

# Predição de Performance de Colaboradores

## Entendimento do negócio

### Objetivo

Uma empresa de Contact Center e Customer Experience contrata muitas pessoas por ano. Esse ramo de Contact Center é conhecido por ser a porta de entrada de muita gente no mercado, pois não exige muita experiência.

No entanto, o turnover por baixo desempenho dos colaboradores também é muito alto, fazendo com que a empresa tenha que ficar contratando e treinando pessoas constantemente para que a operação de atendimento não fique sem colaboradores e o cliente final não sinta uma demora para ser atendido nos mais diversos canais de atendimento, desde o telefone até o chat e whatsapp.

A empresa gostaria de testar técnicas de People Analytics para ajudar o time de Atração de Talentos a filtrar melhor os perfis de colaboradores no processo seletivo a fim de trazer mais colaboradores com potencial de boa performance no trabalho. Para isso, os analistas de dados devem analisar um histórico de 7.240 colaboradores e sua performance após 6 meses a contratação para responder:

Com base nessas informações, pede-se:

1. Quais fatores o time de Atração de Talentos deve olhar para selecionar candidatos com maior chance de ter boa performance após 6 meses à data da contratação?
2. Com base nisso, qual política de contratação deve passar a valer de forma que a empresa aumente as boas contratações?

## Análise SWOT

### Forças (Strengths):

Potencial para melhorar significativamente a qualidade das contratações através de critérios mais objetivos e estruturados  
Possibilidade de reduzir o tempo do processo seletivo com uma filtragem mais eficiente  
Capacidade de criar um banco de dados com perfis adequados para futuras vagas  
Maior alinhamento entre as expectativas da empresa e o perfil dos candidatos

### Fraquezas (Weaknesses):

Risco de criar filtros muito rígidos que podem excluir talentos potenciais  
Necessidade de investimento em treinamento da equipe para aplicar novos critérios  
Possível resistência interna a mudanças no processo atual  
Dificuldade em estabelecer métricas precisas para prever performance futura

### Oportunidades (Opportunities):

Uso de tecnologias e ferramentas de análise de dados para aprimorar o processo  
Desenvolvimento de um processo mais estruturado que pode ser replicado em diferentes áreas  
Possibilidade de criar um programa de referências mais eficiente Potencial redução nos custos  
de turnover com melhores contratações Fortalecimento da marca empregadora com processos  
mais profissionais

Ameaças (Threats):

Mercado competitivo pode dificultar a atração dos perfis desejados Mudanças rápidas no  
mercado podem tornar alguns critérios de seleção obsoletos Risco de perder bons candidatos  
para concorrentes durante um processo muito longo Possíveis questões legais relacionadas a  
critérios de seleção discriminatórios Dificuldade em manter consistência no processo em  
diferentes departamentos

## Análise SMART

Específico (Specific):

Implementar um sistema estruturado de avaliação de candidatos baseado em:

Competências técnicas alinhadas com cada cargo Soft skills prioritárias para a empresa  
Histórico de performance em experiências anteriores Fit cultural com a organização Potencial de  
desenvolvimento

Mensurável (Measurable):

Definir métricas claras de sucesso:

Redução de 30% no tempo médio de contratação Aumento de 25% na taxa de retenção após  
período probatório Melhoria de 40% nas avaliações de performance do primeiro ano Redução  
de 50% no índice de turnover precoce (primeiros 6 meses) Aumento de 20% na satisfação dos  
gestores com as novas contratações

Alcançável (Attainable):

Estabelecer etapas graduais de implementação:

Fase 1: Mapeamento dos critérios de sucesso por cargo Fase 2: Desenvolvimento de ferramentas  
de avaliação Fase 3: Treinamento da equipe de recrutamento Fase 4: Implementação piloto em  
departamentos específicos Fase 5: Expansão para toda a empresa

Relevante (Relevant):

Alinhar o projeto com objetivos estratégicos:

Redução de custos com recrutamento e treinamento Aumento da produtividade geral Melhoria  
do clima organizacional Fortalecimento da cultura empresarial Desenvolvimento de equipes  
mais eficientes

Temporal (Time-bound):

Estabelecer cronograma de implementação:

Mês 1-2: Desenvolvimento do projeto e critérios    Mês 3: Treinamento da equipe    Mês 4-5: Teste piloto    Mês 6: Avaliação de resultados preliminares    Mês 7-8: Ajustes e melhorias    Mês 9-12: Implementação completa em toda empresa    Mês 12: Avaliação geral dos resultados

## Entendimento dos dados

### 1. Carregando as Bibliotecas

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import statsmodels.api as sm
import scorecardpy as sc
sns.set()
```

### 2. Carregando Dataset - Exploração da estrutura

```
metadados = pd.read_excel('base_atendimento.xlsx', header=1)
metadados = baseoriginal.iloc[:,4]
metadados
```

```
-----
-----
NameError                                Traceback (most recent call
last)
Cell In[2], line 2
      1 metadados = pd.read_excel('base_atendimento.xlsx', header=1)
----> 2 metadados = baseoriginal.iloc[:,4]
      3 metadados

NameError: name 'baseoriginal' is not defined
```

O dataframe metadados identifica as variáveis de nosso dataset base

```
df = pd.read_excel('base_atendimento.xlsx', sheet_name="Base")
df.head()
```

	Id_Agente	Boa_Performance	\
0	1	1	
1	2	1	
2	3	1	
3	4	0	
4	5	0	

P 01: Antes de trabalhar na empresa, você tinha experiência em

Contact Center? \

- |   |     |
|---|-----|
| 0 | Não |
| 1 | Sim |
| 2 | Sim |
| 3 | Não |
| 4 | Não |

P 02: Qual área de atendimento você atua ou atuou por mais tempo? \

- |   |                       |
|---|-----------------------|
| 0 | Sem resposta          |
| 1 | Atendimento Receptivo |
| 2 | Ativo de Vendas       |
| 3 | Sem resposta          |
| 4 | Sem resposta          |

P 03: Qual a sua escolaridade? P 04: Qual seu conhecimento em inglês? \

- |   |  |                                  |
|---|--|----------------------------------|
| 0 | Ensino Superior incompleto Intermediário |                                  |
| 1 | Ensino Médio completo                    | Não tenho conhecimento em inglês |
| 2 | Ensino Superior cursando Intermediário   |                                  |
| 3 | Ensino Superior concluído Básico         |                                  |
| 4 | Ensino Médio completo Avançado           |                                  |

P 05: Qual o tempo gasto no trajeto da sua residência até a empresa? \

- |   |              |
|---|--------------|
| 0 | Sem resposta |
| 1 | Sem resposta |
| 2 | Sem resposta |
| 3 | Sem resposta |
| 4 | Sem resposta |

P 06: Você possui dependentes? \

- |   |     |
|---|-----|
| 0 | Não |
| 1 | Não |
| 2 | Não |
| 3 | Não |
| 4 | Não |

P 07: É bom trabalhar com meus colegas de trabalho. \

- |   |                  |
|---|------------------|
| 0 | 0 - Sem resposta |
| 1 | 0 - Sem resposta |
| 2 | 4 - Concordo     |
| 3 | 4 - Concordo     |
| 4 | 4 - Concordo     |

P 08: Não tenho medo de criticar pessoas em cargos de liderança. \

- |   |                  |
|---|------------------|
| 0 | 0 - Sem resposta |
| 1 | 0 - Sem resposta |
| 2 | 3 - Indiferente  |
| 3 | 4 - Concordo     |
| 4 | 4 - Concordo     |

P 09: Fico tenso quando estou trabalhando. \

- |   |                  |
|---|------------------|
| 0 | 0 - Sem resposta |
| 1 | 0 - Sem resposta |
| 2 | 3 - Indiferente  |
| 3 | 4 - Concordo     |
| 4 | 3 - Indiferente  |

P 10: Não gosto de situações que requerem que eu me comporte de forma controlada. \

- |   |                  |
|---|------------------|
| 0 | 0 - Sem resposta |
| 1 | 0 - Sem resposta |
| 2 | 2 - Discordo     |
| 3 | 4 - Concordo     |
| 4 | 3 - Indiferente  |

P 11: Gosto de dizer o que penso. \

- |   |                  |
|---|------------------|
| 0 | 0 - Sem resposta |
| 1 | 0 - Sem resposta |
| 2 | 2 - Discordo     |
| 3 | 4 - Concordo     |
| 4 | 4 - Concordo     |

P 12: Gosto de fazer as coisas do meu jeito e não ligo para o que possam pensar.

- |   |                  |
|---|------------------|
| 0 | 0 - Sem resposta |
| 1 | 0 - Sem resposta |
| 2 | 2 - Discordo     |
| 3 | 3 - Indiferente  |

Vemos que cada linha contém a informação referente a cada Colaborador, além da informação referente a Boa Performance.

```
df.dtypes

Id_Agente
int64
Boa_Performance
int64
P 01: Antes de trabalhar na empresa, você tinha experiência em Contact
Center?      object
P 02: Qual área de atendimento você atua ou atuou por mais tempo?
object
P 03: Qual a sua escolaridade?
object
P 04: Qual seu conhecimento em inglês?
object
P 05: Qual o tempo gasto no trajeto da sua residência até a empresa?
object
P 06: Você possui dependentes?
object
P 07: É bom trabalhar com meus colegas de trabalho.
object
P 08: Não tenho medo de criticar pessoas em cargos de liderança.
object
P 09: Fico tenso quando estou trabalhando.
object
P 10: Não gosto de situações que requerem que eu me comporte de forma
controlada.   object
P 11: Gosto de dizer o que penso.
object
P 12: Gosto de fazer as coisas do meu jeito e não ligo para o que
possam pensar. object
dtype: object

df.drop('Id_Agente', axis = 1, inplace=True)
```

Como não vou utilizar A Coluna Id\_Agente, retirei a mesma.

```
df.columns

Index(['Boa_Performance',
      'P 01: Antes de trabalhar na empresa, você tinha experiência em
Contact Center?',
      'P 02: Qual área de atendimento você atua ou atuou por mais
```

```

tempo?',
    'P 03: Qual a sua escolaridade?',
    'P 04: Qual seu conhecimento em inglês?',
    'P 05: Qual o tempo gasto no trajeto da sua residência até a
empresa?',
    'P 06: Você possui dependentes?',
    'P 07: É bom trabalhar com meus colegas de trabalho.',
    'P 08: Não tenho medo de criticar pessoas em cargos de
liderança.',
    'P 09: Fico tenso quando estou trabalhando.',
    'P 10: Não gosto de situações que requerem que eu me comporte
de forma controlada.',
    'P 11: Gosto de dizer o que penso.',
    'P 12: Gosto de fazer as coisas do meu jeito e não ligo para o
que possam pensar.'],
    dtype='object')

df.shape

(7240, 13)

df.isnull().sum()

Boa_Performance
0
P 01: Antes de trabalhar na empresa, você tinha experiência em Contact
Center?      0
P 02: Qual área de atendimento você atua ou atuou por mais tempo?
0
P 03: Qual a sua escolaridade?
0
P 04: Qual seu conhecimento em inglês?
0
P 05: Qual o tempo gasto no trajeto da sua residência até a empresa?
0
P 06: Você possui dependentes?
0
P 07: É bom trabalhar com meus colegas de trabalho.
0
P 08: Não tenho medo de criticar pessoas em cargos de liderança.
0
P 09: Fico tenso quando estou trabalhando.
0
P 10: Não gosto de situações que requerem que eu me comporte de forma
controlada.      0
P 11: Gosto de dizer o que penso.
0
P 12: Gosto de fazer as coisas do meu jeito e não ligo para o que
possam pensar.      0
dtype: int64

```

```
df.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 7240 entries, 0 to 7239
Data columns (total 13 columns):
 #   Column
Non-Null Count  Dtype
---  -
0   Boa_Performance
7240 non-null   int64
1   P 01: Antes de trabalhar na empresa, você tinha experiência em
Contact Center? 7240 non-null   object
2   P 02: Qual área de atendimento você atua ou atuou por mais tempo?
7240 non-null   object
3   P 03: Qual a sua escolaridade?
7240 non-null   object
4   P 04: Qual seu conhecimento em inglês?
7240 non-null   object
5   P 05: Qual o tempo gasto no trajeto da sua residência até a
empresa?        7240 non-null   object
6   P 06: Você possui dependentes?
7240 non-null   object
7   P 07: É bom trabalhar com meus colegas de trabalho.
7240 non-null   object
8   P 08: Não tenho medo de criticar pessoas em cargos de liderança.
7240 non-null   object
9   P 09: Fico tenso quando estou trabalhando.
7240 non-null   object
10  P 10: Não gosto de situações que requerem que eu me comporte de
forma controlada. 7240 non-null   object
11  P 11: Gosto de dizer o que penso.
7240 non-null   object
12  P 12: Gosto de fazer as coisas do meu jeito e não ligo para o que
possam pensar. 7240 non-null   object
dtypes: int64(1), object(12)
memory usage: 735.4+ KB
```

O Dataset não possui nenhum tipo de dado faltante/nulo, ou seja, não é necessário realizar nenhum tipo de transformação inicial nos dados atuais.

### 3. Análise Univariada

Nessa parte irei explorar cada variável, afim de entender possíveis causas que colaboram com a Performance, além também de descobrir possíveis padrões.



## 3.1 Antes de trabalhar na empresa, você tinha experiência em Contact Center?

```
# Total de respostas referente a experiência

df["P 01: Antes de trabalhar na empresa, você tinha experiência em
Contact Center?"].value_counts()

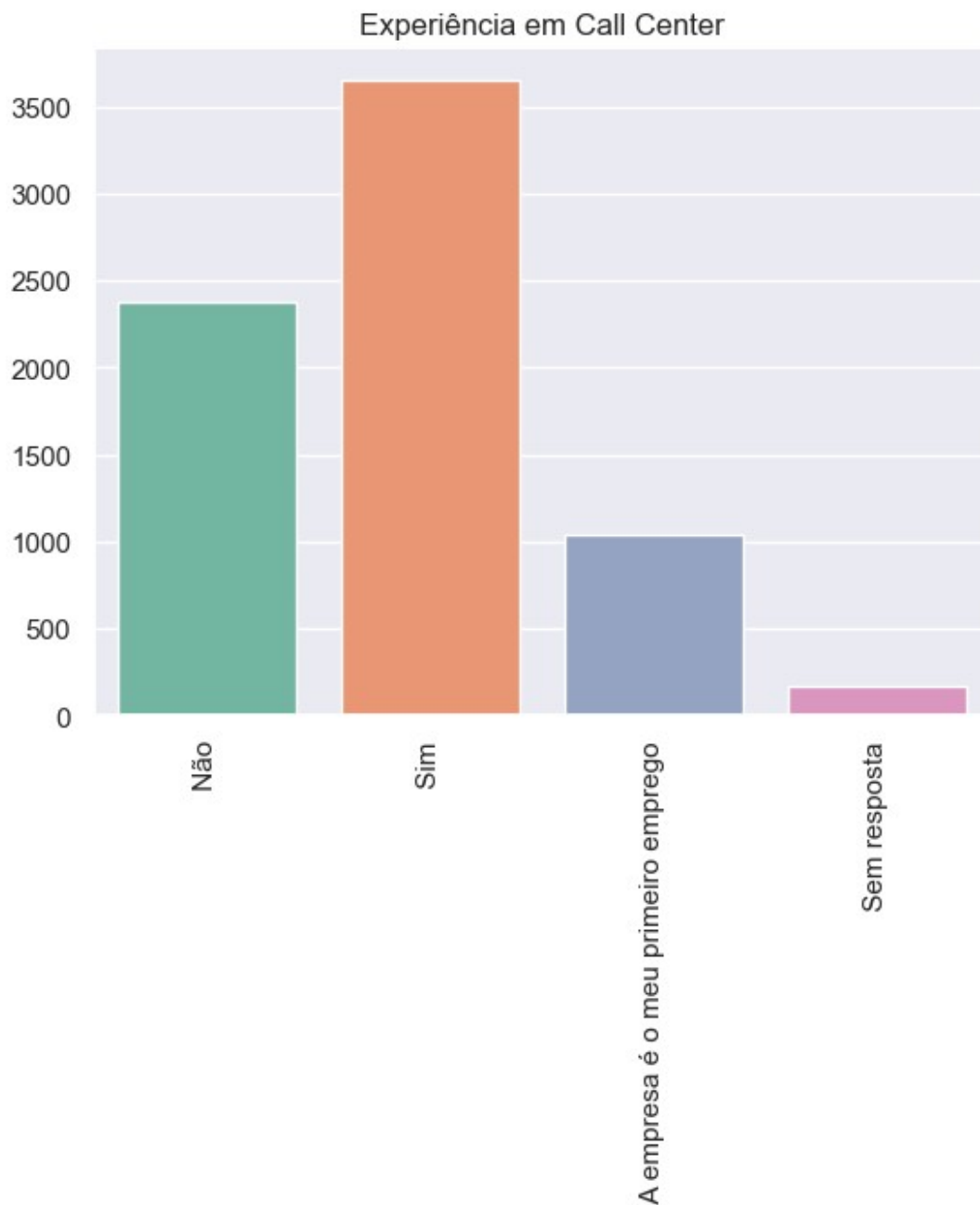
P 01: Antes de trabalhar na empresa, você tinha experiência em Contact
Center?
Sim                3651
Não                2380
A empresa é o meu primeiro emprego    1038
Sem resposta       171
Name: count, dtype: int64

# Frequencia relativa referente a experiência

df["P 01: Antes de trabalhar na empresa, você tinha experiência em
Contact Center?"].value_counts() / sum(df["P 01: Antes de trabalhar na
empresa, você tinha experiência em Contact Center?"].value_counts())

P 01: Antes de trabalhar na empresa, você tinha experiência em Contact
Center?
Sim                0.504282
Não                0.328729
A empresa é o meu primeiro emprego    0.143370
Sem resposta       0.023619
Name: count, dtype: float64

plt.title("Experiência em Call Center")
sns.countplot(
    data=df,
    x="P 01: Antes de trabalhar na empresa, você tinha experiência em
Contact Center?",
    hue="P 01: Antes de trabalhar na empresa, você tinha experiência
em Contact Center?",
    palette="Set2",
    legend=False
)
plt.xticks(rotation=90)
plt.xlabel('')
plt.ylabel('')
sns.despine()
plt.show()
```



Obtemos que:

50% dos colaboradores já possuem experiência em call Center  
32% dos colaboradores não tinham experiência em call Center

## 3.2 Qual área de atendimento você atua ou atuou por mais tempo?

```
# Total

df["P 02: Qual área de atendimento você atua ou atuou por mais tempo?"].value_counts()

P 02: Qual área de atendimento você atua ou atuou por mais tempo?
Sem resposta                                3613
Atendimento Receptivo                      1487
Backoffice                                473
Ativo de Vendas                            353
Cobrança                                   320
Suporte Técnico                           288
Chat ou email                             265
Multiskill                                226
Retenção                                   99
Receptivo de Vendas                        90
Suporte Técnico com Vendas                 26
Name: count, dtype: int64

# Frequencia relativa

df["P 02: Qual área de atendimento você atua ou atuou por mais tempo?"].value_counts() / sum(df["P 02: Qual área de atendimento você atua ou atuou por mais tempo?"].value_counts())

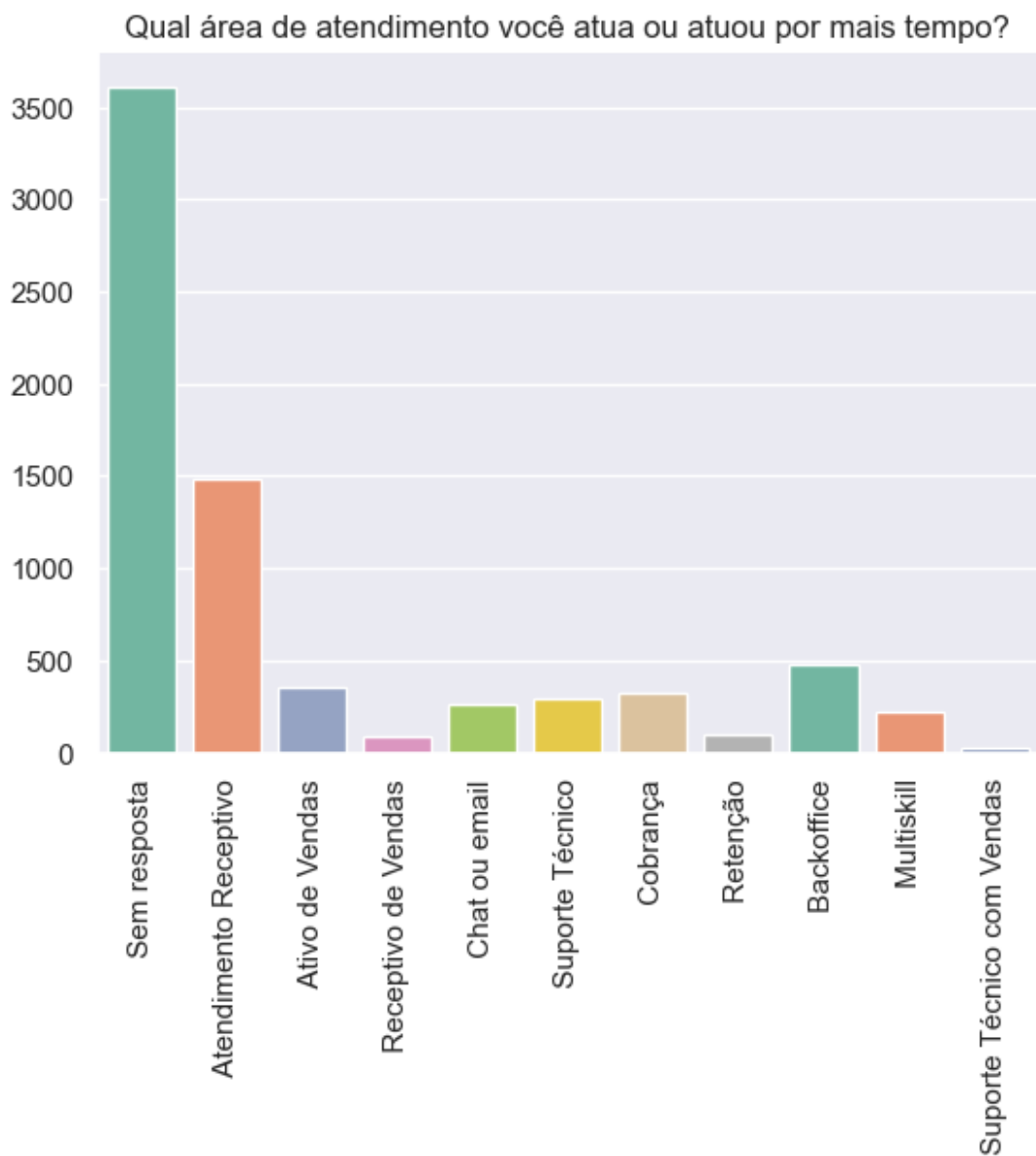
P 02: Qual área de atendimento você atua ou atuou por mais tempo?
Sem resposta                                0.499033
Atendimento Receptivo                      0.205387
Backoffice                                0.065331
Ativo de Vendas                            0.048757
Cobrança                                   0.044199
Suporte Técnico                           0.039779
Chat ou email                             0.036602
Multiskill                                0.031215
Retenção                                   0.013674
Receptivo de Vendas                        0.012431
Suporte Técnico com Vendas                 0.003591
Name: count, dtype: float64

plt.title("Qual área de atendimento você atua ou atuou por mais tempo?")
sns.countplot(
    data=df,
    x="P 02: Qual área de atendimento você atua ou atuou por mais tempo?",
    hue="P 02: Qual área de atendimento você atua ou atuou por mais tempo?",
```

```

tempo?",
    palette="Set2",
    legend=False
)
plt.xticks(rotation=90)
plt.xlabel('')
plt.ylabel('')
sns.despine()
plt.show()

```



Obtemos que:

49% dos colaboradores não responderam ou não tiveram suas respostas computadas.  
20% dos colaboradores atuaram em atendimento receptivo

### 3.3 Qual a sua escolaridade?

*# Total*

```
df["P 03: Qual a sua escolaridade?"].value_counts()
```

P 03: Qual a sua escolaridade?

Ensino Médio completo	2880
Ensino Superior cursando	1799
Ensino Superior concluído	1237
Ensino Superior incompleto	966
Pós Graduação cursando	137
Pós Graduação concluído	129
Sem resposta	92

Name: count, dtype: int64

*# Frequencia relativa*

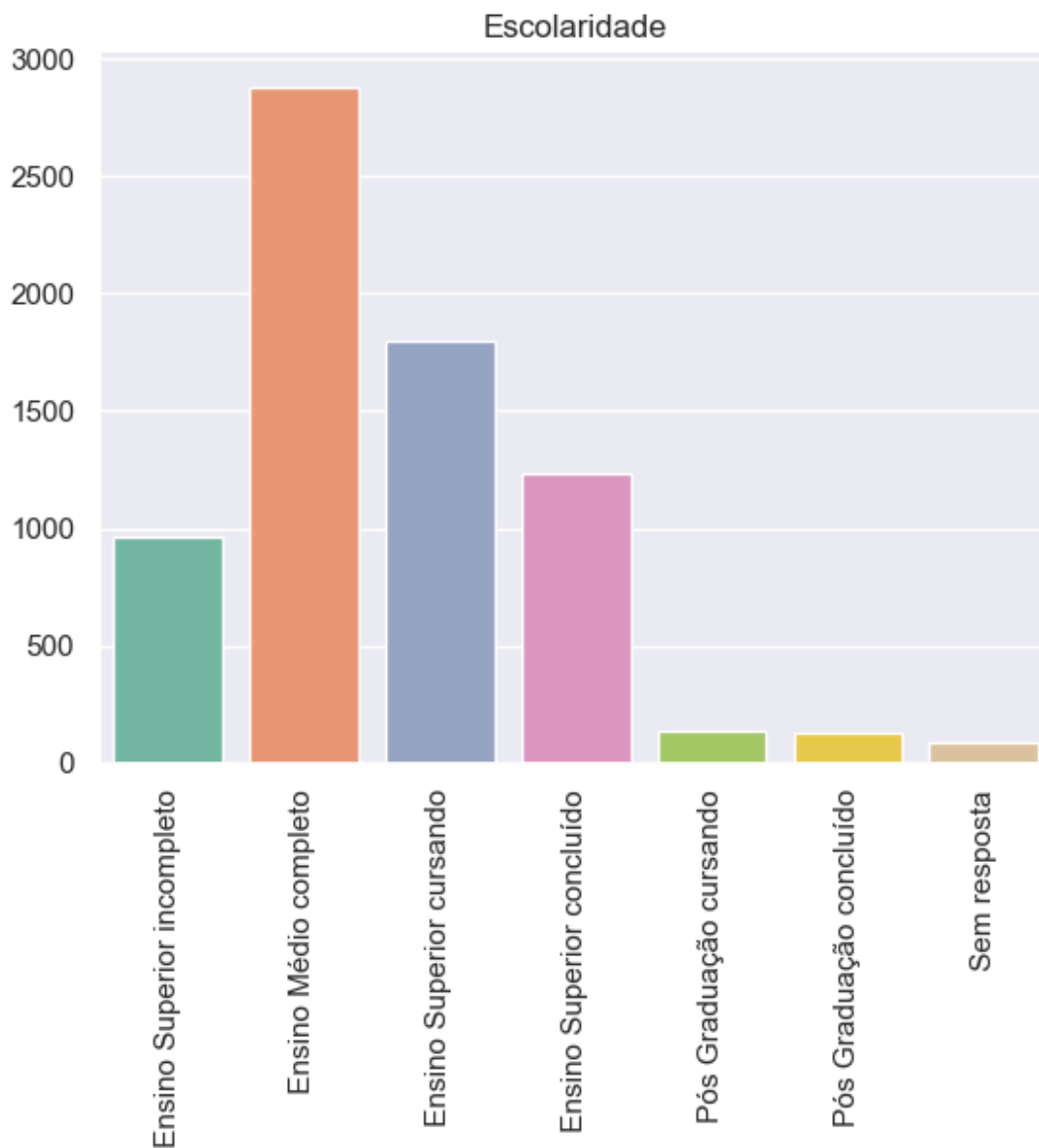
```
df["P 03: Qual a sua escolaridade?"].value_counts() / sum(df["P 03: Qual a sua escolaridade?"].value_counts())
```

P 03: Qual a sua escolaridade?

Ensino Médio completo	0.397790
Ensino Superior cursando	0.248481
Ensino Superior concluído	0.170856
Ensino Superior incompleto	0.133425
Pós Graduação cursando	0.018923
Pós Graduação concluído	0.017818
Sem resposta	0.012707

Name: count, dtype: float64

```
plt.title("Escolaridade")
sns.countplot(
    data=df,
    x="P 03: Qual a sua escolaridade?",
    hue="P 03: Qual a sua escolaridade?",
    palette="Set2",
    legend=False
)
plt.xticks(rotation=90)
plt.xlabel('')
plt.ylabel('')
sns.despine()
plt.show()
```



Obtemos que:

39% dos colaboradores tem Ensino Médio Completo  
24% dos colaboradores estão Cursando o Ensino Superior

### 3.4 Qual seu conhecimento em inglês?

# Total

```
df["P 04: Qual seu conhecimento em inglês?"].value_counts()
```

P 04: Qual seu conhecimento em inglês?  
Básico 3668

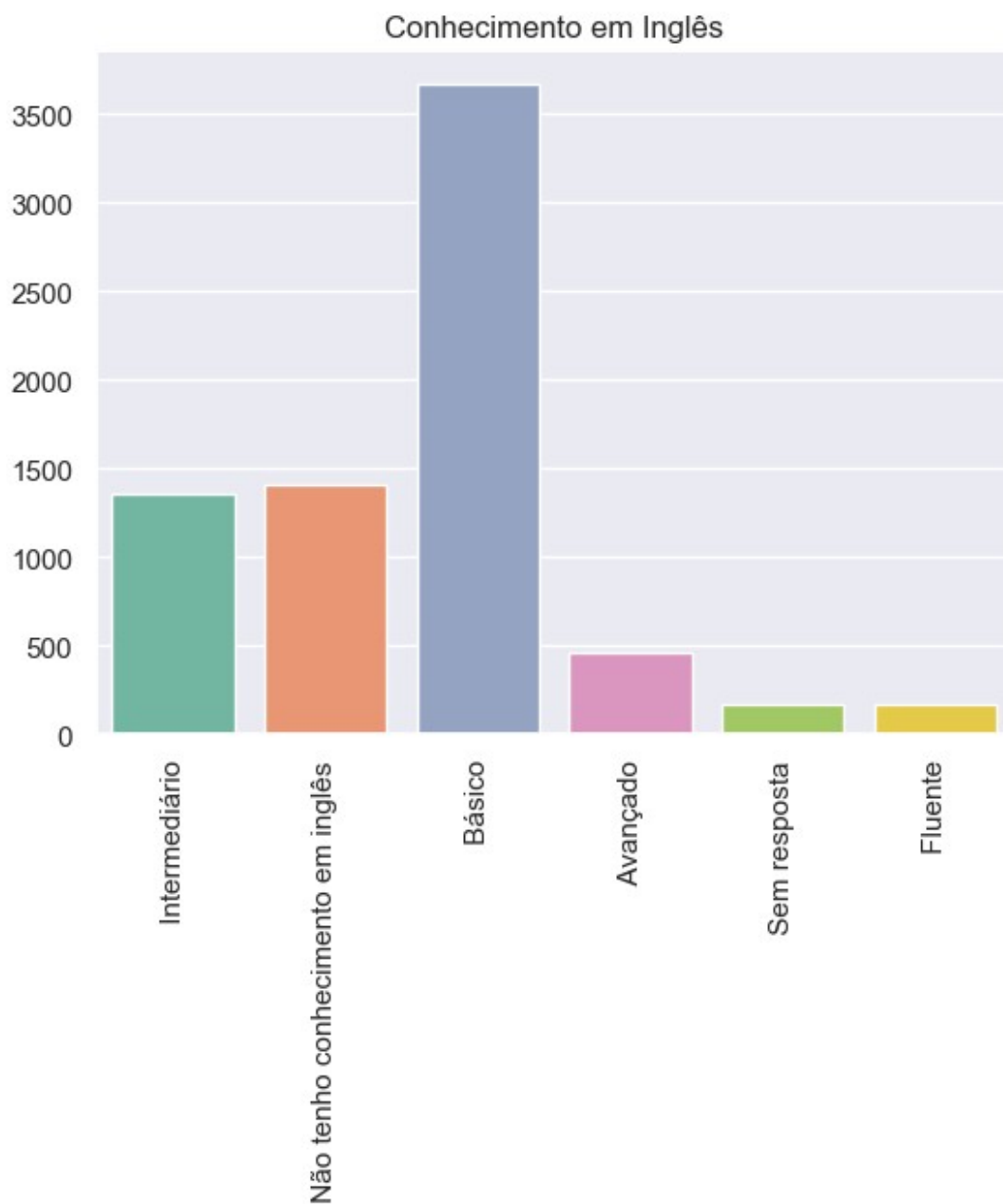
```
Não tenho conhecimento em inglês    1406
Intermediário                        1359
Avançado                             463
Fluente                              173
Sem resposta                         171
Name: count, dtype: int64
```

```
# Frequencia relativa
```

```
df["P 04: Qual seu conhecimento em inglês?"].value_counts() /
sum(df["P 04: Qual seu conhecimento em inglês?"].value_counts())
```

```
P 04: Qual seu conhecimento em inglês?
Básico                                0.506630
Não tenho conhecimento em inglês      0.194199
Intermediário                        0.187707
Avançado                             0.063950
Fluente                              0.023895
Sem resposta                         0.023619
Name: count, dtype: float64
```

```
plt.title("Conhecimento em Inglês")
sns.countplot(
    data=df,
    x="P 04: Qual seu conhecimento em inglês?",
    hue="P 04: Qual seu conhecimento em inglês?",
    palette="Set2",
    legend=False
)
plt.xticks(rotation=90)
plt.xlabel('')
plt.ylabel('')
sns.despine()
plt.show()
```



Obtemos que:

50% dos colaboradores tem Conhecimento básico em Inglês  
19% dos colaboradores não tem Conhecimento básico em Inglês



## 3.5 Qual o tempo gasto no trajeto da sua residência até a empresa?

```
# Total

df["P 05: Qual o tempo gasto no trajeto da sua residência até a
empresa?"].value_counts()

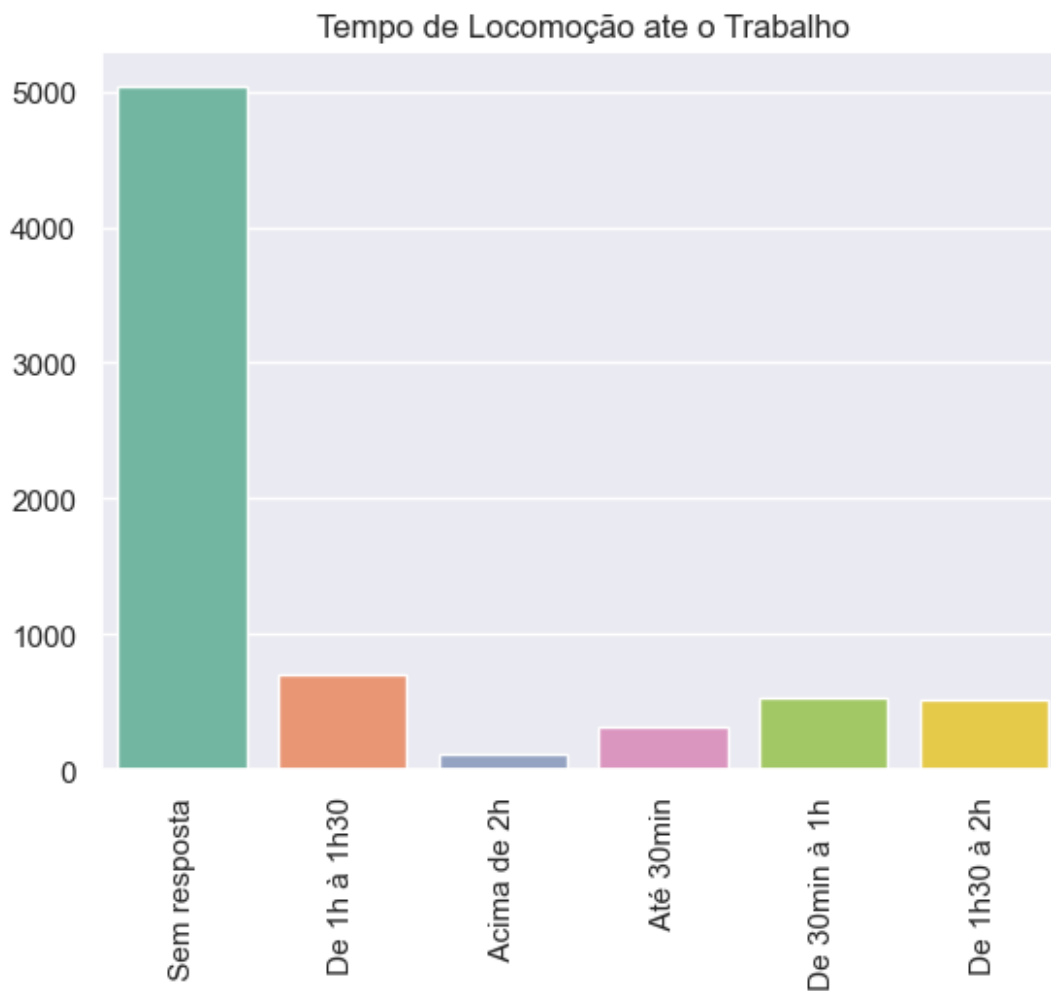
P 05: Qual o tempo gasto no trajeto da sua residência até a empresa?
Sem resposta      5033
De 1h à 1h30      712
De 30min à 1h     540
De 1h30 à 2h      523
Até 30min         314
Acima de 2h       118
Name: count, dtype: int64

# Frequencia relativa

df["P 05: Qual o tempo gasto no trajeto da sua residência até a
empresa?"].value_counts() / sum(df["P 05: Qual o tempo gasto no
trajeto da sua residência até a empresa?"].value_counts())

P 05: Qual o tempo gasto no trajeto da sua residência até a empresa?
Sem resposta      0.695166
De 1h à 1h30      0.098343
De 30min à 1h     0.074586
De 1h30 à 2h      0.072238
Até 30min         0.043370
Acima de 2h       0.016298
Name: count, dtype: float64

plt.title("Tempo de Locomoção ate o Trabalho")
sns.countplot(
    data=df,
    x="P 05: Qual o tempo gasto no trajeto da sua residência até a
empresa?",
    hue="P 05: Qual o tempo gasto no trajeto da sua residência até a
empresa?",
    palette="Set2",
    legend=False
)
plt.xticks(rotation=90)
plt.xlabel('')
plt.ylabel('')
sns.despine()
plt.show()
```



Obtemos que:

69% dos colaboradores não responderam ou não tiveram suas respostas computadas.

### 3.6 Você possui dependentes?

# Total

```
df["P 06: Você possui dependentes?"].value_counts()
```

P 06: Você possui dependentes?

Não 4838

Sim 2269

Sem resposta 133

Name: count, dtype: int64

```
# Frequencia relativa
```

```
df["P 06: Você possui dependentes?"].value_counts() / sum(df["P 06: Você possui dependentes?"].value_counts())
```

```
P 06: Você possui dependentes?
```

```
Não          0.668232
```

```
Sim           0.313398
```

```
Sem resposta  0.018370
```

```
Name: count, dtype: float64
```

```
plt.title("Tem Dependentes")
```

```
sns.countplot(  
    data=df,  
    x="P 06: Você possui dependentes?",  
    hue="P 06: Você possui dependentes?",  
    palette="Set2",  
    legend=False  
)
```

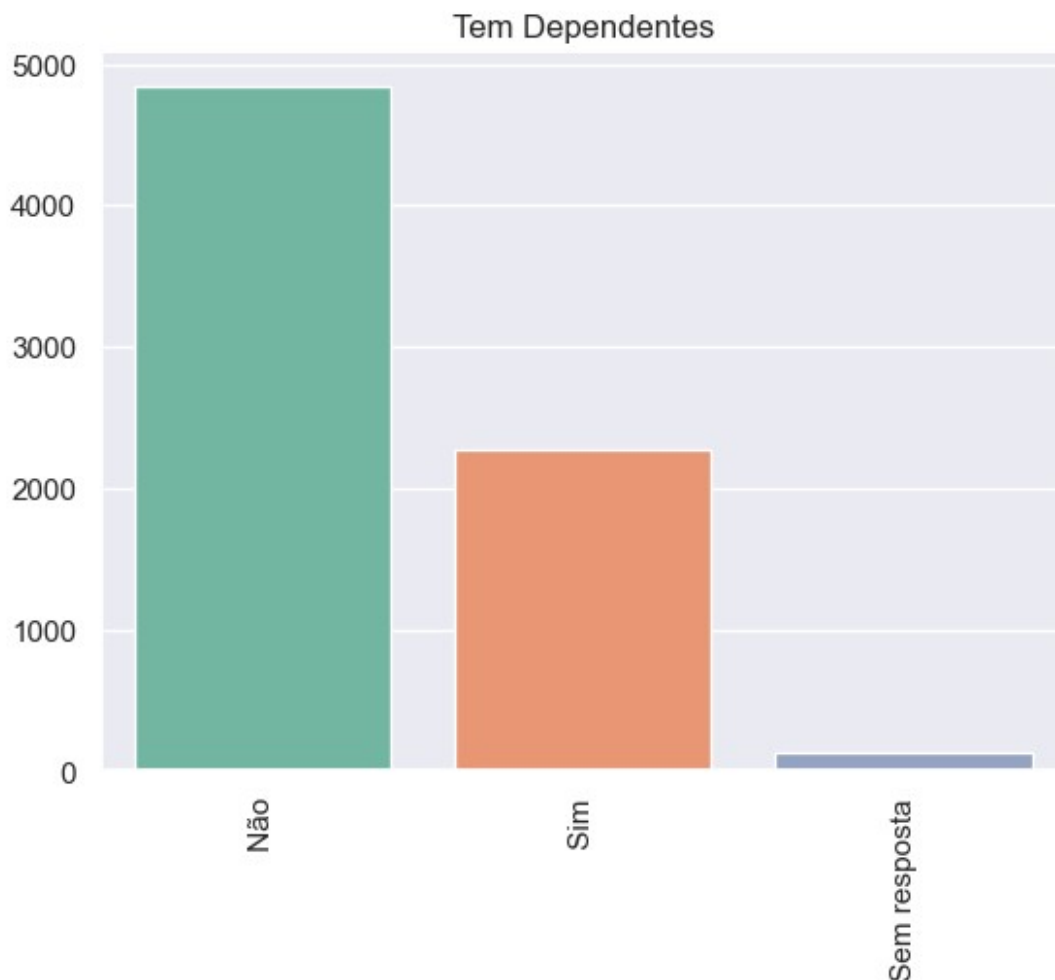
```
plt.xticks(rotation=90)
```

```
plt.xlabel('')
```

```
plt.ylabel('')
```

```
sns.despine()
```

```
plt.show()
```



Obtemos que:

66% dos colaboradores não possuem dependentes.  
31% dos colaboradores possuem dependentes.

### 3.7 É bom trabalhar com meus colegas de trabalho?

# Total

```
df["P 07: É bom trabalhar com meus colegas de trabalho."].value_counts()
```

P 07: É bom trabalhar com meus colegas de trabalho.

5 - Concordo totalmente	3296
4 - Concordo	2687

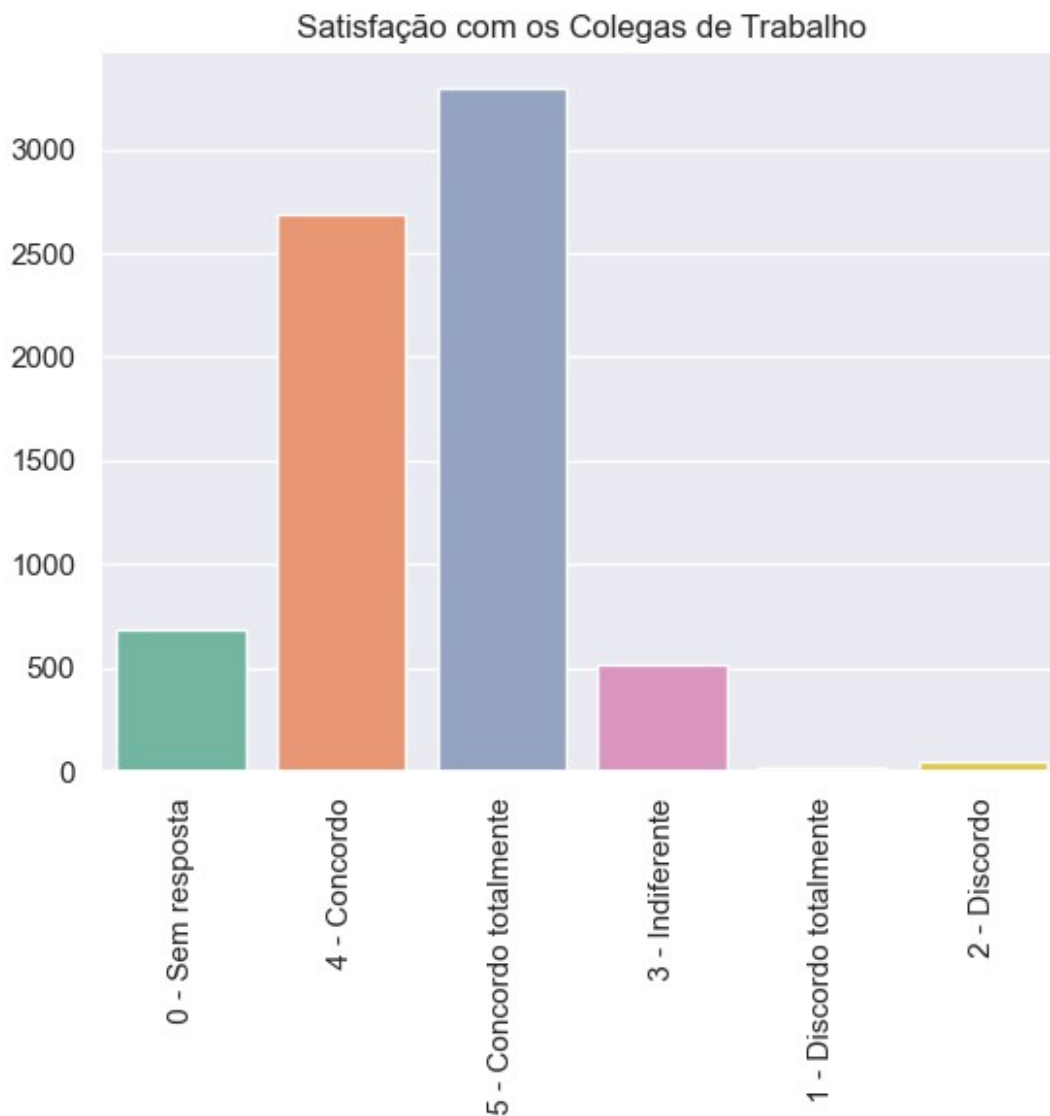
```
0 - Sem resposta          684
3 - Indiferente           514
2 - Discordo              45
1 - Discordo totalmente   14
Name: count, dtype: int64
```

```
# Frequencia relativa
```

```
df["P 07: É bom trabalhar com meus colegas de
trabalho."].value_counts() / sum(df["P 07: É bom trabalhar com meus
colegas de trabalho."].value_counts())
```

```
P 07: É bom trabalhar com meus colegas de trabalho.
5 - Concordo totalmente    0.455249
4 - Concordo               0.371133
0 - Sem resposta           0.094475
3 - Indiferente            0.070994
2 - Discordo               0.006215
1 - Discordo totalmente    0.001934
Name: count, dtype: float64
```

```
plt.title("Satisfação com os Colegas de Trabalho")
sns.countplot(
    data=df,
    x="P 07: É bom trabalhar com meus colegas de trabalho.",
    hue="P 07: É bom trabalhar com meus colegas de trabalho.",
    palette="Set2",
    legend=False
)
plt.xticks(rotation=90)
plt.xlabel('')
plt.ylabel('')
sns.despine()
plt.show()
```



Obtemos que:

82% dos colaboradores concordam ou concordam totalmente que estão satisfeitos com os seus colegas.

### 3.8 Não tenho medo de criticar pessoas em cargos de liderança

# Total

```
df["P 08: Não tenho medo de criticar pessoas em cargos de liderança."].value_counts()
```

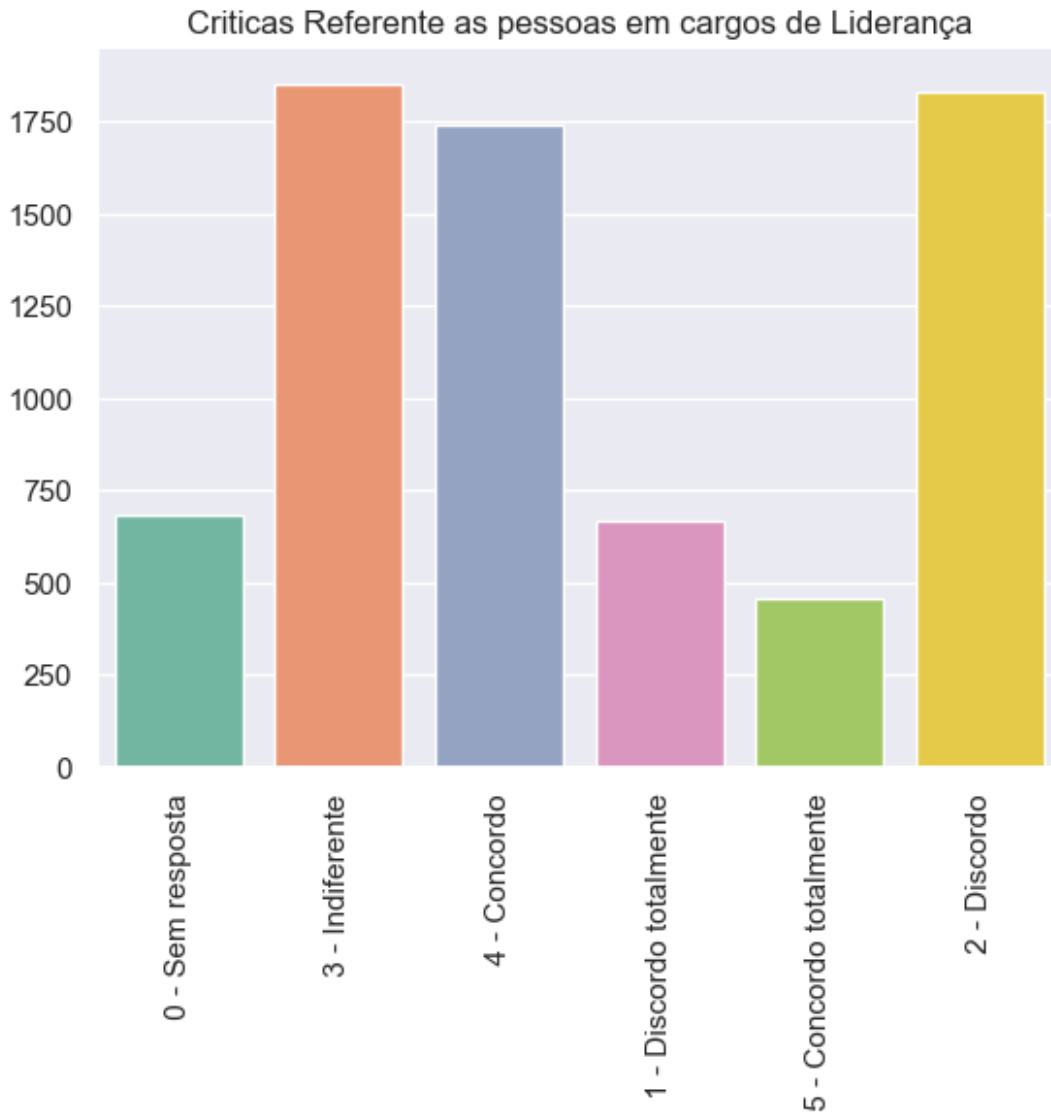
```
P 08: Não tenho medo de criticar pessoas em cargos de liderança.
3 - Indiferente      1853
2 - Discordo         1831
4 - Concordo         1744
0 - Sem resposta     684
1 - Discordo totalmente 668
5 - Concordo totalmente 460
Name: count, dtype: int64
```

```
# Frequencia relativa
```

```
df["P 08: Não tenho medo de criticar pessoas em cargos de
liderança."].value_counts() / sum(df["P 08: Não tenho medo de criticar
pessoas em cargos de liderança."].value_counts())
```

```
P 08: Não tenho medo de criticar pessoas em cargos de liderança.
3 - Indiferente      0.255939
2 - Discordo         0.252901
4 - Concordo         0.240884
0 - Sem resposta     0.094475
1 - Discordo totalmente 0.092265
5 - Concordo totalmente 0.063536
Name: count, dtype: float64
```

```
plt.title("Criticas Referente as pessoas em cargos de Liderança")
sns.countplot(
    data=df,
    x="P 08: Não tenho medo de criticar pessoas em cargos de
liderança.",
    hue="P 08: Não tenho medo de criticar pessoas em cargos de
liderança.",
    palette="Set2",
    legend=False
)
plt.xticks(rotation=90)
plt.xlabel('')
plt.ylabel('')
sns.despine()
plt.show()
```



Obtemos que:

25% dos colaboradores tem medo de criticar pessoas em cargos de Liderança.  
24% dos colaboradores não tem medo de criticar pessoas em cargos de Liderança.  
25% dos colaboradores são indiferentes em criticar pessoas em cargos de Liderança.

## 3.9 Fico tenso quando estou trabalhando

# Total

```
df["P 09: Fico tenso quando estou trabalhando."].value_counts()
```



```
P 09: Fico tenso quando estou trabalhando.
```

```
2 - Discordo                2725
```

```
3 - Indiferente             1798
```

```
1 - Discordo totalmente     1066
```

```
4 - Concordo                780
```

```
0 - Sem resposta            684
```

```
5 - Concordo totalmente     187
```

```
Name: count, dtype: int64
```

```
# Frequencia relativa
```

```
df["P 09: Fico tenso quando estou trabalhando."].value_counts() /  
sum(df["P 09: Fico tenso quando estou trabalhando."].value_counts())
```

```
P 09: Fico tenso quando estou trabalhando.
```

```
2 - Discordo                0.376381
```

```
3 - Indiferente             0.248343
```

```
1 - Discordo totalmente     0.147238
```

```
4 - Concordo                0.107735
```

```
0 - Sem resposta            0.094475
```

```
5 - Concordo totalmente     0.025829
```

```
Name: count, dtype: float64
```

```
plt.title("Tensão enquanto trabalha")
```

```
sns.countplot(  
    data=df,
```

```
    x="P 09: Fico tenso quando estou trabalhando.",  
    hue="P 09: Fico tenso quando estou trabalhando.",
```

```
    palette="Set2",  
    legend=False
```

```
)
```

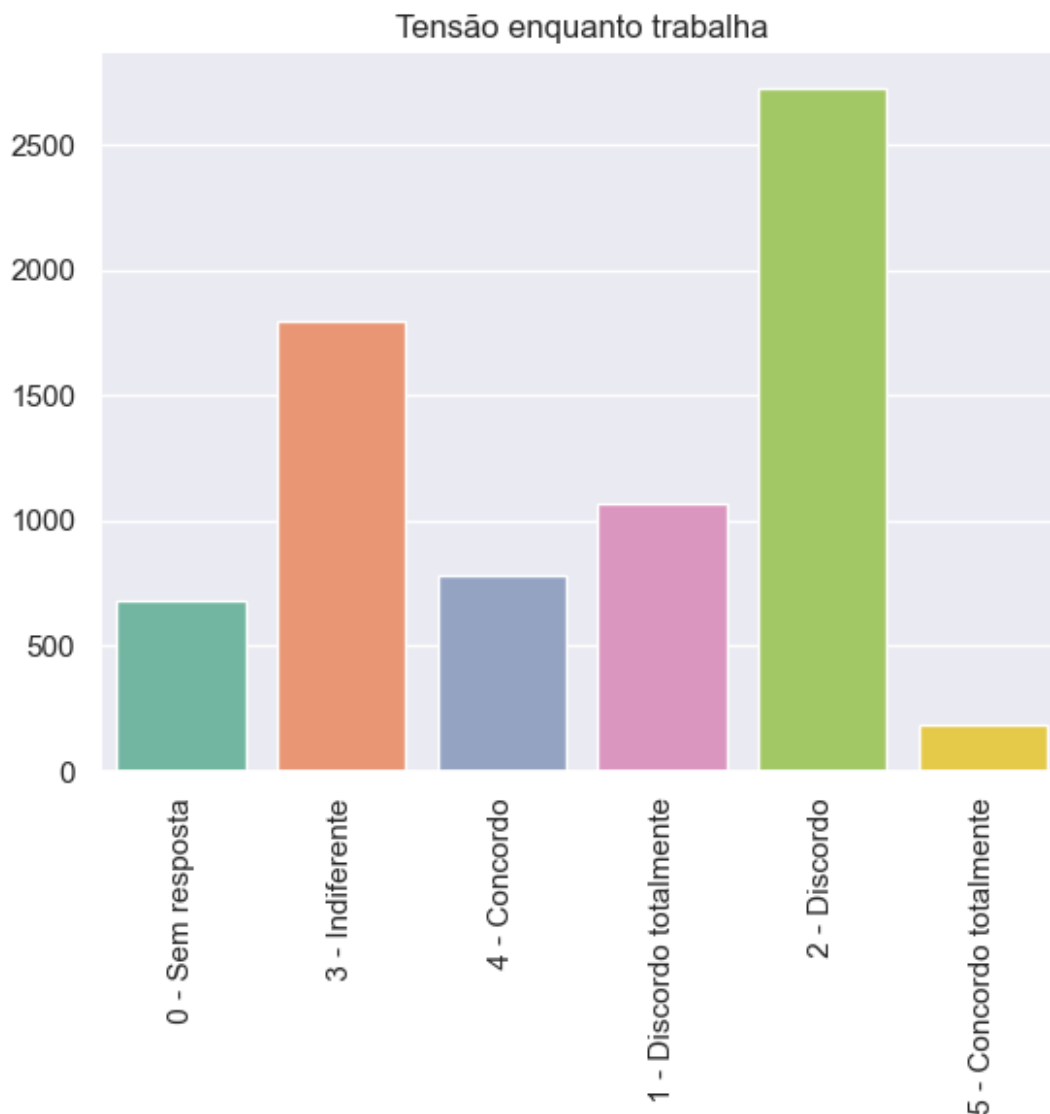
```
plt.xticks(rotation=90)
```

```
plt.xlabel('')
```

```
plt.ylabel('')
```

```
sns.despine()
```

```
plt.show()
```



Obtemos que:

51% dos colaboradores discordam ou discordam totalmente que ficam tensos enquanto trabalham.  
24% dos colaboradores indiferem que ficam tensos enquanto trabalham.

### 3.10 Não gosto de situações que requerem que eu me comporte de forma controlada

# Total

```
df["P 10: Não gosto de situações que requerem que eu me comporte de forma controlada."].value_counts()
```

```
P 10: Não gosto de situações que requerem que eu me comporte de forma controlada.
```

3 - Indiferente	2282
2 - Discordo	2115
1 - Discordo totalmente	1037
4 - Concordo	887
0 - Sem resposta	684
5 - Concordo totalmente	235

```
Name: count, dtype: int64
```

```
# Frequencia relativa
```

```
df["P 10: Não gosto de situações que requerem que eu me comporte de forma controlada."].value_counts() / sum(df["P 10: Não gosto de situações que requerem que eu me comporte de forma controlada."].value_counts())
```

```
P 10: Não gosto de situações que requerem que eu me comporte de forma controlada.
```

3 - Indiferente	0.315193
2 - Discordo	0.292127
1 - Discordo totalmente	0.143232
4 - Concordo	0.122514
0 - Sem resposta	0.094475
5 - Concordo totalmente	0.032459

```
Name: count, dtype: float64
```

```
plt.title(" Forma de trabalho referente a controle")
```

```
sns.countplot(  
    data=df,  
    x="P 10: Não gosto de situações que requerem que eu me comporte de forma controlada.",  
    hue="P 10: Não gosto de situações que requerem que eu me comporte de forma controlada.",  
    palette="Set2",  
    legend=False  
)
```

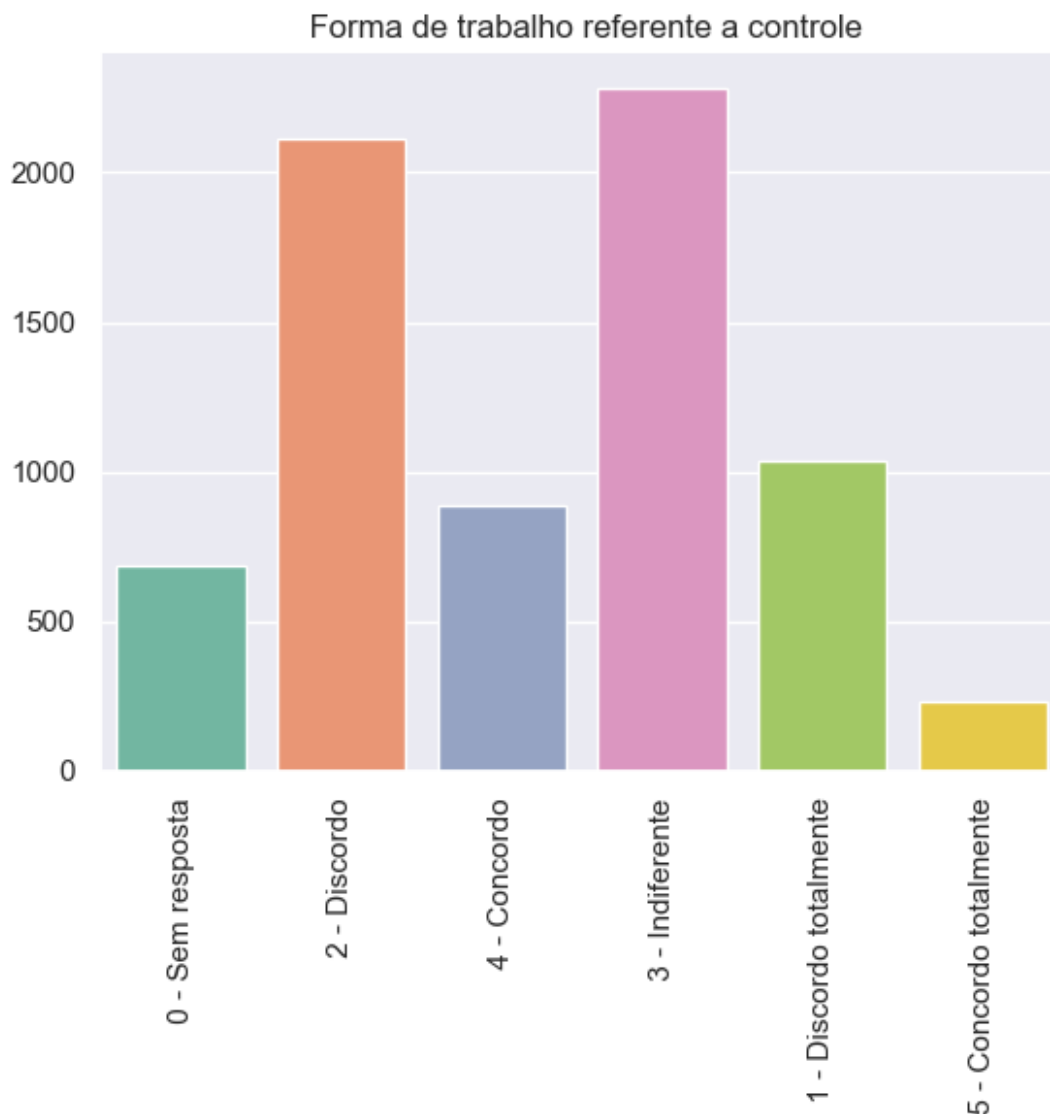
```
plt.xticks(rotation=90)
```

```
plt.xlabel('')
```

```
plt.ylabel('')
```

```
sns.despine()
```

```
plt.show()
```



Obtemos que:

43% dos colaboradores discordam ou discordam totalmente que não gostam de situações que requerem que se comportem de forma controlada. 31% dos colaboradores indiferem que não gostam de situações que requerem que se comportem de forma controlada.

## 3.11 Gosto de dizer o que penso.

# Total

```
df["P 11: Gosto de dizer o que penso."].value_counts()
```

```
P 11: Gosto de dizer o que penso.
4 - Concordo          3482
3 - Indiferente       1417
```

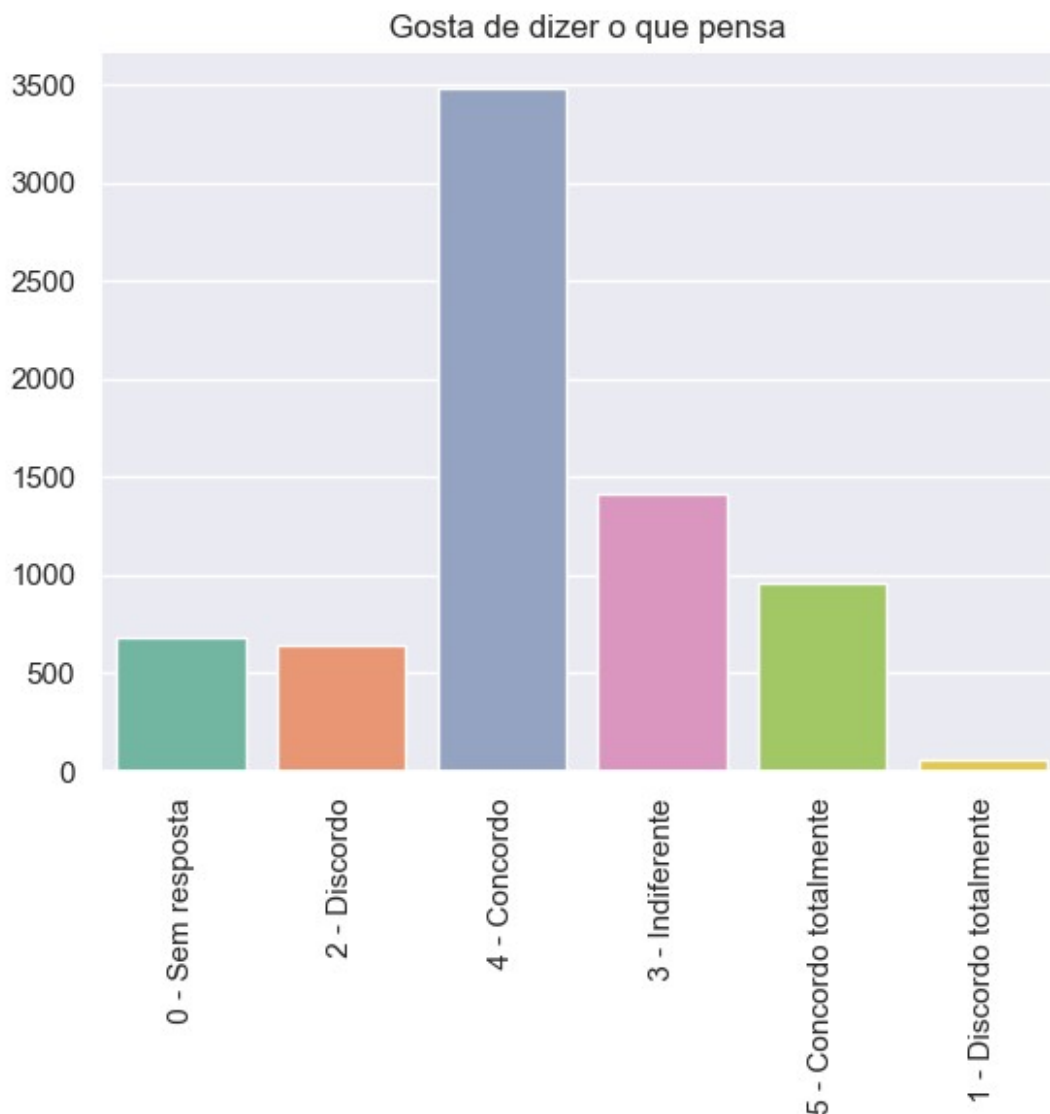
```
5 - Concordo totalmente      957
0 - Sem resposta             684
2 - Discordo                 641
1 - Discordo totalmente      59
Name: count, dtype: int64
```

```
# Frequencia relativa
```

```
df["P 11: Gosto de dizer o que penso."].value_counts() / sum(df["P 11:
Gosto de dizer o que penso."].value_counts())
```

```
P 11: Gosto de dizer o que penso.
4 - Concordo                0.480939
3 - Indiferente             0.195718
5 - Concordo totalmente     0.132182
0 - Sem resposta            0.094475
2 - Discordo                0.088536
1 - Discordo totalmente     0.008149
Name: count, dtype: float64
```

```
plt.title(" Gosta de dizer o que pensa")
sns.countplot(
    data=df,
    x="P 11: Gosto de dizer o que penso.",
    hue="P 11: Gosto de dizer o que penso.",
    palette="Set2",
    legend=False
)
plt.xticks(rotation=90)
plt.xlabel('')
plt.ylabel('')
sns.despine()
plt.show()
```



Obtemos que:

61% dos colaboradores concordam ou concordam totalmente que gostam de dizer o que pensam. 19% dos colaboradores indiferem que gostam de dizer o que pensam.

### 3.12 Gosto de fazer as coisas do meu jeito e não ligo para o que possam pensar.

```
# Total
```

```
df["P 12: Gosto de fazer as coisas do meu jeito e não ligo para o que  
possam pensar."].value_counts()
```

```
P 12: Gosto de fazer as coisas do meu jeito e não ligo para o que possam pensar.
```

```
2 - Discordo          3335
1 - Discordo totalmente 1456
3 - Indiferente        1134
0 - Sem resposta       684
4 - Concordo           507
5 - Concordo totalmente 124
```

```
Name: count, dtype: int64
```

```
# Frequencia relativa
```

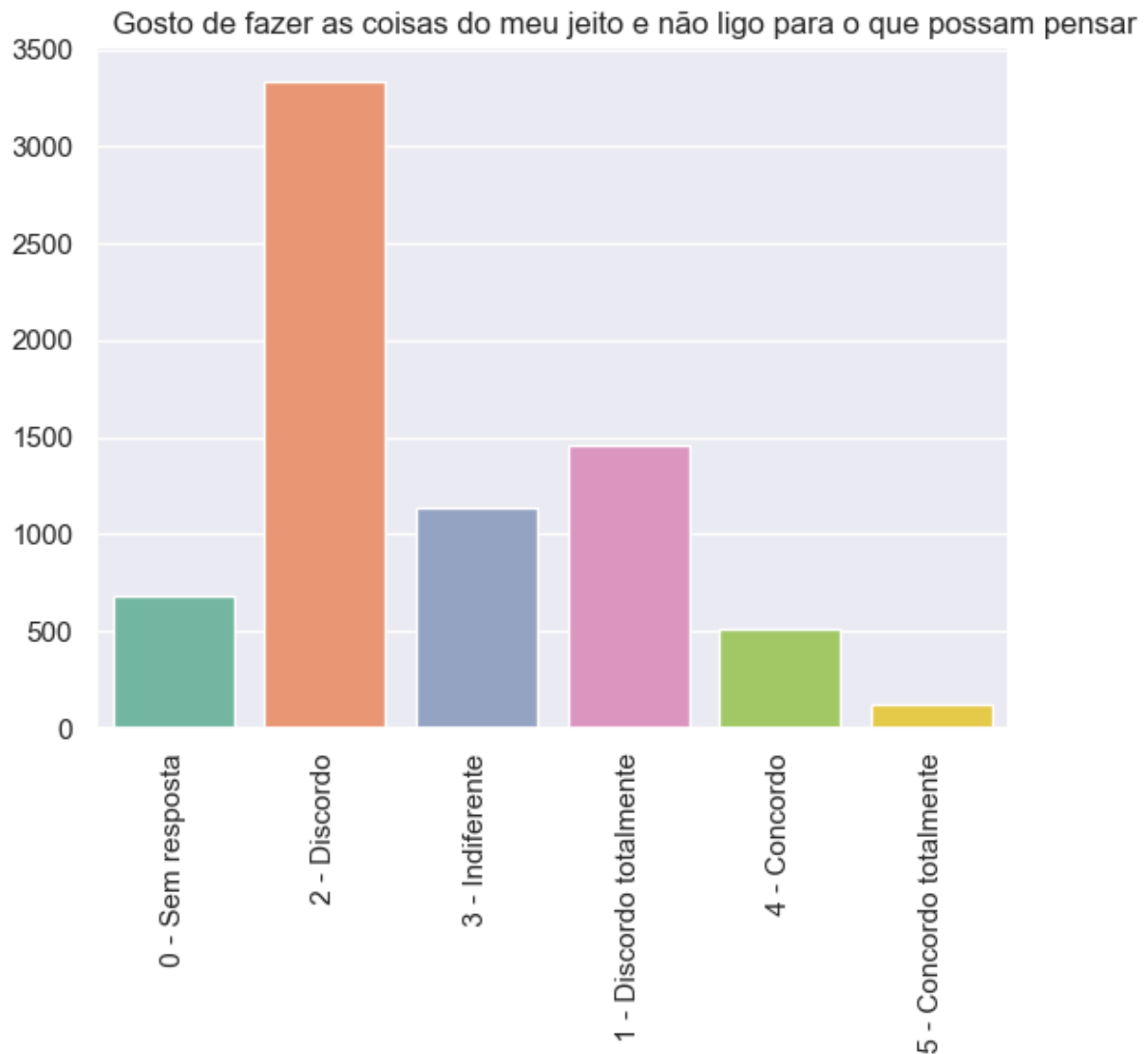
```
df["P 12: Gosto de fazer as coisas do meu jeito e não ligo para o que possam pensar."].value_counts() / sum(df["P 12: Gosto de fazer as coisas do meu jeito e não ligo para o que possam pensar."].value_counts())
```

```
P 12: Gosto de fazer as coisas do meu jeito e não ligo para o que possam pensar.
```

```
2 - Discordo          0.460635
1 - Discordo totalmente 0.201105
3 - Indiferente        0.156630
0 - Sem resposta       0.094475
4 - Concordo           0.070028
5 - Concordo totalmente 0.017127
```

```
Name: count, dtype: float64
```

```
plt.title("Gosto de fazer as coisas do meu jeito e não ligo para o que possam pensar")
sns.countplot(
    data=df,
    x="P 12: Gosto de fazer as coisas do meu jeito e não ligo para o que possam pensar.",
    hue="P 12: Gosto de fazer as coisas do meu jeito e não ligo para o que possam pensar.",
    palette="Set2",
    legend=False
)
plt.xticks(rotation=90)
plt.xlabel('')
plt.ylabel('')
sns.despine()
plt.show()
```



Obtemos que:

66% dos colaboradores discordam ou discordam totalmente que não gostam de fazer as coisas do seu jeito e não ligam para o que possam pensar. 15% dos colaboradores indiferem que gostam de fazer as coisas do seu jeito e não ligam para o que possam pensar.

### 3.13 Boa\_Performance .

```
# Total
```

```
df["Boa_Performance"].value_counts()
```

```
Boa_Performance
```

```
1      4089
```



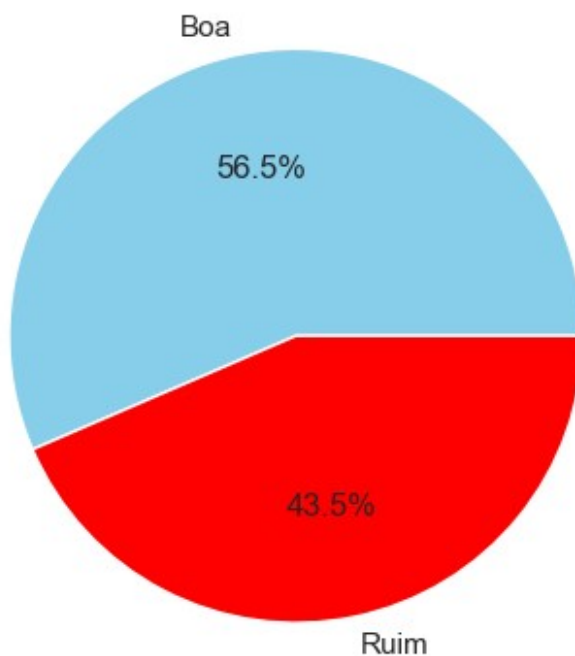
```
0    3151
Name: count, dtype: int64

plt.title("Performance apos 6 meses de empresa")

valores = df["Boa_Performance"].value_counts()

plt.pie(valores, labels=["Boa", "Ruim"], colors=["skyblue", "red"],
autopct='%1.1f%%')
plt.show()
```

Performance apos 6 meses de empresa



No conjunto de dados, contém uma Performance ruim de 43.5%, na análise bivariada irei procurar entender o que está associado a esse grande número.

## 4 Análise Bivariada

### Cálculo do Information Value

O cálculo do Information Value (IV), uma técnica estatística utilizada para avaliar o poder preditivo de variáveis independentes em relação a uma variável de resultado.

Quanto maior o IV de uma variável, maior é a sua capacidade de discriminar entre as categorias da variável resultado.

Os valores de IV podem ser interpretados da seguinte forma:

<b>IV Total</b>	<b>Poder de separação</b>
$< 0,02$	Muito fraco
0,02 a 0,1	Fraco
0,1 a 0,3	Médio
0,3 a 0,5	Forte
$> 0,5$	Muito bom pra ser verdade...Verifique!

IV Total	Poder de separação
< 0,02	Muito fraco
0,02 a 0,1	Fraco
0,1 a 0,3	Médio
0,3 a 0,5	Forte
> 0,5	Muito bom pra ser verdade...Verifique!

```
values_iv = sc.iv(df, "Boa_Performance")
values_iv
```

	variable	info_value
3	P 02: Qual área de atendimento você atua ou at...	0.033429
1	P 07: É bom trabalhar com meus colegas de trab...	0.026451
10	P 01: Antes de trabalhar na empresa, você tinh...	0.025222
0	P 06: Você possui dependentes?	0.022636
9	P 03: Qual a sua escolaridade?	0.015219
2	P 05: Qual o tempo gasto no trajeto da sua res...	0.011531
11	Id_Agente	0.010905
8	P 11: Gosto de dizer o que penso.	0.008098
6	P 08: Não tenho medo de criticar pessoas em ca...	0.008058
12	P 04: Qual seu conhecimento em inglês?	0.005125
4	P 12: Gosto de fazer as coisas do meu jeito e ...	0.004003

```

7          P 09: Fico tenso quando estou trabalhando.      0.003282
5    P 10: Não gosto de situações que requerem que ...      0.003200

# Função para calcular WoE, IV e Probabilidade de Boa_Performance
def calculate_woe_iv_prob(df, feature, target):
    # Agrupa os dados por categoria da variável
    lst = []
    for i in range(df[feature].nunique()):
        val = list(df[feature].unique())[i]
        lst.append({
            'Value': val,
            'All': df[df[feature] == val].count()[feature],
            'Good': df[(df[feature] == val) & (df[target] ==
1)]]).count()[feature],
            'Bad': df[(df[feature] == val) & (df[target] ==
0)]]).count()[feature]
        })
    dset = pd.DataFrame(lst)

    # Calcula as proporções e o WoE
    dset['Distr_Good'] = dset['Good'] / dset['Good'].sum()
    dset['Distr_Bad'] = dset['Bad'] / dset['Bad'].sum()
    dset['WoE'] = np.log(dset['Distr_Good'] / dset['Distr_Bad'])

    # Calcula a probabilidade de Boa_Performance = 1 (em porcentagem)
    dset['Prob_Boa_Performance'] = (dset['Good'] / (dset['Good'] +
dset['Bad'])) * 100).round(2).astype(str) + '%'

    # Formata as colunas Distr_Good e Distr_Bad como porcentagens
    dset['Distr_Good'] = (dset['Distr_Good'] *
100).round(2).astype(str) + '%'
    dset['Distr_Bad'] = (dset['Distr_Bad'] * 100).round(2).astype(str)
+ '%'

    # Calcula o IV
    dset['IV'] = (dset['Good'] / dset['Good'].sum() - dset['Bad'] /
dset['Bad'].sum()) * dset['WoE']
    iv = dset['IV'].sum()

    return dset, iv

# Carregar o arquivo
df = pd.read_excel('base_atendimento.xlsx', sheet_name='Base')

# Lista de variáveis categóricas (excluindo a variável alvo e
variáveis numéricas)
categorical_vars = [
    'P 01: Antes de trabalhar na empresa, você tinha experiência em
Contact Center?',
    'P 02: Qual área de atendimento você atua ou atuou por mais

```

```

tempo?',
    'P 03: Qual a sua escolaridade?',
    'P 04: Qual seu conhecimento em inglês?',
    'P 05: Qual o tempo gasto no trajeto da sua residência até a
empresa?',
    'P 06: Você possui dependentes?',
    'P 07: É bom trabalhar com meus colegas de trabalho.',
    'P 08: Não tenho medo de criticar pessoas em cargos de
liderança.',
    'P 09: Fico tenso quando estou trabalhando.',
    'P 10: Não gosto de situações que requerem que eu me comporte de
forma controlada.',
    'P 11: Gosto de dizer o que penso.',
    'P 12: Gosto de fazer as coisas do meu jeito e não ligo para o que
possam pensar.'
]

```

*# Dicionário para armazenar os resultados*

```
results = {}
```

*# Calcular WoE, IV e Probabilidade de Boa\_Performance para cada
variável categórica*

```

for var in categorical_vars:
    woe_iv_prob_df, iv = calculate_woe_iv_prob(df, var,
'Boa_Performance')
    results[var] = {
        'WoE_Table': woe_iv_prob_df,
        'IV': iv
    }

```

*# Exibir os resultados*

```

for var, result in results.items():
    print(f"Variável: {var}")
    print(result['WoE_Table'])
    print(f"Information Value (IV): {result['IV']}")
    print("\n" + "="*50 + "\n")

```

Variável: P 01: Antes de trabalhar na empresa, você tinha experiência em Contact Center?

	Value	All	Good	Bad	Distr_Good
Distr_Bad \					
0	Não	2380	1389	991	33.97%
31.45%					
1	Sim	3651	1938	1713	47.4%
54.36%					
2	A empresa é o meu primeiro emprego	1038	664	374	16.24%
11.87%					
3	Sem resposta	171	98	73	2.4%
2.32%					

	WoE	Prob_Boa_Performance	IV
0	0.077044	58.36%	0.001941
1	-0.137170	53.08%	0.009558
2	0.313446	63.97%	0.013696
3	0.033927	57.31%	0.000027

Information Value (IV): 0.02522194279387866

Variável: P 02: Qual área de atendimento você atua ou atuou por mais tempo?

		Value	All	Good	Bad	Distr_Good	Distr_Bad
0	Sem resposta	3613	2161	1452	52.85%	46.08%	
1	Atendimento Receptivo	1487	770	717	18.83%	22.75%	
2	Ativo de Vendas	353	192	161	4.7%	5.11%	
3	Receptivo de Vendas	90	47	43	1.15%	1.36%	
4	Chat ou email	265	148	117	3.62%	3.71%	
5	Suporte Técnico	288	138	150	3.37%	4.76%	
6	Cobrança	320	167	153	4.08%	4.86%	
7	Retenção	99	39	60	0.95%	1.9%	
8	Backoffice	473	284	189	6.95%	6.0%	
9	Multiskill	226	124	102	3.03%	3.24%	
10	Suporte Técnico com Vendas	26	19	7	0.46%	0.22%	

	WoE	Prob_Boa_Performance	IV
0	0.137049	59.81%	0.009276
1	-0.189266	51.78%	0.007426
2	-0.084490	54.39%	0.000350
3	-0.171633	52.22%	0.000369
4	-0.025542	55.85%	0.000024
5	-0.343962	47.92%	0.004766
6	-0.173025	52.19%	0.001335
7	-0.691363	39.39%	0.006571
8	0.146647	60.04%	0.001389
9	-0.065272	54.87%	0.000134
10	0.737948	73.08%	0.001790

Information Value (IV): 0.033428710596706615

Variável: P 03: Qual a sua escolaridade?

	Value	All	Good	Bad	Distr_Good	Distr_Bad
0	Ensino Superior incompleto	966	530	436	12.96%	13.84%
1	Ensino Médio completo	2880	1566	1314	38.3%	41.7%
2	Ensino Superior cursando	1799	1004	795	24.55%	25.23%
3	Ensino Superior concluído	1237	766	471	18.73%	14.95%
4	Pós Graduação cursando	137	87	50	2.13%	1.59%
5	Pós Graduação concluído	129	81	48	1.98%	1.52%
6	Sem resposta	92	55	37	1.35%	1.17%

	WoE	Prob_Boa_Performance	IV
0	-0.065346	54.87%	0.000572
1	-0.085132	54.37%	0.002897
2	-0.027175	55.81%	0.000184
3	0.225743	61.92%	0.008546
4	0.293305	63.5%	0.001586
5	0.262668	62.79%	0.001202
6	0.135835	59.78%	0.000232

Information Value (IV): 0.015218989705983533

=====

Variável: P 04: Qual seu conhecimento em inglês?

	Value	All	Good	Bad	Distr_Good
0	Intermediário	1359	782	577	19.12%
1	Não tenho conhecimento em inglês	1406	763	643	18.66%
2	Básico	3668	2058	1610	50.33%
3	Avançado	463	281	182	6.87%
4	Sem resposta	171	98	73	2.4%
5	Fluente	173	107	66	2.62%

	WoE	Prob_Boa_Performance	IV
0	0.043432	57.54%	0.000353
1	-0.089467	54.27%	0.001562



2	-0.015080	56.11%	0.000115
3	0.173767	60.69%	0.001905
4	0.033927	57.31%	0.000027
5	0.222594	61.85%	0.001162

Information Value (IV): 0.005125079942321091

=====

Variável: P 05: Qual o tempo gasto no trajeto da sua residência até a empresa?

	Value	All	Good	Bad	Distr_Good	Distr_Bad	WoE \
0	Sem resposta	5033	2846	2187	69.6%	69.41%	0.002803
1	De 1h à 1h30	712	408	304	9.98%	9.65%	0.033659
2	Acima de 2h	118	53	65	1.3%	2.06%	-0.464676
3	Até 30min	314	199	115	4.87%	3.65%	0.287792
4	De 30min à 1h	540	317	223	7.75%	7.08%	0.091149
5	De 1h30 à 2h	523	266	257	6.51%	8.16%	-0.226160

	Prob_Boa_Performance	IV
0	56.55%	0.000005
1	57.3%	0.000111
2	44.92%	0.003563
3	63.38%	0.003503
4	58.7%	0.000616
5	50.86%	0.003734

Information Value (IV): 0.011531102462205949

=====

Variável: P 06: Você possui dependentes?

	Value	All	Good	Bad	Distr_Good	Distr_Bad	WoE \
0	Não	4838	2854	1984	69.8%	62.96%	0.103026
1	Sim	2269	1157	1112	28.3%	35.29%	-0.220910
2	Sem resposta	133	78	55	1.91%	1.75%	0.088795

	Prob_Boa_Performance	IV
0	58.99%	0.007040
1	50.99%	0.015453
2	58.65%	0.000144

Information Value (IV): 0.022636137469178152

=====

Variável: P 07: É bom trabalhar com meus colegas de trabalho.

	Value	All	Good	Bad	Distr_Good	Distr_Bad	WoE \
0	0 - Sem resposta	684	392	292	9.59%	9.27%	0.033927
1	4 - Concordo	2687	1478	1209	36.15%	38.37%	-0.059684

2	5 - Concordo totalmente	3296	1961	1335	47.96%	42.37%
0.123943						
3	3 - Indiferente	514	235	279	5.75%	8.85% -
0.432207						
4	1 - Discordo totalmente	14	7	7	0.17%	0.22% -
0.260581						
5	2 - Discordo	45	16	29	0.39%	0.92% -
0.855288						

	Prob_Boa_Performance	IV
0	57.31%	0.000108
1	55.01%	0.001327
2	59.5%	0.006929
3	45.72%	0.013430
4	50.0%	0.000133
5	35.56%	0.004525

Information Value (IV): 0.026451463258253885

=====

Variável: P 08: Não tenho medo de criticar pessoas em cargos de liderança.

		Value	All	Good	Bad	Distr_Good	Distr_Bad
WoE \							
0	0 - Sem resposta	684	392	292	9.59%	9.27%	
0.033927							
1	3 - Indiferente	1853	1091	762	26.68%	24.18%	
0.098323							
2	4 - Concordo	1744	982	762	24.02%	24.18% -	
0.006936							
3	1 - Discordo totalmente	668	372	296	9.1%	9.39% -	
0.032046							
4	5 - Concordo totalmente	460	227	233	5.55%	7.39% -	
0.286669							
5	2 - Discordo	1831	1025	806	25.07%	25.58% -	
0.020216							

	Prob_Boa_Performance	IV
0	57.31%	0.000108
1	58.88%	0.002457
2	56.31%	0.000012
3	55.69%	0.000095
4	49.35%	0.005283
5	55.98%	0.000103

Information Value (IV): 0.008058468881502854

=====

Variável: P 09: Fico tenso quando estou trabalhando.

	Value	All	Good	Bad	Distr_Good	Distr_Bad
--	-------	-----	------	-----	------------	-----------

WoE \						
0	0 - Sem resposta	684	392	292	9.59%	9.27%
0.033927						
1	3 - Indiferente	1798	1023	775	25.02%	24.6%
0.017051						
2	4 - Concordo	780	440	340	10.76%	10.79% -
0.002751						
3	1 - Discordo totalmente	1066	610	456	14.92%	14.47%
0.030386						
4	2 - Discordo	2725	1534	1191	37.52%	37.8% -
0.007495						
5	5 - Concordo totalmente	187	90	97	2.2%	3.08% -
0.335482						

	Prob_Boa_Performance	IV
0	57.31%	1.084989e-04
1	56.9%	7.212219e-05
2	56.41%	8.157602e-07
3	57.22%	1.356644e-04
4	56.29%	2.115433e-05
5	48.13%	2.943386e-03

Information Value (IV): 0.003281641717049869

=====

Variável: P 10: Não gosto de situações que requerem que eu me comporte de forma controlada.

		Value	All	Good	Bad	Distr_Good	Distr_Bad
WoE \							
0	0 - Sem resposta	684	392	292	9.59%	9.27%	
0.033927							
1	2 - Discordo	2115	1188	927	29.05%	29.42%	-
0.012508							
2	4 - Concordo	887	499	388	12.2%	12.31%	-
0.008980							
3	3 - Indiferente	2282	1313	969	32.11%	30.75%	
0.043225							
4	1 - Discordo totalmente	1037	580	457	14.18%	14.5%	-
0.022236							
5	5 - Concordo totalmente	235	117	118	2.86%	3.74%	-
0.269091							

	Prob_Boa_Performance	IV
0	57.31%	0.000108
1	56.17%	0.000046
2	56.26%	0.000010
3	57.54%	0.000587
4	55.93%	0.000071
5	49.79%	0.002377

Information Value (IV): 0.003199644087812544

Variável: P 11: Gosto de dizer o que penso.

	Value	All	Good	Bad	Distr_Good	Distr_Bad
WoE \						
0	0 - Sem resposta	684	392	292	9.59%	9.27%
0.033927						
1	2 - Discordo	641	348	293	8.51%	9.3% -
0.088551						
2	4 - Concordo	3482	1974	1508	48.28%	47.86%
0.008697						
3	3 - Indiferente	1417	836	581	20.45%	18.44%
0.103297						
4	5 - Concordo totalmente	957	500	457	12.23%	14.5% -
0.170656						
5	1 - Discordo totalmente	59	39	20	0.95%	0.63%
0.407249						

	Prob_Boa_Performance	IV
0	57.31%	0.000108
1	54.29%	0.000698
2	56.69%	0.000036
3	59.0%	0.002073
4	52.25%	0.003883
5	66.1%	0.001299

Information Value (IV): 0.008097773935183485

Variável: P 12: Gosto de fazer as coisas do meu jeito e não ligo para o que possam pensar.

	Value	All	Good	Bad	Distr_Good	Distr_Bad
WoE \						
0	0 - Sem resposta	684	392	292	9.59%	9.27%
0.033927						
1	2 - Discordo	3335	1842	1493	45.05%	47.38% -
0.050516						
2	3 - Indiferente	1134	659	475	16.12%	15.07%
0.066828						
3	1 - Discordo totalmente	1456	845	611	20.67%	19.39%
0.063659						
4	4 - Concordo	507	289	218	7.07%	6.92%
0.021351						
5	5 - Concordo totalmente	124	62	62	1.52%	1.97% -
0.260581						

	Prob_Boa_Performance	IV
0	57.31%	0.000108
1	55.23%	0.001179

2	58.11%	0.000696
3	58.04%	0.000811
4	57.0%	0.000032
5	50.0%	0.001176

Information Value (IV): 0.00400323441321808

=====

## Conclusão

Os Maiores IV foram da seguintes Categorias:P 02: Qual área de atendimento você atua ou atuou por mais tempo?;P 07: É bom trabalhar com meus colegas de trabalho.;P 01: Antes de trabalhar na empresa, você tinha experiência em Contact Center? e P 06: Você possui dependentes? e com isso foram retirados os seguintes Insights:

Entre os colaboradores com experiência, destaquei as experiências em setores que tiveram a maior taxa de boa performance: Backoffice, Ativo de Vendas, Chat ou email, Multiskill e Suporte Técnico c/ Vendas. Observei ainda os setores com a menor taxa de boa performance como Suporte Técnico e Retenção.

Focar em colaboradores que gostam de trabalho em equipe, visto que a Taxa de Boa Performance nesse grupo, foi de 55% (concordo) e 59,5% (concordo plenamente). O IV da variável "Gosta de trabalhar em equipe", em relação a boa performance foi de 0,029.

Na variável "Possui Dependentes", como já foi mostrado anteriormente, os colaboradores que não possuem dependentes, representam 66,8% dos colaboradores e nessa análise podemos ver que ela apresenta ainda uma melhor taxa de boa performance (59%), em relação a quem possui dependentes (51%). O IV dessa variável em relação a boa performance é de 0,023.

## Respondendo as Perguntas do Negócio

Quais fatores o time de Atração de Talentos deve olhar para selecionar candidatos com maior chance de ter boa performance após 6 meses à data da contratação?

1. Experiência em Setores com Alta Taxa de Boa Performance Setores com maior taxa de boa performance:

Backoffice

Ativo de Vendas

Chat ou email

Multiskill

Suporte Técnico com Vendas

Ação sugerida:

Priorizar candidatos com experiência nesses setores, pois eles apresentam maior probabilidade de sucesso.

Durante o processo de recrutamento, avaliar a experiência prévia do candidato e dar preferência àqueles que já atuaram em áreas com alta taxa de boa performance.

1. Evitar Setores com Baixa Taxa de Boa Performance Setores com menor taxa de boa performance:

Suporte Técnico (sem vendas)

Retenção

Ação sugerida:

Reduzir o foco em candidatos com experiência predominante nesses setores, a menos que haja outros fatores compensatórios (como habilidades comportamentais ou técnicas relevantes).

1. Preferência por Colaboradores que Gostam de Trabalho em Equipe Dados:

Colaboradores que concordam ou concordam plenamente que gostam de trabalhar em equipe têm taxas de boa performance de 55% e 59,5%, respectivamente.

O IV da variável "Gosta de trabalhar em equipe" é 0,029, indicando uma influência positiva, embora moderada.

Ação sugerida:

Incluir perguntas ou dinâmicas no processo seletivo para avaliar a afinidade do candidato com trabalho em equipe.

Priorizar candidatos que demonstram habilidades de colaboração e preferência por ambientes colaborativos.

1. Colaboradores sem Dependentes Dados:

Colaboradores sem dependentes representam 66,8% do total e têm uma taxa de boa performance de 59%, comparado a 51% para aqueles com dependentes.

O IV da variável "Possui Dependentes" é 0,023, indicando uma influência positiva, mas menor.

Ação sugerida:

Dar preferência a candidatos sem dependentes, pois eles apresentam uma taxa de boa performance ligeiramente maior.

No entanto, esse fator deve ser considerado com cautela, pois pode envolver questões éticas e de diversidade.

1. Avaliação de Habilidades Comportamentais Dados:

Colaboradores que gostam de trabalhar em equipe e não possuem dependentes têm maior probabilidade de boa performance.

Ação sugerida:

Incluir avaliações comportamentais no processo seletivo, como testes de perfil comportamental ou dinâmicas de grupo, para identificar candidatos com essas características.

Focar em competências como colaboração, resiliência e adaptabilidade.

## Com base nisso, qual política de contratação deve passar a valer de forma que a empresa aumente as boas contratações?

Estruturando a política de contratação utilizando o método 5W2H, que é uma ferramenta eficaz para planejamento e execução de ações. O método consiste em responder às seguintes perguntas:

What (O que será feito?)

Why (Por que será feito?)

Where (Onde será feito?)

When (Quando será feito?)

Who (Quem fará?)

How (Como será feito?)

How much (Quanto custará?)

Aqui está a política de contratação estruturada com o método 5W2H:

### Política de Contratação para Aumentar as Boas Contratações

1. What (O que será feito?) Implementar uma política de contratação focada em selecionar candidatos com maior probabilidade de boa performance após 6 meses, com base em dados e análises de desempenho.
2. Why (Por que será feito?) Objetivo: Aumentar a taxa de boas contratações, reduzir o turnover e melhorar a produtividade da equipe.

Justificativa: A análise dos dados mostrou que colaboradores com experiência em setores específicos (Backoffice, Ativo de Vendas, Chat ou email, Multiskill e Suporte Técnico com Vendas), que gostam de trabalhar em equipe e não possuem dependentes, têm maior probabilidade de boa performance.

1. Where (Onde será feito?) Local: Em todos os processos de recrutamento e seleção da empresa, tanto para vagas presenciais quanto remotas.

Setores prioritários: Backoffice, Ativo de Vendas, Chat ou email, Multiskill e Suporte Técnico com Vendas.

1. When (Quando será feito?) Início: Imediato, a partir da próxima rodada de contratações.

Frequência: Contínuo, com revisões trimestrais para ajustar os critérios com base em novos dados de performance.

1. Who (Quem fará?) Responsáveis:

Equipe de Recrutamento e Seleção: Aplicar os novos critérios de seleção.

Gestores de Área: Avaliar a adequação dos candidatos às necessidades específicas dos setores.

Analistas de Dados: Monitorar e atualizar os indicadores de performance e IV das variáveis.

Envolvidos:

Candidatos: Serão avaliados com base nos novos critérios.

Líderes de equipe: Darão feedback sobre o desempenho dos novos contratados.

1. How (Como será feito?) Passo a Passo:

Definir critérios de seleção:

Priorizar candidatos com experiência em setores de alta performance (Backoffice, Ativo de Vendas, Chat ou email, Multiskill e Suporte Técnico com Vendas).

Avaliar a afinidade com trabalho em equipe por meio de perguntas comportamentais e dinâmicas de grupo.

Considerar a ausência de dependentes como um fator adicional, mas não exclusivo.

Incluir avaliações comportamentais:

Utilizar testes de perfil comportamental para identificar colaboradores com habilidades como colaboração, resiliência e adaptabilidade.

Avaliar conhecimentos técnicos:

Realizar testes práticos ou entrevistas técnicas para verificar o domínio de ferramentas e processos relevantes.

Monitorar o desempenho:

Acompanhar a performance dos novos contratados após 6 meses para validar os critérios de seleção.

Ajustar os critérios:

Revisar trimestralmente os dados de performance e atualizar os critérios de seleção com base nos novos insights.

1. How much (Quanto custará?) Investimentos necessários:

Ferramentas de avaliação comportamental: R 5.000 a R 5.000aR 10.000 por ano (dependendo do fornecedor).

Treinamento da equipe de recrutamento: R 2.000 a R 2.000aR 5.000 (para capacitação em técnicas de seleção baseadas em dados).



Tempo da equipe: Aproximadamente 20 horas por semana dedicadas à implementação e monitoramento da nova política.

Retorno esperado:

Aumento da taxa de boas contratações em 10% a 20%.

Redução do turnover em 15% a 25%.

Melhoria na produtividade geral da equipe.

Resumo da Política de Contratação A política de contratação proposta visa selecionar candidatos com maior probabilidade de boa performance após 6 meses, com base em dados e análises. Para isso, serão priorizados candidatos com experiência em setores de alta performance, habilidades comportamentais alinhadas (como trabalho em equipe) e características demográficas relevantes (como ausência de dependentes). A implementação será feita pela equipe de recrutamento, com monitoramento contínuo e ajustes trimestrais.

Benefícios Esperados: Aumento da taxa de boas contratações.

Redução do turnover.

Melhoria na produtividade e engajamento da equipe.

Tomada de decisão mais assertiva no recrutamento.

