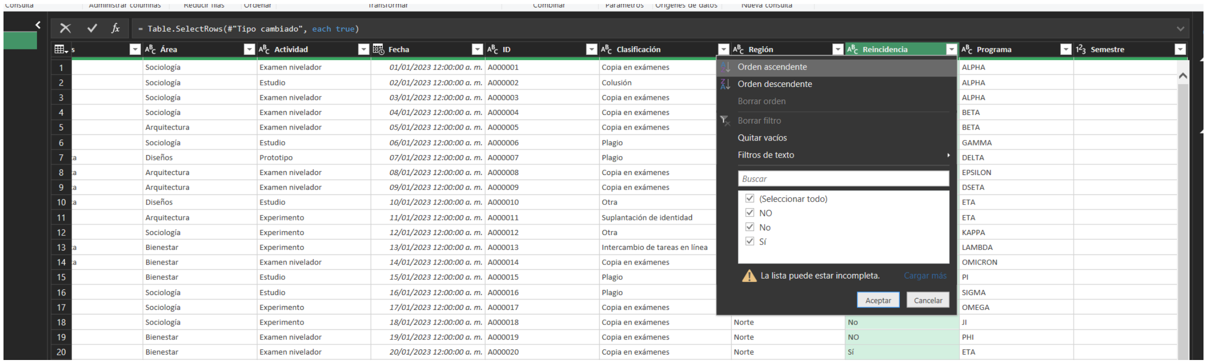
In [13]:

```
import dash
import dash_html_components as htmla
import base64
import numpy as np
#import dash_core_components as dcc
from dash import html as html
from dash.dependencies import Input, Output
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import plotly.express as px
from jupyter_dash import JupyterDash
```

1. Verify the integrity of the data in Excel Power Query

There is an anomaly in the column "Reincidencia"



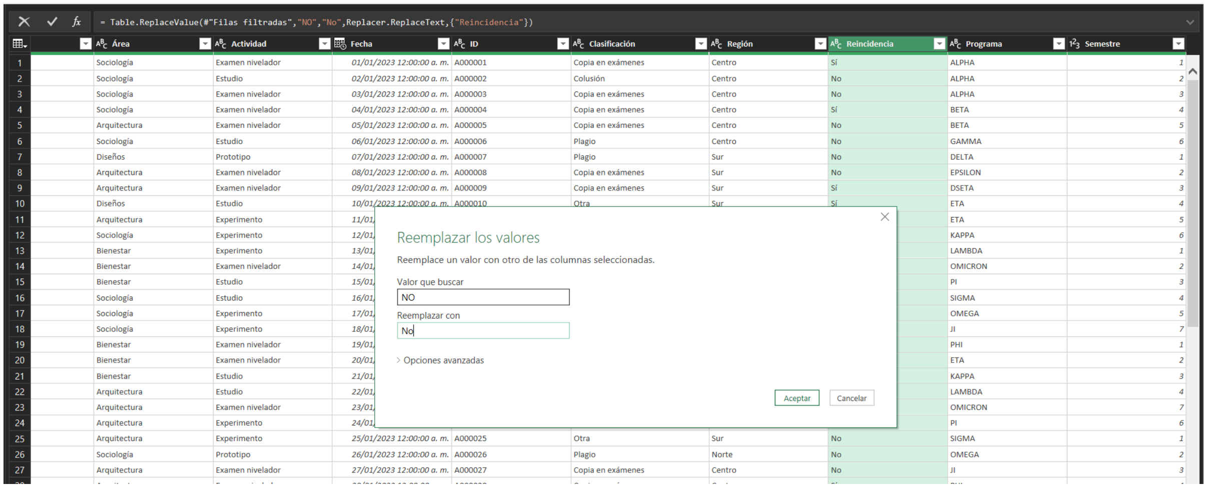
```
= Table.SelectRows(#"Tipo cambiado", each true)
```

