



82%

01

## Distplot, boxplot e violinplot

ATIVIDADES  
1 de 7

### Transcrição

Nesta etapa, faremos a **última análise** deste curso.

Falta-nos analisar a **hora do dia** em relação ao valor da conta e percentual de gorjeta. Neste campo, temos dois períodos diferentes a serem analisados: Almoço e Jantar, visíveis por

```
gorjetas.hora_do_dia.unique()
```

Começaremos adicionando um novo link em uma célula textual da mesma maneira que fizemos para organizar bem nosso relatório. Nesta, escreveremos "#Análise 4 - Hora do dia". Poderemos fechar as outras partes para facilitar o trabalho.

Aplicaremos tanto análises visuais quanto utilizando `ranksums()` para os dois grupos em questão. Nosso primeiro gráfico estatístico será **categórico** cruzando agora do dia com o valor da conta, ou seja, usaremos `catplot()` passando `x` como `'hora_do_dia'` e `y` como `'valor_da_conta'`.

```
sns.catplot(x='hora_do_dia', y='valor_da_conta', data=gorjetas)
```

COPIAR CÓDIGO



82%

01

## Distplot, boxplot e violinplot

ATIVIDADES  
1 de 7

### Transcrição

Nesta etapa, faremos a **última análise** deste curso.

Falta-nos analisar a **hora do dia** em relação ao valor da conta e percentual de gorjeta. Neste campo, temos dois períodos diferentes a serem analisados: Almoço e Jantar, visíveis por

```
gorjetas.hora_do_dia.unique()
```

Começaremos adicionando um novo link em uma célula textual da mesma maneira que fizemos para organizar bem nosso relatório. Nesta, escreveremos "#Análise 4 - Hora do dia". Poderemos fechar as outras partes para facilitar o trabalho.

Aplicaremos tanto análises visuais quanto utilizando `ranksums()` para os dois grupos em questão. Nosso primeiro gráfico estatístico será **categórico** cruzando agora do dia com o valor da conta, ou seja, usaremos `catplot()` passando `x` como `'hora_do_dia'` e `y` como `'valor_da_conta'`.

```
sns.catplot(x='hora_do_dia', y='valor_da_conta', data=gorjetas)
```

