

Tutorial



Reemplazo de Datos con Pandas

{desafío}
latam_





¿Por qué es importante reemplazar datos?



El reemplazo de datos tiene el objetivo principal de **preparar los datos para el análisis**. Los métodos que debemos manejar son los siguientes:

- Reemplazar todos los **valores nulos** con un mismo valor:
 - El método `.fillna(valor)`
- Reemplazar cada **valor nulo** según su índice:
 - El método `.fillna({index1: valor1, index2: valor2})`
- Reemplazos con `.replace()`:
 - El método `.replace(valor_1: valor_2)`
 - El método `.replace({valor_a1: valor_a2, valor_b1: valor_b2})`
- Reemplazos con regex:
 - El método `.replace({r'regex1': 'valor1', r'regex2': 'valor2'}, regex = True)`



Datos Iniciales



```
data = {'producto': ['A', 'B', 'C', 'D', np.nan, np.nan, 'G', 'H'],  
        'precio': [100, 200, 300, 400, 500, 600, 700, 800]}
```

```
df = pd.DataFrame(data)  
df
```

	producto	precio
0	A	100
1	B	200
2	C	300
3	D	400
4	NaN	500
5	NaN	600
6	G	700
7	H	800



Valores Nulos: Opción 1

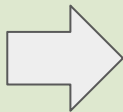


- Reemplazar todos los **valores nulos** con un mismo valor:
 - El método **.fillna(valor)**

```
data = {'producto': ['A', 'B', 'C', 'D', np.nan, np.nan, 'G', 'H'],  
        'precio': [100, 200, 300, 400, 500, 600, 700, 800]}
```

```
df = pd.DataFrame(data)  
df
```

	producto	precio
0	A	100
1	B	200
2	C	300
3	D	400
4	NaN	500
5	NaN	600
6	G	700
7	H	800



```
data = {'producto': ['A', 'B', 'C', 'D', np.nan, np.nan, 'G', 'H'],  
        'precio': [100, 200, 300, 400, 500, 600, 700, 800]}
```

```
df = pd.DataFrame(data)  
df['producto'].fillna('sin informacion')
```

```
0          A  
1          B  
2          C  
3          D  
4  sin informacion  
5  sin informacion  
6          G  
7          H  
Name: producto, dtype: object
```



Valores Nulos: Opción 2

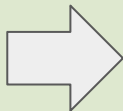


- Reemplazar cada **valor nulo** según su índice:
 - El método `.fillna({index1: valor1, index2: valor2})`

```
data = {'producto': ['A', 'B', 'C', 'D', np.nan, np.nan, 'G', 'H'],  
        'precio': [100, 200, 300, 400, 500, 600, 700, 800]}
```

```
df = pd.DataFrame(data)  
df
```

	producto	precio
0	A	100
1	B	200
2	C	300
3	D	400
4	NaN	500
5	NaN	600
6	G	700
7	H	800



```
data = {'producto': ['A', 'B', 'C', 'D', np.nan, np.nan, 'G', 'H'],  
        'precio': [100, 200, 300, 400, 500, 600, 700, 800]}
```

```
df = pd.DataFrame(data)  
df['producto'].fillna({4:'E', 5:'F'})
```

```
0    A  
1    B  
2    C  
3    D  
4    E  
5    F  
6    G  
7    H  
Name: producto, dtype: object
```



Datos Limpios



```
data = {'producto': ['A', 'B', 'C', 'D', np.nan, np.nan, 'G', 'H'],  
        'precio': [100, 200, 300, 400, 500, 600, 700, 800]}
```

```
df = pd.DataFrame(data)  
df['producto'] = df['producto'].fillna({'4': 'E', 5: 'F'})  
df
```

	producto	precio
0	A	100
1	B	200
2	C	300
3	D	400
4	E	500
5	F	600
6	G	700
7	H	800



Reemplazo de datos: Opción 1

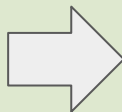


- Reemplazos con `.replace()`:
 - El método `.replace(valor_1: valor_2)`

```
data = {'producto': ['A', 'B', 'C', 'D', np.nan, np.nan, 'G', 'H'],  
       'precio': [100, 200, 300, 400, 500, 600, 700, 800]}
```

```
df = pd.DataFrame(data)  
df['producto'] = df['producto'].fillna({4:'E', 5:'F'})  
df
```

	producto	precio
0	A	100
1	B	200
2	C	300
3	D	400
4	E	500
5	F	600
6	G	700
7	H	800



```
data = {'producto': ['A', 'B', 'C', 'D', np.nan, np.nan, 'G', 'H'],  
       'precio': [100, 200, 300, 400, 500, 600, 700, 800]}
```

```
df = pd.DataFrame(data)  
df['producto'] = df['producto'].fillna({4:'E', 5:'F'})  
df.replace('A', 'Producto A')
```

	producto	precio
0	Producto A	100
1	B	200
2	C	300
3	D	400
4	E	500
5	F	600
6	G	700
7	H	800