	s	Área	Actividad	Fecha	ID	Clasificación	Región	Reincidencia	Programa	Semestre
1		Sociología	Examen nivelador	01/01/2023 12:00:00 a. m.	A000001	Copia en exámenes			ALPHA	
2		Sociología	Estudio	02/01/2023 12:00:00 a. m.	A000002	Colusión			ALPHA	
3		Sociología	Examen nivelador	03/01/2023 12:00:00 a. m.	A000003	Copia en exámenes			ALPHA	
4		Sociología	Examen nivelador	04/01/2023 12:00:00 a. m.	A000004	Copia en exámenes			BETA	
5		Arquitectura	Examen nivelador	05/01/2023 12:00:00 a. m.	A000005	Copia en exámenes			BETA	
6		Sociología	Estudio	06/01/2023 12:00:00 a. m.	A000006	Plagio			GAMMA	
7	a	Diseños	Prototipo	07/01/2023 12:00:00 a. m.	A000007	Plagio			DELTA	
8	a	Arquitectura	Examen nivelador	08/01/2023 12:00:00 a. m.	A000008	Copia en exámenes			EPSILON	
9	a	Arquitectura	Examen nivelador	09/01/2023 12:00:00 a. m.	A000009	Copia en exámenes			DSETA	
10	a	Diseños	Estudio	10/01/2023 12:00:00 a. m.	A000010	Otra			ETA	
11		Arquitectura	Experimento	11/01/2023 12:00:00 a. m.	A000011	Suplantación de identidad			ETA	
12		Sociología	Experimento	12/01/2023 12:00:00 a. m.	A000012	Otra			KAPPA	
13	a	Bienestar	Experimento	13/01/2023 12:00:00 a. m.	A000013	Intercambio de tareas en línea			LAMBDA	
14	a	Bienestar	Examen nivelador	14/01/2023 12:00:00 a. m.	A000014	Copia en exámenes			OMICRON	
15		Bienestar	Estudio	15/01/2023 12:00:00 a. m.	A000015	Plagio			PI	
16		Sociología	Estudio	16/01/2023 12:00:00 a. m.	A000016	Plagio			SIGMA	
17		Sociología	Experimento	17/01/2023 12:00:00 a. m.	A000017	Copia en exámenes			OMEGA	
18		Sociología	Experimento	18/01/2023 12:00:00 a. m.	A000018	Copia en exámenes	Norte	No	JI	
19		Bienestar	Examen nivelador	19/01/2023 12:00:00 a. m.	A000019	Copia en exámenes	Norte	NO	PHI	
20		Bienestar	Examen nivelador	20/01/2023 12:00:00 a. m.	A000020	Copia en exámenes	Norte	Sí	ETA	

☑ (Seleccionar todo)
☑ NO
☑ No
☑ Sí

La lista puede estar incompleta. Cargar más

Replaced values:



```
= Table.ReplaceValue(#"Filas filtradas","NO","No",Replacer.ReplaceText,{"Reincidencia"})
```

