

Análise Salarial de Desenvolvedores

- MySQL

- **Consulta de salário médio em USD por ano de trabalho e nível de experiencia:**

```
SELECT
  ano_de_trabalho,
  nivel_de_experiencia,
  CONCAT('U$ ', FORMAT(AVG(salario_em_usd), 2)) AS media_salarial,
  COUNT(*) AS total_funcionarios
FROM
  dataset_salary
WHERE
  salario_em_usd IS NOT NULL
GROUP BY
  ano_de_trabalho,
  nivel_de_experiencia
ORDER BY
  ano_de_trabalho,
  AVG(salario_em_usd) ASC;
```

ano_de_trabalho	nivel_de_experiencia	media_salarial	total_funcionari...
2020	Iniciante	U\$ 69,840.67	21
2020	Intermediario	U\$ 91,870.94	31
2020	Senior	U\$ 138,648.89	19
2020	Exevutivo	U\$ 179,958.25	4
2021	Iniciante	U\$ 63,463.07	46
2021	Intermediario	U\$ 86,818.39	87
2021	Senior	U\$ 125,989.75	75
2021	Exevutivo	U\$ 186,128.00	10
2022	Iniciante	U\$ 80,018.70	116
2022	Intermediario	U\$ 102,745.58	359
2022	Senior	U\$ 147,903.35	1139
2022	Exevutivo	U\$ 188,260.29	41
2023	Iniciante	U\$ 91,166.31	464
2023	Intermediario	U\$ 123,821.31	1716
2023	Senior	U\$ 165,421.03	6097
2023	Exevutivo	U\$ 191,314.19	242
2024	Iniciante	U\$ 97,882.79	678
2024	Intermediario	U\$ 134,804.00	1845
2024	Senior	U\$ 166,815.83	3340
2024	Exevutivo	U\$ 202,105.81	204

Na análise fornecida, percebemos que, com o decorrer dos anos, não só a quantidade de funcionários aumentou, mas o salário também cresceu.

- Cálculo de salário médio em USD para cada combinação de cargo e residencia_do_funcionario:

```
SELECT
  cargo,
  residencia_do_funcionario,
  CONCAT('U$', FORMAT(AVG(salario_em_usd), 2)) AS media_salarial,
  COUNT(*) AS total_funcionarios
FROM
  dataset_salary
WHERE
  salario_em_usd IS NOT NULL
GROUP BY
  cargo,
  residencia_do_funcionario
ORDER BY
  cargo;
```

cargo	residencia_do_funcionario	media_salarial	total_funcionari...
Admin & Data Analyst	DE	U\$ 50,000.00	1
Admin & Data Analyst	ES	U\$ 42,222.00	1
Admin & Data Analyst	IN	U\$ 60,000.00	1
Admin & Data Analyst	US	U\$ 95,000.00	1
AI Architect	BE	U\$ 215,936.00	1
AI Architect	CA	U\$ 800,000.00	1
AI Architect	HR	U\$ 180,000.00	1
AI Architect	US	U\$ 235,694.23	26
AI Developer	BA	U\$ 120,000.00	1
AI Developer	CA	U\$ 275,000.00	1
AI Developer	DE	U\$ 140,358.50	2
AI Developer	FI	U\$ 74,883.00	1
AI Developer	FR	U\$ 97,642.50	2
AI Developer	IT	U\$ 53,984.00	1
AI Developer	SE	U\$ 80,000.00	1

Neste cálculo, notamos que a média salarial varia de acordo com o local de residência do funcionário.

- Calculo da média da proporcao_remota agrupada por cargo:

```
SELECT
  cargo,
  FORMAT(AVG(proporcao_remota), 2) AS media_proporcao_remota,
  COUNT(*) AS total_vagas
FROM
  dataset_salary
GROUP BY
  cargo
ORDER BY
  AVG(proporcao_remota);
```

cargo	media_proporcao_rem...	total_vagas
Consultant Data Engineer	100.00	2
Manager Data Management	100.00	2
CRM Data Analyst	100.00	1
ETL Engineer	100.00	3
Data Strategy Manager	100.00	4
Principal Data Engineer	100.00	3
Power BI Developer	100.00	1
Machine Learning Specialist	100.00	2
Compliance Data Analyst	100.00	2
AWS Data Architect	100.00	1
BI Data Engineer	100.00	1
Marketing Data Analyst	100.00	2
Principal Data Architect	100.00	1
Principal Data Analyst	100.00	2
Principal Data Scientist	90.00	10

A análise mostra que são poucos os cargos que são 100% remoto e a todos da área da tecnologia, enquanto a maioria é híbrida ou remota.

- Consulta dos 10 países com mais empresas por ano:

```
WITH CountryCounts AS (  
  SELECT  
    ano_de_trabalho,  
    localizacao_da_empresa,  
    COUNT(*) AS empresa_contagem  
  FROM  
    dataset_salary  
  GROUP BY  
    ano_de_trabalho,  
    localizacao_da_empresa  
)  
RankedCountries AS (  
  SELECT  
    ano_de_trabalho,  
    localizacao_da_empresa,  
    empresa_contagem,  
    ROW_NUMBER() OVER(PARTITION BY ano_de_trabalho ORDER BY empresa_contagem  
DESC) AS ranking  
  FROM  
    CountryCounts  
)  
SELECT  
  ano_de_trabalho,  
  localizacao_da_empresa,  
  empresa_contagem,  
  ranking  
FROM  
  RankedCountries  
WHERE  
  ranking <= 10  
ORDER BY  
  ano_de_trabalho,  
  ranking;
```

ano_de_trabalho	localizacao_da_empresa	empresa_contage...	ranking
2020	US	33	1
2020	DE	8	2
2020	FR	5	3
2020	GB	4	4
2020	CA	3	5
2020	PT	2	6
2020	AT	2	7
2020	ES	2	8
2020	JP	2	9
2020	MU	1	10
2021	US	99	1
2021	IN	16	2
2021	DE	16	3
2021	GB	15	4
2021	CA	11	5

Nesta análise, podemos notar que o país com mais empresas em todos os anos é o Estados Unidos.

- Python

- Importações usadas:

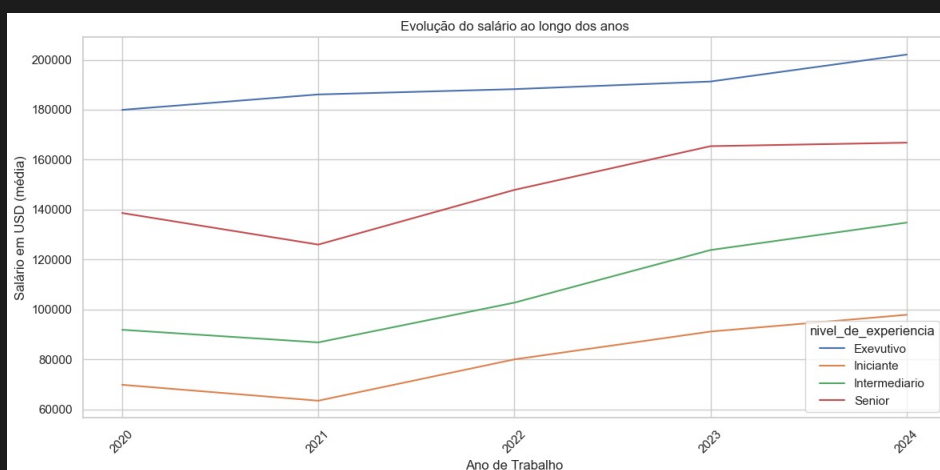
```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

- **Análise de tendências salariais ao longo do tempo:**

```
plt.figure(figsize=(12,6))
sns.lineplot(data=df_group, x='ano_de_trabalho',
y='salario_em_usd', hue='nivel_de_experiencia')
plt.xlabel('Ano de Trabalho', fontsize=12)
plt.ylabel('Salário em USD (média)', fontsize=12)
plt.title('Evolução do salário ao longo dos anos')
plt.xticks(ticks=range(2020, 2025), rotation=45)
plt.tight_layout()
plt.show()
```