Reemplazo de datos: Opción 2

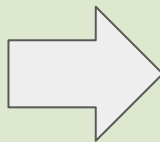


- Reemplazos con `.replace()`:
 - El método `.replace({valor_a1: valor_a2, valor_b1: valor_b2})`

```
data = {'producto': ['A', 'B', 'C', 'D', np.nan, np.nan, 'G', 'H'],  
       'precio': [100, 200, 300, 400, 500, 600, 700, 800]}
```

```
df = pd.DataFrame(data)  
df['producto'] = df['producto'].fillna({4:'E', 5:'F'})  
df
```

	producto	precio
0	A	100
1	B	200
2	C	300
3	D	400
4	E	500
5	F	600
6	G	700
7	H	800



```
data = {'producto': ['A', 'B', 'C', 'D', np.nan, np.nan, 'G', 'H'],  
       'precio': [100, 200, 300, 400, 500, 600, 700, 800]}
```

```
df = pd.DataFrame(data)  
df['producto'] = df['producto'].fillna({4:'E', 5:'F'})  
df['producto'].replace({'A': 'Producto A',  
                       'B': 'Producto B',  
                       'C': 'Producto C',  
                       'D': 'Producto D',  
                       'E': 'Producto E',  
                       'F': 'Producto F',  
                       'G': 'Producto G',  
                       'H': 'Producto H'})
```

```
0    Producto A  
1    Producto B  
2    Producto C  
3    Producto D  
4    Producto E  
5    Producto F  
6    Producto G  
7    Producto H  
Name: producto, dtype: object
```




Reemplazo de datos: Opción 2

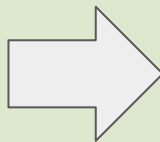


- Reemplazos con **.replace()**:
 - El método **.replace({valor_a1: valor_a2, valor_b1: valor_b2})**

```
data = {'producto': ['A', 'B', 'C', 'D', np.nan, np.nan, 'G', 'H'],  
       'precio': [100, 200, 300, 400, 500, 600, 700, 800]}
```

```
df = pd.DataFrame(data)  
df['producto'] = df['producto'].fillna({4:'E', 5:'F'})  
df
```

	producto	precio
0	A	100
1	B	200
2	C	300
3	D	400
4	E	500
5	F	600
6	G	700
7	H	800



```
data = {'producto': ['A', 'B', 'C', 'D', np.nan, np.nan, 'G', 'H'],  
       'precio': [100, 200, 300, 400, 500, 600, 700, 800]}
```

```
df = pd.DataFrame(data)  
df['producto'] = df['producto'].fillna({4:'E', 5:'F'})  
df['producto'].replace({'A': 'Producto A',  
                       'B': 'Producto B',  
                       'C': 'Producto C',  
                       'D': 'Producto D',  
                       'E': 'Producto E',  
                       'F': 'Producto F',  
                       'G': 'Producto G',  
                       'H': 'Producto H'})
```

```
0    Producto A  
1    Producto B  
2    Producto C  
3    Producto D  
4    Producto E  
5    Producto F  
6    Producto G  
7    Producto H  
Name: producto, dtype: object
```

Tutorial



Expresiones Regulares

{desafío}
latam_





Expresiones regulares



Expresiones numéricas:

- `\d`: representa cualquier dígito individual (0-9).
- `\d+`: representa uno o más dígitos consecutivos.
- `\d*`: representa cero o más dígitos (es comodín)

```
1 import re
2
3 texto = """Hola, me encantaría participar en el concurso.
4         Mi nombre es Ana y mi correo es ana123@gmail.com.
5         Mi cuenta de Instagram es @ana_martinez y mi número de teléfono es +56987654321.
6         Estoy ansiosa por participar."""
7
8 print(re.findall(r'\d', texto))
9 print(re.findall(r'\d+', texto))
10 print(re.findall(r'\+\d+', texto))
```

`['1', '2', '3', '5', '6', '9', '8', '7', '6', '5', '4', '3', '2', '1']`
`['123', '56987654321']`
`['+56987654321']`



Expresiones regulares

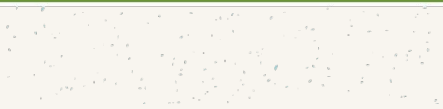


Para alfanuméricos:

