

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

Informe Técnico de Avance Núm. 7

RESULTADOS DE LAS BIOPSIAS PROSTÁTICAS

1 Análisis de datos

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. El reporte fue generado usando RStudio versión 1.3.1093 a través del paquete *Reticulate*. Los datos fueron analizados por el Investigador Principal del proyecto, Prof. Dr. Alcides Chaux.

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

```
library(reticulate)
use_condaenv("r-reticulate")

import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
df = pd.read_pickle('ITA_7.pkl')
```

2 Biopsia prostática

2.1 Biopsias realizadas

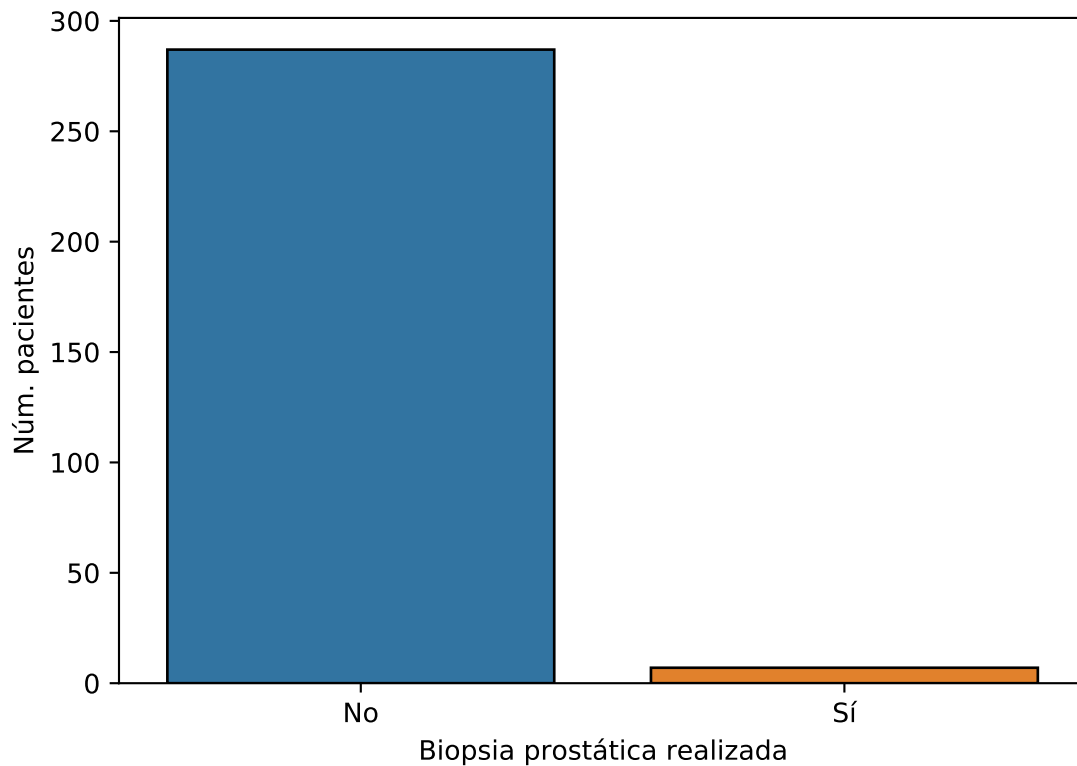
```
var = 'bx'
df[var].value_counts()

## No      287
## Sí       7
## Name: bx, dtype: int64

df[var].value_counts(normalize=True).round(2)

## No      0.98
## Sí      0.02
## Name: bx, dtype: float64
```

```
sns.countplot(x=var, data=df, edgecolor='black')
plt.xlabel('Biopsia prostática realizada')
plt.ylabel('Núm. pacientes')
plt.show()
```



2.2 Diagnóstico de las biopsias

```
df['bx_categoria'].replace({
    'Proceso inflamatorio/hiperplásico': 'Inflamatorio/hiperplásico',
    'Proceso neoplásico maligno': 'Neoplásico maligno',
    'Proceso preneoplásico': 'Preneoplásico'}, inplace=True)
df['bx_categoria'] = df['bx_categoria'].astype('category')
df['bx_categoria'].cat.reorder_categories([
    'Inflamatorio/hiperplásico',
    'Preneoplásico',
    'Neoplásico maligno'], inplace=True)

var = 'bx_categoria'
df[var].value_counts(sort=False)
```

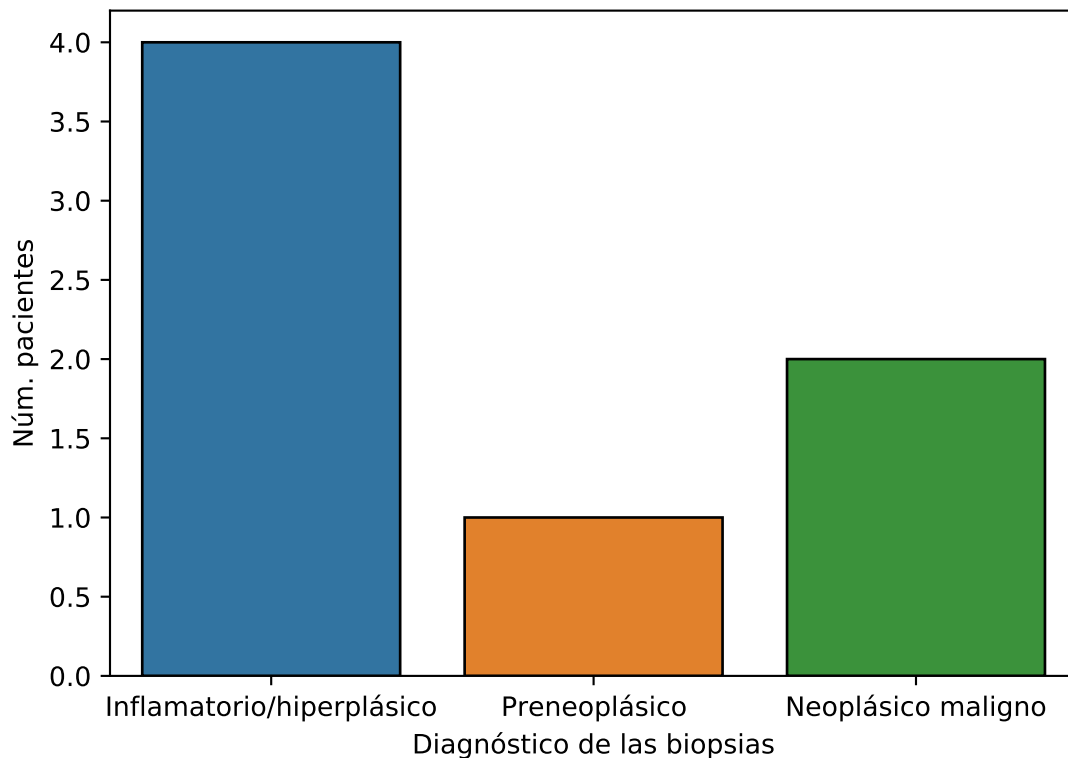
```
## Inflamatorio/hiperplásico    4
## Preneoplásico                1
## Neoplásico maligno          2
```

```
## Name: bx_categoria, dtype: int64
```

```
df[var].value_counts(sort=False, normalize=True).round(2)
```

```
## Inflamatorio/hiperplásico    0.57
## Preneoplásico                0.14
## Neoplásico maligno          0.29
## Name: bx_categoria, dtype: float64
```

```
sns.countplot(x=var, data=df, edgecolor='black')
plt.xlabel('Diagnóstico de las biopsias')
plt.ylabel('Núm. pacientes')
plt.show()
```



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

```
## ADENOCARCINOMA INFILTRANTE DE PRÓSTATA VARIEDAD ACIR                2
## HIPERPLASIA POST ATRÓFICA / PIN DE ALTO GRADO                      1
## HIPERPLASIA PROSTÁTICA / PROSTATITIS ATRÓFICA                      1
## PROSTATITIS CRÓNICA / HIPERPLASIA PROSTÁTICA BENIG                 1
## HIPERPLASIA PROSTÁTICA / PROSTATITIS CRÓNICA                       1
## HIPERPLASIA GLANDULAR Y FIBROMUSCULAR DE PRÓSTATA / PROSTATITIS CRÓNICA INESPECÍFICA 1
## Name: bx_dx, dtype: int64
```