

▼ Prueba #1

- materia: Simulacion
- autora: Helen Companioni vargas

```
#librerias
import matplotlib.pyplot as plt
import numpy as np
import pandas as pd
import io
from google.colab import files
import plotly.graph_objects as go
```

▼ Orden

• Diseñe y desarrolle un modelo y/o script que permita describir el siguiente caso aplicado: ◦ Se tiene los datos del ecuador
(https://www.ecuadorencifras.gob.ec/documentos/web-inec/EMPLEO/2021/Octubre-2021/202110_Tabulados_Mercado_Laboral_CSV.zip):

• Generar graficas y reportes del total de personas empleadas y desempleadas por año.

archivo_fifa = files.upload()

Elegir archivos | 1.Poblaciones.csv

- **1.Poblaciones.csv**(application/vnd.ms-excel) - 85911 bytes, last modified: 1/12/2021 - 100% done

Saving 1.Poblaciones.csv to 1.Poblaciones.csv

```
import pandas as pd
df = pd.read_csv('1.Poblaciones.csv', sep = ';', encoding = 'latin1',header=1, error_bad_lines=False) #error_bad_lines is avoid single line error
#df['Área'] = df[['Área', 'Unnamed: 5']].apply(' '.join, axis=1)
#df.drop(df.columns[df.columns.str.contains('unnamed',case = False)],axis = 1, inplace = True)
df = df.rename(columns={'Área':'Área urbana ', 'Unnamed: 5':'Área rural','Nacional':'Nacional/Total', 'Sexo':'Sexo-Hombre', 'Unnamed: 7':'Sexo-Mujer'})
df.drop(df.columns[df.columns.str.contains('unnamed',case = False)],axis = 1, inplace = True)
df
```

	Encuesta	Periodo	Indicadores	Nacional/Total	Área urbana	Área rural	Sexo-Hombre	Sexo-Mujer
0	NaN	NaN	NaN	Total	Urbana	Rural	Hombre	Mujer
1	ENEMDU	dic-07	Población Total	13.682.302	9.066.209	4.616.093	6.768.646	6.913.656
2	ENEMDU	dic-07	Población menor de 15 años	4.372.812	2.723.124	1.649.688	2.226.618	2.146.194
3	ENEMDU	dic-07	Población en Edad de Trabajar (PET)	9.309.490	6.343.085	2.966.404	4.542.028	4.767.462
4	ENEMDU	dic-07	Población Económicamente Activa	6.336.029	4.227.702	2.108.328	3.777.232	2.558.798
...
896	ENEMDU*	oct-21	Desempleo Abierto	336.101	291.606	44.495	150.276	185.824
897	ENEMDU*	oct-21	Desempleo Oculto	48.103	27.600	20.503	25.975	22.128
898	ENEMDU*	oct-21	Desempleo Cesante	298.846	257.856	40.991	140.223	158.623
899	ENEMDU*	oct-21	Desempleo Nuevo	85.358	61.351	24.007	36.028	49.330
900	ENEMDU*	oct-21	Población Económicamente Inactiva	4.330.241	3.307.420	1.022.821	1.324.745	3.005.495

901 rows x 8 columns

```
df.drop([0])
```

	Encuesta	Periodo	Indicadores	Nacional/Total	Área urbana	Área rural	Sexo-Hombre	Sexo-Mujer
1	ENEMDU	dic-07	Población Total	13.682.302	9.066.209	4.616.093	6.768.646	6.913.656
2	ENEMDU	dic-07	Población menor de 15 años	4.372.812	2.723.124	1.649.688	2.226.618	2.146.194
3	ENEMDU	dic-07	Población en Edad de Trabajar (PET)	9.309.490	6.343.085	2.966.404	4.542.028	4.767.462
4	ENEMDU	dic-07	Población Económicamente Activa	6.336.029	4.227.702	2.108.328	3.777.232	2.558.798
5	ENEMDU	dic-07	Empleo	6.019.332	3.971.040	2.048.292	3.632.314	2.387.018
...
896	ENEMDU*	oct-21	Desempleo Abierto	336.101	291.606	44.495	150.276	185.824

df.columns

```
Index(['Encuesta', 'Periodo', 'Indicadores', 'Nacional/Total', 'Área urbana ',
      'Área rural', 'Sexo-Hombre', 'Sexo-Mujer'],
      dtype='object')
```

```
df1= df.drop([0]).copy()
df1
```

	Encuesta	Periodo	Indicadores	Nacional/Total	Área urbana	Área rural	Sexo-Hombre	Sexo-Mujer
1	ENEMDU	dic-07	Población Total	13.682.302	9.066.209	4.616.093	6.768.646	6.913.656
2	ENEMDU	dic-07	Población menor de 15 años	4.372.812	2.723.124	1.649.688	2.226.618	2.146.194
3	ENEMDU	dic-07	Población en Edad de Trabajar (PET)	9.309.490	6.343.085	2.966.404	4.542.028	4.767.462
4	ENEMDU	dic-07	Población Económicamente Activa	6.336.029	4.227.702	2.108.328	3.777.232	2.558.798
5	ENEMDU	dic-07	Empleo	6.019.332	3.971.040	2.048.292	3.632.314	2.387.018
...
896	ENEMDU*	oct-21	Desempleo Abierto	336.101	291.606	44.495	150.276	185.824
897	ENEMDU*	oct-21	Desempleo Oculto	48.103	27.600	20.503	25.975	22.128
898	ENEMDU*	oct-21	Desempleo Cesante	298.846	257.856	40.991	140.223	158.623
899	ENEMDU*	oct-21	Desempleo Nuevo	85.358	61.351	24.007	36.028	49.330
900	ENEMDU*	oct-21	Población Económicamente Inactiva	4.330.241	3.307.420	1.022.821	1.324.745	3.005.495

900 rows × 8 columns

```
df6 = df1.copy()
```

```
df1['Periodo']
```

```
1      dic-07
2      dic-07
3      dic-07
4      dic-07
5      dic-07
...
896    oct-21
897    oct-21
898    oct-21
899    oct-21
900    oct-21
Name: Periodo, Length: 900, dtype: object
```

```
df1['Periodo']=df1['Periodo'].replace('\-', '20',regex=True)
df1
```

	Encuesta	Periodo	Indicadores	Nacional/Total	Área urbana	Área rural	Sexo-Hombre	Sexo-Mujer
1	ENEMDU	dic2007	Población Total	13.682.302	9.066.209	4.616.093	6.768.646	6.913.656
2	ENEMDU	dic2007	Población menor de 15 años	4.372.812	2.723.124	1.649.688	2.226.618	2.146.194
3	ENEMDU	dic2007	Población en Edad de Trabajar (PET)	9.309.490	6.343.085	2.966.404	4.542.028	4.767.462
4	ENEMDU	dic2007	Población Económicamente Activa	6.336.029	4.227.702	2.108.328	3.777.232	2.558.798
5	ENEMDU	dic2007	Empleo	6.019.332	3.971.040	2.048.292	3.632.314	2.387.018
...
896	ENEMDU*	oct2021	Desempleo Abierto	336.101	291.606	44.495	150.276	185.824

```
df1['Periodo'] = df1['Periodo'].str.replace(r"[a-zA-Z]", '')
df1
```

	Encuesta	Periodo	Indicadores	Nacional/Total	Área urbana	Área rural	Sexo-Hombre	Sexo-Mujer
1	ENEMDU	2007	Población Total	13.682.302	9.066.209	4.616.093	6.768.646	6.913.656
2	ENEMDU	2007	Población menor de 15 años	4.372.812	2.723.124	1.649.688	2.226.618	2.146.194
3	ENEMDU	2007	Población en Edad de Trabajar (PET)	9.309.490	6.343.085	2.966.404	4.542.028	4.767.462
4	ENEMDU	2007	Población Económicamente Activa	6.336.029	4.227.702	2.108.328	3.777.232	2.558.798
5	ENEMDU	2007	Empleo	6.019.332	3.971.040	2.048.292	3.632.314	2.387.018
...
896	ENEMDU*	2021	Desempleo Abierto	336.101	291.606	44.495	150.276	185.824
897	ENEMDU*	2021	Desempleo Oculto	48.103	27.600	20.503	25.975	22.128
898	ENEMDU*	2021	Desempleo Cesante	298.846	257.856	40.991	140.223	158.623
899	ENEMDU*	2021	Desempleo Nuevo	85.358	61.351	24.007	36.028	49.330
900	ENEMDU*	2021	Población Económicamente Inactiva	4.330.241	3.307.420	1.022.821	1.324.745	3.005.495

900 rows × 8 columns

```
df2= df1.copy()
```

```
df2['Nacional/Total']=df2['Nacional/Total'].replace('\.',',',regex=True)
df2
```

	Encuesta	Periodo	Indicadores	Nacional/Total	Área urbana	Área rural	Sexo-Hombre	Sexo-Mujer
1	ENEMDU	2007	Población Total	13682302	9.066.209	4.616.093	6.768.646	6.913.656
2	ENEMDU	2007	Población menor de 15 años	4372812	2.723.124	1.649.688	2.226.618	2.146.194
3	ENEMDU	2007	Población en Edad de Trabajar (PET)	9309490	6.343.085	2.966.404	4.542.028	4.767.462
4	ENEMDU	2007	Población Económicamente Activa	6336029	4.227.702	2.108.328	3.777.232	2.558.798
5	ENEMDU	2007	Empleo	6019332	3.971.040	2.048.292	3.632.314	2.387.018
...
896	ENEMDU*	2021	Desempleo Abierto	336101	291.606	44.495	150.276	185.824
897	ENEMDU*	2021	Desempleo Oculto	48103	27.600	20.503	25.975	22.128
898	ENEMDU*	2021	Desempleo Cesante	298846	257.856	40.991	140.223	158.623
899	ENEMDU*	2021	Desempleo Nuevo	85358	61.351	24.007	36.028	49.330
900	ENEMDU*	2021	Población Económicamente Inactiva	4330241	3.307.420	1.022.821	1.324.745	3.005.495

900 rows × 8 columns

```
df2['Nacional/Total'] = df2['Nacional/Total'].astype(float, errors = 'raise')
```

```
df2['Nacional/Total']
```

1	13682302.0
2	4372812.0
3	9309490.0

```
4      6336029.0
5      6019332.0
...
896     336101.0
897     48103.0
898     298846.0
899      85358.0
900     4330241.0
Name: Nacional/Total, Length: 900, dtype: float64
```

```
list = ['Población Total','Población Económicamente Inactiva','Población en Edad de Trabajar (PET)','Poblacion Total','Población menor de 15 años',
conte= df2.shape[0]
```

```
for i in range(len(list)):
    for j in range(conte):
        df2 = df2.drop(df2[df2['Indicadores']== list[i]].index)
```

df2

	Encuesta	Periodo	Indicadores	Nacional/Total	Área urbana	Área rural	Sexo-Hombre	Sexo-Mujer
5	ENEMDU	2007	Empleo	6019332.0	3.971.040	2.048.292	3.632.314	2.387.018
6	ENEMDU	2007	Empleo Adecuado/Pleno	2737158.0	2.236.440	500.719	1.907.451	829.707
7	ENEMDU	2007	Subempleo	1155872.0	687.714	468.158	679.118	476.753
8	ENEMDU	2007	Subempleo por insuficiencia de tiempo de trabajo	860315.0	519.627	340.688	474.355	385.960
9	ENEMDU	2007	Subempleo por insuficiencia de ingresos	295556.0	168.087	127.470	204.763	90.793
...
895	ENEMDU*	2021	Desempleo	384204.0	319.206	64.997	176.251	207.953
896	ENEMDU*	2021	Desempleo Abierto	336101.0	291.606	44.495	150.276	185.824
897	ENEMDU*	2021	Desempleo Oculto	48103.0	27.600	20.503	25.975	22.128
898	ENEMDU*	2021	Desempleo Cesante	298846.0	257.856	40.991	140.223	158.623
899	ENEMDU*	2021	Desempleo Nuevo	85358.0	61.351	24.007	36.028	49.330