- `\w`: representa cualquier carácter alfanumérico (letra, dígito y guión bajo).
- `\w+`: representa una o más secuencias consecutivas de los caracteres alfanuméricos.
- `\w*`: representa cero o más secuencias consecutivas de los caracteres alfanuméricos.

```
1 import re
2
3 texto = """Hola, me encantaría participar en el concurso.
4         Mi nombre es Ana y mi correo es ana123@gmail.com.
5         Mi cuenta de Instagram es @ana_martinez y mi número de teléfono es +56987654321.
6         Estoy ansiosa por participar."""
7
8 print(re.findall(r'\w+', texto))
9 print(re.findall(r'@\w+', texto))
10 print(re.findall(r'\w*@\w+', texto))
```

['Hola', 'me', 'encantaría', 'participar', 'en', 'el', 'concurso', 'Mi', 'nombre', 'es', 'Ana', 'y',
ticipar']
['@gmail', '@ana_martinez']
['ana123@gmail', '@ana_martinez']





La Librería re I



```
1 import re
2
3 texto = """Hola, me encantaría participar en el concurso.
4         Mi nombre es Ana y mi correo es ana123@gmail.com.
5         Mi cuenta de Instagram es @ana_martinez y mi número de teléfono es +56987654321.
6         Estoy ansiosa por participar."""
7
8 re.findall(r'\+\d+', texto)
```

['+56987654321']

Número de telefono

- `\+`: signo + literal.
- `\d+`: secuencia de dígitos largo 1 o más.



La Librería re II



```
1 import re
2
3 texto = """Hola, me encantaría participar en el concurso.
4         Mi nombre es Ana y mi correo es ana123@gmail.com.
5         Mi cuenta de Instagram es @ana_martinez y mi número de teléfono es +56987654321.
6         Estoy ansiosa por participar."""
7
8 re.findall(r'\@\w+', texto)
```

```
['@gmail', '@ana_martinez']
```

Usuarios o emails

- \@: arroba literal.
- \w+: alfanumerico largo 1 o más.



La Librería re II



```
1 import re
2
3 texto = """Hola, me encantaría participar en el concurso.
4         Mi nombre es Ana y mi correo es ana123@gmail.com.
5         Mi cuenta de Instagram es @ana_martinez y mi número de teléfono es +56987654321.
6         Estoy ansiosa por participar."""
7
8 re.findall(r'\w*\@\w+|\+\d+', texto)
```

['ana123@gmail', '@ana_martinez', '+56987654321']

email | usuario

- `\w*`: alfanumérico opcional (comodín).
- `\@`: arroba literal.
- `\w+`: alfanumerico largo 1 o más.

telefono

- `\+`: signo + literal.
- `\d+`: secuencia de dígitos largo 1 o más.



Caso de negocio



Asunto: Análisis de texto.

Estimado equipo ADL:



Necesitamos obtener información sobre los precios, hashtags y menciones de usuarios dentro del texto en posts de redes sociales de influencers más de moda, les adjunto una muestra de las publicaciones a las que estamos tratando de hacer seguimiento.

Esperamos que puedan agregar columnas a la base que tengan el precio, los hashtags y las menciones de usuario que encuentren en el texto de las publicaciones.

Además es importante que cuando nos envíen de vuelta la base, el texto original traiga estos datos censurados.

Por favor, háganos saber si necesita información adicional.

Saludos cordiales,
Lulu LaRue
@tiendaonline



Caso de negocio



Asunto: Requerimientos para el proyecto

Hola [Tu nombre],

Espero que estés teniendo un buen día. Quería compartir contigo los requerimientos para el proyecto que estamos trabajando para nuestro cliente @tiendaonline.

En este proyecto, necesitamos utilizar **expresiones regulares** para obtener información específica de un conjunto de textos y **guardar los resultados en columnas adicionales** del DataFrame. Para este requerimiento utiliza por favor la librería **re** de Python y extrae las coincidencias con el método **.findall()**.

Además, **utilizando las mismas expresiones regulares anteriores** debemos censurar los datos de precio, hashtag y usuarios del texto de la columna original. Para este requerimiento utiliza por favor el método **.replace()** de Pandas con el parámetro **regex= True**.

Por favor, hazme saber si tienes alguna pregunta o si necesitas más detalles sobre los requerimientos.

Saludos cordiales,

[Tu líder técnico]





Los Datos



df			
		texto	influencer
0		¡Me encanta este vestido de verano Talla 7! 🥰 Lo compré por \$50 en una tienda local. #ModaDeVerano #OutfitDelDia	Influencer1
1		Esta chaqueta de cuero es impresionante. 😍 La compré por \$150 en una venta de liquidación. #EstiloInvernal #Moda	Influencer2
2		Estoy enamorada de estos pantalones son número 36. 🥰 Compré estos por \$80 y valen cada centavo. #PantalonesDeMezclilla #ModaDeOtoño	Influencer3
3		Este vestido es todo lo que necesito para la temporada de bodas. 💍 Compré este por \$120 en una tienda de segunda mano. #VestidoDeNovia #ModaDeBoda	Influencer2
4		Estas botas son un verdadero robo. 🙌 Las compré por \$30 en una venta de garaje. #BotasDeInvierno #Liquidacion	Influencer3
5		Acabo de encontrar el top perfecto para mi próximo viaje. 🌴 Compré esto por \$45 en una tienda local. #ModaDeViaje #Outfit	Influencer4
6		¡Hola chicos! Hoy les quiero hablar de mi nuevo vestido favorito. Es perfecto para cualquier ocasión y lo compré en @tiendaonline a un súper precio de 25.99 dólares. #moda #vestido	Influencer4
7		¡Hola! Les cuento que encontré los jeans más cómodos y económicos en @tiendaonline. Solo me costaron 18.50 dólares y me encanta cómo me quedan. ¡Son perfectos para un look casual! #moda #fabulous	Influencer3
8		¿Tiene solo 1 o 2 camisas? Esta camisa que compré en @tiendaonline es perfecta. Solo me costó 14.99 dólares y el diseño es increíble. #ropitanueva	Influencer2
9		¡Chicas! Les cuento que encontré la chaqueta perfecta para este clima frío en @tiendaonline. Me costó 45.00 dólares y creo que es un precio justo por la calidad de la prenda. ¡Me encanta!	Influencer1
10		¡Hola a todos! Hoy quiero recomendarles estos pantalones que encontré en @tiendaonline. Solo me costaron 22.99 dólares, ¡un súper precio! #moda #pantalón	Influencer1



texto influencer

0	¡Me encanta este vestido de verano Talla 7! 🥰 Lo compré por \$50 en una tienda local. #ModaDeVerano #OutfitDelDía	Influencer1
1	Esta chaqueta de cuero es impresionante. 🥰 La compré por \$150 en una venta de liquidación. #EstiloInvernal #Moda	Influencer2
2	Estoy enamorada de estos pantalones son número 36. 🥰 Compré estos por \$80 y valen cada centavo. #PantalonesDeMezclilla #ModaDeOtono	Influencer3
3	Este vestido es todo lo que necesito para la temporada de bodas. 🥰 Compré este por \$120 en una tienda de segunda mano. #VestidoDeNovia #ModaDeBoda	Influencer2
4	Estas botas son un verdadero robo. 🥰 Las compré por \$30 en una venta de garaje. #BotasDelInvierno #Liquidacion	Influencer3
5	Acabo de encontrar el top perfecto para mi próximo viaje. 🥰 Compré esto por \$45 en una tienda local. #ModaDeViaje #Outfit	Influencer4
6	¡Hola chicos! Hoy les quiero hablar de mi nuevo vestido favorito. Es perfecto para cualquier ocasión y lo compré en @tiendaonline a un súper precio de 25.99 dólares. #moda #vestido	Influencer4
7	¡Hola! Les cuento que encontré los jeans más cómodos y económicos en @tiendaonline. Solo me costaron 18.50 dólares y me encanta cómo me quedan. ¡Son perfectos para un look casual. #moda #fabulous	Influencer3
8	¿Tiene solo 1 o 2 camisas? Esta camisa que compré en @tiendaonline es perfecta. Solo me costó 14.99 dólares y el diseño es increíble. #ropitanueva	Influencer2
9	¡Chicas! Les cuento que encontré la chaqueta perfecta para este clima frío en @tiendaonline. Me costó 45.00 dólares y creo que es un precio justo por la calidad de la prenda. ¡Me encanta!	Influencer1
10	¡Hola a todos! Hoy quiero recomendarles estos pantalones que encontré en @tiendaonline. Solo me costaron 22.99 dólares, ¡un súper precio! #moda #pantalón	Influencer1



Precios con signo de dolar



```
regex_precio_signo_dolar = r'\$\d+'  
df['texto'].replace(to_replace = regex_precio_signo_dolar, value = '$x', regex = True).to_frame()
```

	texto
0	¡Me encanta este vestido de verano Talla 7! 🥰 Lo compré por \$x en una tienda local. #ModaDeVerano #OutfitDelDia
1	Esta chaqueta de cuero es impresionante. 😍 La compré por \$x en una venta de liquidación. #EstiloInvernal #Moda
2	Estoy enamorada de estos pantalones son número 36. 🥰 Compré estos por \$x y valen cada centavo. #PantalonesDeMezclilla #ModaDeOtoño
3	Este vestido es todo lo que necesito para la temporada de bodas. 🥰 Compré este por \$x en una tienda de segunda mano. #VestidoDeNovia #ModaDeBoda
4	Estas botas son un verdadero robo. 🥰 Las compré por \$x en una venta de garaje. #BotasDelInvierno #Liquidacion
5	Acabo de encontrar el top perfecto para mi próximo viaje. 🥰 Compré esto por \$x en una tienda local. #ModaDeViaje #Outfit
6	¡Hola chicos! Hoy les quiero hablar de mi nuevo vestido favorito. Es perfecto para cualquier ocasión y lo compré en @tiendaonline a un súper precio de 25.99 dólares. #moda #vestido
7	¡Hola! Les cuento que encontré los jeans más cómodos y económicos en @tiendaonline. Solo me costaron 18.50 dólares y me encanta cómo me quedan. ¡Son perfectos para un look casual! #moda #fabulous
8	¿Tiene solo 1 o 2 camisas? Esta camisa que compré en @tiendaonline es perfecta. Solo me costó 14.99 dólares y el diseño es increíble. #ropitanueva
9	¡Chicas! Les cuento que encontré la chaqueta perfecta para este clima frío en @tiendaonline. Me costó 45.00 dólares y creo que es un precio justo por la calidad de la prenda. ¡Me encanta!
10	¡Hola a todos! Hoy quiero recomendarles estos pantalones que encontré en @tiendaonline. Solo me costaron 22.99 dólares, ¡un súper precio! #moda #pantalón



Precios en decimal



```
regex_precio_formato_decimal = r'\d+\.\d+'  
df['texto'].replace(to_replace = regex_precio_formato_decimal, value = '$x', regex = True).to_frame()
```

	texto
0	¡Me encanta este vestido de verano Talla 7! 🥰 Lo compré por \$50 en una tienda local. #ModaDeVerano #OutfitDelDia
1	Esta chaqueta de cuero es impresionante. 😍 La compré por \$150 en una venta de liquidación. #EstiloInvernal #Moda
2	Estoy enamorada de estos pantalones son número 36. 🥰 Compré estos por \$80 y valen cada centavo. #PantalonesDeMezclilla #ModaDeOtoño
3	Este vestido es todo lo que necesito para la temporada de bodas. 💍 Compré este por \$120 en una tienda de segunda mano. #VestidoDeNovia #ModaDeBoda
4	Estas botas son un verdadero robo. 🤑 Las compré por \$30 en una venta de garaje. #BotasDelInvierno #Liquidacion
5	Acabo de encontrar el top perfecto para mi próximo viaje. 🌴 Compré esto por \$45 en una tienda local. #ModaDeViaje #Outfit
6	¡Hola chicos! Hoy les quiero hablar de mi nuevo vestido favorito. Es perfecto para cualquier ocasión y lo compré en @tiendaonline a un súper precio de \$x dólares. #moda #vestido
7	¡Hola! Les cuento que encontré los jeans más cómodos y económicos en @tiendaonline. Solo me costaron \$x dólares y me encanta cómo me quedan. ¡Son perfectos para un look casual! #moda #fabulous
8	¿Tiene solo 1 o 2 camisas? Esta camisa que compré en @tiendaonline es perfecta. Solo me costó \$x dólares y el diseño es increíble. #ropitanueva
9	¡Chicas! Les cuento que encontré la chaqueta perfecta para este clima frío en @tiendaonline. Me costó \$x dólares y creo que es un precio justo por la calidad de la prenda. ¡Me encanta!
10	¡Hola a todos! Hoy quiero recomendarles estos pantalones que encontré en @tiendaonline. Solo me costaron \$x dólares, ¡un súper precio! #moda #pantalón



Hastags



```
regex_hashtag = r'\#\w+'  
df['texto'].replace(to_replace = regex_hashtag, value = '#x', regex = True).to_frame()
```

	texto
0	¡Me encanta este vestido de verano Talla 7! 🍷 Lo compré por \$50 en una tienda local. #x #x
1	Esta chaqueta de cuero es impresionante. 😊 La compré por \$150 en una venta de liquidación. #x #x
2	Estoy enamorada de estos pantalones son número 36. 🍷 Compré estos por \$80 y valen cada centavo. #x #x
3	Este vestido es todo lo que necesito para la temporada de bodas. 🍷 Compré este por \$120 en una tienda de segunda mano. #x #x
4	Estas botas son un verdadero robo. 🍷 Las compré por \$30 en una venta de garaje. #x #x
5	Acabo de encontrar el top perfecto para mi próximo viaje. 🍷 Compré esto por \$45 en una tienda local. #x #x
6	¡Hola chicos! Hoy les quiero hablar de mi nuevo vestido favorito. Es perfecto para cualquier ocasión y lo compré en @tiendaonline a un súper precio de 25.99 dólares. #x #x
7	¡Hola! Les cuento que encontré los jeans más cómodos y económicos en @tiendaonline. Solo me costaron 18.50 dólares y me encanta cómo me quedan. ¡Son perfectos para un look casual! #x #x
8	¿Tiene solo 1 o 2 camisas? Esta camisa que compré en @tiendaonline es perfecta. Solo me costó 14.99 dólares y el diseño es increíble. #x
9	¡Chicas! Les cuento que encontré la chaqueta perfecta para este clima frío en @tiendaonline. Me costó 45.00 dólares y creo que es un precio justo por la calidad de la prenda. ¡Me encanta!
10	¡Hola a todos! Hoy quiero recomendarles estos pantalones que encontré en @tiendaonline. Solo me costaron 22.99 dólares, ¡un súper precio! #x #x



