

**Universidade Federal de Minas Gerais – UFMG**  
**Instituto de Ciências Exatas – ICEx**  
**Departamento de Estatística – DEST**  
**Centro de Estudos em Estatística e Ciências de Dados – CECiDa**

**Projeto LabEst: Consultoria em Análise Estatística para Alunos de Pós-Graduação**

**Relatório Final da Análise Estatística:**  
**Além das Telas: práticas de lazer de turistas motivado pelo audiovisual**

Clientes:

**João Lucas de Almeida Campos**

Doutorando Estudos do Lazer - UFMG – Escola de Educação Física, Fisioterapia e  
Terapia Ocupacional (EEFFTO/UFMG)

**Christianne Luce Gomes**

Orientador

Consultores:

**Isabelle Fernandes de Oliveira**

**Joaquim Pimenta Leite Neto**

Alunos da Disciplina “Laboratório de Estatística I”

**Orientação: Roberto da Costa Quinino**

**02/12/2023**

# Sumário

<b>1. Introdução</b>	<b>2</b>
<b>2 Metodologia</b>	<b>3</b>
<b>3 Resultados</b>	<b>4</b>
3.1 Análise exploratória - Estatísticas descritivas . . . . .	4
3.2 - Modelo de regressão logística . . . . .	53
<b>4. Conclusão</b>	<b>59</b>

# 1. Introdução

Este trabalho propõe uma análise das experiências dos turistas motivados por produções audiovisuais (**turistas específicos**), explorando como as produções audiovisuais impactam suas motivações, expectativas e percepções durante a vivência do turismo como experiência de lazer.

O autor definiu os conceitos da classificação de turistas como:

- Casual: estava em um local relacionado ao universo das produções audiovisual por acaso;
- Geral: o local não foi escolhido especificamente motivado pelas produções audiovisuais, mas aproveitou para realizar alguma atividade relacionadas ao universo cinematográfico;
- Específica: quando a escolha do destino foi exclusivamente induzida/motivada por uma produção audiovisual.

Como o interesse é entender a relação do turista motivado pelo audiovisual e as variáveis da viagem, o trabalho focalizou no grupo de turistas específicos, já que são eles que viajam motivados exclusivamente pelo audiovisual.

A coleta de dados foi concretizada por meio da aplicação de um questionário disponibilizado pelo Google Forms, entre os meses de julho e agosto de 2022. O questionário incluiu 27 questões ao todo, contemplando informações sobre os dados sociodemográficos e a caracterização do grupo de voluntários: idade, gênero, educação, ocupação, país e região do residente; questões relacionadas ao turismo induzido por produções audiovisuais; questões relacionadas à viagem realizada pelos participantes da pesquisa, contemplando as seguintes categorias: motivação, percepção e sentimentos. Ao todo, participaram 121 voluntários, recrutados por meio de divulgação da pesquisa em redes sociais e convites enviados por e-mail. Das 121 respostas, 117 foram válidas e e quatro descartadas, visto que os participantes não responderam à pergunta sobre a viagem realizada, um requisito essencial para participar da pesquisa.

## 1.1 Objetivos Gerais do estudo

Identificar as variáveis que impactam mais quando a viagem é motivada exclusivamente pelo audiovisual.

## 1.2 Objetivos Específicos

- Fazer uma análise por grupo de entrevistados;
- Mensurar o impacto das variáveis que estão mais associadas com o turista específico

## 2 Metodologia

O presente estudo busca uma compreensão aprofundada das motivações dos turistas ao visitarem uma localidade específica, com especial ênfase na influência do audiovisual.

**1. Tratamento do Banco de Dados:** Inicialmente, o banco de dados, coletado pelo nosso cliente, passou por um processo de tratamento. Essa etapa foi crucial para garantir a integridade e relevância dos dados, permitindo uma análise estatística mais precisa.

**2. Análise Exploratória:** Utilizando estatísticas descritivas e gráficos para obter e apresentar insights preliminares sobre as variáveis presentes no banco de dados. Isso proporcionou uma compreensão inicial da distribuição e tendências dos dados, destacando padrões e anomalias que orientaram as etapas subsequentes da pesquisa.

**3. Testes Qui-Quadrado:** Para investigar as relações entre a variável de interesse, representada pelo grupo dos participantes, e outras variáveis relevantes, empregamos testes qui-quadrado. Com isso determinamos se existem associações significativas entre as variáveis, revelando fatores que influenciam a motivação dos turistas. Os resultados orientaram a seleção das variáveis mais relevantes para o modelo de regressão logística.

**4. Modelo de Regressão Logística:** Com base nas variáveis identificadas como estatisticamente significativas nos testes qui-quadrado, desenvolvemos um modelo de regressão logística explorando a relação entre as variáveis independentes selecionadas e a variável dependente de interesse.

### 3 Resultados

Os resultados são apresentados de maneira abrangente, com gráficos, testes estatísticos para avaliação das variáveis e tabelas que descrevem os dados no geral. Além disso, foram empregadas a regressão logística para compreender a relação das variáveis explicativas com a variável de interesse.

#### Teste qui-quadrado de independência

Foi aplicado o teste qui-quadrado para verificar se existe associação entre as variáveis grupo com cada uma das outras variáveis qualitativas.

Hipóteses:

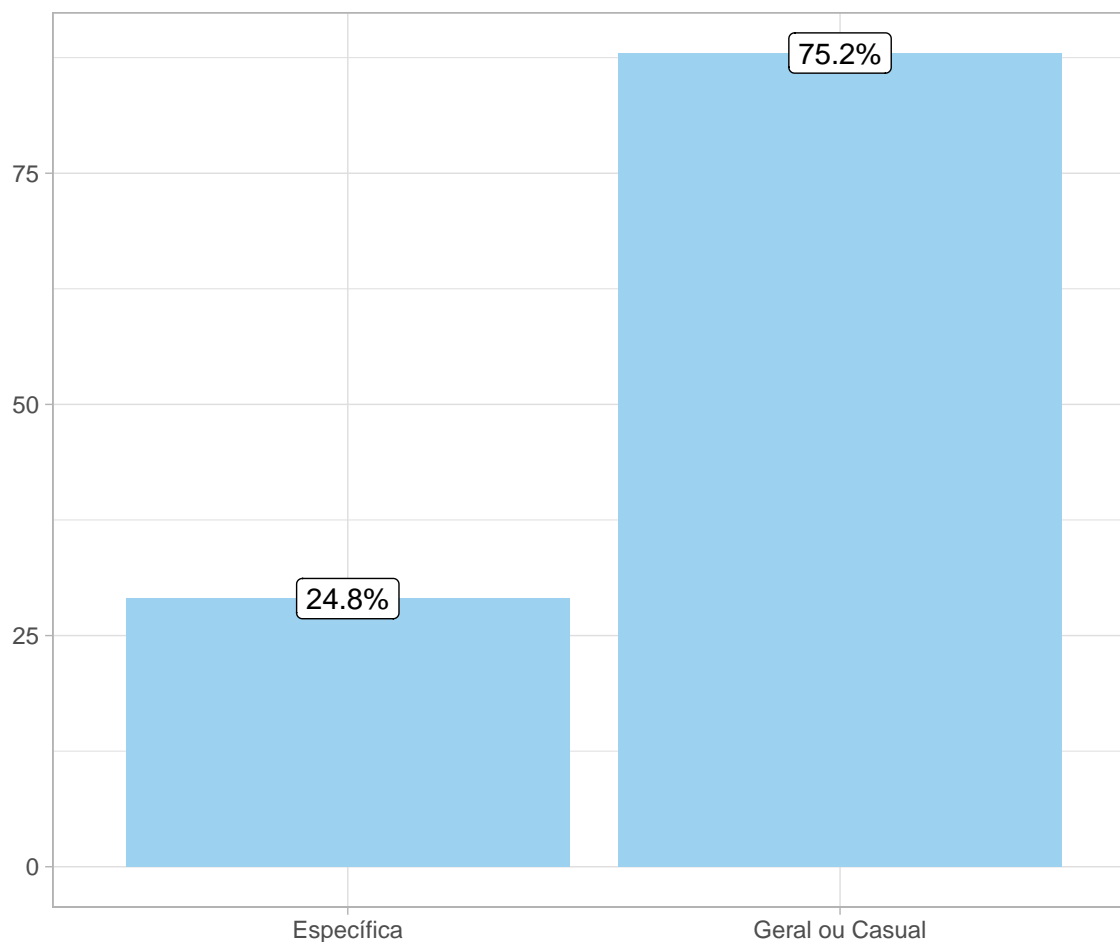
$H_0$ : As variáveis **grupo** com outra variável de interesse são independentes

$H_1$ : As variáveis **grupo** com outra variável de interesse estão associadas

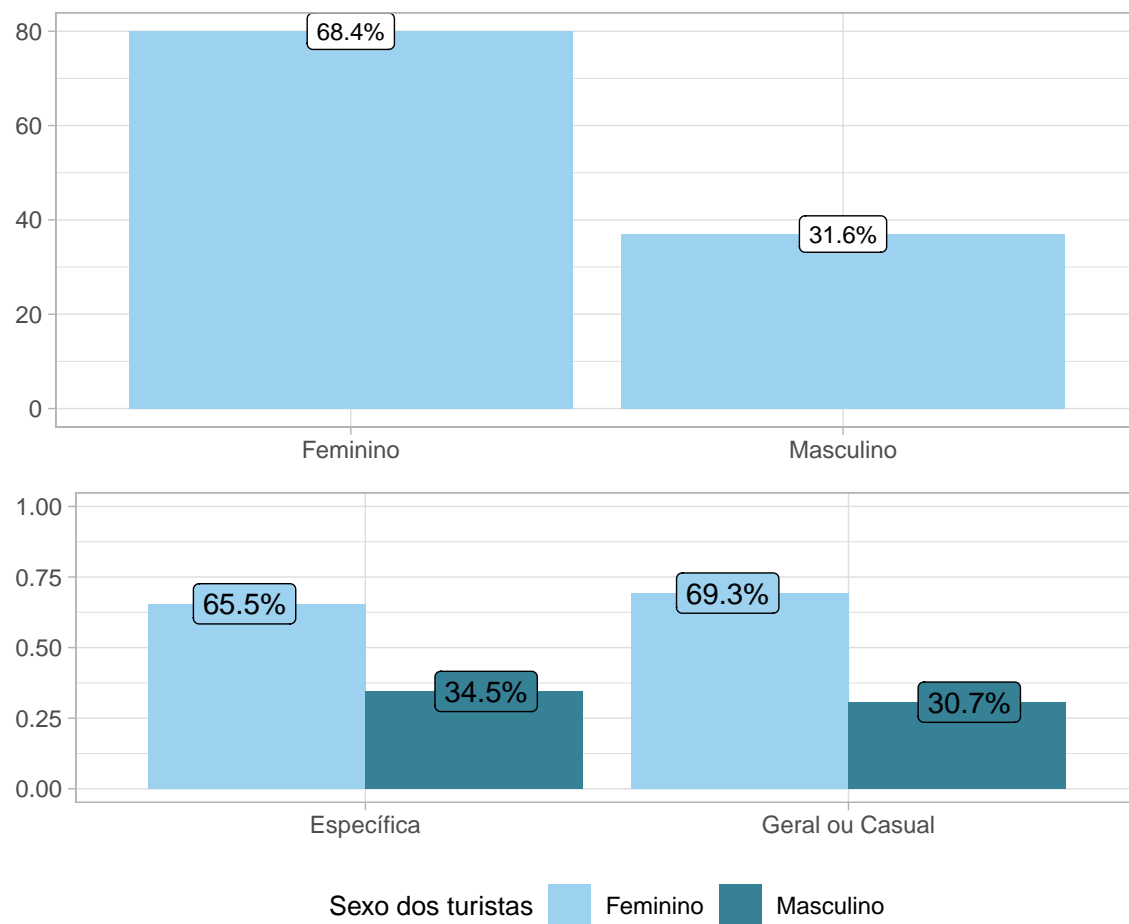
Caso o valor-p do teste seja inferior a 0.05, então rejeitamos  $H_0$  e acredita-se com que as variáveis analisadas estejam associadas. Em outras palavras, significa dizer que a proporção do cruzamento das duas variáveis é significativa.

#### 3.1 Análise exploratória - Estatísticas descritivas

Gráfico 1: Turistas classificados segundo o grupo

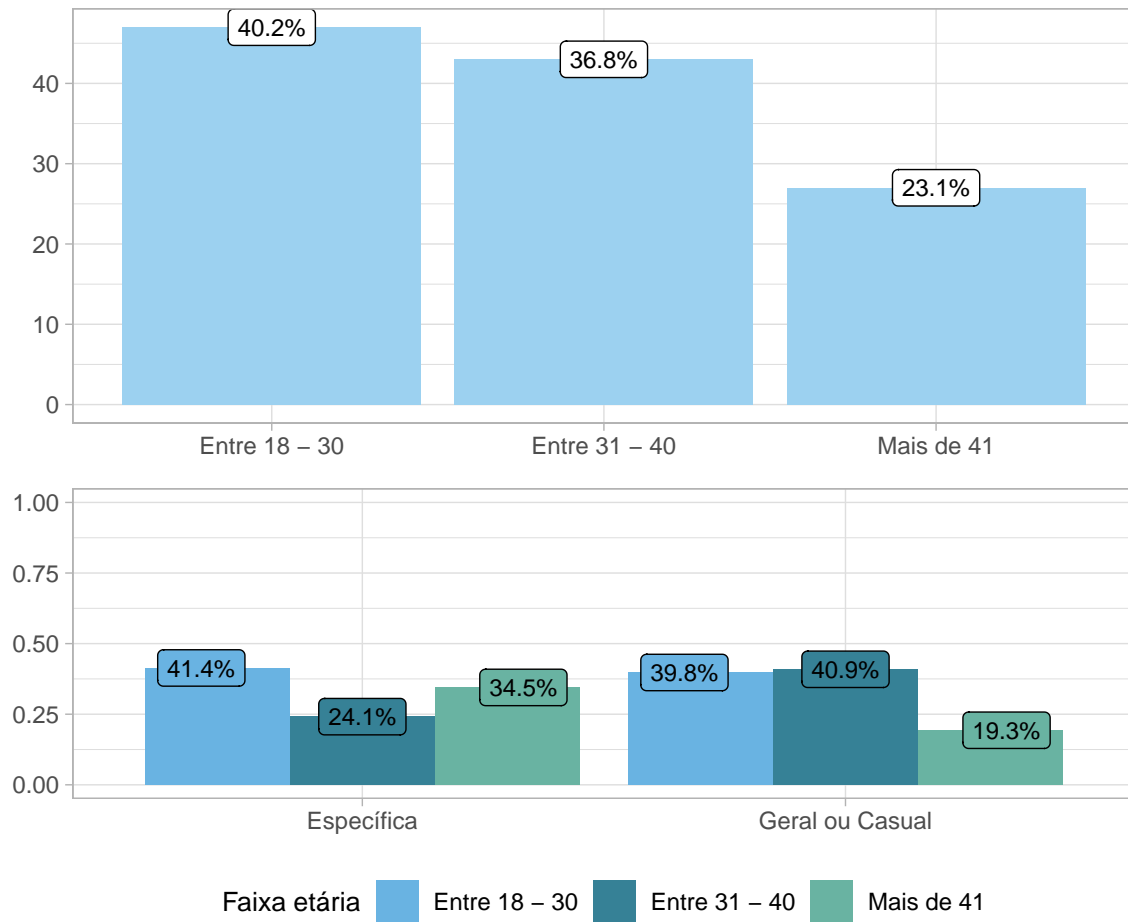


**Gráfico 2: Sexo por grupo**



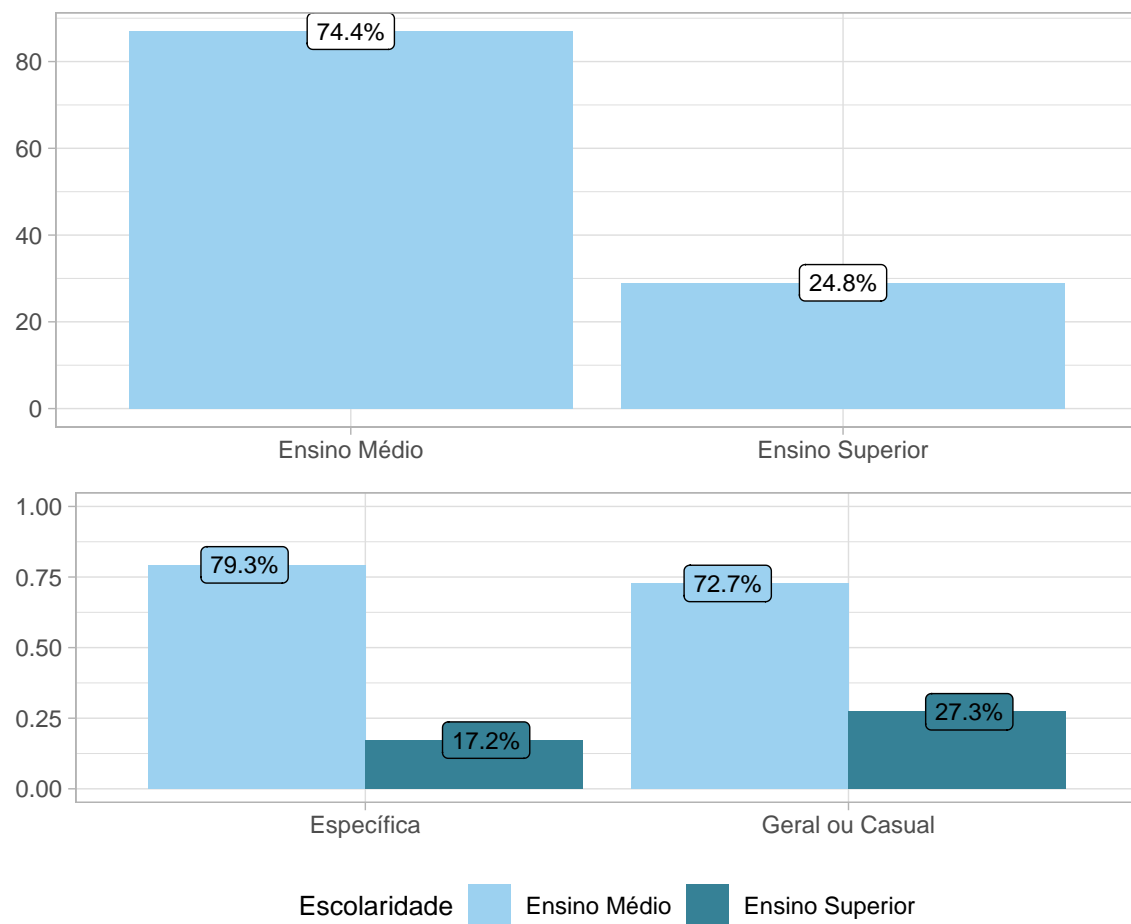
## Valor-p do teste qui-quadrado: 0.8796. Como o valor-p é superior a 0.05, não rejeitamos a hipótese  
## nula e acredita-se que as variáveis são independentes.

**Gráfico 3: Faixa etária por grupo**



## Valor-p do teste qui-quadrado: 0.1454. Como o valor-p é superior a 0.05, não rejeitamos a hipótese  
## nula e acredita-se que as variáveis são independentes.

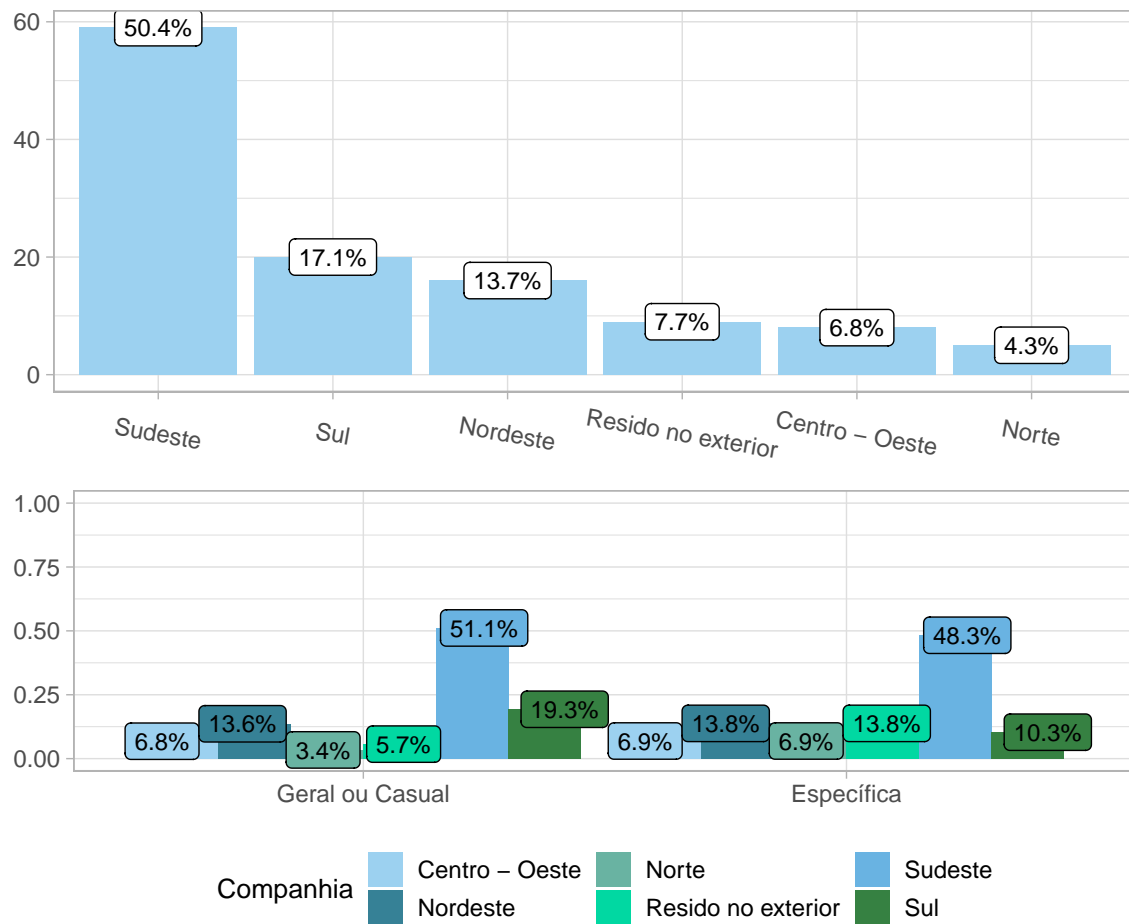
**Gráfico 4: Escolaridade por grupo**



## Valor-p do teste qui-quadrado: 0.4523. Como o valor-p é superior a 0.05, não rejeitamos a hipótese nula e acredita-se que as variáveis são independentes.

