

```

import pandas as pd
import matplotlib.pyplot as plt

data = pd.read_csv("/dataset-anonimizado.csv")

semestres_restantes = data['Quantos semestres faltam para você se
formar?']
linguagens_programacao = data['Marque as linguagens de programação que
você já teve algum contato prático:'].apply(lambda x: len(x.split('\
n')))
experiencia_pd = data['Você já teve experiência de participação em
projetos de Pesquisa e Desenvolvimento?']

cores = ['green' if exp == 'Sim' else 'red' for exp in experiencia_pd]

transparencia = 1
tamanho_pontos = 200

semestres_restantes_sorted, linguagens_programacao_sorted,
cores_sorted = zip(*sorted(zip(semestres_restantes,
linguagens_programacao, cores), reverse=True))

plt.figure(figsize=(10, 6))

plt.scatter(semestres_restantes_sorted[::-1],
linguagens_programacao_sorted[::-1], c=cores_sorted[::-1],
alpha=transparencia, s=tamanho_pontos)

plt.xlabel('Semestres Restantes')
plt.ylabel('Quantidade de Linguagens de Programação')
plt.title('Relação entre experiência P&D com tempo de curso')

plt.legend(handles=[plt.Line2D([0], [0], marker='o', color='w',
markerfacecolor='green', markersize=10, label='Com Experiência em
P&D'),
plt.Line2D([0], [0], marker='o', color='w',
markerfacecolor='red', markersize=10, label='Sem Experiência em
P&D')],
title='Experiência em P&D')

plt.grid(True)
plt.tight_layout()
plt.show()

```