COPIAR CÓDIGO

DISCORD  
ALURA

FÓRUM DO  
CURSO

VOLTAR  
PARA  
DASHBOARD

MODOS  
NOTURNO



14.7k xp

a



82%

ATIVIDADES  
1 de 7

DISCORD  
ALURA

FÓRUM DO  
CURSO

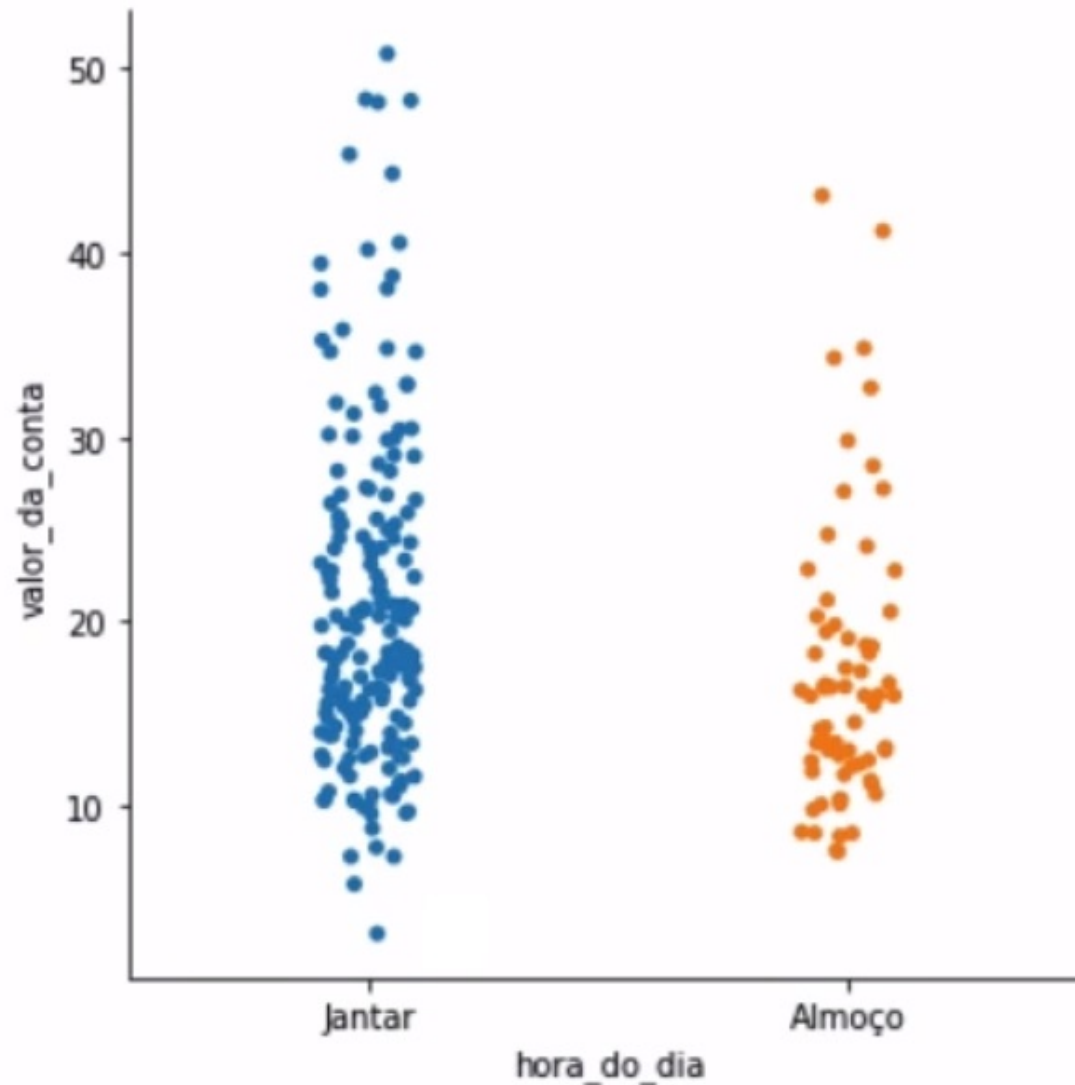
VOLTAR  
PARA  
DASHBOARD

MODO  
NOTURNO



14.7k xp

a



