

Lendo documentos

Arquivos XLSX
(Microsoft Excel)

Lendo documentos

Arquivos XLSX

Os arquivos *XLSX* (*Microsoft Excel Open XML Spreadsheet*) são criados usando o padrão *Open XML*, que foi introduzido no *Microsoft Office* 2007. Eles são armazenados como um arquivo compactado (similar ao arquivo *DOCX* apresentado na aula anterior), que contém uma coleção de arquivos individuais. O arquivo *XLSX* inclui o arquivo *[Content_Types].xml*, que descreve a pasta de trabalho e contém também arquivos *XML* para cada planilha desta pasta de trabalho.

Lendo documentos

Arquivos XLSX

Para trabalharmos com arquivos *XLSX* podemos utilizar a biblioteca *OpenPyXL* que pode ser instalada utilizando-se o *pip*.

```
$ pip install openpyxl
```

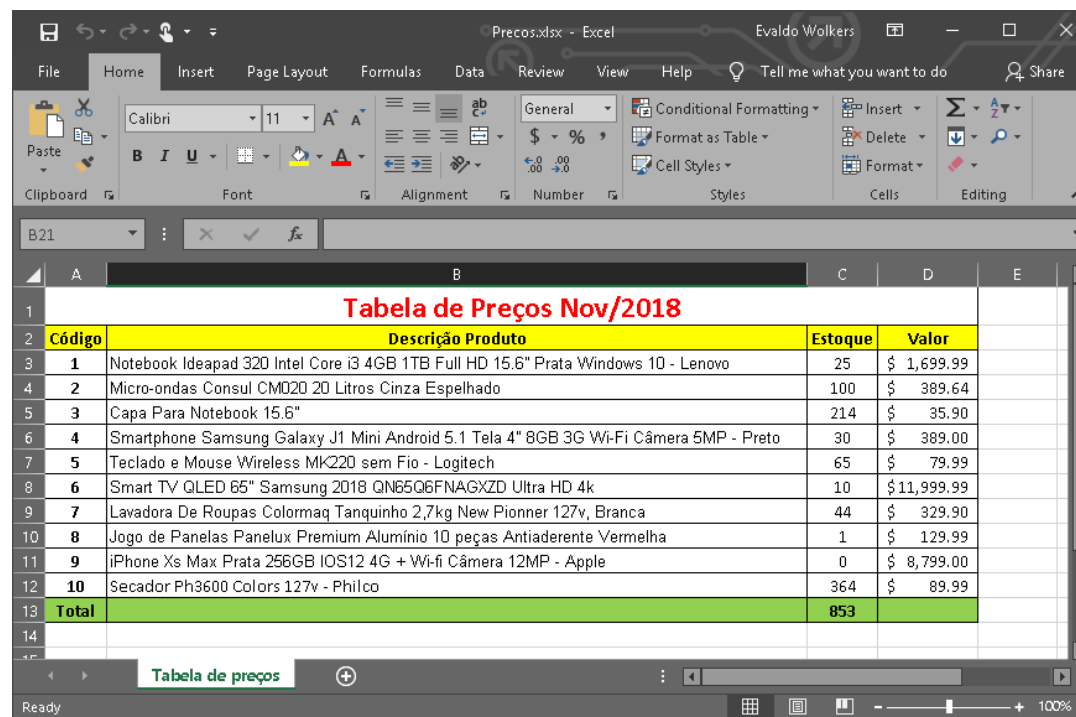
Lendo documentos

Arquivos XLSX

Vamos utilizar a biblioteca *OpenPyXL* para ler o arquivo “Precos.xlsx” que foi disponibilizado junto aos arquivos da aula.

O arquivo “Precos.xlsx” é uma pasta de trabalho, ou **workbook**.

“Tabela de preços” é uma planinha, ou **worksheet**.