653 rows x 8 columns

df4=df2.copy()

df5=df2.copy()

```
df3=df2.copy()
empleo=['Empleo','Empleo Adecuado/Pleno','Otro Empleo no pleno','Empleo no clasificado','Empleo no remunerado','Subempleo','Subempleo por insuficien
desempleo=['Desempleo','Desempleo Cesante','Desempleo Abierto','Desempleo Nuevo','Desempleo Oculto']
df3['Indicadores'] = df3['Indicadores'].str.replace('Subempleo por insuficiencia de tiempo de trabajo','Empleo')
df3['Indicadores'] = df3['Indicadores'].str.replace('Subempleo por insuficiencia de ingresos','Empleo')
df3.drop(df3[df3['Indicadores']=='Población Total'].index)
```

```
for j in empleo:
    for i in df3['Indicadores']:
        if j==i:
            df3['Indicadores']=df3['Indicadores'].replace(i,'Empleo',regex=True)
```

```
for j in desempleo:
    for i in df3['Indicadores']:
        if j==i:
            df3['Indicadores']=df3['Indicadores'].replace(i,'Desempleo',regex=True)
```

df3

	Encuesta	Periodo	Indicadores	Nacional/Total	Área urbana	Área rural	Sexo-Hombre	Sexo-Mujer
5	ENEMDU	2007	Empleo	6019332.0	3.971.040	2.048.292	3.632.314	2.387.018
6	ENEMDU	2007	Empleo	2737158.0	2.236.440	500.719	1.907.451	829.707
7	ENEMDU	2007	Empleo	1155872.0	687.714	468.158	679.118	476.753
8	ENEMDU	2007	Empleo	860315.0	519.627	340.688	474.355	385.960
9	ENEMDU	2007	Empleo	295556.0	168.087	127.470	204.763	90.793
...
895	ENEMDU*	2021	Desempleo	384204.0	319.206	64.997	176.251	207.953
896	ENEMDU*	2021	Desempleo	336101.0	291.606	44.495	150.276	185.824

df3['Nacional/Total'] = df3['Nacional/Total'].astype('int64', errors = 'raise')

df3

	Encuesta	Periodo	Indicadores	Nacional/Total	Área urbana	Área rural	Sexo-Hombre	Sexo-Mujer
5	ENEMDU	2007	Empleo	6019332	3.971.040	2.048.292	3.632.314	2.387.018
6	ENEMDU	2007	Empleo	2737158	2.236.440	500.719	1.907.451	829.707
7	ENEMDU	2007	Empleo	1155872	687.714	468.158	679.118	476.753
8	ENEMDU	2007	Empleo	860315	519.627	340.688	474.355	385.960
9	ENEMDU	2007	Empleo	295556	168.087	127.470	204.763	90.793
...
895	ENEMDU*	2021	Desempleo	384204	319.206	64.997	176.251	207.953
896	ENEMDU*	2021	Desempleo	336101	291.606	44.495	150.276	185.824
897	ENEMDU*	2021	Desempleo	48103	27.600	20.503	25.975	22.128
898	ENEMDU*	2021	Desempleo	298846	257.856	40.991	140.223	158.623
899	ENEMDU*	2021	Desempleo	85358	61.351	24.007	36.028	49.330

653 rows × 8 columns

```
empleo = df3.loc[df3['Indicadores'] == 'Empleo']
empleo
```

	Encuesta	Periodo	Indicadores	Nacional/Total	Área urbana	Área rural	Sexo-Hombre	Sexo-Mujer
5	ENEMDU	2007	Empleo	6019332	3.971.040	2.048.292	3.632.314	2.387.018
6	ENEMDU	2007	Empleo	2737158	2.236.440	500.719	1.907.451	829.707
7	ENEMDU	2007	Empleo	1155872	687.714	468.158	679.118	476.753
8	ENEMDU	2007	Empleo	860315	519.627	340.688	474.355	385.960
9	ENEMDU	2007	Empleo	295556	168.087	127.470	204.763	90.793
...
890	ENEMDU*	2021	Empleo	1789372	1.161.744	627.628	1.062.086	727.286
891	ENEMDU*	2021	Empleo	216021	108.169	107.852	149.922	66.098
892	ENEMDU*	2021	Empleo	947623	289.744	657.879	311.341	636.281
893	ENEMDU*	2021	Empleo	2325638	1.397.591	928.046	1.314.700	1.010.938
894	ENEMDU*	2021	Empleo	51233	46.170	5.063	29.327	21.907

400 rows × 8 columns

```
desempleo = df3.loc[df3['Indicadores'] == 'Desempleo']
desempleo
```

	Encuesta	Periodo	Indicadores	Nacional/Total	Área urbana	Área rural	Sexo-Hombre	Sexo-Mujer
13	ENEMDU	2007	Desempleo	316697	256.662	60.036	144.918	171.780
14	ENEMDU	2007	Desempleo	193225	168.728	24.497	97.565	95.660
15	ENEMDU	2007	Desempleo	123472	87.934	35.539	47.353	76.119
16	ENEMDU	2007	Desempleo	190044	160.289	29.755	99.490	90.554
17	ENEMDU	2007	Desempleo	126653	96.372	30.281	45.427	81.226
...
895	ENEMDU*	2021	Desempleo	384204	319.206	64.997	176.251	207.953
896	ENEMDU*	2021	Desempleo	336101	291.606	44.495	150.276	185.824
897	ENEMDU*	2021	Desempleo	48103	27.600	20.503	25.975	22.128
898	ENEMDU*	2021	Desempleo	298846	257.856	40.991	140.223	158.623

```
df3['Periodo']
```

```

5      2007
6      2007
7      2007
8      2007
9      2007
...
895    2021
896    2021
897    2021
898    2021
899    2021
Name: Periodo, Length: 653, dtype: object
```

```
print(df3.loc[(df3['Indicadores'] == 'Empleo') | (df3['Indicadores'] == 'Desempleo')].groupby(['Periodo', 'Indicadores'])['Nacional/Total'].agg([np.s
```

		sum
2007	Desempleo	950091
	Empleo	13194534
2008	Desempleo	2172508
	Empleo	26589916
2009	Desempleo	1271486
	Empleo	13321895
2010	Desempleo	2194036
	Empleo	26585264
2011	Desempleo	1817976
	Empleo	26546819
2012	Desempleo	1676235
	Empleo	27295094
2013	Desempleo	1688088
	Empleo	28407467
2014	Desempleo	3654660
	Empleo	57837089
2015	Desempleo	3895417
	Empleo	61345230
2016	Desempleo	5086464
	Empleo	65609785
2017	Desempleo	4301597
	Empleo	68087727
2018	Desempleo	3963190
	Empleo	68203728
2019	Desempleo	4381268
	Empleo	69107123
2020	Desempleo	5477492
	Empleo	69299209
2021	Desempleo	12942078
	Empleo	177691736

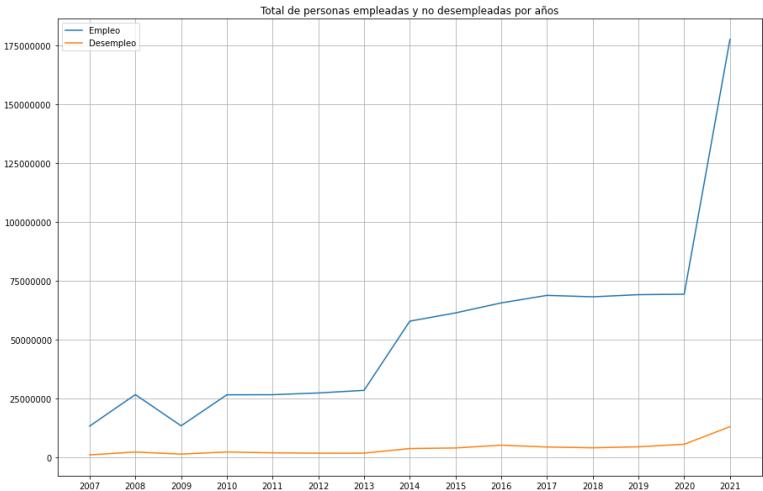
```

tdesempleo=df3.loc[df3['Indicadores'] == 'Desempleo'].groupby(['Periodo'])['Nacional/Total'].sum().reset_index()
templeo=df3.loc[df3['Indicadores'] == 'Empleo'].groupby(['Periodo'])['Nacional/Total'].sum().reset_index()
```

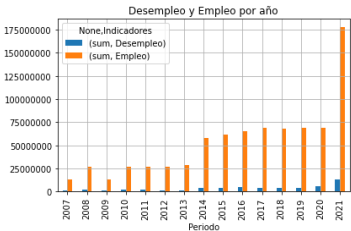
```

plt.figure(figsize=(15,10))
plt.grid(True)
plt.plot(templeo['Periodo'], templeo['Nacional/Total'], label='Empleo')
plt.plot(tdesempleo['Periodo'], tdesempleo['Nacional/Total'], label='Desempleo')
plt.title('Total de personas empleadas y no desempleadas por años')
try:
    plt.ticklabel_format(axis='y',style='plain')
```

```
except AttributeError:
    print('')
plt.legend(loc='upper left')
plt.show()
```



```
templeodesempleo = df3.loc[(df3['Indicadores'] == 'Empleo') | (df3['Indicadores'] == 'Desempleo')]
stack_selection = templeodesempleo.groupby(['Periodo', 'Indicadores'])['Nacional/Total'].agg([np.sum]).unstack()
stack_selection.plot(kind="bar")
plt.grid()
plt.ticklabel_format(axis='y', style='plain')
plt.title('Desempleo y Empleo por año')
plt.tight_layout()
```



• Generar grafico de pie por personas basadas en la sectorización de empleo.

```
df4.head()
```

	Encuesta	Periodo	Indicadores	Nacional/Total	Área urbana	Área rural	Sexo-Hombre	Sexo-Mujer
5	ENEMDU	2007	Empleo	6019332.0	3.971.040	2.048.292	3.632.314	2.387.018
6	ENEMDU	2007	Empleo Adecuado/Pleno	2737158.0	2.236.440	500.719	1.907.451	829.707
7	ENEMDU	2007	Subempleo	1155872.0	687.714	468.158	679.118	476.753
8	ENEMDU	2007	Subempleo por insuficiencia de tiempo de trabajo	860315.0	519.627	340.688	474.355	385.960

df4['Área urbana ']=df4['Área urbana '].replace('\.','',regex=True)

df4['Área rural']=df4['Área rural'].replace('\.','',regex=True)

df4['Área urbana ']=df4['Área urbana '].astype(int, errors = 'raise')

df4['Área rural']=df4['Área rural'].astype(int, errors = 'raise')

dfrural= df4['Área rural'].sum()
dfrural

299813628

dfurbana= df4['Área urbana '].sum()
dfurbana

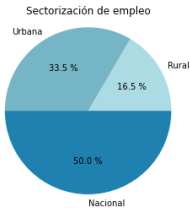
605967957

dftotal= df4['Nacional/Total'].sum()
dftotal

905781585.0

import matplotlib.pyplot as plt

secto=[dfrural,dfurbana,dftotal]
nombres = ["Rural","Urbana","Nacional"]
colors = ['#abdbe3', '#76b5c5', '#1e81b0']
plt.pie(secto, labels=nombres, autopct="%0.1f %%",colors=colors)
plt.axis("equal")
plt.title('Sectorización de empleo')
plt.show()



• Generar histogramas subempleo, empleo pleno y empleo no pleno por año.

seleccionado = df5.loc[(df5['Indicadores'] == 'Subempleo') | (df5['Indicadores'] == 'Subempleo por insuficiencia de tiempo de trabajo') | (df5['Indi
seleccionado

	Encuesta	Periodo	Indicadores	Nacional/Total	Área urbana	Área rural	Sexo-Hombre	Sexo-Mujer
6	ENEMDU	2007	Empleo Adecuado/Pleno	2737158.0	2.236.440	500.719	1.907.451	829.707
7	ENEMDU	2007	Subempleo	1155872.0	687.714	468.158	679.118	476.753
8	ENEMDU	2007	Subempleo por insuficiencia de tiempo de trabajo	860315.0	519.627	340.688	474.355	385.960
9	ENEMDU	2007	Subempleo por insuficiencia de ingresos	295556.0	168.087	127.470	204.763	90.793
11	ENEMDU	2007	Otro Empleo no pleno	1504000.0	785.461	718.539	830.341	673.659
...
888	ENEMDU*	2021	Empleo Adecuado/Pleno	2684258.0	2.237.457	446.800	1.851.954	832.303
889	ENEMDU*	2021	Subempleo	2005393.0	1.269.913	735.480	1.212.008	793.384
890	ENEMDU*	2021	Subempleo por insuficiencia de tiempo de trabajo	1789372.0	1.161.744	627.628	1.062.086	727.286

```
seleccionado['Indicadores'] = seleccionado['Indicadores'].str.replace('Subempleo por insuficiencia de tiempo de trabajo','Subempleo')
seleccionado['Indicadores'] = seleccionado['Indicadores'].str.replace('Subempleo por insuficiencia de ingresos','Subempleo')
```

```
seleccionado
```

```
/usr/local/lib/python3.7/dist-packages/ipykernel_launcher.py:1: SettingWithCopyWarning:
```

```
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

```
See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user\_guide/indexing.html#returning-a-view-versus-a-copy.
```

```
/usr/local/lib/python3.7/dist-packages/ipykernel_launcher.py:2: SettingWithCopyWarning:
```

```
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

```
See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user\_guide/indexing.html#returning-a-view-versus-a-copy.
```

	Encuesta	Periodo	Indicadores	Nacional/Total	Área urbana	Área rural	Sexo-Hombre	Sexo-Mujer
6	ENEMDU	2007	Empleo Adecuado/Pleno	2737158.0	2.236.440	500.719	1.907.451	829.707
7	ENEMDU	2007	Subempleo	1155872.0	687.714	468.158	679.118	476.753
8	ENEMDU	2007	Subempleo	860315.0	519.627	340.688	474.355	385.960
9	ENEMDU	2007	Subempleo	295556.0	168.087	127.470	204.763	90.793
11	ENEMDU	2007	Otro Empleo no pleno	1504000.0	785.461	718.539	830.341	673.659
...
888	ENEMDU*	2021	Empleo Adecuado/Pleno	2684258.0	2.237.457	446.800	1.851.954	832.303
889	ENEMDU*	2021	Subempleo	2005393.0	1.269.913	735.480	1.212.008	793.384
890	ENEMDU*	2021	Subempleo	1789372.0	1.161.744	627.628	1.062.086	727.286
891	ENEMDU*	2021	Subempleo	216021.0	108.169	107.852	149.922	66.098
893	ENEMDU*	2021	Otro Empleo no pleno	2325638.0	1.397.591	928.046	1.314.700	1.010.938

```
250 rows × 8 columns
```

```
dfseleccionado= seleccionado['Indicadores'].value_counts().reset_index()
dfseleccionado
```

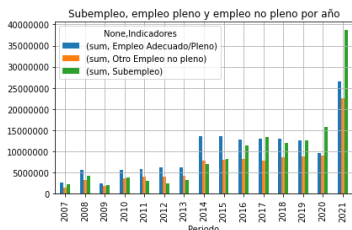
	index	Indicadores
0	Subempleo	150
1	Otro Empleo no pleno	50
2	Empleo Adecuado/Pleno	50

```
print(seleccionado.groupby(['Periodo','Indicadores'])['Nacional/Total'].agg([np.sum]))
```

	Periodo	Indicadores	sum
2007	Empleo Adecuado/Pleno	2737158.0	

	Otro Empleo no pleno	1504000.0
	Subempleo	2311743.0
2008	Empleo Adecuado/Pleno	5609254.0
	Otro Empleo no pleno	3338042.0
	Subempleo	4177354.0
2009	Empleo Adecuado/Pleno	2565691.0
	Otro Empleo no pleno	1778578.0
	Subempleo	2143230.0
2010	Empleo Adecuado/Pleno	5591075.0
	Otro Empleo no pleno	3577581.0
	Subempleo	3861043.0
2011	Empleo Adecuado/Pleno	5786091.0
	Otro Empleo no pleno	4089874.0
	Subempleo	2975966.0
2012	Empleo Adecuado/Pleno	6222998.0
	Otro Empleo no pleno	4081973.0
	Subempleo	2537744.0
2013	Empleo Adecuado/Pleno	6326529.0
	Otro Empleo no pleno	4164807.0
	Subempleo	3254790.0
2014	Empleo Adecuado/Pleno	13567061.0
	Otro Empleo no pleno	7935607.0
	Subempleo	7123557.0
2015	Empleo Adecuado/Pleno	13617561.0
	Otro Empleo no pleno	8130755.0
	Subempleo	8272233.0
2016	Empleo Adecuado/Pleno	12755136.0
	Otro Empleo no pleno	8280823.0
	Subempleo	11502229.0
2017	Empleo Adecuado/Pleno	13101364.0
	Otro Empleo no pleno	7945160.0
	Subempleo	13354748.0
2018	Empleo Adecuado/Pleno	12998778.0
	Otro Empleo no pleno	8651232.0
	Subempleo	11944242.0
2019	Empleo Adecuado/Pleno	12592867.0
	Otro Empleo no pleno	8830974.0
	Subempleo	12563450.0
2020	Empleo Adecuado/Pleno	9583775.0
	Otro Empleo no pleno	9012175.0
	Subempleo	15824933.0
2021	Empleo Adecuado/Pleno	26463801.0
	Otro Empleo no pleno	22659899.0
	Subempleo	38671862.0

```
stack_seleccion = seleccionado.groupby(['Periodo', 'Indicadores'])['Nacional/Total'].agg([np.sum]).unstack()
stack_seleccion.plot(kind="bar")
plt.grid()
plt.ticklabel_format(axis='y', style='plain')
plt.title('Subempleo, empleo pleno y empleo no pleno por año')
plt.tight_layout()
#hist
```



• Generar un reporte parametrizado que permita ingresar los datos de las fechas inicio y fin para obtener la información de las graficas vistas en el primer punto.

```
import pandas as pd
dfparametri = pd.read_csv('1.Poblaciones.csv', sep = ';', encoding = 'latin1', header=1, error_bad_lines=False) #error_bad_lines is avoid single line
```

```

#df['Área'] = df[['Área', 'Unnamed: 5']].apply(' '.join, axis=1)
#df.drop(df.columns[df.columns.str.contains('unnamed',case = False)],axis = 1, inplace = True)
dfparametri = dfparametri.rename(columns={'Área':'Área urbana ', 'Unnamed: 5':'Área rural', 'Nacional':'Nacional/Total', 'Sexo':'Sexo-Hombre', 'Unnamed: 6':'Sexo-Mujer'})
dfparametri.drop(dfparametri.columns[dfparametri.columns.str.contains('unnamed',case = False)],axis = 1, inplace = True)
parametriselecc=dfparametri.drop([0])

parametriselecc['Periodo'] = parametriselecc['Periodo'].str.replace('-', '/20')
for i in parametriselecc['Periodo']:
    i=i.split(sep='/')
    #print(i[0])
    if (i[0]=='ene'):
        parametriselecc['Periodo'] = parametriselecc['Periodo'].str.replace(i[0], '1/01')
    if (i[0]=='feb'):
        parametriselecc['Periodo'] = parametriselecc['Periodo'].str.replace(i[0], '1/02')
    if (i[0]=='mar'):
        parametriselecc['Periodo'] = parametriselecc['Periodo'].str.replace(i[0], '1/03')
    if (i[0]=='abr'):
        parametriselecc['Periodo'] = parametriselecc['Periodo'].str.replace(i[0], '1/04')
    if (i[0]=='may'):
        parametriselecc['Periodo'] = parametriselecc['Periodo'].str.replace(i[0], '1/05')
    if (i[0]=='jun'):
        parametriselecc['Periodo'] = parametriselecc['Periodo'].str.replace(i[0], '1/06')
    if (i[0]=='jul'):
        parametriselecc['Periodo'] = parametriselecc['Periodo'].str.replace(i[0], '1/07')
    if (i[0]=='ago'):
        parametriselecc['Periodo'] = parametriselecc['Periodo'].str.replace(i[0], '1/08')
    if (i[0]=='sep'):
        parametriselecc['Periodo'] = parametriselecc['Periodo'].str.replace(i[0], '1/09')
    if (i[0]=='oct'):
        parametriselecc['Periodo'] = parametriselecc['Periodo'].str.replace(i[0], '1/10')
    if (i[0]=='nov'):
        parametriselecc['Periodo'] = parametriselecc['Periodo'].str.replace(i[0], '1/11')
    if (i[0]=='dic'):
        parametriselecc['Periodo'] = parametriselecc['Periodo'].str.replace(i[0], '1/12')

#dataParametrizado['Periodo'] = dataParametrizado['Periodo'].str.replace(r"[a-zA-Z]", '')
parametriselecc.head(30)