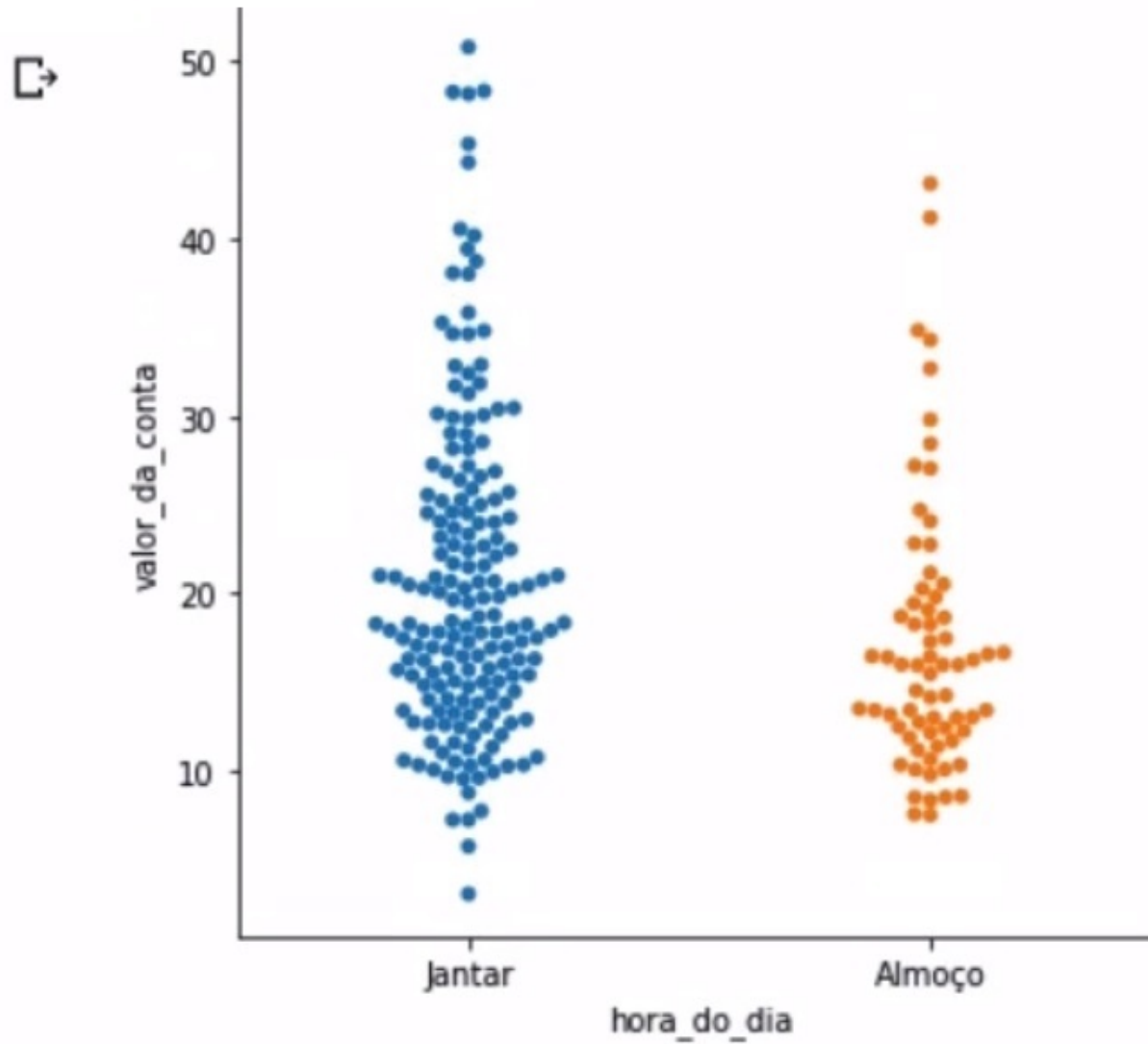
Aparentemente, o jantar é o momento do dia com os maiores valores de conta. Perceberemos que existem alguns pontos com quantias muito parecidas, e poderemos aplicar um parâmetro que nos permitirá visualizar melhor todos os pontos separadamente.

Copiaremos a última sentença de código e colaremos na célula seguinte, passando um quarto parâmetro `kind` sendo igual a `'swarm'` .



82%

```
sns.catplot(x='hora_do_dia', y='valor_da_conta', kind='swarm', data=gorjetas)
```

[COPIAR CÓDIGO](#)

Desta forma, conseguiremos visualizar melhor cada um dos pontos do gráfico.

ATIVIDADES  
1 de 7

DISCORD  
ALURA

FÓRUM DO  
CURSO

VOLTAR  
PARA  
DASHBOARD

MODO  
NOTURNO



14.7k xp

a

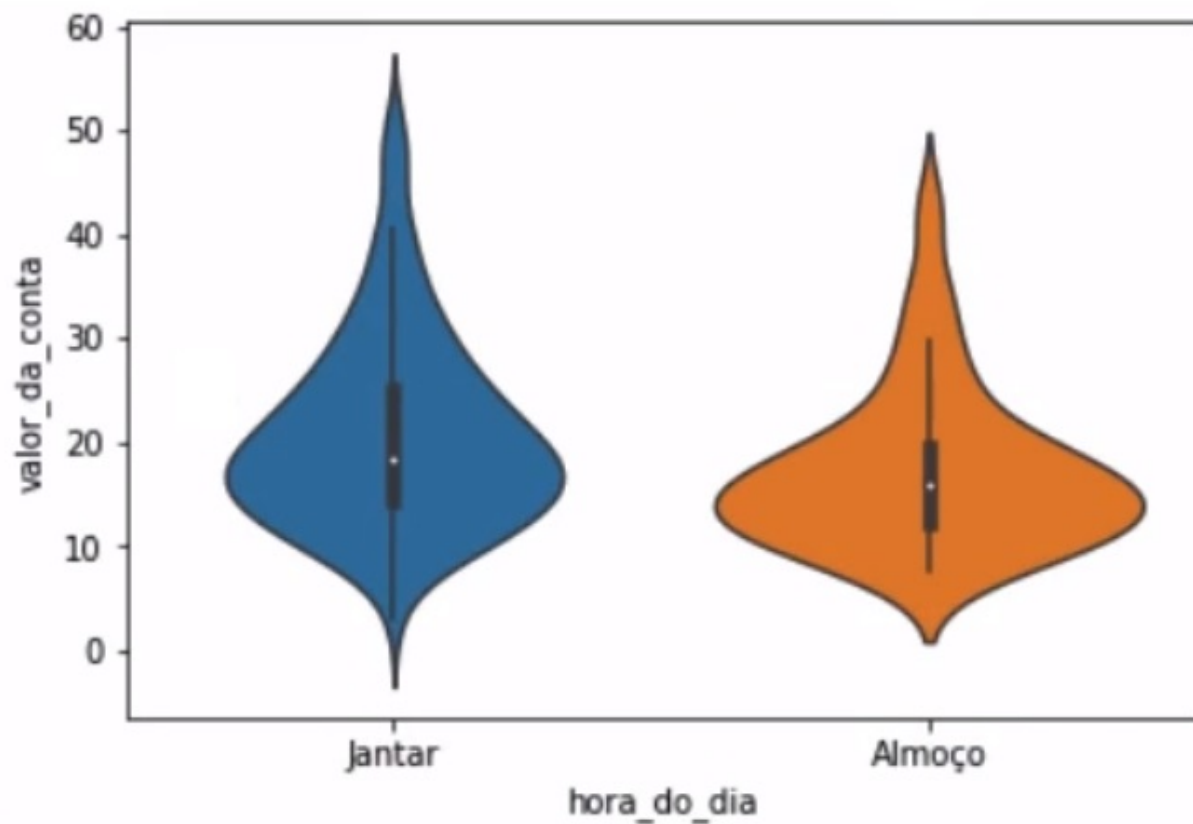


82%

Geraremos um outro gráfico interessante por meio de `violinplot()` , sendo justamente a tradução de "violino" para o inglês.

Como parâmetro, passaremos `x` sendo a `'hora_do_dia'` , `y` como `'valor_da_conta'` e `data=gorjetas` .

```
sns.violinplot(x='hora_do_dia', y='valor_da_conta', data=gorjetas)
```

[COPIAR CÓDIGO](#)

ATIVIDADES  
1 de 7

DISCORD  
ALURA

FÓRUM DO  
CURSO

VOLTAR  
PARA  
DASHBOARD

MODOS  
NOTURNO



14.7k xp

a



82%

ATIVIDADES  
1 de 7

DISCORD  
ALURA

FÓRUM DO  
CURSO

VOLTAR  
PARA  
DASHBOARD

MODOS  
NOTURNO



14.7k xp

a

Neste gráfico, conseguiremos ver que há um grupo bem grande com valor da conta próximo de \$20 no jantar, e outro maior ainda por volta dos \$15 na hora do almoço.

Em seguida, visualizaremos os dois grupos lado a lado, analisando a média dos valores para cada horário do dia.

Este novo gráfico utilizará o comando `boxplot()` recebendo os mesmos valores nos eixos anteriores.

```
sns.boxplot(x='hora_do_dia', y='valor_da_conta', data=gorjetas)
```