The screenshot shows an Excel spreadsheet titled "Precos.xlsx" with a worksheet named "Tabela de preços". The table contains the following data:

Código	Descrição Produto	Estoque	Valor
1	Notebook Ideapad 320 Intel Core i3 4GB 1TB Full HD 15.6" Prata Windows 10 - Lenovo	25	\$ 1,699.99
2	Micro-ondas Consul CM020 20 Litros Cinza Espelhado	100	\$ 389.64
3	Capa Para Notebook 15.6"	214	\$ 35.90
4	Smartphone Samsung Galaxy J1 Mini Android 5.1 Tela 4" 8GB 3G Wi-Fi Câmera 5MP - Preto	30	\$ 389.00
5	Teclado e Mouse Wireless MK220 sem Fio - Logitech	65	\$ 79.99
6	Smart TV QLED 65" Samsung 2018 QN65Q6FNAGXZD Ultra HD 4k	10	\$ 11,999.99
7	Lavadora De Roupas Colomaaq Tanquinho 2,7kg New Pioneer 127v, Branca	44	\$ 329.90
8	Jogo de Panelas Panelux Premium Alumínio 10 peças Antiaderente Vermelha	1	\$ 129.99
9	iPhone Xs Max Prata 256GB IOS12 4G + Wi-fi Câmera 12MP - Apple	0	\$ 8,799.00
10	Secador Ph3600 Colors 127v - Philco	364	\$ 89.99
Total		853	

Lendo documentos

Arquivos XLSX

Para abrir um arquivo *XLSX* temos que utilizar a função *load_workbook* do módulo *openpyxl* passando o caminho completo do arquivo como parâmetro.

Esta função retorna um objeto do tipo *workbook* (pasta de trabalho).

```
from openpyxl import load_workbook  
  
wb = load_workbook('Precos.xlsx')
```

Acima estou considerando que o arquivo está na mesma pasta do meu arquivo *Python*.

Lendo documentos

Arquivos XLSX

Após abrir o arquivo, podemos trabalhar com as planilhas que estão dentro deste arquivo. Podemos criar um objeto que receba uma planilha da pasta de trabalho da seguinte forma:

```
planilha = wb.active
```

Desta forma estamos considerando a planilha ativa.

Ou podemos informar o nome da planilha:

```
planilha = wb["Tabela de preços"]
```



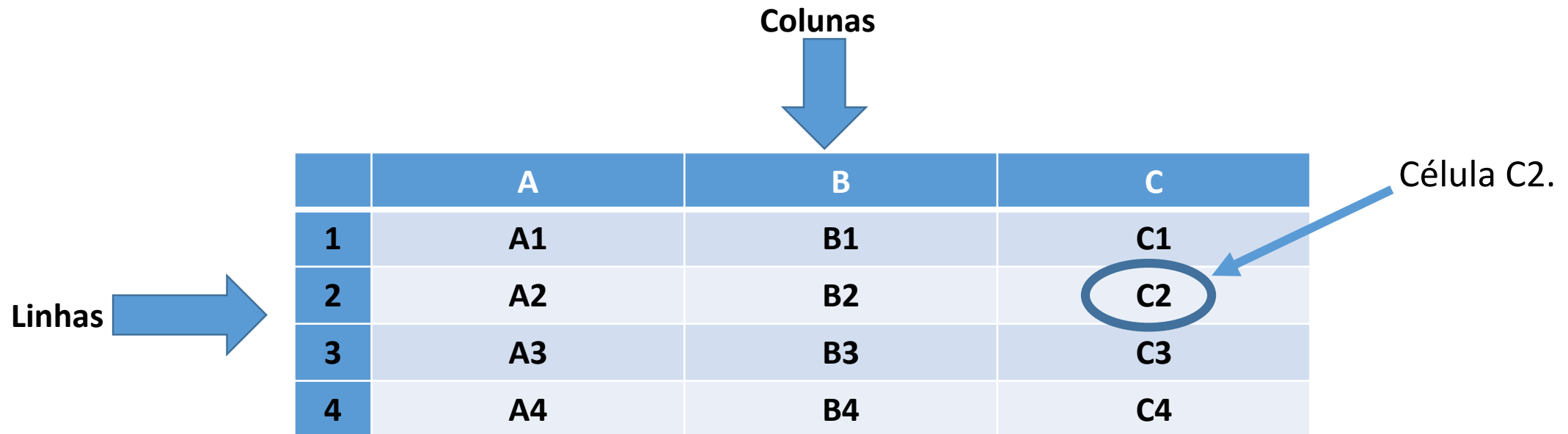
10	8	Jogo de Painelas Panelux Premium Alumínio 10 peç
11	9	iPhone Xs Max Prata 256GB IOS12 4G + Wi-fi Câ
12	10	Secador Ph3600 Colors 127v - Philco
13	Total	
14		
15		

Ready

Lendo documentos

Arquivos XLSX

Para capturar o conteúdo de uma célula da planilha podemos informar o seu endereço, que é composto pela junção da coluna e linha de localização da célula.