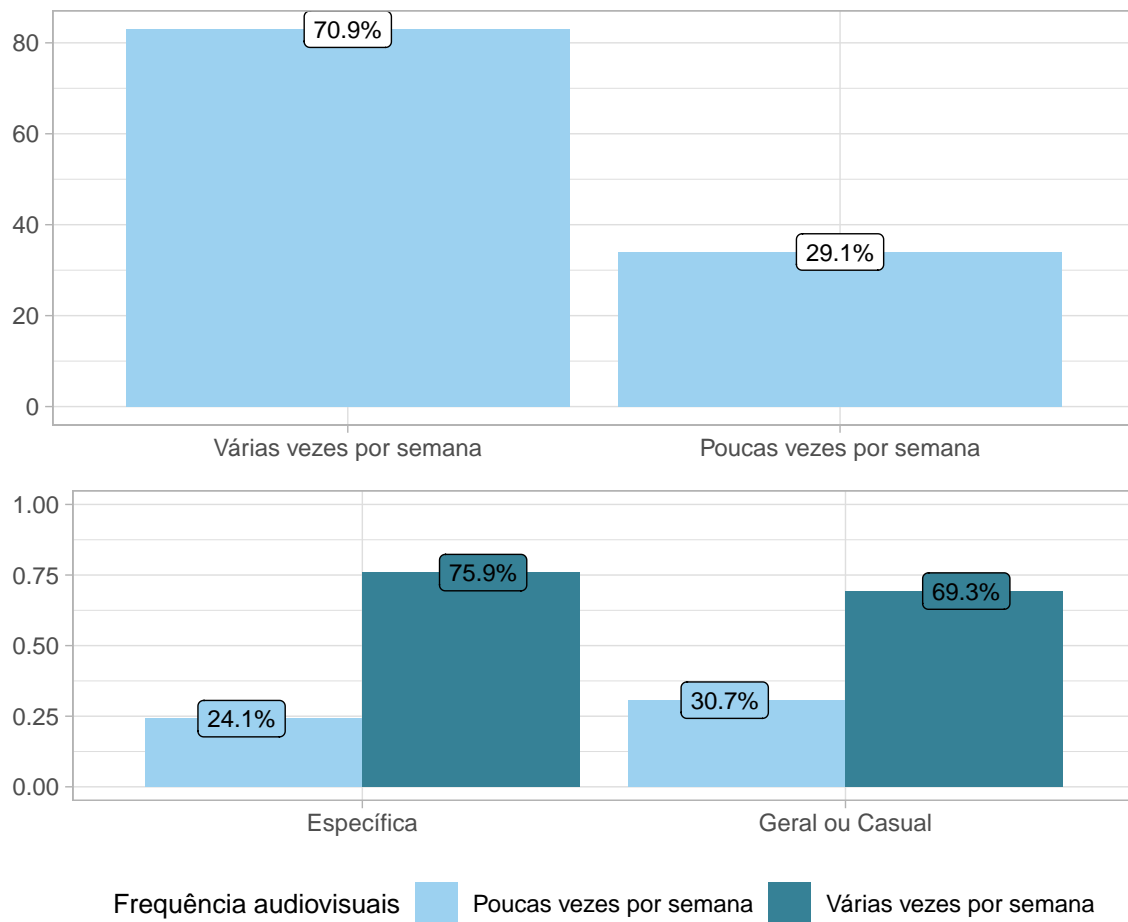


**Gráfico 5: Região de residência por grupo**



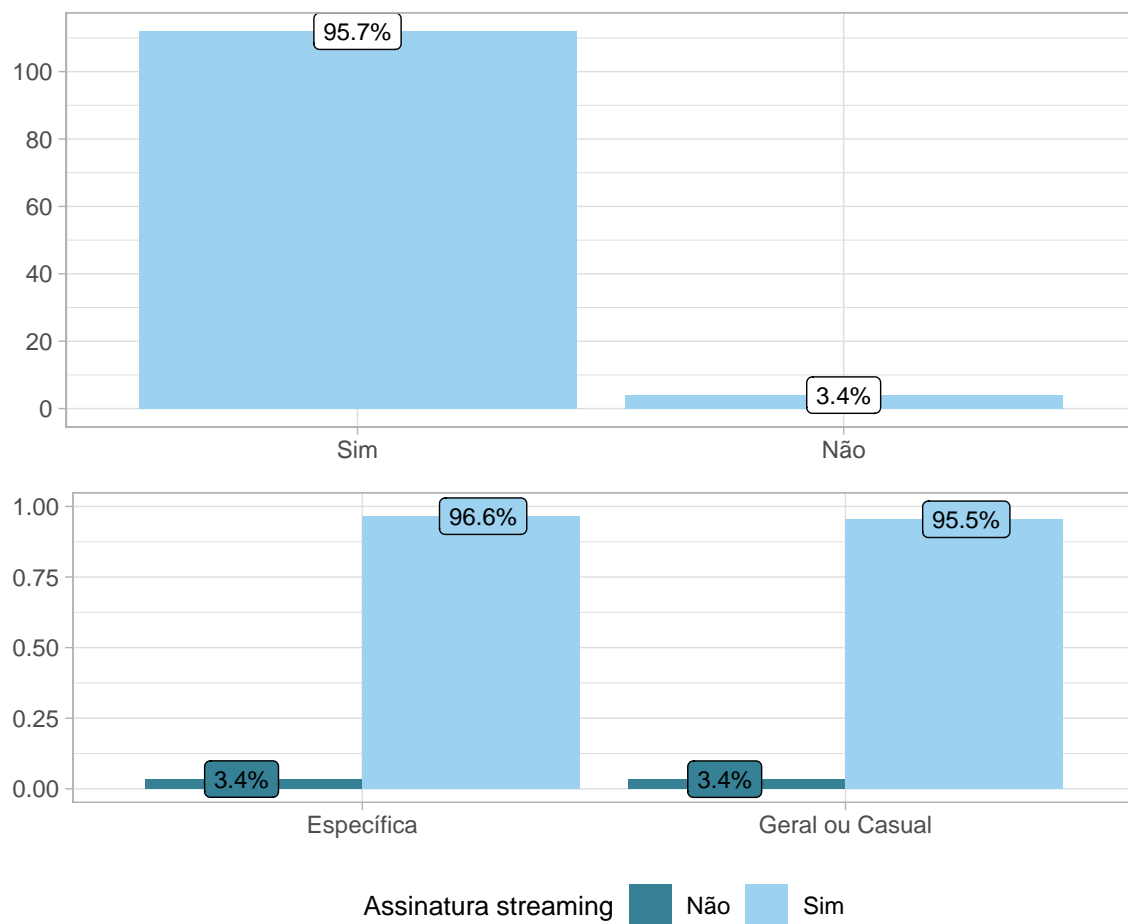
## Valor-p do teste qui-quadrado: 0.6159. Como o valor-p é superior a 0.05, não rejeitamos a hipótese nula e acredita-se que as variáveis são independentes.

**Gráfico 6: Frequência de consumo audiovisual por grupo**



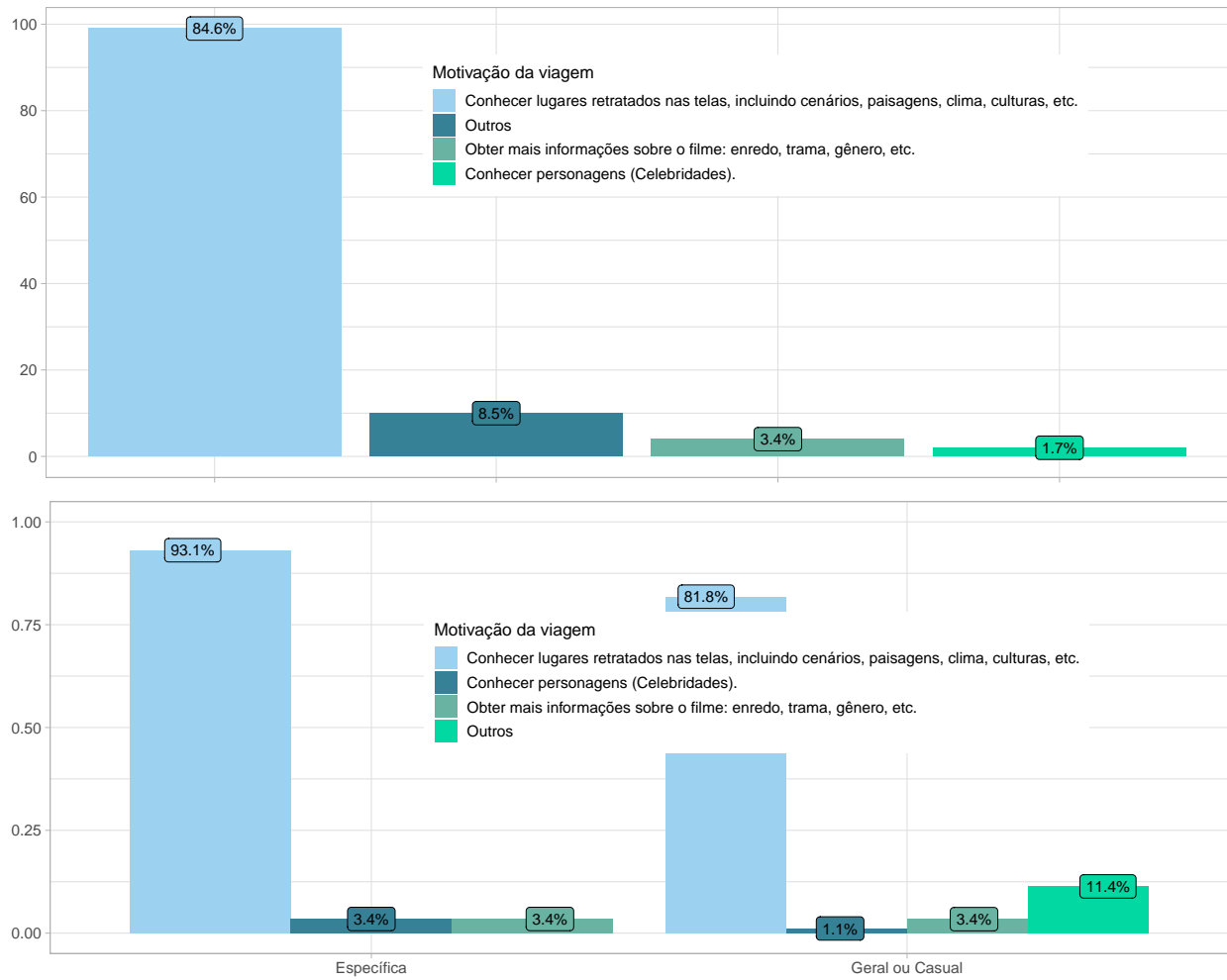
## Valor-p do teste qui-quadrado: 0.6619. Como o valor-p é superior a 0.05, não rejeitamos a hipótese nula e acredita-se que as variáveis são independentes.

**Gráfico 7: Assinatura streaming por grupo**



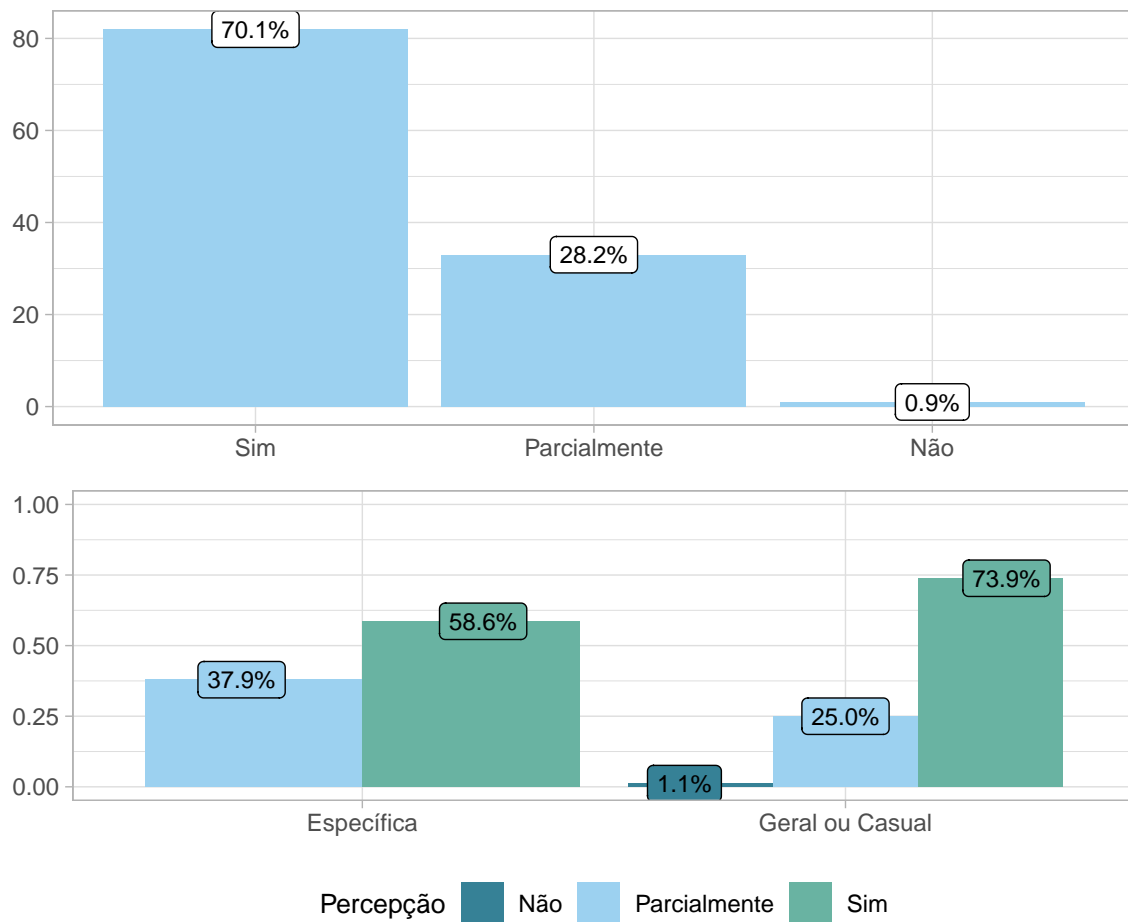
## Valor-p do teste qui-quadrado: 1. Como o valor-p é superior a 0.05, não rejeitamos a hipótese  
## nula e acredita-se que as variáveis são independentes.

**Gráfico 8: Motivação da viagem por grupo**



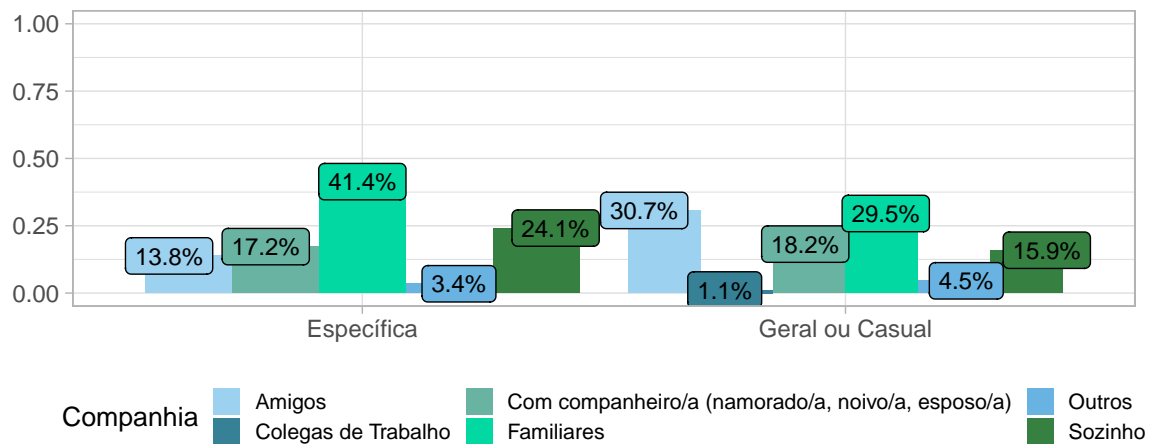
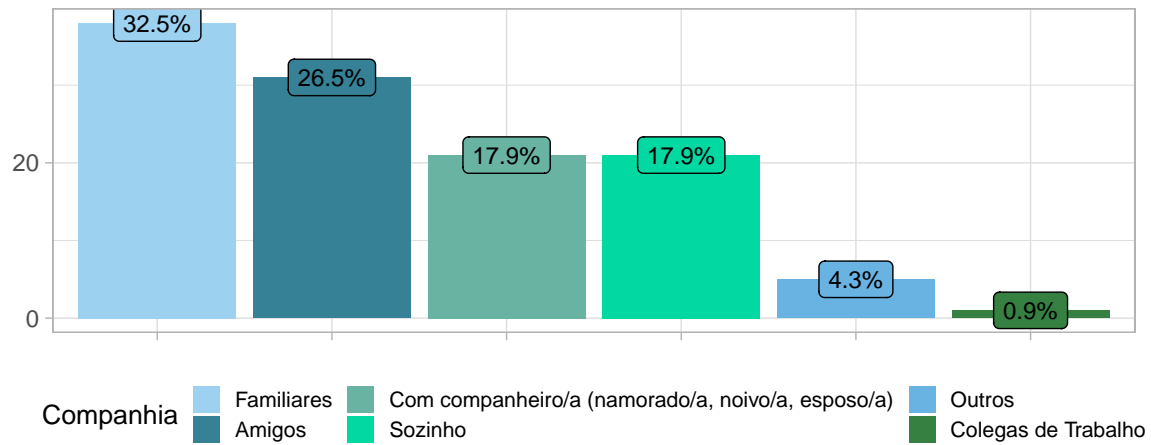
## Valor-p do teste qui-quadrado: 0.2362. Como o valor-p é superior a 0.05, não rejeitamos a hipótese nula e acredita-se que as variáveis são independentes.

**Gráfico 9: Locais visitados estavam de acordo com os retratados em produção audiovisual por grupo**



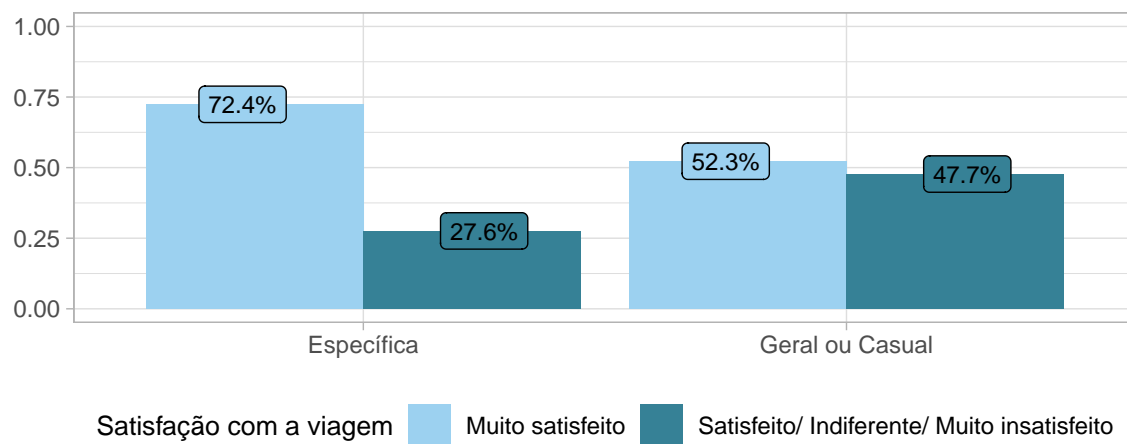
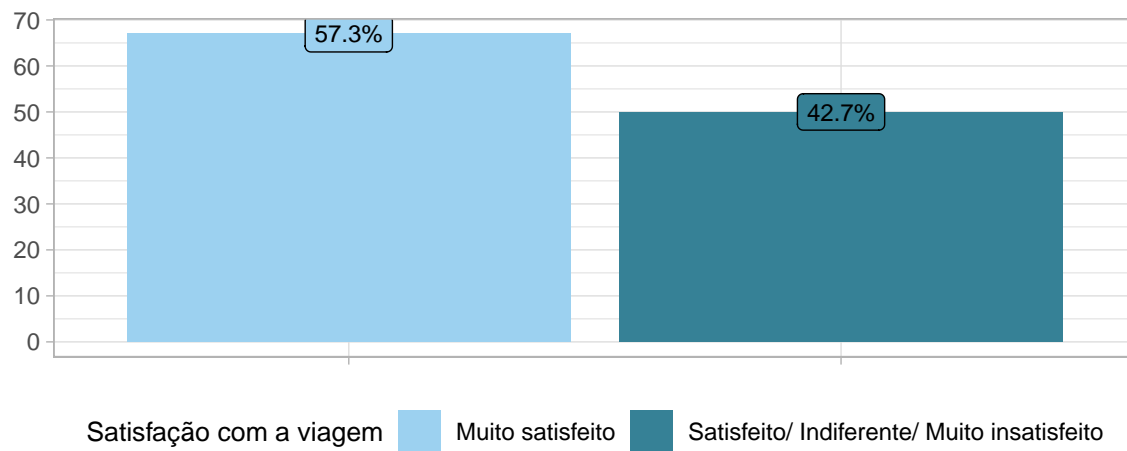
## Valor-p do teste qui-quadrado: 0.307. Como o valor-p é superior a 0.05, não rejeitamos a hipótese nula e acredita-se que as variáveis são independentes.

**Gráfico 10: Companhia da viagem por grupo**



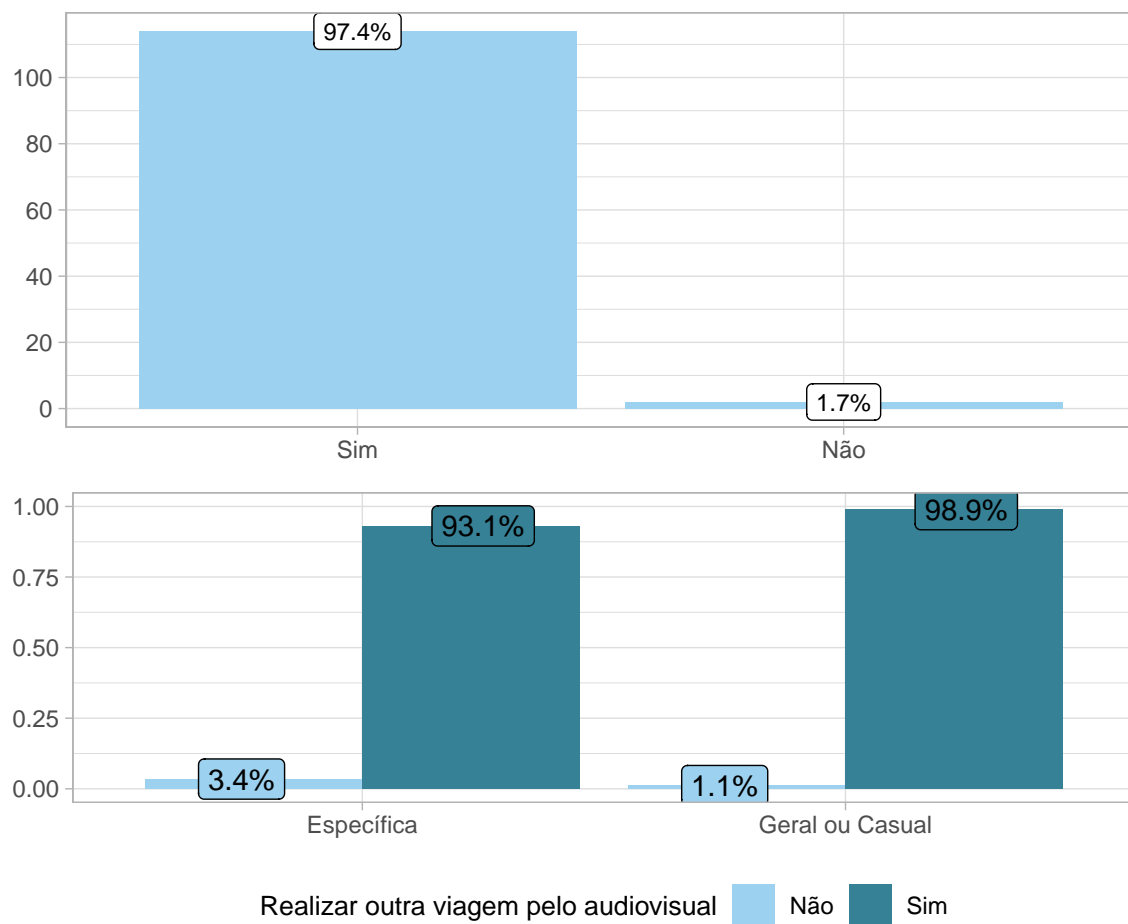
## Valor-p do teste qui-quadrado: 0.4781. Como o valor-p é superior a 0.05, não rejeitamos a hipótese nula e acredita-se que as variáveis são independentes.

**Gráfico 11: Satisfação com viagem induzida pelo audiovisual por grupo**



## Valor-p do teste qui-quadrado: 0.092. Como o valor-p é superior a 0.05, não rejeitamos a hipótese  
## nula e acredita-se que as variáveis são independentes.

