

10 - Parte EDA & Limpieza - Sección NPCKP

October 31, 2022

1 Análisis de la Encuesta Multiproposito

1.1 Importación de Paquetes y carga del archivo

```
[181]: import pandas as pd
from matplotlib import pyplot as plt
import seaborn as sns
from scipy import stats
import numpy as np
```

```
[182]: data = pd.read_excel('Encuesta_Multiproposito_Suba.xlsx')
```

```
[183]: data.shape
```

```
[183]: (24536, 498)
```

Hay en total 24536 encuestados en la Localidad de Suba

2 Décima Sección (NPCKP)

2.1 NPCKP1

2.1.1 1. ¿En qué actividad ocupó ...la mayor parte del tiempo la SEMANA PASADA?

1. Trabajando
2. Buscando trabajo
3. Estudiando
4. Oficios del hogar
5. Incapacitado(a) permanente para trabajar
6. Otra actividad

Datos: 22213

```
[184]: data['NPCKP1'].count()
```

```
[184]: 22213
```

```
[185]: data['NPCKP1'].value_counts()
```

```
[185]: 1.0    11336
        4.0    4324
        3.0    3379
        6.0    1358
        2.0    1127
        5.0     689
        Name: NPCKP1, dtype: int64
```

```
[186]: data['NPCKP1'].describe().apply("{0:.2f}".format)
```

```
[186]: count    22213.00
        mean       2.37
        std       1.60
        min       1.00
        25%       1.00
        50%       1.00
        75%       4.00
        max       6.00
        Name: NPCKP1, dtype: object
```

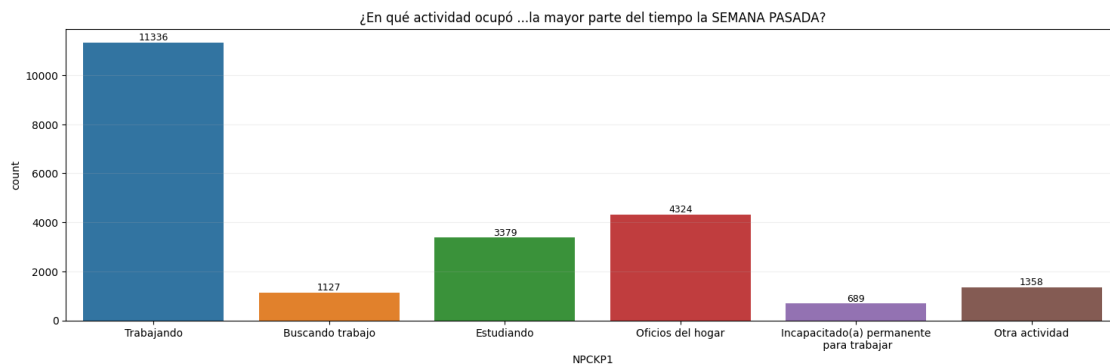
```
[187]: fig, ax = plt.subplots(figsize=(18, 5))

        g = sns.countplot(ax=ax, data = data, x = 'NPCKP1')

        for bars in ax.containers:
            ax.bar_label(bars, fmt='%.0f', fontsize=9)

        ax.set_title('¿En qué actividad ocupó ...la mayor parte del tiempo la SEMANA_
        ↪PASADA?')
        ax.set_xticklabels(['Trabajando', 'Buscando trabajo', 'Estudiando', 'Oficios del_
        ↪hogar',
                            'Incapacitado(a) permanente\n para trabajar', 'Otra_
        ↪actividad'])

        plt.grid(alpha = 0.2, axis = 'y')
        plt.show()
```



2.2 NPCKP6

2.2.1 6. ¿Qué hizo principalmente en las ÚLTIMAS 4 SEMANAS ...para conseguir un trabajo o instalar un negocio?

1. Pidió ayuda a familiares, amigos/as o colegas
2. Visitó, llevó o envió hojas de vida a empresas o empleadores
3. Visitó, llevó o envió hojas de vida a bolsas de empleo o intermediarios
4. Puso o consultó avisos clasificados
5. Se presentó a convocatorias
6. Hizo preparativos para iniciar un negocio
7. Utilizó el servicio público de empleo (SENA, cajas de compensación, Agencias Públicas de Empleo)
8. Buscó en internet
9. Otro medio

Datos: 1172

```
[188]: data['NPCKP6'].count()
```

```
[188]: 1172
```

```
[189]: data['NPCKP6'].value_counts()
```

```
[189]: 1.0    345
      8.0    312
      2.0    265
      5.0    104
      3.0     87
      6.0     20
      4.0     19
      7.0     19
      9.0      1
      Name: NPCKP6, dtype: int64
```

```
[190]: data['NPCKP6'].describe().apply("{0:.2f}".format)
```

```
[190]: count    1172.00
      mean      3.83
      std       2.86
      min       1.00
      25%       1.00
      50%       2.00
      75%       8.00
      max       9.00
      Name: NPCKP6, dtype: object
```

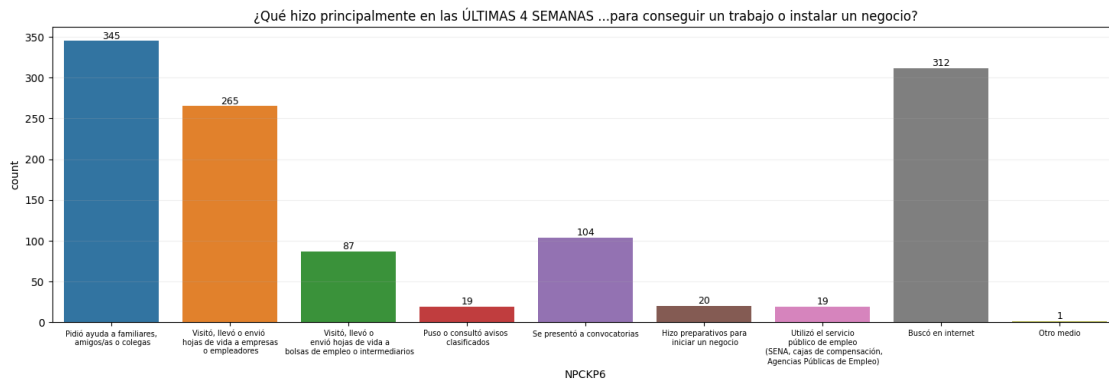
```
[191]: fig, ax = plt.subplots(figsize=(18, 5))

g = sns.countplot(ax=ax, data = data, x = 'NPCKP6')

for bars in ax.containers:
    ax.bar_label(bars, fmt='%.0f', fontsize=9)

ax.set_title('¿Qué hizo principalmente en las ÚLTIMAS 4 SEMANAS ...para
↳conseguir un trabajo o instalar un negocio?')
ax.set_xticklabels(['Pidió ayuda a familiares,\n amigos/as o colegas', 'Visitó,\n
↳llevó o envió \nhojas de vida a empresas\n o empleadores',
                    'Visitó, llevó o\n envió hojas de vida a\n bolsas de empleo\
↳o intermediarios', 'Puso o consultó avisos \n clasificados',
                    'Se presentó a convocatorias', 'Hizo preparativos para\n
↳iniciar un negocio',
                    'Utilizó el servicio\n público de empleo \n (SENA, cajas de\
↳compensación,\n Agencias Públicas de Empleo)',
                    'Buscó en internet', 'Otro medio'], fontsize=7)

plt.grid(alpha = 0.2, axis = 'y')
plt.show()
```



2.3 NPCKP7

2.3.1 7. ¿Desea ...conseguir un trabajo remunerado o instalar un negocio?

1. Sí
2. No

Datos: 8348

```
[192]: data['NPCKP7'].count()
```

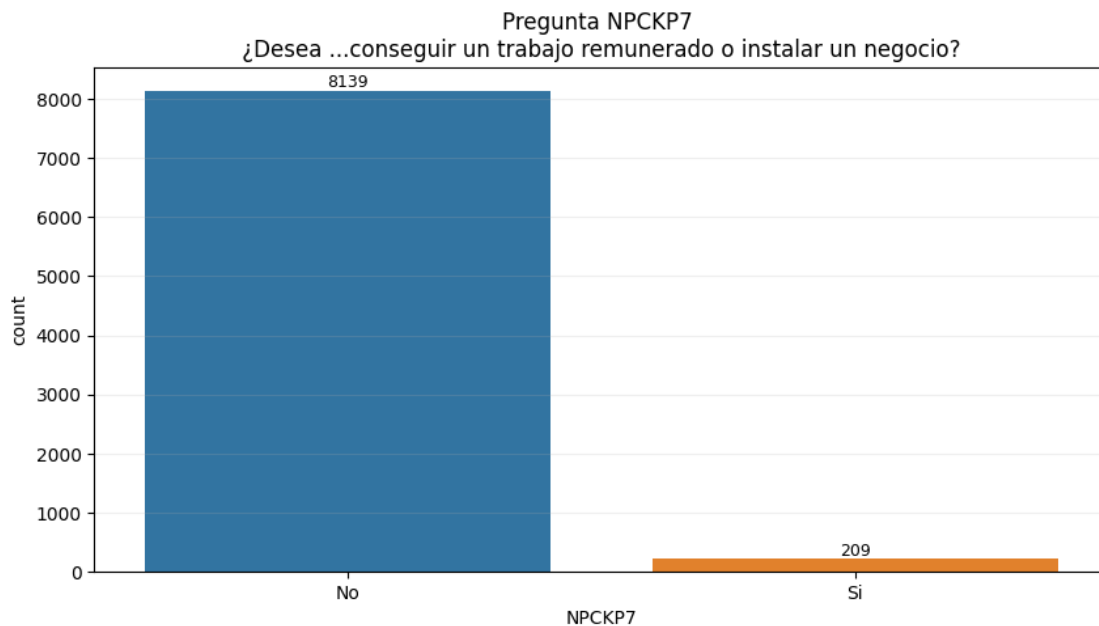
[192]: 8348

```
[193]: data['NPCKP7'].value_counts()
```

```
[193]: 2.0    8139  
      1.0     209  
      Name: NPCKP7, dtype: int64
```

```
[194]: data = data.replace({'NPCKP7':2},0)
```

```
[195]: fig, ax = plt.subplots(figsize=(10, 5))  
      g = sns.countplot(ax=ax, data = data, x ='NPCKP7')  
  
      for bars in ax.containers:  
          ax.bar_label(bars, fmt='%.0f', fontsize=9)  
  
      ax.set_title('Pregunta NPCKP7 \n ¿Desea ...conseguir un trabajo remunerado o  
      ↪instalar un negocio? ')  
      ax.set_xticklabels(['No', 'Si'])  
  
      plt.grid(alpha = 0.2, axis = 'y')  
      plt.show()
```



2.4 NPCKP8

2.4.1 8. Aunque ...desea trabajar, ¿por qué motivo principal no hizo diligencias para buscar un trabajo o instalar un negocio, en las ÚLTIMAS 4 SEMANAS?