		Área	Actividad	Fecha	ID	Clasificación	Región	Reincidencia	Programa	Semestre
1		Sociología	Examen nivelador	01/01/2023 12:00:00 a. m.	A000001	Copia en exámenes	Centro	Sí	ALPHA	1
2		Sociología	Estudio	02/01/2023 12:00:00 a. m.	A000002	Colusión	Centro	No	ALPHA	2
3		Sociología	Examen nivelador	03/01/2023 12:00:00 a. m.	A000003	Copia en exámenes	Centro	No	ALPHA	3
4		Sociología	Examen nivelador	04/01/2023 12:00:00 a. m.	A000004	Copia en exámenes	Centro	Sí	BETA	4
5		Arquitectura	Examen nivelador	05/01/2023 12:00:00 a. m.	A000005	Copia en exámenes	Centro	No	BETA	5
6		Sociología	Estudio	06/01/2023 12:00:00 a. m.	A000006	Plagio	Centro	No	GAMMA	6
7		Diseños	Prototipo	07/01/2023 12:00:00 a. m.	A000007	Plagio	Sur	No	DELTA	1
8		Arquitectura	Examen nivelador	08/01/2023 12:00:00 a. m.	A000008	Copia en exámenes	Sur	No	EPSILON	2
9		Arquitectura	Examen nivelador	09/01/2023 12:00:00 a. m.	A000009	Copia en exámenes	Sur	Sí	DSETA	3
10		Diseños	Estudio	10/01/2023 12:00:00 a. m.	A000010	Otra	Sur	Sí	ETA	4
11		Arquitectura	Experimento	11/01					ETA	5
12		Sociología	Experimento	12/01					KAPPA	6
13		Bienestar	Experimento	13/01					LAMBDA	1
14		Bienestar	Examen nivelador	14/01					OMICRON	2
15		Bienestar	Estudio	15/01					PI	3
16		Sociología	Estudio	16/01					SIGMA	4
17		Sociología	Experimento	17/01					OMEGA	5
18		Sociología	Experimento	18/01					JI	7
19		Bienestar	Examen nivelador	19/01					PHI	1
20		Bienestar	Examen nivelador	20/01					ETA	2
21		Bienestar	Estudio	21/01					KAPPA	3
22		Arquitectura	Estudio	22/01					LAMBDA	4
23		Arquitectura	Examen nivelador	23/01					OMICRON	7
24		Arquitectura	Experimento	24/01					PI	6
25		Arquitectura	Experimento	25/01/2023 12:00:00 a. m.	A000025	Otra	Sur	No	SIGMA	1
26		Sociología	Prototipo	26/01/2023 12:00:00 a. m.	A000026	Plagio	Norte	No	OMEGA	2
27		Arquitectura	Examen nivelador	27/01/2023 12:00:00 a. m.	A000027	Copia en exámenes	Centro	No	JI	3

Reemplazar los valores

Reemplace un valor con otro de las columnas seleccionadas.

Valor que buscar: NO

Reemplazar con: No

2. Analize the data

This dataset shows reports from the Ethics Management Department

In [2]:

```
df = pd.read_csv('datasets/Assessment_Data_Analyst_Cleaned.csv')
df.head()
```

Out[2]:

	Folio	Nivel de estudios	Campus	Área	Actividad	Fecha	ID	Clasificación	Región
0	R0001	Kindergardiano	Narnia	Sociología	Examen nivelador	01/01/2023 00:00	A000001	Copia en exámenes	Centro
1	R0002	La Secu	Distrito 13	Sociología	Estudio	02/01/2023 00:00	A000002	Colusión	Centro
2	R0003	La Secu	Narnia	Sociología	Examen nivelador	03/01/2023 00:00	A000003	Copia en exámenes	Centro
3	R0004	La Secu	Narnia	Sociología	Examen nivelador	04/01/2023 00:00	A000004	Copia en exámenes	Centro
4	R0005	Big Bang Theory	Narnia	Arquitectura	Examen nivelador	05/01/2023 00:00	A000005	Copia en exámenes	Centro

Delete null values

In [3]:

```
df.dropna()
```

Out[3]:

