



Search



ERIC MENDES - 10D AGO - 62 VIEWS



2

Copy &amp; Edit



# Novos Talentos - Um Estudo Sobre Pessoas Júnior

Python · State of Data Brazil 2021[Notebook](#)   [Data](#)   [Logs](#)   [Comments \(0\)](#)

Run

56.3s

Version 3 of 3

## Novos talentos



Sou um jovem engenheiro de dados que acabou de começar a carreira.

Tive dificuldades arranjando o primeiro emprego na área, então me interessa saber mais sobre colegas de nível júnior e é neste ponto onde focarei a análise.

Como veremos mais na seção "Motivação", também é importante para as empresas no geral entenderem sobre a atração e man-

## Dependências

In [1]:

```
!pip install matplotlib
import matplotlib.pyplot as plt

!pip install pandas
import pandas as pd
pd.set_option('display.max_columns', 500)

!pip install seaborn
import seaborn as sns
```

```
Requirement already satisfied: matplotlib in /opt/conda/lib/python3.7/site-packages (3.5.2)
```

```
Requirement already satisfied: numpy>=1.17 in /opt/conda/lib/python3.7/site-packages (from matplotlib) (1.21.6)
```

```
Requirement already satisfied: fonttools>=4.22.0 in /opt/conda/lib/python3.7/site-packages (from matplotlib) (4.30.0)
```

```
Requirement already satisfied: packaging>=20.0 in /opt/conda/lib/python3.7/site-pa
```

```
ckages (from matplotlib) (21.3)
Requirement already satisfied: kiwisolver>=1.0.1 in /opt/conda/lib/python3.7/site-
packages (from matplotlib) (1.4.0)
Requirement already satisfied: python-dateutil>=2.7 in /opt/conda/lib/python3.7/site-
packages (from matplotlib) (2.8.2)
Requirement already satisfied: cycler>=0.10 in /opt/conda/lib/python3.7/site-pac-
ges (from matplotlib) (0.11.0)
Requirement already satisfied: pyparsing>=2.2.1 in /opt/conda/lib/python3.7/site-pa-
ckages (from matplotlib) (3.0.7)
Requirement already satisfied: pillow>=6.2.0 in /opt/conda/lib/python3.7/site-pac-
kages (from matplotlib) (9.0.1)
Requirement already satisfied: typing-extensions in /opt/conda/lib/python3.7/site-
packages (from kiwisolver>=1.0.1->matplotlib) (4.2.0)
Requirement already satisfied: six>=1.5 in /opt/conda/lib/python3.7/site-packages
(from python-dateutil>=2.7->matplotlib) (1.16.0)
WARNING: Running pip as the 'root' user can result in broken permissions and con-
flicting behaviour with the system package manager. It is recommended to use a virtu-
al environment instead: https://pip.pypa.io/warnings/venv
Requirement already satisfied: pandas in /opt/conda/lib/python3.7/site-packages
(1.3.5)
Requirement already satisfied: python-dateutil>=2.7.3 in /opt/conda/lib/python3.7/
site-packages (from pandas) (2.8.2)
Requirement already satisfied: pytz>=2017.3 in /opt/conda/lib/python3.7/site-pac-
kages (from pandas) (2021.3)
Requirement already satisfied: numpy>=1.17.3 in /opt/conda/lib/python3.7/site-pac-
kages (from pandas) (1.21.6)
Requirement already satisfied: six>=1.5 in /opt/conda/lib/python3.7/site-packages
(from python-dateutil>=2.7.3->pandas) (1.16.0)
WARNING: Running pip as the 'root' user can result in broken permissions and con-
flicting behaviour with the system package manager. It is recommended to use a virtu-
al environment instead: https://pip.pypa.io/warnings/venv
Requirement already satisfied: seaborn in /opt/conda/lib/python3.7/site-packages
(0.11.2)
Requirement already satisfied: matplotlib>=2.2 in /opt/conda/lib/python3.7/site-pa-
ckages (from seaborn) (3.5.2)
Requirement already satisfied: numpy>=1.15 in /opt/conda/lib/python3.7/site-pac-
kages (from seaborn) (1.21.6)
Requirement already satisfied: scipy>=1.0 in /opt/conda/lib/python3.7/site-package
s (from seaborn) (1.7.3)
Requirement already satisfied: pandas>=0.23 in /opt/conda/lib/python3.7/site-pac-
kages (from seaborn) (1.3.5)
Requirement already satisfied: packaging>=20.0 in /opt/conda/lib/python3.7/site-pa-
ckages (from matplotlib>=2.2->seaborn) (21.3)
Requirement already satisfied: pyparsing>=2.2.1 in /opt/conda/lib/python3.7/site-pac-
kages (from matplotlib>=2.2->seaborn) (3.0.7)
Requirement already satisfied: pillow>=6.2.0 in /opt/conda/lib/python3.7/site-pac-
kages (from matplotlib>=2.2->seaborn) (9.0.1)
Requirement already satisfied: python-dateutil>=2.7 in /opt/conda/lib/python3.7/site-
packages (from matplotlib>=2.2->seaborn) (2.8.2)
Requirement already satisfied: fonttools>=4.22.0 in /opt/conda/lib/python3.7/site-
packages (from matplotlib>=2.2->seaborn) (4.30.0)
Requirement already satisfied: kiwisolver>=1.0.1 in /opt/conda/lib/python3.7/site-
packages (from matplotlib>=2.2->seaborn) (1.4.0)
Requirement already satisfied: cycler>=0.10 in /opt/conda/lib/python3.7/site-pac-
kages (from matplotlib>=2.2->seaborn) (0.11.0)
Requirement already satisfied: pytz>=2017.3 in /opt/conda/lib/python3.7/site-pac-
kages (from pandas>=0.23->seaborn) (2021.3)
```

```
Requirement already satisfied: typing-extensions in /opt/conda/lib/python3.7/site-packages (from kiwisolver>=1.0.1->matplotlib>=2.2->seaborn) (4.2.0)
Requirement already satisfied: six>=1.5 in /opt/conda/lib/python3.7/site-packages (from python-dateutil>=2.7->matplotlib>=2.2->seaborn) (1.16.0)
WARNING: Running pip as the 'root' user can result in broken permissions and conflicting behaviour with the system package manager. It is recommended to use a virtual environment instead: https://pip.pypa.io/warnings/venv
```

