

# CAPES - Average Time to Graduate

October 15, 2023

```
[3]: import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
```

```
[4]: def plot(df,title,descr,ymin,ymax,yexp):
    plt.figure(figsize=(12, 6)) # Adjust the figure size as needed
    plt.ylim(ymin, ymax)
    plt.axhline(y=yexp, color='#FF8080', linestyle='--',zorder=0)
    sns.violinplot(x='AnoBase', y='MesesParaGraduacao', data=df, palette=sns.
    →color_palette('husl', 10),zorder=1)
    # Add labels and title
    # plt.yscale('log')
    plt.xlabel('Ano')
    plt.ylabel('Meses')
    plt.title('Distribuição do Tempo para Graduação por Ano - '+title+' -
    →'+descr)
    data_counts = df['AnoBase'].value_counts()
    x_positions = range(len(data_counts))
    for x, (year, count) in zip(x_positions, data_counts.items()):
        bbox_props = dict(boxstyle='round, pad=0.3', facecolor='white', alpha=0.
    →9)
        plt.text(x, ymax*.96, f'n={count}', fontsize=10, color='#000080',
    →ha='center',bbox=bbox_props)
    plt.show()
```

```
[5]: df = pd.read_csv("Titulado.csv")
df = df[(df['AnoBase'] >= 2012) & (df['AnoBase'] <= 2021)]
#####
dfma = df[df['TitulacaoDiscente'] == "MESTRADO"]
dfma = dfma[dfma['MesesParaGraduacao'] <= 48]
dfma = dfma[dfma['MesesParaGraduacao'] >= 12]
plot(dfma,"Mestrado Acadêmico","(12 a 48 meses)",12,48,24)
#####
dfda = df[df['TitulacaoDiscente'] == "DOCTORADO"]
dfda = dfda[dfda['MesesParaGraduacao'] <= 96]
dfda = dfda[dfda['MesesParaGraduacao'] >= 12]
plot(dfda,"Doutorado Acadêmico","(12 a 96 meses)",12,96,48)
```

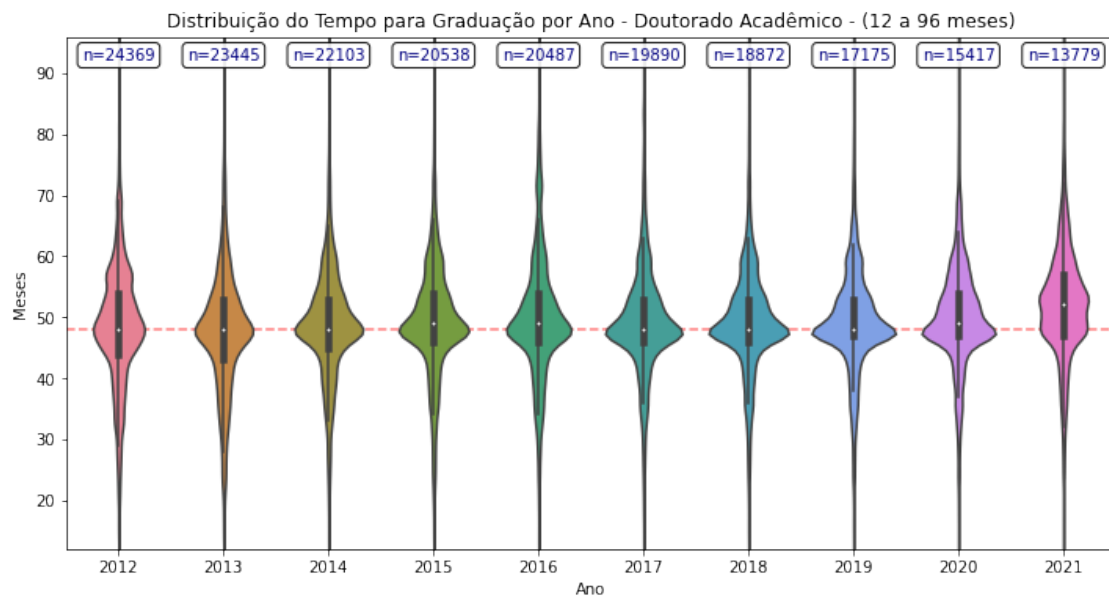
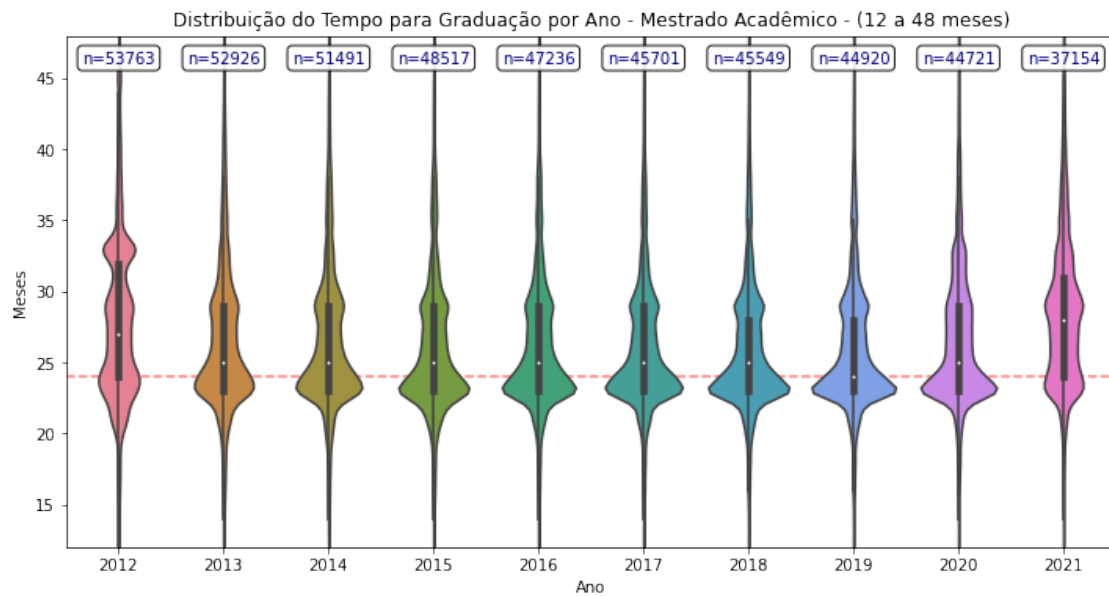
```
#####
```

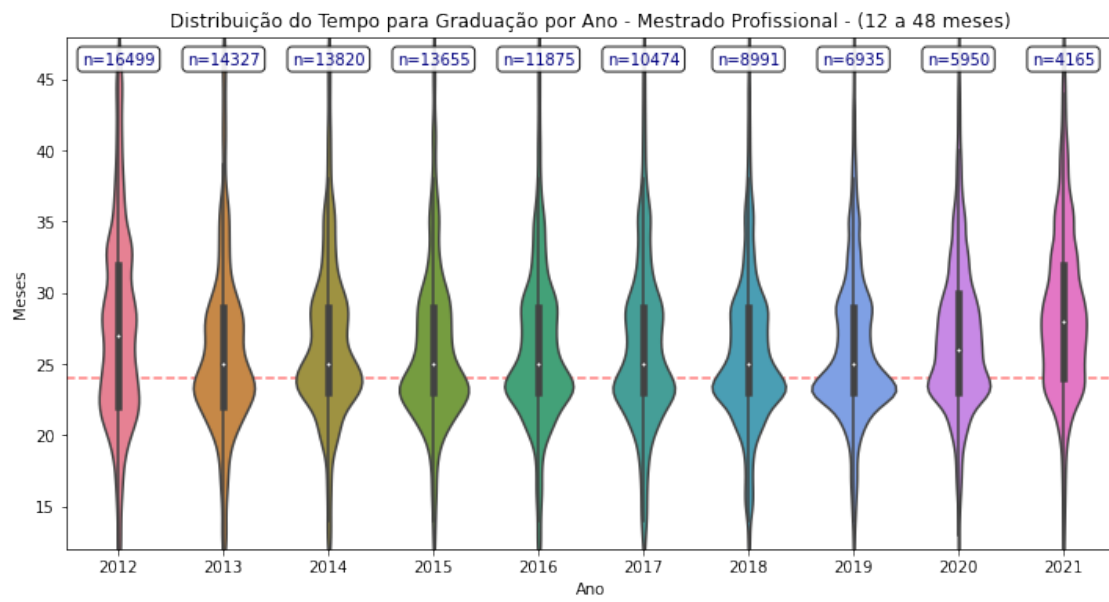
```
dfmp = df[df['TitulacaoDiscente'] == "MESTRADO PROFISSIONAL"]
```

```
dfmp = dfmp[dfmp['MesesParaGraduacao'] <= 48]
```

```
dfmp = dfmp[dfmp['MesesParaGraduacao'] >= 12]
```

```
plot(dfmp,"Mestrado Profissional","(12 a 48 meses)",12,48,24)
```





[ ]: