In [ ]:
```
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
import seaborn as sns
```

In [ ]:
```

```

In [ ]:
```
salud=pd.read_csv("salud1.csv",encoding='latin-1',sep=";",decimal=",")
```

In [ ]:
```
salud.head()
```

Out[ ]:

	ID	COLESTEROL	HDL	EDAD	FUMA	NEDU	SEXO	PAS	PAD	TALLA
0	1	161.6146	43.2533	21	No	> 12 ños	Femenino	109.3657	69.7861	159.224(
1	2	221.9609	45.9553	59	No	8 - 12 años	Masculino	149.1800	89.8402	171.264{
2	3	254.8755	87.2984	51	No	8 - 12 años	Femenino	139.1653	82.8118	163.593
3	4	187.3437	57.3342	50	No	> 12 ños	Masculino	174.1770	113.4944	166.872{
4	5	252.3962	31.7201	60	Si	< 8 años	Masculino	150.7024	78.7364	147.698{

In [ ]:
```
salud.tail()
```

Out[ ]:

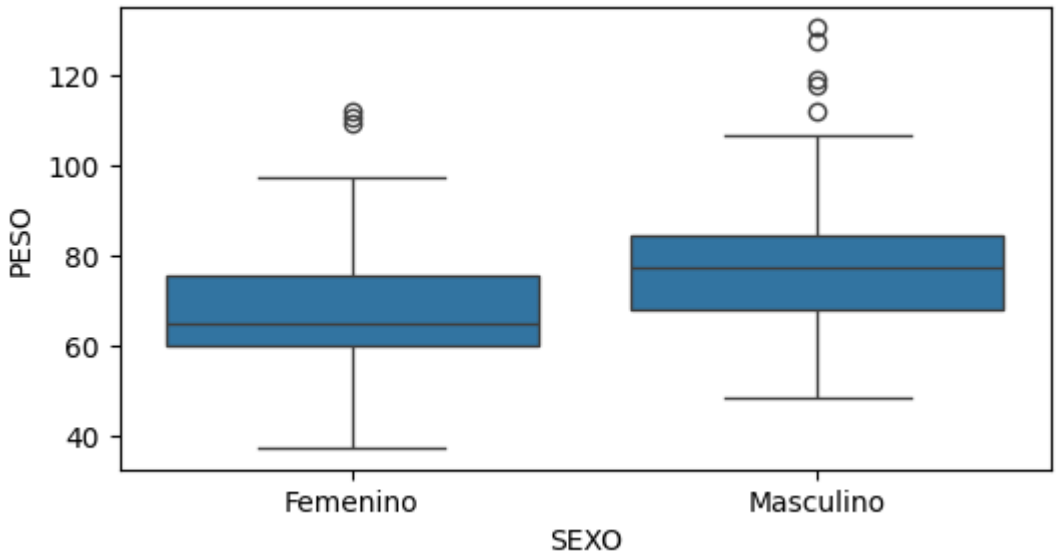
	ID	COLESTEROL	HDL	EDAD	FUMA	NEDU	SEXO	PAS	PAD	TAI
345	346	128.8601	28.3728	17	No	< 8 años	Masculino	127.5279	75.3712	145.5
346	347	114.3015	28.5680	25	Si	8 - 12 años	Masculino	115.9097	56.6316	159.3
347	348	146.3468	25.9187	61	No	8 - 12 años	Masculino	141.0092	80.3581	166.5
348	349	172.3351	52.3295	22	Si	< 8 años	Femenino	125.2207	76.2409	151.9
349	350	231.9689	51.8565	66	No	< 8 años	Femenino	138.1179	77.8108	178.4