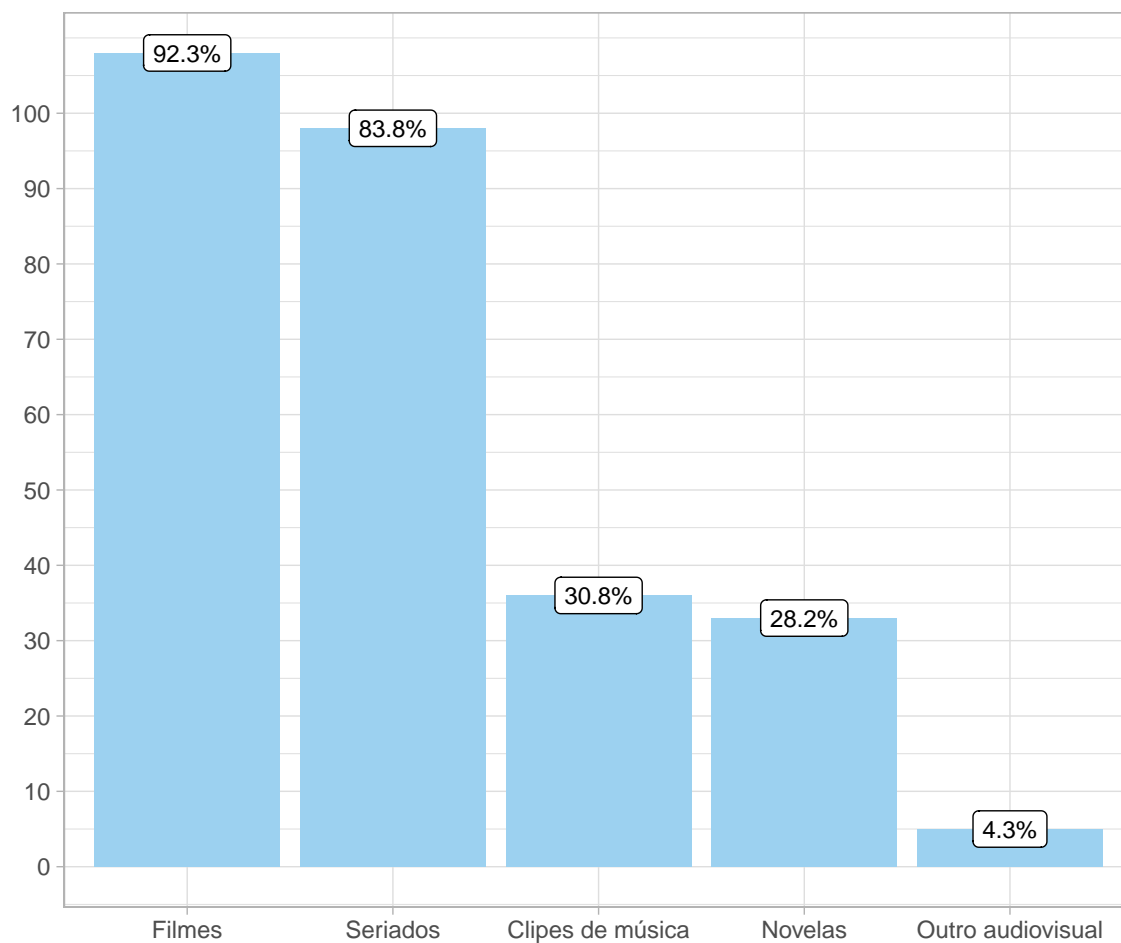
**Gráfico 12: Realizar outra viagem pelo audiovisual por grupo**



## Valor-p do teste qui-quadrado: 0.9771. Como o valor-p é superior a 0.05, não rejeitamos a hipótese nula e acredita-se que as variáveis são independentes.

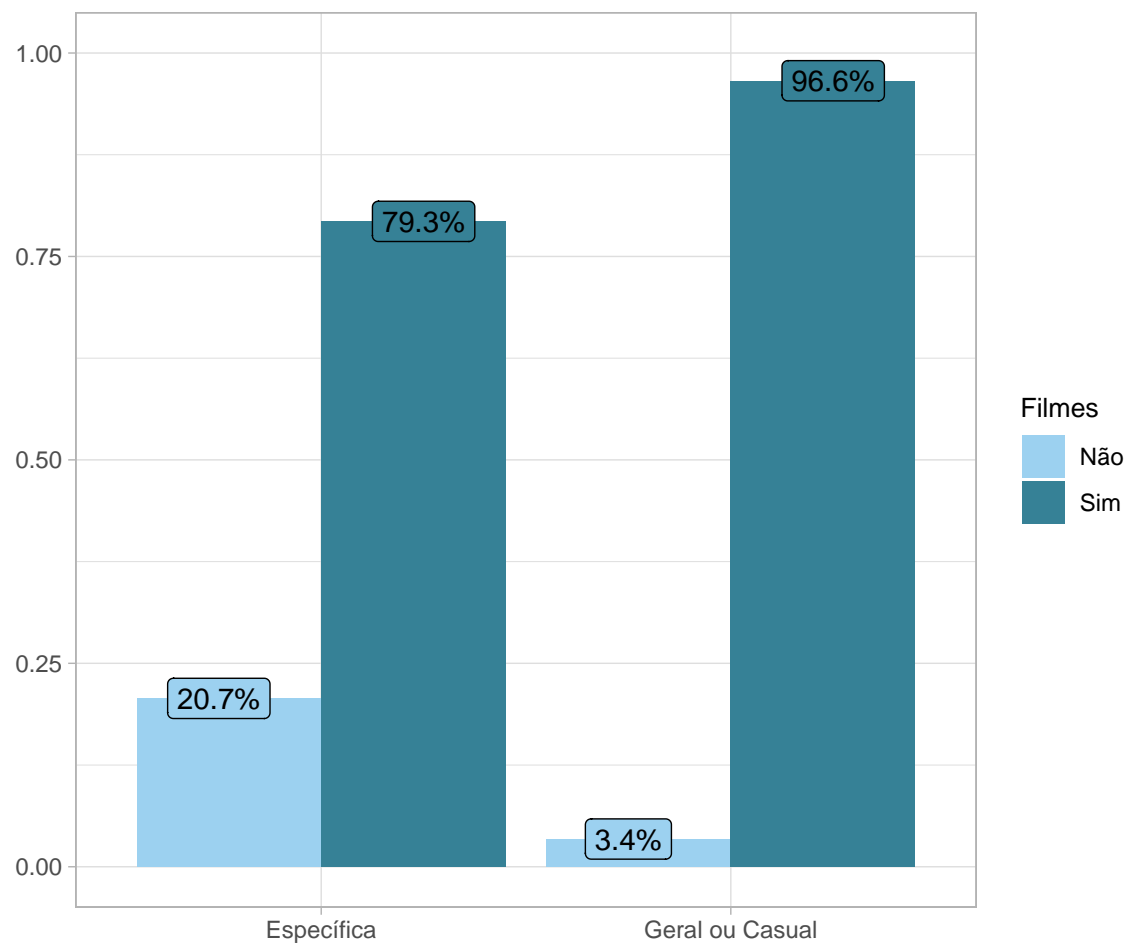


**Gráfico 13: Consumo de produção audiovisual**



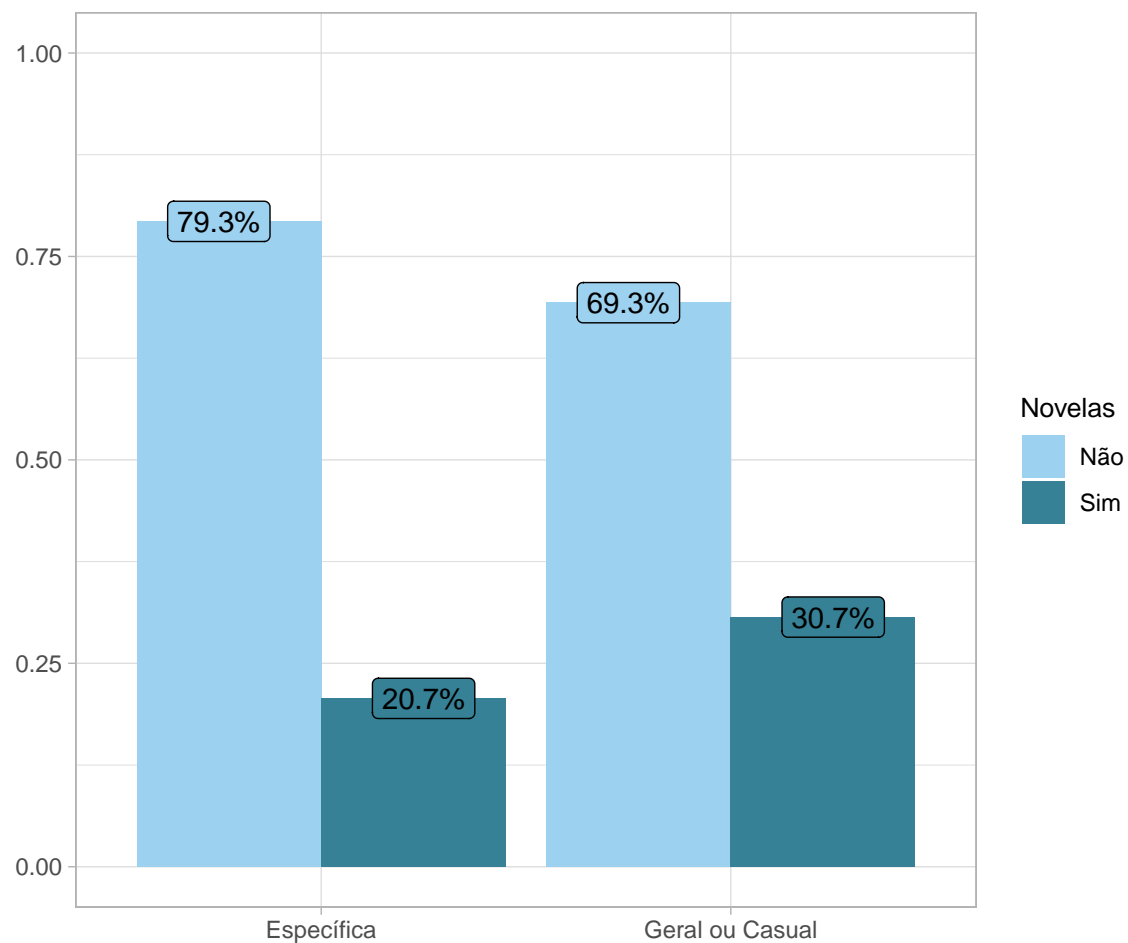
Entre os que citaram “Outro audiovisual”, descreveram: documentários, curtas, propagandas criativas, produtores de conteúdo nas redes sociais.

**Gráfico 14: Turistas que assistem filmes por grupo**



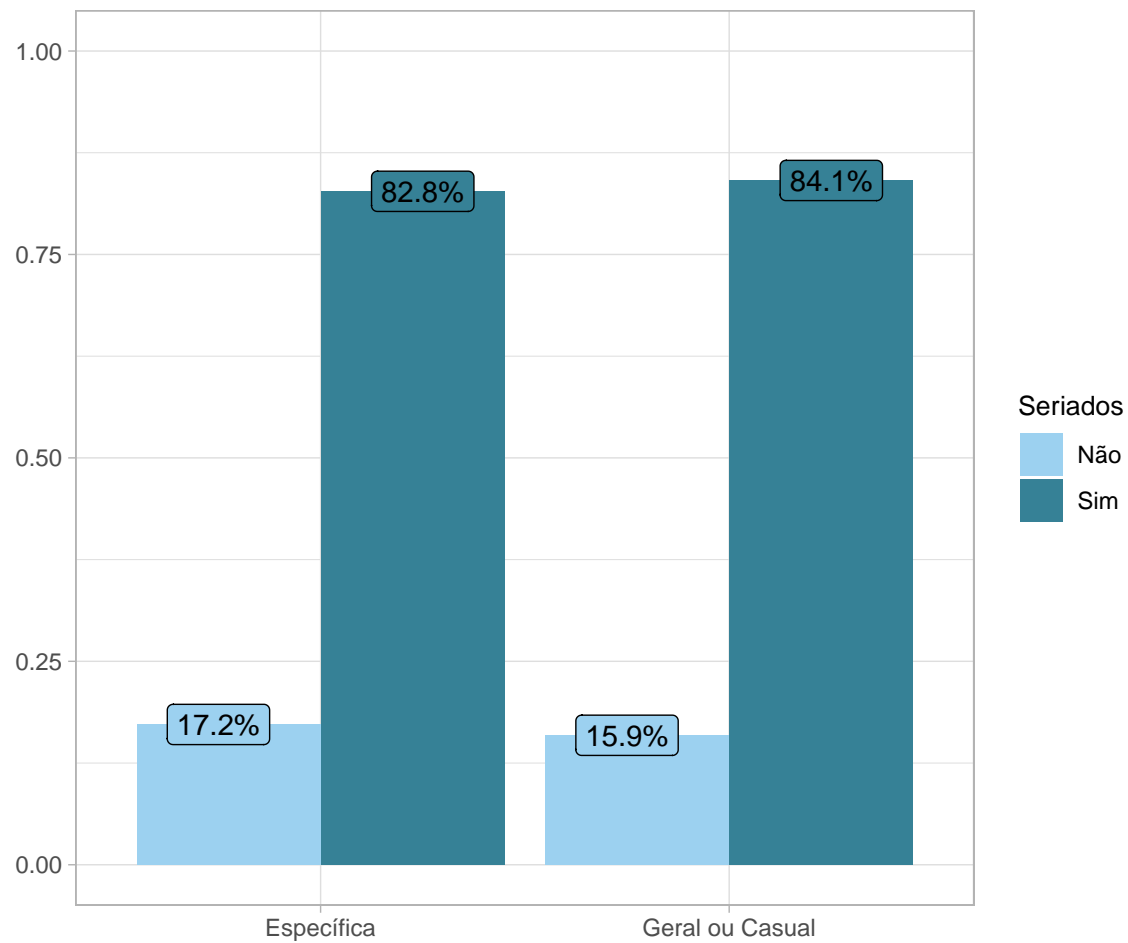
## Valor-p do teste qui-quadrado: 0.0086. Como o valor-p é inferior a 0.05, rejeitamos a hipótese nula  
## e acredita-se que as variáveis estão associadas.

**Gráfico 15: Turistas que assistem novelas por grupo**



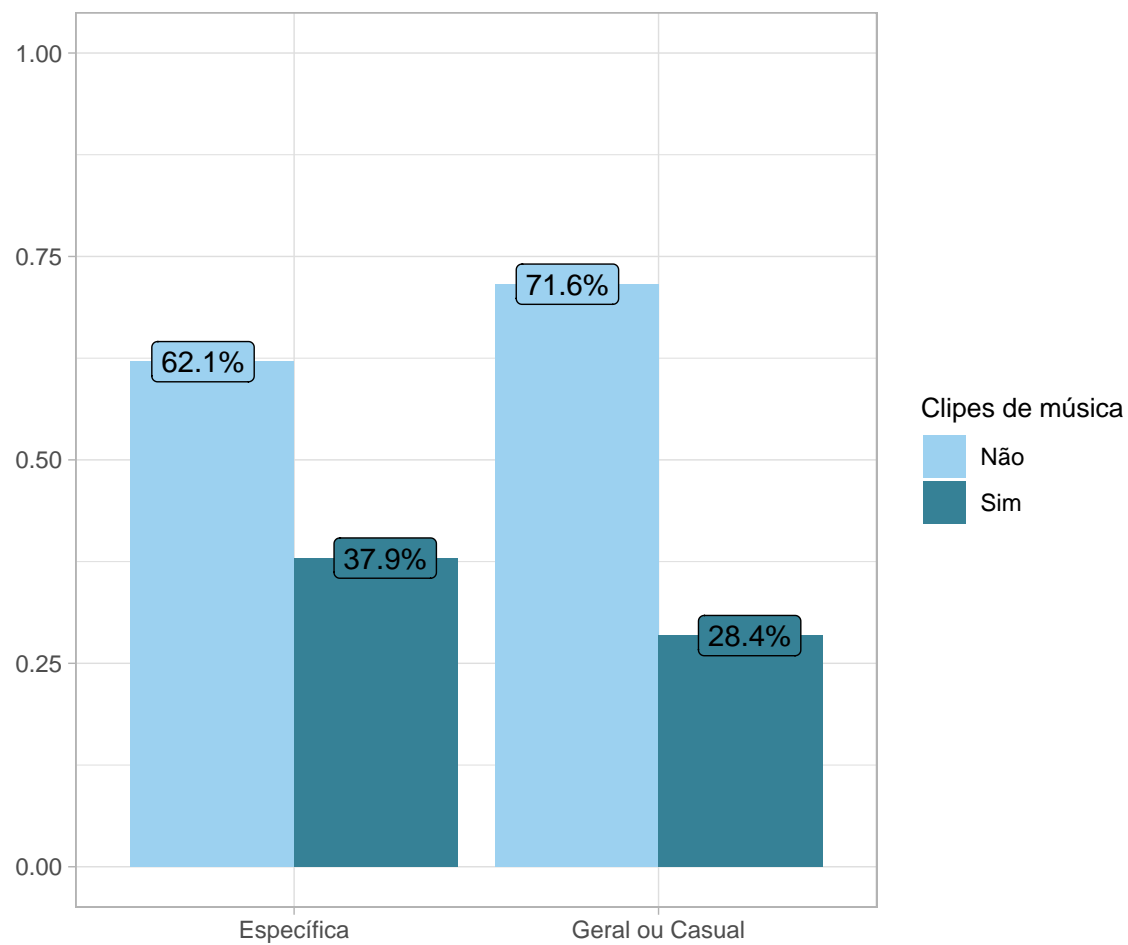
## Valor-p do teste qui-quadrado: 0.4242. Como o valor-p é superior a 0.05, não rejeitamos a hipótese  
## nula e acredita-se que as variáveis são independentes.

**Gráfico 16: Turistas que assistem seriados por grupo**



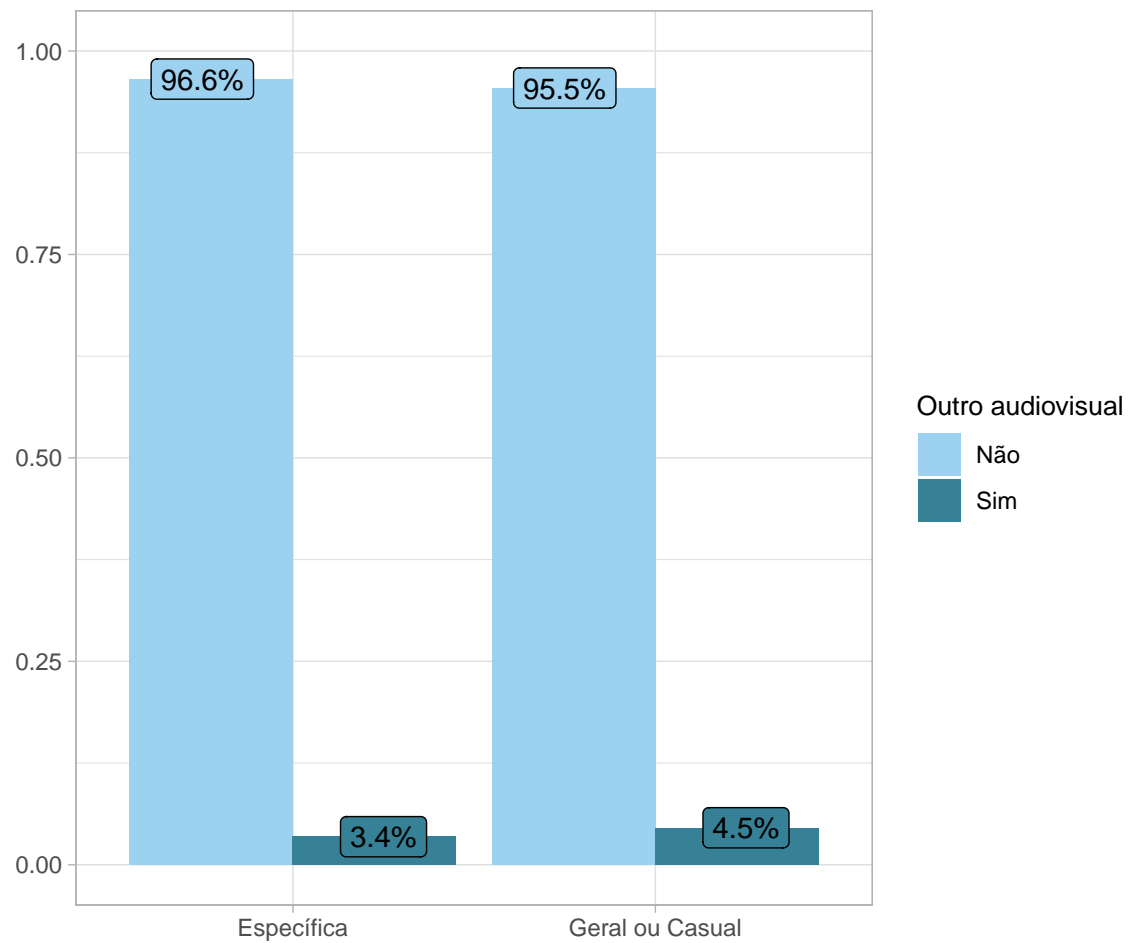
## Valor-p do teste qui-quadrado: 1. Como o valor-p é superior a 0.05, não rejeitamos a hipótese  
## nula e acredita-se que as variáveis são independentes.

**Gráfico 17: Turistas que assistem clipes de música por grupo**



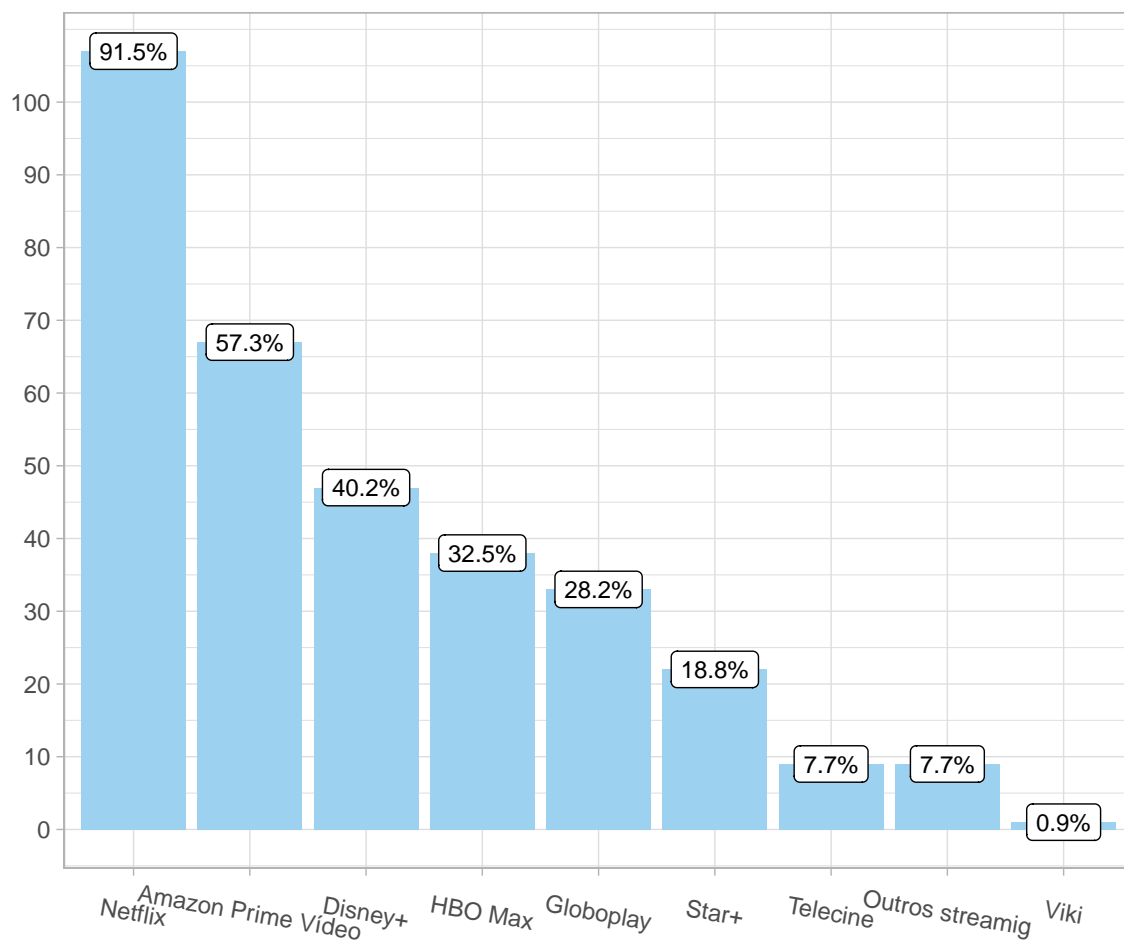
## Valor-p do teste qui-quadrado: 0.4644. Como o valor-p é superior a 0.05, não rejeitamos a hipótese  
## nula e acredita-se que as variáveis são independentes.