	A	B	C
1	A1	B1	C1
2	A2	B2	C2
3	A3	B3	C3
4	A4	B4	C4

Lendo documentos

Arquivos XLSX

Podemos criar um objeto do tipo célula (*Cell*) informando a referência da célula desta forma:

```
celula = planilha['A1']
```

Ou desta forma:

```
celula = planilha.cell(linha=númerolinha, coluna=númerocoluna)
```

Ficando assim:

```
celula = planilha.cell(row=1, column=1) # equivale a "A1".
```

Para obter o valor da célula usamos *celula.value*.

Lendo documentos

Arquivos XLSX

Veja agora o exemplo completo:

```
from openpyxl import load_workbook

wb = load_workbook('Precos.xlsx')

planilha = wb['Tabela de preços']

a1 = planilha['A1']
print(a1.value)

a1 = planilha.cell(row=1, column=1)
print(a1.value)
```

Foi impresso o conteúdo da célula "A1":
Tabela de Preços Nov/2018

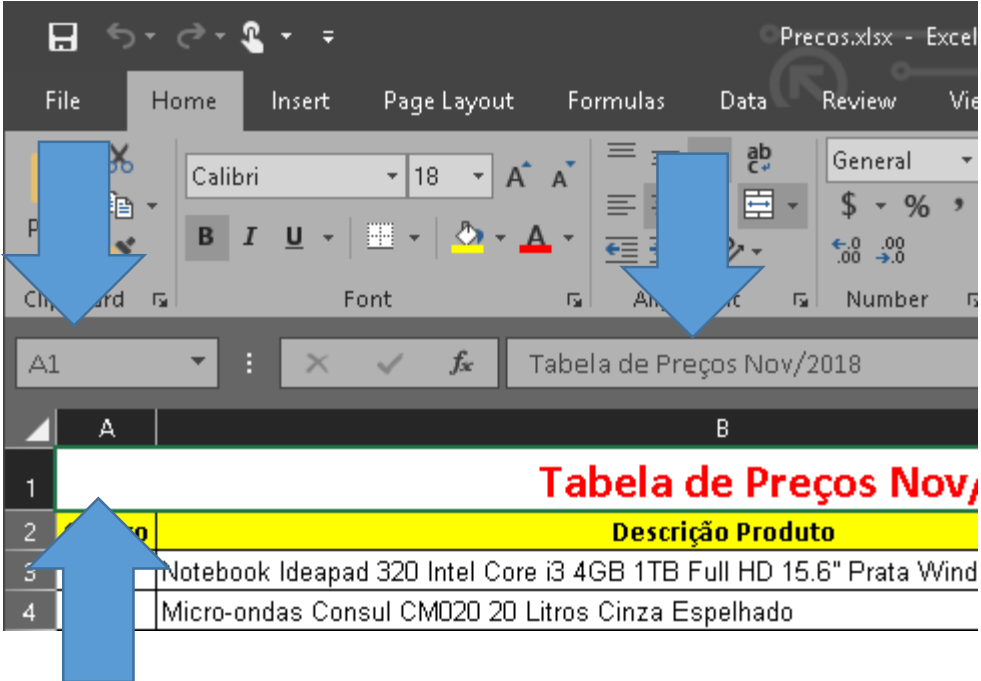


Tabela de Preços Nov/2018	
	Descrição Produto
1	
2	Notebook Ideapad 320 Intel Core i3 4GB 1TB Full HD 15.6" Prata Wind
3	
4	Micro-ondas Consul CM020 20 Litros Cinza Espelhado

Lendo documentos

Arquivos XLSX

Podemos também criar um objeto de célula contendo um intervalo de células. Vou por exemplo armazenar o conteúdo das células de *A3* até *D12* em uma variável e imprimir cada célula do intervalo.