Menciones de usuarios



```
regex_usuario = r'\@\w+'\n df['texto'].replace(to_replace = regex_usuario, value = '@x', regex = True).to_frame()
```

	texto
0	¡Me encanta este vestido de verano Talla 7! 🥰 Lo compré por \$50 en una tienda local. #ModaDeVerano #OutfitDelDia
1	Esta chaqueta de cuero es impresionante. 😊 La compré por \$150 en una venta de liquidación. #EstiloInvernal #Moda
2	Estoy enamorada de estos pantalones son número 36. 🥰 Compré estos por \$80 y valen cada centavo. #PantalonesDeMezclilla #ModaDeOtoño
3	Este vestido es todo lo que necesito para la temporada de bodas. 🥰 Compré este por \$120 en una tienda de segunda mano. #VestidoDeNovia #ModaDeBoda
4	Estas botas son un verdadero robo. 🥰 Las compré por \$30 en una venta de garaje. #BotasDelInvierno #Liquidacion
5	Acabo de encontrar el top perfecto para mi próximo viaje. 🌿 Compré esto por \$45 en una tienda local. #ModaDeViaje #Outfit
6	¡Hola chicos! Hoy les quiero hablar de mi nuevo vestido favorito. Es perfecto para cualquier ocasión y lo compré en @x a un súper precio de 25.99 dólares. #moda #vestido
7	¡Hola! Les cuento que encontré los jeans más cómodos y económicos en @x. Solo me costaron 18.50 dólares y me encanta cómo me quedan. ¡Son perfectos para un look casual! #moda #fabulous
8	¿Tiene solo 1 o 2 camisas? Esta camisa que compré en @x es perfecta. Solo me costó 14.99 dólares y el diseño es increíble. #ropitanueva
9	¡Chicas! Les cuento que encontré la chaqueta perfecta para este clima frío en @x. Me costó 45.00 dólares y creo que es un precio justo por la calidad de la prenda. ¡Me encanta!
10	¡Hola a todos! Hoy quiero recomendarles estos pantalones que encontré en @x. Solo me costaron 22.99 dólares, ¡un súper precio! #moda #pantalón



Reemplazo de datos



```
regex_precio_signo_dolar = r'\\$\\d+'
regex_precio_formato_decimal = r'\\d+\\.\\d+'
regex_hashtag = r'\\#\\w+'
regex_usuario = r'\\@\\w+'

df.replace({regex_precio_signo_dolar: '$x',
            regex_precio_formato_decimal: '$x',
            regex_hashtag: '#x',
            regex_usuario: '@x'
            }, regex=True)
```

	texto	influencer
0	¡Me encanta este vestido de verano Talla 7! 🥰 Lo compré por \$x en una tienda local. #x #x	Influencer1
1	Esta chaqueta de cuero es impresionante. 🥰 La compré por \$x en una venta de liquidación. #x #x	Influencer2
2	Estoy enamorada de estos pantalones son número 36. 🥰 Compré estos por \$x y valen cada centavo. #x #x	Influencer3
3	Este vestido es todo lo que necesito para la temporada de bodas. 🥰 Compré este por \$x en una tienda de segunda mano. #x #x	Influencer2
4	Estas botas son un verdadero robo. 🥰 Las compré por \$x en una venta de garaje. #x #x	Influencer3
5	Acabo de encontrar el top perfecto para mi próximo viaje. 🥰 Compré esto por \$x en una tienda local. #x #x	Influencer4
6	¡Hola chicos! Hoy les quiero hablar de mi nuevo vestido favorito. Es perfecto para cualquier ocasión y lo compré en @x a un súper precio de \$x dólares. #x #x	Influencer4
7	¡Hola! Les cuento que encontré los jeans más cómodos y económicos en @x. Solo me costaron \$x dólares y me encanta cómo me quedan. ¡Son perfectos para un look casual! #x #x	Influencer3
8	¿Tiene solo 1 o 2 camisas? Esta camisa que compré en @x es perfecta. Solo me costó \$x dólares y el diseño es increíble. #x	Influencer2
9	¡Chicas! Les cuento que encontré la chaqueta perfecta para este clima frío en @x. Me costó \$x dólares y creo que es un precio justo por la calidad de la prenda. ¡Me encanta!	Influencer1
10	¡Hola a todos! Hoy quiero recomendarles estos pantalones que encontré en @x. Solo me costaron \$x dólares, ¡un súper precio! #x #x	Influencer1



Extracción a nuevas columnas



```
import re

# Visualización del DataFrame mostrando solo 50 caracteres por celda.
pd.set_option('display.max_colwidth', 50)

# Expresiones regulares.
regexPrecioSignoDolar = r'\$\d+'
regexPrecioFormatoDecimal = r'\d+\.\d+'
regexHashtag = r'#\w+'
regexUsuario = r'\@\w+'

# Columnas nuevas.
df['regexPrecioSignoDolar'] = df['texto'].apply(lambda x: re.findall(regexPrecioSignoDolar, x))
df['regexPrecioFormatoDecimal'] = df['texto'].apply(lambda x: re.findall(regexPrecioFormatoDecimal, x))
df['regexHashtag'] = df['texto'].apply(lambda x: re.findall(regexHashtag, x))
df['regexUsuario'] = df['texto'].apply(lambda x: re.findall(regexUsuario, x))

df
```

	texto	influencer	regexPrecioSignoDolar	regexPrecioFormatoDecimal	regexHashtag	regexUsuario
0	¡Me encanta este vestido de verano Talla 7! 🍷 ...	Influencer1	[\$50]	[]	[#ModaDeVerano, #OutfitDelDia]	[]
1	Esta chaqueta de cuero es impresionante. 🍷 La ...	Influencer2	[\$150]	[]	[#EstiloInvernal, #Moda]	[]
2	Estoy enamorada de estos pantalones son número...	Influencer3	[\$80]	[]	[#PantalonesDeMezclilla, #ModaDeOtoño]	[]
3	Este vestido es todo lo que necesito para la t...	Influencer2	[\$120]	[]	[#VestidoDeNovia, #ModaDeBoda]	[]
4	Estas botas son un verdadero robo. 🍷 Las compr...	Influencer3	[\$30]	[]	[#BotasDelInvierno, #Liquidacion]	[]
5	Acabo de encontrar el top perfecto para mi pró...	Influencer4	[\$45]	[]	[#ModaDeViaje, #Outfit]	[]
6	¡Hola chicos! Hoy les quiero hablar de mi nuev...	Influencer4	[]	[25.99]	[#moda, #vestido]	[@tiendaonline]
7	¡Hola! Les cuento que encontré los jeans más c...	Influencer3	[]	[18.50]	[#moda, #fabulous]	[@tiendaonline]
8	¿Tiene solo 1 o 2 camisas? Esta camisa que com...	Influencer2	[]	[14.99]	[#ropitanueva]	[@tiendaonline]
9	¡Chicas! Les cuento que encontré la chaqueta p...	Influencer1	[]	[45.00]	[]	[@tiendaonline]
10	¡Hola a todos! Hoy quiero recomendarles estos ...	Influencer1	[]	[22.99]	[#moda, #pantalón]	[@tiendaonline]



Extracción a nuevas columnas



```
import re

regexPrecioSignoDolar = r'\${d+}'
regexPrecioFormatoDecimal = r'\d+\.\d+'
regexHashtag = r'\#[w+]'
regexUsuario = r'\@[w+]'

regexGeneral = rf'{regexPrecioSignoDolar}|{regexPrecioFormatoDecimal}|{regexHashtag}|{regexUsuario}'

df['regex_general'] = df['texto'].apply(lambda x: re.findall(regexGeneral, x)).to_frame()
df
```