1. Ya encontró trabajo

2. No hay trabajo disponible en la ciudad o región/No encuentra trabajo en su oficio o profesión
3. Está esperando que le llamen o esperando temporada alta
4. No sabe cómo buscarlo
5. Está cansado(a) de buscar
6. Carece de la experiencia necesaria
7. No tiene recursos para instalar un negocio
8. Los empleadores le consideran muy joven o muy viejo/a
9. Usted se considera muy joven o muy viejo/a
10. Responsabilidades familiares
11. Problemas de salud
12. Está estudiando
13. Otro

Datos: 209

```
[196]: data['NPCKP8'].count()
```

```
[196]: 209
```

```
[197]: data['NPCKP8'].value_counts()
```

```
[197]: 12.0    34
      10.0    33
      7.0    24
      2.0    19
      11.0   17
      3.0    16
      5.0    15
      6.0    14
      13.0   12
      8.0     8
      9.0     7
      1.0     5
      4.0     5
      Name: NPCKP8, dtype: int64
```

```
[198]: data['NPCKP8'].describe().apply("{0:.2f}".format)
```

```
[198]: count      209.00
      mean        7.88
      std         3.64
      min         1.00
      25%         5.00
      50%         8.00
      75%        11.00
      max        13.00
      Name: NPCKP8, dtype: object
```

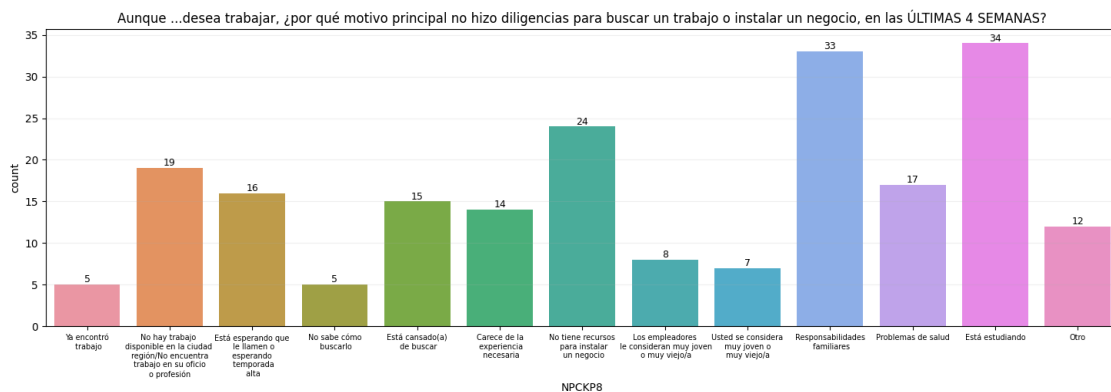
```
[199]: fig, ax = plt.subplots(figsize=(18, 5))

g = sns.countplot(ax=ax, data = data, x ='NPCKP8')

for bars in ax.containers:
    ax.bar_label(bars, fmt='%.0f', fontsize=9)

ax.set_title('Aunque ...desea trabajar, ¿por qué motivo principal no hizo
↳diligencias para buscar un trabajo o instalar un negocio, en las ÚLTIMAS 4
↳SEMANAS?')
ax.set_xticklabels(['Ya encontré\n trabajo',
'No hay trabajo \ndisponible en la ciudad \nregión/No encuentra\n trabajo en su
↳oficio \nno profesión',
'Está esperando que\n le llamen o\n esperando \n temporada\n alta',
'No sabe cómo\n buscarlo',
'Está cansado(a)\n de buscar',
'Carece de la\n experiencia\n necesaria',
'No tiene recursos\n para instalar\n un negocio',
'Los empleadores\n le consideran muy joven\n o muy viejo/a',
'Usted se considera\n muy joven o \nmuy viejo/a',
'Responsabilidades\n familiares',
'Problemas de salud',
'Está estudiando',
'Otro'],fontsize=7)

plt.grid(alpha = 0.2, axis = 'y')
plt.show()
```



2.5 NPCKP9

2.5.1 9. Durante los ÚLTIMOS 12 MESES, ¿ ...trabajó por lo menos DOS SEMANAS consecutivas?

1. Sí

2. No

Datos: 101

```
[200]: data['NPCKP9'].count()
```

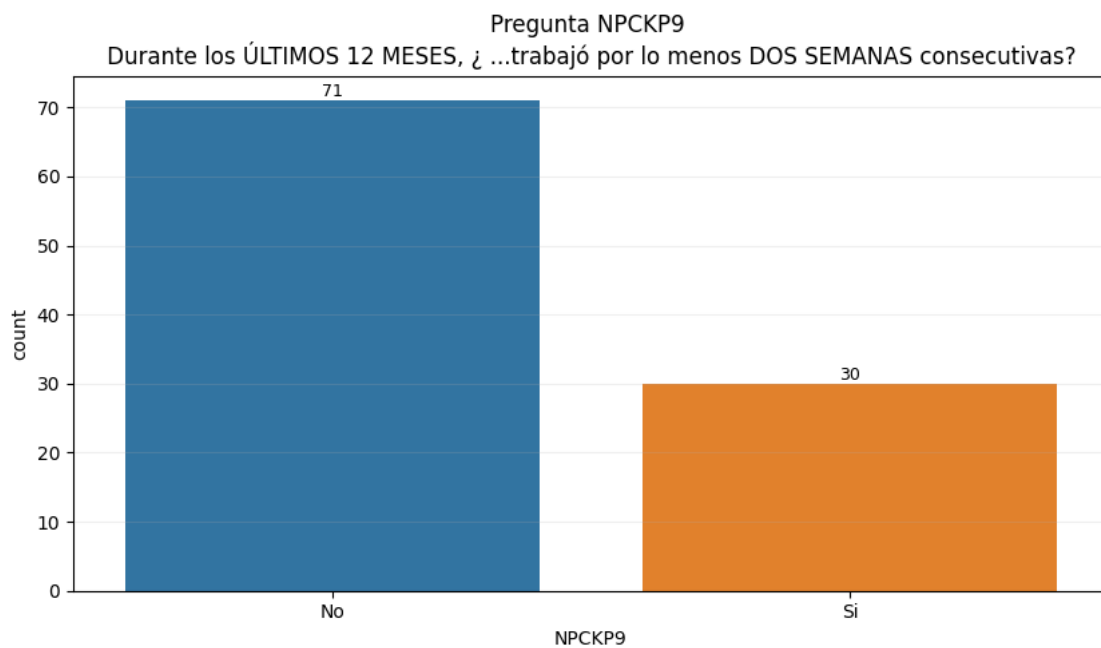
```
[200]: 101
```

```
[201]: data['NPCKP9'].value_counts()
```

```
[201]: 2.0    71  
      1.0    30  
      Name: NPCKP9, dtype: int64
```

```
[202]: data = data.replace({'NPCKP9':2},0)
```

```
[203]: fig, ax = plt.subplots(figsize=(10, 5))  
      g = sns.countplot(ax=ax, data = data, x = 'NPCKP9')  
  
      for bars in ax.containers:  
          ax.bar_label(bars, fmt='%.0f', fontsize=9)  
  
      ax.set_title('Pregunta NPCKP9 \n Durante los ÚLTIMOS 12 MESES, ¿ ...trabajó por lo  
      menos DOS SEMANAS consecutivas? ')  
      ax.set_xticklabels(['No', 'Si'])  
  
      plt.grid(alpha = 0.2, axis = 'y')  
      plt.show()
```



2.6 NPCKP10

2.6.1 10. Después de su ÚLTIMO empleo, ¿ ...ha hecho alguna diligencia para conseguir trabajo o instalar un negocio?

1. Sí
2. No

Datos: 30

```
[204]: data['NPCKP10'].count()
```

```
[204]: 30
```

```
[205]: data['NPCKP10'].value_counts()
```

```
[205]: 1.0    17
      2.0    13
      Name: NPCKP10, dtype: int64
```

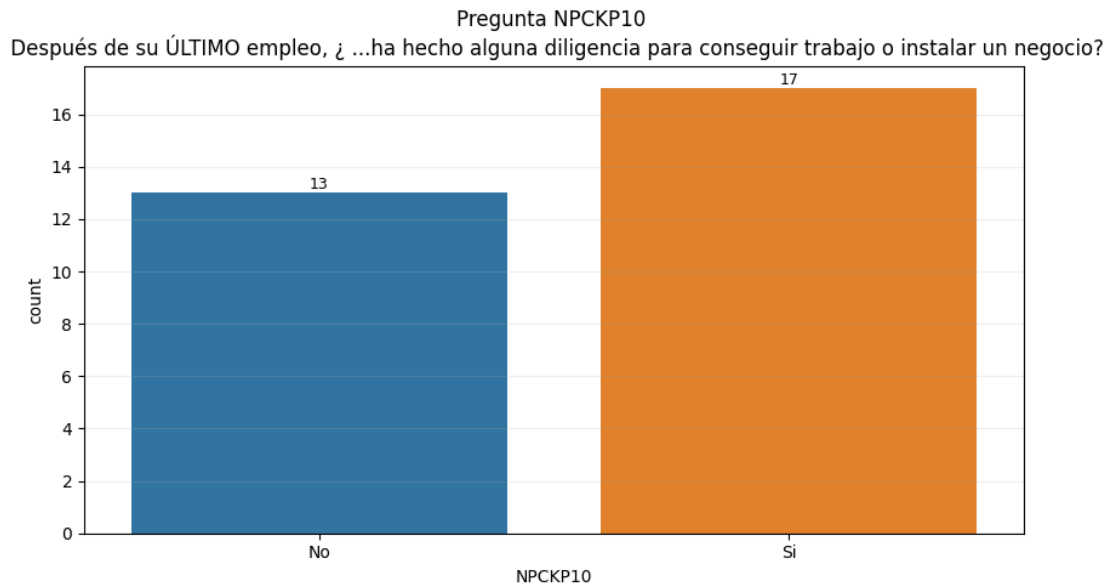
```
[206]: data = data.replace({'NPCKP10':2},0)
```

```
[207]: fig, ax = plt.subplots(figsize=(10, 5))
      g = sns.countplot(ax=ax, data = data, x ='NPCKP10')

      for bars in ax.containers:
          ax.bar_label(bars, fmt='%0f', fontsize=9)

      ax.set_title('Pregunta NPCKP10 \n Después de su ÚLTIMO empleo, ¿ ...ha hecho
          ↪ alguna diligencia para conseguir trabajo o instalar un negocio?')
      ax.set_xticklabels(['No','Si'])

      plt.grid(alpha = 0.2, axis = 'y')
      plt.show()
```



2.7 NPCKP11

2.7.1 11. Durante los ÚLTIMOS 12 MESES, ¿ ...ha hecho alguna diligencia para conseguir trabajo o instalar un negocio?

1. Sí
2. No

Datos: 71

```
[ ]: data['NPCKP11'].count()
```

```
[ ]: 71
```

```
[ ]: data['NPCKP11'].value_counts()
```

```
[ ]: 2.0    48
     1.0    23
     Name: NPCKP11, dtype: int64
```

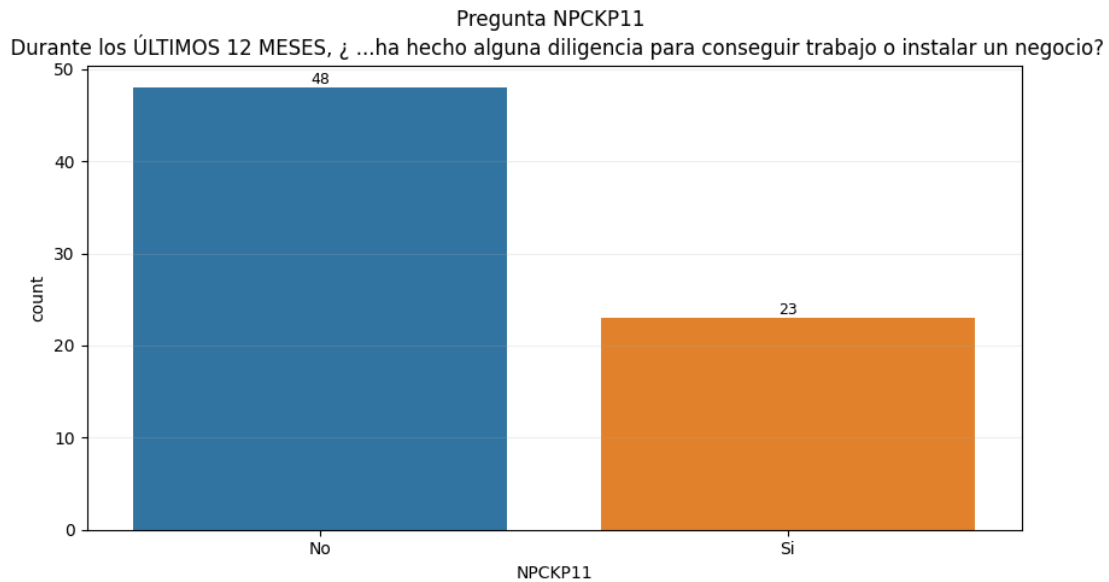
```
[ ]: data = data.replace({'NPCKP11':2},0)
```

```
[ ]: fig, ax = plt.subplots(figsize=(10, 5))
     g = sns.countplot(ax=ax, data = data, x ='NPCKP11')

     for bars in ax.containers:
         ax.bar_label(bars, fmt='%.0f', fontsize=9)
```

```
ax.set_title('Pregunta NPCKP11 \n Durante los ÚLTIMOS 12 MESES, ¿ ...ha hecho_
↪ alguna diligencia para conseguir trabajo o instalar un negocio?')
ax.set_xticklabels(['No', 'Si'])

plt.grid(alpha = 0.2, axis = 'y')
plt.show()
```