COPIAR CÓDIGO



82%

ATIVIDADES  
1 de 7

DISCORD  
ALURA

FÓRUM DO  
CURSO

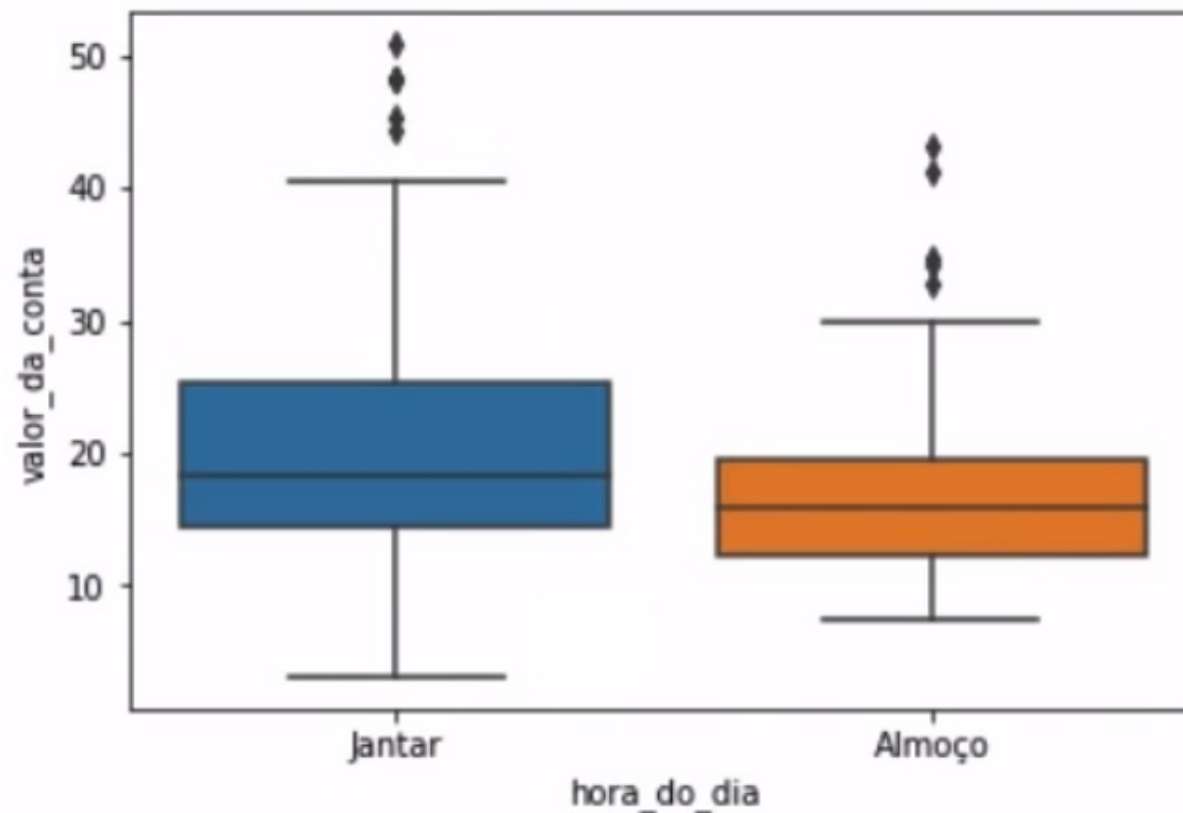
VOLTAR  
PARA  
DASHBOARD

MODOS  
NOTURNO



14.7k xp

a



Desta forma, poderemos ver mais claramente que a mediana dos valores da conta no horário do jantar é diferente da mediana do almoço.

Outro gráfico bastante conhecido e tradicional que utilizaremos em nossa análise é o **histograma** por meio de `distplot()`. Para que funcione, precisaremos pegar os valores da conta de todas as pessoas que almoçaram no restaurante através de uma `query()`.

Passaremos como parâmetro `hora_do_dia` sendo `==` a `'Almoço'` entre aspas. Por fim, aplicaremos `valor_da_conta` e armazenaremos em uma nova variável `almoco`.

```
almoco = gorjetas.query("hora_do_dia == 'Almoço']").valor_da_conta
```



82%

[COPIAR CÓDIGO](#)

Com esta variável, chamaremos a biblioteca Seaborn com `distplot()` passando `almoco` .

ATIVIDADES  
1 de 7

```
sns.distplot(almoco)
```

[COPIAR CÓDIGO](#)