	texto	influencer	regexPrecioSignoDolar	regexPrecioFormatoDecimal	regexHashtag	regexUsuario	regexGeneral
0	¡Me encanta este vestido de verano Talla 7! 🥰 ...	Influencer1	[\$50]	[]	[#ModaDeVerano, #OutfitDelDia]	[]	[\$50, #ModaDeVerano, #OutfitDelDia]
1	Esta chaqueta de cuero es impresionante. 🥰 La ...	Influencer2	[\$150]	[]	[#EstiloInvernal, #Moda]	[]	[\$150, #EstiloInvernal, #Moda]
2	Estoy enamorada de estos pantalones son número...	Influencer3	[\$80]	[]	[#PantalonesDeMezclilla, #ModaDeOtoño]	[]	[\$80, #PantalonesDeMezclilla, #ModaDeOtoño]
3	Este vestido es todo lo que necesito para la t...	Influencer2	[\$120]	[]	[#VestidoDeNovia, #ModaDeBoda]	[]	[\$120, #VestidoDeNovia, #ModaDeBoda]
4	Estas botas son un verdadero robo. 🥰 Las compr...	Influencer3	[\$30]	[]	[#BotasDelInvierno, #Liquidacion]	[]	[\$30, #BotasDelInvierno, #Liquidacion]
5	Acabo de encontrar el top perfecto para mi pró...	Influencer4	[\$45]	[]	[#ModaDeViaje, #Outfit]	[]	[\$45, #ModaDeViaje, #Outfit]
6	¡Hola chicos! Hoy les quiero hablar de mi nuev...	Influencer4	[]	[25.99]	[#moda, #vestido]	[@tiendaonline]	[@tiendaonline, 25.99, #moda, #vestido]
7	¡Hola! Les cuento que encontré los jeans más c...	Influencer3	[]	[18.50]	[#moda, #fabulous]	[@tiendaonline]	[@tiendaonline, 18.50, #moda, #fabulous]
8	¿Tiene solo 1 o 2 camisas? Esta camisa que com...	Influencer2	[]	[14.99]	[#ropitanueva]	[@tiendaonline]	[@tiendaonline, 14.99, #ropitanueva]
9	¡Chicas! Les cuento que encontré la chaqueta p...	Influencer1	[]	[45.00]	[]	[@tiendaonline]	[@tiendaonline, 45.00]
10	¡Hola a todos! Hoy quiero recomendarles estos ...	Influencer1	[]	[22.99]	[#moda, #pantalón]	[@tiendaonline]	[@tiendaonline, 22.99, #moda, #pantalón]



Extracción a nuevas columnas



```
import re

regexPrecioSignoDolar = r'\${d+}'
regexPrecioFormatoDecimal = r'\d+\.\d+'
regexHashtag = r'\#[w+]'
regexUsuario = r'\@[w+]'

regexGeneral = rf'{regexPrecioSignoDolar}|{regexPrecioFormatoDecimal}|{regexHashtag}|{regexUsuario}'

df['regex_general'] = df['texto'].apply(lambda x: re.findall(regexGeneral, x)).to_frame()
df
```

	texto	influencer	regexPrecioSignoDolar	regexPrecioFormatoDecimal	regexHashtag	regexUsuario	regexGeneral
0	¡Me encanta este vestido de verano Talla 7! 🥰 ...	Influencer1	[\$50]	[]	[#ModaDeVerano, #OutfitDelDia]	[]	[\$50, #ModaDeVerano, #OutfitDelDia]
1	Esta chaqueta de cuero es impresionante. 🥰 La ...	Influencer2	[\$150]	[]	[#EstiloInvernal, #Moda]	[]	[\$150, #EstiloInvernal, #Moda]
2	Estoy enamorada de estos pantalones son número...	Influencer3	[\$80]	[]	[#PantalonesDeMezclilla, #ModaDeOtoño]	[]	[\$80, #PantalonesDeMezclilla, #ModaDeOtoño]
3	Este vestido es todo lo que necesito para la t...	Influencer2	[\$120]	[]	[#VestidoDeNovia, #ModaDeBoda]	[]	[\$120, #VestidoDeNovia, #ModaDeBoda]
4	Estas botas son un verdadero robo. 🥰 Las compr...	Influencer3	[\$30]	[]	[#BotasDelInvierno, #Liquidacion]	[]	[\$30, #BotasDelInvierno, #Liquidacion]
5	Acabo de encontrar el top perfecto para mi pró...	Influencer4	[\$45]	[]	[#ModaDeViaje, #Outfit]	[]	[\$45, #ModaDeViaje, #Outfit]
6	¡Hola chicos! Hoy les quiero hablar de mi nuev...	Influencer4	[]	[25.99]	[#moda, #vestido]	[@tiendaonline]	[@tiendaonline, 25.99, #moda, #vestido]
7	¡Hola! Les cuento que encontré los jeans más c...	Influencer3	[]	[18.50]	[#moda, #fabulous]	[@tiendaonline]	[@tiendaonline, 18.50, #moda, #fabulous]
8	¿Tiene solo 1 o 2 camisas? Esta camisa que com...	Influencer2	[]	[14.99]	[#ropitanueva]	[@tiendaonline]	[@tiendaonline, 14.99, #ropitanueva]
9	¡Chicas! Les cuento que encontré la chaqueta p...	Influencer1	[]	[45.00]	[]	[@tiendaonline]	[@tiendaonline, 45.00]
10	¡Hola a todos! Hoy quiero recomendarles estos ...	Influencer1	[]	[22.99]	[#moda, #pantalón]	[@tiendaonline]	[@tiendaonline, 22.99, #moda, #pantalón]

Tutorial



¡Esperamos sea de utilidad!



Reemplazo de Datos con Pandas

{desafío}
latam_