In [2]:

```
# Carregando o conjunto de dados
df = pd.read_csv('/kaggle/input/state-of-data-2021/State of Data 2021 - Dataset - Pgina1.csv', low_memory=False)
```

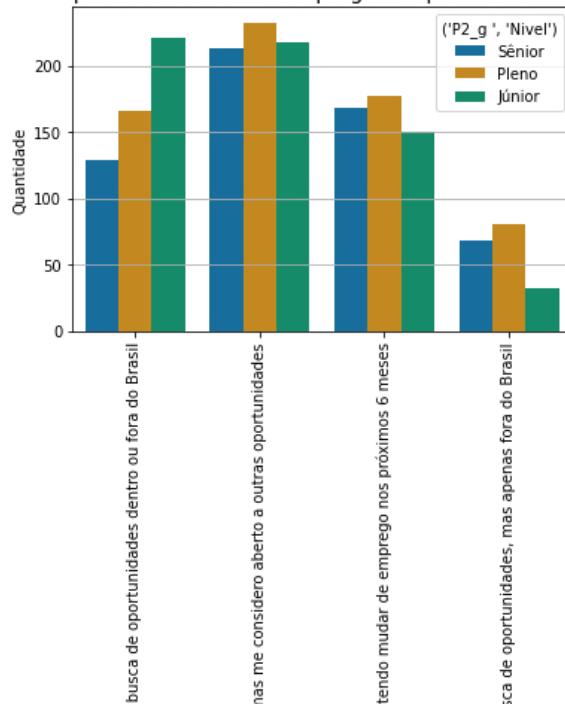
## Motivação

Para motivar a análise para além dos meus próprios interesses, visando também o interesse do leitor, gostaria de chamar a atenção para os números abaixo:

In [3]:

```
plt.title("Você pretende mudar de emprego nos próximos 6 meses?", fontsize=14)
sns.countplot(
    x=("(P2_n ', 'Você pretende mudar de emprego nos próximos 6 meses?'"),
    hue="('P2_g ', 'Nível')",
    data=df,
    palette=sns.color_palette("colorblind")
)
plt.xticks(rotation="vertical")
plt.ylabel("Quantidade")
plt.xlabel(None)
plt.grid(axis="y")
plt.show()
```

Você pretende mudar de emprego nos próximos 6 meses?



Estou em

Não estou buscando, n

Estou em bu:

Quando comparado com a situação de pessoas sênior, o número de pessoas de nível júnior que manifestaram estarem buscando oportunidades dentro ou fora do Brasil é significativamente maior.

Portanto também é de interesse empresarial entender mais sobre a atração e manutenção de pessoas júnior para a sustentabilidade dos seus próprios departamentos de dados.

## Exploração do conjunto de dados

Vamos começar respondendo à seguinte pergunta: quem é o júnior no Brasil?

In [4]:

```
# Minha análise será feita para pessoas júnior,  
# mas esta parametrização facilitará a análise para  
# outros níveis, caso o leitor deseje  
nivel_alvo = "Júnior"  
df_nivel_alvo = df.loc[df[("P2_g ", 'Nivel')] == nivel_alvo]
```

In [5]:

```
mediana_de_idade = df_nivel_alvo[("P1_a ", 'Idade")].median()  
genero = df_nivel_alvo[("P1_b ", 'Genero")].value_counts().index[0]  
regiao_onde_mora = df_nivel_alvo[("P1_e_b ", 'Regiao onde mora")].value_counts().in  
dex[0]  
nivel_de_ensino = df_nivel_alvo[("P1_h ", 'Nivel de Ensino")].value_counts().index  
[0]  
area_de_formacao = df_nivel_alvo[("P1_i ", 'Área de Formação")].value_counts().inde  
x[0]  
  
print(f"O {nivel_alvo} no Brasil é do gênero {genero} e possui {int(mediana_de_idad  
e)} anos de idade.\n\nMora na região {regiao_onde_mora}, é formado em {area_de_formacao}, e possui {nivel_d  
e_ensino} .")
```

O Júnior no Brasil é do gênero Masculino e possui 26 anos de idade.

Mora na região Sudeste, é formado em Computação / Engenharia de Software / Sistema  
s de Informação/ TI, e possui Graduação/Bacharelado.

Chamo atenção para a questão de gênero:

In [6]:

```
fig, ax = plt.subplots(1, 2, figsize=(17, 5))  
  
fig.suptitle(
```

```

'Porcentagem de presença do gênero masculino e não-masculino por nível'.title(),
fontsize=16
)

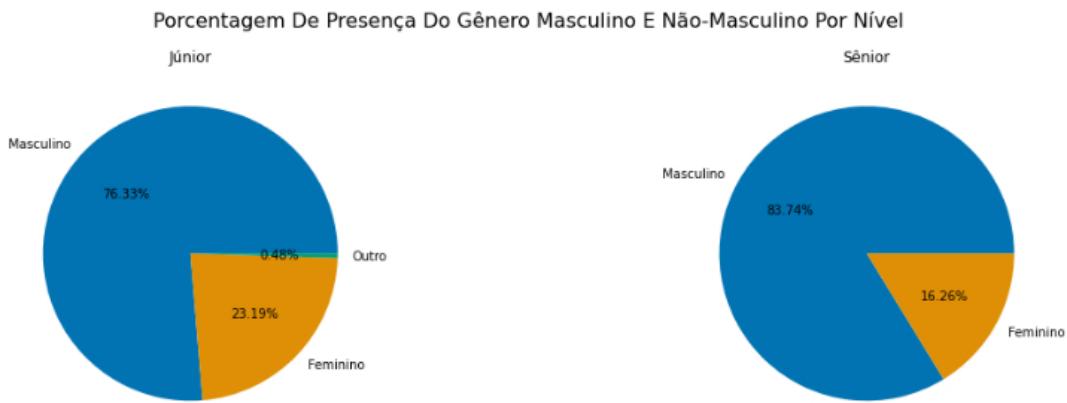
ax[0].pie(x=df_nivel_alvo[("P1_b ", 'Genero')].value_counts(), labels=df_nivel_alvo[("P1_b ", 'Genero')].value_counts().index, autopct=".2f%%", colors=sns.color_palette("colorblind"))
ax[0].set_title(nivel_alvo)

ax[1].pie(x=df.loc[df[("P2_g ", 'Nivel')] == "Sênior"][("P1_b ", 'Genero')].value_counts(), labels=df.loc[df[("P2_g ", 'Nivel')] == "Sênior"][("P1_b ", 'Genero')].value_counts().index, autopct=".2f%%", colors=sns.color_palette("colorblind"))
ax[1].set_title("Sênior")

fig.tight_layout()
fig.subplots_adjust(top=0.85)

plt.show()

```



Note que a presença de pessoas que não são do sexo masculino aumentam drasticamente no nível júnior (quase 50% comparado com o nível sênior), o que talvez indique que a área esteja caminhando para uma maior diversidade de gênero, mas note como o total ainda sequer chega a 1/4 da força de trabalho.

Fica o desafio às empresas: acelerem estes passos tímidos rumo à diversidade de gênero. Qualquer trabalhador da área sabe a importância da diversidade para eliminar viéses nos processos de obtenção e processamento de dados. Diversidade não é apenas importante, como também é necessário.

## Manutenção de bons funcionários júnior

Já vimos então quem são, e qual a importância da contratação de pessoas no nível júnior. Vamos ver agora o que valorizam e do que são críticos em uma empresa para que não sejam apenas atraídos, mas também sejam lapidados e bem geridos hoje, para se tornarem os gestores que a área de dados precisará amanhã.

In [7]:

```

from collections import Counter

def principais_criterios_que_levam_em_conta(df) -> Counter:
    criterios = [str(criterio).split(", ") for criterio in df.loc[~df[("P2_o ", 'Qua

```

```
is os principais critérios que você leva em consideração no momento de decidir onde trabalhar?").isna()][("P2_o ', 'Quais os principais critérios que você leva em consideração no momento de decidir onde trabalhar?")]]  
criterios_tratado = []  
for conjunto_criterio in criterios:  
    criterios_tratado += conjunto_criterio  
  
return Counter(criterios_tratado)
```

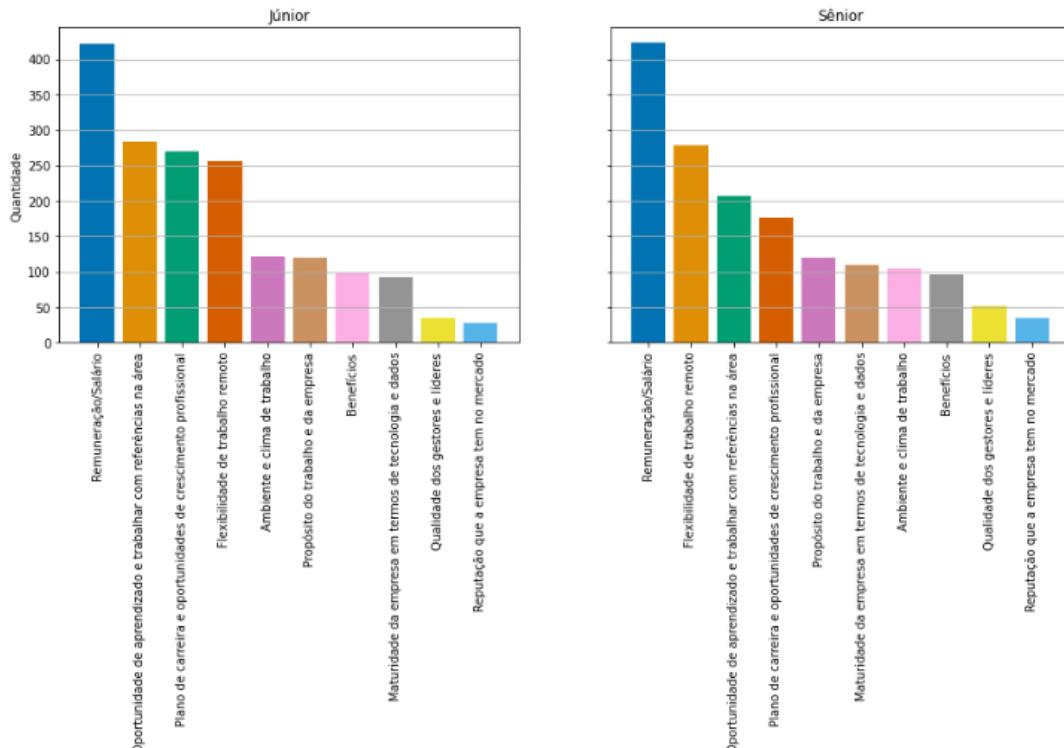
```
In [8]:  
criterios_nivel_alvo = principais_criterios_que_levam_em_conta(df_nivel_alvo)  
criterios_senior = principais_criterios_que_levam_em_conta(df.loc[df[("P2_g ', 'Nível')] == "Sênior"]]  
  
# Ordenando os resultados para facilitar a visualização  
criterios_nivel_alvo = sorted(criterios_nivel_alvo.items(), key=lambda tup: tup[1], reverse=True)[:10]  
criterios_senior = sorted(criterios_senior.items(), key=lambda tup: tup[1], reverse=True)[:10]  
  
criterios_nivel_alvo = list(zip(*criterios_nivel_alvo))  
criterios_senior = list(zip(*criterios_senior))
```

```
In [9]:  
fig, ax = plt.subplots(1, 2, figsize=(15, 5), sharey=True)  
  
fig.suptitle(  
    'Os 10 critérios mais levados em consideração na hora de decidir onde trabalhar'.  
    title(),  
    fontsize=16  
)  
  
ax[0].bar(x=criterios_nivel_alvo[0], height=criterios_nivel_alvo[1], color=sns.color_palette("colorblind"))  
ax[0].set_title(nivel_alvo.title())  
ax[0].grid(axis="y")  
ax[0].set_ylabel("Quantidade")  
ax[0].set_xticklabels(criterios_nivel_alvo[0], rotation=90)  
  
ax[1].bar(x=criterios_senior[0], height=criterios_senior[1], color=sns.color_palette("colorblind"))  
ax[1].set_title("Sênior".title())  
ax[1].grid(axis="y")  
ax[1].set_xticklabels(criterios_senior[0], rotation=90)  
  
fig.tight_layout()  
fig.subplots_adjust(top=0.85)  
plt.show()
```

```
/opt/conda/lib/python3.7/site-packages/ipykernel_launcher.py:12: UserWarning: FixedFormatter should only be used together with FixedLocator  
    if sys.path[0] == '':  
/opt/conda/lib/python3.7/site-packages/ipykernel_launcher.py:17: UserWarning: FixedFormatter should only be used together with FixedLocator  
/opt/conda/lib/python3.7/site-packages/ipykernel_launcher.py:19: UserWarning: Tigh
```

t layout not applied. The bottom and top margins cannot be made large enough to accommodate all axes decorations.

### Os 10 Critérios Mais Levados Em Consideração Na Hora De Decidir Onde Trabalhar



Note como a remuneração é o fator mais importante tanto para pessoas júnior como para pessoas sênior.  
Então nada de oferecer "salário emocional" só porque a pessoa está começando agora na carreira!

Chamo atenção também para como critérios que levam em conta crescimento profissional são dois dos três mais importantes para as pessoas de nível júnior. Nota em redes sociais corporativas por aí como algumas empresas ainda tem uma ideia chata e ultrapassada de que júnior só entra pra fazer o trabalho que um sênior não quer ou não pode.

Quer um júnior feliz hoje? Pague-o de forma justa e o ajude a se tornar o sênior especialista de amanhã!

### License

This Notebook has been released under the [Apache 2.0](#) open source license.

### Continue exploring



#### Data

1 input and 0 output



#### Logs

56.3 second run - successful





Comments

0 comments