DISCORD  
ALURA

FÓRUM DO  
CURSO

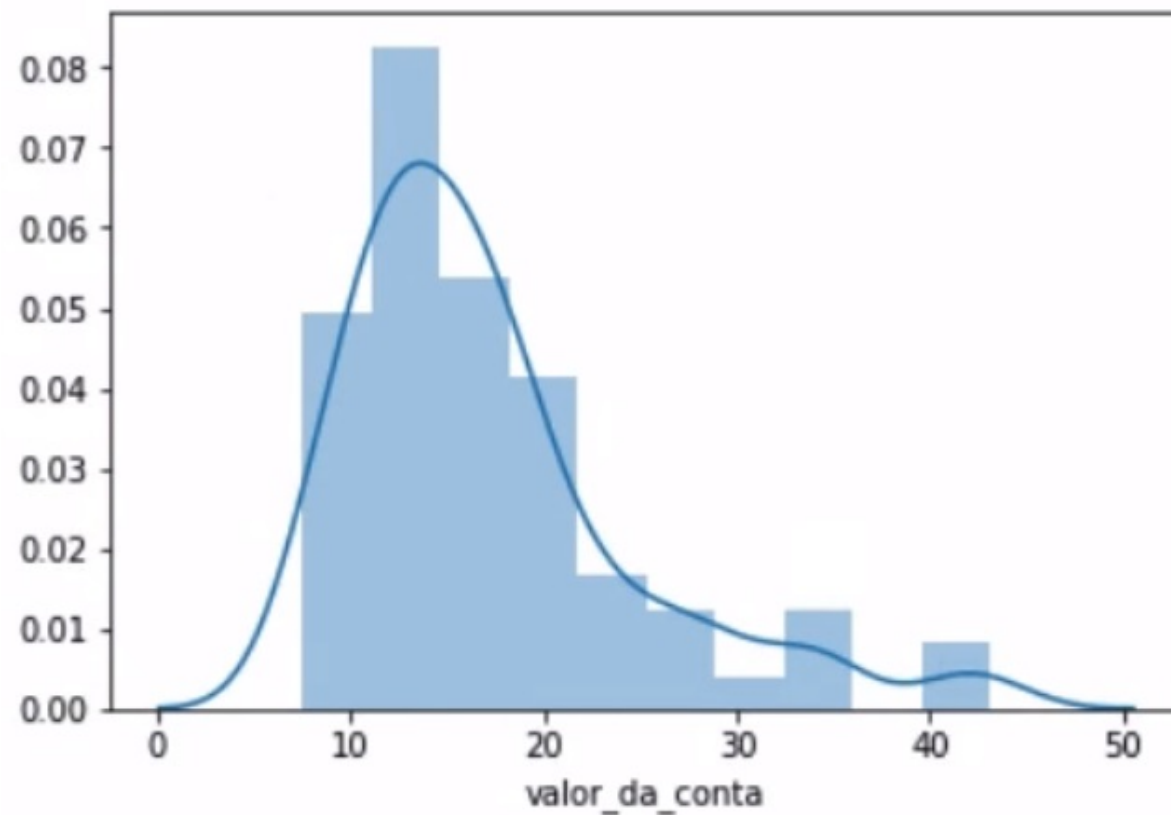
VOLTAR  
PARA  
DASHBOARD

MODOS  
NOTURNO



14.7k xp

a

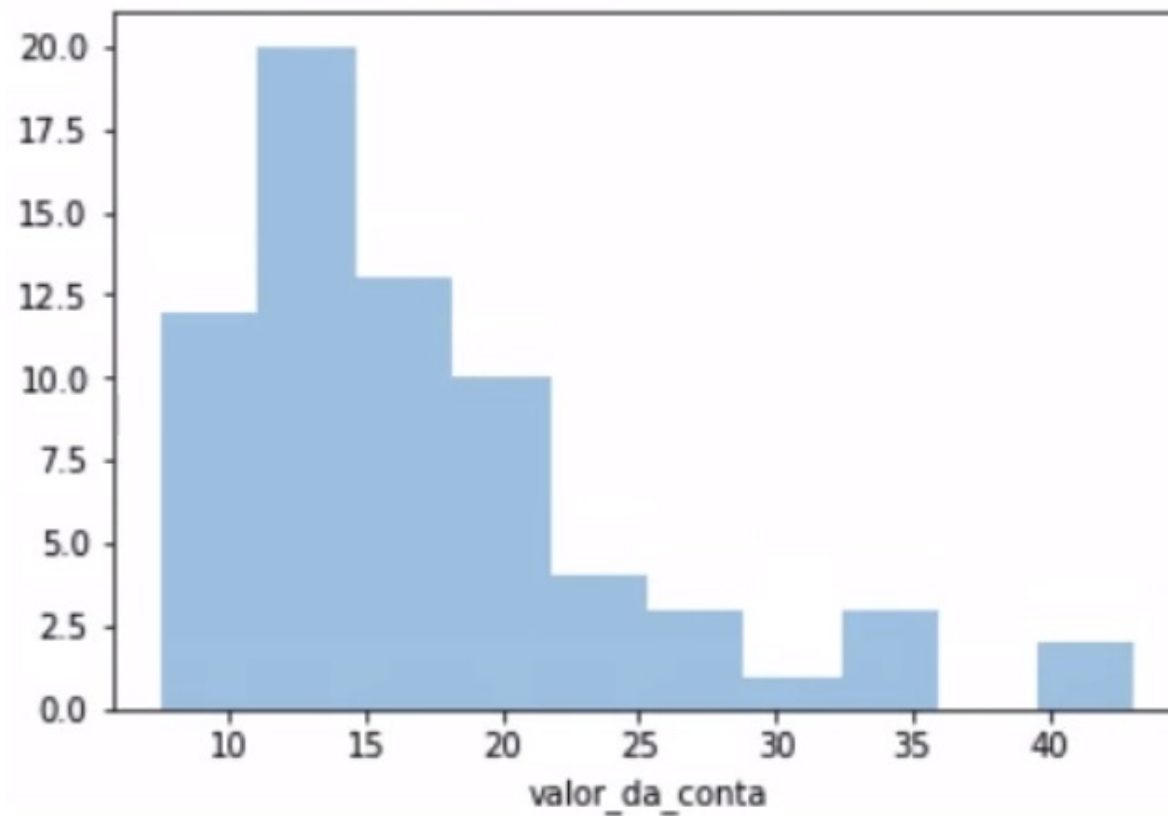


Poderemos passar um parâmetro `kde` igual a `False` para visualizarmos um gráfico atrás deste último, facilitando nossa leitura.



