

RESULTADOS DE LAS BIOPSIAS DE PRÓSTATA

Proyecto 14INV-457 "Programa de diagnóstico precoz del cáncer de próstata"

Informe técnico de avance núm. 8

El análisis de datos se realizó con Python 3.8 (Anaconda Distribution 2020-07, Anaconda, Inc., Austin, TX), usando las librerías *Pandas* y *Numpy* para el análisis estadístico, y las librerías *Matplotlib* y *Seaborn* para los gráficos. Los datos fueron analizados por el investigador principal del proyecto Prof. Dr. Alcides Chaux.

- **Fecha de cierre del periodo de análisis:** 10 de septiembre de 2019
- **Número de biopsias a la fecha:** 7 pacientes

```
In [1]: import numpy as np
import pandas as pd

from matplotlib import pyplot as plt
import seaborn as sns
sns.set_theme(rc={'figure.figsize': [12,8]})
```

```
In [2]: df = pd.read_pickle('../Datos/ITA_8.pkl')
```

```
In [3]: df.shape
```

```
Out[3]: (447, 37)
```

Biopsia de próstata

Realización de biopsia de próstata

```
In [4]: df['bx'].value_counts()
```

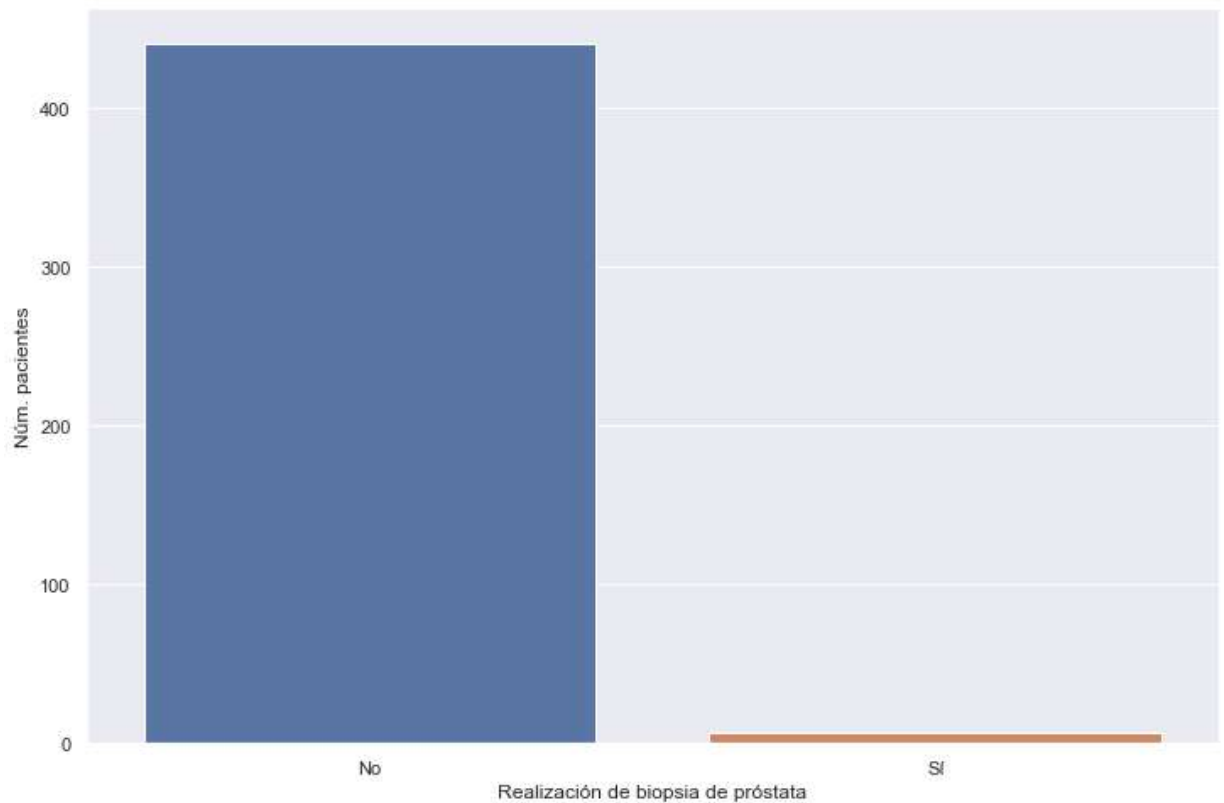
```
Out[4]: No      440
Sí         7
Name: bx, dtype: int64
```

```
In [5]: df['bx'].value_counts(normalize=True).round(2)
```

```
Out[5]: No      0.98
Sí       0.02
Name: bx, dtype: float64
```

```
In [6]: sns.countplot(x='bx', data=df)
plt.xlabel('Realización de biopsia de próstata')
```

```
plt.ylabel('Núm. pacientes')
plt.show()
```



Categoría diagnóstica

```
In [7]: df['bx_categoria'] = df['bx_categoria'].astype('category')
df['bx_categoria'].cat.reorder_categories(
    ['Proceso inflamatorio/hiperplásico', 'Proceso preneoplásico', 'Proceso neoplásico
    ordered=True,
    inplace=True
)
```

```
In [8]: df['bx_categoria'].value_counts(sort=False)
```

```
Out[8]: Proceso inflamatorio/hiperplásico    4
Proceso preneoplásico                      1
Proceso neoplásico maligno                 2
Name: bx_categoria, dtype: int64
```

```
In [9]: df['bx_categoria'].value_counts(sort=False, normalize=True).round(2)
```

```
Out[9]: Proceso inflamatorio/hiperplásico    0.57
Proceso preneoplásico                      0.14
Proceso neoplásico maligno                 0.29
Name: bx_categoria, dtype: float64
```

```
In [10]: sns.countplot(x='bx_categoria', data=df)
plt.xlabel('Categoría diagnóstica de biopsia de próstata')
plt.ylabel('Núm. pacientes')
plt.show()
```