```

	Encuesta	Periodo	Indicadores	Nacional/Total	Área urbana	Área rural	Sexo-Hombre	Sexo-Mujer
1	ENEMDU	1/12/2007	Población Total	13.682.302	9.066.209	4.616.093	6.768.646	6.913.656
2	ENEMDU	1/12/2007	Población menor de 15 años	4.372.812	2.723.124	1.649.688	2.226.618	2.146.194
3	ENEMDU	1/12/2007	Población en Edad de Trabajar (PET)	9.309.490	6.343.085	2.966.404	4.542.028	4.767.462
4	ENEMDU	1/12/2007	Población Económicamente Activa	6.336.029	4.227.702	2.108.328	3.777.232	2.558.798
5	ENEMDU	1/12/2007	Empleo	6.019.332	3.971.040	2.048.292	3.632.314	2.387.018
6	ENEMDU	1/12/2007	Empleo Adecuado/Pleno	2.737.158	2.236.440	500.719	1.907.451	829.707
7	ENEMDU	1/12/2007	Subempleo	1.155.872	687.714	468.158	679.118	476.753
8	ENEMDU	1/12/2007	Subempleo por insuficiencia de tiempo de trabajo	860.315	519.627	340.688	474.355	385.960
9	ENEMDU	1/12/2007	Subempleo por insuficiencia de ingresos	295.556	168.087	127.470	204.763	90.793

```

from datetime import datetime
FMT = '%d/%m/%Y'
parametriselecc['Periodo'] = parametriselecc['Periodo'].map(lambda x : (datetime.strptime(x, FMT)))

```

parametriselecc

	Encuesta	Periodo	Indicadores	Nacional/Total	Área urbana	Área rural	Sexo-Hombre	Sexo-Mujer
1	ENEMDU	2007-12-01	Población Total	13.682.302	9.066.209	4.616.093	6.768.646	6.913.656
2	ENEMDU	2007-12-01	Población menor de 15 años	4.372.812	2.723.124	1.649.688	2.226.618	2.146.194
3	ENEMDU	2007-12-01	Población en Edad de Trabajar (PET)	9.309.490	6.343.085	2.966.404	4.542.028	4.767.462
4	ENEMDU	2007-12-01	Población Económicamente Activa	6.336.029	4.227.702	2.108.328	3.777.232	2.558.798
5	ENEMDU	2007-12-01	Empleo	6.019.332	3.971.040	2.048.292	3.632.314	2.387.018
...
896	ENEMDU*	2021-10-01	Desempleo Abierto	336.101	291.606	44.495	150.276	185.824
897	ENEMDU*	2021-10-01	Desempleo Oculto	48.103	27.600	20.503	25.975	22.128
898	ENEMDU*	2021-10-01	Desempleo Cesante	298.846	257.856	40.991	140.223	158.623
899	ENEMDU*	2021-10-01	Desempleo Nuevo	85.358	61.351	24.007	36.028	49.330
900	ENEMDU*	2021-10-01	Población Económicamente Inactiva	4.330.241	3.307.420	1.022.821	1.324.745	3.005.495
900 rows × 8 columns								
98	ENEMDU I	1/10/2008	Empleo no remunerado	614.179	267.690	346.649	204.693	409.648

```

parametriselecc['Nacional/Total']=parametriselecc['Nacional/Total'].replace('\.', '', regex=True)

```

```

parametriselecc['Nacional/Total'] = parametriselecc['Nacional/Total'].astype(float, errors = 'raise')

```

```

parametriseleccpie = parametriselecc.copy()

```

```

list = ['Población Total','Población Económicamente Inactiva','Población en Edad de Trabajar (PET)','Poblacion Total','Población menor de 15 años',
conte= parametriselecc.shape[0]

```

```

for i in range(len(list)):
    for j in range(conte):
        parametriselecc = parametriselecc.drop(parametriselecc[parametriselecc['Indicadores']== list[i]].index)

```

```

empleo=['Empleo','Empleo Adecuado/Pleno','Otro Empleo no pleno','Empleo no clasificado','Empleo no remunerado','Subempleo','Subempleo por insuficien
desempleo=['Desempleo','Desempleo Cesante','Desempleo Abierto','Desempleo Nuevo','Desempleo Oculto']
parametriselecc['Indicadores'] = parametriselecc['Indicadores'].str.replace('Subempleo por insuficiencia de tiempo de trabajo','Empleo')
parametriselecc['Indicadores'] = parametriselecc['Indicadores'].str.replace('Subempleo por insuficiencia de ingresos','Empleo')
parametriselecc.drop(parametriselecc[parametriselecc['Indicadores']=='Población Total'].index)

```

```

for j in empleo:
    for i in parametriselecc['Indicadores']:
        if j==i:
            parametriselecc['Indicadores']=parametriselecc['Indicadores'].replace(i,'Empleo',regex=True)

```

```
for j in desempleo:
    for i in parametriselecc['Indicadores']:
        if j==1:
            parametriselecc['Indicadores']=parametriselecc['Indicadores'].replace(i,'Desempleo',regex=True)
```

parametriselecc

	Encuesta	Periodo	Indicadores	Nacional/Total	Área urbana	Área rural	Sexo-Hombre	Sexo-Mujer
5	ENEMDU	2007-12-01	Empleo	6019332.0	3.971.040	2.048.292	3.632.314	2.387.018
6	ENEMDU	2007-12-01	Empleo	2737158.0	2.236.440	500.719	1.907.451	829.707
7	ENEMDU	2007-12-01	Empleo	1155872.0	687.714	468.158	679.118	476.753
8	ENEMDU	2007-12-01	Empleo	860315.0	519.627	340.688	474.355	385.960
9	ENEMDU	2007-12-01	Empleo	295556.0	168.087	127.470	204.763	90.793
...
895	ENEMDU*	2021-10-01	Desempleo	384204.0	319.206	64.997	176.251	207.953
896	ENEMDU*	2021-10-01	Desempleo	336101.0	291.606	44.495	150.276	185.824
897	ENEMDU*	2021-10-01	Desempleo	48103.0	27.600	20.503	25.975	22.128
898	ENEMDU*	2021-10-01	Desempleo	298846.0	257.856	40.991	140.223	158.623
899	ENEMDU*	2021-10-01	Desempleo	85358.0	61.351	24.007	36.028	49.330

653 rows × 8 columns

```
parametriselecc['Nacional/Total'] = parametriselecc['Nacional/Total'].astype('int64', errors = 'raise')
```

parametriselecc

	Encuesta	Periodo	Indicadores	Nacional/Total	Área urbana	Área rural	Sexo-Hombre	Sexo-Mujer
5	ENEMDU	2007-12-01	Empleo	6019332	3.971.040	2.048.292	3.632.314	2.387.018
6	ENEMDU	2007-12-01	Empleo	2737158	2.236.440	500.719	1.907.451	829.707
7	ENEMDU	2007-12-01	Empleo	1155872	687.714	468.158	679.118	476.753
8	ENEMDU	2007-12-01	Empleo	860315	519.627	340.688	474.355	385.960
9	ENEMDU	2007-12-01	Empleo	295556	168.087	127.470	204.763	90.793
...
895	ENEMDU*	2021-10-01	Desempleo	384204	319.206	64.997	176.251	207.953
896	ENEMDU*	2021-10-01	Desempleo	336101	291.606	44.495	150.276	185.824
897	ENEMDU*	2021-10-01	Desempleo	48103	27.600	20.503	25.975	22.128
898	ENEMDU*	2021-10-01	Desempleo	298846	257.856	40.991	140.223	158.623
899	ENEMDU*	2021-10-01	Desempleo	85358	61.351	24.007	36.028	49.330

653 rows × 8 columns

```
print(parametriselecc.loc[(parametriselecc['Indicadores'] == 'Empleo') | (parametriselecc['Indicadores'] == 'Desempleo')].groupby(['Periodo','Indica
```

			sum
Periodo	Indicadores		
2007-12-01	Desempleo		950091
	Empleo		13194534
2008-06-01	Desempleo		1032430
	Empleo		13621148
2008-12-01	Desempleo		1140078
...			...
2021-08-01	Empleo		18137695
2021-09-01	Desempleo		1237323
	Empleo		18054703
2021-10-01	Desempleo		1152612
	Empleo		18033682

[100 rows × 1 columns]

```
parametridos=parametriselecc.loc[(parametriselecc['Indicadores'] == 'Empleo') | (parametriselecc['Indicadores'] == 'Desempleo')].groupby(['Periodo',
```

```
parametridos
```

		sum
Periodo	Indicadores	
2007-12-01	Desempleo	950091
	Empleo	13194534
2008-06-01	Desempleo	1032430
	Empleo	13621148
2008-12-01	Desempleo	1140078
...
2021-08-01	Empleo	18137695
2021-09-01	Desempleo	1237323
	Empleo	18054703
2021-10-01	Desempleo	1152612
	Empleo	18033682

100 rows × 3 columns

```
parametritres=parametridos['sum'].reset_index()
```

```
from datetime import datetime
f1, f2 = input('Ingreso fecha inicial: '), input('Ingreso fecha final: ')

baseinicial = [parametritres.index.get_indexer_for((parametritres[parametritres['Periodo'] == f1].index))]
valorinicial = baseinicial[0][0]
print(valorinicial)

basefinal = [parametritres.index.get_indexer_for((parametritres[parametritres['Periodo'] == f2].index))]
valorfinal = basefinal[0][1]
print(valorfinal)
```

```
Ingreso fecha inicial: 2007-12-01
Ingreso fecha final: 2017-12-01
0
55
```

```
parametricuatro=parametritres.iloc[valorinicial:valorfinal]
parametricuatro
```

	Periodo	Indicadores	sum
0	2007-12-01	Desempleo	950091
1	2007-12-01	Empleo	13194534
2	2008-06-01	Desempleo	1032430
3	2008-06-01	Empleo	13621148
4	2008-12-01	Desempleo	1140078
5	2008-12-01	Empleo	12968768
6	2009-12-01	Desempleo	1271406
7	2009-12-01	Empleo	13321885
8	2010-06-01	Desempleo	1224955
9	2010-06-01	Empleo	13389548
10	2010-12-01	Desempleo	969081
11	2010-12-01	Empleo	13115716
12	2011-06-01	Desempleo	987615
13	2011-06-01	Empleo	13230693
14	2011-12-01	Desempleo	830361
15	2011-12-01	Empleo	13316126
16	2012-06-01	Desempleo	847713
17	2012-06-01	Empleo	13841524
18	2012-12-01	Desempleo	828522
19	2012-12-01	Empleo	13453570
20	2013-06-01	Desempleo	821853
21	2013-06-01	Empleo	14269716
22	2013-12-01	Desempleo	866235
23	2013-12-01	Empleo	14137751
24	2014-03-01	Desempleo	1026288
25	2014-03-01	Empleo	14302989
26	2014-06-01	Desempleo	972867
27	2014-06-01	Empleo	14138855
28	2014-09-01	Desempleo	835264
29	2014-09-01	Empleo	14627257
30	2014-12-01	Desempleo	820241
31	2014-12-01	Empleo	14767988
32	2015-03-01	Desempleo	848902
33	2015-03-01	Empleo	15167931
34	2015-06-01	Desempleo	996351
35	2015-06-01	Empleo	15174952
36	2015-09-01	Desempleo	976488
37	2015-09-01	Empleo	15670429
38	2015-12-01	Desempleo	1073676
39	2015-12-01	Empleo	15331918
40	2016-03-01	Desempleo	1346972
41	2016-03-01	Empleo	16173573
42	2016-06-01	Desempleo	1250648

parametricinco= parametricuatro.copy()

44 2016-06-01 Desempleo 1257510
parametricinco

	Periodo	Indicadores	sum
0	2007-12-01	Desempleo	950091
1	2007-12-01	Empleo	13194534
2	2008-06-01	Desempleo	1032430
3	2008-06-01	Empleo	13621148
4	2008-12-01	Desempleo	1140078
5	2008-12-01	Empleo	12968768
6	2009-12-01	Desempleo	1271406
7	2009-12-01	Empleo	13321885
8	2010-06-01	Desempleo	1224955
9	2010-06-01	Empleo	13389548
10	2010-12-01	Desempleo	969081
11	2010-12-01	Empleo	13115716
12	2011-06-01	Desempleo	987615
13	2011-06-01	Empleo	13230693
14	2011-12-01	Desempleo	830361
15	2011-12-01	Empleo	13316126
16	2012-06-01	Desempleo	847713
17	2012-06-01	Empleo	13841524
18	2012-12-01	Desempleo	828522
19	2012-12-01	Empleo	13453570
20	2013-06-01	Desempleo	821853
21	2013-06-01	Empleo	14269716
22	2013-12-01	Desempleo	866235
23	2013-12-01	Empleo	14137751
24	2014-03-01	Desempleo	1026288
25	2014-03-01	Empleo	14302989
26	2014-06-01	Desempleo	972867
27	2014-06-01	Empleo	14138855
28	2014-09-01	Desempleo	835264
29	2014-09-01	Empleo	14627257
30	2014-12-01	Desempleo	820241
31	2014-12-01	Empleo	14767988
32	2015-03-01	Desempleo	848902
33	2015-03-01	Empleo	15167931
34	2015-06-01	Desempleo	996351
35	2015-06-01	Empleo	15174952
36	2015-09-01	Desempleo	976488
37	2015-09-01	Empleo	15670429
38	2015-12-01	Desempleo	1073676
39	2015-12-01	Empleo	15331918
40	2016-03-01	Desempleo	1346972
41	2016-03-01	Empleo	16173573
42	2016-06-01	Desempleo	1250648
43	2016-06-01	Empleo	16107915
44	2016-09-01	Desempleo	1257519
45	2016-09-01	Empleo	16836314


```

46 2016-12-01    Desempleo    1231325
47 2016-12-01      Empleo    16491983
48 2017-03-01    Desempleo    1066241
49 2017-03-01      Empleo    17183966
50 2017-06-01    Desempleo    1098011
51 2017-06-01      Empleo    17231697
52 2017-09-01    Desempleo    1015733
53 2017-09-01      Empleo    17364800
54 2017-12-01    Desempleo    1121612

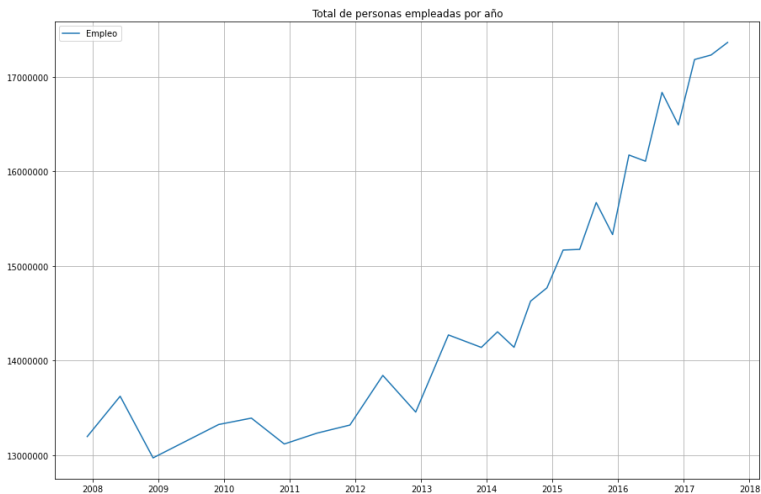
```

```
templeo=parametricinco.loc[parametricinco['Indicadores'] == 'Empleo'].groupby(['Periodo'])['sum'].sum().reset_index()
```

```

plt.figure(figsize=(15,10))
plt.grid(True)
plt.plot(templeo['Periodo'], templeo['sum'], label='Empleo')
plt.title('Total de personas empleadas por año')
try:
    plt.ticklabel_format(axis='y',style='plain')
except AttributeError:
    print('')
plt.legend(loc='upper left')
plt.show()

```



```
tdeempleo=parametricinco.loc[parametricinco['Indicadores'] == 'Desempleo'].groupby(['Periodo'])['sum'].sum().reset_index()
```

```

plt.figure(figsize=(15,10))
plt.grid(True)
plt.plot(tdeempleo['Periodo'], tdeempleo['sum'], label='Desempleo',color= 'orange')
plt.title('Total de personas desempleadas por año')
try:
    plt.ticklabel_format(axis='y',style='plain')
except AttributeError:

```

```
print('')
plt.legend(loc='upper left')
plt.show()
```



▼ Segundo grafico

parametriseleccpie

Encuesta	Periodo	Indicadores	Nacional/Total	Área urbana	Área rural	Sexo-Hombre	Sexo-Mujer
1	ENEMDU 2007-12-01	Población Total	13682302.0	9.066.209	4.616.093	6.768.646	6.913.656
2	ENEMDU 2007-12-01	Población menor de 15 años	4372812.0	2.723.124	1.649.688	2.226.618	2.146.194
3	ENEMDU 2007-12-01	Población en Edad de Trabajar (PET)	9309490.0	6.343.085	2.966.404	4.542.028	4.767.462
4	ENEMDU 2007-12-01	Población Económicamente Activa	6336029.0	4.227.702	2.108.328	3.777.232	2.558.798
5	ENEMDU 2007-12-01	Empleo	6019332.0	3.971.040	2.048.292	3.632.314	2.387.018
...
896	ENEMDU* 2021-10-01	Desempleo Abierto	336101.0	291.606	44.495	150.276	185.824
897	ENEMDU* 2021-10-01	Desempleo Oculto	48103.0	27.600	20.503	25.975	22.128
898	ENEMDU* 2021-10-01	Desempleo Cesante	298846.0	257.856	40.991	140.223	158.623
899	ENEMDU* 2021-10-01	Desempleo Nuevo	85358.0	61.351	24.007	36.028	49.330
900	ENEMDU* 2021-10-01	Población Económicamente Inactiva	4330241.0	3.307.420	1.022.821	1.324.745	3.005.495

900 rows × 8 columns

```
#separar por la fecha de inicio y fin
parametricuatro=parametriseleccpie.iloc[valorinicial:valorfinal-1]
parametricuatro
```

Encuesta	Periodo	Indicadores	Nacional/Total	Área urbana	Área rural	Sexo-Hombre	Sexo-Mujer
1	ENEMDU 2007-12-01	Población Total	13682302.0	9.066.209	4.616.093	6.768.646	6.913.656
2	ENEMDU 2007-12-01	Población menor de 15 años	4372812.0	2.723.124	1.649.688	2.226.618	2.146.194
3	ENEMDU 2007-12-01	Población en Edad de Trabajar (PET)	9309490.0	6.343.085	2.966.404	4.542.028	4.767.462
4	ENEMDU 2007-12-01	Población Económicamente Activa	6336029.0	4.227.702	2.108.328	3.777.232	2.558.798
5	ENEMDU 2007-12-01	Empleo	6019332.0	3.971.040	2.048.292	3.632.314	2.387.018
6	ENEMDU 2007-12-01	Empleo Adecuado/Pleno	2737158.0	2.236.440	500.719	1.907.451	829.707
7	ENEMDU 2007-12-01	Subempleo	1155872.0	687.714	468.158	679.118	476.753
8	ENEMDU 2007-12-01	Subempleo por insuficiencia de tiempo de trabajo	860315.0	519.627	340.688	474.355	385.960
9	ENEMDU 2007-12-01	Subempleo por insuficiencia de ingresos	295556.0	168.087	127.470	204.763	90.793
10	ENEMDU 2007-12-01	Empleo no remunerado	557146.0	215.994	341.152	176.122	381.025
11	ENEMDU 2007-12-01	Otro Empleo no pleno	1504000.0	785.461	718.539	830.341	673.659
12	ENEMDU 2007-12-01	Empleo no clasificado	65155.0	45.430	19.725	39.282	25.873
13	ENEMDU 2007-12-01	Desempleo	316697.0	256.662	60.036	144.918	171.780
14	ENEMDU 2007-12-01	Desempleo Abierto	193225.0	168.728	24.497	97.565	95.660
15	ENEMDU 2007-12-01	Desempleo Oculto	123472.0	87.934	35.539	47.353	76.119
16	ENEMDU 2007-12-01	Desempleo Cesante	190044.0	160.289	29.755	99.490	90.554
17	ENEMDU 2007-12-01	Desempleo Nuevo	126653.0	96.372	30.281	45.427	81.226
18	ENEMDU 2007-12-01	Población Económicamente Inactiva	2973460.0	2.115.384	858.077	764.796	2.208.664
19	ENEMDU 2008-06-01	Población Total	13781797.0	9.129.646	4.652.151	6.868.579	6.913.218
20	ENEMDU 2008-06-01	Población menor de 15 años	4268148.0	2.655.282	1.612.866	2.189.753	2.078.395
21	ENEMDU 2008-06-01	Población en Edad de Trabajar (PET)	9513649.0	6.474.364	3.039.285	4.678.826	4.834.823
22	ENEMDU 2008-06-01	Población Económicamente Activa	6589368.0	4.433.491	2.155.877	3.893.496	2.695.873
23	ENEMDU 2008-06-01	Empleo	6245225.0	4.151.002	2.094.223	3.742.213	2.503.012
24	ENEMDU 2008-06-01	Empleo Adecuado/Pleno	2750595.0	2.248.018	502.577	1.937.602	812.993
25	ENEMDU 2008-06-01	Subempleo	1130699.0	696.949	433.750	665.939	464.760
26	ENEMDU 2008-06-01	Subempleo por insuficiencia de tiempo de trabajo	846132.0	527.132	319.001	457.856	388.276
27	ENEMDU 2008-06-01	Subempleo por insuficiencia de ingresos	284567.0	169.817	114.750	208.083	76.484
28	ENEMDU 2008-06-01	Empleo no remunerado	614172.0	257.629	356.542	204.623	409.548
29	ENEMDU 2008-06-01	Otro Empleo no pleno	1688693.0	909.254	779.439	898.349	790.344
30	ENEMDU 2008-06-01	Empleo no clasificado	61065.0	39.152	21.914	35.699	25.366
31	ENEMDU 2008-06-01	Desempleo	344143.0	282.489	61.654	151.283	192.860
32	ENEMDU 2008-06-01	Desempleo Abierto	243614.0	217.324	26.290	123.250	120.364
33	ENEMDU 2008-06-01	Desempleo Oculto	100530.0	65.165	35.365	28.033	72.497
34	ENEMDU 2008-06-01	Desempleo Cesante	238031.0	207.834	30.198	113.592	124.440
35	ENEMDU 2008-06-01	Desempleo Nuevo	106112.0	74.655	31.457	37.691	68.421
36	ENEMDU 2008-06-01	Población Económicamente Inactiva	2924281.0	2.040.873	883.408	785.331	2.138.950
37	ENEMDU 2008-12-01	Población Total	13878704.0	9.202.348	4.676.355	6.831.453	7.047.250
38	ENEMDU 2008-12-01	Población menor de 15 años	4229707.0	2.633.450	1.596.257	2.175.561	2.054.146
39	ENEMDU 2008-12-01	Población en Edad de Trabajar (PET)	9648996.0	6.568.898	3.080.098	4.655.892	4.993.104
40	ENEMDU 2008-12-01	Población Económicamente Activa	6385421.0	4.329.308	2.056.113	3.800.531	2.584.890
41	ENEMDU 2008-12-01	Empleo	6005395.0	4.012.298	1.993.097	3.635.236	2.370.159
42	ENEMDU 2008-12-01	Empleo Adecuado/Pleno	2858659.0	2.339.288	519.370	1.991.157	867.502
43	ENEMDU 2008-12-01	Subempleo	957978.0	588.824	369.155	549.097	408.881
44	ENEMDU 2008-12-01	Subempleo por insuficiencia de tiempo de trabajo	704516.0	445.009	259.507	378.324	326.192
45	ENEMDU 2008-12-01	Subempleo por insuficiencia de ingresos	253462.0	143.815	109.648	170.773	82.689
46	ENEMDU 2008-12-01	Empleo no remunerado	523928.0	211.070	312.858	159.714	364.214