82%

```
sns.distplot(almoco, kde=False)
```

[COPIAR CÓDIGO](#)

Neste gráfico, vemos uma grande faixa entre os valores \$10 e \$15. Faremos o mesmo processo para o jantar.

Copiaremos e colaremos a sentença da variável `almoco`, substituindo 'Almoço' por 'Jantar' e renomeando a variável para '`jantar`'.

```
jantar = gorjetas.query("hora_do_dia == 'Jantar']").valor_da_conta
```

ATIVIDADES  
1 de 7

DISCORD  
ALURA

FÓRUM DO  
CURSO

VOLTAR  
PARA  
DASHBOARD

MODO  
NOTURNO



14.7k xp

a





82%

[COPIAR CÓDIGO](#)

Em seguida, geraremos o gráfico com `distplot()` para `jantar`.

ATIVIDADES  
1 de 7

```
sns.distplot(jantar)
```

[COPIAR CÓDIGO](#)

DISCORD  
ALURA

FÓRUM DO  
CURSO

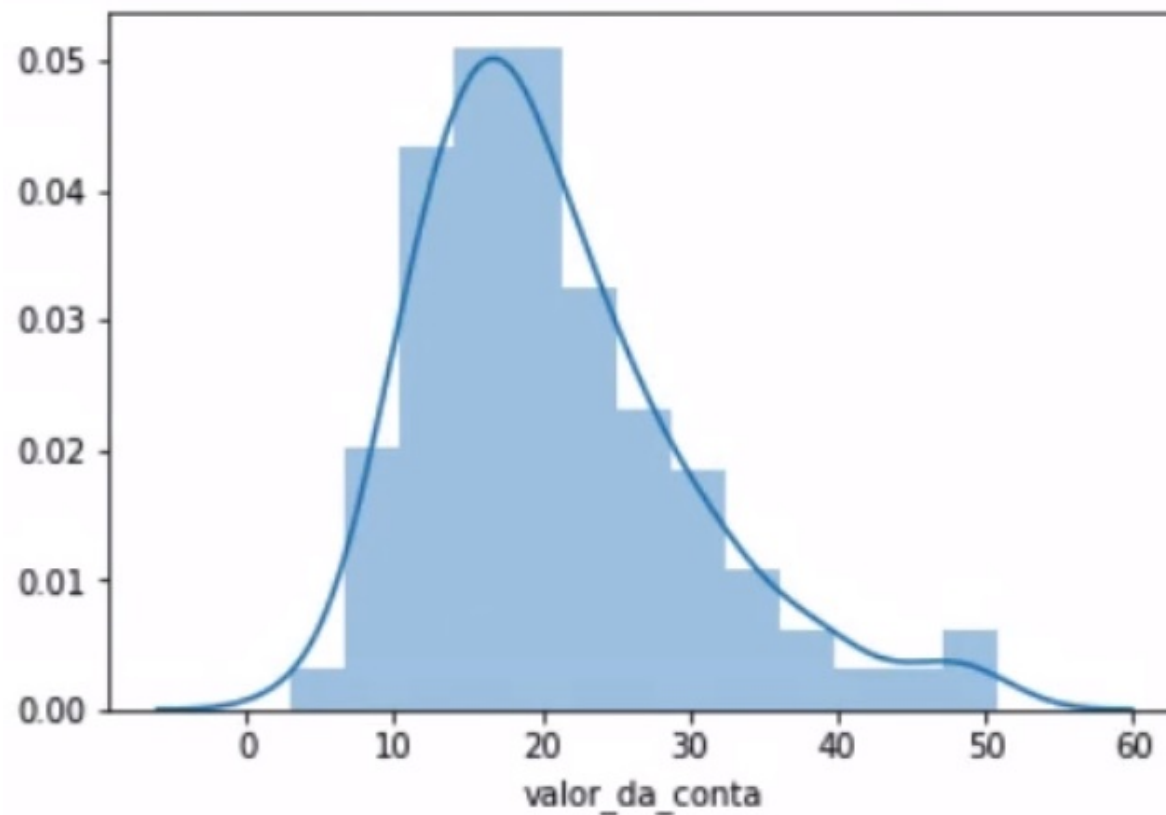
VOLTAR  
PARA  
DASHBOARD

MODO  
NOTURNO



14.7k xp

a



Faremos a mesma alteração na visualização que aplicamos para o almoço, usando um gráfico que facilite a leitura.



82%

```
sns.distplot(jantar, kde=False)
```

[COPIAR CÓDIGO](#)

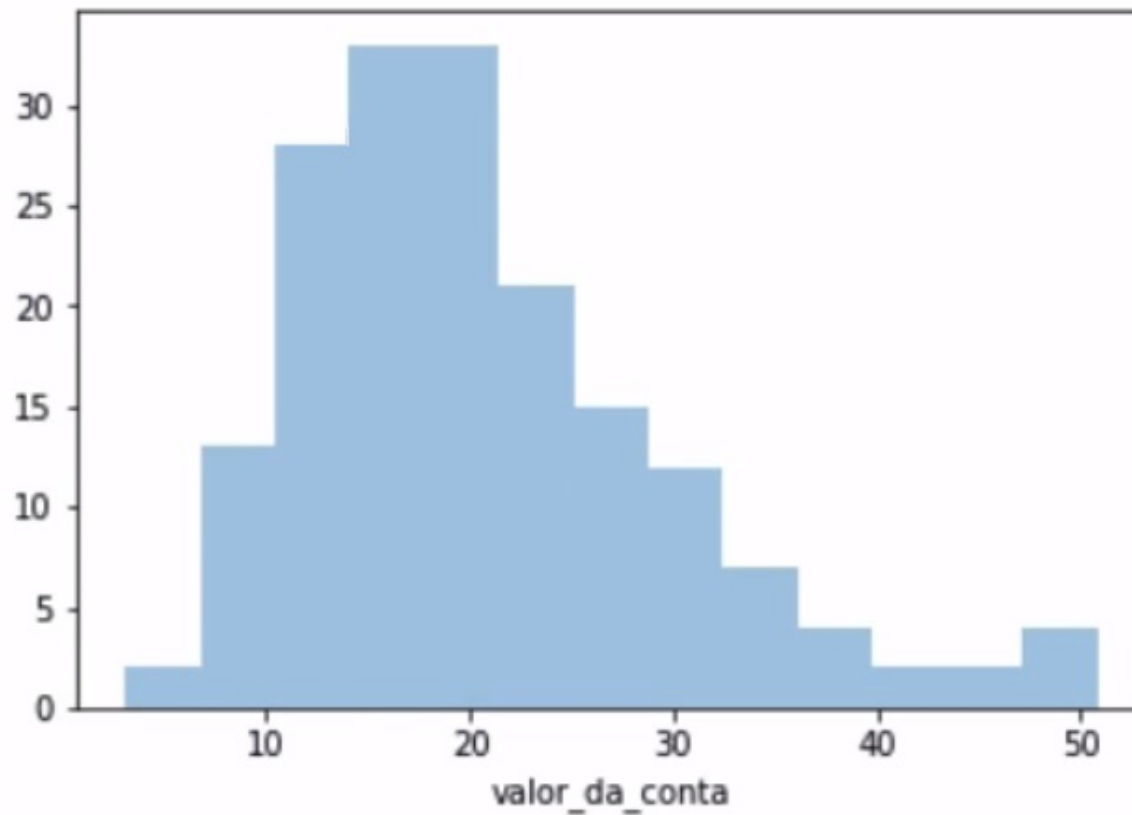
ATIVIDADES  
1 de 7

DISCORD  
ALURA

FÓRUM DO  
CURSO

VOLTAR  
PARA  
DASHBOARD

MODO  
NOTURNO



Com esses gráficos gerados, poderemos ver que são bastante diferentes.

A seguir, verificaremos com testes de hipóteses se o valor das contas de sábado são, de fato, diferente dos valores de domingo.



14.7k xp

a



82%

ATIVIDADES

1 de 7

DISCORD

ALURA

FÓRUM DO

CURSO

VOLTAR

PARA

DASHBOARD

**MODO  
NOTURNO**



14.7k xp

a