	A	B	C	D
1		Tabela de Preços Nov/2018		
2	Código	Descrição Produto	Estoque	Valor
3	1	Notebook Ideapad 320 Intel Core i3 4GB 1TB Full HD 15.6" Prata Windows 10 - Lenovo	25	\$ 1,699.99
4	2	Micro-ondas Consul CM020 20 Litros Cinza Espelhado	100	\$ 389.64
5	3	Capa Para Notebook 15.6"	214	\$ 35.90
6	4	Smartphone Samsung Galaxy J1 Mini Android 5.1 Tela 4" 8GB 3G Wi-Fi Câmera 5MP - Preto	30	\$ 389.00
7	5	Teclado e Mouse Wireless MK220 sem Fio - Logitech	65	\$ 79.99
8	6	Smart TV QLED 65" Samsung 2018 QN65Q6FNAGXZD Ultra HD 4k	10	\$ 11,999.99
9	7	Lavadora De Roupas Colormaq Tanquinho 2,7kg New Pioneer 127v, Branca	44	\$ 329.90
10	8	Jogo de Panelas Panelux Premium Alumínio 10 peças Antiaderente Vermelha	1	\$ 129.99
11	9	iPhone Xs Max Prata 256GB IOS12 4G + Wi-fi Câmera 12MP - Apple	0	\$ 8,799.00
12	10	Secador Ph3600 Colors 127v - Philco	364	\$ 89.99
13	Total		853	

Lendo documentos

Arquivos XLSX

Veja o programa:

```
from openpyxl import load_workbook

wb = load_workbook('Precos.xlsx')

planilha = wb['Tabela de preços']

cells = planilha['A3': 'D12']

for c1, c2, c3, c4 in cells:
    print("{0:2} {1:90} {2:4} {3:10}".
          format(c1.value, c2.value, c3.value, c4.value))
```

Estamos informando para o método *format* o argumento por posição (0,1,2,3) e definindo a largura que irá ocupar (2, 90, 4, 10).

Lendo documentos

Arquivos XLSX

Veja o resultado:

0	1	2	3
2	Código	Descrição Produto	Estoque
3	1	Notebook Ideapad 320 Intel Core i3 4GB 1TB Full HD 15.6" Prata Windows 10 - Lenovo	25
4	2	Micro-ondas Consul CM020 20 Litros Cinza Espelhado	100
5	3	Capa Para Notebook 15.6"	214

2	90	4	10
1	Notebook Ideapad 320 Intel Core i3 4GB 1TB Full HD 15.6'' Prata Windows 10 - Lenovo	25	1699.99
2	Micro-ondas Consul CM020 20 Litros Cinza Espelhado	100	389.64
3	Capa Para Notebook 15.6"	214	35.9
4	Smartphone Samsung Galaxy J1 Mini Android 5.1 Tela 4" 8GB 3G Wi-Fi Câmera 5MP - Preto	30	389
5	Teclado e Mouse Wireless MK220 sem Fio - Logitech	65	79.99
6	Smart TV QLED 65" Samsung 2018 QN65Q6FNAGXZD Ultra HD 4k	10	11999.99
7	Lavadora De Roupas Colormaq Tanquinho 2,7kg New Pioneer 127v, Branca	44	329.9
8	Jogo de Panelas Panelux Premium Alumínio 10 peças Antiaderente Vermelha	1	129.99
9	iPhone Xs Max Prata 256GB IOS12 4G + Wi-fi Câmera 12MP - Apple	0	8799
10	Secador Ph3600 Colors 127v - Philco	364	89.99

Lendo documentos

Arquivos XLSX

Imprimindo as descrições de todos os produtos:

```
from openpyxl import load_workbook

wb = load_workbook('Precos.xlsx')

planilha = wb['Tabela de preços']

for linha in range(3, 13):
    descricao = planilha[f'B{linha}']
    print(descricao.value)
```

Lendo documentos

Arquivos XLSX

Somando a quantidade total em estoque (não pegando somente o conteúdo da célula *C13* que já tem a soma).

```
from openpyxl import load_workbook

wb = load_workbook('Precos.xlsx')
planilha = wb['Tabela de preços']

qtd = 0

for linha in range(3, 13):
    qtd += planilha[f'C{linha}'].value

print(qtd)
```

Lendo documentos

Arquivos XLSX

Para adicionar uma nova planilha à nossa pasta de trabalho podemos utilizar a função `create_sheet('nome')`, onde 'nome' é o nome/título da planilha que será criada. Em seguida o arquivo deve ser salvo.

```
from openpyxl import load_workbook

arquivo = 'Precos.xlsx'
wb = load_workbook(arquivo)

wb.create_sheet('Planilha de carros')

wb.save(arquivo)
```