In [ ]:
```
round(salud.describe(),2)
```

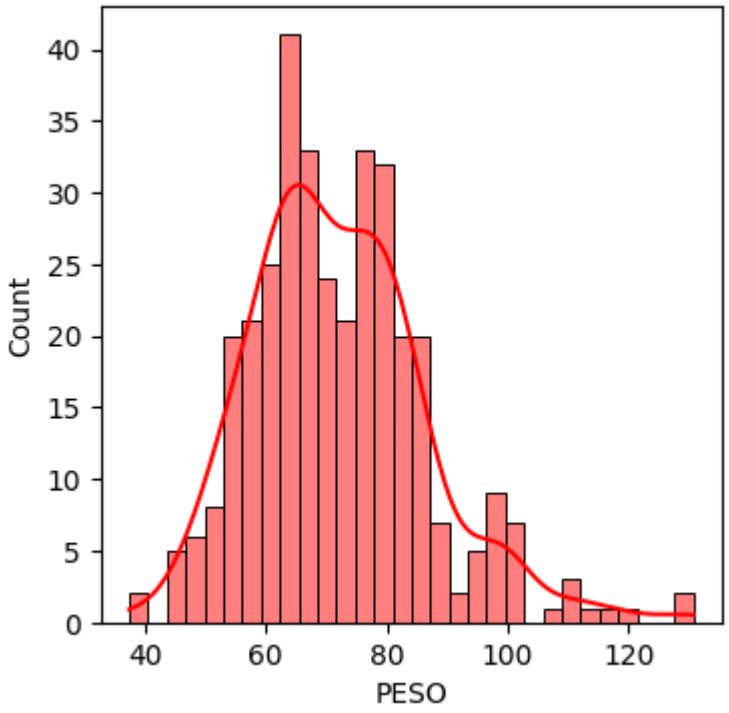
Out[ ]:

	ID	COLESTEROL	HDL	EDAD	PAS	PAD	TALLA	CINTURA	CUELLO	I
count	350.00	350.00	350.00	350.00	350.00	350.00	350.00	350.00	350.00	350
mean	175.50	192.77	48.00	44.65	128.19	76.88	161.48	92.75	37.41	27
std	101.18	43.28	12.95	18.05	21.36	11.25	9.31	13.68	4.87	6
min	1.00	93.67	21.48	15.00	91.21	51.01	139.88	63.11	25.18	11
25%	88.25	162.07	38.78	30.00	112.81	69.37	154.40	83.80	34.67	23
50%	175.50	189.78	46.94	44.00	123.94	75.42	161.08	91.78	37.16	27
75%	262.75	221.84	54.45	58.00	139.00	84.59	167.93	100.32	39.72	31
max	350.00	311.68	97.16	93.00	221.79	116.52	190.90	133.16	92.89	51

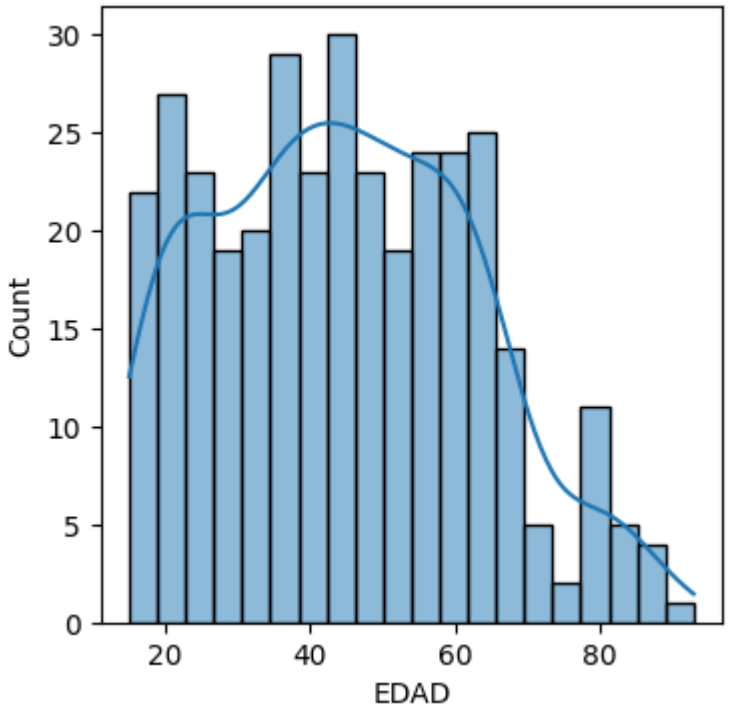
In [ ]:
```
plt.figure(figsize=(6,3))
sns.boxplot(data=salud, y="PESO", x="SEXO")
plt.show()
```



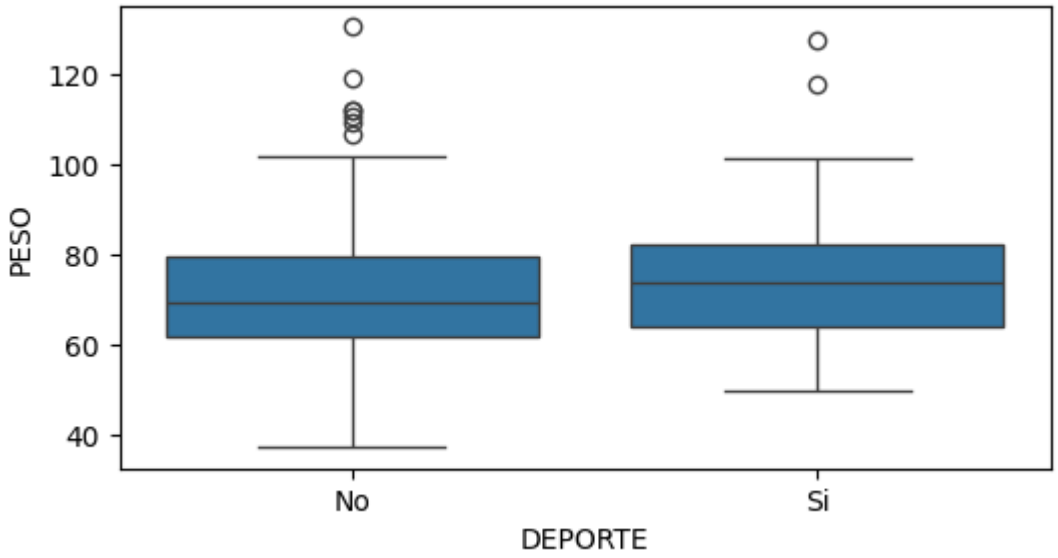
In [ ]:
```
plt.figure(figsize=(4,4))
sns.histplot(data=salud,x="PESO",bins=30,kde=True,color="red")
plt.savefig("histogramaPESO.png")
plt.show()
```



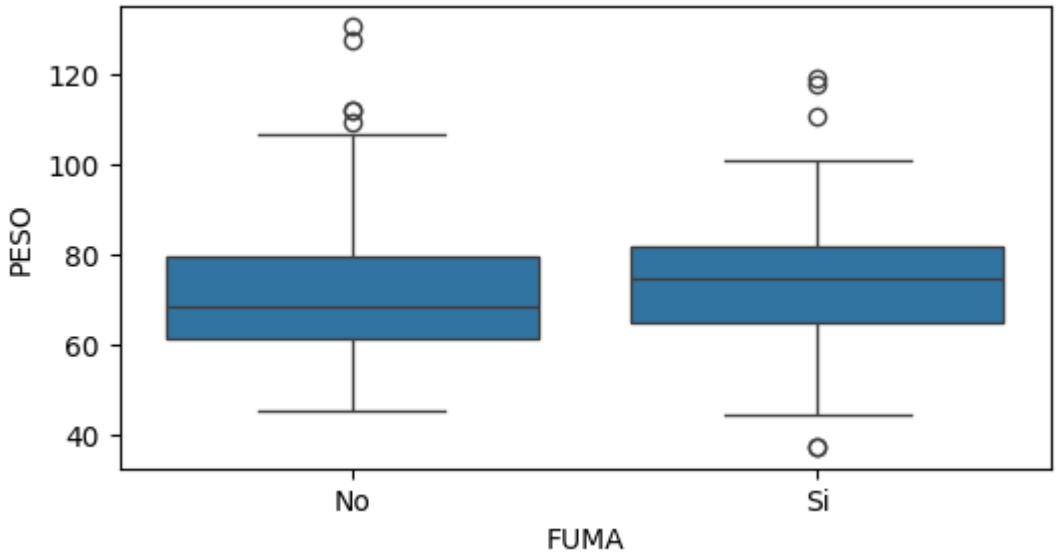
In [ ]:
```
plt.figure(figsize=(4,4))
sns.histplot(data=salud,x="EDAD",bins=20,kde=True)
plt.savefig("histogramaEDAD.png")
plt.show()
```



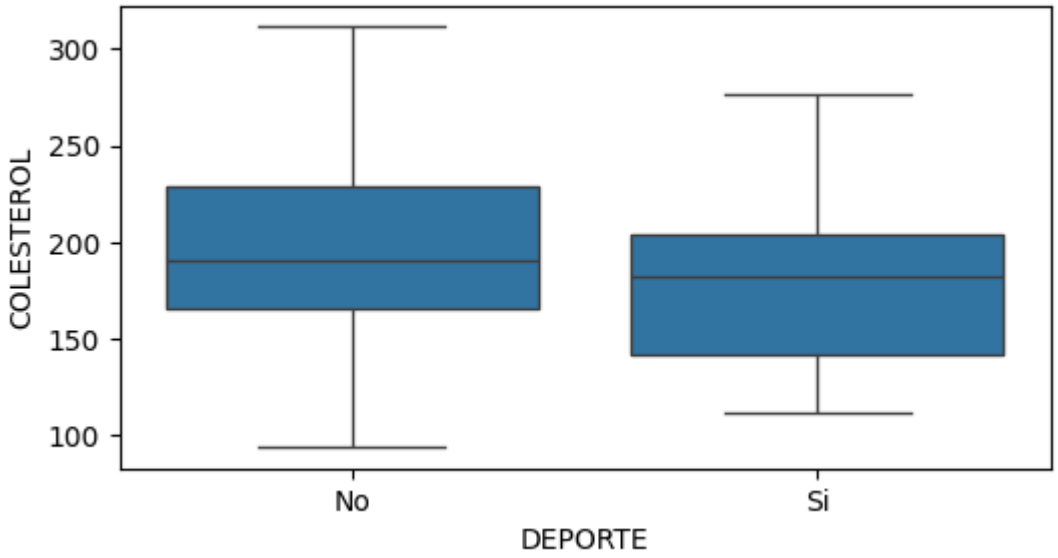
In [ ]:
```
plt.figure(figsize=(6,3))
sns.boxplot(data=salud, y="PESO", x="DEPORTE")
plt.show()
```



In [ ]:
```
plt.figure(figsize=(6,3))
sns.boxplot(data=salud, y="PESO", x="FUMA")
plt.show()
```



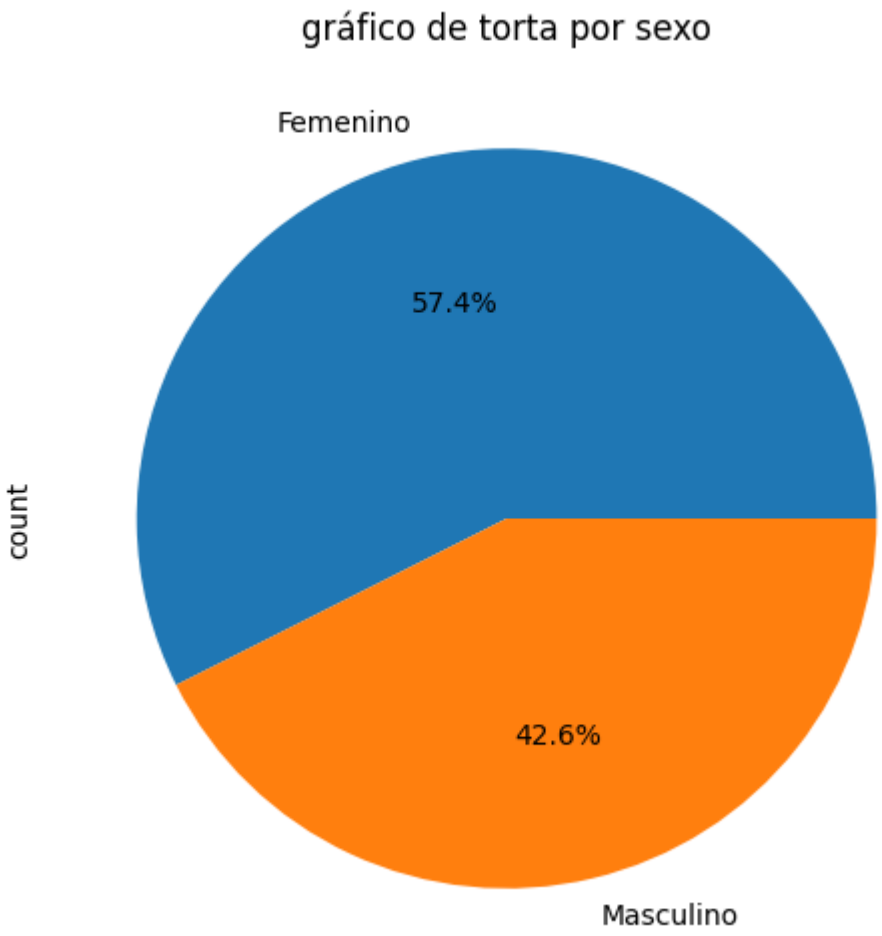
In [ ]:
```
plt.figure(figsize=(6,3))
sns.boxplot(data=salud, y="COLESTEROL", x="DEPORTE")
plt.show()
```



In [ ]:
```
# gráficos de  torta y  barras
serie=salud.SEXO.value_counts()
serie
```