2.8 NPCKP17

2.8.1 17. En este trabajo ...es:

1. Obrero o empleado de empresa particular
2. Obrero o empleado del gobierno
3. Empleado doméstico
4. Profesional independiente
5. Trabajador independiente o por cuenta propia
6. Patrón o empleador
7. Trabajador de su propia finca o de finca en arriendo o aparcería
8. Trabajador familiar sin remuneración
9. Trabajador sin remuneración en empresas o negocios de otros hogares
10. Jornalero o peón

Datos: 12204

```
[212]: list_names = ['Obrero o \n empleado de\n empresa particular',
'Obrero o empleado\n del gobierno',
'Empleado doméstico',
'Profesional \nindependiente',
```

```
'Trabajador \nindependiente o\n por cuenta propia',
'Patrón o empleador',
'Trabajador \nde su propia finca\n o de finca en\n arriendo o aparcería',
'Trabajador \nfamiliar sin\n remuneración',
'Trabajador sin\n remuneración en\n empresas o negocios\n de otros hogares',
'Jornalero \no peón']
```

```
[213]: data['NPCKP17'].count()
```

```
[213]: 12004
```

```
[214]: data['NPCKP17'].value_counts()
```

```
[214]: 1.0      6783
5.0      2232
4.0      1501
2.0       810
6.0       370
3.0       218
10.0       51
7.0        28
8.0         8
9.0         3
Name: NPCKP17, dtype: int64
```

```
[215]: data['NPCKP17'].describe().apply("{0:.2f}".format)
```

```
[215]: count      12004.00
mean          2.44
std           1.85
min           1.00
25%           1.00
50%           1.00
75%           4.00
max           10.00
Name: NPCKP17, dtype: object
```

```
[216]: fig, ax = plt.subplots(figsize=(18, 5))

g = sns.countplot(ax=ax, data = data, x ='NPCKP17')

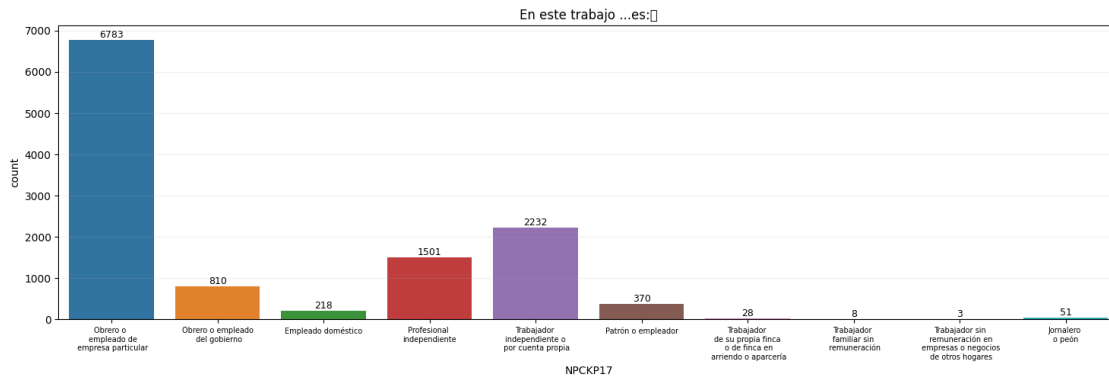
for bars in ax.containers:
    ax.bar_label(bars, fmt='%.0f', fontsize=9)

ax.set_title('En este trabajo ...es:      ')
ax.set_xticklabels(list_names,fontsize=7)
```

```
plt.grid(alpha = 0.2, axis = 'y')
plt.show()
```

c:\Users\andre\OneDrive\Encuesta_MultiProposito\venv\lib\site-packages\IPython\core\pylabtools.py:151: UserWarning: Glyph 9 () missing from current font.

```
fig.canvas.print_figure(bytes_io, **kw)
```



2.9 NPCKP22

2.9.1 22. ¿Por qué medio principal, ...consiguió su empleo o trabajo actual?

1. Pidió ayuda a familiares, amigos/as, colegas
2. Visitó, llevó o envió hojas de vida a empresas o empleadores
3. Visitó, llevó o envió hojas de vida a bolsas de empleo o intermediarios
4. Puso o consultó avisos clasificados
5. Por convocatorias
6. Por el servicio público de empleo (SENA, cajas de compensación, Agencias Públicas de Empleo)
7. A través de internet

Datos:

```
[217]: list_names = ['Pidió ayuda a\n familiares, amigos/as,\n colegas',
'Visitó, llevó o \nenvió hojas de\n vida a empresas\n o empleadores',
'Visitó, llevó o\n envió hojas de\n vida a bolsas de\n empleo o intermediarios',
'Puso o consultó\n avisos clasificados',
'Por convocatorias',
'Por el servicio\n público de empleo\n (SENA,\n cajas de compensación,\n ↪Agencias Públicas de Empleo)',
'A través\n de internet'
]
```

```
[218]: data['NPCKP22'].count()
```

```
[218]: 7862
```

```
[219]: data['NPCKP22'].value_counts()
```

```
[219]: 1.0    3607
      5.0    1625
      2.0    1253
      7.0     774
      3.0     372
      4.0     139
      6.0      92
      Name: NPCKP22, dtype: int64
```

```
[220]: data['NPCKP22'].describe().apply("{0:.2f}".format)
```

```
[220]: count    7862.00
      mean      2.78
      std      2.11
      min      1.00
      25%      1.00
      50%      2.00
      75%      5.00
      max      7.00
      Name: NPCKP22, dtype: object
```

```
[221]: fig, ax = plt.subplots(figsize=(18, 5))

      g = sns.countplot(ax=ax, data = data, x = 'NPCKP22')

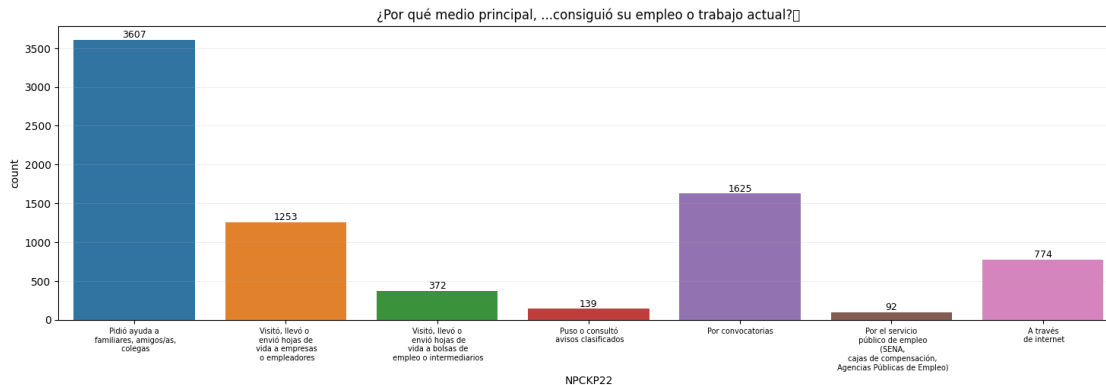
      for bars in ax.containers:
          ax.bar_label(bars, fmt='%.0f', fontsize=9)

      ax.set_title('¿Por qué medio principal, ...consiguió su empleo o trabajo actual?
      ↪      ')
      ax.set_xticklabels(list_names,fontsize=7)

      plt.grid(alpha = 0.2, axis = 'y')
      plt.show()
```

```
c:\Users\andre\OneDrive\Encuesta_MultiProposito\venv\lib\site-
packages\IPython\core\pylabtools.py:151: UserWarning: Glyph 9 ( ) missing from
current font.
```

```
    fig.canvas.print_figure(bytes_io, **kw)
```



2.10 NPCKP45{A-O}

2.10.1 45. En un trayecto normal a su sitio de trabajo, ¿cuáles medios de transporte utiliza ...:

1. Transmilenio?
2. Buses del SITP?
3. Bus, buseta o colectivo?
4. Automóvil de uso particular?
5. Taxi?
6. Motocicleta?
7. Bicicleta?
8. Transporte de la empresa?
9. A pie?
10. Bus intermunicipal?
11. Bicitaxi o mototaxi?
12. Caballo? (solo para rural Bogotá y municipios)
13. Transporte en vehículos particulares a través de plataformas o aplicaciones?
14. Patineta, bicicleta o moto eléctricos?
15. Otro?

Datos: 6318

```
[222]: disc = [i for i in data.columns if 'NPCKP45' in str(i)]
#disc.remove('NPCIP2DA')
print(disc)
```

```
['NPCKP45A', 'NPCKP45B', 'NPCKP45C', 'NPCKP45D', 'NPCKP45E', 'NPCKP45F',
'NPCKP45G', 'NPCKP45H', 'NPCKP45I', 'NPCKP45J', 'NPCKP45K', 'NPCKP45L',
'NPCKP45M', 'NPCKP45N', 'NPCKP45O']
```

```
[223]: for i in disc:
        print(i, data[i].count())
```

NPCKP45A 6318

```

NPCKP45B 6318
NPCKP45C 6318
NPCKP45D 6318
NPCKP45E 6318
NPCKP45F 6318
NPCKP45G 6318
NPCKP45H 6318
NPCKP45I 6318
NPCKP45J 6318
NPCKP45K 6318
NPCKP45L 412
NPCKP45M 6318
NPCKP45N 6318
NPCKP45O 6318

```

```

[224]: for i in disc:
        data = data.replace({i:2},0)

```

```

[225]: list_disc = ['Transmilenio',
                  'Buses del SITP',
                  'Bus, buseta o colectivo',
                  'Automóvil de uso particular',
                  'Taxi',
                  'Motocicleta',
                  'Bicicleta',
                  'Transporte de la empresa',
                  'A pie',
                  'Bus intermunicipal',
                  'Bicitaxi o mototaxi',
                  'Caballo (solo para rural Bogotá y municipios)',
                  'Transporte en vehículos particulares a través de plataformas o aplicaciones',
                  'Patineta, bicicleta o moto eléctricos',
                  'Otro' ]