5	3	Capa Para Notebook 15.6"
6	4	Smartphone Samsung Galaxy S Mini Android 5.1 Tela 4" 8GB 3G
7	5	Teclado e Mouse Wireless MK220 sem Fio - Logitech
		Tabela de preços
		Planilha de carros
Ready		

Lendo documentos

Arquivos XLSX

Para remover uma planilha usamos o método *remove* do objeto *workbook*.

```
from openpyxl import load_workbook

arquivo = 'Precos.xlsx'
wb = load_workbook(arquivo)

wb.remove(wb['Planilha de carros'])

wb.save(arquivo)
```


Lendo documentos

Arquivos XLSX

Para copiar uma planilha usamos o método *copy_worksheet* do objeto *workbook*.

```
from openpyxl import load_workbook

arquivo = 'Precos.xlsx'
wb = load_workbook(arquivo)

origem = wb['Tabela de preços']
destino = wb.copy_worksheet(origem)
destino.title = 'Nova tabela de preços'

wb.save(arquivo)
```

Lendo documentos

Arquivos XLSX

Para alterar o conteúdo de uma célula devemos passar um valor usando o endereço da célula:

```
from openpyxl import load_workbook
```

```
arquivo = 'Precos.xlsx'
```

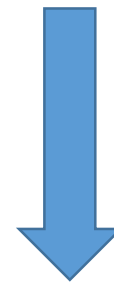
```
wb = load_workbook(arquivo)
```

```
planilha = wb['Nova tabela de preços']
```

```
planilha['A1'] = 'Tabela de Preços Dez/2018'
```

```
wb.save(arquivo)
```

	A	B	C	D
1		Tabela de Preços Nov/2018		
2	Código	Descrição Produto	Estoque	Valor



	A	B	C	D
1		Tabela de Preços Dez/2018		
2	Código	Descrição Produto	Estoque	Valor
3	1	Notebook Ideapad 320 Intel Core i2 4GB 1TB Full HD 15.6" Preto Windows 10 - Lenovo	25	R\$ 1.699,00

Lendo documentos

Arquivos XLSX

Para criar um arquivo temos que criar um objeto da classe Workbook e salvar informando o nome do arquivo.

```
from openpyxl import Workbook  
wb = Workbook()  
wb.save('MeusNumeros.xlsx')
```

Lendo documentos

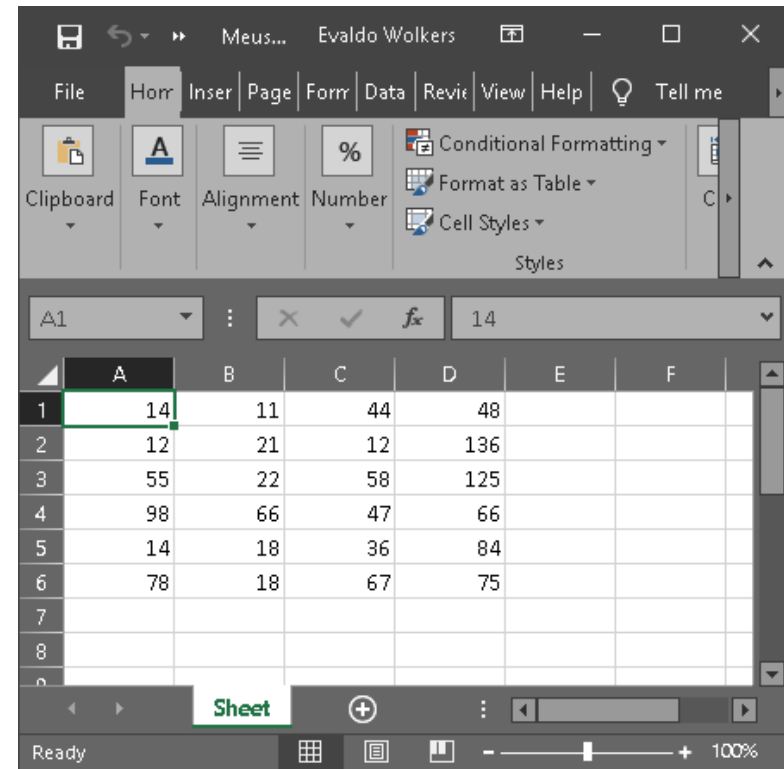
Arquivos XLSX

Vamos agora criar este arquivo e adicionar conteúdo usando o método *append* do objeto *Worksheet*.

```
from openpyxl import Workbook

wb = Workbook()
planilha = wb.active
linhas = (
    (14, 11, 44, 48),
    (12, 21, 12, 136),
    (55, 22, 58, 125),
    (98, 66, 47, 66),
    (14, 18, 36, 84),
    (78, 18, 67, 75)
)

for linha in linhas:
    planilha.append(linha)
wb.save('MeusNumeros.xlsx')
```



Lendo documentos

Arquivos XLSX

Vamos inserir uma linha e uma coluna em nossa planilha.

Para inserir uma linha usamos o método `insert_rows(x)` e para inserir uma coluna o `insert_cols(x)`.

“x” é a posição da linha ou coluna.

	A	B	C	D	E	F
1	14	11	44	48		
2	12	21	12	136		
3	55	22	58	125		
4	98	66	47	66		
5	14	18	36	84		
6	78	18	67	75		
7						
8						



```
planilha.insert_rows(2)  
planilha.insert_cols(3)
```



	A	B	C	D	E	F
1	14	11		44	48	
2						
3	12	21		12	136	
4	55	22		58	125	
5	98	66		47	66	
6	14	18		36	84	
7	78	18		67	75	
8						

Lendo documentos

Arquivos XLSX

Vamos inserir uma linha e uma coluna para colocarmos cabeçalhos em nossa planilha de números.

```
from openpyxl import load_workbook

wb = load_workbook('MeusNumeros.xlsx')

planilha = wb['Sheet']

planilha.insert_rows(1)
planilha.insert_cols(1)

wb.save('MeusNumeros.xlsx')
```

	A	B	C	D	E	F
1						
2		14	11	44	48	
3		12	21	12	136	
4		55	22	58	125	
5		98	66	47	66	
6		14	18	36	84	
7		78	18	67	75	

Lendo documentos

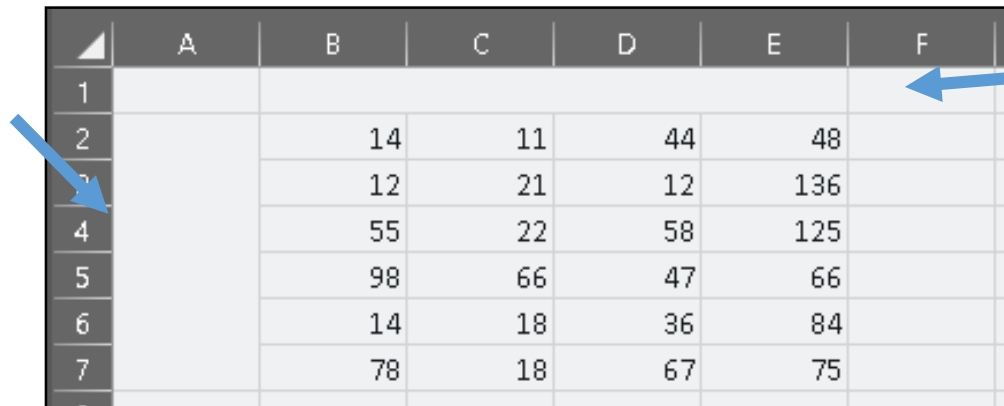
Arquivos XLSX

Agora vamos mesclar as células.

```
from openpyxl import load_workbook

wb = load_workbook('MeusNumeros.xlsx')
planilha = wb['Sheet']
planilha.merge_cells(start_row=1, start_column=2, end_row=1, end_column=5)
planilha.merge_cells("A2:A7")
wb.save('MeusNumeros.xlsx')
```

Mesclado de:
A2 até A7



	A	B	C	D	E	F
1						
2		14	11	44	48	
3		12	21	12	136	
4		55	22	58	125	
5		98	66	47	66	
6		14	18	36	84	
7		78	18	67	75	

Mesclado de:
Linha 1, coluna 2 até
linha 1, coluna 5.

Lendo documentos

Arquivos XLSX

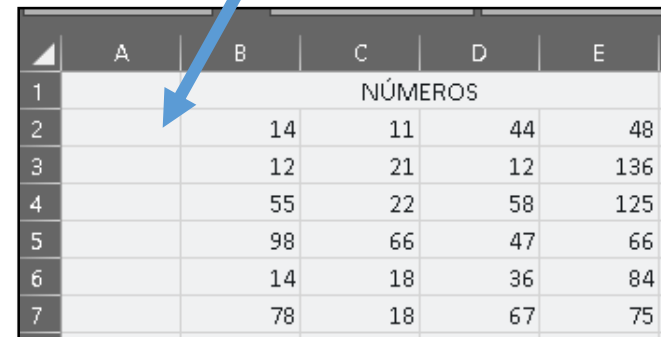
Agora vamos separar as células mescladas da primeira coluna. Vamos alterar o alinhamento da célula onde ficará o título “NÚMEROS” para centralizado.

```
from openpyxl import load_workbook
from openpyxl.styles import Alignment
wb = load_workbook('MeusNumeros.xlsx')

planilha = wb['Sheet']

planilha.unmerge_cells("A2:A7")
planilha['B1'] = "NÚMEROS"
planilha['B1'].alignment = Alignment(horizontal='center')
wb.save('MeusNumeros.xlsx')
```

O merge foi desfeito



	A	B	C	D	E
1		NÚMEROS			
2		14	11	44	48
3		12	21	12	136
4		55	22	58	125
5		98	66	47	66
6		14	18	36	84
7		78	18	67	75

Lendo documentos

Arquivos XLSX

Utilizando fórmulas do *Excel*.

```
from openpyxl import load_workbook

wb = load_workbook('MeusNumeros.xlsx')
planilha = wb['Sheet']

planilha['B8'] = "=SUM(B2:B7) "
planilha['C8'] = "=SUM(C2:C7) "
planilha['D8'] = "=SUM(D2:D7) "
planilha['E8'] = "=SUM(E2:E7) "

planilha['F2'] = "=AVERAGE(B2:E2) "
planilha['F3'] = "=AVERAGE(B3:E3) "
planilha['F4'] = "=AVERAGE(B4:E4) "
...
```

```
...
planilha['F5'] = "=AVERAGE(B5:E5) "
planilha['F6'] = "=AVERAGE(B6:E6) "
planilha['F7'] = "=AVERAGE(B7:E7) "

planilha['G2'] = '=IF(F2>50,"LIKE","DISLIKE") '
planilha['G3'] = '=IF(F3>50,"LIKE","DISLIKE") '
planilha['G4'] = '=IF(F4>50,"LIKE","DISLIKE") '
planilha['G5'] = '=IF(F5>50,"LIKE","DISLIKE") '
planilha['G6'] = '=IF(F6>50,"LIKE","DISLIKE") '
planilha['G7'] = '=IF(F7>50,"LIKE","DISLIKE") '

wb.save('MeusNumeros.xlsx')
```

	A	B	C	D	E	F	G
1		NÚMEROS					
2		14	11	44	48	29.25	DISLIKE
3		12	21	12	136	45.25	DISLIKE
4		55	22	58	125	65	LIKE
5		98	66	47	66	69.25	LIKE
6		14	18	36	84	38	LIKE
7		78	18	67	75	59.5	LIKE
8		271	156	264	534		

Lendo documentos

Arquivos XLSX

Formatando células

	A	B	C	D	E	F	G
1			NÚMEROS				
2		14	11	44	48	29.25	DISLIKE
3		12	21	12	136	45.25	DISLIKE
4		55	22	58	125	65	LIKE
5		98	66	47	66	69.25	LIKE
6		14	18	36	84	38	LIKE
7		78	18	67	75	59.5	LIKE
8		271	156	264	534		

```
from openpyxl import load_workbook
from openpyxl.styles import Font
from openpyxl.styles.colors import BLUE, RED

wb = load_workbook('MeusNumeros.xlsx')

planilha = wb['Sheet']

fonte_titulo = Font(bold=True, size=12, color=RED)
fonte_total = Font(bold=True, color=BLUE)

planilha['B1'].font = fonte_titulo
planilha['B8'].font = fonte_total
planilha['C8'].font = fonte_total
planilha['D8'].font = fonte_total
planilha['E8'].font = fonte_total

wb.save('MeusNumeros.xlsx')
```

FIM