Out[ ]:
```
SEXO
Femenino     201
Masculino    149
Name: count, dtype: int64
```

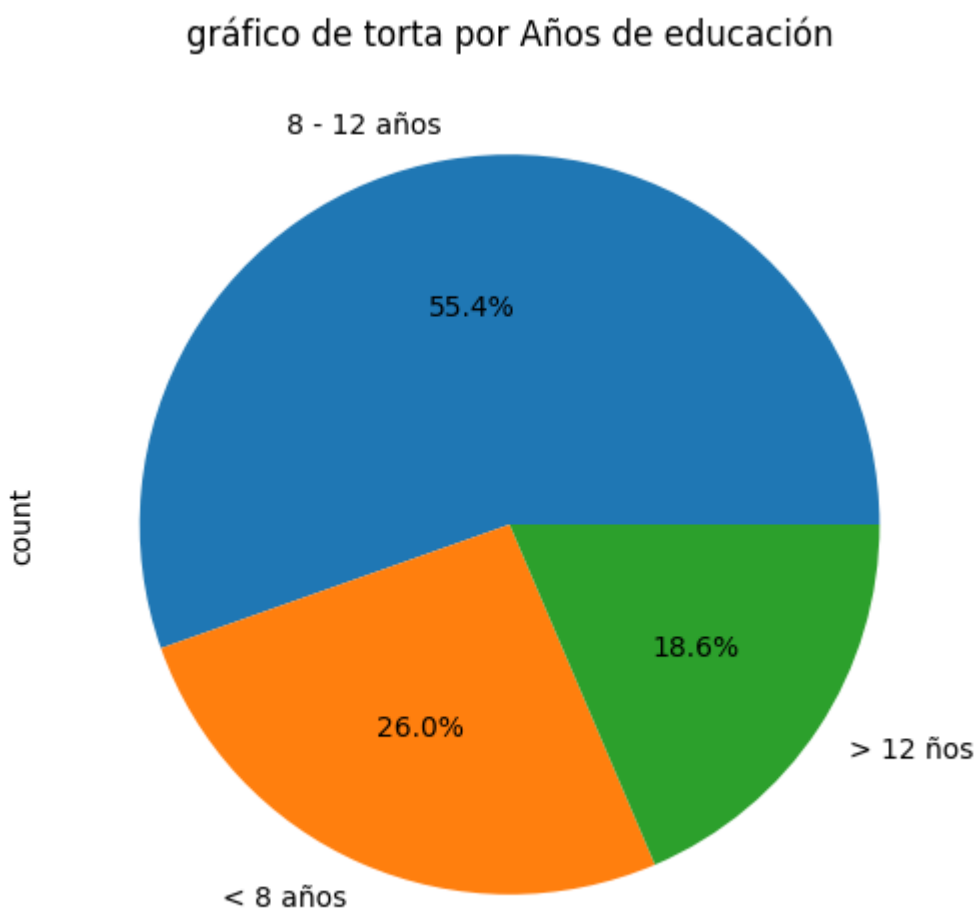
In [ ]:
```
plt.figure(figsize=(6,6))
serie.plot.pie(autopct='%1.1f%%')
plt.title("gráfico de torta por sexo")
plt.show()
```



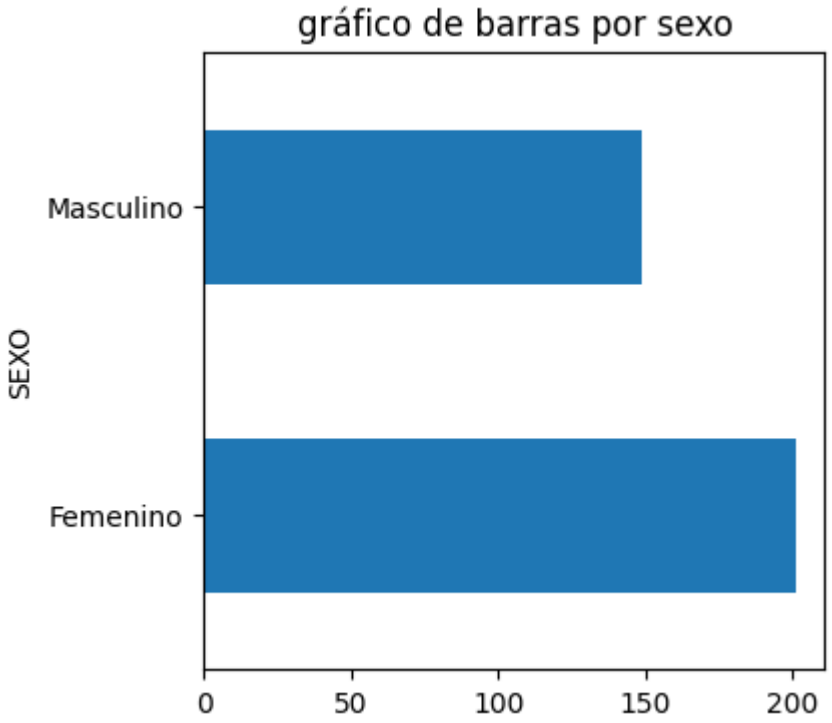
In [ ]:
```
serie1=salud.NEDU.value_counts()
serie1
```