```

```

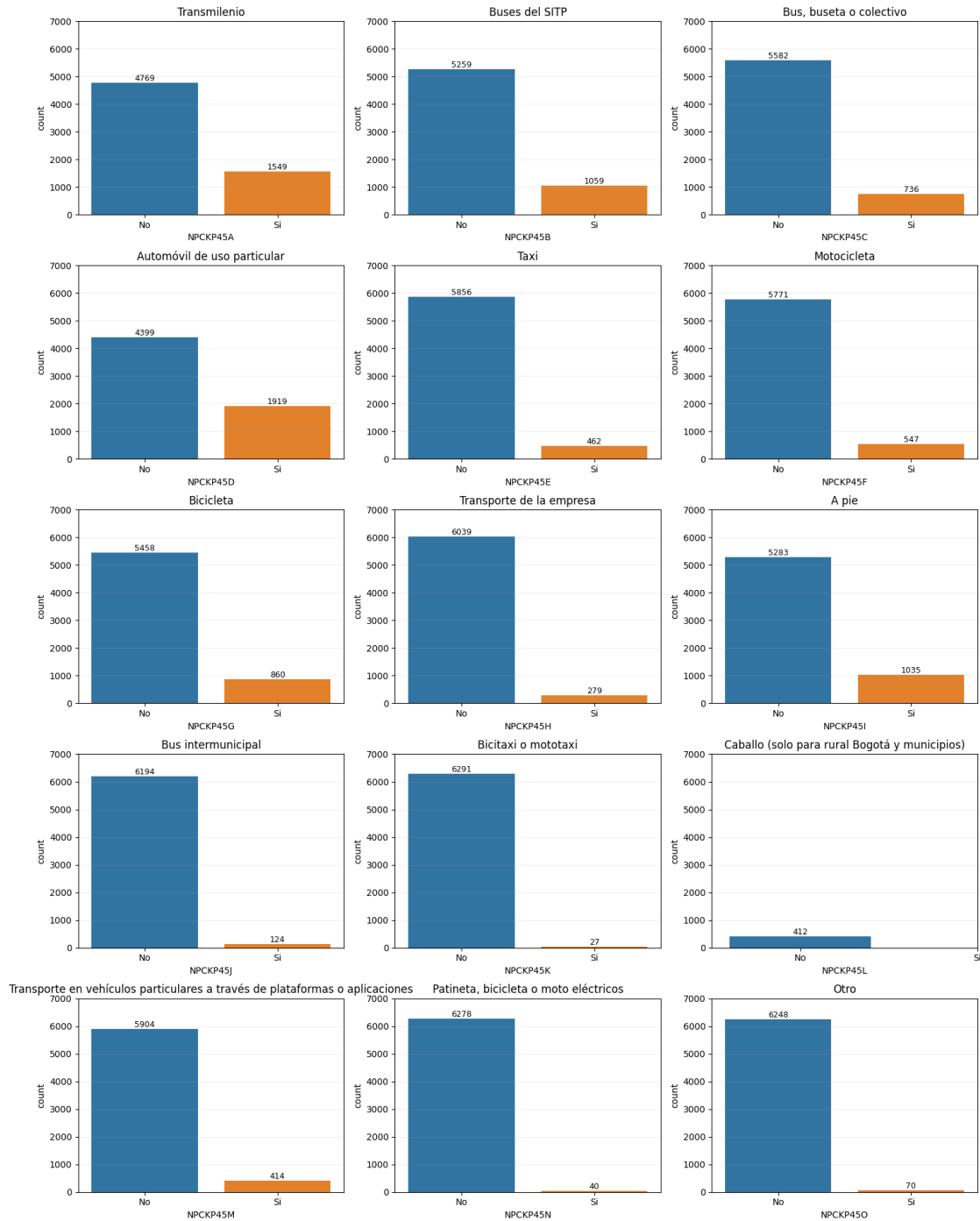
[226]: fig, axes = plt.subplots(5,3, figsize = (15,20), squeeze=False)
        axli = axes.flatten()
        fig.subplots_adjust(top=0.9)
        fig.suptitle('En un trayecto normal a su sitio de trabajo, ¿cuáles medios de_
        ↪transporte utiliza ...', fontsize= 16)
        for ax,cols,names in zip(axli,disc,list_disc):
            sns.countplot(x = cols, data = data, ax = ax)
            ax.grid(alpha = 0.2, axis = 'y')
            ax.set_title(f'{names}')
            ax.set_xticks([0,1],['No','Si'])
            ax.set_ylim(0,7000)
            ax.margins(y=0.1) # make room for the labels
            for bars in ax.containers:

```



```
ax.bar_label(bars, fmt='%0f', fontsize=9)
plt.tight_layout()
plt.subplots_adjust(top=0.94)
plt.show()
```

En un trayecto normal a su sitio de trabajo, ¿cuáles medios de transporte utiliza ...



2.11 NPCKP46B

2.11.1 46. ¿Cuánto tiempo se demora ... en su viaje de ida al trabajo (incluya tiempo de espera del medio de transporte)? Minutos

Datos: 6318

```
[227]: data['NPCKP46B'].count()
```

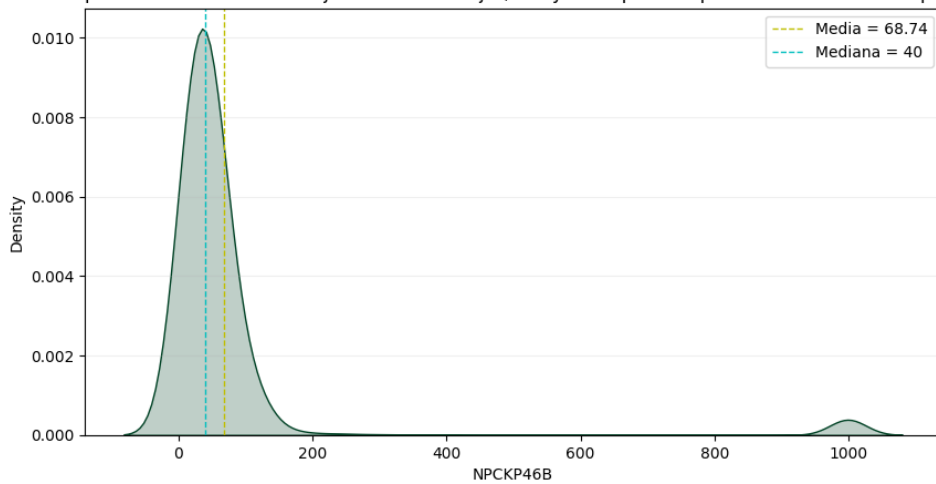
```
[227]: 6318
```

```
[228]: data['NPCKP46B'].describe().apply("{0:.2f}".format)
```

```
[228]: count      6318.00
      mean        68.74
      std        156.22
      min         0.00
      25%        20.00
      50%        40.00
      75%        60.00
      max        999.00
      Name: NPCKP46B, dtype: object
```

```
[229]: plt.figure(figsize=(10,5))
      sns.kdeplot(data = data, x = 'NPCKP46B', color = '#004225', fill = True)
      plt.title('¿Cuánto tiempo se demora ... en su viaje de ida al trabajo (incluya
      ↪ tiempo de espera del medio de transporte)? Minutos')
      plt.grid(alpha = 0.2, axis = 'y')
      plt.axvline(data['NPCKP46B'].mean(), color='y', linestyle='dashed',
      ↪ linewidth=1, label = f'Media = {data["NPCKP46B"].mean():.2f}')
      plt.axvline(data['NPCKP46B'].median(), color='c', linestyle='dashed',
      ↪ linewidth=1, label = f'Mediana = {data["NPCKP46B"].median():.0f}')
      #plt.xticks(range(0,110000,10000),[f'${(i / 1000):.0f} mil' for i in
      ↪ range(0,110000,10000)], fontsize=7)
      plt.legend()
      plt.show()
```

¿Cuánto tiempo se demora ... en su viaje de ida al trabajo (incluya tiempo de espera del medio de transporte)? Minutos



2.12 NPCKPA46

2.12.1 47. El lugar donde trabaja ...está ubicado en:

1. Este municipio
2. Otro municipio

Datos: 6318

```
[230]: list_names = ['Este Municipio', 'Otro Municipio']
```

```
[231]: data['NPCKPA46'].count()
```

```
[231]: 6318
```

```
[232]: data['NPCKPA46'].value_counts()
```

```
[232]: 1.0    5816
       2.0    502
       Name: NPCKPA46, dtype: int64
```

```
[233]: data['NPCKPA46'].describe().apply("{0:.2f}".format)
```

```
[233]: count    6318.00
       mean      1.08
       std      0.27
       min      1.00
       25%      1.00
       50%      1.00
       75%      1.00
       max      2.00
```

Name: NPCKPA46, dtype: object

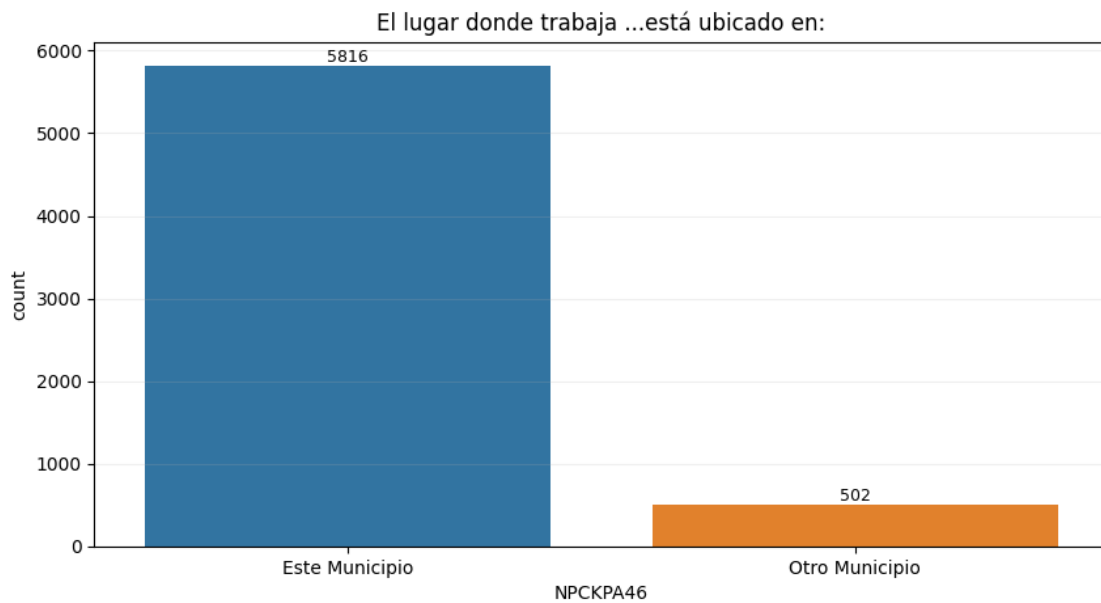
```
[234]: fig, ax = plt.subplots(figsize=(10, 5))

g = sns.countplot(ax=ax, data = data, x = 'NPCKPA46')

for bars in ax.containers:
    ax.bar_label(bars, fmt='%.0f', fontsize=9)

ax.set_title('El lugar donde trabaja ...está ubicado en:')
ax.set_xticklabels(list_names)

plt.grid(alpha = 0.2, axis = 'y')
plt.show()
```



2.13 NPCKP46AB

2.13.1 Otro municipio - Departamento

Datos:

2.14 NPCKP46AC

2.14.1 Otro municipio - Municipio

Datos:

2.15 NPCKNP48

2.15.1 48a. ¿En los ÚLTIMOS 12 MESES ha trabajado bajo la modalidad de teletrabajo o trabajo en casa?

1. Sí
2. No

Datos: 12004

```
[235]: data['NPCKNP48'].count()
```

```
[235]: 12004
```

```
[236]: data['NPCKNP48'].value_counts()
```

```
[236]: 2.0    6076
      1.0    5928
      Name: NPCKNP48, dtype: int64
```

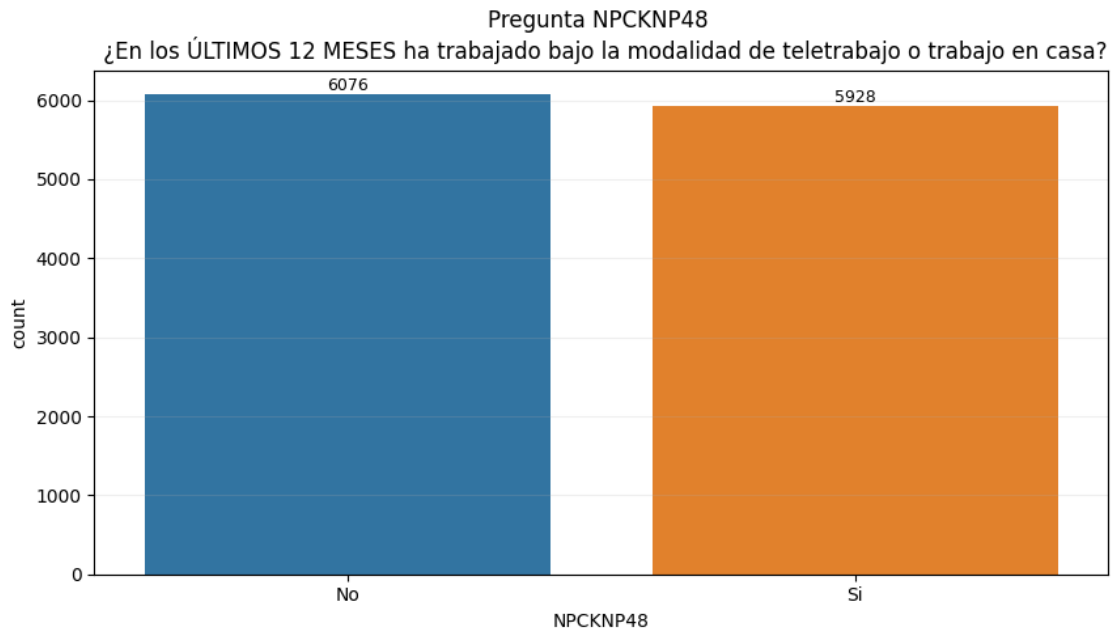
```
[237]: data = data.replace({'NPCKNP48':2},0)
```

```
[238]: fig, ax = plt.subplots(figsize=(10, 5))
      g = sns.countplot(ax=ax, data = data, x = 'NPCKNP48')

      for bars in ax.containers:
          ax.bar_label(bars, fmt='%.0f', fontsize=9)

      ax.set_title('Pregunta NPCKNP48 \n ¿En los ÚLTIMOS 12 MESES ha trabajado bajo la
      modalidad de teletrabajo o trabajo en casa?')
      ax.set_xticklabels(['No', 'Si'])

      plt.grid(alpha = 0.2, axis = 'y')
      plt.show()
```



2.16 NPCKNP48{A-D}

2.16.1 ¿Por qué motivo?

1. Por la pandemia del coronavirus (COVID-19)
2. Su empresa lo promueve permanentemente
3. Trabaja como independiente desde su casa
4. Otro

Datos: 5928

```
[239]: disc = [i for i in data.columns if 'NPCKNP48' in str(i)]
disc.remove('NPCKNP48')
disc.remove('NPCKNP48D1')
print(disc)
```

```
['NPCKNP48A', 'NPCKNP48B', 'NPCKNP48C', 'NPCKNP48D']
```

```
[240]: for i in disc:
        print(i, data[i].count())
```

```
NPCKNP48A 5928
```

```
NPCKNP48B 5928
```

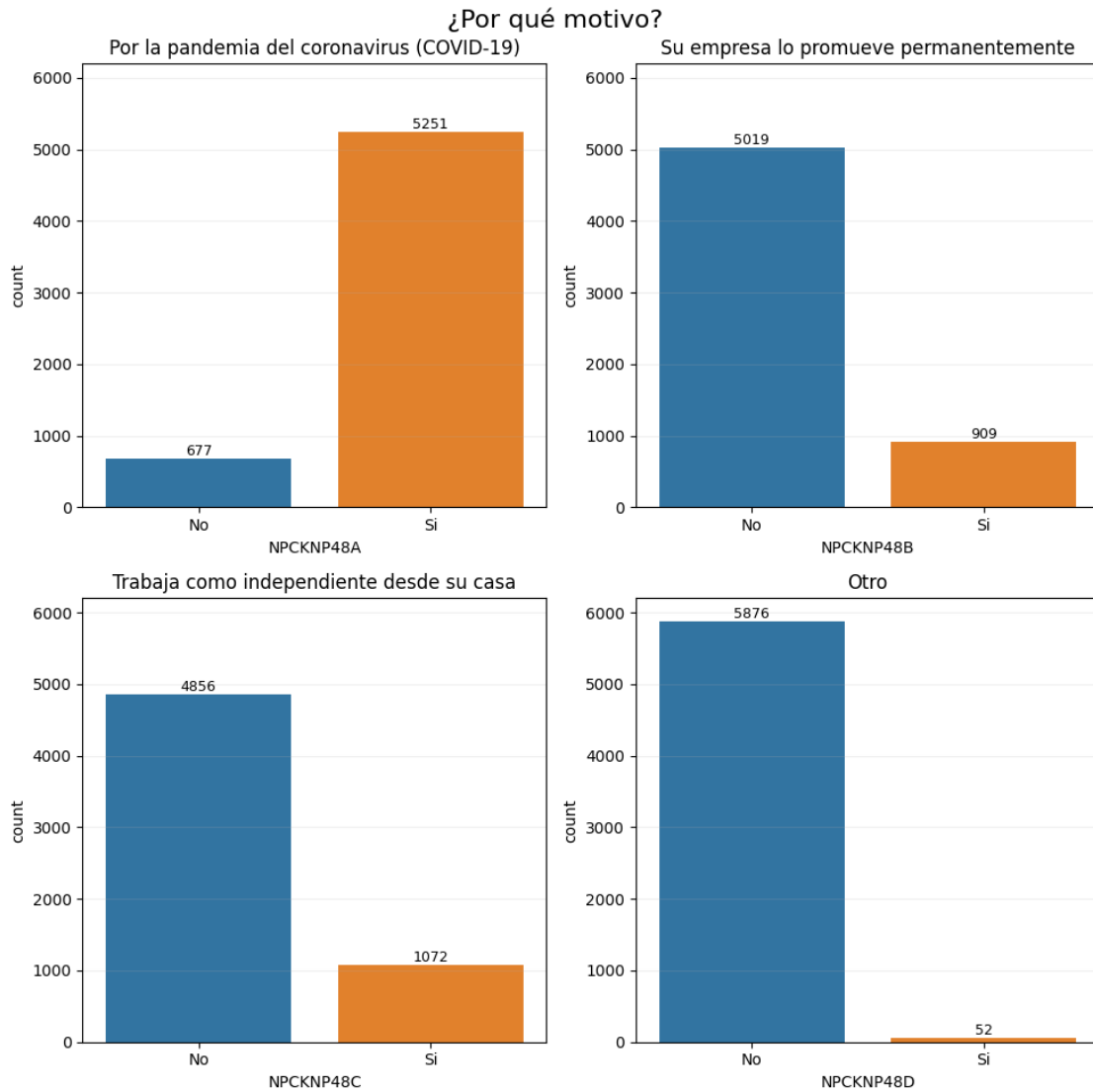
```
NPCKNP48C 5928
```

```
NPCKNP48D 5928
```

```
[241]: for i in disc:
        data = data.replace({i:2},0)
```

```
[242]: list_disc = ['Por la pandemia del coronavirus (COVID-19)',
'Su empresa lo promueve permanentemente',
'Trabaja como independiente desde su casa',
'Otro' ]
```

```
[243]: fig, axes = plt.subplots(2,2, figsize = (10,10), squeeze=False)
axli = axes.flatten()
fig.subplots_adjust(top=0.9)
fig.suptitle('¿Por qué motivo?', fontsize= 16)
for ax,cols,names in zip(axli,disc,list_disc):
    sns.countplot(x = cols, data = data, ax = ax)
    ax.grid(alpha = 0.2, axis = 'y')
    ax.set_title(f'{names}')
    ax.set_xticks([0,1],['No','Si'])
    ax.set_ylim(0,6200)
    ax.margins(y=0.1) # make room for the labels
    for bars in ax.containers:
        ax.bar_label(bars, fmt='%.0f', fontsize=9)
plt.tight_layout()
plt.subplots_adjust(top=0.93)
plt.show()
```



2.17 NPCKNP48D1

2.17.1 ¿Cuántos días a la semana teletrabaja?

Datos: 1856

```
[244]: data['NPCKNP48D1'].count()
```

```
[244]: 1856
```

```
[245]: data['NPCKNP48D1'].describe().apply("{0:.2f}".format)
```

```
[245]: count    1856.00
      mean      5.10
```



```

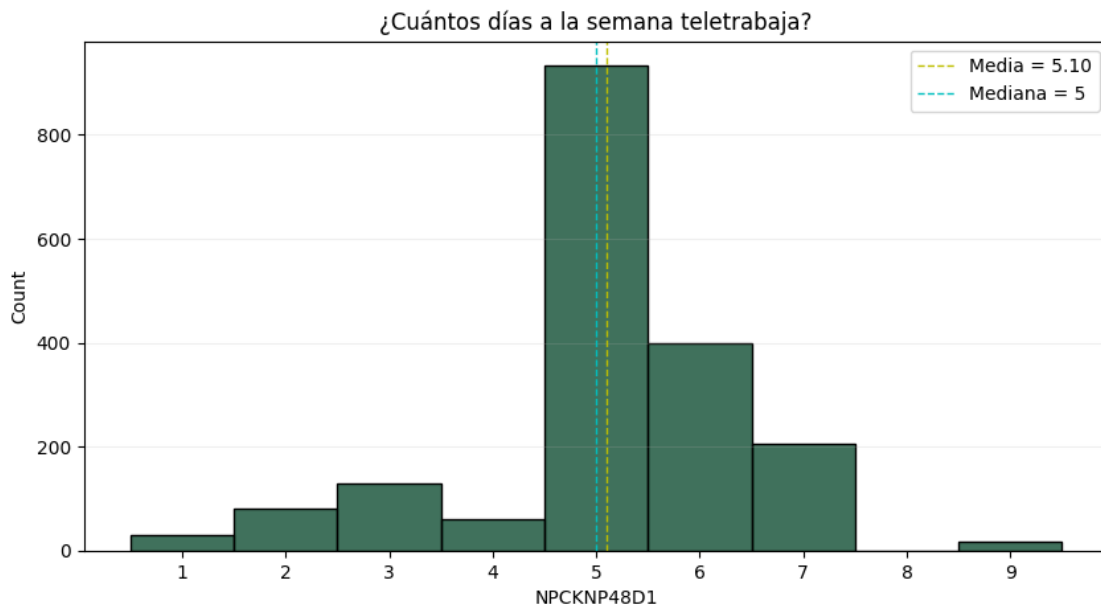
std          1.33
min          1.00
25%          5.00
50%          5.00
75%          6.00
max          9.00
Name: NPCKNP48D1, dtype: object

```

```

[246]: plt.figure(figsize=(10,5))
sns.histplot(data = data, x = 'NPCKNP48D1', color = '#004225', bins= np.
        ↳arange(1,11)-0.5)
plt.title('¿Cuántos días a la semana teletrabaja?')
plt.grid(alpha = 0.2, axis = 'y')
plt.axvline(data['NPCKNP48D1'].mean(), color='y', linestyle='dashed',
        ↳linewidth=1, label = f'Media = {data["NPCKNP48D1"].mean():.2f}')
plt.axvline(data['NPCKNP48D1'].median(), color='c', linestyle='dashed',
        ↳linewidth=1, label = f'Mediana = {data["NPCKNP48D1"].median():.0f}')
plt.xticks(range(1,10))
plt.legend()
plt.show()

```



2.18 NPCKP56

2.18.1 59. Durante los ÚLTIMOS 12 MESES, ¿... recibió algún ingreso por concepto de ayudas en dinero proveniente de otros hogares o instituciones (padres, hijos/as, familiares, amigos/as)?

1. Sí
2. No

Datos: 22213

```
[247]: data['NPCKP56'].count()
```

```
[247]: 22213
```

```
[248]: data['NPCKP56'].value_counts()
```

```
[248]: 2.0    21348
      1.0     865
      Name: NPCKP56, dtype: int64
```

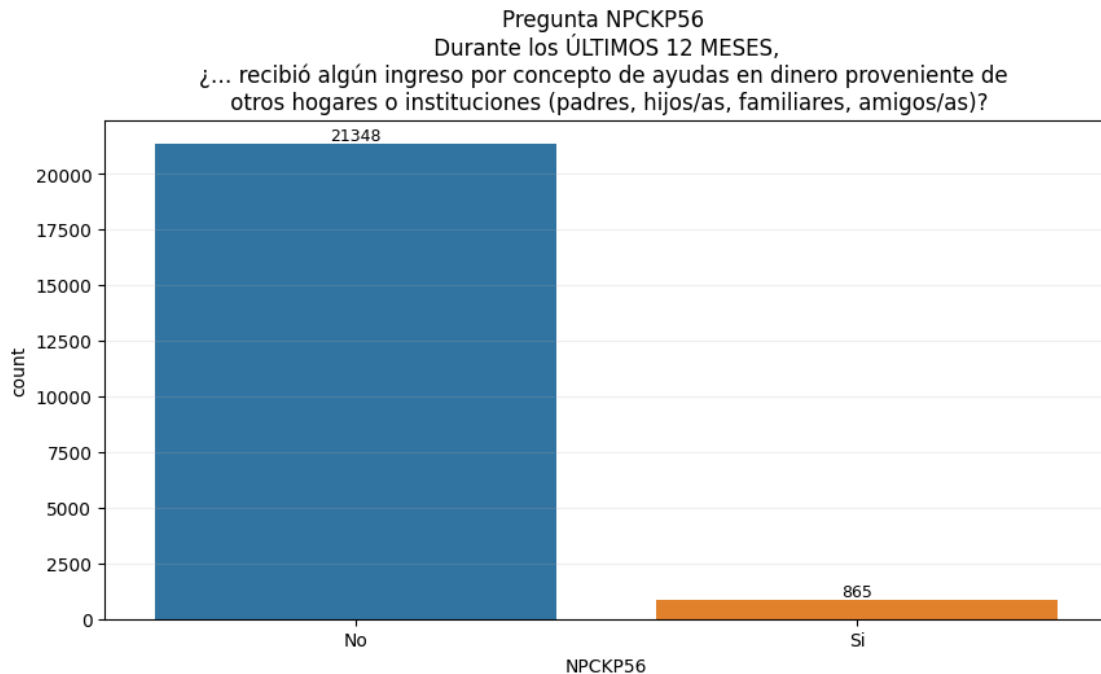
```
[249]: data = data.replace({'NPCKP56':2},0)
```

```
[250]: fig, ax = plt.subplots(figsize=(10, 5))
      g = sns.countplot(ax=ax, data = data, x = 'NPCKP56')

      for bars in ax.containers:
          ax.bar_label(bars, fmt='%.0f', fontsize=9)

      ax.set_title('Pregunta NPCKP56 \n Durante los ÚLTIMOS 12 MESES, \n ¿... recibió
      ↪ algún ingreso por concepto de ayudas en dinero proveniente de \n otros
      ↪ hogares o instituciones (padres, hijos/as, familiares, amigos/as)?')
      ax.set_xticklabels(['No', 'Si'])

      plt.grid(alpha = 0.2, axis = 'y')
      plt.show()
```



2.19 NPCKP56A

2.19.1 El dinero provino de:

1. Fuera del país
2. Dentro del país
3. Ambas partes

Datos: 865

```
[251]: list_names = ['Fuera del país', 'Dentro del país', 'Ambas partes']
```

```
[252]: data['NPCKP56A'].count()
```

```
[252]: 865
```

```
[253]: data['NPCKP56A'].value_counts()
```

```
[253]: 2.0    715
      1.0    101
      3.0     49
      Name: NPCKP56A, dtype: int64
```

```
[254]: data['NPCKPA46'].describe().apply("{0:.2f}".format)
```

```
[254]: count    6318.00
      mean      1.08
      std       0.27
      min       1.00
      25%       1.00
      50%       1.00
      75%       1.00
      max       2.00
      Name: NPCKPA46, dtype: object
```

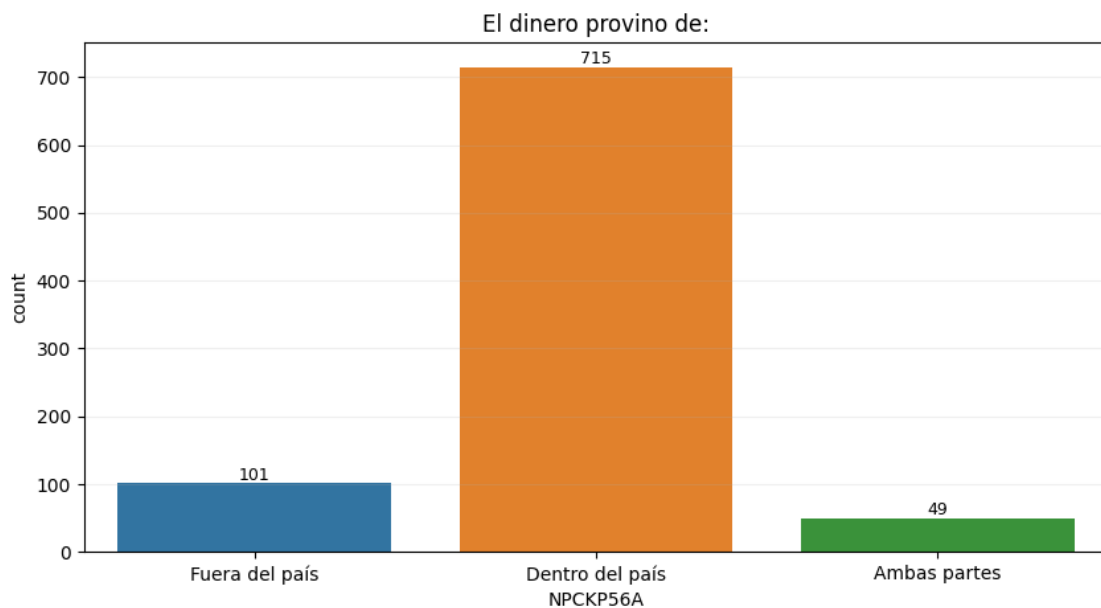
```
[255]: fig, ax = plt.subplots(figsize=(10, 5))

g = sns.countplot(ax=ax, data = data, x = 'NPCKP56A')

for bars in ax.containers:
    ax.bar_label(bars, fmt='%.0f', fontsize=9)

ax.set_title('El dinero provino de:')
ax.set_xticklabels(list_names)

plt.grid(alpha = 0.2, axis = 'y')
plt.show()
```



2.20 NPCKP56B

2.20.1 Valor recibido en los últimos 12 meses

Datos: 865

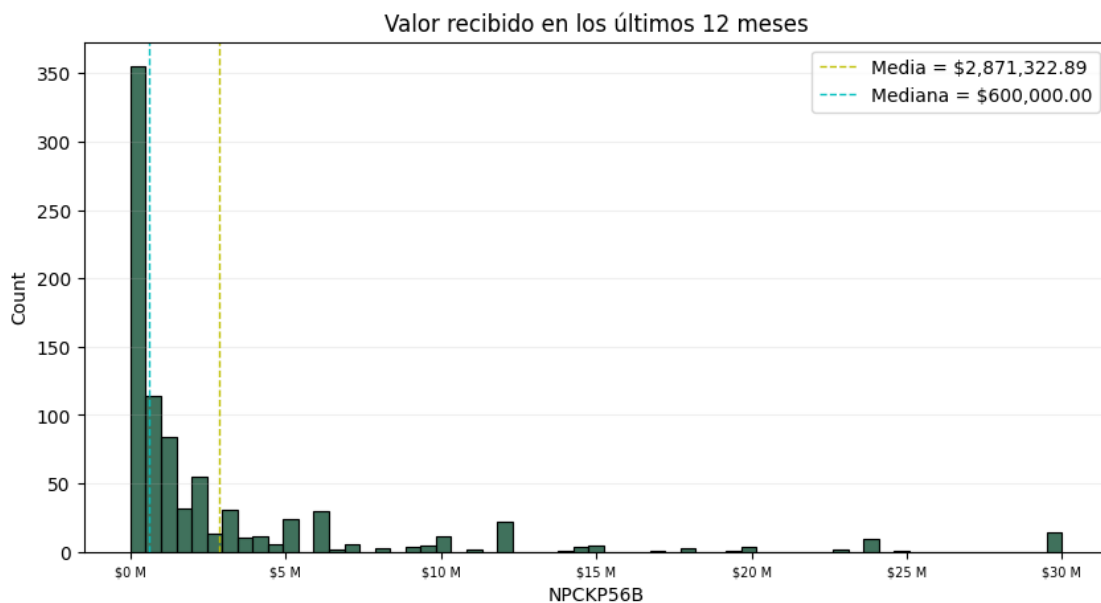
```
[256]: data['NPCKP56B'].count()
```

```
[256]: 865
```

```
[257]: data['NPCKP56B'].describe().apply("{0:.2f}".format)
```

```
[257]: count          865.00  
mean         2871322.89  
std         5541835.07  
min           98.00  
25%        150000.00  
50%        600000.00  
75%       2500000.00  
max       30000000.00  
Name: NPCKP56B, dtype: object
```

```
[258]: plt.figure(figsize=(10,5))  
sns.histplot(data = data, x = 'NPCKP56B', color = '#004225', fill = True)  
plt.title('Valor recibido en los últimos 12 meses')  
plt.grid(alpha = 0.2, axis = 'y')  
plt.axvline(data['NPCKP56B'].mean(), color='y', linestyle='dashed',  
            ↳linewidth=1, label = f'Media = ${data["NPCKP56B"].mean():.2f}')  
plt.axvline(data['NPCKP56B'].median(), color='c', linestyle='dashed',  
            ↳linewidth=1, label = f'Mediana = ${data["NPCKP56B"].median():.2f}')  
plt.xticks(range(0,31000000,5000000),[f'${(i / 1000000):.0f} M' for i in  
            ↳range(0,31000000,5000000)], fontsize=7)  
plt.legend()  
plt.show()
```



2.21 NPCKP57

2.21.1 60. Durante los ÚLTIMOS 12 MESES, ¿ ...recibió dinero por venta de propiedades (casas, edificios, lotes, maquinaria, vehículos, electrodomésticos, etc.)? }

1. Sí
2. No

Datos: 22213

```
[259]: data['NPCKP57'].count()
```

```
[259]: 22213
```

```
[260]: data['NPCKP57'].value_counts()
```

```
[260]: 2.0    22057
      1.0     156
      Name: NPCKP57, dtype: int64
```

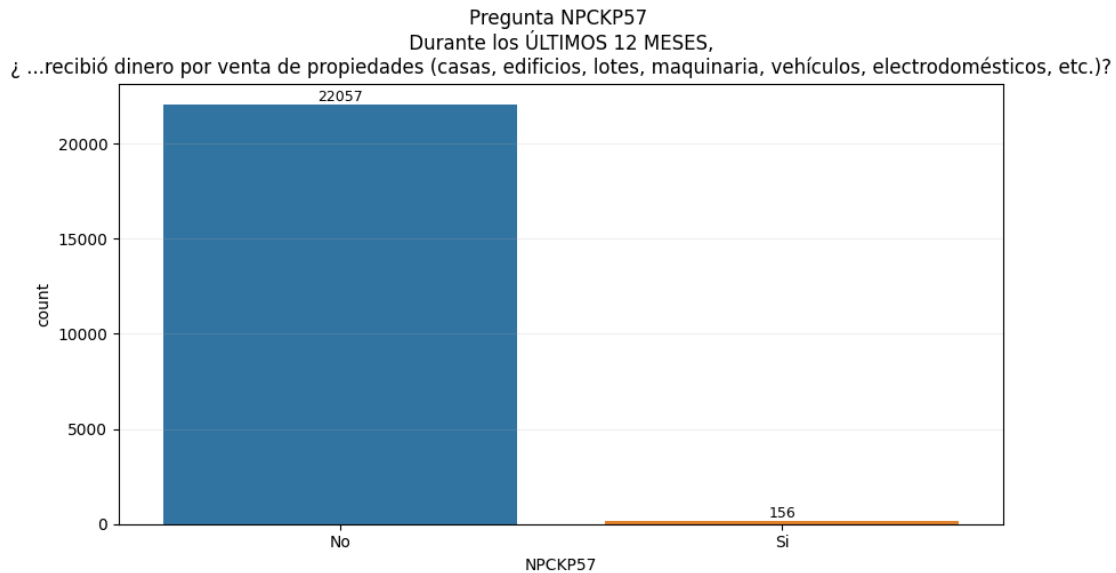
```
[261]: data = data.replace({'NPCKP57':2},0)
```

```
[262]: fig, ax = plt.subplots(figsize=(10, 5))
      g = sns.countplot(ax=ax, data = data, x ='NPCKP57')

      for bars in ax.containers:
          ax.bar_label(bars, fmt='%.0f', fontsize=9)

      ax.set_title('Pregunta NPCKP57 \n Durante los ÚLTIMOS 12 MESES, \n ¿ ...recibió
          ↪ dinero por venta de propiedades (casas, edificios, lotes, maquinaria,
          ↪ vehículos, electrodomésticos, etc.)?')
      ax.set_xticklabels(['No','Si'])

      plt.grid(alpha = 0.2, axis = 'y')
      plt.show()
```



2.22 NPCKP57A

2.22.1 Valor

Datos: 156

```
[263]: data['NPCKP57A'].count()
```

```
[263]: 156
```

```
[264]: data['NPCKP57A'].describe().apply("{0:.2f}".format)
```

```
[264]: count          156.00
      mean          46334959.99
      std           89195857.12
      min             98.00
      25%           350000.00
      50%          12000000.00
      75%           50000000.00
      max          518000000.00
      Name: NPCKP57A, dtype: object
```

```
[265]: plt.figure(figsize=(10,5))
      sns.histplot(data = data, x = 'NPCKP57A', color = '#004225', fill = True)
      plt.title('Valor recibido en los últimos 12 meses')
      plt.grid(alpha = 0.2, axis = 'y')
      plt.axvline(data['NPCKP57A'].mean(), color='y', linestyle='dashed',
      ↪ linewidth=1, label = f'Media = ${data["NPCKP57A"].mean():.2f}')
```

```
plt.axvline(data['NPCKP57A'].median(), color='c', linestyle='dashed',
            linewidth=1, label = f'Mediana = ${data["NPCKP57A"].median():.2f}')
plt.xticks(range(0,600000000,50000000),[f'${(i / 1000000):.0f} M' for i in
            range(0,600000000,50000000)], fontsize=7)
plt.legend()
plt.show()
```



2.23 NPCKP58

2.23.1 61. Durante los ÚLTIMOS 12 MESES, ¿ ...recibió dinero por otros conceptos (cesantías, intereses de cesantías, intereses por préstamos o CDT, rifas, etc.)?

1. Sí
2. No

Datos: 22213

```
[266]: data['NPCKP58'].count()
```

```
[266]: 22213
```

```
[267]: data['NPCKP58'].value_counts()
```

```
[267]: 2.0    21400
       1.0     813
       Name: NPCKP58, dtype: int64
```

```
[268]: data = data.replace({'NPCKP58':2},0)
```

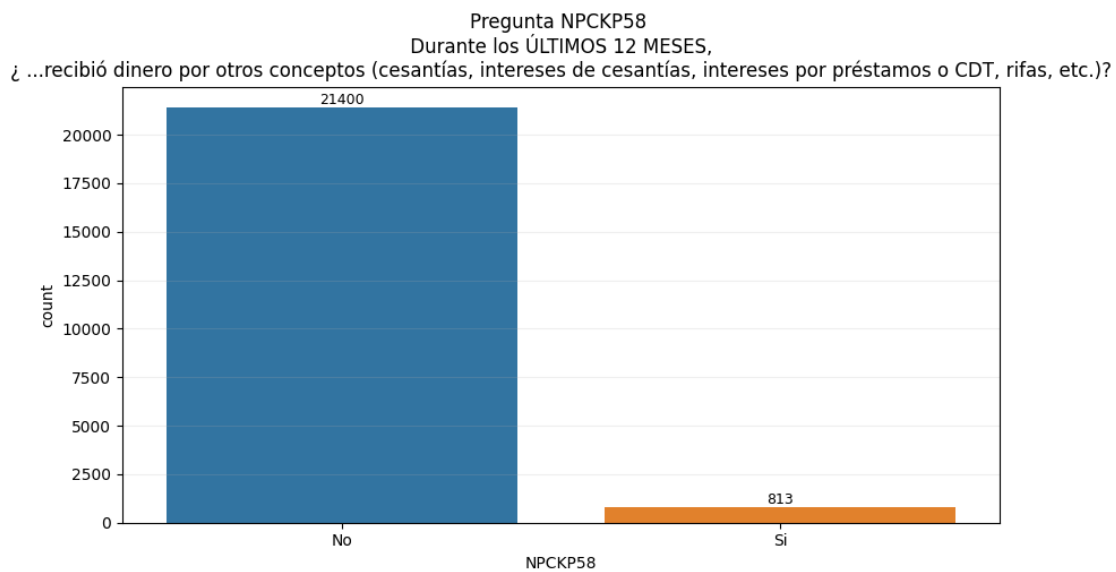


```
[269]: fig, ax = plt.subplots(figsize=(10, 5))
g = sns.countplot(ax=ax, data = data, x = 'NPCKP58')

for bars in ax.containers:
    ax.bar_label(bars, fmt='%.0f', fontsize=9)

ax.set_title('Pregunta NPCKP58 \n Durante los ÚLTIMOS 12 MESES, \n ¿ ...recibió
↳ dinero por otros conceptos (cesantías, intereses de cesantías, intereses por
↳ préstamos o CDT, rifas, etc.)?')
ax.set_xticklabels(['No', 'Si'])

plt.grid(alpha = 0.2, axis = 'y')
plt.show()
```



2.24 NPCKP58A

2.24.1 Valor

Datos:

```
[270]: data['NPCKP58A'].count()
```

```
[270]: 813
```

```
[271]: data['NPCKP58A'].describe().apply("{0:.2f}".format)
```

```
[271]: count          813.00
mean        3770258.39
std        12396268.68
```

```

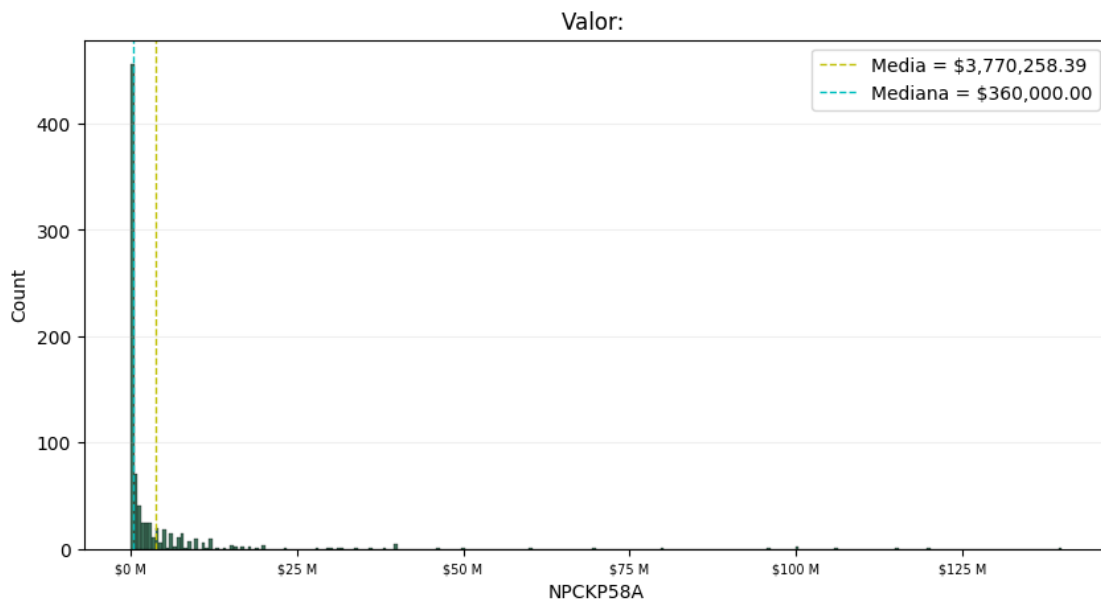
min          98.00
25%          99.00
50%        360000.00
75%        2500000.00
max        140000000.00
Name: NPCKP58A, dtype: object

```

```

[272]: plt.figure(figsize=(10,5))
sns.histplot(data = data, x = 'NPCKP58A', color = '#004225', fill = True)
plt.title('Valor: ')
plt.grid(alpha = 0.2, axis = 'y')
plt.axvline(data['NPCKP58A'].mean(), color='y', linestyle='dashed',
↳linewidth=1, label = f'Media = ${data["NPCKP58A"].mean():.2f}')
plt.axvline(data['NPCKP58A'].median(), color='c', linestyle='dashed',
↳linewidth=1, label = f'Mediana = ${data["NPCKP58A"].median():.2f}')
plt.xticks(range(0,150000000,25000000),[f'${(i / 1000000):.0f} M' for i in
↳range(0,150000000,25000000)], fontsize=7)
plt.legend()
plt.show()

```



2.25 NPCKP59{A-H}

2.25.1 63. Adicional a las actividades que me informó haber realizado la SEMANA PASADA ¿cuáles de las siguientes labores no remuneradas hizo:

1. Realizar oficios en su hogar?
2. Realizar oficios en otros hogares o instituciones?

3. Cuidar o atender niños(as)?
4. Cuidar o atender personas de la tercera edad, enfermas o en condición de Discapacidad?
5. Asistir a cursos o eventos de capacitación?
6. Trabajar en la autoconstrucción de vivienda?
7. Realizar trabajos comunitarios o voluntarios en edificaciones u obras comunitarias o públicas?
8. Participar en otras actividades cívicas, sociales, comunales o de trabajo voluntario?