**Gráfico 18: Turistas que assistem outro produto audiovisual por grupo**



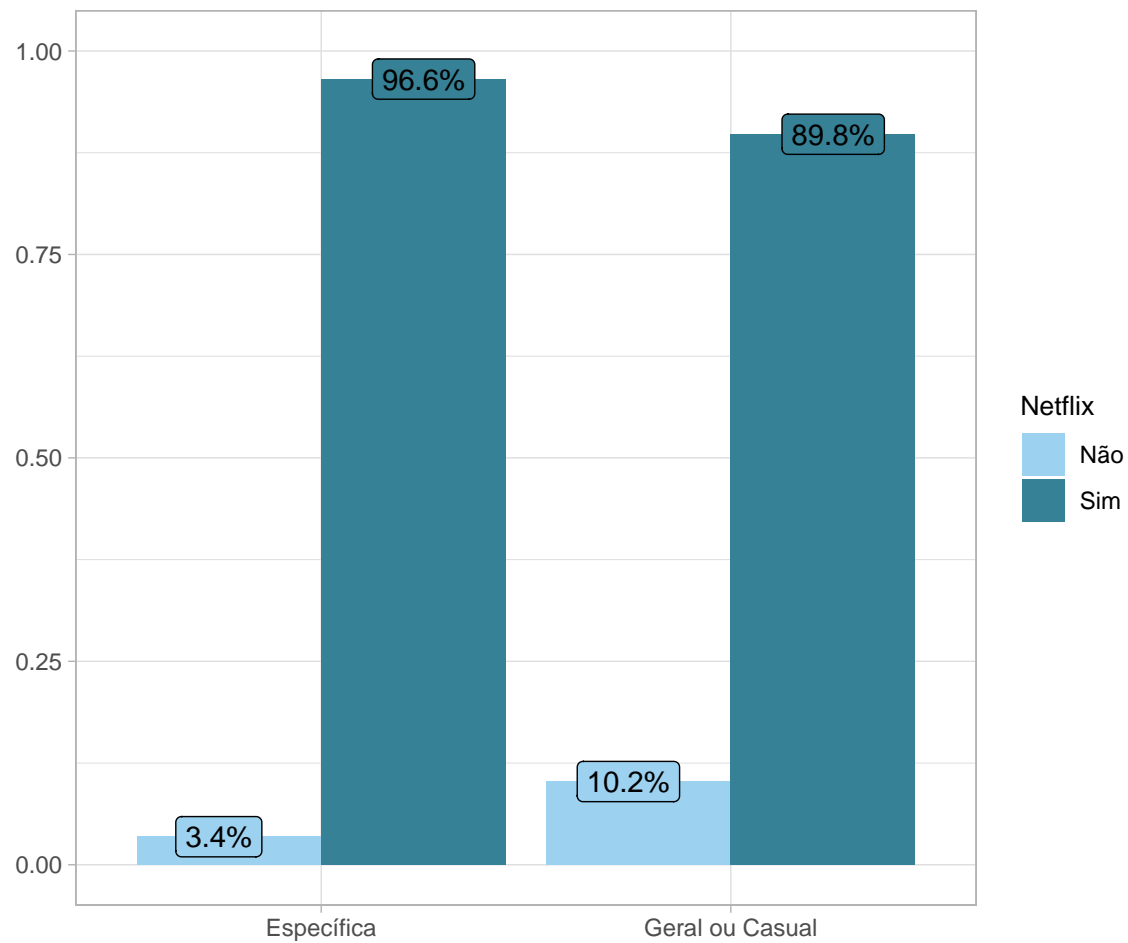
## Valor-p do teste qui-quadrado: 1. Como o valor-p é superior a 0.05, não rejeitamos a hipótese  
## nula e acredita-se que as variáveis são independentes.

**Gráfico 19: Produção audiovisual consumida pelos turistas**



Entre os que citaram “Outros streaming”, descreveram: Pluto TV, Apple TV, Hulu, My Família TV.

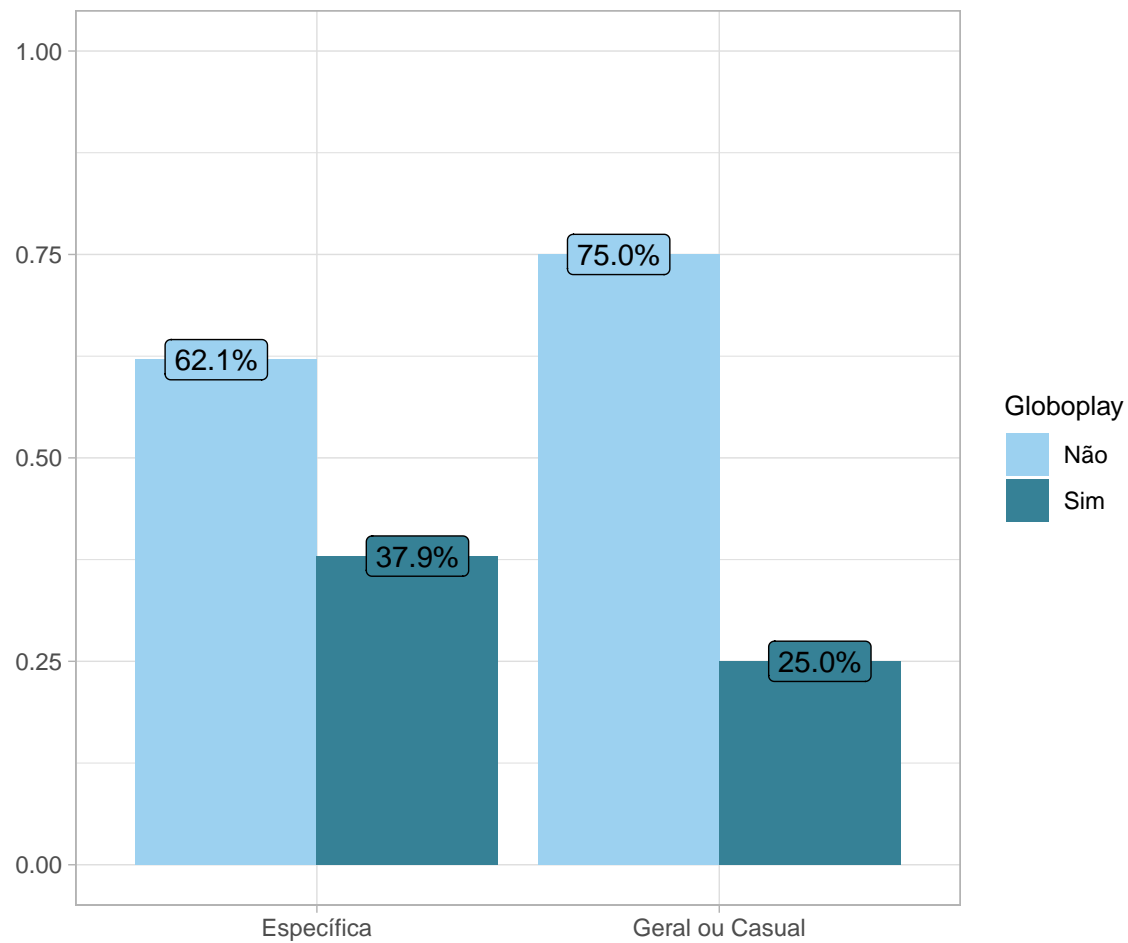
**Gráfico 20: Turistas que assinam Netflix por grupo**



## Valor-p do teste qui-quadrado: 0.4536. Como o valor-p é superior a 0.05, não rejeitamos a hipótese  
## nula e acredita-se que as variáveis são independentes.

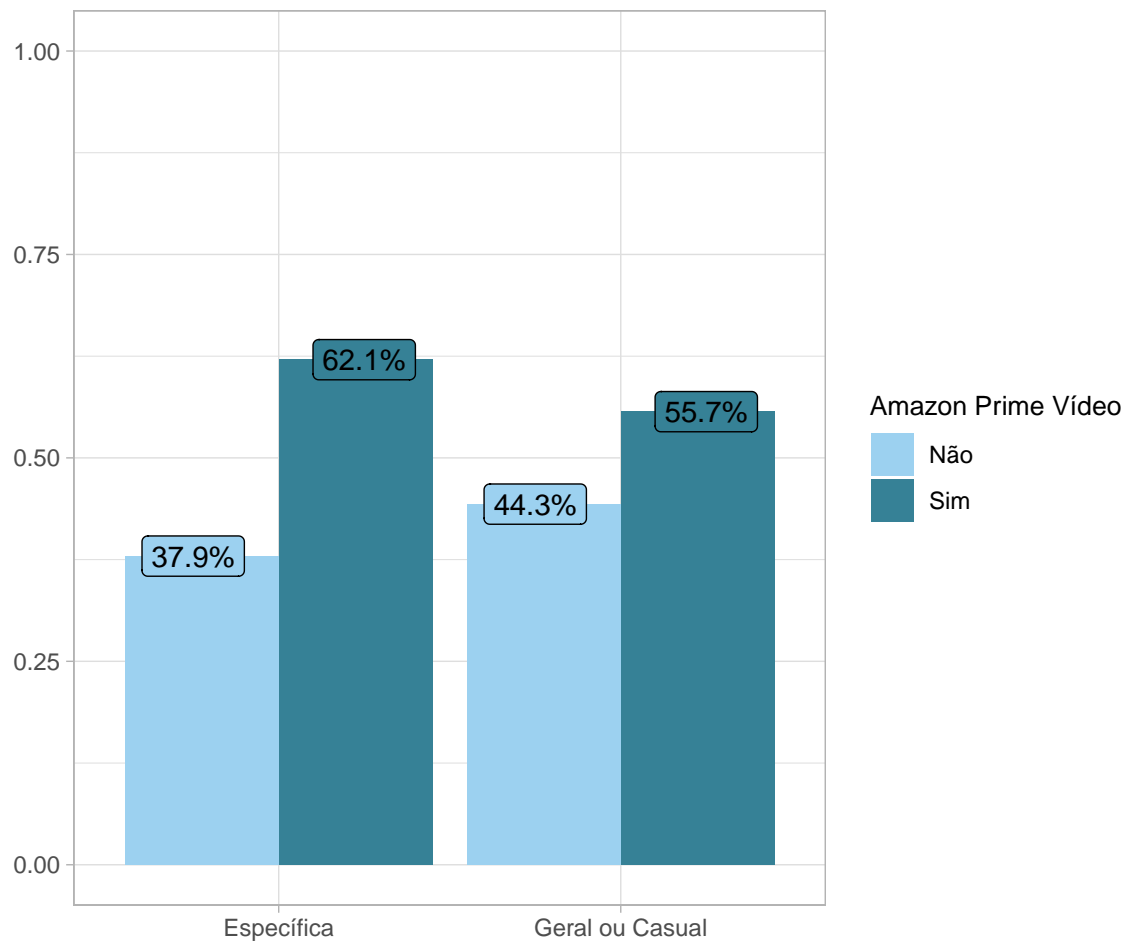


**Gráfico 21: Turistas que assinam Globoplay por grupo**



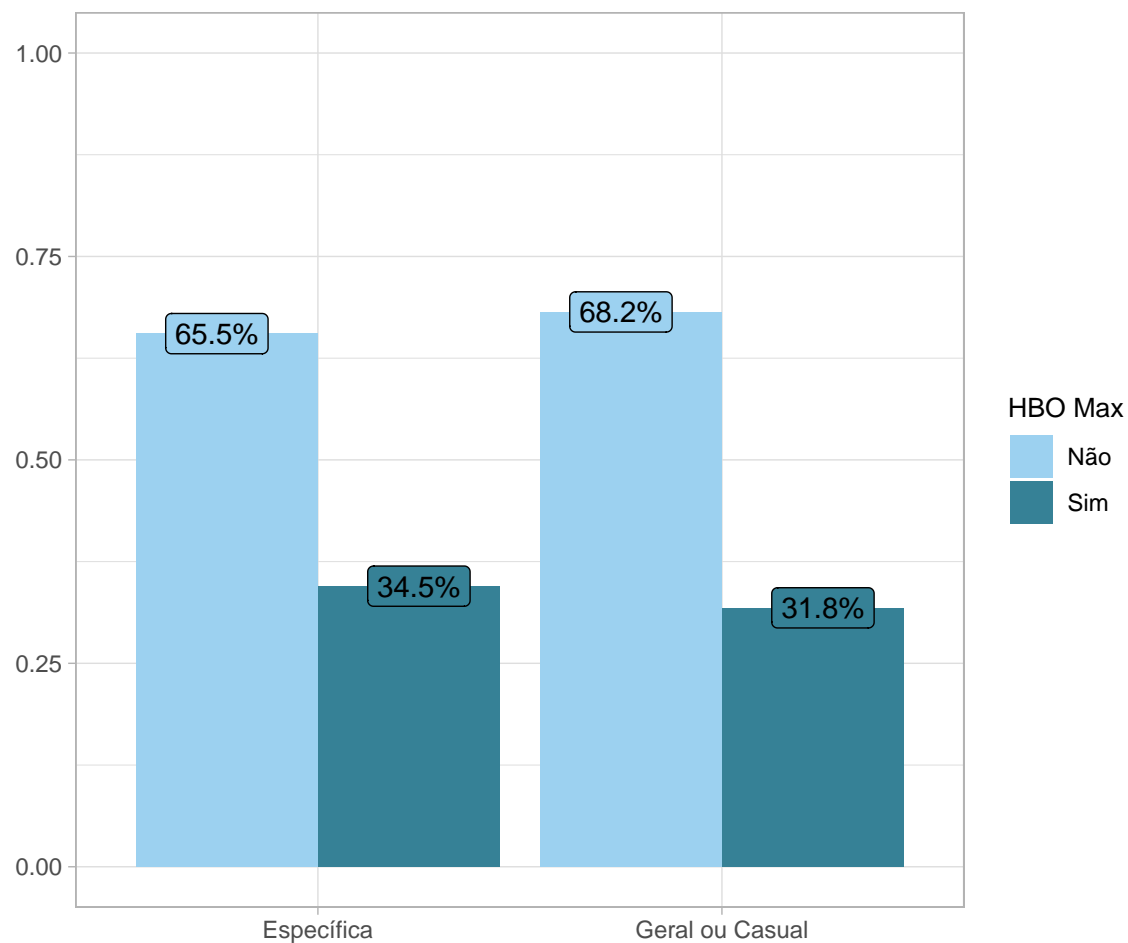
## Valor-p do teste qui-quadrado: 0.2695. Como o valor-p é superior a 0.05, não rejeitamos a hipótese  
## nula e acredita-se que as variáveis são independentes.

**Gráfico 22: Turistas que assinam Amazon Prime Video por grupo**



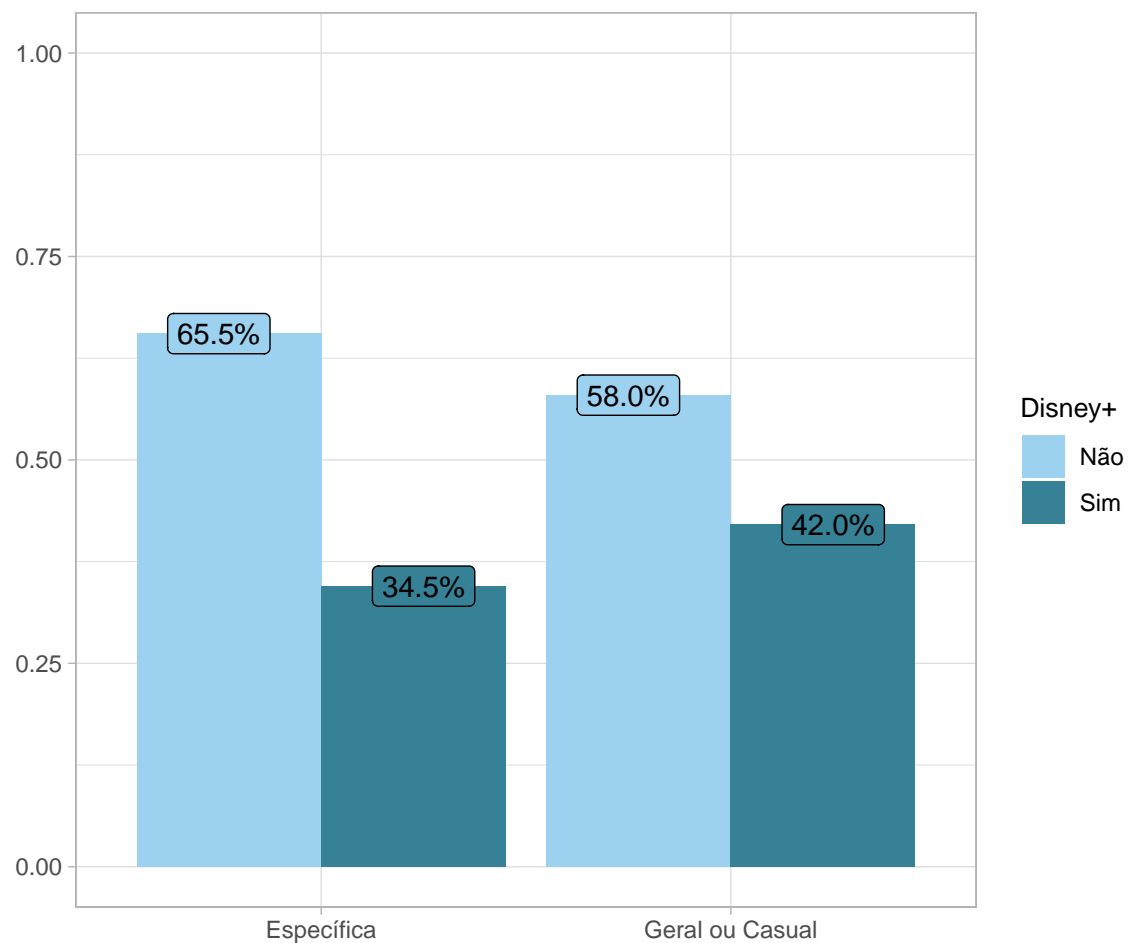
## Valor-p do teste qui-quadrado: 0.6991. Como o valor-p é superior a 0.05, não rejeitamos a hipótese  
## nula e acredita-se que as variáveis são independentes.

**Gráfico 23: Turistas que assinam HBO Max por grupo**



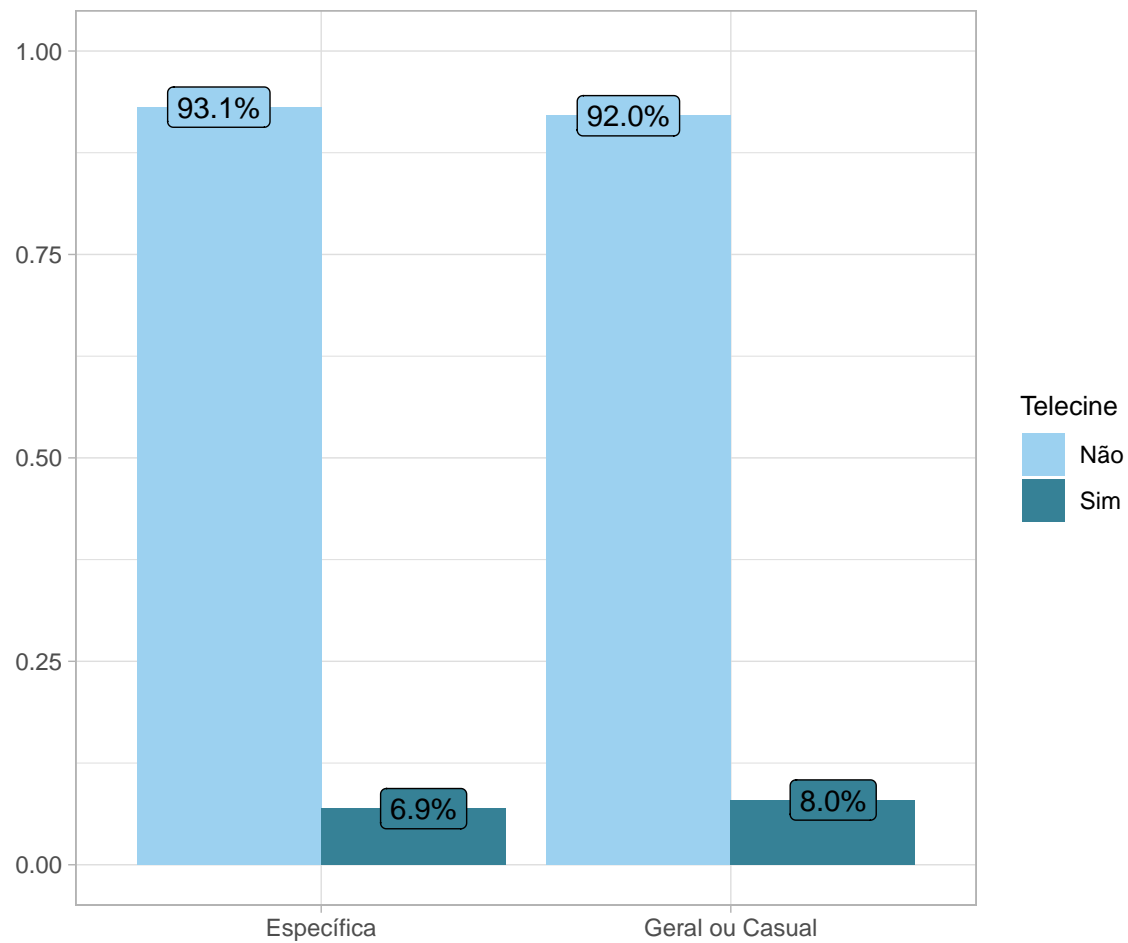
## Valor-p do teste qui-quadrado: 0.9704. Como o valor-p é superior a 0.05, não rejeitamos a hipótese  
## nula e acredita-se que as variáveis são independentes.

**Gráfico 24: Turistas que assinam Disney+ por grupo**



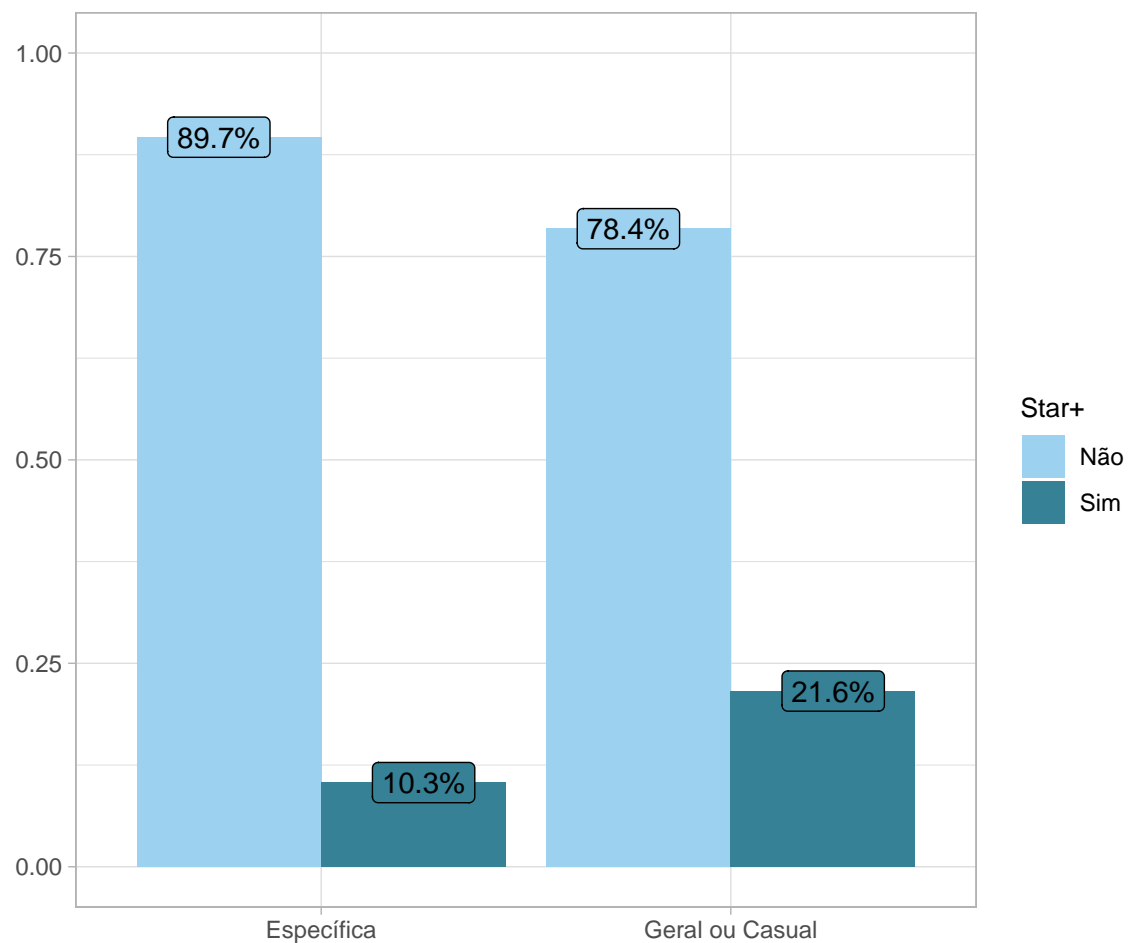
## Valor-p do teste qui-quadrado: 0.6156. Como o valor-p é superior a 0.05, não rejeitamos a hipótese  
## nula e acredita-se que as variáveis são independentes.