Out[ ]:
```
NEDU
8 - 12 años    194
< 8 años        91
> 12 ños        65
Name: count, dtype: int64
```

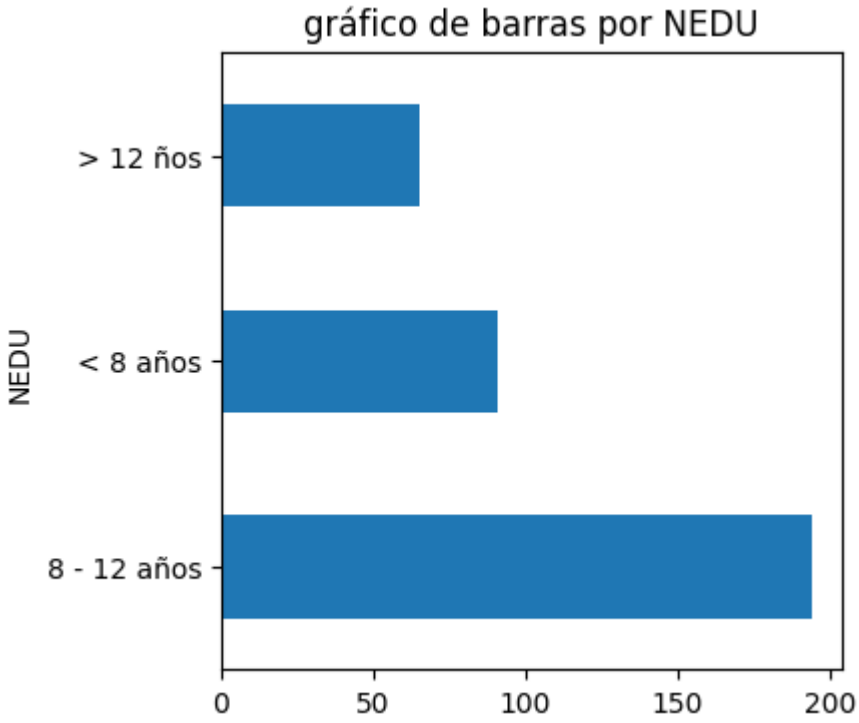
In [ ]:
```
plt.figure(figsize=(6,6))
serie1.plot.pie(autopct='%1.1f%%')
plt.title("gráfico de torta por Años de educación")
plt.show()
```



In [ ]:
```
plt.figure(figsize=(4,4))
serie.plot.barh()
plt.title("gráfico de barras por sexo")
plt.show()
```



In [ ]:
```
plt.figure(figsize=(4,4))
serie1.plot.barh()
plt.title("gráfico de barras por NEDU")
plt.show()
```



In [ ]:
```
salud['FUMA'].describe()
```

Out[ ]:
```
count     350
unique      2
top        No
freq      230
Name: FUMA, dtype: object
```

medidas descriptivas

In [ ]:
```
mediana=salud["EDAD"].median()
mediana
```

Out[ ]: 44.0

In [ ]:
```
varianza=salud['EDAD'].var(ddof=1)
varianza
```

Out[ ]: 325.82084322554243

In [ ]:
```
DE=np.sqrt(salud['EDAD'].var(ddof=1))
DE
```

Out[ ]: 18.050508115439367

In [ ]:
```
DE=np.sqrt(varianza)
DE
```

Out[ ]: 18.050508115439367

In [ ]:
```
varianza1=(DE)**2
varianza1
```

Out[ ]: 325.82084322554243

In [ ]:
```
promedioedad=salud['EDAD'].mean()
promedioedad
```

Out[ ]: 44.651428571428575

In [ ]:
```
CVEDAD=(DE/promedioedad)*100 # coefienciente de  variación edad
CVEDAD
```

Out[ ]: 40.42537650629497

In [ ]:
```
salud['EDAD'].quantile(0.18)
```

Out[ ]: 24.82