```
df_group.groupby('nivel_de_experiencia')['salario_em_usd']
.mean()
.sort_values(ascending=False)
)
```

```
media_salario_experiencia = media_salario_experiencia.apply(lambda
x: f"U$ {x:,.2f}")
print("Média de salário em USD por nível de experiência (em ordem
decrecente):")
print(media_salario_experiencia)
```



Média salarial dos 4 anos:

Executivo: U\$ 189, 553. 31

Sênior: U\$ 148, 955. 77

Intermediário: U\$ 108, 012. 04

Iniciante: U\$ 80, 474. 31

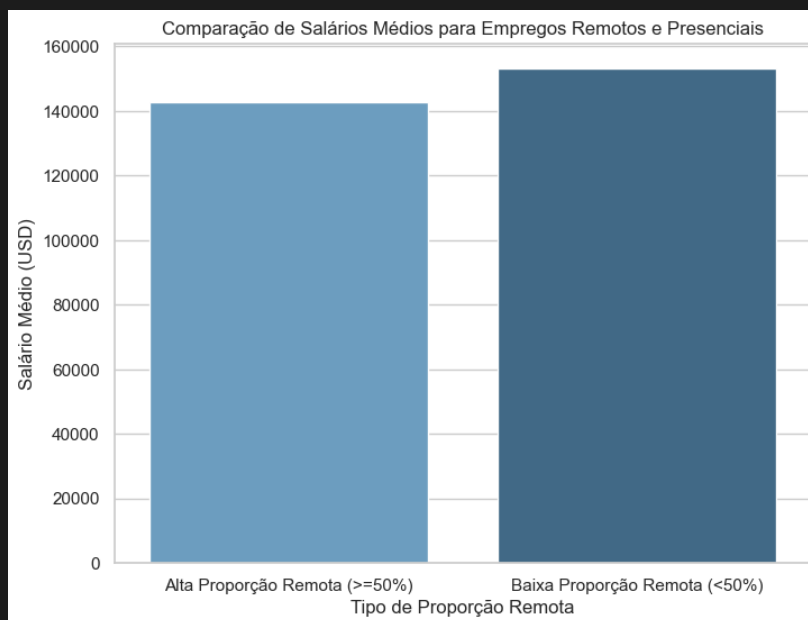
- Comparação de salários remotos e presenciais:

```
alta_remota = df[df['proporcao_remota'] >= 50]
baixa_remota = df[df['proporcao_remota'] < 50]

media_alta_remota = alta_remota['salario_em_usd'].mean()
media_baixa_remota = baixa_remota['salario_em_usd'].mean()

dados_grafico = pd.DataFrame({
    'Tipo de Proporção Remota': ['Alta Proporção Remota (>=50%)',
    'Baixa Proporção Remota (<50%)'],
    'Salário Médio (USD)': [media_alta_remota, media_baixa_remota]
})

sns.set(style="whitegrid")
plt.figure(figsize=(8, 6))
sns.barplot(x='Tipo de Proporção Remota', y='Salário Médio (USD)',
data=dados_grafico, palette='Blues_d')
plt.title('Comparação de Salários Médios para Empregos Remotos e Presenciais')
plt.xlabel('Tipo de Proporção Remota')
plt.ylabel('Salário Médio (USD)')
plt.show()
```



O gráfico mostra que empregos remotos pode ter o salário melhor de quem trabalha remotamente.

- Detecção de outliers salariais:

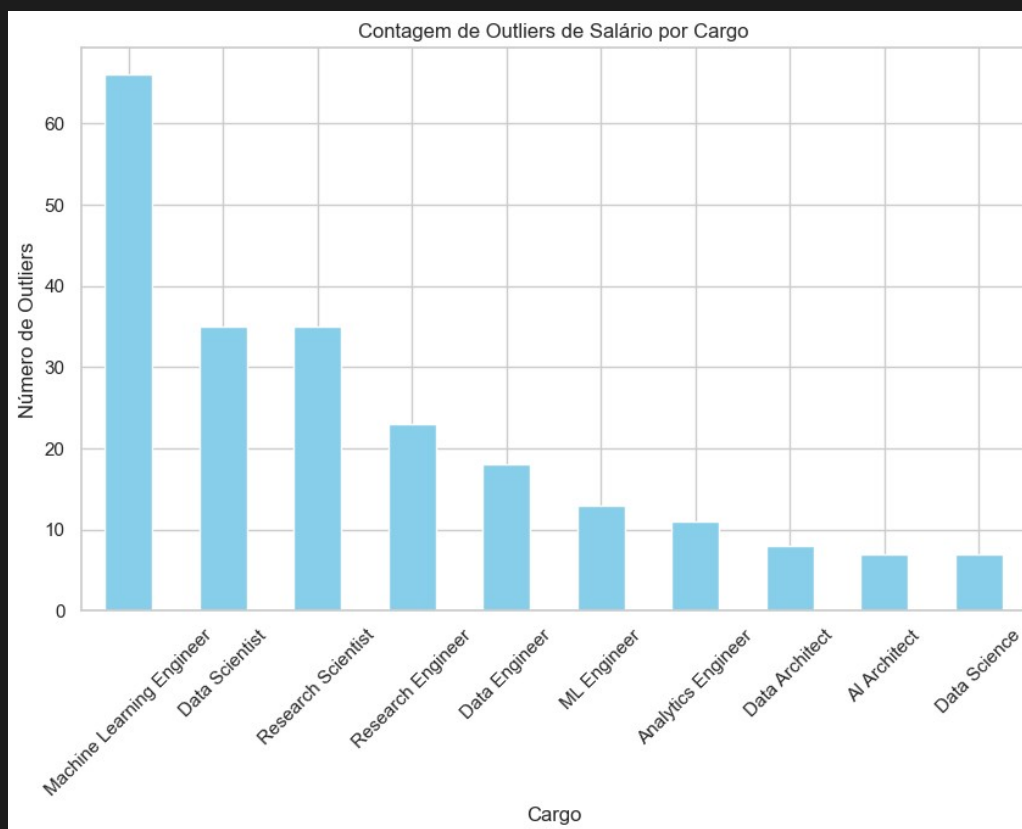
```
Q1 = df['salario_em_usd'].quantile(0.25)
Q3 = df['salario_em_usd'].quantile(0.75)
IQR = Q3 - Q1

lower_bound = Q1 - 1.5 * IQR
upper_bound = Q3 + 1.5 * IQR

outliers = df[(df['salario_em_usd'] < lower_bound) | (df['salario_em_usd'] >
upper_bound)]

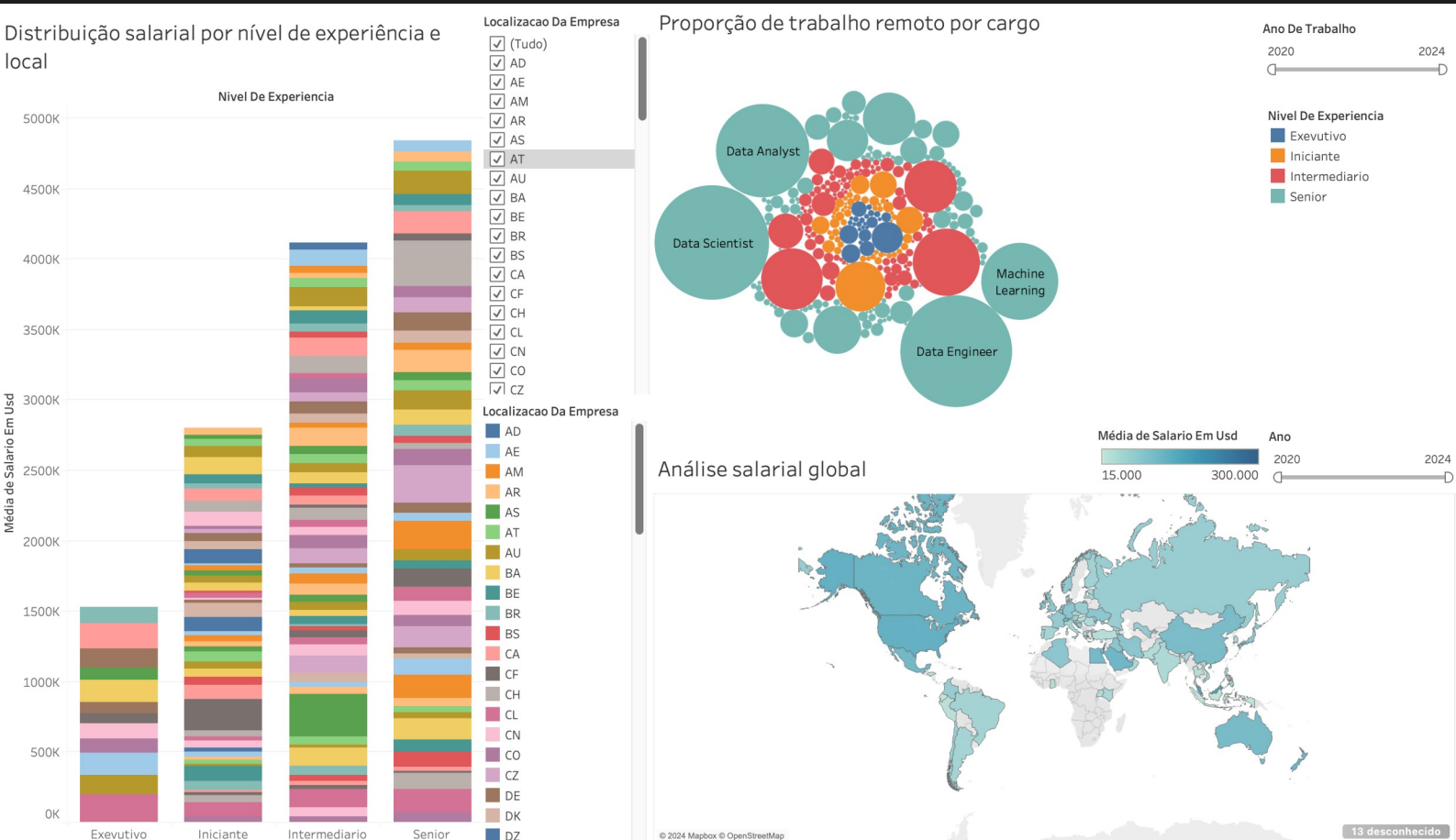
outliers_count = outliers['cargo'].value_counts().head(10)

plt.figure(figsize=(10, 6))
outliers_count.plot(kind='bar', color='skyblue')
plt.title('Contagem de Outliers de Salário por Cargo')
plt.xlabel('Cargo')
plt.ylabel('Número de Outliers')
plt.xticks(rotation=45)
plt.show()
```



- Salários entre U\$ 101,125.00 e 185,900.00 estão dentro da faixa central do conjunto de dados e Salários acima de U\$ 313,062.5 são identificados como possíveis outliers, representando valores significativamente altos comparados com o restante dos dados.

•Tableau



- A disparidade salarial entre regiões é significativa, e níveis de experiência influenciam fortemente os salários médios.
- Cargos de maior flexibilidade (alto trabalho remoto) estão concentrados em setores com forte presença de tecnologia.
- O mapa interativo é útil para identificar as regiões mais competitivas em termos de remuneração e analisar a distribuição geográfica dos salários.
- Essas visualizações permitem uma análise holística das tendências salariais e da estrutura de trabalho remoto, fornecendo insights valiosos para estratégias de recursos humanos e políticas de remuneração.