	Folio	Nivel de estudios	Campus	Área	Actividad	Fecha	ID	Clasificación
0	R0001	Kindergardiano	Narnia	Sociología	Examen nivelador	01/01/2023 00:00	A000001	Copia en exámenes
1	R0002	La Secu	Distrito 13	Sociología	Estudio	02/01/2023 00:00	A000002	Colusión
2	R0003	La Secu	Narnia	Sociología	Examen nivelador	03/01/2023 00:00	A000003	Copia en exámenes
3	R0004	La Secu	Narnia	Sociología	Examen nivelador	04/01/2023 00:00	A000004	Copia en exámenes
4	R0005	Big Bang Theory	Narnia	Arquitectura	Examen nivelador	05/01/2023 00:00	A000005	Copia en exámenes
...	...	...	...	...	...	...	...	...
1615	R1616	High School Musical	Pizza Planeta	Sociología	Experimento	23/09/2023 00:00	A000266	Copia en exámenes
1616	R1617	High School Musical	Pizza Planeta	Sociología	Experimento	24/09/2023 00:00	A000267	Copia en exámenes
1617	R1618	High School Musical	Pizza Planeta	Sociología	Experimento	25/09/2023 00:00	A000268	Copia en exámenes
1618	R1619	High School Musical	Pizza Planeta	Sociología	Experimento	26/09/2023 00:00	A000269	Copia en exámenes
1619	R1620	High School Musical	Pizza Planeta	Sociología	Prototipo	27/09/2023 00:00	A000270	Plagio

1620 rows × 12 columns

Drop not relevant columns

In [4]:

```
df.drop(['Folio'],axis = 1, inplace=True)
```

In [5]:

```
df['Campus'].value_counts()

#daa=['Wonderland','Genovia','Springfield','Distrito 13','Pizza Planeta','Pan
 #'El País de Nunca Jamás','Narnia','Muy, muy lejano','Hogwarts','Parque Jurá
 #'Ciudad Gótica','Rarotonga','Empire']
```

Out[5]:

```
Wonderland              354
Genovia                 288
Springfield             156
Distrito 13             138
Pizza Planeta           132
Pandora                 108
El País de Nunca Jamás   84
Narnia                   78
Muy, muy lejano          78
Hogwarts                 72
Parque Jurásico          48
Ciudad Gótica            42
Rarotonga                36
Empire                    6
Name: Campus, dtype: int64
```

Add a query for drop cases without Recidivism

In [6]:

```
options = ['Sí']
# selecting rows based on condition
df_s = df.loc[df['Reincidencia'].isin(options)]
df_s.head()
```

Out[6]:

	Nivel de estudios	Campus	Área	Actividad	Fecha	ID	Clasificación	Región	Re
0	Kindergardiano	Narnia	Sociología	Examen nivelador	01/01/2023 00:00	A000001	Copia en exámenes	Centro	
3	La Secu	Narnia	Sociología	Examen nivelador	04/01/2023 00:00	A000004	Copia en exámenes	Centro	
8	Kindergardiano	Pizza Planeta	Arquitectura	Examen nivelador	09/01/2023 00:00	A000009	Copia en exámenes	Sur	
9	Big Bang Theory	Pizza Planeta	Diseños	Estudio	10/01/2023 00:00	A000010	Otra	Sur	
11	La Secu	Pandora	Sociología	Experimento	12/01/2023 00:00	A000012	Otra	Norte	

Questions

What campus has more recidivsm?

What study level has more recivism?

What activity by region has more recivism?

Graph1 Recidivism by "Campus"

In [7]:

```
import plotly.express as px
da=df['Campus'].value_counts()
#df = px.data.wind()
fig = px.line_polar(da, r="Campus", theta=['Wonderland','Genovia','Springfiel
 'El País de Nunca Jamás','Narnia','Muy, muy lejano','Hogwarts','Parque Jurás
 'Ciudad Gótica','Rarotonga','Empire'], line_close=True,
             color_discrete_sequence=px.colors.sequential.Plasma_r)
fig.show()
```

Graph2 Recidivism by "Nivel de estudios"

In [8]:

```
da=df['Nivel de estudios'].value_counts()
fig2 = px.bar(da, y='Nivel de estudios')
fig2.show()
```

Create new dataframe to filter data

In [9]:

```
index = ['Sur', 'Centro', 'Norte']
dba = df.query('Región == "Sur"')
dbb = df.query('Región == "Centro"')
dbc = df.query('Región == "Norte"')
dba = dba['Actividad'].value_counts()
dbb = dbb['Actividad'].value_counts()
dbc = dbc['Actividad'].value_counts()
dbd = [dba,dbb,dbc]
ud = np.array(dbd)
dbe = pd.DataFrame(ud,index=index)
dbe.columns = ['Experimento','Examen regulador','Prototipo','Examen nivelador
dbe
```

