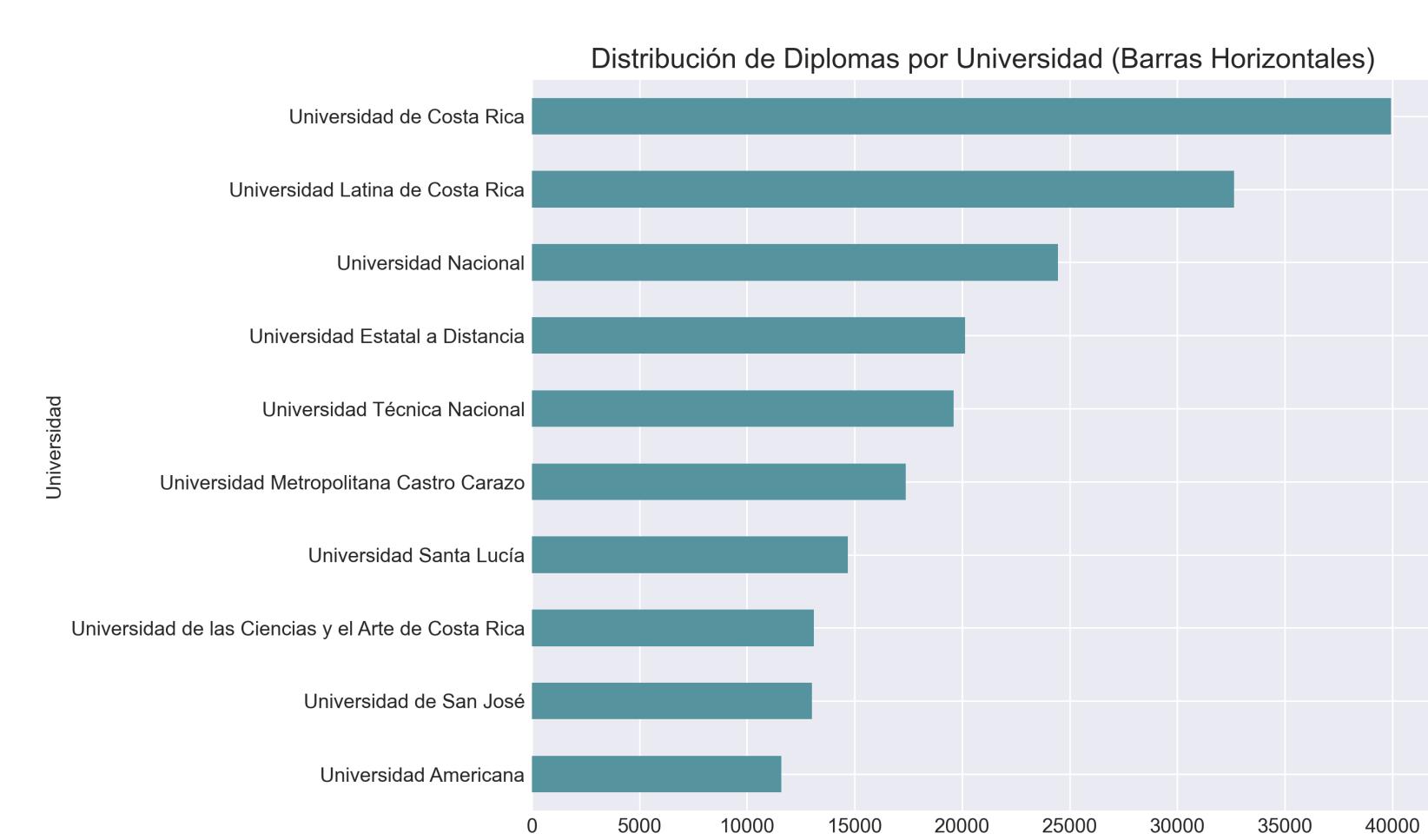


# Guía Visual de Gráficos Estadísticos en Python

Trabajo Corto #2: Gráficos para Estadística Descriptiva  
Colegio Universitario de Cartago | Kendall Solano y Roberto Coto  
Big Data - III Cuatrimestre 2025 | Prof. David Martínez Salazar