Datos: 22213

```
[273]: disc1=['NPCKP59A','NPCKP59B','NPCKP59C','NPCKP59D','NPCKP59E','NPCKP59F','NPCKP59G','NPCKP59H',
disc2=['NPCKP59AA','NPCKP59BA','NPCKP59BA','NPCKP59CA',
↪'NPCKP59DA','NPCKP59EA','NPCKP59FA','NPCKP59GA','NPCKP59HA']
```

```
[274]: for i in disc1:
        print(i, data[i].count())
```

```
NPCKP59A 22213
NPCKP59B 22213
NPCKP59C 22213
NPCKP59D 22213
NPCKP59E 22213
NPCKP59F 22213
NPCKP59G 22213
NPCKP59H 22213
```

```
[275]: for i in disc1:
        data = data.replace({i:2},0)
```

```
[276]: list_disc = ['Realizar oficios en su hogar?',
'Realizar oficios en otros hogares o instituciones?',
'Cuidar o atender niños(as)?',
'Cuidar o atender personas de la tercera edad, enfermas o en condición de
↪Discapacidad?',
'Asistir a cursos o eventos de capacitación?',
'Trabajar en la autoconstrucción de vivienda?',
'Realizar trabajos comunitarios o voluntarios en edificaciones u obras
↪comunitarias o públicas?',
'Participar en otras actividades cívicas, sociales, comunales o de trabajo
↪voluntario?']
```

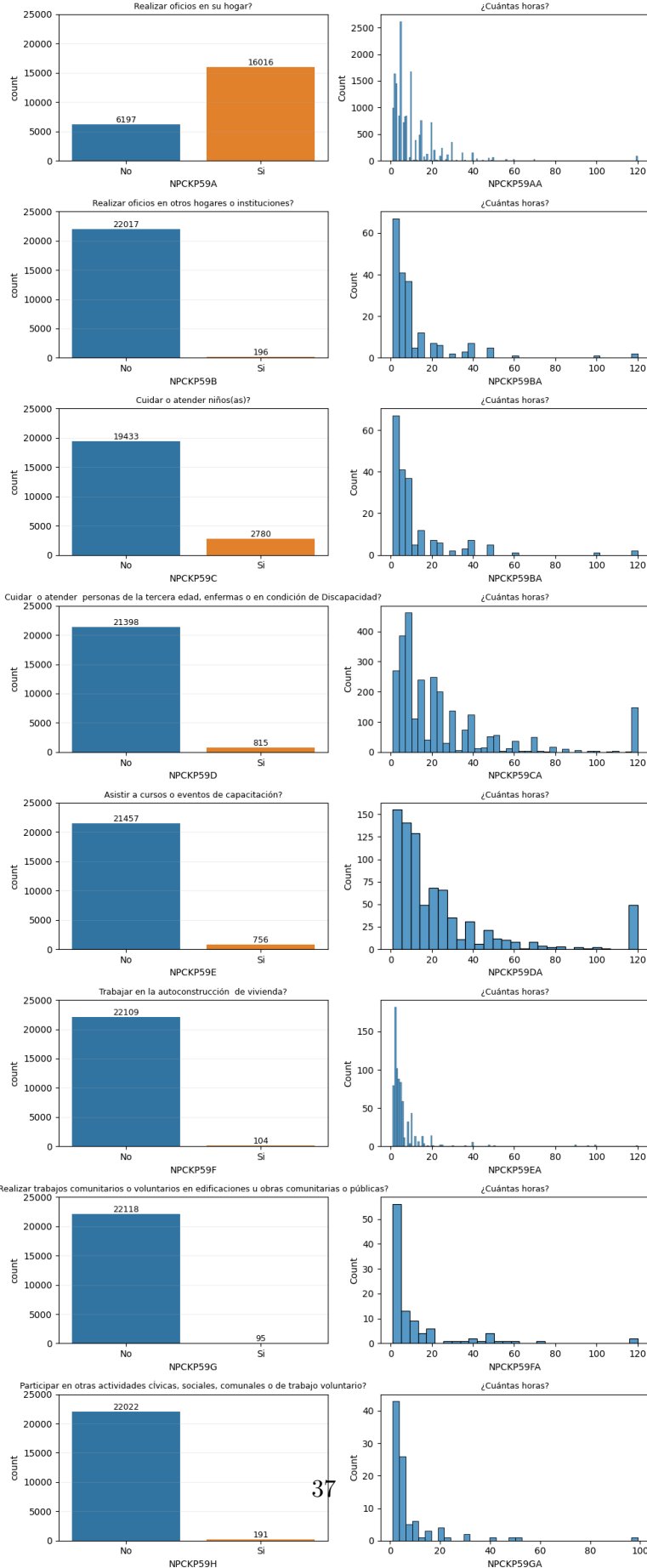
```
[277]: for col, title, hitcol in zip(disc1,list_disc,disc2):
        print(col, title, hitcol)
```

```
NPCKP59A Realizar oficios en su hogar? NPCKP59AA
NPCKP59B Realizar oficios en otros hogares o instituciones? NPCKP59BA
NPCKP59C Cuidar o atender niños(as)? NPCKP59BA
NPCKP59D Cuidar o atender personas de la tercera edad, enfermas o en condición
de Discapacidad? NPCKP59CA
NPCKP59E Asistir a cursos o eventos de capacitación? NPCKP59DA
```

NPCKP59F Trabajar en la autoconstrucción de vivienda? NPCKP59EA
NPCKP59G Realizar trabajos comunitarios o voluntarios en edificaciones u obras comunitarias o públicas? NPCKP59FA
NPCKP59H Participar en otras actividades cívicas, sociales, comunales o de trabajo voluntario? NPCKP59GA

```
[278]: fig, axes = plt.subplots(8,2, figsize = (10,25))
fig.subplots_adjust(top=0.9)
fig.suptitle('Adicional a las actividades que me informó haber realizado la_
↳SEMANA PASADA \n ¿cuáles de las siguientes labores no remuneradas hizo:',
↳fontsize= 16)
for i in range(len(axes)):
    sns.countplot(x = disc1[i], data = data, ax = axes[i][0])
    axes[i][0].grid(alpha = 0.2, axis = 'y')
    axes[i][0].set_title(f'{list_disc[i]}',fontsize=9)
    axes[i][0].set_xticks([0,1],['No','Si'])
    axes[i][0].set_ylim(0,25000)
    axes[i][0].margins(y=0.1) # make room for the labels
    for bars in axes[i][0].containers:
        axes[i][0].bar_label(bars, fmt='%.0f', fontsize=9)
    sns.histplot(data = data, x = disc2[i], ax = axes[i][1])
    axes[i][1].set_title('¿Cuántas horas?',fontsize=9)
plt.tight_layout()
plt.subplots_adjust(top=0.95)
plt.show()
```

Adicional a las actividades que me informó haber realizado la SEMANA PASADA
¿cuáles de las siguientes labores no remuneradas hizo:



2.26 NPCKP64{A-F}

2.26.1 64. ¿A raíz de las medidas tomadas por la pandemia del coronavirus (COVID-19) a ...se le presentó alguna de las siguientes situaciones:

1. Perdió su empleo
2. Tuvo que cerrar o interrumpir las actividades de su negocio
3. Tuvo que ir a licencia no remunerada
4. Tuvo que tomar vacaciones forzosas
5. Registró una pérdida sustancial de sus ingresos
6. Otra

Datos: 22213

```
[279]: disc = [i for i in data.columns if 'NPCKP64' in str(i)]
#disc.remove('NPCKNP48')
#disc.remove('NPCKNP48D1')
print(disc)
```

```
['NPCKP64A', 'NPCKP64B', 'NPCKP64C', 'NPCKP64D', 'NPCKP64E', 'NPCKP64F']
```

```
[280]: for i in disc:
        print(i, data[i].count())
```

```
NPCKP64A 22213
NPCKP64B 22213
NPCKP64C 22213
NPCKP64D 22213
NPCKP64E 22213
NPCKP64F 22213
```

```
[281]: for i in disc:
        data = data.replace({i:2},0)
```

```
[282]: list_disc = ['Perdió su empleo',
'Tuvo que cerrar o interrumpir las actividades de su negocio',
'Tuvo que ir a licencia no remunerada',
'Tuvo que tomar vacaciones forzosas',
'Registró una pérdida sustancial de sus ingresos',
'Otra']
```

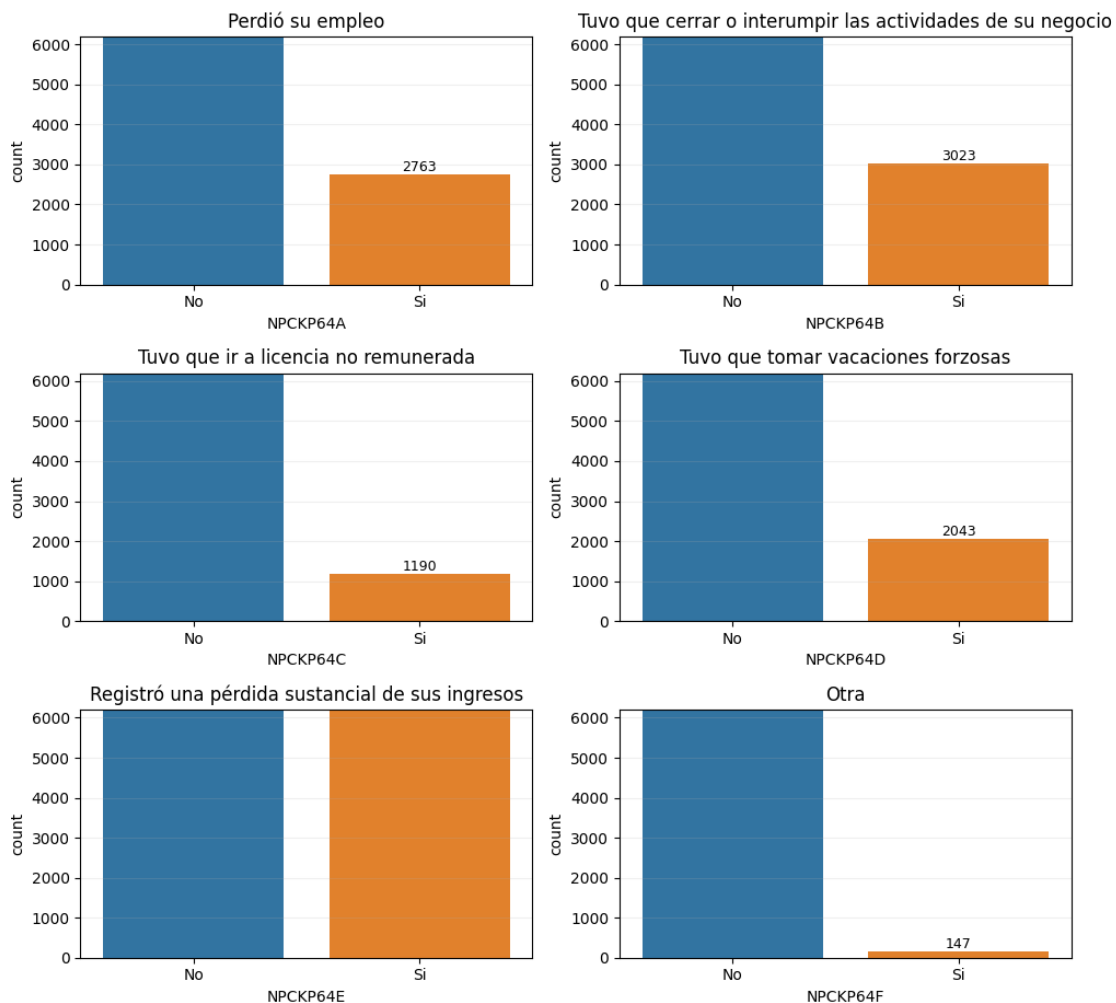
```
[283]: fig, axes = plt.subplots(3,2, figsize = (10,10), squeeze=False)
axli = axes.flatten()
fig.subplots_adjust(top=0.9)
fig.suptitle('¿A raíz de las medidas tomadas por la pandemia del coronavirus_
↳(COVID-19) a ...se le presentó alguna de las siguientes situaciones:',
↳fontsize= 16)
```

```

for ax,cols,names in zip(axli,disc,list_disc):
    sns.countplot(x = cols, data = data, ax = ax)
    ax.grid(alpha = 0.2, axis = 'y')
    ax.set_title(f'{names}')
    ax.set_xticks([0,1],['No','Si'])
    ax.set_ylim(0,6200)
    ax.margins(y=0.1) # make room for the labels
    for bars in ax.containers:
        ax.bar_label(bars, fmt='%.0f', fontsize=9)
plt.tight_layout()
plt.subplots_adjust(top=0.9)
plt.show()

```

¿A raíz de las medidas tomadas por la pandemia del coronavirus (COVID-19) a ...
se le presentó alguna de las siguientes situaciones:



```
[284]: data.to_excel('Encuesta_Multiproposito_Suba.xlsx', index = False)
```