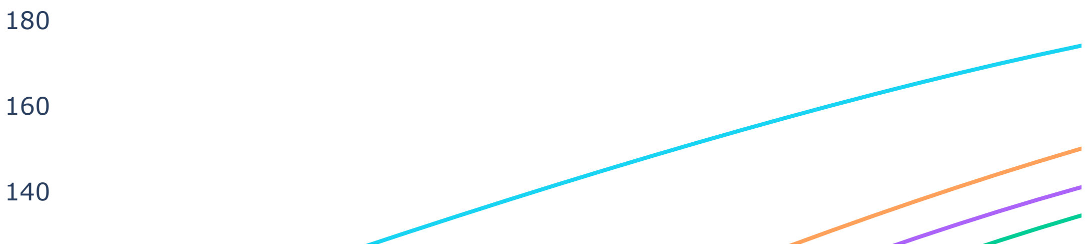
Out[9]:

	Experimento	Examen regulador	Prototipo	Examen nivelador	Examen diagnóstico	Estudio
Sur	102	54	42	30	24	24
Centro	186	168	156	150	96	48
Norte	144	138	84	72	66	36

Graph3 Recidivism by "Actividad" compared with "Región"

In [10]:

```
db=df['Actividad'].value_counts()
colors=['Estudio','Examen diagnóstico','Examen nivelador','Prototipo','Examen
#'auto', 'svg' or 'webgl'
fig3 = px.line(dbe, y=['Estudio','Examen diagnóstico','Examen nivelador','Prot
        line_shape="spline", render_mode="svg")
fig3.update_layout(
    showlegend=False,
    plot_bgcolor="white",
    #margin=dict(t=10,l=10,b=10,r=10)
)
fig3.show()
```



In [12]:

```
external_stylesheets = ['https://codepen.io/chriddyp/pen/bWLwgP.css']

app = JupyterDash(__name__, external_stylesheets=external_stylesheets)

# assume you have a "long-form" data frame
# see https://plotly.com/python/px-arguments/ for more options
df_bar = pd.DataFrame({
    "Fruit": ["Apples", "Oranges", "Bananas", "Apples", "Oranges", "Bananas"],
    "Amount": [4, 1, 2, 2, 4, 5],
    "City": ["SF", "SF", "SF", "Montreal", "Montreal", "Montreal"]
})

#fig = px.bar(df_bar, x="Fruit", y="Amount", color="City", barmode="group")

app.layout = html.Div(children=[
    # All elements from the top of the page
    html.Div([
        html.Div([
            html.H2(children='Ethics Management Department: Recidivism of AD(

            html.Div(children='''
                Recidivism by Activity
            '''),

            dcc.Graph(
                id='graph1',
                figure=fig
            ),
        ], className='six columns'),

        html.Div([
            html.Img(src=app.get_asset_url('logo.png')),

            html.Div(children='''
                Recidivism by study level.
            '''),

            dcc.Graph(
                id='graph2',
                figure=fig2
            ),
        ], className='six columns'),
    ], className='row'),
    # New Div for all elements in the new 'row' of the page
    html.Div([
        html.H1(children=' '),

        html.Div(children='''
            Recidivism by Activity compared with Región.
        '''),
        html.Div(children='''
            Blue= Estudio||
            Red= Examen diagnóstico||
            Green= Examen nivelador||
            Purple=Prototipo||
            Orange = Examen regulador||
            Light_blue = Experimento||
        '''),
        dcc.Graph(
            id='graph3',
            figure=fig3
        ),
    ], className='row'),
])

if __name__ == '__main__':
    app.run_server(debug=True,port=8052)
```

```
Dash app running on http://127.0.0.1:8052/
```

In [ ]: