

Meu Painel

Exercício de Revisão | Tratando valores discrepantes

QUIZ - 1 PERGUNTA

4.4 - Intervalo de confiança e distribuição amostral

VIDEO - 11 MIN

Exercício de Código | Intervalo de confiança e distribuição amostral

TEXT

4.5 - Featring engineering: criando novas variáveis para facilitar a análise (parte 1)

VIDEO - 18 MIN

4.5 - Featring engineering: criando novas variáveis para facilitar a análise (parte 2)

VIDEO - 10 MIN

Exercício de Código | Featring engineering

TEXT

Faça você mesmo | Featring engineering

TEXT

Exercício de revisão | Featring engineering

QUIZ - 1 PERGUNTA

4.6 - Correlação, diferentes funções para dados discretos e contínuos

VIDEO - 16 MIN

Exercício de Código | Featring engineering

5 COMENTÁRIOS/DÚVIDAS

Exercício de Código

Agora é a sua vez! Crie as colunas assim como a Instrutora Carol

Depois, confira o passo a passo / gabarito!

Parte 1

Crie a função:

Featring engineering

```
def preencher_nivel(gestor,nivel):
    if gestor==1:
        return "Pessoa Gestora"
    else:
        return nivel
```

Aplice a função:

```
[267] 1 dados['NOVO_NIVEL'] = dados.apply(lambda x: preencher_nivel(x['GESTOR'], x['NIVEL']), axis=1)
```

```
1 dados['NOVO_NIVEL'].value_counts()
```

| NOVO_NIVEL | count |
|----------------|-------|
| Pessoa | 1860 |
| Junior | 1813 |
| Sênior | 898 |
| Pessoa Gestora | 713 |

Nome: count, dtype: int64

Vamos fazer a conversão de variáveis categóricas:

```
[272] 1 dados = pd.get_dummies(dados, columns=['NIVEL'])
```

```
1 dados.columns
```

Index(['Unnamed: 0', 'ID', 'IDADE', 'FAIXA IDADE', 'GÊNERO', 'COR/RACA/ETNIA', 'PCD', 'EXPERIENCIA_PROFISSIONAL_PREVISTO/REAL', 'ASPECTOS_PREVISTO/REAL', 'VIVE_NO_BRASIL', 'ESTADO_ONDE Mora', 'UF ONDE Mora', 'PRECISA ONDE Mora', 'MODO DE ESTADO', 'MODO DE ORIENT', 'NIVEL DE ENSINO', 'ÁREA DE FORMAÇÃO', 'QUAL SUA SITUAÇÃO ATUAL DE TRABALHO', 'SETOR', 'NÚMERO DE FUNCIONÁRIOS', 'GESTÃO', 'CARGO COMO GESTOR', 'CARGO ATUAL', 'FAIXA SALARIAL', 'QUANTO TEMPO DE EXPERIÊNCIA NA ÁREA DE DADOS VOCÊ TEM?', 'QUANTO TEMPO DE EXPERIÊNCIA NA ÁREA DE TI/ENGENHARIA DE SOFTWARE VOCÊ TEM ANTES DE COMEÇAR A TRABALHAR NA ÁREA DE DADOS SALARIO', 'NOVO_NIVEL', 'NIVEL_Junior', 'NIVEL_Pessoa', 'NIVEL_Sênior'], dtype='object')

Vamos criar a coluna geração:

```
[274] 1 def determinar_geracao(idade):
2     if 39<idade<=58:
3         return "Geração X"
4     elif 29<idade<=39:
5         return 'Millenial'
6     elif 13<idade<=29:
7         return "Geração Z"
8     else:
9         return "Outra geração"
```

```
[275] 1 dados['GERACAO'] = dados['IDADE'].apply(determinar_geracao)
```

```
1 dados['GERACAO'].value_counts()
```

| GERACAO | count |
|-----------|-------|
| Geração Z | 1999 |
| Millenial | 1761 |
| Geração X | 511 |

Parte 2

Faça uma cópia da [planilha dessa aula](#)

Faça um cópia da planilha, salve no seu drive e procure:

Arquivos

programaria

arquivos

notebooks

teste

Fazer download

Renomear arquivo

Excluir arquivo

Copiar caminho

Atualizar

soma_ingredientes...

searchbooks

```
1 dados2 = pd.read_excel('/content/drive/MyDrive/programaria/planilha_parte2.xlsx')
```

Vamos juntar os dados em comum:

```
1 dados.merge(dados2, on='ID', how='left')
```

```
1 dados = dados.merge(dados2, on='ID', how='left')
```

Identifique a intenção da mudança de emprego:

```
1 dados['Você pretende mudar de emprego nos próximos 6 meses?'].value_counts()
```

| Você pretende mudar de emprego nos próximos 6 meses? | |
|---|------|
| Não estou buscando, mas me considero aberto a outras oportunidades | 1354 |
| Estou em busca de oportunidades dentro ou fora do Brasil | 1011 |
| Não estou buscando e não pretendo mudar de emprego nos próximos 6 meses | 978 |
| Estou em busca de oportunidades, mas apenas fora do Brasil | 353 |

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

```
1284 | dados['EM_BUSCA'] = dados['Você pretende mudar de emprego nos próximos 6 meses?'].str.contains('em busca', case=False)
```

```
1 dados['EM_BUSCA'].value_counts()
```

| EM_BUSCA | |
|----------|------|
| False | 2332 |
| True | 1364 |

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

Identifique quem está aberto à oportunidades:

```
1285 | dados['ABERTO_OPORTUNIDADES'] = dados['Você pretende mudar de emprego nos próximos 6 meses?'].str.contains('aberto', case=False)
```

```
1 dados['ABERTO_OPORTUNIDADES'].value_counts()
```

| ABERTO_OPORTUNIDADES | |
|----------------------|------|
| False | 2342 |
| True | 1354 |

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