## 1. Gráfico de Barras Horizontales

Definición: Barras horizontales donde la longitud representa el valor de cada categoría.  
Cuándo usar: Series cualitativas o geográficas con nombres largos.  
Ejemplo: Top 10 universidades por diplomas emitidos (2014-2020)

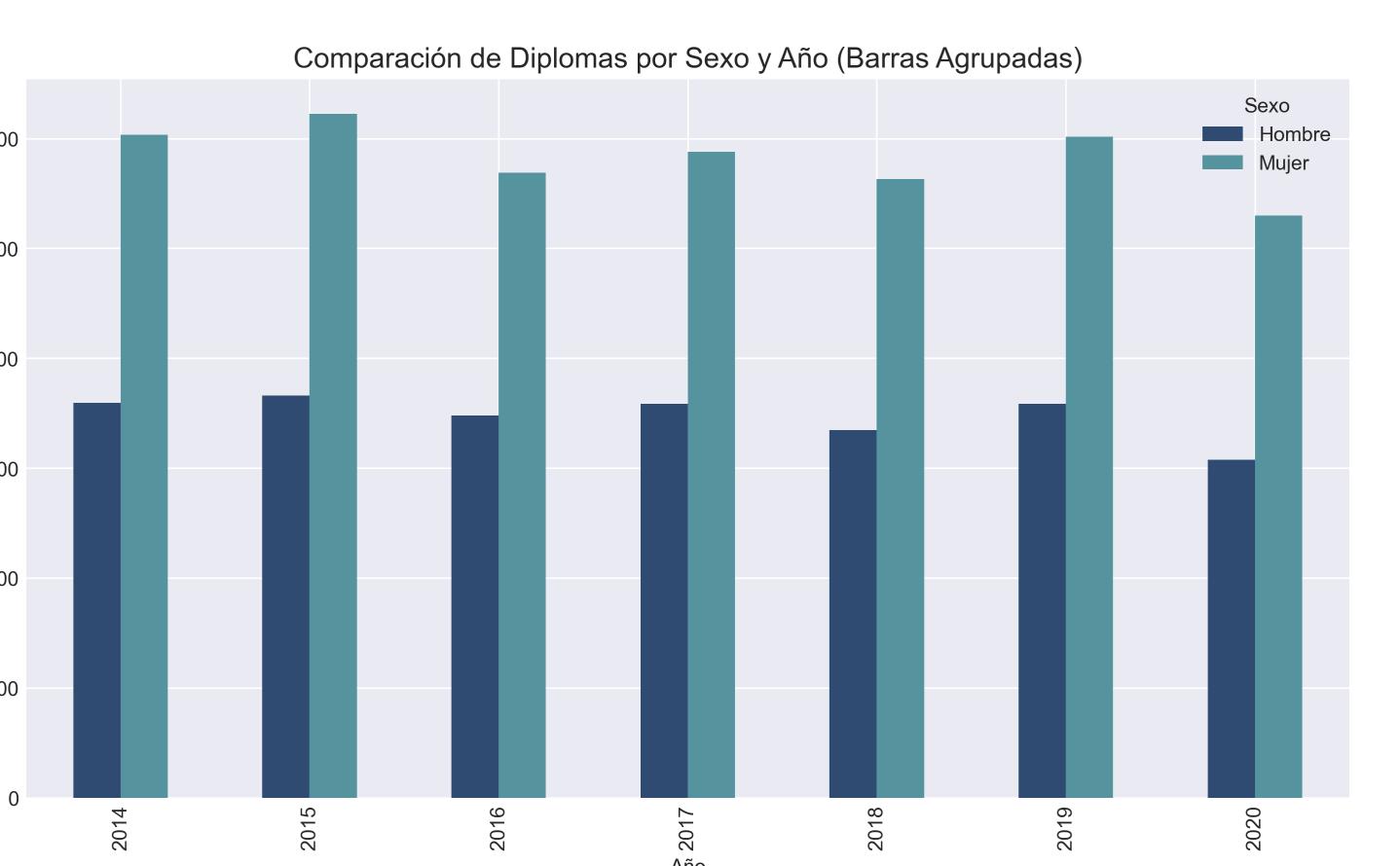


Universidad de Costa Rica lidera con >40,000 diplomas

```
diplomas = pd.read_csv('diplomas.csv')
diplomas.set_index('Universidad', inplace=True)
diplomas['CantDiplomas'] = diplomas['CantDiplomas'].apply(lambda x: int(x))
diplomas['CantDiplomas'].plot(kind='barh', color='teal')
plt.xlabel('Cantidad de Diplomas')
plt.title('Distribución de Diplomas por Universidad (Barras Horizontales)')
plt.show()
```

## 2. Gráfico de Barras Agrupadas

Definición: Barras lado a lado para comparación directa entre subcategorías.  
Cuándo usar: Comparar magnitudes absolutas entre grupos.  
Ejemplo: Comparación directa hombres vs mujeres por año

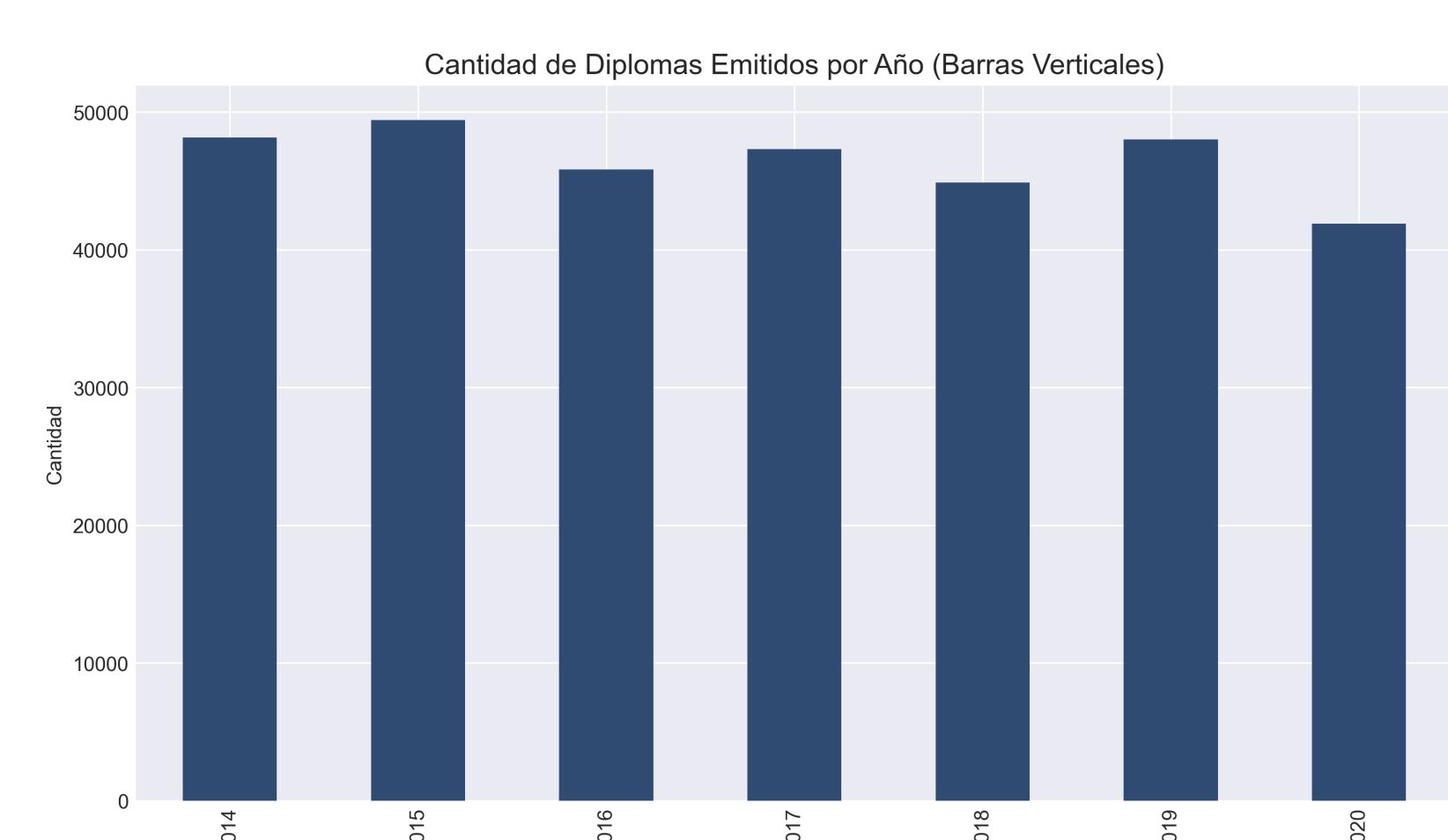


Diferencia constante de ~10,000 diplomas

```
diplomas = pd.read_csv('diplomas.csv')
diplomas.set_index('Año', inplace=True)
diplomas['CantDiplomas'].groupby('Sexo').plot(kind='bar', color=['darkblue', 'teal'])
plt.title('Comparación de Diplomas por Sexo y Año (Barras Agrupadas)')
plt.show()
```

## 2. Gráfico de Barras Verticales

Definición: Barras que crecen verticalmente desde el eje horizontal.  
Cuándo usar: Comparar categorías discretas, especialmente datos temporales.  
Ejemplo: Diplomas emitidos por año (2014-2020)

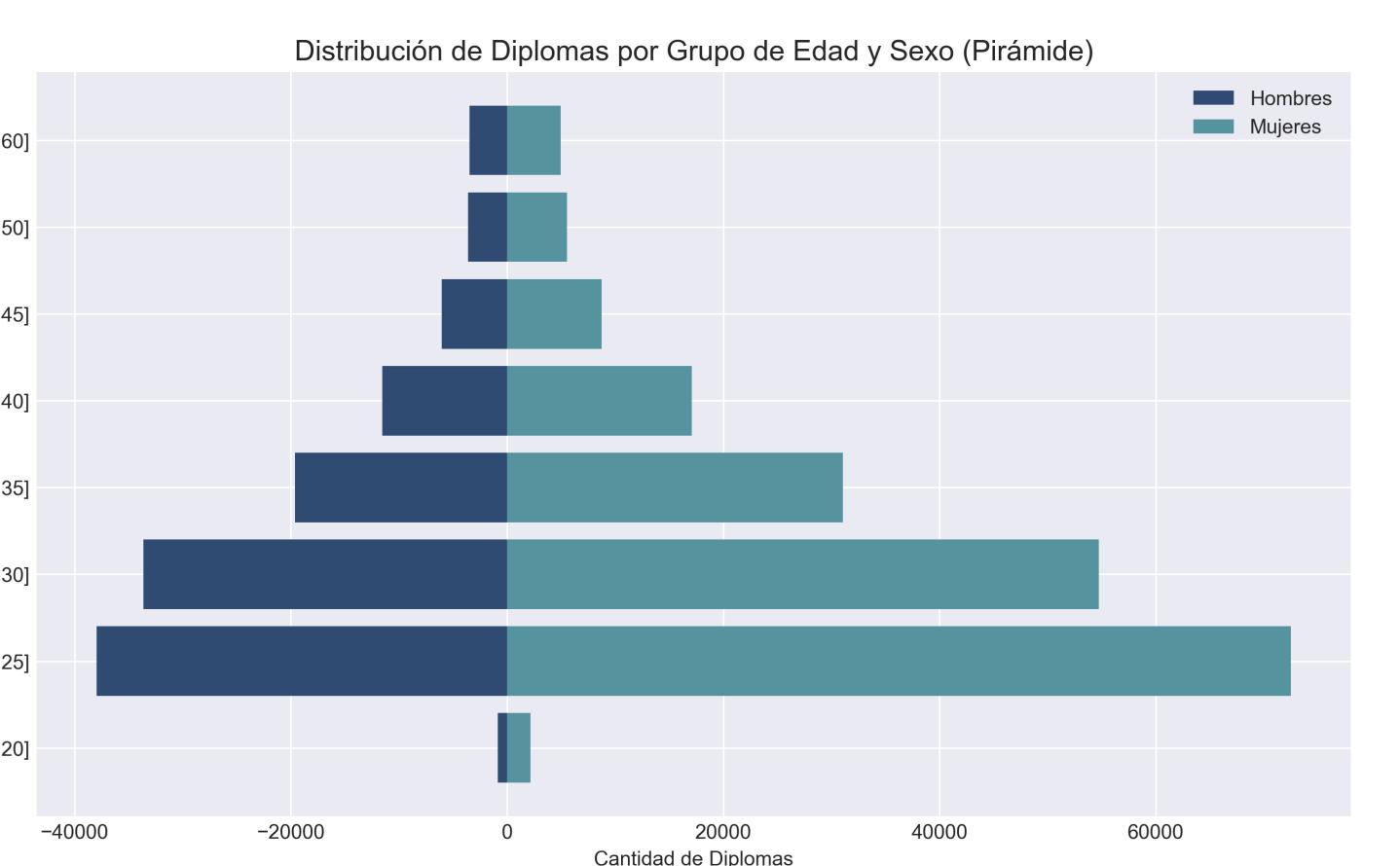


Pico en 2015 (49,500), caída en 2020 (41,900)

```
diplomas = pd.read_csv('diplomas.csv')
diplomas.set_index('Año', inplace=True)
diplomas['CantDiplomas'].plot(kind='bar', color='darkblue')
plt.title('Cantidad de Diplomas Emitidos por Año (Barras Verticales)')
plt.show()
```

## 4. Gráfico de Pirámide Poblacional

Definición: Barras horizontales bidireccionales que muestran dos grupos por intervalos.  
Cuándo usar: Análisis demográfico, estructura por edad y género.  
Ejemplo: Distribución de diplomas por grupos de edad y sexo

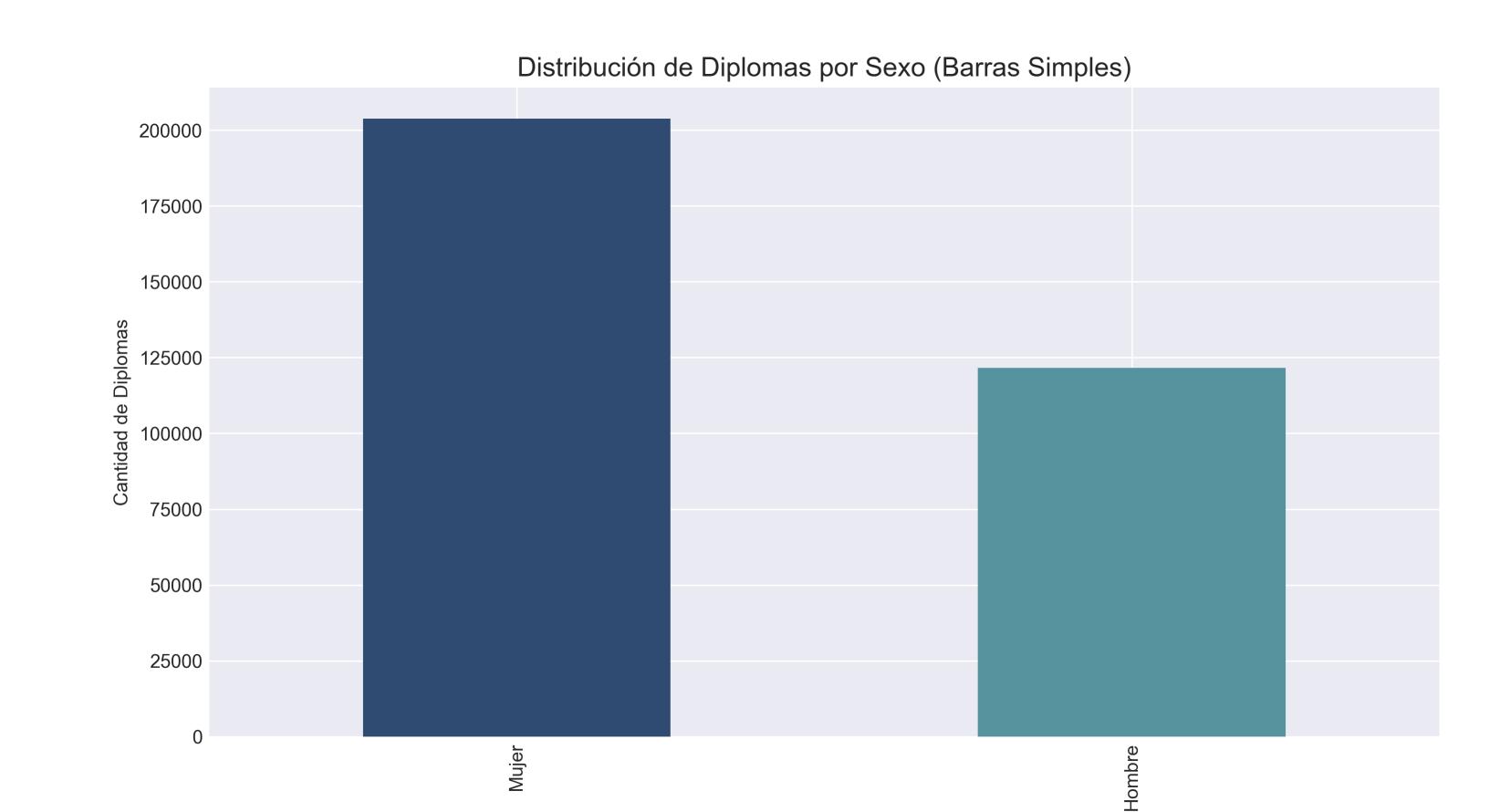


Mayor concentración: 20-30 años

```
diplomas = pd.read_csv('diplomas.csv')
diplomas.set_index('Año', inplace=True)
diplomas['CantDiplomas'].groupby(['Sexo', 'Edad']).plot(kind='bar', color=['darkblue', 'teal'])
plt.title('Distribución de Diplomas por Grupo de Edad y Sexo (Pirámide)')
plt.show()
```

## 3. Gráfico de Barras Simples

Definición: Representación básica de frecuencias para una variable cualitativa.  
Cuándo usar: Comparar categorías mutuamente excluyentes.  
Ejemplo: Distribución de diplomas por sexo

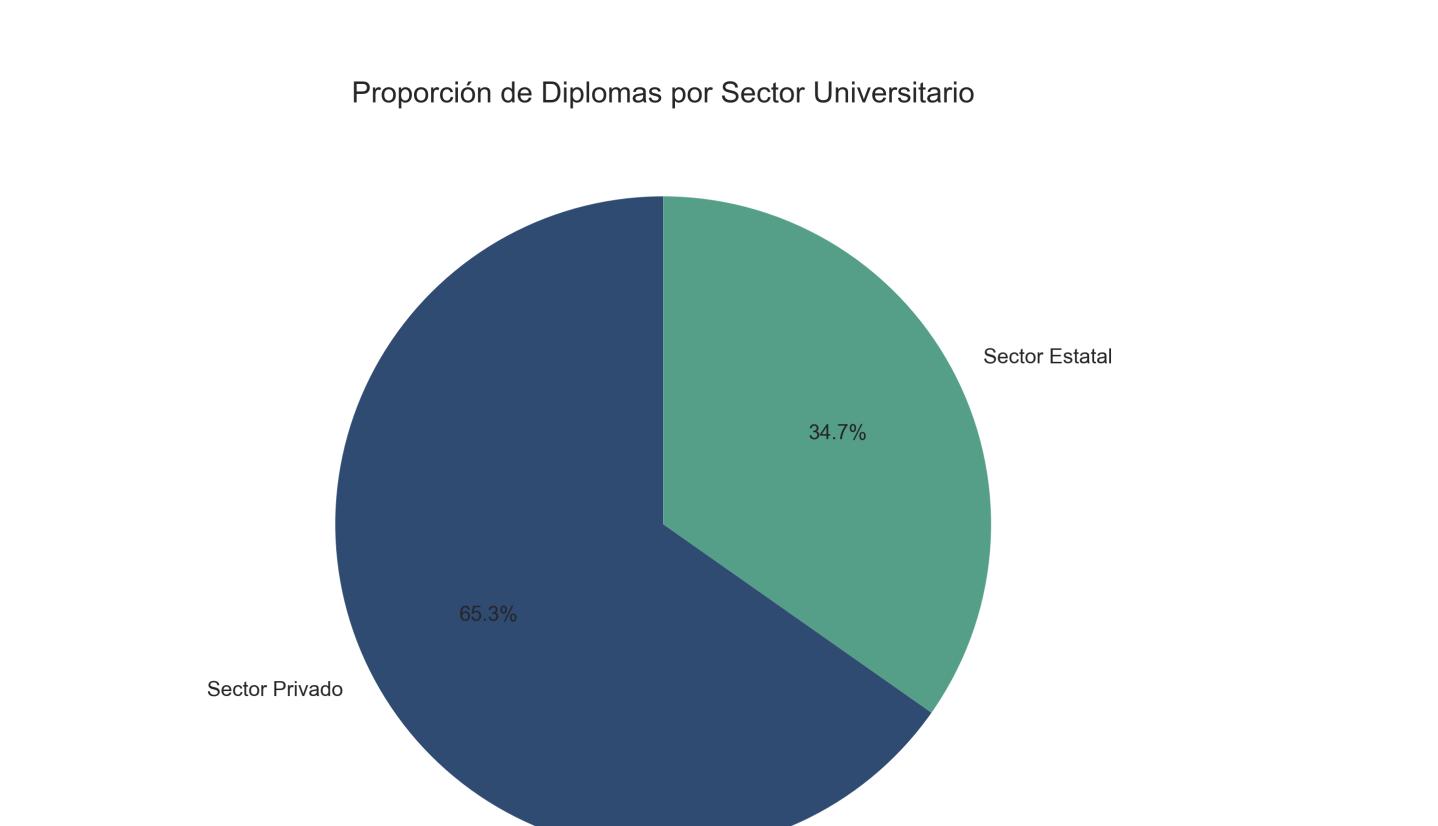


Mujeres: 62% (204,000) | Hombres: 38% (121,000)

```
diplomas = pd.read_csv('diplomas.csv')
diplomas['Sexo'].value_counts().plot(kind='bar', color='teal')
plt.title('Distribución de Diplomas por Sexo (Barras Simples)')
plt.show()
```

## 7. Gráfico de Pastel (Circular)

Definición: Circulo dividido en sectores proporcionalmente.  
Cuándo usar: Mostrar composición porcentual (2-5 categorías).  
Ejemplo: Proporción diplomas sector estatal vs privado



Privado: 65.3% | Estatal: 34.7%

```
diplomas = pd.read_csv('diplomas.csv')
diplomas['Sector'].value_counts().plot(kind='pie', autopct='%1.1f%%')
plt.title('Proporción de Diplomas por Sector Universitario')
plt.show()
```

## 4. Gráfico de Barras Apiladas

Definición: Barras divididas en segmentos apilados que muestran composición interna.  
Cuándo usar: Comparar totales y ver composición simultáneamente.  
Ejemplo: Diplomas por año diferenciando sexo (valores absolutos)

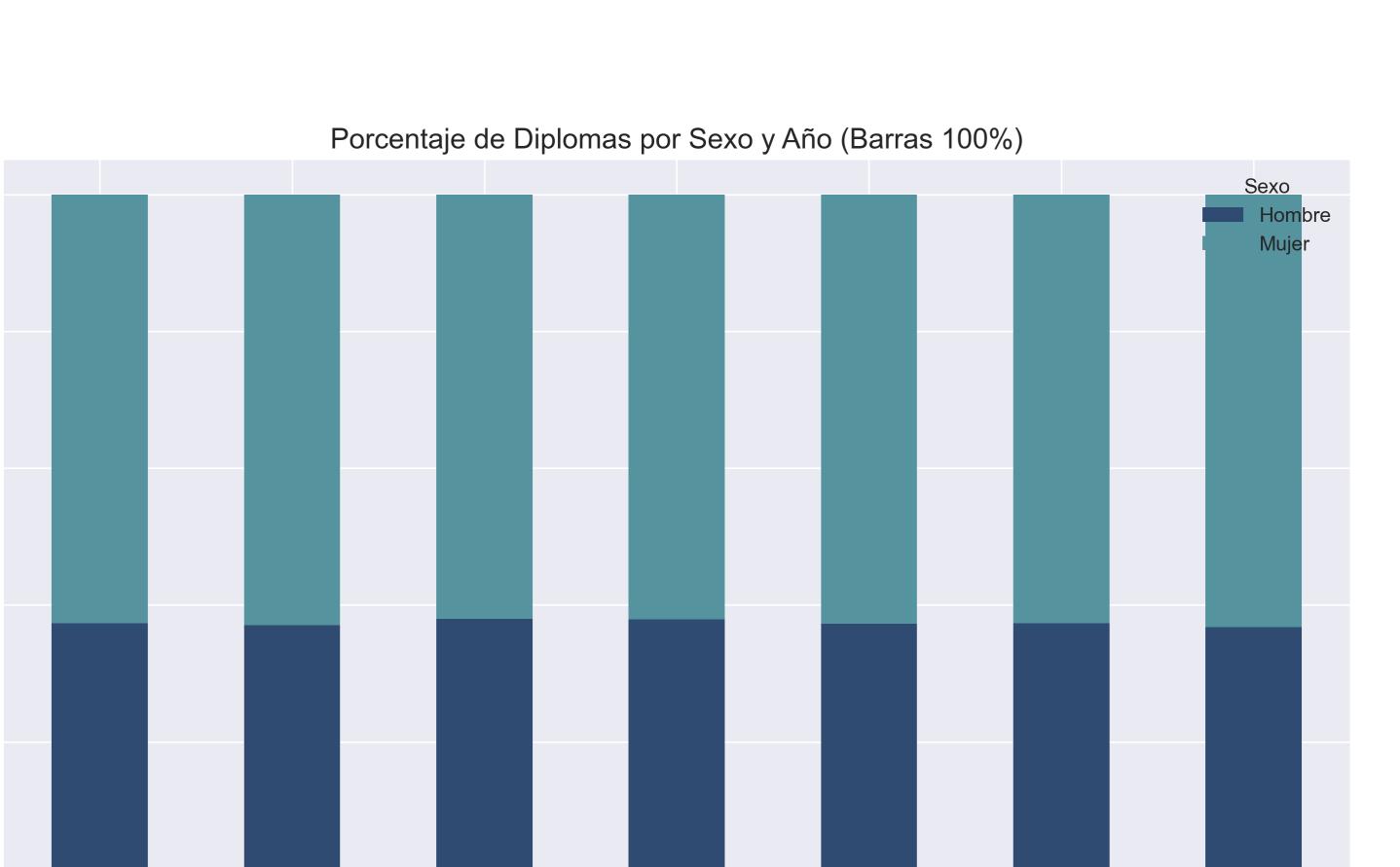


Dominancia femenina constante en todos los años

```
diplomas = pd.read_csv('diplomas.csv')
diplomas.set_index('Año', inplace=True)
diplomas['CantDiplomas'].groupby(['Sexo']).plot(kind='bar', stacked=True)
plt.title('Distribución por Sexo y Año (Barras Apiladas)')
plt.show()
```

## 8. Gráfico de Barras 100 %

Definición: Barras normalizadas al 100 % mostrando proporciones relativas.  
Cuándo usar: Comparar composiciones porcentuales.  
Ejemplo: Evolución porcentual de género por año



Proporción estable: 62% mujeres, 38% hombres

```
diplomas = pd.read_csv('diplomas.csv')
diplomas.set_index('Año', inplace=True)
diplomas['CantDiplomas'].groupby(['Sexo']).plot(kind='bar', stacked=True, normalize=True)
plt.title('Porcentaje de Diplomas por Sexo y Año (Barras 100%)')
plt.show()
```

###### 9. Gráfico Lineal

Definición: Línea continua que conecta puntos mostrando tendencias.  
Cuándo usar: Series temporales o patrones de cambio.  
Ejemplo: Tendencia de diplomas emitidos anualmente



###### 10. Gráfico de Araña (Radar)

Definición: Gráfico circular con ejes radiales.  
Cuándo usar: Comparar múltiples variables simultáneamente.  
Ejemplo: Distribución de diplomas por área de conocimiento

