

 **Generate**


10 random numbers using numpy



Close

```
import pandas as pd
import numpy as np
import os
from google.colab import files



uploaded = files.upload()
df = pd.read_csv('data_sales.csv', header='infer')
df.head(5)
```

 Choose Files

 data_sales.csv


- **data_sales.csv**(text/csv) - 1426 bytes, last modified: 11/25/2024 - 100% done

Saving data_sales.csv to data_sales (1).csv

	Date	Region	Item	Unit Price	Units	
0	04-Jul-14	East	Pen Set	4.99	62	
1	12-Jul-14	East	Binder	1.99	29	
2	21-Jul-14	Central	Pen Set	12.49	55	
3	29-Jul-14	East	Binder	19.99	81	
4	07-Aug-14	Central	Pen Set	23.95	42	

Next steps:


Generate code with df



 View recommended plots

New interactive sheet

1) Añade una nueva variable Amount = UnitPrice x Units y muestra las 5 primeras observaciones


```
df['Amount'] = df['Unit Price'] * df['Units']
df.head(5)
```



	Date	Region	Item	Unit Price	Units	Amount	
0	04-Jul-14	East	Pen Set	4.99	62	309.38	
1	12-Jul-14	East	Binder	1.99	29	57.71	
2	21-Jul-14	Central	Pen Set	12.49	55	686.95	
3	29-Jul-14	East	Binder	19.99	81	1619.19	
4	07-Aug-14	Central	Pen Set	23.95	42	1005.90	

Next steps:

Generate code with df

 View recommended plots

New interactive sheet

2) Calcula el precio unitario promedio de cada región. Utiliza el método groupby()

```
avg = df.groupby('Region')['Unit Price'].mean()
print('Usando groupby')
avg
```

↗ Usando groupby

Unit Price	
Region	
Central	18.018750
East	9.143846
West	53.658333

dtype: float64

3) Calcula el precio unitario promedio de cada región. Utiliza el método pivot_table()

```
avg_up = df.pivot_table(values='Unit Price', index='Region', aggfunc='mean')
print('Usando Pivot table')
avg_up
```

↗ Usando Pivot table

Unit Price	
Region	
Central	18.018750
East	9.143846
West	53.658333

Next steps:

[Generate code with avg_up](#)[View recommended plots](#)[New interactive sheet](#)

4) Calcula el precio unitario promedio para cada región en una línea de código. Utiliza el método groupby()

```
print('Usando una línea de código ')
print(df.groupby('Region')['Unit Price'].mean())
```

```

➞ Usando una linea de codigo
Region
Central    18.018750
East       9.143846
West      53.658333
Name: Unit Price, dtype: float64

```

- 5) Calcula el precio unitario promedio y las unidades para cada región en una línea de código. Utiliza el método `pivot_table()`

```

avg_unit_price_region = df.pivot_table(values=['Unit Price', 'Units'], index='Region', aggfunc='mean')
print('Usando una linea de codigo ')
print(avg_unit_price_region)

```

```

➞ Usando una linea de codigo
      Unit Price      Units
Region
Central    18.018750  49.958333
East       9.143846  53.153846
West      53.658333  38.500000

```

- 6) Calcula el total de unidades para cada region y tipo de artículo en una línea de código. Utiliza el método `pivot_table()` y rellena los valores faltantes con 0

```

avgupubr = df.pivot_table(values='Units', index='Region', columns='Item', aggfunc='sum', fill_value=0)
print('Usando una linea de codigo ')
print(avgupubr)

```

```

➞ Usando una linea de codigo
Item      Binder  Desk  Pen  Pen Set  Pencil
Region
Central      424    7   27      243   498
East         234    0  175      152   130
West          64    3   76         0    88

```

- 7) Calcula el importe total de ventas para cada región y el tipo de artículo en una línea de código. Utiliza el método `pivot_table()` y rellena los valores faltantes con 0

```

total_sale = df.pivot_table(values='Amount', index = 'Region', columns = 'Item', aggfunc='sum', fill_value=0)
print('Usando una linea de codigo ')
print(total_sale)

```



Usando una linea de codigo

Item	Binder	Desk	Pen	Pen Set	Pencil
Region					
Central	5762.63	875.0	539.73	2421.39	1540.32
East	2535.66	0.0	1354.25	1748.48	363.70
West	1279.36	825.0	151.24	0.00	231.12