47	ENEMDU	2008-12-01	Otro Empleo no pleno	1649349.0	859.126	790.223	923.415	725.935
48	ENEMDU	2008-12-01	Empleo no clasificado	15481.0	13.990	1.491	11.853	3.628
49	ENEMDU	2008-12-01	Desempleo	380026.0	317.011	63.015	165.296	214.730
50	ENEMDU	2008-12-01	Desempleo Abierto	252924.0	226.983	25.942	119.858	133.067
51	ENEMDU	2008-12-01	Desempleo Oculto	127102.0	90.028	37.074	45.438	81.664
52	ENEMDU	2008-12-01	Desempleo Cesante	225489.0	196.643	28.847	109.477	116.013
53	ENEMDU	2008-12-01	Desempleo Nuevo	154537.0	120.368	34.169	55.819	98.718

```
parametricuatro['Área urbana ']=parametricuatro['Área urbana '].replace('\.','',regex=True)
```

```
parametricuatro['Área rural']=parametricuatro['Área rural'].replace('\.','',regex=True)
```

```
parametricuatro['Área urbana ']=parametricuatro['Área urbana '].astype(int, errors = 'raise')
```

```
parametricuatro['Área rural']=parametricuatro['Área rural'].astype(int, errors = 'raise')
```

```
parametricuatro['Nacional/Total']=parametricuatro['Nacional/Total'].replace('\.','',regex=True)
```

```
parametricuatro['Nacional/Total']=parametricuatro['Nacional/Total'].astype(int, errors = 'raise')
```

```
/usr/local/lib/python3.7/dist-packages/ipykernel_launcher.py:1: SettingWithCopyWarning:
```

```
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

```
See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user\_guide/indexing.html#returning-a-view-versus-a-copy.
```

```
/usr/local/lib/python3.7/dist-packages/ipykernel_launcher.py:3: SettingWithCopyWarning:
```

```
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

```
See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user\_guide/indexing.html#returning-a-view-versus-a-copy.
```

```
/usr/local/lib/python3.7/dist-packages/ipykernel_launcher.py:5: SettingWithCopyWarning:
```

```
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

```
See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user\_guide/indexing.html#returning-a-view-versus-a-copy.
```

```
/usr/local/lib/python3.7/dist-packages/ipykernel_launcher.py:7: SettingWithCopyWarning:
```

```
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

```
See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user\_guide/indexing.html#returning-a-view-versus-a-copy.
```

```
/usr/local/lib/python3.7/dist-packages/ipykernel_launcher.py:9: SettingWithCopyWarning:
```

```
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

```
See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user\_guide/indexing.html#returning-a-view-versus-a-copy.
```

```
/usr/local/lib/python3.7/dist-packages/ipykernel_launcher.py:11: SettingWithCopyWarning:
```

```
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

```
See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user\_guide/indexing.html#returning-a-view-versus-a-copy.
```

```
dfruralpara= parametricuatro['Área rural'].sum()
dfruralpara
```

```
51071393
```

```
dfurbanapara= parametricuatro['Área urbana '].sum()
dfurbanapara

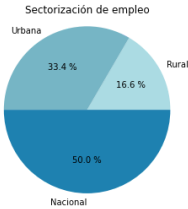
102993405

dftotalpara= parametricuatro['Nacional/Total'].sum()
dftotalpara

154064788

import matplotlib.pyplot as plt

secto=[dfrruralpara,dfurbanapara,dftotalpara]
nombres = ["Rural", "Urbana", "Nacional"]
colors = ['#abdbe3', '#76b5c5', '#1e81b0']
plt.pie(secto, labels=nombres, autopct="%0.1f %%", colors=colors)
plt.axis("equal")
plt.title('Sectorización de empleo')
plt.show()
```



▼ Tercer grafico

```
parametriselecchist=parametriseleccpie.copy()
parametriselecchist
```

	Encuesta	Período	Indicadores	Nacional/Total	Área urbana	Área rural	Sexo-Hombre	Sexo-Mujer
1	ENEMDU	2007-12-01	Población Total	13682302.0	9066209	4616093	6.768.646	6.913.656
2	ENEMDU	2007-12-01	Población menor de 15 años	4372812.0	2723124	1649688	2.226.618	2.146.194
3	ENEMDU	2007-12-01	Población en Edad de Trabajar (PET)	9309490.0	6343085	2966404	4.542.028	4.767.462
4	ENEMDU	2007-12-01	Población Económicamente Activa	6336029.0	4227702	2108328	3.777.232	2.558.798
5	ENEMDU	2007-12-01	Empleo	6019332.0	3971040	2048292	3.632.314	2.387.018
...
896	ENEMDU*	2021-10-01	Desempleo Abierto	336101.0	291.606	44.495	150.276	185.824
897	ENEMDU*	2021-10-01	Desempleo Oculto	48103.0	27.600	20.503	25.975	22.128
898	ENEMDU*	2021-10-01	Desempleo Cesante	298846.0	257.856	40.991	140.223	158.623
899	ENEMDU*	2021-10-01	Desempleo Nuevo	85358.0	61.351	24.007	36.028	49.330
900	ENEMDU*	2021-10-01	Población Económicamente Inactiva	4330241.0	3.307.420	1.022.821	1.324.745	3.005.495

900 rows × 8 columns

```
#separar por fecha de inicio y fecha fin
parametrihist=parametriselecchist.iloc[valorinicial:valorfinal-1]
parametrihist
```

Encuesta	Periodo	Indicadores	Nacional/Total	Área urbana	Área rural	Sexo-Hombre	Sexo-Mujer
1	ENEMDU 2007-12-01	Población Total	13682302.0	9066209	4616093	6.768.646	6.913.656
2	ENEMDU 2007-12-01	Población menor de 15 años	4372812.0	2723124	1649688	2.226.618	2.146.194
3	ENEMDU 2007-12-01	Población en Edad de Trabajar (PET)	9309490.0	6343085	2966404	4.542.028	4.767.462
4	ENEMDU 2007-12-01	Población Económicamente Activa	6336029.0	4227702	2108328	3.777.232	2.558.798
5	ENEMDU 2007-12-01	Empleo	6019332.0	3971040	2048292	3.632.314	2.387.018
6	ENEMDU 2007-12-01	Empleo Adecuado/Pleno	2737158.0	2236440	500719	1.907.451	829.707
7	ENEMDU 2007-12-01	Subempleo	1155872.0	687714	468158	679.118	476.753
8	ENEMDU 2007-12-01	Subempleo por insuficiencia de tiempo de trabajo	860315.0	519627	340688	474.355	385.960
9	ENEMDU 2007-12-01	Subempleo por insuficiencia de ingresos	295556.0	168087	127470	204.763	90.793
10	ENEMDU 2007-12-01	Empleo no remunerado	557146.0	215994	341152	176.122	381.025
11	ENEMDU 2007-12-01	Otro Empleo no pleno	1504000.0	785461	718539	830.341	673.659
12	ENEMDU 2007-12-01	Empleo no clasificado	65155.0	45430	19725	39.282	25.873
13	ENEMDU 2007-12-01	Desempleo	316697.0	256662	60036	144.918	171.780
14	ENEMDU 2007-12-01	Desempleo Abierto	193225.0	168728	24497	97.565	95.660
15	ENEMDU 2007-12-01	Desempleo Oculto	123472.0	87934	35539	47.353	76.119
16	ENEMDU 2007-12-01	Desempleo Cesante	190044.0	160289	29755	99.490	90.554
17	ENEMDU 2007-12-01	Desempleo Nuevo	126653.0	96372	30281	45.427	81.226
18	ENEMDU 2007-12-01	Población Económicamente Inactiva	2973460.0	2115384	858077	764.796	2.208.664
19	ENEMDU 2008-06-01	Población Total	13781797.0	9129646	4652151	6.868.579	6.913.218
20	ENEMDU 2008-06-01	Población menor de 15 años	4268148.0	2655282	1612866	2.189.753	2.078.395
21	ENEMDU 2008-06-01	Población en Edad de Trabajar (PET)	9513649.0	6474364	3039285	4.678.826	4.834.823
22	ENEMDU 2008-06-01	Población Económicamente Activa	6589368.0	4433491	2155877	3.893.496	2.695.873
23	ENEMDU 2008-06-01	Empleo	6245225.0	4151002	2094223	3.742.213	2.503.012
24	ENEMDU 2008-06-01	Empleo Adecuado/Pleno	2750595.0	2248018	502577	1.937.602	812.993
25	ENEMDU 2008-06-01	Subempleo	1130699.0	696949	433750	665.939	464.760
26	ENEMDU 2008-06-01	Subempleo por insuficiencia de tiempo de trabajo	846132.0	527132	319001	457.856	388.276
27	ENEMDU 2008-06-01	Subempleo por insuficiencia de ingresos	284567.0	169817	114750	208.083	76.484
28	ENEMDU 2008-06-01	Empleo no remunerado	614172.0	257629	356542	204.623	409.548
29	ENEMDU 2008-06-01	Otro Empleo no pleno	1688693.0	909254	779439	898.349	790.344
30	ENEMDU 2008-06-01	Empleo no clasificado	61065.0	39152	21914	35.699	25.366
31	ENEMDU 2008-06-01	Desempleo	344143.0	282489	61654	151.283	192.860
32	ENEMDU 2008-06-01	Desempleo Abierto	243614.0	217324	26290	123.250	120.364
33	ENEMDU 2008-06-01	Desempleo Oculto	100530.0	65165	35365	28.033	72.497
34	ENEMDU 2008-06-01	Desempleo Cesante	238031.0	207834	30198	113.592	124.440
35	ENEMDU 2008-06-01	Desempleo Nuevo	106112.0	74655	31457	37.691	68.421
36	ENEMDU 2008-06-01	Población Económicamente Inactiva	2924281.0	2040873	883408	785.331	2.138.950
37	ENEMDU 2008-12-01	Población Total	13878704.0	9202348	4676355	6.831.453	7.047.250
38	ENEMDU 2008-12-01	Población menor de 15 años	4229707.0	2633450	1596257	2.175.561	2.054.146
39	ENEMDU 2008-12-01	Población en Edad de Trabajar (PET)	9648996.0	6568898	3080098	4.655.892	4.993.104
40	ENEMDU 2008-12-01	Población Económicamente Activa	6385421.0	4329308	2056113	3.800.531	2.584.890
41	ENEMDU 2008-12-01	Empleo	6005395.0	4012298	1993097	3.635.236	2.370.159
42	ENEMDU 2008-12-01	Empleo Adecuado/Pleno	2858659.0	2339288	519370	1.991.157	867.502
43	ENEMDU 2008-12-01	Subempleo	957978.0	588824	369155	549.097	408.881
44	ENEMDU 2008-12-01	Subempleo por insuficiencia de tiempo de trabajo	704516.0	445009	259507	378.324	326.192
45	ENEMDU 2008-12-01	Subempleo por insuficiencia de ingresos	253462.0	143815	109648	170.773	82.689
46	ENEMDU 2008-12-01	Empleo no remunerado	523928.0	211070	312858	159.714	364.214

47	ENEMDU	2008-12-01	Otro Empleo no pleno	1649349.0	859126	790223	923.415	725.935
48	ENEMDU	2008-12-01	Empleo no clasificado	15481.0	13990	1491	11.853	3.628
49	ENEMDU	2008-12-01	Desempleo	380026.0	317011	63015	165.296	214.730
50	ENEMDU	2008-12-01	Desempleo Abierto	252924.0	226983	25942	119.858	133.067
51	ENEMDU	2008-12-01	Desempleo Oculto	127102.0	90028	37074	45.438	81.664
52	ENEMDU	2008-12-01	Desempleo Cesante	225489.0	196643	28847	109.477	116.013
53	ENEMDU	2008-12-01	Desempleo Nuevo	154537.0	120368	34169	55.819	98.718

```
seleccionadohist = parametrihist.loc[(parametrihist['Indicadores'] == 'Subempleo') | (parametrihist['Indicadores'] == 'Subempleo por insuficiencia de tiempo de trabajo')]
seleccionadohist
```

Encuesta	Periodo	Indicadores	Nacional/Total	Área urbana	Área rural	Sexo-Hombre	Sexo-Mujer	
6	ENEMDU	2007-12-01	Empleo Adecuado/Pleno	2737158.0	2236440	500719	1.907.451	829.707
7	ENEMDU	2007-12-01	Subempleo	1155872.0	687714	468158	679.118	476.753
8	ENEMDU	2007-12-01	Subempleo por insuficiencia de tiempo de trabajo	860315.0	519627	340688	474.355	385.960
9	ENEMDU	2007-12-01	Subempleo por insuficiencia de ingresos	295556.0	168087	127470	204.763	90.793
11	ENEMDU	2007-12-01	Otro Empleo no pleno	1504000.0	785461	718539	830.341	673.659
24	ENEMDU	2008-06-01	Empleo Adecuado/Pleno	2750595.0	2248018	502577	1.937.602	812.993
25	ENEMDU	2008-06-01	Subempleo	1130699.0	696949	433750	665.939	464.760
26	ENEMDU	2008-06-01	Subempleo por insuficiencia de tiempo de trabajo	846132.0	527132	319001	457.856	388.276
27	ENEMDU	2008-06-01	Subempleo por insuficiencia de ingresos	284567.0	169817	114750	208.083	76.484
29	ENEMDU	2008-06-01	Otro Empleo no pleno	1688693.0	909254	779439	898.349	790.344
42	ENEMDU	2008-12-01	Empleo Adecuado/Pleno	2858659.0	2339288	519370	1.991.157	867.502
43	ENEMDU	2008-12-01	Subempleo	957978.0	588824	369155	549.097	408.881
44	ENEMDU	2008-12-01	Subempleo por insuficiencia de tiempo de trabajo	704516.0	445009	259507	378.324	326.192
45	ENEMDU	2008-12-01	Subempleo por insuficiencia de ingresos	253462.0	143815	109648	170.773	82.689
47	ENEMDU	2008-12-01	Otro Empleo no pleno	1649349.0	859126	790223	923.415	725.935

```
seleccionadohist['Indicadores'] = seleccionadohist['Indicadores'].str.replace('Subempleo por insuficiencia de tiempo de trabajo','Subempleo')
seleccionadohist['Indicadores'] = seleccionadohist['Indicadores'].str.replace('Subempleo por insuficiencia de ingresos','Subempleo')
seleccionadohist
```

```
/usr/local/lib/python3.7/dist-packages/ipykernel_launcher.py:1: SettingWithCopyWarning:
```

```
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row_indexer,col_indexer] = value instead
```

```
See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user\_guide/indexing.html#returning-a-view-versus-a-copy.
```

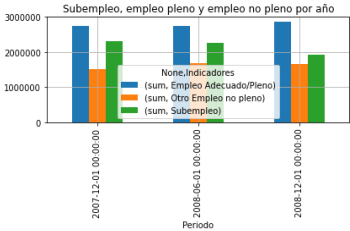
```
/usr/local/lib/python3.7/dist-packages/ipykernel_launcher.py:2: SettingWithCopyWarning:
```

```
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row_indexer,col_indexer] = value instead
```

```
dfseleccionadohist= seleccionadohist['Indicadores'].value_counts().reset_index()  
dfseleccionadohist
```

	index	Indicadores
0	Subempleo	9
1	Empleo Adecuado/Pleno	3
2	Otro Empleo no pleno	3

```
stack_seleccion = seleccionadohist.groupby(['Periodo', 'Indicadores'])['Nacional/Total'].agg([np.sum]).unstack()  
stack_seleccion.plot(kind="bar")  
plt.grid()  
plt.ticklabel_format(axis='y',style='plain')  
plt.title('Subempleo, empleo pleno y empleo no pleno por año')  
plt.tight_layout()
```



• Describir con estadística descriptiva los datos que se encuentran en el archivos.

▼ Archivo Poblaciones

```
df6seleccionado= df6.iloc[:,3:8]  
  
df6seleccionado=df6seleccionado.replace('\.','',regex=True)  
  
df6seleccionado= df6seleccionado.astype(int, errors = 'raise')  
  
df6seleccionado.describe()
```


	Nacional/Total	Área urbana	Área rural	Sexo-Hombre	Sexo-Mujer
count	9.000000e+02	9.000000e+02	9.000000e+02	9.000000e+02	9.000000e+02
mean	3.424082e+06	2.320600e+06	1.103482e+06	1.770175e+06	1.653907e+06
std	4.522188e+06	3.081895e+06	1.454728e+06	2.299806e+06	2.269156e+06
min	1.294700e+04	1.231800e+04	4.240000e+02	8.393000e+03	3.628000e+03
25%	2.507870e+06	1.070288e+06	4.544525e+04	1.438025e+06	1.152755e+06
50%	4.399430e+06	3.270104e+06	1.112270e+06	2.009110e+06	2.020070e+06

▼ Archivo Tasas

```
import pandas as pd
dftasas = pd.read_csv('2.Tasas.csv', sep = ';', encoding = 'latin1',header=1, error_bad_lines=False) #error_bad_lines is avoid single line error
dftasas = dftasas.rename(columns={'Área':'Área urbana ','Unnamed: 5':'Área rural','Nacional':'Nacional/Total','Sexo':'Sexo-Hombre','Unnamed: 7':'Sexo-Mujer'})

dftasassele = dftasas.drop([0]).iloc[:,3:8]
dftasassele
```

	Nacional/Total	Área urbana	Área rural	Sexo-Hombre	Sexo-Mujer
1	64,7	62,6	69,0	80,0	50,1
2	95,0	93,9	97,2	96,2	93,3
3	43,2	52,9	23,7	50,5	32,4
4	18,2	16,3	22,2	18,0	18,6
5	13,6	12,3	16,2	12,6	15,1
...
696	4,6	5,7	2,3	3,6	5,9
697	4,0	5,2	1,6	3,1	5,3
698	0,6	0,5	0,7	0,5	0,6
699	66,0	62,7	73,5	78,7	53,8
700	46,9	45,6	49,8	55,9	38,3

700 rows x 5 columns

```
dftasassele=dftasassele.replace('\ ','',regex=True)

dftasassele = dftasassele.astype(int, errors = 'raise')

dftasassele.describe()
```

	Nacional/Total	Área urbana	Área rural	Sexo-Hombre	Sexo-Mujer
count	700.000000	700.000000	700.000000	700.000000	700.000000
mean	280.738571	275.884286	291.610000	307.172857	255.324286
std	287.191993	285.384382	302.381077	324.321160	255.665541
min	2.000000	2.000000	0.000000	2.000000	1.000000
25%	40.000000	46.000000	23.000000	35.000000	46.750000
50%	163.000000	155.000000	196.500000	163.500000	180.500000
75%	469.000000	487.250000	485.250000	558.000000	384.750000
max	963.000000	955.000000	986.000000	969.000000	956.000000

▼ Archivo Categorizacion empleados

```
import pandas as pd
dfcempleados = pd.read_csv('3.1. Caracterización empleados.csv', sep = ';', encoding = 'latin1',header=1, error_bad_lines=False)
dfcempleados.drop(dfcempleados.columns[dfcempleados.columns.str.contains('unnamed',case = False)],axis = 1, inplace = True)
dfcempleacc=dfcempleados.iloc[0:28,1:51]
```

```
dfcemselecc = dfcemselecc.replace('\.', '.', regex=True)

dfcemselecc = dfcemselecc.replace('\%', '', regex=True)

dfcemselecc = dfcemselecc.replace('\-', '', regex=True)

dfcemselecc = dfcemselecc.replace('', 0.0, regex=True)

for i in range(50):
    va=dfcemselecc.columns[i]
    dfcemselecc[va]=(dfcemselecc[va]).apply(lambda x: float(x))
```

```
dfcemselecc.describe()
```

	dic-07	jun-08	dic-08	dic-09	jun-10	dic-10	jun-11	dic-11	jun-12
count	28.000000	28.000000	28.000000	28.000000	28.000000	28.000000	28.000000	28.000000	28.000000
mean	29.446429	29.500000	29.517857	29.446429	29.317857	29.442857	29.392857	29.357143	29.257143
std	24.850527	24.710562	24.599805	24.612892	23.719519	24.471532	24.301576	24.337890	24.383592
min	0.000000	0.000000	0.000000	0.000000	0.000000	0.100000	0.100000	0.100000	0.200000
25%	7.475000	6.875000	7.850000	7.900000	8.200000	7.250000	8.225000	8.200000	8.175000
50%	22.600000	23.000000	22.650000	22.550000	22.350000	22.600000	23.000000	23.750000	23.750000
75%	42.800000	43.375000	43.600000	43.725000	42.700000	43.125000	42.600000	42.950000	41.650000
max	92.400000	91.600000	91.800000	91.800000	90.700000	90.600000	90.200000	90.800000	91.600000

▼ Archivo Categorizacion empleo adecuado

```
import pandas as pd
dfcempleadosa = pd.read_csv('3.2. Caracterización adec_pleno.csv', sep = ';', encoding = 'latin1',header=1, error_bad_lines=False) #error_bad_lines
dfcempleadosa.drop(dfcempleadosa.columns[dfcempleadosa.columns.str.contains('unnamed',case = False)],axis = 1, inplace = True)
dfcempleadosa=dfcempleadosa.iloc[0:25,1:51]

dfcempleadosa = dfcempleadosa.replace('\.', '.', regex=True)

dfcempleadosa = dfcempleadosa.replace('\%', '', regex=True)

dfcempleadosa = dfcempleadosa.replace('\-', '', regex=True)

dfcempleadosa = dfcempleadosa.replace('', 0.0, regex=True)

for i in range(50):
    va=dfcempleadosa.columns[i]
    dfcempleadosa[va]=(dfcempleadosa[va]).apply(lambda x: float(x))
```

```
dfcempleadosa.describe()
```

	dic-07	jun-08	dic-08	dic-09	jun-10	dic-10	jun-11	dic-11	jun-12
count	25.00000	25.00000	25.00000	25.00000	25.00000	25.00000	25.00000	25.00000	25.00000
mean	27.99200	27.99600	27.97600	27.97600	27.98000	28.00000	27.99200	27.98000	28.00000
std	26.96357	27.18151	27.04672	27.35177	27.03984	27.57393	27.54995	27.64932	27.78247
min	0.00000	0.00000	0.00000	0.00000	0.10000	0.10000	0.10000	0.10000	0.10000
25%	4.10000	4.10000	4.30000	4.40000	4.60000	4.10000	4.10000	4.10000	4.20000

▼ Archivo Categorizacion Subempleo

```

import pandas as pd
dfcsubempleado = pd.read_csv('3.3. Caracterización subempleo.csv', sep = ';', encoding = 'latin1',header=1, error_bad_lines=False) #error_bad_lines
dfcsubempleado.drop(dfcsubempleado.columns[dfcsubempleado.columns.str.contains('unnamed',case = False)],axis = 1, inplace = True)
dfcsubempleado=dfcsubempleado.iloc[0:22,1:51]

dfcsubempleado = dfcsubempleado.replace('\,','.',regex=True)

dfcsubempleado = dfcsubempleado.replace('\%', '', regex=True)

dfcsubempleado = dfcsubempleado.replace('\-', '', regex=True)

dfcsubempleado = dfcsubempleado.replace('',0.0,regex=True)

for i in range(50):
    va=dfcsubempleado.columns[i]
    dfcsubempleado[va]=(dfcsubempleado[va]).apply(lambda x: float(x))

dfcsubempleado.describe()
```

	dic-07	jun-08	dic-08	dic-09	jun-10	dic-10	jun-11	dic-11	jun-12
count	21.00000	21.00000	21.00000	21.00000	21.00000	21.00000	21.00000	21.00000	21.00000
mean	28.571429	28.576190	28.576190	28.576190	28.571429	28.566667	28.580952	28.580952	28.566667
std	28.151113	28.752981	27.899837	28.333353	27.209468	28.166191	28.251666	28.232758	28.582098
min	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.500000	0.200000	0.100000
25%	4.800000	5.400000	5.500000	5.200000	6.300000	4.800000	4.800000	5.300000	5.600000
50%	21.300000	21.700000	22.900000	22.300000	21.200000	20.400000	20.800000	21.200000	19.000000
75%	47.600000	49.200000	47.600000	48.000000	48.700000	49.800000	49.100000	46.400000	45.500000
max	97.700000	98.000000	98.100000	98.000000	97.400000	97.800000	98.400000	99.200000	99.100000

▼ Archivo Categorizacion Otro empleo no pleno

```

import pandas as pd
dfcbempleadooe = pd.read_csv('3.4 Caracterización Ot. no pleno.csv', sep = ';', encoding = 'latin1',header=1, error_bad_lines=False) #error_bad_line
dfcbempleadooe.drop(dfcbempleadooe.columns[dfcbempleadooe.columns.str.contains('unnamed',case = False)],axis = 1, inplace = True)
dfcbempleadooe=dfcbempleadooe.iloc[0:21,1:51]

dfcbempleadooe = dfcbempleadooe.replace('\,','.',regex=True)

dfcbempleadooe = dfcbempleadooe.replace('\%', '', regex=True)

dfcbempleadooe = dfcbempleadooe.replace('\-', '', regex=True)

dfcbempleadooe = dfcbempleadooe.replace('',0.0,regex=True)
```

```
for i in range(50):
    va=dfcbempleadooe.columns[i]
    dfcbempleadooe[va]=(dfcbempleadooe[va]).apply(lambda x: float(x))
```

dfcbempleadooe.describe()

	dic-07	jun-08	dic-08	dic-09	jun-10	dic-10	jun-11	dic-11	jun-12
count	21.000000	21.000000	21.000000	21.000000	21.000000	21.000000	21.000000	21.000000	21.000000
mean	28.580952	28.576190	28.571429	28.566667	28.576190	28.571429	28.566667	28.571429	28.566667
std	27.804201	27.832641	27.501003	27.909646	26.952827	27.937522	27.900902	28.139850	27.951750
min	0.000000	0.000000	0.000000	0.000000	0.100000	0.100000	0.100000	0.000000	0.200000
25%	6.400000	6.500000	6.000000	5.500000	5.700000	6.500000	5.500000	5.700000	6.300000
50%	18.900000	18.000000	18.000000	17.500000	17.600000	17.900000	17.700000	16.500000	17.400000
75%	47.700000	48.900000	49.700000	46.100000	44.300000	42.300000	42.500000	42.700000	42.300000
max	98.800000	98.000000	98.600000	98.400000	98.600000	98.400000	98.500000	98.800000	98.700000

▼ Archivo Categorizacion Desempleo

```
import pandas as pd
dfcdesempleo = pd.read_csv('3.5 Caracterización Desempleo.csv', sep = ';', encoding = 'latin1',header=1, error_bad_lines=False) #error_bad_lines is
dfcdesempleo.drop(dfcdesempleo.columns[dfcdesempleo.columns.str.contains('unnamed',case = False)],axis = 1, inplace = True)
dfcdesempleo=dfcdesempleo.iloc[0:12,1:51]
```

```
dfcdesempleo = dfcdesempleo.replace('\,',',',regex=True)
```

```
dfcdesempleo = dfcdesempleo.replace('\%',',',regex=True)
```

```
dfcdesempleo = dfcdesempleo.replace('\-',',',regex=True)
```

```
dfcdesempleo = dfcdesempleo.replace(',',0.0,regex=True)
```

```
for i in range(50):
    va=dfcdesempleo.columns[i]
    dfcdesempleo[va]=(dfcdesempleo[va]).apply(lambda x: float(x))
```

dfcdesempleo.describe()

	dic-07	jun-08	dic-08	dic-09	jun-10	dic-10	jun-11	dic-11	jun-12
count	11.000000	11.000000	11.000000	11.000000	11.000000	11.000000	11.000000	11.000000	11.000000
mean	36.363636	36.363636	36.363636	36.363636	36.354545	36.363636	36.354545	36.363636	36.363636
std	19.800165	23.061582	21.215102	22.073345	21.816983	21.903163	23.322108	21.953099	24.663020
min	3.100000	2.000000	2.200000	1.400000	2.200000	1.600000	1.700000	1.400000	2.200000
25%	18.700000	19.300000	18.350000	20.400000	21.400000	20.600000	19.050000	20.900000	16.350000
50%	40.000000	30.800000	40.700000	34.600000	38.000000	33.100000	33.700000	38.100000	31.900000
75%	50.000000	51.800000	52.650000	50.000000	50.000000	50.000000	50.000000	50.000000	50.000000
max	61.000000	70.800000	66.600000	69.500000	72.400000	69.500000	76.000000	70.900000	79.300000

▼ Archivo Sectorizacion del empleo

```
import pandas as pd
dfsectempleo = pd.read_csv('4. Sectorización del empleo.csv', sep = ';', encoding = 'latin1',header=1, error_bad_lines=False) #error_bad_lines is av
dfsectempleo.drop(dfsectempleo.columns[dfsectempleo.columns.str.contains('unnamed',case = False)],axis = 1, inplace = True)
dfsectempleo=dfsectempleo.iloc[0:22,1:67]
```

```
dfsectempleo = dfsectempleo.replace('\,','.',regex=True)
```

```
dfsectempleo = dfsectempleo.replace('%','',regex=True)
```

```
dfsectempleo = dfsectempleo.replace('\-','',regex=True)
```

```
dfsectempleo = dfsectempleo.replace(' ',0.0,regex=True)
```

```
for i in range(65):
    va=dfsectempleo.columns[i]
    dfsectempleo[va]=(dfsectempleo[va]).apply(lambda x: float(x))
```

```
dfsectempleo.describe()
```

	jun-07	sep-07	dic-07	mar-08	jun-08	sep-08	dic-08	mar-09	jun-09	sep-09	dic-09	mar-10	jun-10	sep-10	dic-10
count	12.000000	12.000000	12.000000	12.000000	12.000000	12.000000	12.000000	12.0000	12.000000	12.000000	12.000000	12.000000	12.000000	12.000000	12.000000
mean	8.333333	8.333333	25.008333	8.333333	25.000000	8.333333	25.008333	8.3250	8.333333	8.325000	25.000000	8.333333	24.983333	8.333333	24.991667
std	17.979096	18.105716	22.260706	17.854988	23.113003	18.012336	22.237824	18.2304	18.135968	18.428492	22.674776	18.110486	22.546229	18.618385	23.756127
min	0.000000	0.000000	1.500000	0.000000	1.800000	0.000000	1.900000	0.0000	0.000000	0.000000	2.000000	0.000000	1.900000	0.000000	1.800000
25%	0.000000	0.000000	6.850000	0.000000	5.650000	0.000000	5.850000	0.0000	0.000000	0.000000	5.525000	0.000000	5.700000	0.000000	4.300000
50%	0.000000	0.000000	15.900000	0.000000	15.050000	0.000000	16.950000	0.0000	0.000000	0.000000	16.200000	0.000000	16.800000	0.000000	16.200000
75%	4.100000	4.100000	42.025000	4.200000	43.225000	4.200000	43.600000	3.9750	4.375000	4.125000	43.725000	4.425000	42.950000	3.650000	43.900000
max	54.700000	55.700000	66.700000	54.400000	68.100000	56.000000	64.200000	57.4000	57.700000	59.400000	65.600000	57.500000	64.500000	59.400000	65.700000

Conclusiones

Los graficos de barras, líneas y el grafico de pastel fueron clave para visualizar de mejor manera la forma en que se distribuye el Mercado Laboral en Ecuador.Por las graficas se evidencio que apesar del duro golpe de la pandemia en estos dos ultimos años predomino el empleo, hubo formas de las personas sustentarse y el 2021 fue un año de levantar el mercado en comparacion de otros años.

Recomendaciones

Como trabajo futuro analizar quizas las formas en que ha cambiado la sectorizacion del empleo poder clasificar las edades de las personas y a que Indicadores estan asignados para poder comprender mejor a que grupo de personas estan alcanzando mejor oportunidades y brindar herramientas para los ciudadanos que estan siendo dejados.

✓ 0 s completado a las 3:19

