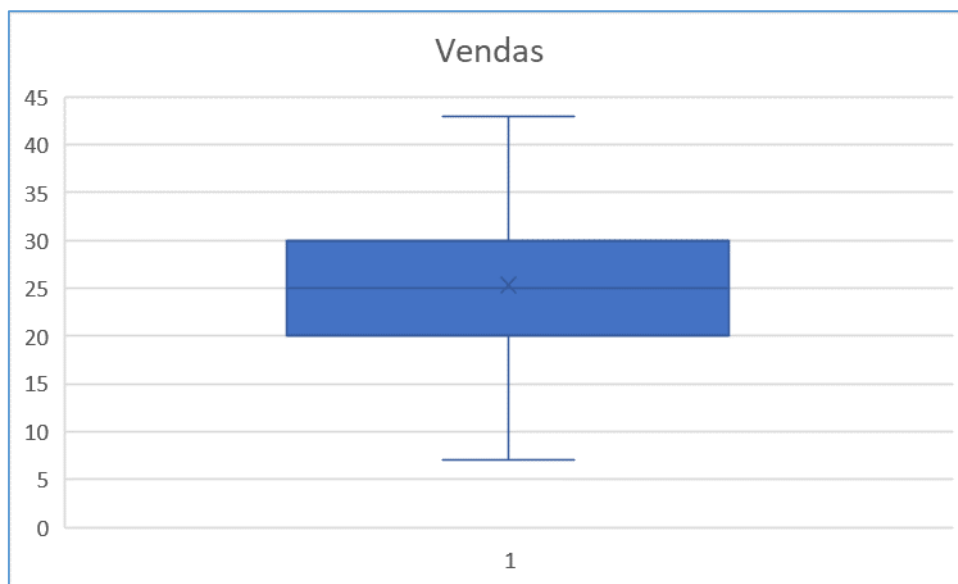
3. Examine o gráfico e observe o seguinte:

- O histograma mostra a frequência de valores diferentes para valores de vendas agrupados em intervalos ou caixas. Por exemplo, há cerca de 15 dias com um valor de vendas entre 10,4 e 13,8; e há cerca de 20 dias com um valor de vendas entre 34,2 e 37,6.
- Os valores de vendas mais frequentes estão entre 24 e 27,5. Esse intervalo inclui as estatísticas de média, mediana e moda que você calculou anteriormente. Em outras palavras, na maioria dos dias, o número de vendas foi mais ou menos no meio dos dias com menor e o maior número de vendas.
- A distribuição é aproximadamente simétrica em torno dos valores médios, formando um “sino” que afila uniformemente em direção às extremidades; onde há poucas ocorrências de valores extremos para vendas. Os estatísticos referem-se a este tipo de distribuição como uma distribuição normal.
- O desvio padrão calculado anteriormente é de pouco menos de 6,9. Esta estatística fornece uma unidade padrão de variação em torno da média (que é um pouco acima de 25,3). Os dados dentro de 1 desvio-padrão acima ou abaixo da média, incluem, portanto, valores de aproximadamente 18,4 a 32,2, como mostrado aqui:



Em uma distribuição normal, cerca de 68,26% dos dados se enquadram em um único desvio padrão; assim, neste caso, o número de vendas foi entre 18,4 e 32,2 em 68,26% dos dias. Cerca de 95,45% dos valores estão dentro de 2 desvios padrão em uma distribuição normal, portanto, havia entre 11,5 e 39,1 vendas em 95,4% dos dias.

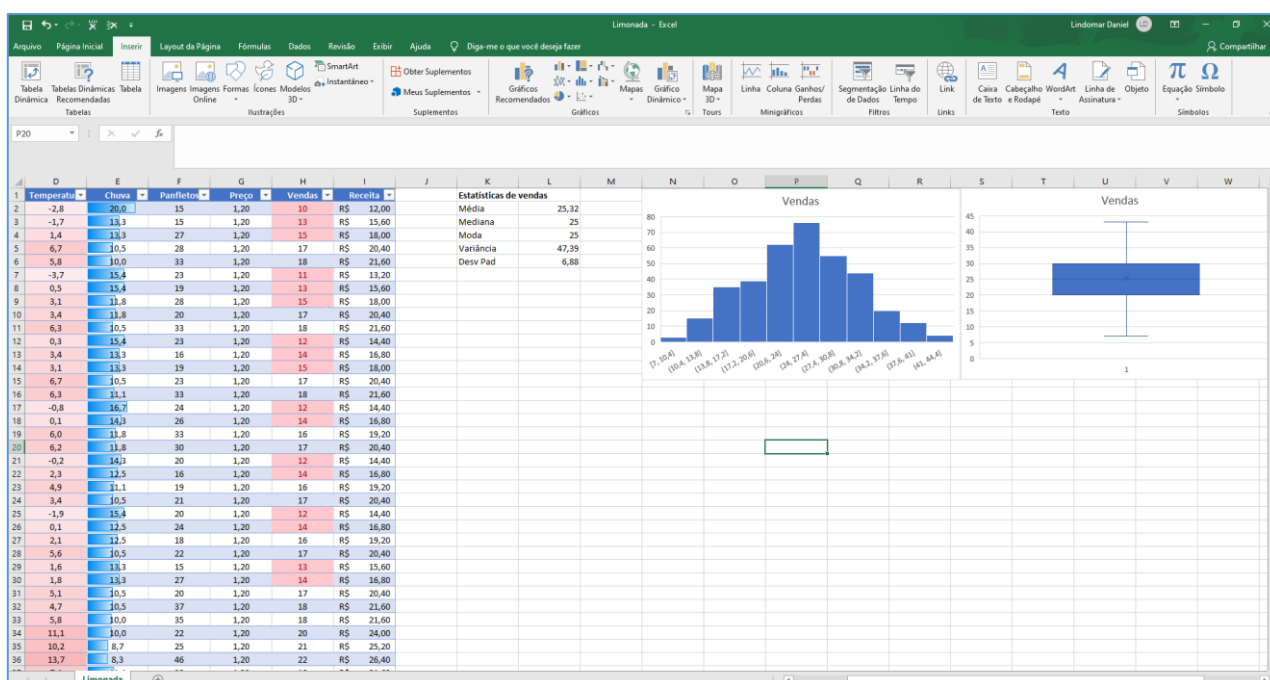
4. Selecione todos os dados na coluna **Vendas**, incluindo o cabeçalho. Em seguida, na guia **Inserir** da faixa de opções, na lista suspensa **Gráficos**, clique no gráfico **Caixa Estreita** ou Boxplot (que é o terceiro na seção **Inserir Gráfico de Estatística**).
5. Selecione o gráfico que é produzido e edite o título do gráfico para alterá-lo para **Vendas**. Então mova o gráfico para que fique à direita do histograma que você criou anteriormente. O gráfico deve se parecer com esse:



6. Examine o gráfico e observe o seguinte:

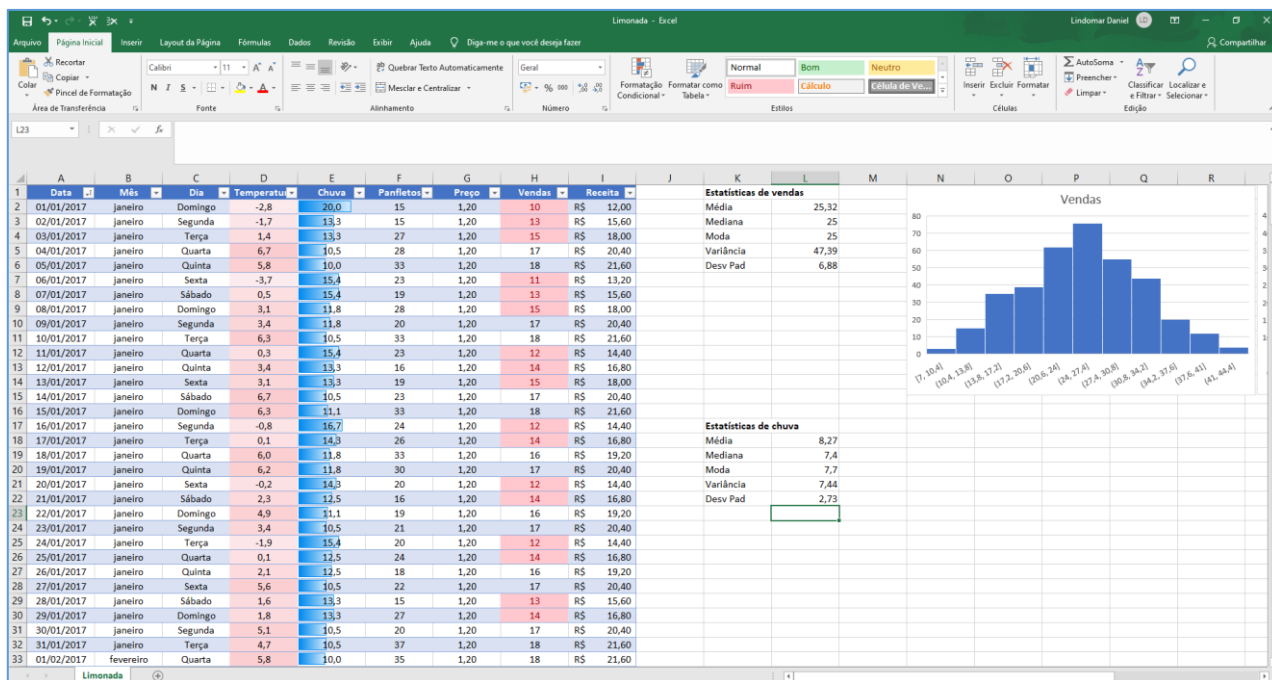
- A linha horizontal no meio indica o valor mediano para vendas. Esse é o percentil de 50%, em outras palavras, 50% dos valores são maiores que isso e 50% são menores.
- O X na caixa indica a média - ela é ligeiramente superior à mediana.
- A caixa preenchida indica o intervalo de valores no segundo e terceiro quartis - em outras palavras, do percentil 25 ao percentil 75. Esses valores variam entre 20 a 30, indicando que o número de vendas na metade dos dias estava dentro dessa faixa,
- As linhas que se estendem da caixa (conhecidas como bigodes) mostram o intervalo para o primeiro e quarto quartis, no qual houve um número maior ou menor de vendas que no segundo e terceiro quartis.

Agora, sua planilha deve ser semelhante a essa (talvez seja necessário rolar para a direita para ver os gráficos.)



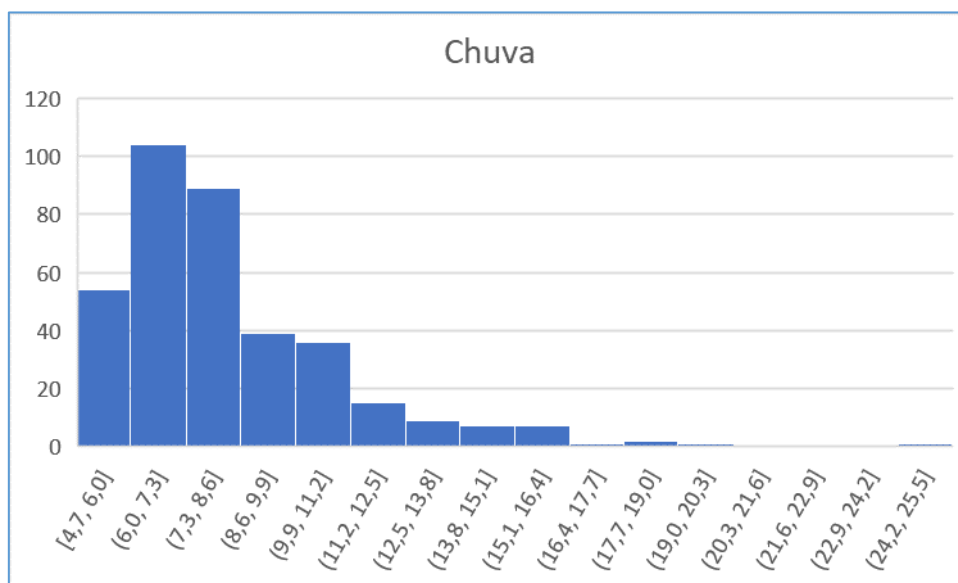
Análise as estatísticas para chuva

1. Deixe algum espaço sob as estatísticas existentes e, na célula **K17**, insira o texto **Estatísticas de Chuva** e formate-o como negrito.
2. Na célula **K18**, insira o texto Média e, em seguida, na célula **L18**, insira a seguinte fórmula: = MÉDIA (E2: E366)
Isso calcula a média aritmética da precipitação, que deve ser de 8,27.
3. Nas células **K19** a **L22**, calcule a mediana, a moda, a variância populacional e o desvio padrão populacional para a chuva, da mesma forma que você fez para as vendas. Quando terminar, sua planilha deve ficar assim:

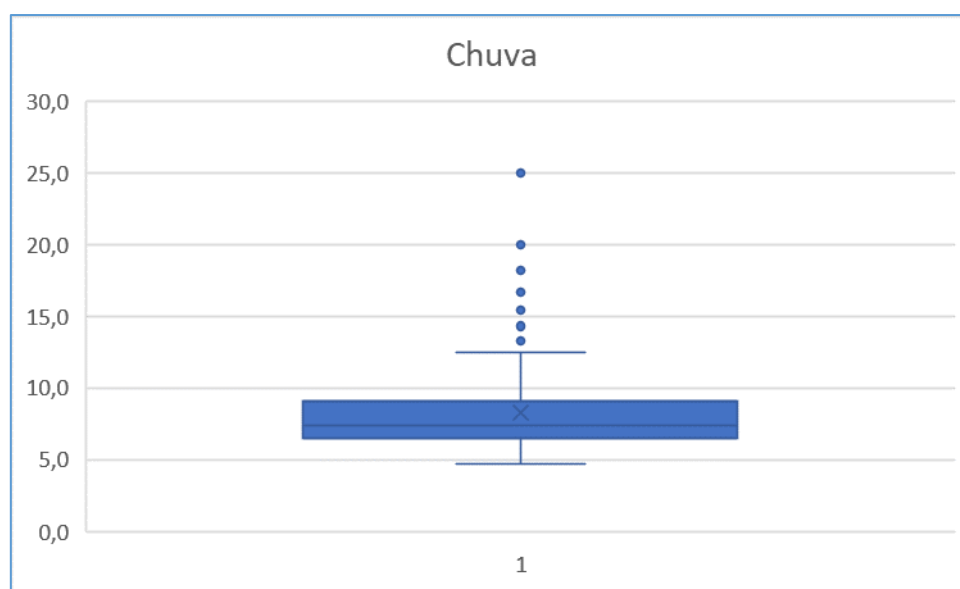


Note que a média de chuva é pouco superior a moda e a mediana.

4. Crie um histograma para chuva e, em seguida, adicione um título de gráfico apropriado e mova o gráfico para a direita das estatísticas de precipitação. O histograma da chuva deve ficar assim:



- Examine o histograma e observe que a distribuição dos dados de precipitação não é normal. O valor mediano está em torno de 7,4, assim, na metade dos dias houve menos chuva do que isso, e na metade houve mais. No entanto, em alguns dias raros, houve muito mais chuva do que isso - entre 24,2 a 25,5. Esses dias infrequentes de chuvas extremamente altas estão distorcendo a distribuição “puxando” a média para a direita. Isso resulta em uma cauda longa de valores que ocorrem raramente e que se inclinam para a direita. Nos referimos a isso como uma distribuição com assimetria à direita (se a cauda tivesse puxado a média para a esquerda, seria uma distribuição com assimetria à esquerda).
- Crie uma Caixa Estreita ou boxplot com um título apropriado; e mova-o para a direita do histograma de chuva. O gráfico de Caixa Estreita deve ficar assim:



- Examine o gráfico e observe o seguinte:
 - A linha que indica a mediana (percentil 50) está visivelmente mais baixa que o X, que indica a média.
 - A caixa preenchida, representando o 2º e 3º quartis, representa 50% dos dados - ou seja, na metade dos dias, a precipitação ficou entre 6,0 e 8,0.
 - Os pontos indicam outliers; valores raros que são considerados extremos em comparação com a faixa típica de valores, que está dentro dos bigodes.
 - Mesmo descontando os outliers, a faixa de valores no quarto quartil é maior do que a dos outros quartis.

Desafio: Analise as estatísticas de temperatura

- Calcule a média, a mediana, a moda, a variância populacional e o desvio padrão populacional da temperatura.
- Crie um histograma e uma caixa estreita (boxplot) para visualizar a distribuição da temperatura.
- Responda a seguinte questão do Enade:

No Brasil, o Banco Central procura alcançar uma meta de inflação com um valor central, por exemplo, de 4% a.a. em torno da qual se define uma faixa tolerável de $\pm 2\%$ a.a. Se tal intervalo ocorrer de fato e as probabilidades puderem ser aproximadas pela distribuição uniforme entre + 2% e + 6%,

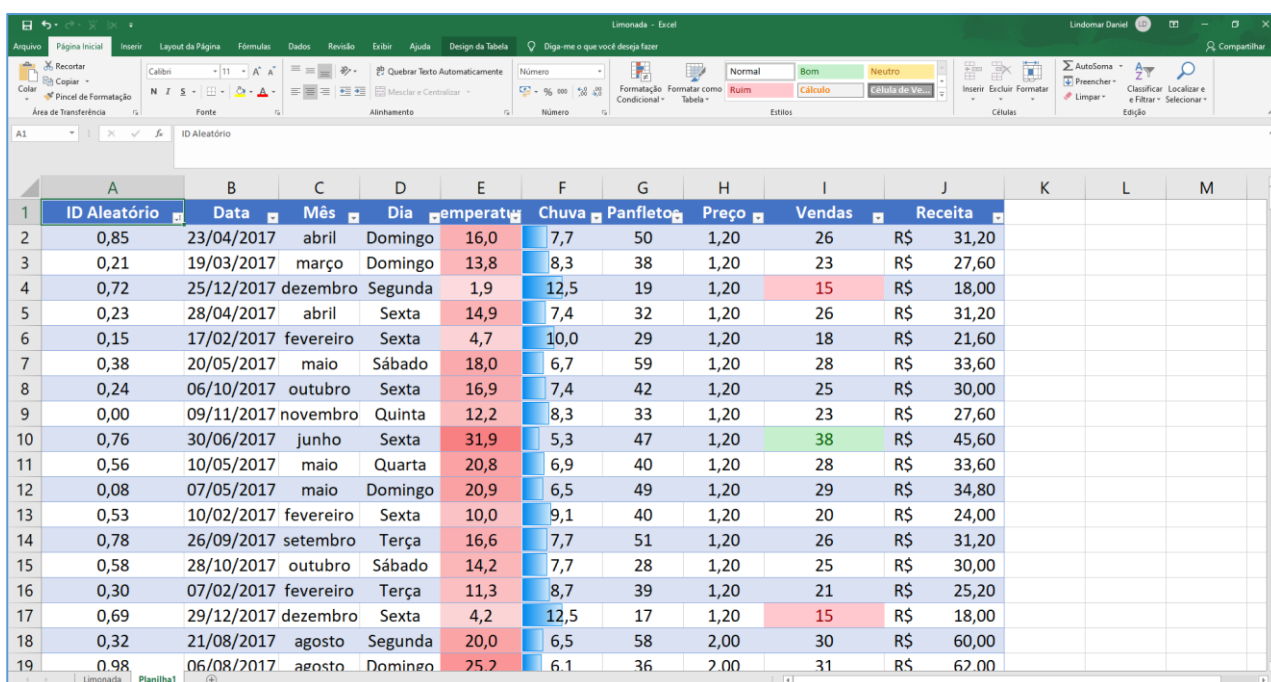
- (A) a probabilidade de que a inflação venha a se situar entre 2 e 2.5% a.a. será de 10%.
 - (B) a mediana da distribuição será 3.5% a.a.
 - (C) a distribuição será bimodal.
 - (D) a amplitude total da distribuição será de 2% a.a.
 - (E) o quartil inferior da distribuição irá de 2 a 3% a.a.
4. Lembre-se de anotar essas informações, elas serão úteis ao responder os exercícios posteriormente.

EXERCÍCIO 2: Trabalhando com amostras

Até agora, trabalhamos com toda a população de dados - em outras palavras, todos os dados de vendas de limonada de Rosie eram usados. Na realidade, é mais comum trabalhar com uma amostra de dados. Por exemplo, suponha que você precise realizar alguma pesquisa para determinar a cor dos olhos mais comum no Brasil. Seria irreal examinar os olhos de todas as pessoas do Brasil, então você abordaria esse problema pesquisando uma amostra representativa de pessoas e usando as estatísticas coletadas como aproximações para os parâmetros da população.

Criando uma amostra aleatória

1. Na planilha **Limonada**, clique na célula **A1** e pressione CTRL + T para selecionar a tabela inteira de dados de vendas da limonada. Em seguida, na guia **Página Inicial** da faixa de opções, clique em **Copiar**.
2. Adicione uma nova planilha e clique na célula **A1**. Em seguida, na guia **Página Inicial** da faixa de opções, clique em **Colar** para colar a tabela copiada na nova planilha.
3. Clique na célula **A1** (cabeçalho de **Data**) e, na guia **Página Inicial** da faixa de opções, na seção **Células**, clique na lista suspensa **Inserir** e selecione **Inserir Colunas de Tabela à Esquerda**. Isso insere uma nova coluna para um campo de tabela chamado **Coluna1**.
4. Na nova célula **A1**, renomeie **Coluna1** para **ID Aleatório** e, em seguida, selecione a coluna **A** e, na lista suspensa **Formato** na seção **Número** da guia **Início** da faixa de opções, clique em **Número**.
5. na célula **A2**, insira a seguinte fórmula:
`=ALEATÓRIO()`
Isso irá gerar números aleatórios na coluna **ID Aleatório**.
6. Clique na seta suspensa no cabeçalho da coluna **ID Aleatório** e clique em **Classificar do Menor para o Maior**. Os dados na tabela são então classificados pelo campo **ID Aleatório**, que torna aleatória a ordem dos registros de dados. Isso facilita a seleção de uma amostra aleatória de registros (é mais provável que uma amostra aleatória seja representativa da população do que uma amostra baseada em alguma ordem inerente aos próprios dados). Sua planilha agora deve parecer com a seguinte, contudo não será idêntica devido ao sorteio aleatório dos dados:



	A	B	C	D	E	F	G	H	I	J	K	L	M
1	ID Aleatório	Data	Mês	Dia	temperatura	Chuva	Panfletos	Preço	Vendas	Receita			
2	0,85	23/04/2017	abril	Domingo	16,0	7,7	50	1,20	26	R\$ 31,20			
3	0,21	19/03/2017	março	Domingo	13,8	8,3	38	1,20	23	R\$ 27,60			
4	0,72	25/12/2017	dezembro	Segunda	1,9	12,5	19	1,20	15	R\$ 18,00			
5	0,23	28/04/2017	abril	Sexta	14,9	7,4	32	1,20	26	R\$ 31,20			
6	0,15	17/02/2017	fevereiro	Sexta	4,7	10,0	29	1,20	18	R\$ 21,60			
7	0,38	20/05/2017	maio	Sábado	18,0	6,7	59	1,20	28	R\$ 33,60			
8	0,24	06/10/2017	outubro	Sexta	16,9	7,4	42	1,20	25	R\$ 30,00			
9	0,00	09/11/2017	novembro	Quinta	12,2	8,3	33	1,20	23	R\$ 27,60			
10	0,76	30/06/2017	junho	Sexta	31,9	5,3	47	1,20	38	R\$ 45,60			
11	0,56	10/05/2017	maio	Quarta	20,8	6,9	40	1,20	28	R\$ 33,60			
12	0,08	07/05/2017	maio	Domingo	20,9	6,5	49	1,20	29	R\$ 34,80			
13	0,53	10/02/2017	fevereiro	Sexta	10,0	9,1	40	1,20	20	R\$ 24,00			
14	0,78	26/09/2017	setembro	Terça	16,6	7,7	51	1,20	26	R\$ 31,20			
15	0,58	28/10/2017	outubro	Sábado	14,2	7,7	28	1,20	25	R\$ 30,00			
16	0,30	07/02/2017	fevereiro	Terça	11,3	8,7	39	1,20	21	R\$ 25,20			
17	0,69	29/12/2017	dezembro	Sexta	4,2	12,5	17	1,20	15	R\$ 18,00			
18	0,32	21/08/2017	agosto	Segunda	20,0	6,5	58	2,00	30	R\$ 60,00			
19	0,98	06/08/2017	agosto	Domingo	25,2	6,1	36	2,00	31	R\$ 62,00			

7. Na célula **M1**, insira o texto **Média de Chuva** e, na célula **N1**, insira o texto **Desv. Pad. da Chuva**.
8. Na célula **L2**, insira o texto **População**.
9. Em seguida, na célula **M2**, insira a seguinte fórmula:
= MÉDIA (F2: F366)
10. Na célula **N2**, insira a seguinte fórmula:
= DESVPAD.P (F2: F366)

Isso fornece parâmetros completos da população para a média e o desvio padrão da chuva, para que você possa compará-los com as estatísticas da amostra. Sua planilha agora deve ficar assim:

The screenshot shows an Excel spreadsheet with the following data:

	D	E	F	G	H	I	J	K	L	M	N	O	P
1	Dia	temperatura	Chuva	Panfleto	Preço	Vendas	Receita			Média chuva	Desv Pad Chuva		
2	Domingo	16,0	7,7	50	1,20	26	R\$ 31,20		População	8,3	2,73		
3	Domingo	13,8	8,3	38	1,20	23	R\$ 27,60						
4	Segunda	1,9	12,5	19	1,20	15	R\$ 18,00						
5	Sexta	14,9	7,4	32	1,20	26	R\$ 31,20						
6	Sexta	4,7	10,0	29	1,20	18	R\$ 21,60						
7	Sábado	18,0	6,7	59	1,20	28	R\$ 33,60						
8	Sexta	16,9	7,4	42	1,20	25	R\$ 30,00						
9	Quinta	12,2	8,3	33	1,20	23	R\$ 27,60						
10	Sexta	31,9	5,3	47	1,20	38	R\$ 45,60						
11	Quarta	20,8	6,9	40	1,20	28	R\$ 33,60						
12	Domingo	20,9	6,5	49	1,20	29	R\$ 34,80						
13	Sexta	10,0	9,1	40	1,20	20	R\$ 24,00						
14	Terça	16,6	7,7	51	1,20	26	R\$ 31,20						
15	Sábado	14,2	7,7	28	1,20	25	R\$ 30,00						
16	Terça	11,3	8,7	39	1,20	21	R\$ 25,20						
17	Sexta	4,2	12,5	17	1,20	15	R\$ 18,00						
18	Segunda	20,0	6,5	58	2,00	30	R\$ 60,00						
19	Domingo	25,2	6,1	36	2,00	31	R\$ 62,00						

11. Na célula **L3**, insira o texto **Amostra1**.
12. Em seguida, na célula **M3**, insira a seguinte fórmula:
= MÉDIA (F2: F41)
13. Na célula **N3**, insira a seguinte fórmula:
= DESVPAD.A (F2: F41)

Isso fornece estatísticas da amostra para a média e o desvio padrão da chuva com base nas primeiras 40 linhas aleatórias de dados. Observe que você usa a mesma função **MÉDIA** para calcular uma média de amostra ou média de população, mas usa a função **DESVPAD.A** para calcular o desvio padrão de uma amostra - isso incorpora algumas variações adicionais para permitir o viés de amostra. Agora, sua planilha deve ser semelhante a essa (os números podem não ser exatamente os mesmos, pois os dados foram sorteados de forma aleatória):

	D	E	F	G	H	I	J	K	L	M	N	O	P
1	Dia	temperatura	Chuva	Panfleto	Preço	Vendas	Receita			Média chuva	Desv Pad Chuva		
2	Domingo	16,0	7,7	50	1,20	26	R\$ 31,20		População	8,3	2,73		
3	Domingo	13,8	8,3	38	1,20	23	R\$ 27,60		Amostra1	8,0	1,99		
4	Segunda	1,9	12,5	19	1,20	15	R\$ 18,00						
5	Sexta	14,9	7,4	32	1,20	26	R\$ 31,20						
6	Sexta	4,7	10,0	29	1,20	18	R\$ 21,60						
7	Sábado	18,0	6,7	59	1,20	28	R\$ 33,60						
8	Sexta	16,9	7,4	42	1,20	25	R\$ 30,00						
9	Quinta	12,2	8,3	33	1,20	23	R\$ 27,60						
10	Sexta	31,9	5,3	47	1,20	38	R\$ 45,60						
11	Quarta	20,8	6,9	40	1,20	28	R\$ 33,60						
12	Domingo	20,9	6,5	49	1,20	29	R\$ 34,80						
13	Sexta	10,0	9,1	40	1,20	20	R\$ 24,00						
14	Terça	16,6	7,7	51	1,20	26	R\$ 31,20						
15	Sábado	14,2	7,7	28	1,20	25	R\$ 30,00						
16	Terça	11,3	8,7	39	1,20	21	R\$ 25,20						
17	Sexta	4,2	12,5	17	1,20	15	R\$ 18,00						
18	Segunda	20,0	6,5	58	2,00	30	R\$ 60,00						
19	Domingo	25,2	6,1	36	2,00	31	R\$ 62,00						

Compare as estatísticas da amostra com os parâmetros da população.

14. Na célula **L4**, insira o texto **Amostra2**.

15. Em seguida, na célula **M4**, digite a seguinte fórmula:

= MÉDIA (F35: F74)

16. Na célula **N4**, insira a seguinte fórmula:

= DESVPAD.A (F35: F74)

Isso produz estatísticas de uma amostra diferente. Observe que a proximidade das estatísticas da amostra com os parâmetros da população varia dependendo da amostra. Nesse caso, as duas amostras incluem 40 observações - o uso de amostras maiores geralmente resulta em estatísticas mais próximas dos parâmetros reais da população.

Criar uma distribuição amostral

1. Selecione as células **L3** a **N4** (as estatísticas **Amostra1** e **Amostra2** criadas anteriormente, mas não os parâmetros da população) e arraste o “quadrado” pequeno na parte inferior direita das células selecionadas até a linha 292. Isso cria 290 amostras como mostrado aqui:

	C	D	E	F	G	H	I	J	K	L	M	N	O
1	Mês	Dia	temperatura	Chuva	Panfleto	Preço	Vendas	Receita			Média chuva	Desv Pad Chuva	
2	abril	Domingo	16,0	7,7	50	1,20	26	R\$ 31,20		População	8,3	2,73	
3	março	Domingo	13,8	8,3	38	1,20	23	R\$ 27,60		Amostra1	8,0	1,99	
4	dezembro	Segunda	1,9	12,5	19	1,20	15	R\$ 18,00		Amostra2	8,2	2,76	
5	abril	Sexta	14,9	7,4	32	1,20	26	R\$ 31,20		Amostra3	8,0	2,01	
6	fevereiro	Sexta	4,7	10,0	29	1,20	18	R\$ 21,60		Amostra4	8,3	2,72	
7	maio	Sábado	18,0	6,7	59	1,20	28	R\$ 33,60		Amostra5	7,8	1,90	
8	outubro	Sexta	16,9	7,4	42	1,20	25	R\$ 30,00		Amostra6	8,3	2,71	
9	novembro	Quinta	12,2	8,3	33	1,20	23	R\$ 27,60		Amostra7	7,8	1,94	
10	junho	Sexta	31,9	5,3	47	1,20	38	R\$ 45,60		Amostra8	8,3	2,67	
11	maio	Quarta	20,8	6,9	40	1,20	28	R\$ 33,60		Amostra9	8,0	2,29	
12	maio	Domingo	20,9	6,5	49	1,20	29	R\$ 34,80		Amostra10	8,4	2,70	
13	fevereiro	Sexta	10,0	9,1	40	1,20	20	R\$ 24,00		Amostra11	8,1	2,29	
14	setembro	Terça	16,6	7,7	51	1,20	26	R\$ 31,20		Amostra12	8,4	2,70	
15	outubro	Sábado	14,2	7,7	28	1,20	25	R\$ 30,00		Amostra13	8,2	2,41	
16	fevereiro	Terça	11,3	8,7	39	1,20	21	R\$ 25,20		Amostra14	8,5	2,64	
17	dezembro	Sexta	4,2	12,5	17	1,20	15	R\$ 18,00		Amostra15	8,1	2,45	
18	agosto	Segunda	20,0	6,5	58	2,00	30	R\$ 60,00		Amostra16	8,7	2,75	
19	agosto	Domingo	25,2	6,1	36	2,00	31	R\$ 62,00		Amostra17	8,2	2,45	

As médias das amostras formam uma **distribuição amostral da média** - em outras palavras, uma nova distribuição de dados que consiste nas médias amostrais.

- Na célula **O1**, insira o texto **Média de amostragem**. Em seguida, na célula **O2**, insira a seguinte fórmula:

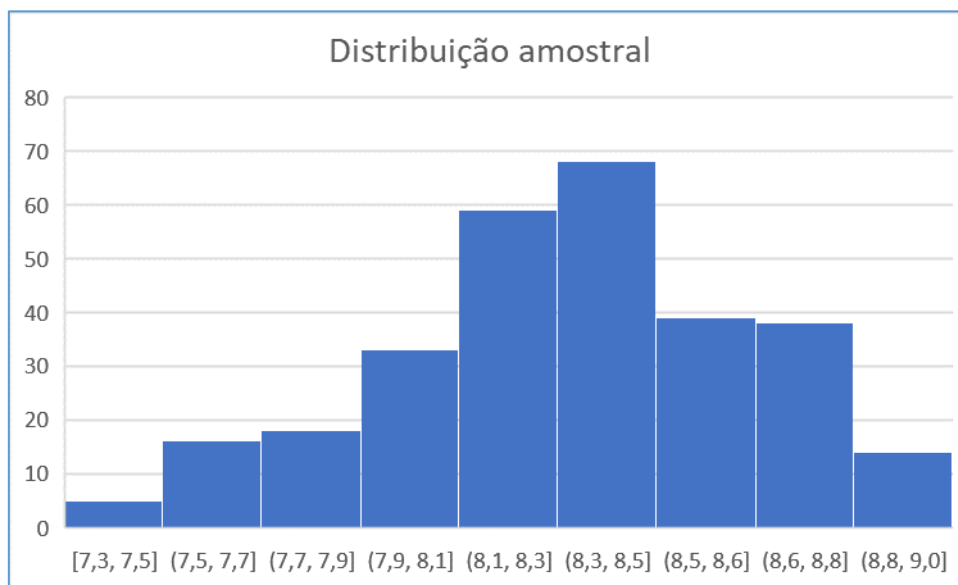
=MÉDIA (M3: M292)

Isso calcula a média das médias da amostra, em outras palavras, a média da distribuição amostral. Isso deve ser bastante próximo da média da população, como mostrado aqui (o seu pode não ser exatamente o mesmo que a média da população, mas deve ser próximo!)

	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	temperatura	Chuva	Panfleto	Preço	Vendas	Receita			Média chuva	Desv Pad Chuva	Média de amostragem		
2	16,0	7,7	50	1,20	26	R\$ 31,20		População	8,3	2,73	8,3		
3	13,8	8,3	38	1,20	23	R\$ 27,60		Amostra1	8,0	1,99			
4	1,9	12,5	19	1,20	15	R\$ 18,00		Amostra2	8,2	2,76			
5	14,9	7,4	32	1,20	26	R\$ 31,20		Amostra3	8,0	2,01			
6	4,7	10,0	29	1,20	18	R\$ 21,60		Amostra4	8,3	2,72			
7	18,0	6,7	59	1,20	28	R\$ 33,60		Amostra5	7,8	1,90			
8	16,9	7,4	42	1,20	25	R\$ 30,00		Amostra6	8,3	2,71			
9	12,2	8,3	33	1,20	23	R\$ 27,60		Amostra7	7,8	1,94			
10	31,9	5,3	47	1,20	38	R\$ 45,60		Amostra8	8,3	2,67			
11	20,8	6,9	40	1,20	28	R\$ 33,60		Amostra9	8,0	2,29			
12	20,9	6,5	49	1,20	29	R\$ 34,80		Amostra10	8,4	2,70			
13	10,0	9,1	40	1,20	20	R\$ 24,00		Amostra11	8,1	2,29			
14	16,6	7,7	51	1,20	26	R\$ 31,20		Amostra12	8,4	2,70			
15	14,2	7,7	28	1,20	25	R\$ 30,00		Amostra13	8,2	2,41			
16	11,3	8,7	39	1,20	21	R\$ 25,20		Amostra14	8,5	2,64			
17	4,2	12,5	17	1,20	15	R\$ 18,00		Amostra15	8,1	2,45			
18	20,0	6,5	58	2,00	30	R\$ 60,00		Amostra16	8,7	2,75			
19	25,2	6,1	36	2,00	31	R\$ 62,00		Amostra17	8,2	2,45			

- Clique na célula **M3** (a média da **Amostra1**) e segure as teclas Shift e Ctrl e pressione a tecla de seta para baixo para selecionar todas as outras médias de amostra.

4. Na guia **Inserir** da faixa de opções, na lista suspensa **Gráficos**, clique no gráfico **Histograma** (que é o primeiro na seção **Inserir Gráfico de Estatística**) e visualize o histograma criado, como mostrado aqui (o seu pode parecer um pouco diferente):



O histograma pode não parecer exatamente simétrico; mas quando você cria uma distribuição amostral a partir de um número suficientemente grande de amostras de tamanho razoável, verá que ela tem uma aparência de sino. Nós vamos discutir isso mais adiante, mas é útil saber que com amostras aleatórias suficientes, uma distribuição amostral geralmente assume uma distribuição normal devido a algo chamado de teorema central do limite - mesmo quando (como neste caso), os dados populacionais dos quais as médias amostrais são derivadas não são normalmente distribuídos.

Desafio: Analise amostras de temperatura

1. Crie uma distribuição amostral com base em 290 amostras de temperatura média. Cada amostra deve ser baseada em 40 observações aleatórias.
2. Calcule a média da distribuição amostral de temperatura.
3. Lembre-se de que essas informações serão úteis ao responder os exercícios posteriormente.

EXERCÍCIO 3: Inferência estatística e teste de hipóteses

Até agora, exploramos estatísticas descritivas que mostram a distribuição dos dados em uma população completa ou em uma amostra. Estatísticas inferenciais, como o nome sugere, são usadas para fazer inferências, previsões ou chegar a conclusões, a partir de dados baseados em relações estatísticas entre campos (ou recursos) dos dados.

Calcule a correlação

1. Volte para a planilha **Limonada** e selecione a célula **A1** (o cabeçalho da coluna **Data**).
2. Na guia **Inserir** da faixa de opções, clique em **Tabela Dinâmica** e insira uma tabela dinâmica para a tabela de vendas de limonada em uma nova planilha.
3. No painel **Campos da Tabela Dinâmica**, arraste **Data** para a área **LINHAS** e arraste **Temperatura** e **Vendas** para a área **VALORES** para que sua planilha seja semelhante a esta:

	A	B	C	D	E	F	G	H
1								
2								
3	Rótulos de Linha	Soma de Temperatura	Soma de Vendas					
4	01/jan	-2,78	10					
5	02/jan	-1,72	13					
6	03/jan	1,39	15					
7	04/jan	6,72	17					
8	05/jan	5,78	18					
9	06/jan	-3,72	11					
10	07/jan	0,50	13					
11	08/jan	3,06	15					
12	09/jan	3,39	17					
13	10/jan	6,33	18					
14	11/jan	0,33	12					
15	12/jan	3,44	14					
16	13/jan	3,06	15					
17	14/jan	6,72	17					
18	15/jan	6,33	18					

4. Desloque-se para a parte inferior da tabela dinâmica até conseguir ver a linha **Total Geral**. Você precisa copiar os valores de temperatura e vendas, exceto este **Total Geral**, para uma nova planilha.
5. Clique na célula **B368**, que deve conter o valor de Temperatura para 31 de dezembro de 2017 (acima do **Total Geral**). Em seguida, pressione SHIFT + CTRL + ↑ para selecionar a coluna de valores de temperatura e, em seguida, pressione SHIFT + → para estender a seleção para incluir os valores de vendas. Finalmente, copie as células selecionadas para a área de transferência.
6. Crie uma nova planilha e, na nova planilha, clique na célula **A1** e cole os dados copiados. Em seguida, altere os cabeçalhos das colunas para **Temperatura** e **Vendas**, conforme mostrado aqui:

Salvamento Automático Limonada - Excel Lindomar Daniel

Arquivo Página Inicial Inserir Layout da Página Fórmulas Dados Revisão Exibir Ajuda Pesquisar

Calibri 11 A A

Quebrar Texto Automaticamente

Fonte Alinhamento Número Estilos

Formatação Condicional Formatar como Tabela Estilos de Célula Inserir Excluir Formatar Células

AutoSoma Preencher Limpar Classificar e Filtrar Localizar e Selecionar Edição

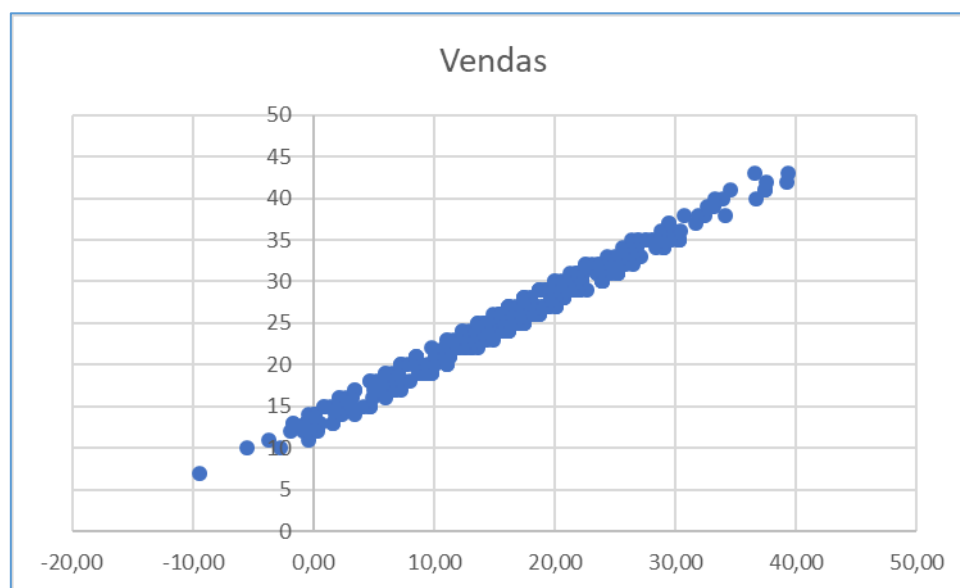
Área de Transferência

A1 Temperatura

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
1	Temperatura	Vendas																				
2	-2,78	10																				
3	-1,72	13																				
4	1,39	15																				
5	6,72	17																				
6	5,78	18																				
7	-3,72	11																				
8	0,50	13																				
9	3,06	15																				
10	3,39	17																				
11	6,33	18																				
12	0,33	12																				
13	3,44	14																				
14	3,06	15																				
15	6,72	17																				
16	6,33	18																				
17	-0,78	12																				
18	0,11	14																				
19	6,00	16																				
20	6,17	17																				
21	-0,22	12																				
22	2,33	14																				
23	4,89	16																				
24	3,39	17																				
25	-1,89	12																				
26	0,11	14																				
27	2,11	16																				
28	5,61	17																				
29	1,61	13																				

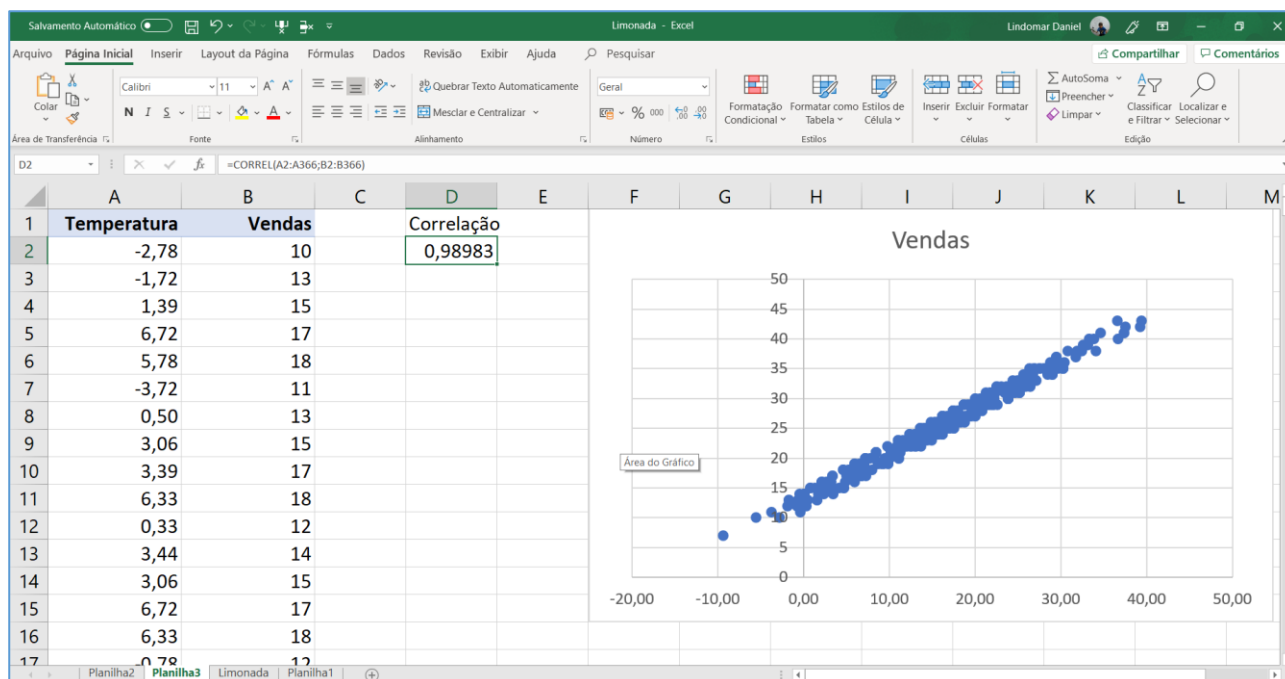
Planilha2 Planilha3 Limonada Planilha1

7. Selecione a célula **A1** (o cabeçalho da coluna **Temperatura**) e pressione CTRL + T para selecionar os dados e, na guia **Inserir** da faixa de opções, na lista suspensa **Dispersão**, clique em **Gráfico de Dispersão**. Isso insere um gráfico de dispersão que se parece com isso:



Observe que o gráfico de dispersão parece refletir uma relação linear entre temperatura e vendas, em que quanto maior a temperatura, maior o número de vendas.

8. Mova o gráfico para o lado da planilha para criar algum espaço e, em seguida, na célula **D1**, insira o texto **Correlação** e, na célula **D2**, insira a seguinte fórmula:
 =CORREL (A2: A366; B2: B366)
 Isso calcula a correlação entre temperatura e vendas; e deve produzir um valor em torno de 0,989832. Correlação é uma medida estatística da força de uma relação aparente entre duas variáveis numéricas - neste caso, temperatura e vendas.
9. Sua planilha deve ficar assim:



A correlação é medida como um valor entre -1 e 1. Um valor próximo de 1 indica uma correlação positiva; em outras palavras, valores altos para uma variável parecem corresponder a valores altos para a outra variável (padrão alto-alto baixo-baixo). Um valor próximo a -1, por outro lado, indica uma correlação negativa, na qual valores altos para uma variável correspondem a valores baixos para a outra variável (padrão alto-baixo baixo-alto). Um valor próximo de 0 indica a falta de qualquer relação discernível entre as variáveis.

Com uma correlação de quase 0,99, há uma forte relação positiva entre temperatura e vendas.

Nota: Estatísticos frequentemente citam o mantra “correlação não é causalidade”. Podemos usar a correlação para determinar que os dias com altos volumes de vendas tendem a ter altas temperaturas; mas não podemos dizer que Rosie vendeu muita limonada em um determinado dia porque a temperatura estava alta - assim como não podemos dizer que um dia foi particularmente quente porque Rosie vendeu muita limonada. Na verdade, pode haver um terceiro fator “oculto” que afeta tanto as vendas de limonada quanto a temperatura. Por exemplo, durante os meses de verão, quando os limões estão na estação, a temperatura tende a ser mais alta.

Desafio: Calcule Correlação de Chuva / Vendas

1. Calcule a correlação entre Chuva e Vendas.
2. O que a correlação indica?
3. Lembre-se de anotar as respostas, elas serão úteis para responder os exercícios.

Faça um teste de hipóteses

Quando temos a intuição de que algum fator, situação ou variável tem relação com algum acontecimento ou fenômeno, podemos utilizar um teste de hipóteses para “confirmar” ou não a nossa suspeita. Por exemplo, será que as vendas de limonada são afetadas pelo marketing? Em outras palavras, será que a distribuição de panfletos está relacionada com maior volume de vendas de limonada? Vejamos como podemos utilizar o Excel para realizar um teste de hipóteses.

1. Na planilha **Limonada**, limpe quaisquer filtros da tabela de dados de vendas de limonada. Em seguida, clique na célula **A1** e pressione CTRL + T para selecionar a tabela inteira de dados de vendas de limonada. Em seguida, na guia **Início** da faixa de opções, clique em **Copiar**.
2. Adicione uma nova planilha e cole os dados copiados na célula **A11** (deixando dez linhas em branco acima dos dados colados).

Você formula a hipótese de que nos dias em que Rosie distribui um número de panfletos acima da média, as vendas são maiores. Você precisa testar essa hipótese para determinar se qualquer aumento nas vendas em dias com distribuição de panfletos acima da média pode ser explicado pelo acaso, ou se a variação nas vendas é muito improvável para ser explicada apenas pelo acaso. Para ajudar a determinar isso, você precisa criar uma amostra de dados contendo vendas para dias com distribuição de panfletos acima da média.

3. Na lista suspensa do cabeçalho da coluna **Panfletos**, no submenu **Filtros de números**, clique em **Acima da média**.
4. Selecione a coluna **Vendas** (incluindo o cabeçalho) e, em seguida, selecione a célula **K1** e, na lista suspensa **Colar**, clique em **Colar Valores**. Isso cola a amostra de dados de vendas filtrada (que contém as observações de dias com distribuição de folhetos acima da média), conforme mostrado aqui:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1											Vendas			
2											22			
3											22			
4											22			
5											23			
6											25			
7											25			
8											23			
9											25			
10											23			
11	Data	Mês	Dia	Temperat	Chuva	Panfletos	Preço	Vendas	Receita					
46	04/02/2017	fevereiro	Sábado	13,7	8,3	46	1,20	22	R\$ 26,40					
54	12/02/2017	fevereiro	Domingo	13,1	8,3	41	1,20	22	R\$ 26,40					
70	28/02/2017	fevereiro	Terça	9,8	9,1	45	1,20	22	R\$ 26,40					
71	01/03/2017	março	Quarta	14,4	8,7	46	1,20	23	R\$ 27,60					
78	08/03/2017	março	Quarta	14,7	7,7	43	1,20	25	R\$ 30,00					
82	12/03/2017	março	Domingo	16,4	7,4	47	1,20	25	R\$ 30,00					
83	13/03/2017	março	Segunda	13,3	8,7	48	1,20	23	R\$ 27,60					

5. Limpe o filtro da coluna **Panfletos** para que a tabela que começa na linha 11 mostre o preenchimento total dos dados de vendas.
6. Na célula **K1**, altere o cabeçalho da coluna para Amostra. Em seguida, selecione os dados da amostra (incluindo o cabeçalho) e mova-os para a célula **K11**, como mostrado aqui:

Salvamento Automático

Limônada - Excel

Lindomar Daniel

Arquivo Página Inicial Inserir Layout da Página Fórmulas Dados Revisão Exibir Ajuda Pesquisar

Calibri 11 A A

Quebrar Texto Automaticamente

Formato % 000 00 00

Formatação Condicional

Formatar como Tabela

Estilos de Célula

Inserir Excluir Formatar

Compartilhar Comentários

AutoSoma

Preencher

Limpar

Classificar e Filtrar

Localizar e Selecionar

Edição

Área de Transferência

Fonte

Alinhamento

Número

Estilos

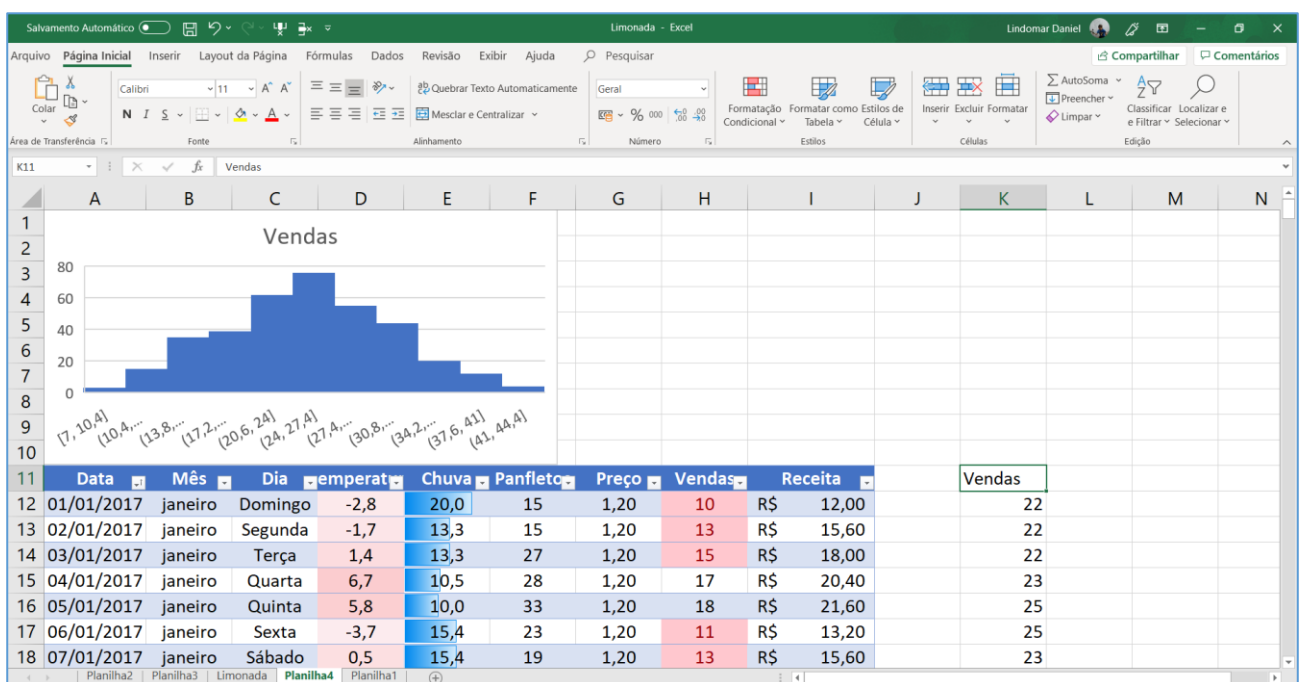
Células

L14

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1														
2														
3														
4														
5														
6														
7														
8														
9														
10														
11		Data	Mês	Dia	temperat	Chuva	Panfleto	Preço	Vendas	Receita		Vendas		
12		01/01/2017	janeiro	Domingo	-2,8	20,0	15	1,20	10	R\$ 12,00		22		
13		02/01/2017	janeiro	Segunda	-1,7	13,3	15	1,20	13	R\$ 15,60		22		
14		03/01/2017	janeiro	Terça	1,4	13,3	27	1,20	15	R\$ 18,00		22		
15		04/01/2017	janeiro	Quarta	6,7	10,5	28	1,20	17	R\$ 20,40		23		
16		05/01/2017	janeiro	Quinta	5,8	10,0	33	1,20	18	R\$ 21,60		25		
17		06/01/2017	janeiro	Sexta	-3,7	15,4	23	1,20	11	R\$ 13,20		25		
18		07/01/2017	janeiro	Sábado	0,5	15,4	19	1,20	13	R\$ 15,60		23		

Planilha2 Planilha3 Limônada Planilha4 Planilha1

7. Selecione a coluna **Vendas** (incluindo o cabeçalho) e insira um histograma mostrando a distribuição de vendas. Altere o título do gráfico para **Vendas** e, em seguida, redimensione e posicione o histograma no espaço acima dos dados, da seguinte forma:



8. Na célula **G2**, insira o texto **Média** e, em seguida, na célula **H2**, digite a seguinte fórmula:
= MÉDIA (H12: H376)
9. Na célula **G3**, insira o texto DesvPad e, na célula **H3**, insira a seguinte fórmula:
= DESVPAD.P (H12: H376)
10. Na célula **G4**, insira o texto **Amostra** e, na célula **H4**, insira a seguinte fórmula para calcular a média da amostra de vendas para dias com distribuição de panfletos acima da média (dos quais existem 172):
= MÉDIA (K12: K183)

Nas células **G2** a **H4**, você deve agora ter os seguintes valores:

Média	25,32329
Desvio Padrão	6,884139
Amostra	29,99419

A partir disso, você pode ver que a média da amostra de aproximadamente 29,99 (o número médio de vendas em dias com distribuição de panfletos acima da média) é de fato maior do que a média populacional de cerca de 25,32 (o número médio de vendas em todos os dias). Você também pode ver no histograma que a média da população está no meio e a média da amostra está à direita. No entanto, há alguma variação nos dados da população, resultando em um desvio padrão de cerca de 6,88. Assim, as vendas mais altas nos dias em que Rosie distribui mais panfletos seriam explicadas simplesmente por essa variância normal (vamos chamar isso de nossa hipótese nula), ou há evidências suficientes para rejeitar essa explicação em favor de uma hipótese alternativa de que as vendas de fato aumentam nesses dias?

H₀: (hipótese nula) as vendas mais altas nos dias em que são distribuídos mais panfletos são explicadas pela variância normal das vendas;

H₁: (hipótese alternativa) as vendas mais altas nos dias em que são distribuídos mais panfletos não podem ser explicadas apenas pela variância normal das vendas.

Para determinar isso, vamos realizar algo chamado teste Z. O teste consiste basicamente em comparar a diferença média de vendas com o desvio padrão. Se o desvio padrão for “grande o suficiente” podemos dizer que as vendas acima da média são normais, caso contrário, se o desvio padrão for “pequeno o suficiente” podemos dizer que as vendas acima da média registrada nos dias específicos é improvável, ou seja, estariam associados a um fator específico, nesse caso, à distribuição de mais panfletos.

11. Na célula **G5**, insira o texto **Valor-P** e, em seguida, na célula **H5**, digite a seguinte fórmula: =TESTE.Z (K12:K183; H2; H3)

Isso calcula um **valor p**, que é a probabilidade de observar uma média de amostra pelo menos tão alta quanto nosso valor de 29,99 em uma distribuição de amostra de 172 de uma população com uma média de 25,32 e um desvio padrão de 6,88.

Provavelmente, o valor p é exibido em notação científica, com um valor semelhante a 2.83128E-19. Para visualizá-lo como um número decimal regular, selecione o valor p (na célula **H4**) e na guia **Início** da faixa de opções, na seção **Número**, selecione o formato **Número** e clique repetidamente no botão **Aumentar Decimal** até o primeiro não-zero a casa decimal é mostrada. Deve estar perto de 0,000000000000000003.

Esta é claramente uma probabilidade muito pequena e, de fato, é comum usar um valor de 0,05 (ou 5%) como o limite para rejeitar a hipótese nula em favor da hipótese alternativa. Em outras palavras, quando o valor-p é maior que 0,05 o desvio padrão é “grande o suficiente” e, quando é menor que 0,05 ele é “pequeno o suficiente”. Portanto, neste caso, o valor-p é muito menor do que esse limite e podemos rejeitar a hipótese de que o aumento nas vendas pode ser explicado pela variância aleatória e concluir que há algum fator não-aleatório em ação aqui. Note que não podemos afirmar categoricamente que o aumento nas vendas se deve ao maior número de panfletos distribuídos, mas podemos dizer que nos dias em que mais panfletos foram distribuídos, houve um aumento estatisticamente significativo nas vendas.

Vale a pena pensar um pouco sobre o que isso significa. O que isso mostra? Nós mostramos que, se os resultados fossem devidos ao acaso (ou seja, H_0 = verdadeiro), poderíamos ainda obter esses resultados. Em outras palavras, não refutamos o acaso como explicação. No entanto, nós mostramos que a probabilidade de obter esses resultados é muito, muito, muito pequena ... apenas

0,00000000000000000003. Estamos seguros em rejeitar o acaso como explicação para nossos resultados aqui.

Desafio: Teste de hipóteses de chuva

1. Teste as seguintes hipóteses:
 - H_0 (hipótese nula): Vendas mais altas em dias com precipitação menor do que a média pode ser explicada por variância aleatória.
 - H_1 (hipótese alternativa): as vendas médias em dias com precipitação pluviométrica abaixo da média são significativamente mais altas do que a média da população e não podem ser explicadas pela variância aleatória.
2. Você deve rejeitar a hipótese nula se o valor p para seu teste for menor que 0,05.
3. Lembre-se de anotar os resultados, eles serão necessários para responder os exercícios.

REFERÊNCIAS

GUJARATI, D. **Econometria: princípios, teoria e aplicações práticas**. São Paulo: Saraiva, 2019. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9788553131952/pageid/4> .

GUJARATI, D. N. **Econometria básica**. 5. ed. Porto Alegre: AMGH, 2011. Disponível em: <https://integrada.minhabiblioteca.com.br/books/9788580550511>.

MICROSOFT PROFESSIONAL PROGRAM. **Introduction to data Science**. 2018. Disponível em: <https://academy.microsoft.com/en-us/professional-program/>.

SARTORIS, A. **Estatística e introdução à econometria**. 2. ed. São Paulo: Saraiva, 2013. Disponível em: <https://integrada.minhabiblioteca.com.br/books/9788502199835>.

WOOLDRIDGE, J. M. **Introdução à econometria: uma abordagem moderna**. São Paulo: Cengage Learning, 2016. Disponível em: <https://integrada.minhabiblioteca.com.br/books/9788522126996>.

EXERCÍCIOS

EXERCÍCIO LABORATÓRIO 3

No Laboratório 3, você calculou estatísticas descritivas sobre a temperatura.

- 1) Quais os valores calculados para medidas de tendência central?
 - a) Média:
 - b) Mediana:
 - c) Moda:
- 2) Quais os valores calculados para medidas de variância?
 - a) Variância populacional:
 - b) Desvio-Padrão populacional:
- 3) No Laboratório 3, você criou um histograma e um gráfico Caixa Estreita para temperatura. Qual a melhor descrição para a distribuição da temperatura?
 - a) Assimétrica à esquerda;
 - b) Aproximadamente normal;
 - c) Assimétrica à direita;
 - d) Uniforme.
- 4) No Laboratório 3, você calculou a correlação entre chuva e vendas:
 - a) Qual foi o valor calculado para a correlação?
 - b) O que indica o valor calculado para a correlação?
 - i. Altos níveis de chuva causam baixos níveis de vendas;
 - ii. Altos níveis de chuva causam altos níveis de vendas;
 - iii. Dias com baixos níveis de chuva tendem a ter vendas mais elevadas do que dias com mais chuva;
 - iv. Dias com altos níveis de chuva tendem a ter vendas mais elevadas do que dias com menos chuva.
- 5) No Laboratório 3, você realizou um teste de hipóteses com a seguinte hipótese:
 - H_0 : Dias com vendas acima da média e com chuva abaixo da média podem ser explicados pela variância aleatória;
 - H_1 : A média de vendas em dias com chuva abaixo da média são significativamente maiores do que a média populacional e não podem ser explicados pela variância aleatório.Use um valor crítico de 0,05 para determinar se a hipótese nula pode ser rejeitada em favor da hipótese alternativa.
- 6) No Brasil, o Banco Central procura alcançar uma meta de inflação com um valor central, por exemplo, de 4% a.a. em torno da qual se define uma faixa tolerável de $\pm 2\%$ a.a. Se tal intervalo ocorrer de fato e as probabilidades puderem ser aproximadas pela distribuição uniforme entre + 2% e + 6%,

(A) a probabilidade de que a inflação venha a se situar entre 2 e 2.5% a.a. será de 10%.

- (B) a mediana da distribuição será 3.5% a.a.
- (C) a distribuição será bimodal.
- (D) a amplitude total da distribuição será de 2% a.a.
- (E) o quartil inferior da distribuição irá de 2 a 3% a.a.