

✓ Reto 2: Filter

1. Objetivos:

- Practicar el uso de `filter` para filtrar los datos en una `lista`
-

2. Desarrollo:

✓ a) Limpiando datos nulos

Debajo tenemos una `lista` que incluye datos acerca de las edades de las personas que han atendido a un curso de Cocina Medieval (ya sabes: puerco al horno, manzanas asadas, aguardiente, sangre fresca de tus enemigos). Algunas de las personas que atendieron no quisieron dar su edad. Es por eso que algunos de los elementos son `None`:

```
edades = [12, 16, 19, None, 21, 25, 24, None, None, 16, 17, 25, 23, 28, None, 23, 35, 59, 67, None, 34, 21, 23, 15, 14, None, 18, 24, 23, 17]
```

Queremos realizar una pequeña visualización (un histograma, que ya aprenderás a hacer más tarde) con nuestros datos. Pero no nos interesan los datos que vienen como `None`. Escribe una función llamada `valor_no_es_none` que reciba un `valor`, cheque si el `valor` es `None`, regrese `False` si el `valor` es `None` o regrese `True` si el `valor` **no** es `None`. Después úsala para filtrar tus datos:

```
## Tu función va aquí
def valor_no_es_none(valor):
```

```
    if valor == None:
        return False
    else:
        return True
```

```
edades_filtradas = list(filter(valor_no_es_none, edades))
```

Pídele a tu experta la función de verificación `crear_histograma_con` (encontrada en el archivo `helpers.py` de la carpeta donde se encuentra este Reto), pégala debajo y corre la celda para verificar tu resultado:

```
# Pega aquí la función de verificación
def crear_histograma_con(datos):
    import seaborn as sns
    import numpy as np

    sns.histplot(datos, kde=False, bins=len(np.unique(datos)))

crear_histograma_con(edades_filtradas)
```

► Solución

✓ b) Filtrando datos atípicos

Aquí tenemos una lista que contiene datos acerca de los sueldos (cada número representa "miles de pesos") de los empleados de EyePoker Inc. (la empresa donde se producen los mejores picadores de ojos en todo el Hemisferio Occidental):

```
sueldos = [26, 32, 26, 30, 30, 32, 28, 30, 28, 110, 34, 30, 28, 26, 28, 30, 28, 85, 25, 30, 34, 34, 30, 30, 120, 28, 28, 120, 125]
```

En general todos los sueldos se encuentran en un rango bastante restringido, pero tenemos algunos datos sobre sueldos "anormalmente" grandes. Los sueldos tan grandes son los de los ejecutivos, que claramente no tienen ninguna noción de "justicia" (eso pasa cuando tus picadores de ojos son los mejores de todo el Hemisferio Occidental). Nosotros queremos usar el promedio para tener una idea de cuál es el sueldo típico en esta empresa. Nuestros valores atípicos (los sueldos anormalmente grandes) van a arruinar nuestro cálculo.

Mira cuál es el sueldo típico si no filtramos nuestros valores anormalmente grandes:

```
print(f'El sueldo "típico" en EyePoker Inc. es de {sum(sueldos) / len(sueldos)}')
```

```
El sueldo "típico" en EyePoker Inc. es de 43.62068965517241
```

Para corregir esto haz una función llamada `numero_es_menor_que_40` que descarte los números mayores de 40, y úsala para filtrar la lista `sueldos`, para tener un cálculo más apropiado del sueldo típico en esta empresa.

```
## Tu función va aquí
def numero_es_menor_que_40(numero):
```

```
    if numero < 40:
        return True
    else:
        return False
```

```
sueldos_filtrados = list(filter(numero_es_menor_que_40, sueldos))
```

```
print(f'El sueldo "típico" en EyePoker Inc. es de {sum(sueldos_filtrados) / len(sueldos_filtrados)}')
```

```
El sueldo "típico" en EyePoker Inc. es de 29.375
```

► Solución

Comienza a programar o [generar](#) con IA.