

## Entrega 4 - Estatística Inferencial

### Objetivo do Projeto

O projeto de Ciência de Dados do grupo consiste em analisar a diversidade e as tendências do mercado de TI, ligado à área de Dados no Brasil, focando em como as vagas de emprego estão sendo distribuídas entre diferentes grupos, como por exemplo, gênero, raça, orientação sexual e deficiência. Além disso, queremos filtrar quais tecnologias estão sendo mais demandadas e quais níveis de senioridade estão sendo mais buscados.

O objetivo geral do nosso projeto é gerar análises eficientes para entender como as empresas filtram candidatos, gerando assim maior entendimento na hora dos processos seletivos. Além disso, queremos saber como esses profissionais desempenham seu trabalho, seja por senioridade, horas trabalhadas e entre outros aspectos.

Nas entregas da matéria de Análise Inferencial de Dados, resolvemos filtrar nossa pesquisa sobre gênero e nível de ensino dos participantes.

### Teste de Hipóteses

H0: O gênero feminino possui número igual ou maior de pessoas com nível de ensino em comparação aos homens.

H1: Pessoas do gênero masculino são em maior quantidade em nível de ensino em comparação às mulheres.

### Teste qui-quadrado

```
> #ENTREGA 04 - CASO SEJAM DADOS CATEGÓRICOS (TESTE QUI-QUADRADO)
>
> # Contar as frequências de cada categoria por gênero
> frequencias <- table(dados_pesquisa_2023$dsc_genero, dados_pesquisa_2023$dsc_nivel_ensino)
>
> # Exibir a tabela de frequências
> print(frequencias)
```

	Doutorado ou Phd	Estudante de Graduação	Graduação/Bacharelado	Mestrado
Feminino	72	132	411	196
Masculino	137	542	1376	476
Outro	0	3	5	0
Prefiro não informar	1	1	6	4

  

	Não tenho graduação formal	Pós-graduação	Prefiro não informar
Feminino	20	462	0
Masculino	85	1351	8
Outro	0	1	0
Prefiro não informar	0	4	0

```
>
```

## Valor -p

```
>
> # Realizar o teste qui-quadrado de independência usando a tabela de frequências
> teste_qui2 <- chisq.test(frequencias)
warning message:
In chisq.test(frequencias) :
  Aproximação do qui-quadrado pode estar incorreta
>
> # Exibir os resultados
> cat("Estatística Qui-Quadrado:", teste_qui2$statistic, "\n")
Estatística Qui-Quadrado: 46.00697
> cat("Valor-p:", teste_qui2$p.value, "\n")
Valor-p: 0.000295979
> |
```

---

## Interpretação do Resultado

```
> # Interpretação do resultado
> nivel_significancia <- 0.05
> if (teste_qui2$p.value < nivel_significancia) {
+   cat("Resultado: Existe uma relação significativa entre o gênero e o nível de formação.\n")
+ } else {
+   cat("Resultado: Não existe uma relação significativa entre o gênero e o nível de formação de exata
s.\n")
+ }
Resultado: Existe uma relação significativa entre o gênero e o nível de formação.
> |
```

## Conclusão

De acordo com os dados obtidos, podemos concluir que os homens possuem um maior nível de escolaridade em relação às mulheres, ressaltando uma relação significativa entre os gêneros e nível de formação. Isso provavelmente se reflete em diversos fatores sociais já existentes na sociedade atual.