**Gráfico 25: Turistas que assinam Telecine por grupo**



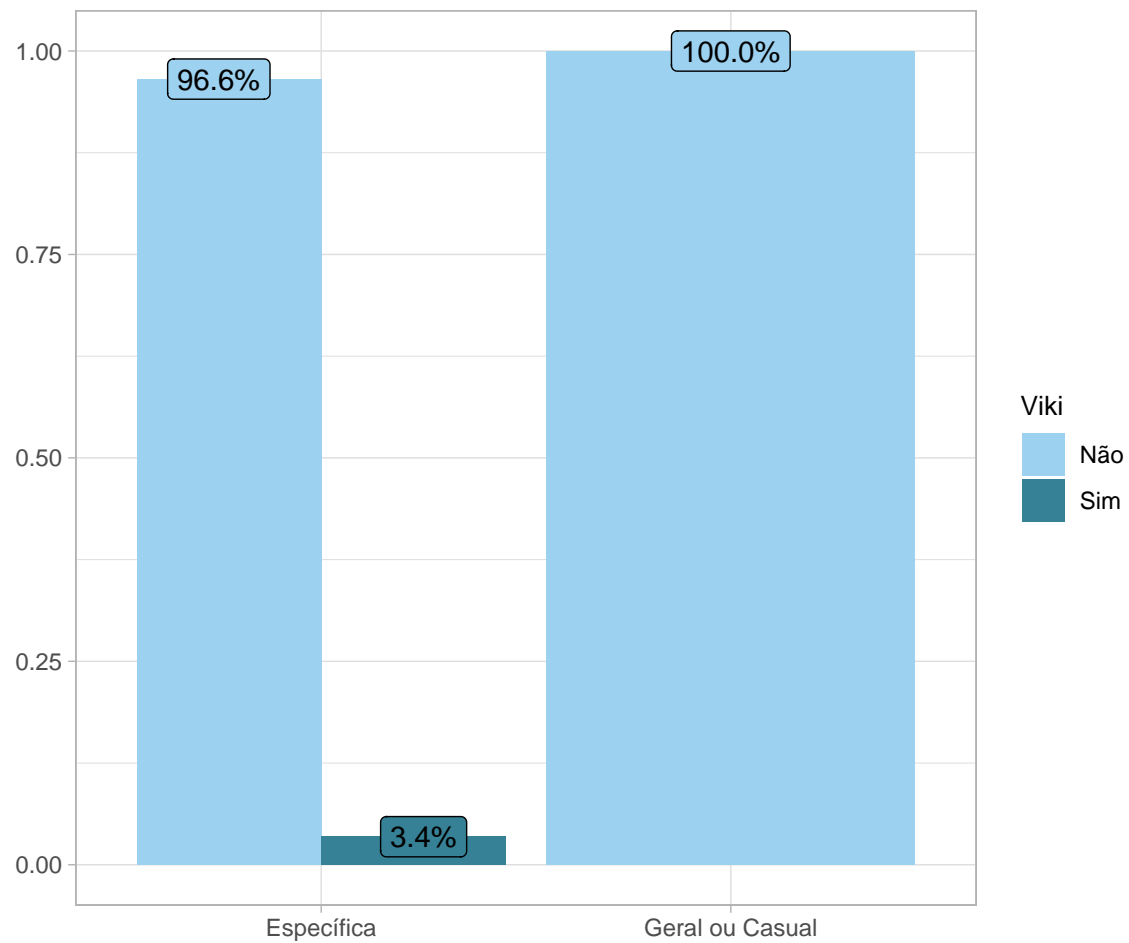
## Valor-p do teste qui-quadrado: 1. Como o valor-p é superior a 0.05, não rejeitamos a hipótese  
## nula e acredita-se que as variáveis são independentes.

**Gráfico 26: Turistas que assinam Star+ por grupo**



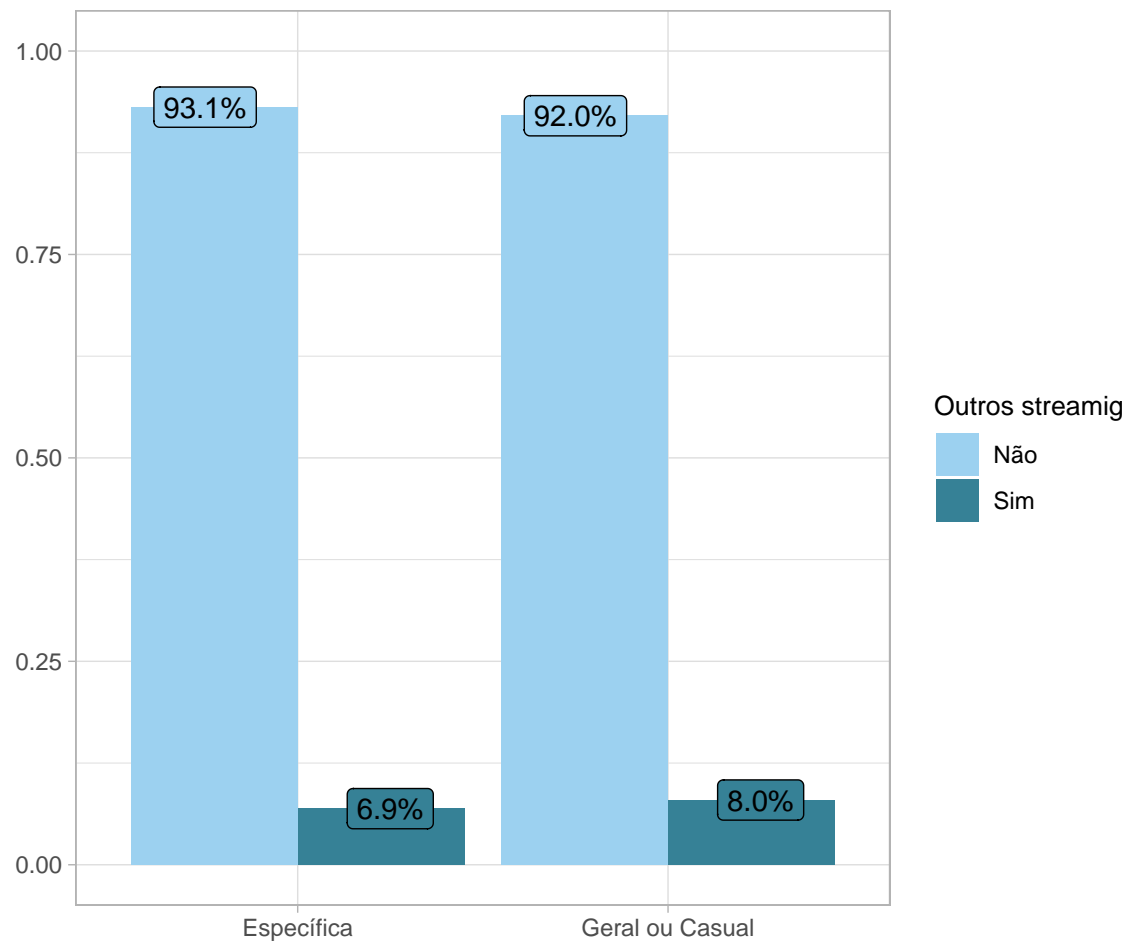
## Valor-p do teste qui-quadrado: 0.2845. Como o valor-p é superior a 0.05, não rejeitamos a hipótese  
## nula e acredita-se que as variáveis são independentes.

**Gráfico 27: Turistas que assinam Viki por grupo**



## Valor-p do teste qui-quadrado: 0.5576. Como o valor-p é superior a 0.05, não rejeitamos a hipótese  
## nula e acredita-se que as variáveis são independentes.

**Gráfico 28: Turistas que assinam Outros streamig por grupo**

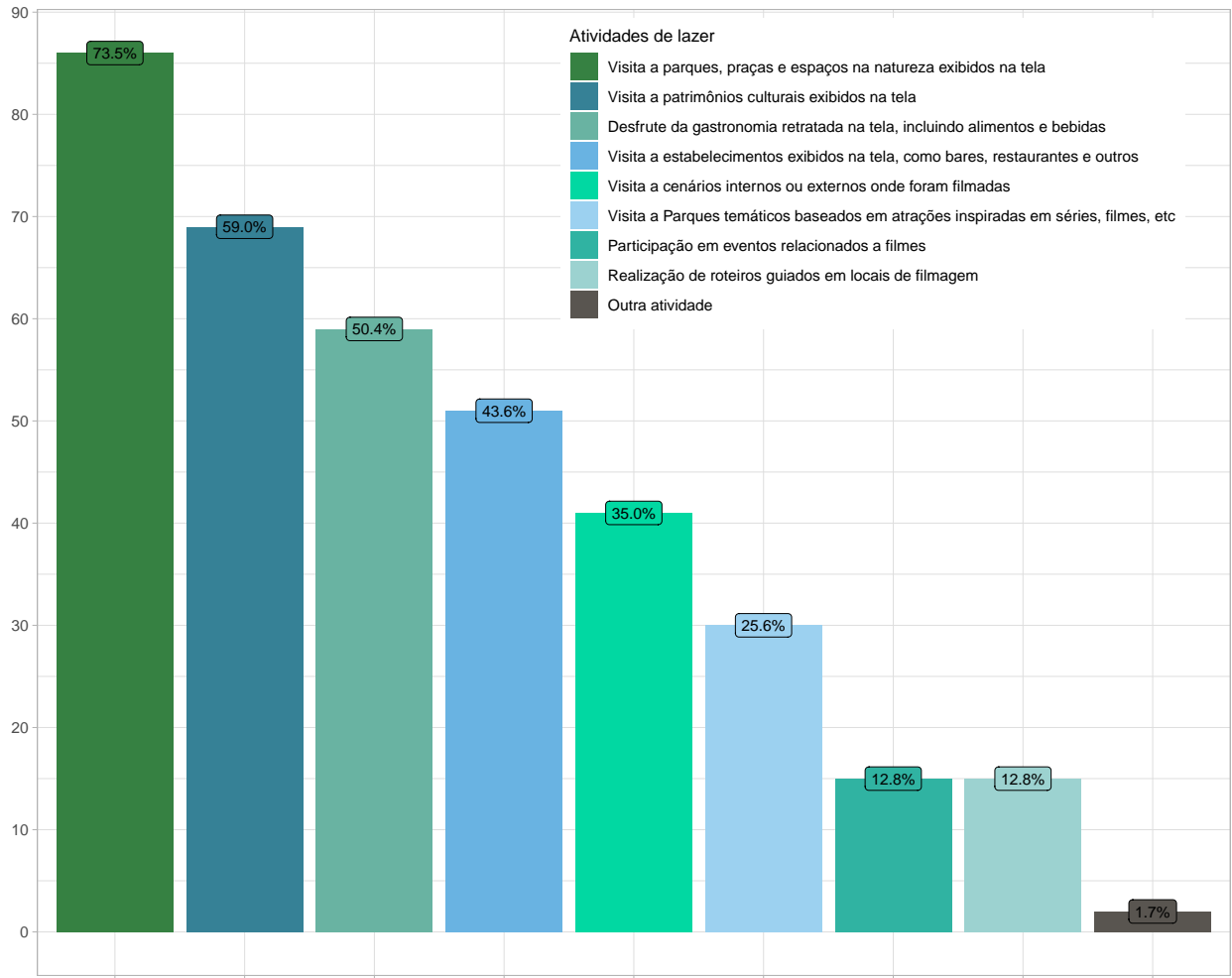


## Valor-p do teste qui-quadrado: 1. Como o valor-p é superior a 0.05, não rejeitamos a hipótese nula e acredita-se que as variáveis são independentes.

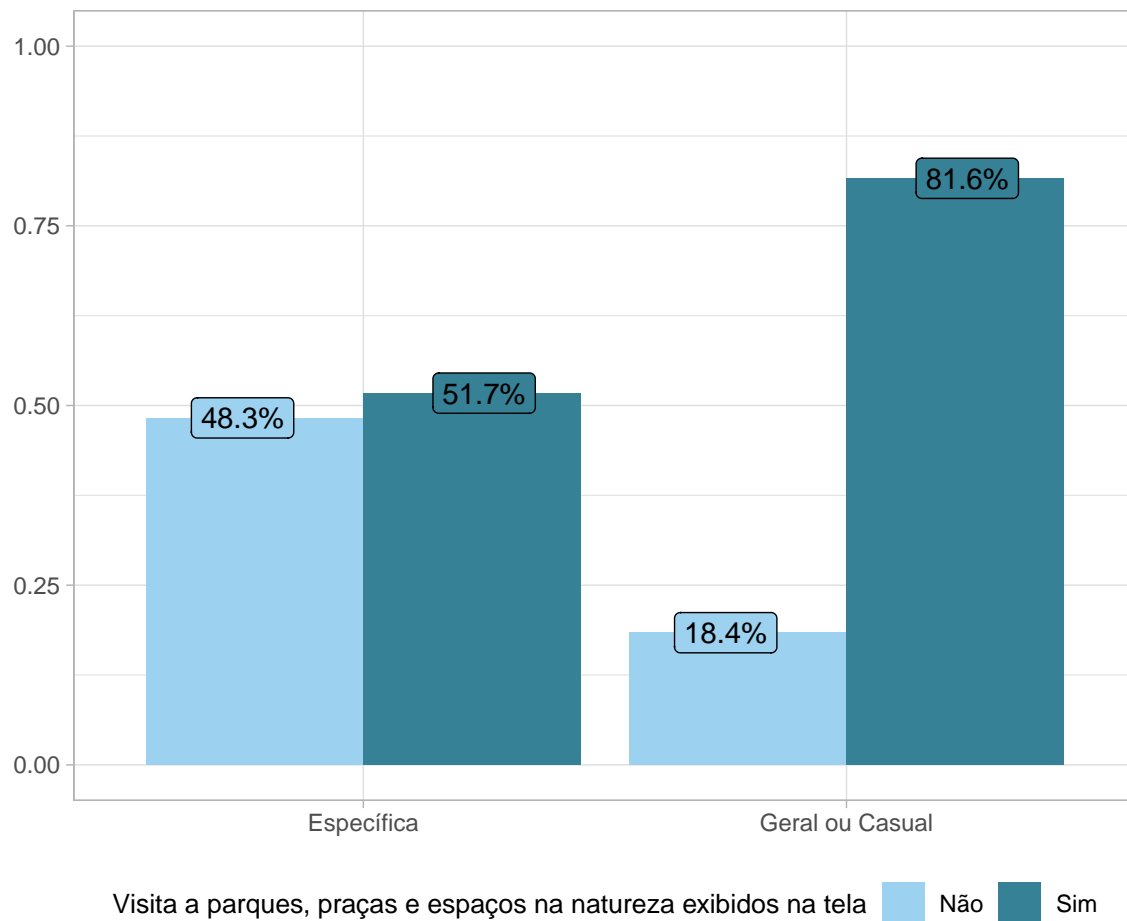
Precisa fazer a mesma coisa feita do result2 para as demais...



Gráfico 29: Atividade de lazer relacionada ao audiovisual realizada pelos turistas

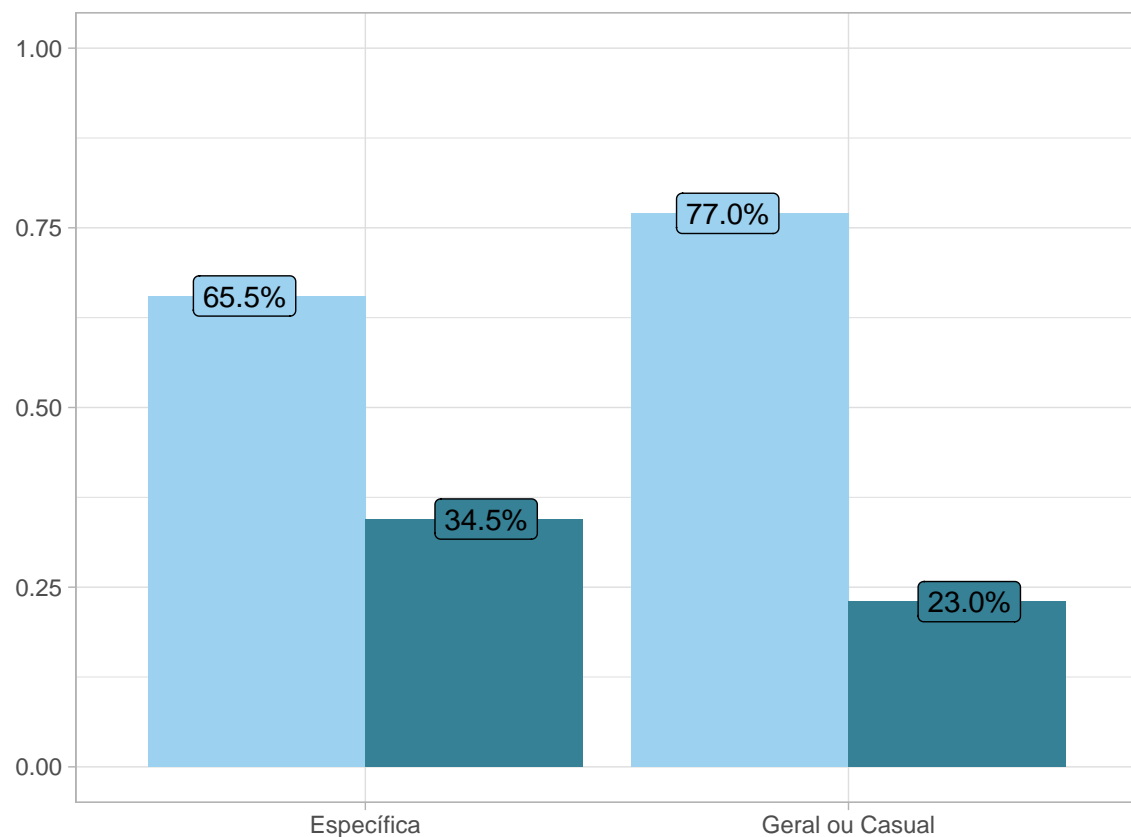


**Gráfico 30: Turistas que realizaram a atividade Visita a parques, praças e espaços na natureza exibidos na tela por grupo**



## Valor-p do teste qui-quadrado: 0.0033. Como o valor-p é inferior a 0.05, rejeitamos a hipótese nula  
## e acredita-se que as variáveis estão associadas.

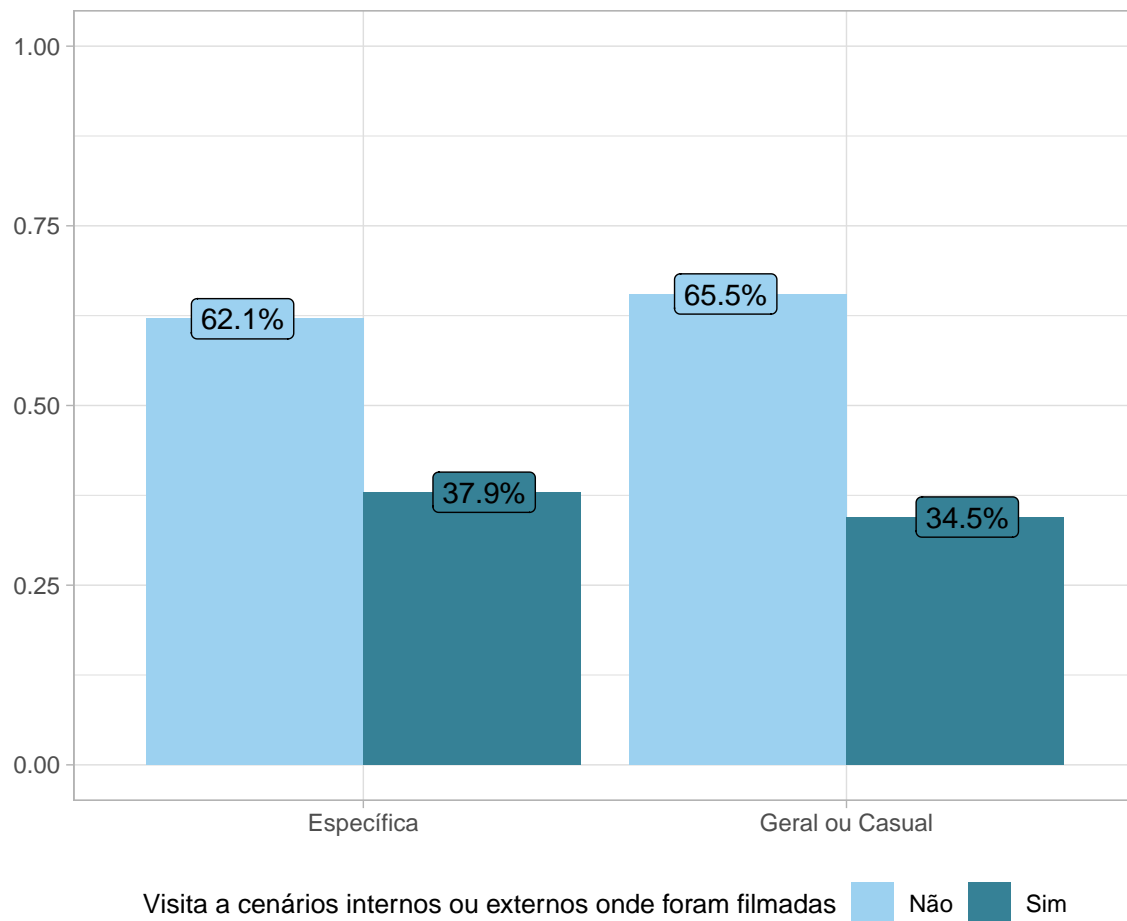
**Gráfico 31: Turistas que realizaram a atividade Visita a Parques temáticos baseados em atrações inspiradas em séries, filmes, etc por grupo**



Visita a Parques temáticos baseados em atrações inspiradas em séries, filmes, etc ■ Não ■ Sim

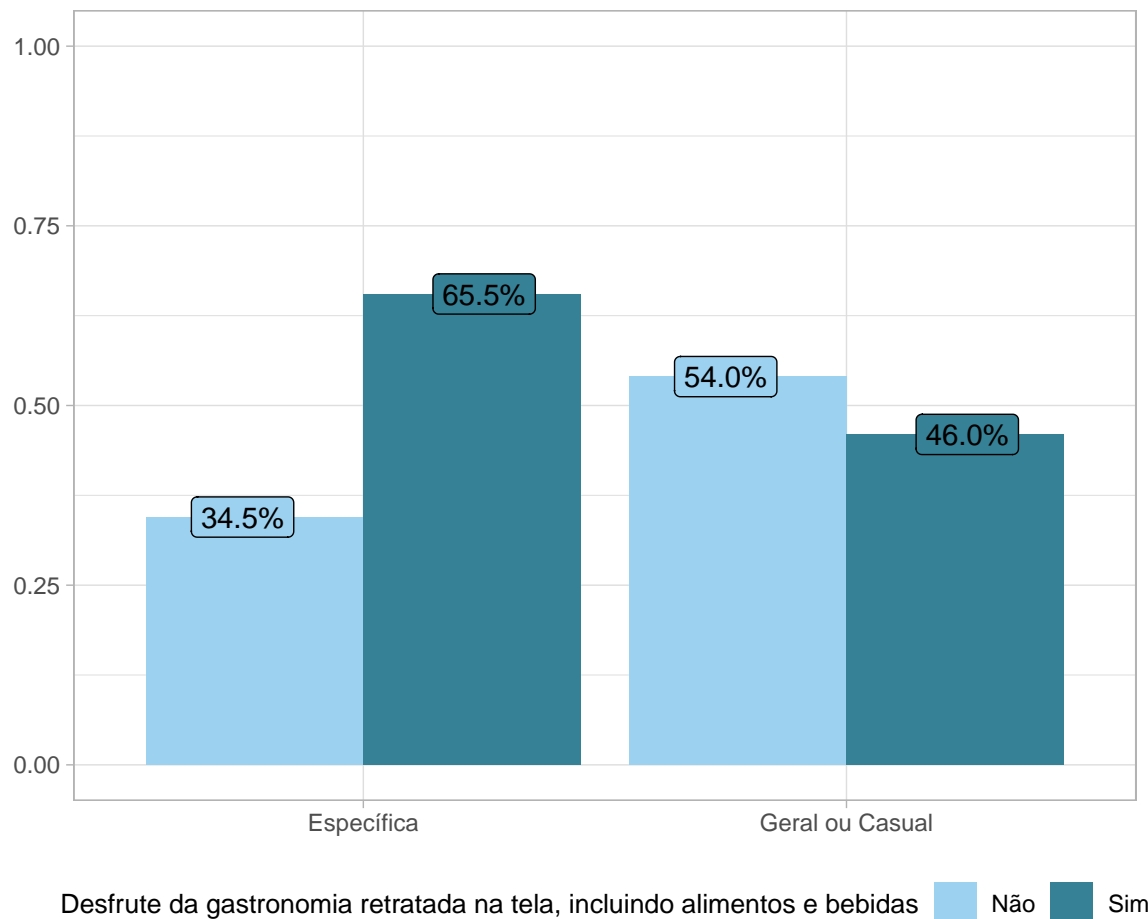
## Valor-p do teste qui-quadrado: 0.3274. Como o valor-p é superior a 0.05, não rejeitamos a hipótese  
## nula e acredita-se que as variáveis são independentes.

**Gráfico 32: Turistas que realizaram a atividade Visita a cenários internos ou externos onde foram filmadas por grupo**



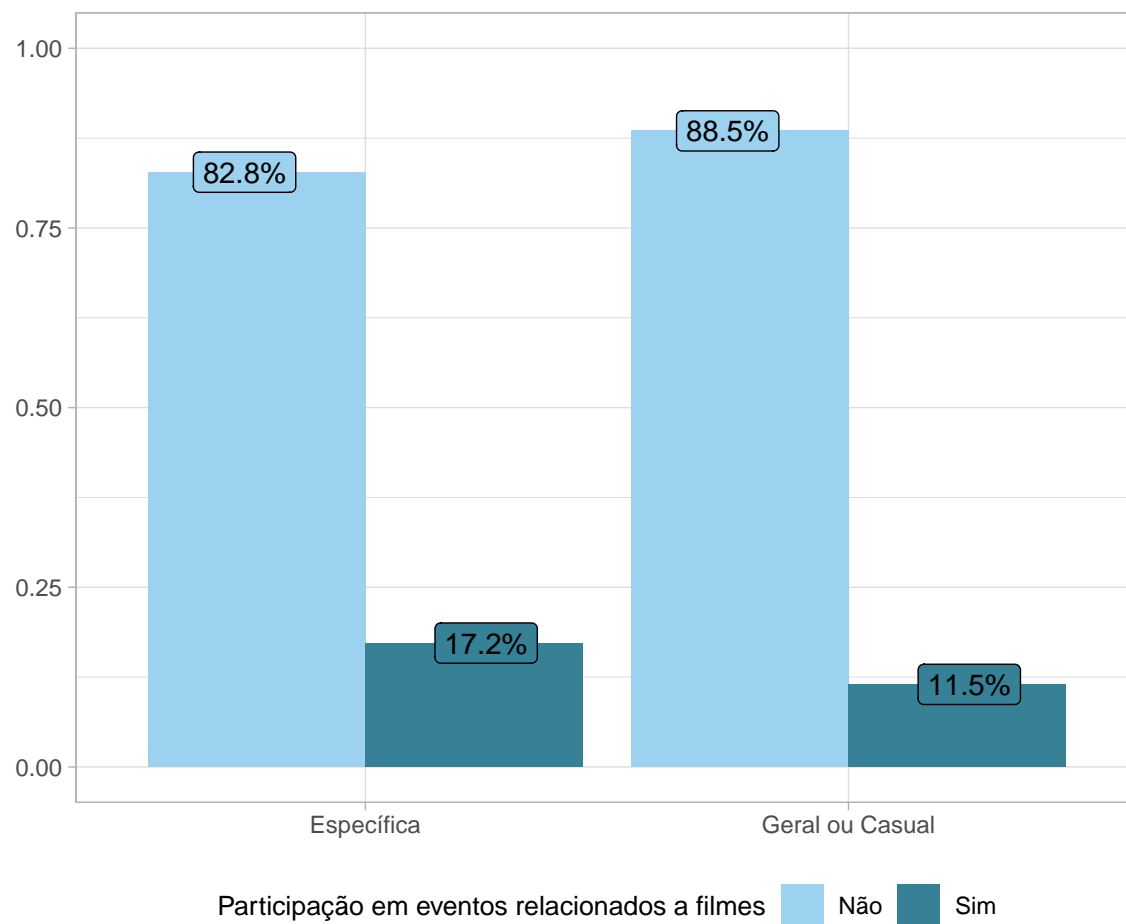
## Valor-p do teste qui-quadrado: 0.9107. Como o valor-p é superior a 0.05, não rejeitamos a hipótese nula e acredita-se que as variáveis são independentes.

**Gráfico 33: Turistas que realizaram a atividade Desfrute da gastronomia retratada na tela, incluindo alimentos e bebidas por grupo**



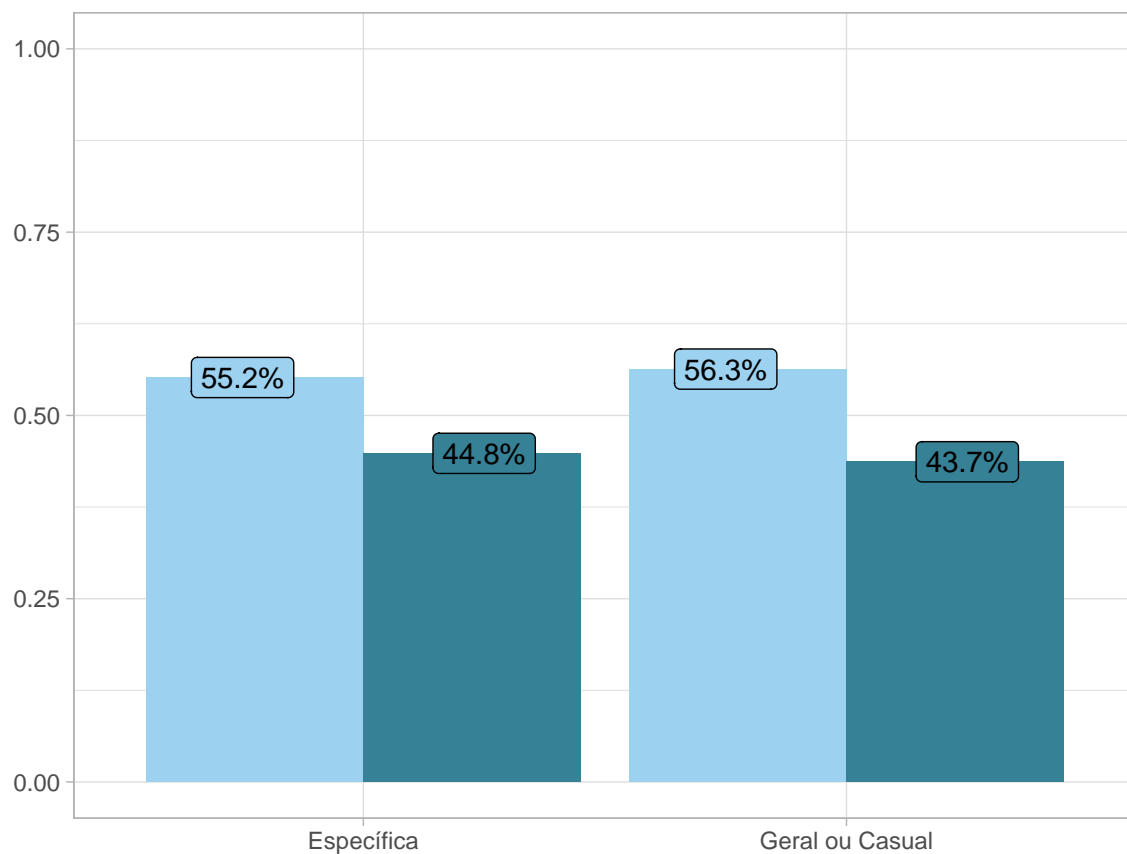
## Valor-p do teste qui-quadrado: 0.1077. Como o valor-p é superior a 0.05, não rejeitamos a hipótese nula e acredita-se que as variáveis são independentes.

**Gráfico 34: Turistas que realizaram a atividade Participação em eventos relacionados a filmes por grupo**



## Valor-p do teste qui-quadrado: 0.6317. Como o valor-p é superior a 0.05, não rejeitamos a hipótese nula e acredita-se que as variáveis são independentes.

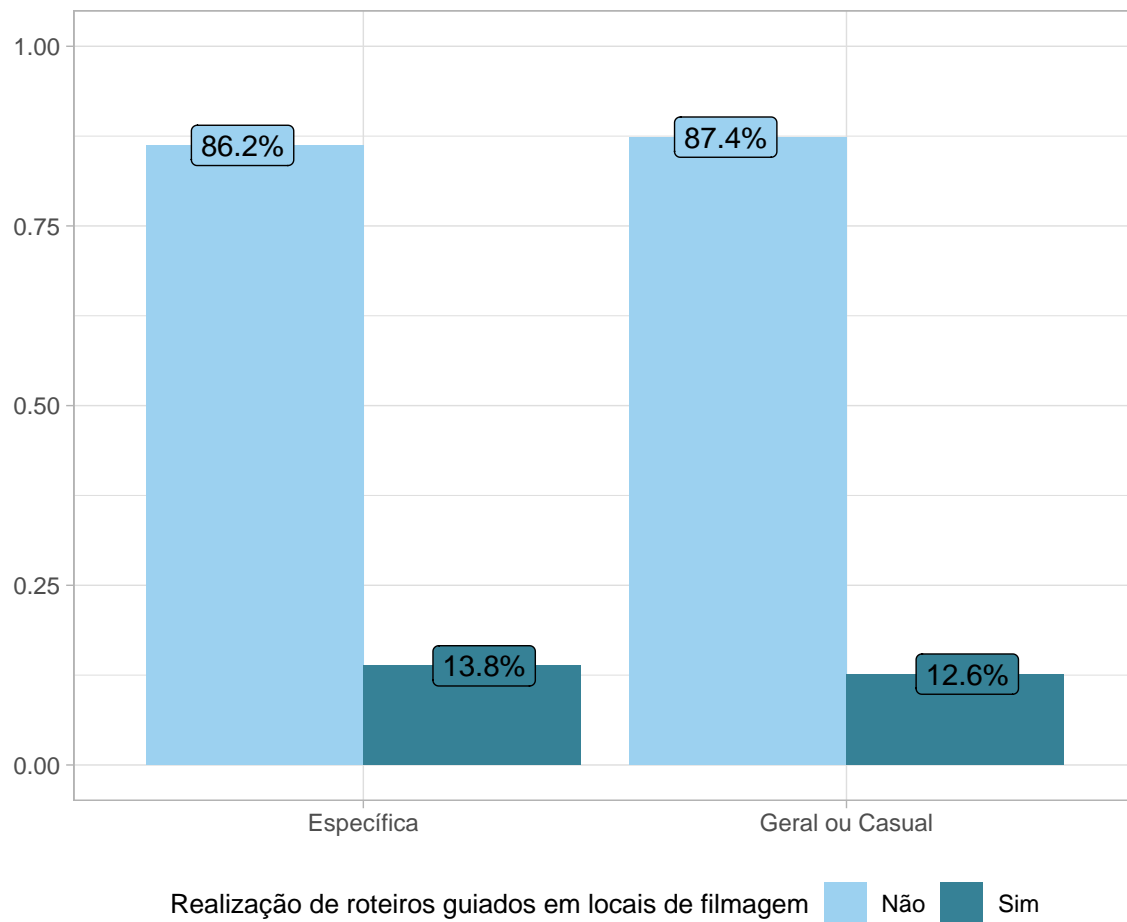
**Gráfico 35: Turistas que realizaram a atividade Visita a estabelecimentos exibidos na tela, como bares, restaurantes e outros por grupo**



Visita a estabelecimentos exibidos na tela, como bares, restaurantes e outros Não Sim

## Valor-p do teste qui-quadrado: 1. Como o valor-p é superior a 0.05, não rejeitamos a hipótese nula e acredita-se que as variáveis são independentes.

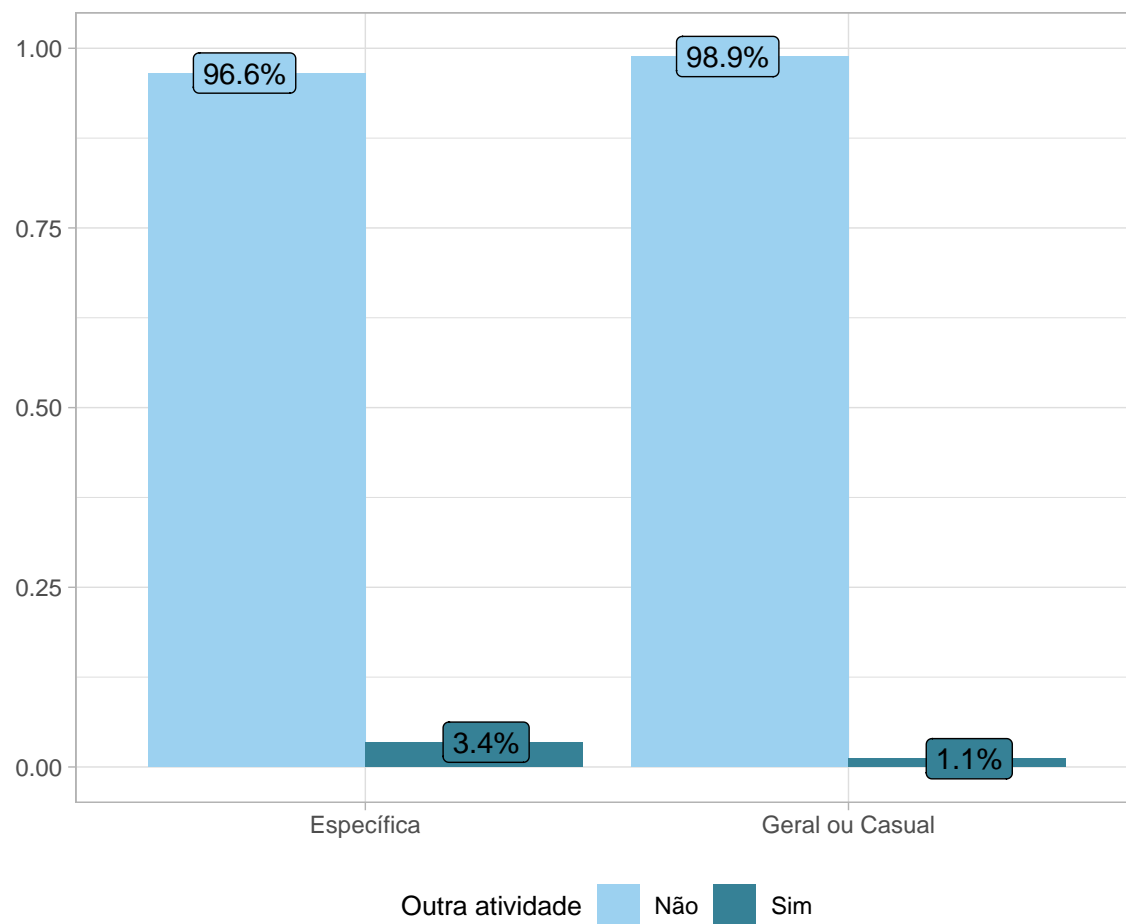
**Gráfico 36: Turistas que realizaram a atividade Realização de roteiros guiados em locais de filmagem por grupo**



## Valor-p do teste qui-quadrado: 1. Como o valor-p é superior a 0.05, não rejeitamos a hipótese nula e acredita-se que as variáveis são independentes.

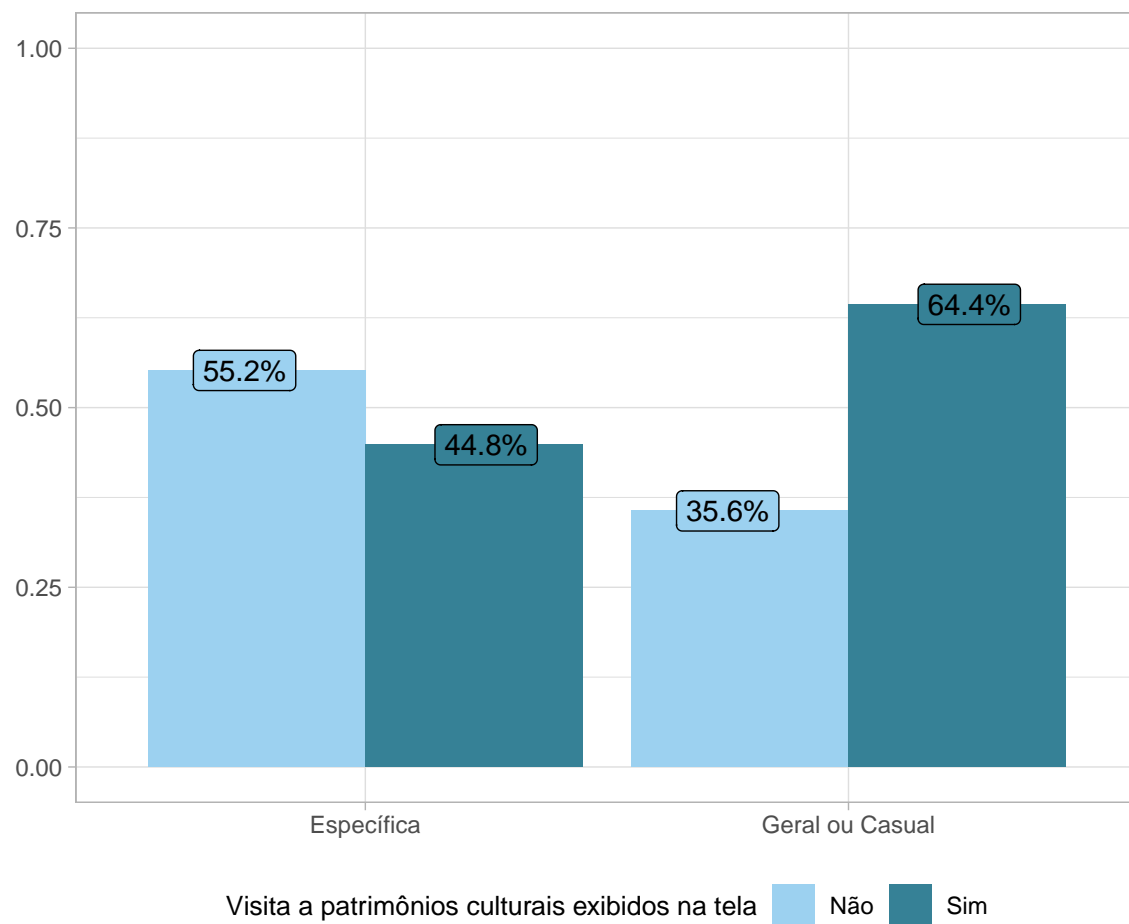


**Gráfico 37: Turistas que realizaram a atividade Outra atividade por grupo**



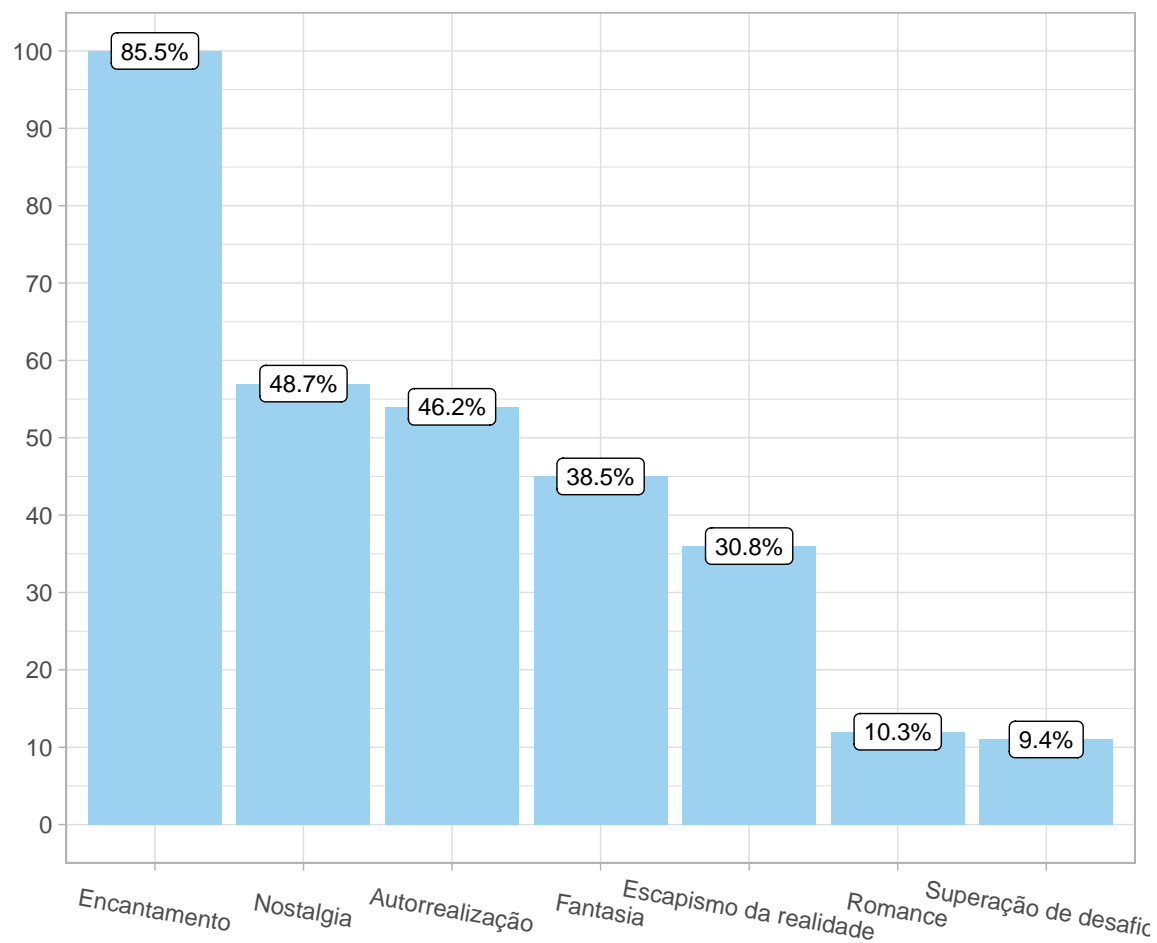
## Valor-p do teste qui-quadrado: 1. Como o valor-p é superior a 0.05, não rejeitamos a hipótese nula e acredita-se que as variáveis são independentes.

**Gráfico 38: Turistas que realizaram a atividade Visita a patrimônios culturais exibidos na tela por grupo**



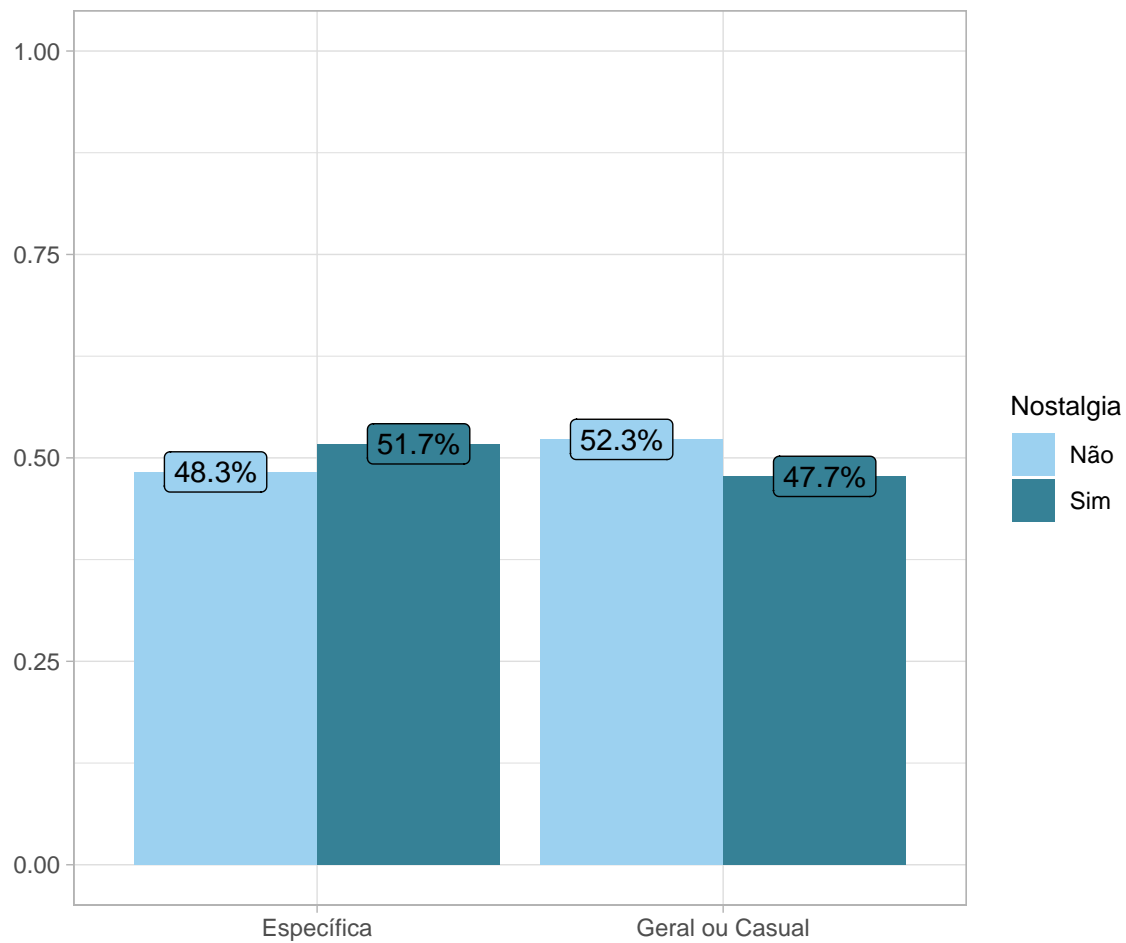
## Valor-p do teste qui-quadrado: 0.1014. Como o valor-p é superior a 0.05, não rejeitamos a hipótese nula e acredita-se que as variáveis são independentes.

**Gráfico 39: Produção audiovisual consumida pelos turistas**



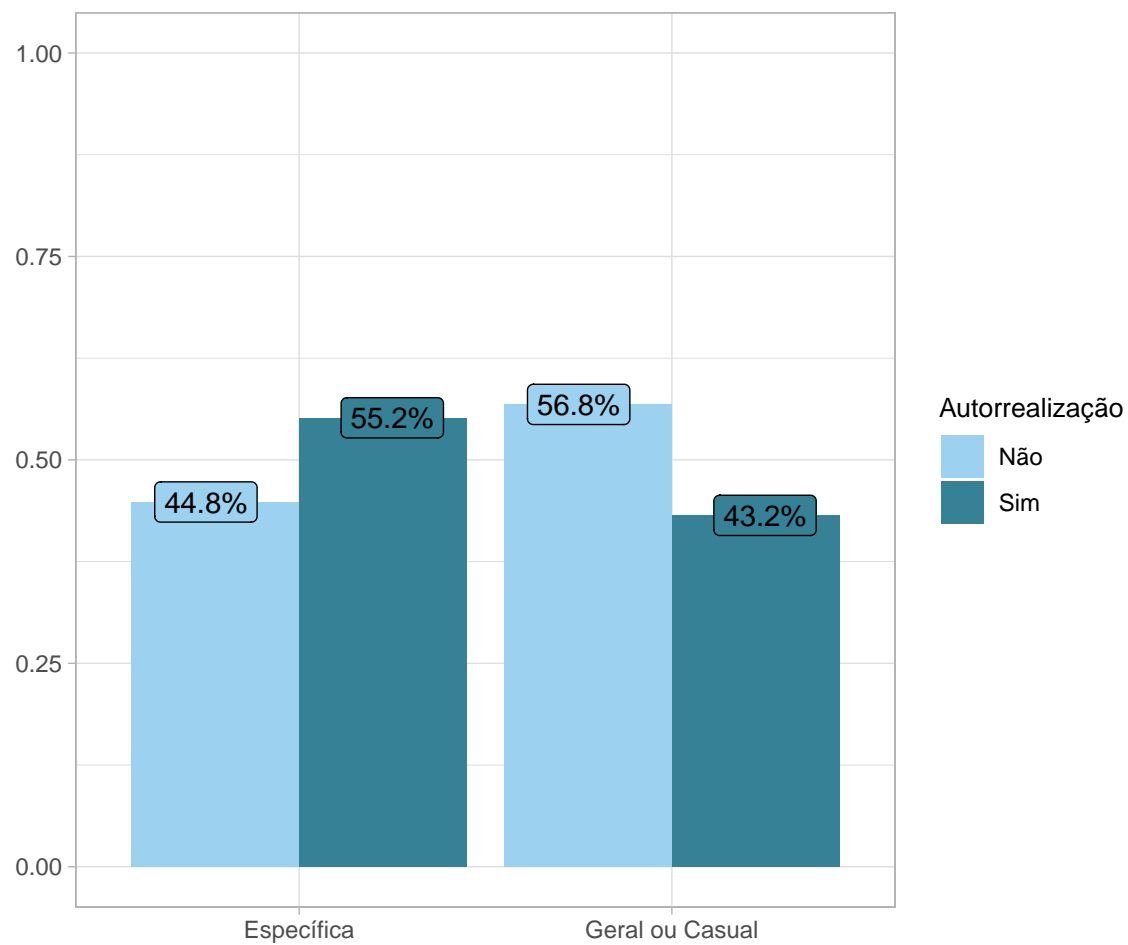
Entre os que citaram “Outros streaming”, descreveram: Pluto TV, Apple TV, Hulu, My Família TV.

**Gráfico 40: Turistas que teve o sentimento nostalgia despertado durante a viagem por grupo**



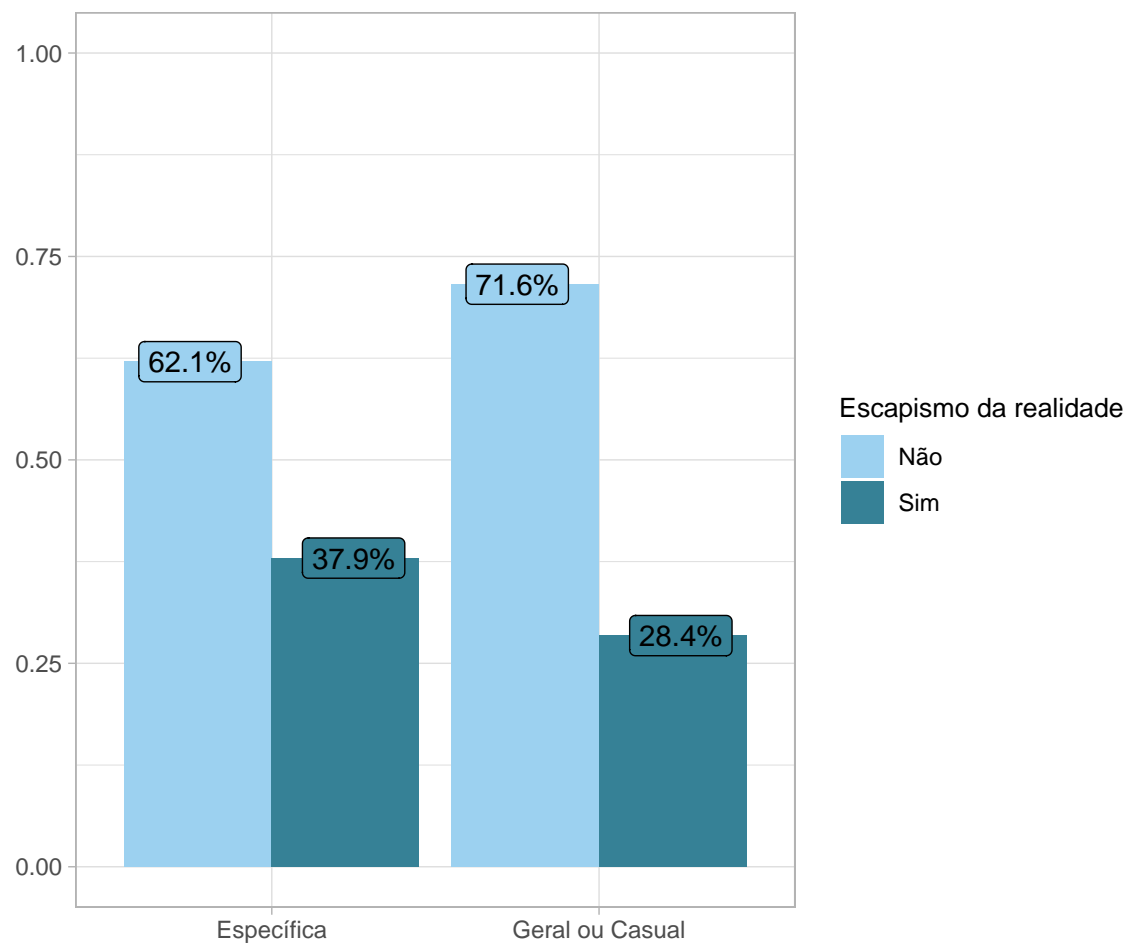
## Valor-p do teste qui-quadrado: 0.8735. Como o valor-p é superior a 0.05, não rejeitamos a hipótese  
## nula e acredita-se que as variáveis são independentes.

**Gráfico 41: Turistas que teve o sentimento autorrealização despertado durante a viagem por grupo**



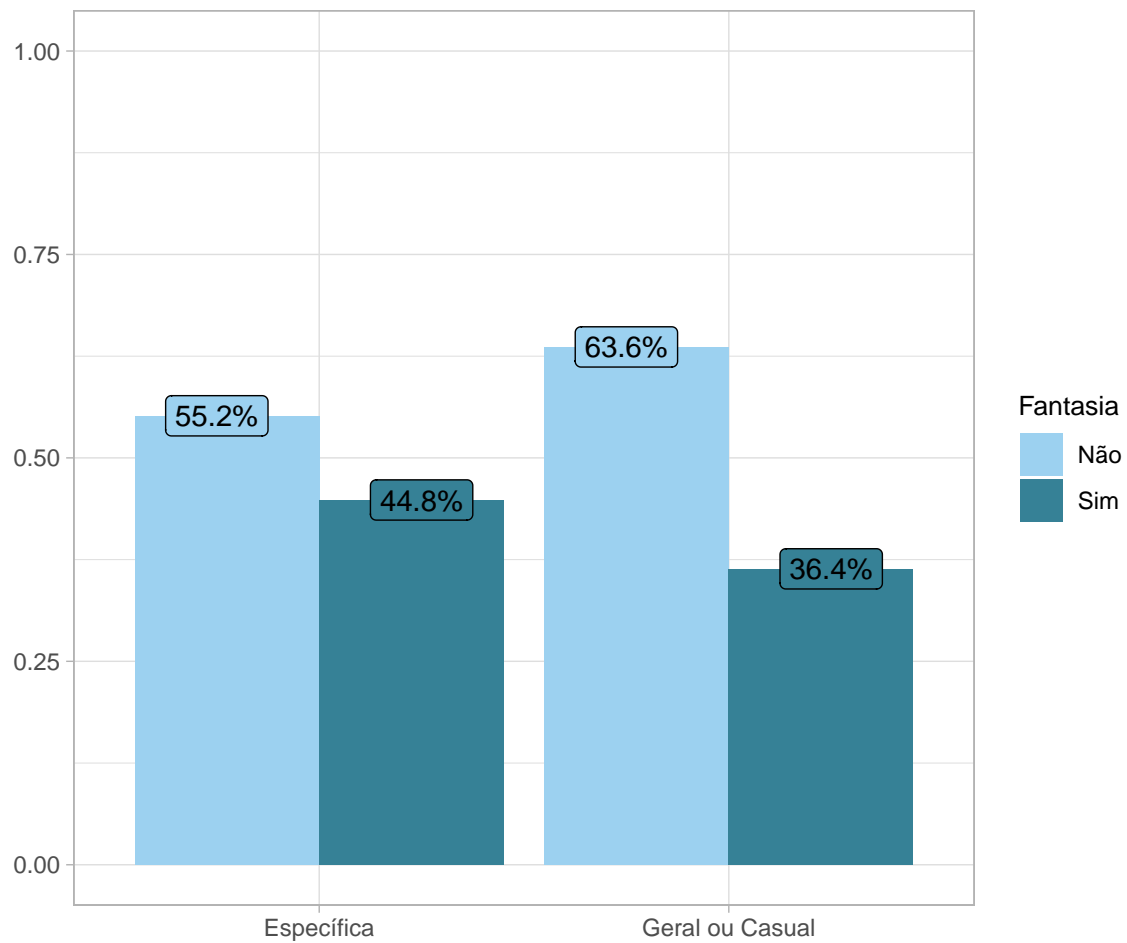
## Valor-p do teste qui-quadrado: 0.3636. Como o valor-p é superior a 0.05, não rejeitamos a hipótese nula e acredita-se que as variáveis são independentes.

**Gráfico 42: Turistas que teve o sentimento escapismo da realidade despertado durante a viagem por grupo**



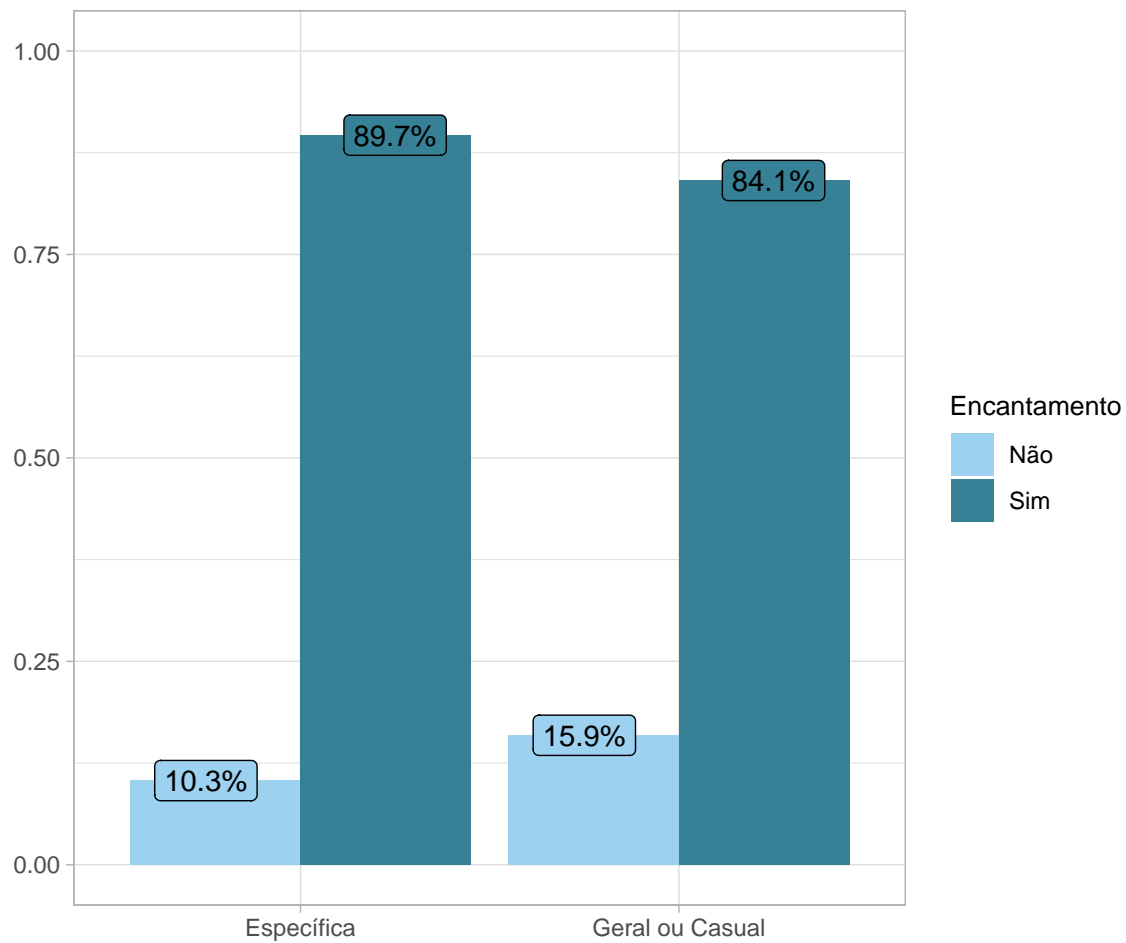
## Valor-p do teste qui-quadrado: 0.4644. Como o valor-p é superior a 0.05, não rejeitamos a hipótese nula e acredita-se que as variáveis são independentes.

**Gráfico 43: Turistas que teve o sentimento fantasia despertado durante a viagem por grupo**



## Valor-p do teste qui-quadrado: 0.5535. Como o valor-p é superior a 0.05, não rejeitamos a hipótese  
## nula e acredita-se que as variáveis são independentes.

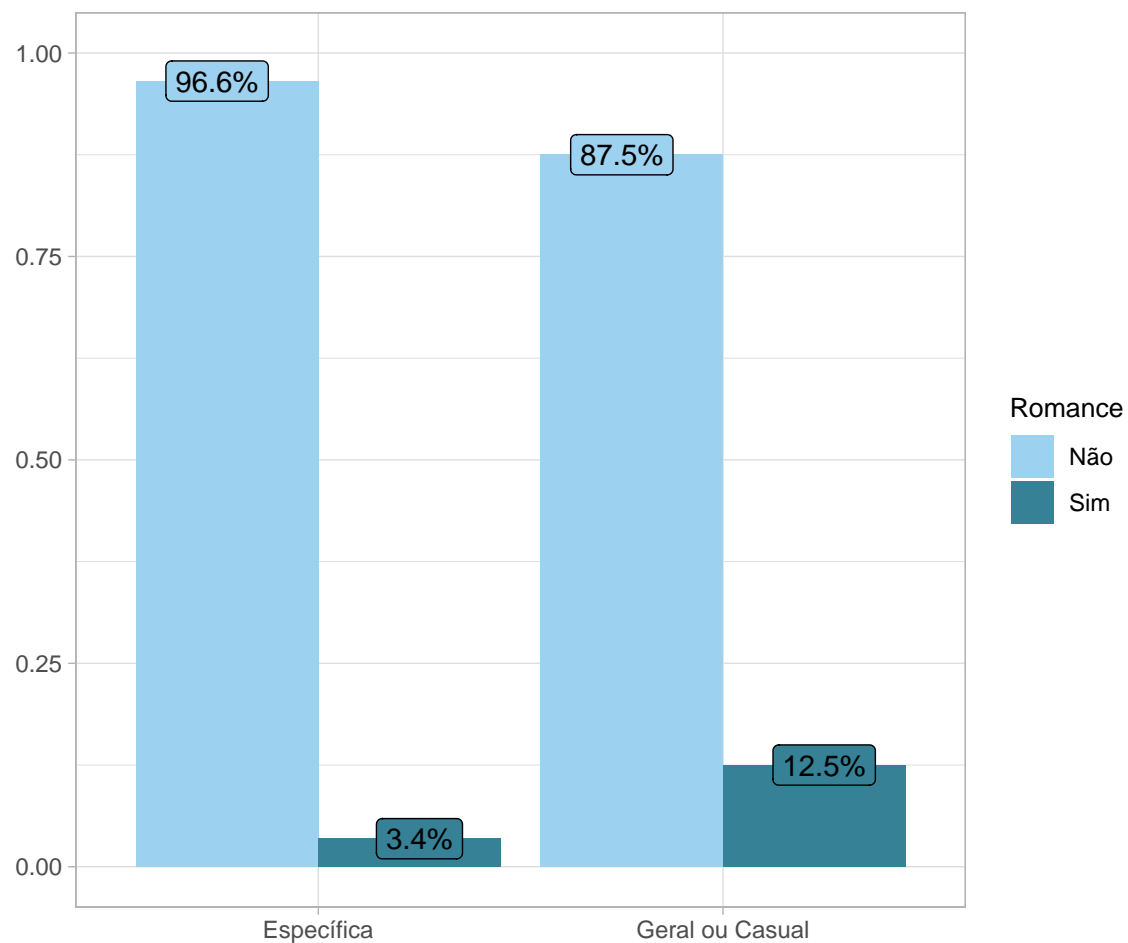
**Gráfico 44: Turistas que teve o sentimento encantamento despertado durante a viagem por grupo**



## Valor-p do teste qui-quadrado: 0.6646. Como o valor-p é superior a 0.05, não rejeitamos a hipótese  
## nula e acredita-se que as variáveis são independentes.

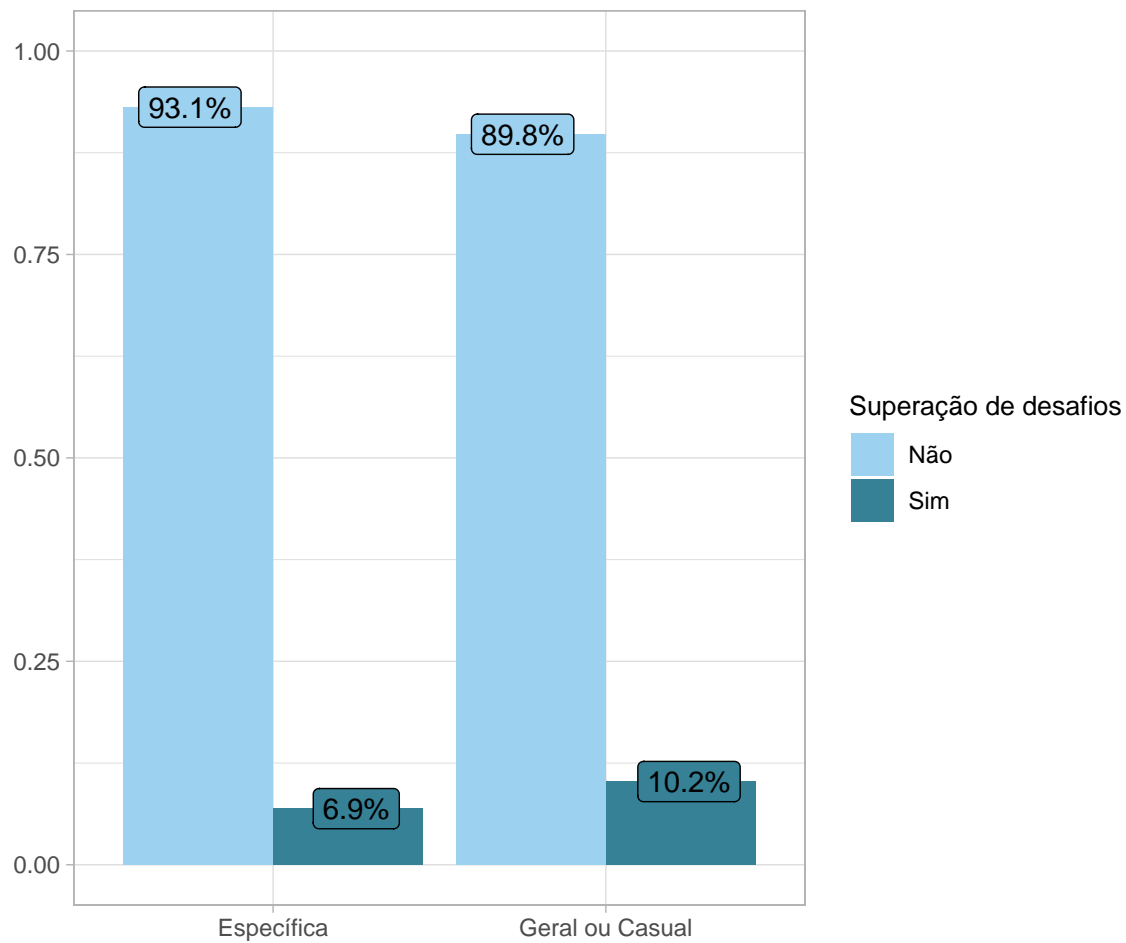


**Gráfico 45: Turistas que teve o sentimento romance despertado durante a viagem por grupo**



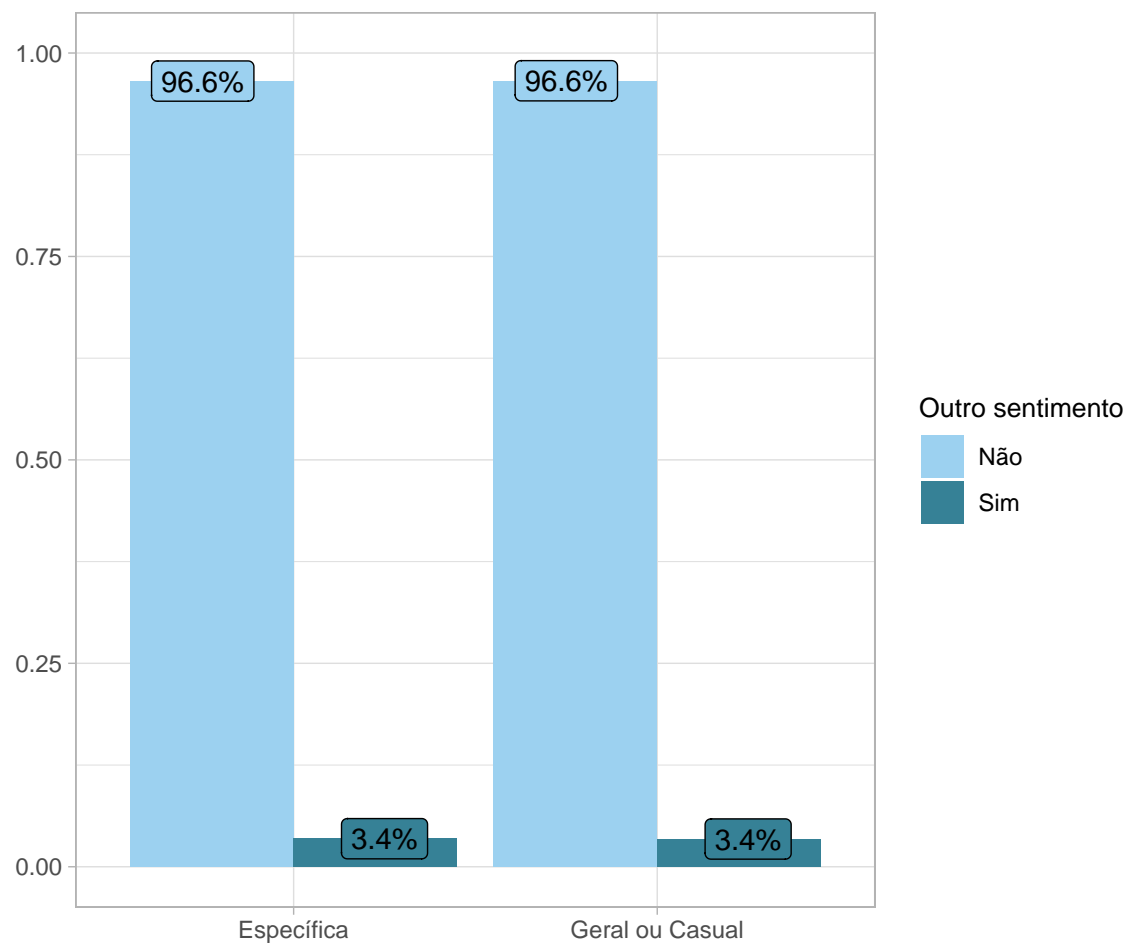
## Valor-p do teste qui-quadrado: 0.2981. Como o valor-p é superior a 0.05, não rejeitamos a hipótese nula e acredita-se que as variáveis são independentes.

**Gráfico 46: Turistas que teve o sentimento supeação de desafios despertados durante a viagem por grupo**



## Valor-p do teste qui-quadrado: 0.868. Como o valor-p é superior a 0.05, não rejeitamos a hipótese  
## nula e acredita-se que as variáveis são independentes.

**Gráfico 47: Turistas que teve o outros sentimentos despertados durante a viagem por grupo**



## Valor-p do teste qui-quadrado: 1. Como o valor-p é superior a 0.05, não rejeitamos a hipótese nula e acredita-se que as variáveis são independentes.

**Tabela 1 - Resumo estatística descritiva para cada variável**

Variável	Nível	Qtde	Percentual
Sexo	Feminino	80	68.4%
Sexo	Masculino	37	31.6%
Idade	Entre 18 - 30	47	40.2%
Idade	Entre 31 - 40	43	36.8%
Idade	Mais de 41	27	23.1%
Escolaridade	Ensino Médio	87	74.4%
Escolaridade	Ensino Superior	29	24.8%
Região residência	Centro - Oeste	8	6.8%
Região residência	Nordeste	16	13.7%
Região residência	Norte	5	4.3%
Região residência	Resido no exterior	9	7.7%
Região residência	Sudeste	59	50.4%

Variável	Nível	Qtde	Percentual
Região residência	Sul	20	17.1%
Frequência audiovisuais	Poucas vezes por semana	34	29.1%
Frequência audiovisuais	Várias vezes por semana	83	70.9%
Assinatura streaming	Não	4	3.4%
Assinatura streaming	Sim	112	95.7%
Grupo	Específica	29	24.8%
Grupo	Geral ou Casual	88	75.2%
Motivação	Conhecer lugares retratados nas telas, incluindo cenários, paisagens, clima, culturas, etc.	99	84.6%
Motivação	Conhecer personagens (Celebidades).	2	1.7%
Motivação	Obter mais informações sobre o filme: enredo, trama, gênero, etc.	4	3.4%
Motivação	Outros	10	8.5%
Percepção	Não	1	0.9%
Percepção	Sim	82	70.1%
Percepção	Parcialmente	33	28.2%
Companhia	Outros	5	4.3%
Companhia	Amigos	31	26.5%
Companhia	Colegas de Trabalho	1	0.9%
Companhia	Com companheiro/a (namorado/a, noivo/a, esposo/a)	21	17.9%
Companhia	Familiares	38	32.5%
Companhia	Sozinho	21	17.9%
Satisfação	Muito satisfeito	67	57.3%
Satisfação	Satisfeito/ Indiferente/ Muito insatisfeito	50	42.7%
Realizar outra viagem pelo audiovisual	Não	2	1.7%
Realizar outra viagem pelo audiovisual	Sim	114	97.4%
Filmes	Não	9	7.7%
Filmes	Sim	108	92.3%
Novelas	Não	84	71.8%
Novelas	Sim	33	28.2%
Seriados	Não	19	16.2%
Seriados	Sim	98	83.8%
Clipes de música	Não	81	69.2%
Clipes de música	Sim	36	30.8%
Outro audiovisual	Não	112	95.7%
Outro audiovisual	Sim	5	4.3%
Netflix	Não	10	8.5%
Netflix	Sim	107	91.5%
Globoplay	Não	84	71.8%
Globoplay	Sim	33	28.2%
Amazon Prime Vídeo	Não	50	42.7%
Amazon Prime Vídeo	Sim	67	57.3%
HBO Max	Não	79	67.5%
HBO Max	Sim	38	32.5%
Disney+	Não	70	59.8%
Disney+	Sim	47	40.2%
Telecine	Não	108	92.3%
Telecine	Sim	9	7.7%
Star+	Não	95	81.2%
Star+	Sim	22	18.8%
Viki	Não	116	99.1%

Variável	Nível	Qtde	Percentual
Viki	Sim	1	0.9%
Outros streamig	Não	108	92.3%
Outros streamig	Sim	9	7.7%
Visita a parques, praças e espaços na natureza exibidos na tela	Não	30	25.6%
Visita a parques, praças e espaços na natureza exibidos na tela	Sim	86	73.5%
Visita a Parques temáticos baseados em atrações inspiradas em séries, filmes, etc	Não	86	73.5%
Visita a Parques temáticos baseados em atrações inspiradas em séries, filmes, etc	Sim	30	25.6%
Visita a cenários internos ou externos onde foram filmadas	Não	75	64.1%
Visita a cenários internos ou externos onde foram filmadas	Sim	41	35.0%
Desfrute da gastronomia retratada na tela, incluindo alimentos e bebidas	Não	57	48.7%
Desfrute da gastronomia retratada na tela, incluindo alimentos e bebidas	Sim	59	50.4%
Participação em eventos relacionados a filmes	Não	101	86.3%
Participação em eventos relacionados a filmes	Sim	15	12.8%
Visita a estabelecimentos exibidos na tela, como bares, restaurantes e outros	Não	65	55.6%
Visita a estabelecimentos exibidos na tela, como bares, restaurantes e outros	Sim	51	43.6%
Realização de roteiros guiados em locais de filmagem	Não	101	86.3%
Realização de roteiros guiados em locais de filmagem	Sim	15	12.8%
Outra atividade	Não	114	97.4%
Outra atividade	Sim	2	1.7%
Visita a patrimônios culturais exibidos na tela	Não	47	40.2%
Visita a patrimônios culturais exibidos na tela	Sim	69	59.0%
Nostalgia	Não	60	51.3%
Nostalgia	Sim	57	48.7%
Autorrealização	Não	63	53.8%
Autorrealização	Sim	54	46.2%
Escapismo da realidade	Não	81	69.2%
Escapismo da realidade	Sim	36	30.8%
Fantasia	Não	72	61.5%
Fantasia	Sim	45	38.5%
Encantamento	Não	17	14.5%
Encantamento	Sim	100	85.5%
Romance	Não	105	89.7%
Romance	Sim	12	10.3%
Superação de desafios	Não	106	90.6%
Superação de desafios	Sim	11	9.4%
Outro sentimento	Não	113	96.6%
Outro sentimento	Sim	4	3.4%

**Tabela 2 - Resumo teste qui-quadrado descritiva para cada variável**

Variável	p-valor
Visita a parques, praças e espaços na natureza exibidos na tela	0.0033
Filmes	0.0086
Satisfação	0.092
Visita a patrimônios culturais exibidos na tela	0.1014
Desfrute da gastronomia retratada na tela, incluindo alimentos e bebidas	0.1077
Idade	0.1454
Motivação	0.2362
Globoplay	0.2695
Star+	0.2845
Romance	0.2981
Percepção	0.307
Visita a Parques temáticos baseados em atrações inspiradas em séries, filmes, etc	0.3274
Autorrealização	0.3636
Novelas	0.4242
Escolaridade	0.4523
Netflix	0.4536
Clipes de música	0.4644
Escapismo da realidade	0.4644
Companhia	0.4781
Fantasia	0.5535
Viki	0.5576
Disney+	0.6156
Região residência	0.6159
Participação em eventos relacionados a filmes	0.6317
Frequência audiovisuais	0.6619
Encantamento	0.6646
Amazon Prime Vídeo	0.6991
Superação de desafios	0.868
Nostalgia	0.8735
Sexo	0.8796
Visita a cenários internos ou externos onde foram filmadas	0.9107
HBO Max	0.9704
Realizar outra viagem pelo audiovisual	0.9771
Assinatura streaming	1
Seriados	1
Outro audiovisual	1
Telecine	1
Outros streamig	1
Visita a estabelecimentos exibidos na tela, como bares, restaurantes e outros	1
Realização de roteiros guiados em locais de filmagem	1
Outra atividade	1
Outro sentimento	1

### 3.2 - Modelo de regressão logística

A análise de regressão é uma ferramenta estatística que nos permite explorar as relações entre variáveis, identificar fatores significativos e quantificar o impacto desses fatores em uma variável de interesse. Como temos uma variável de interesse com duas categorias, vamos utilizar uma regressão logística.

Essa abordagem permite a compreensão de forma mais intuitiva como as variáveis explicativas estão relacionadas à variável de interesse.

### Estimação do modelo de regressão logística

Foram estimados dois modelos em que todos os coeficientes deram significativos. Utilizamos os Testes de de independência para nortear quais variáveis seriam utilizadas a princípio no modelo.

No modelo 1, como mostrado abaixo, a variável binária grupo é explicada pelas variáveis: - Visita a parques, praças e espaços na natureza exibidos na tela; - Filmes; - Satisfação; - Desfrute da gastronomia retratada na tela, incluindo alimentos e bebidas.

No modelo 2, como mostrado abaixo, a variável binária grupo é explicada pelas variáveis: - Visita a parques, praças e espaços na natureza exibidos na tela; - Filmes; - Desfrute da gastronomia retratada na tela, incluindo alimentos e bebidas.

É preciso saber se o modelo maior (modelo 1) é significativamente superior ao modelo menor (modelo 2), ou seja, se a variável “Satisfação” consegue melhorar a previsão de classificação do grupo. Caso a resposta seja não, optaremos pelo modelo 2 por ser mais parcimonioso.

Para isso, será preciso realizar a análise do que é chamado em Estatística de modelos encaixados, e observaremos a deviance dos modelos.

### Análise de modelos encaixados

**Tabela 3 - ANOVA para os modelos de regressão encaixados**

Resid. Df	Resid. Dev	Df	Deviance	Pr(>Chi)
112	106.0773	NA	NA	NA
111	100.0668	1	6.010499	0.014221

A estatística  $D_q$  (deviance do modelo com menos parâmetros)  $D_p$  (deviance do modelo com mais parâmetros), onde q e p é número de parâmetros dos modelos, respectivamente, é útil como medida de variação dos dados explicada pelos termos que estão no modelo maior e não estão no modelo menor.

Assintoticamente, pode-se escrever:

$$\Delta D = D_q - D_p \sim \chi^2_{p-q}$$

Hipóteses:

$H_0$ : a diferença entre os modelos não é significativa

$H_1$ : a diferença entre os modelos é significativa

Como o valor-p deu 0.014, rejeitamos a hipótese nula e ficamos com o modelo com mais parâmetros. Acredita-se que a variável “Satisfação” contribui para explicar a variabilidade dos dados.

### Análise de deviance

Agora, é preciso conferir se a deviance do modelo escolhido é moderada. Para decidir se uma deviance é pequena / moderada / grande, podemos usar como referência:

O percentil  $\chi^2_{n-p}$ , onde “n” é o tamanho da amostra (116) e “p” o número de parâmetros (4).

Se  $P(D > \chi^2_{n-p;\alpha}) = \alpha$ ; neste caso,  $0 < D < \chi^2_{n-p;\alpha}$  indica deviance pequena ou moderada.

O valor esperado  $\chi^2_{n-p;\alpha}$  que é igual (n - p); neste caso, uma deviance moderada estaria próxima de n - p.

Para o modelo,  $D = 100.07$ . A média da distribuição  $\chi^2_{112}$  é 112. D está próximo da média, logo podemos considerá-la moderada. Também, o percentil da distribuição  $\chi^2_{112}$  com  $\alpha = 0.05$  é 137.70. Como  $0 < D < 137.7015$ , o modelo possui deviance moderada.

**Tabela 4 - Regressores, coeficientes, Valor-p e impacto da covariável sobre a chance (odds)**

	Coeficiente	Valor-p	100*(e <sup>{beta}</sup> -1)
Intercepto	1.158200	0.2005	218.42%
Filmes (sim)	-2.084936	0.0005	-87.57%
Visita a parques (sim), praças e espaços na natureza exibidos na tela (sim)	-2.766893	0.0032	-93.71%
Muito satisfeito (sim)	1.356181	0.0228	288.13%
Desfrute da gastronomia retratada na tela (sim)	1.365575	0.0185	291.80%

Evidenciando o modelo de regressão logística para a chance, a equação fica:

$$\frac{\theta}{1-\theta} = e^{1.16-2.08X_1-2.77X_2+1.36X_3+1.37X_4}$$

A chance é a probabilidade do turista ser específico dividido pela probabilidade dele não ser específico (ser geral ou casual). Se a razão der maior que 1, então quer dizer que há mais chance dele ser específico. Se menor que 1, há mais chance dele não ser específico. Por fim, se for igual a 1, as chances são iguais dele ser específico ou ser casual/geral.

Na coluna 100 x (e<sup>β</sup> - 1), tem-se a seguinte interpretação:

- Mantendo todas as covariáveis fixas, exceto a covariável “filme”, se o turista assistir filme, estima-se uma diminuição de 87,57% da chance quando passamos da categoria “não assiste filmes” para “assiste filmes”.

Da mesma forma, interpreta para as demais covariáveis. Outro exemplo:

- Mantendo todas as covariáveis fixas, exceto a covariável “muito satisfeito”, se o turista ficar muito satisfeito com a viagem, estima-se um aumento de 288,13% da chance quando passamos da categoria “não ficou muito satisfeito” para “ficou muito satisfeito” com a viagem.

Agora, o mesmo modelo para prever a probabilidade do turista ser classificado como turista específico:

$$\theta = \frac{1}{1+e^{-(1.16-2.08X_1-2.77X_2+1.36X_3+1.37X_4)}}$$

Para o modelo, as variáveis estão abaixo:

X<sub>1</sub>: Visita a parques, praças e espaços na natureza exibidos na tela (sim)

X<sub>2</sub>: Filmes (sim)

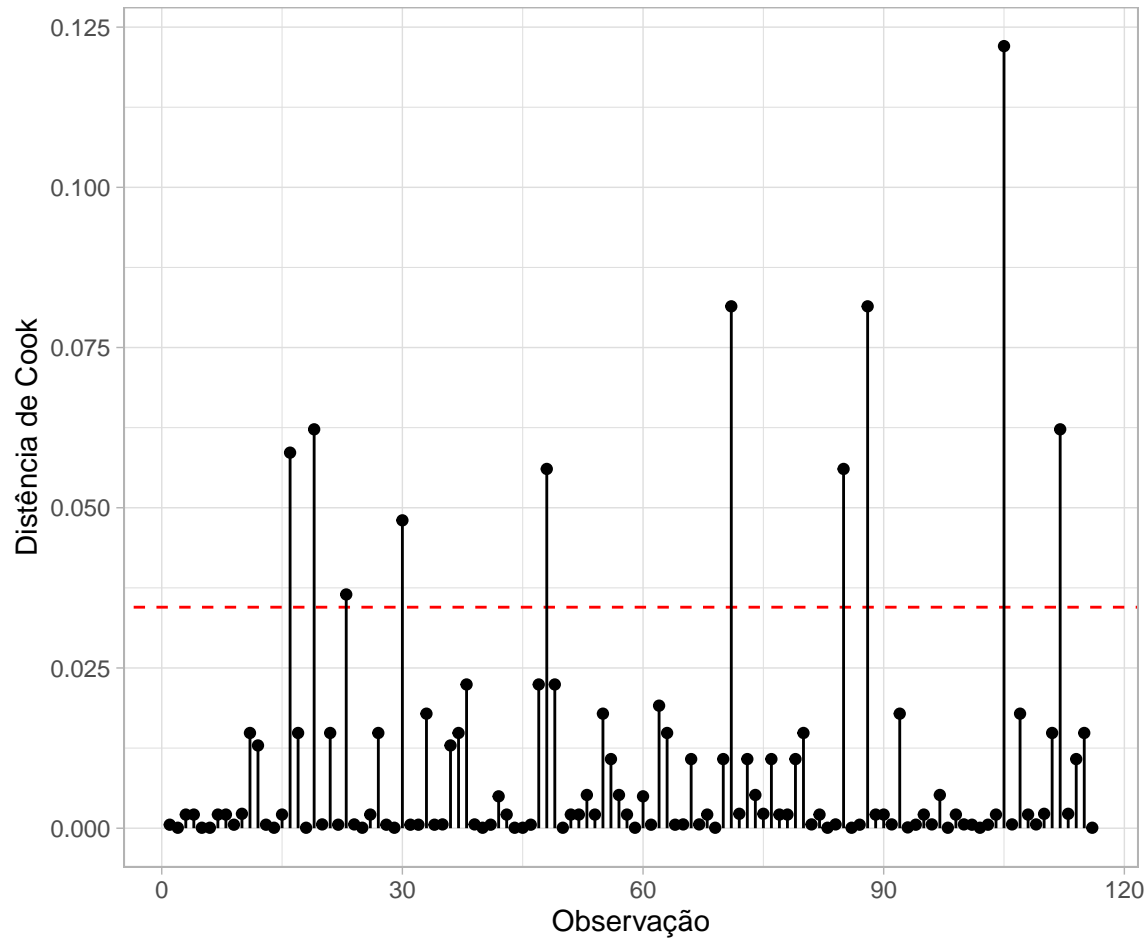
X<sub>3</sub>: Muito satisfeito (sim)

X<sub>4</sub>: Desfrute da gastronomia retratada na tela, incluindo alimentos e bebidas (sim)

#### **Análise de resíduos**



**Gráfico 48 - Distância de Cook**



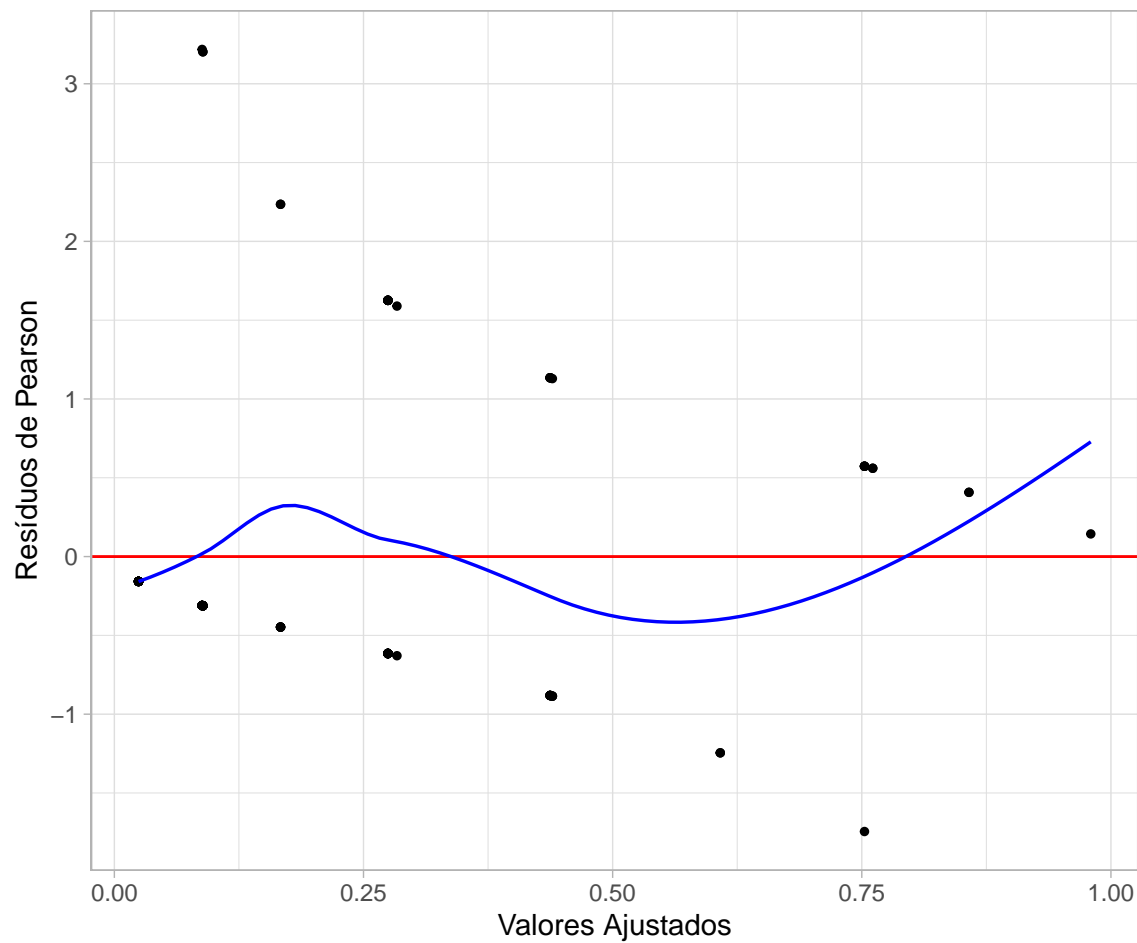
Reajustaremos o modelo sem a observação 106 e será comparado se os coeficientes da regressão alteram muito.

**Tabela 5 - Regressores, coeficientes, Valor-p e impacto da covariável sobre a odds**

	Coeficiente	Valor-p	$100 \cdot (\exp\{\beta\} - 1)$
Intercepto	0.6262247	0.4958	87.05%
Filmes (sim)	-2.2581642	0.0003	-89.55%
Visita a parques (sim), praças e espaços na natureza exibidos na tela (sim)	-2.3150870	0.0132	-90.12%
Muito satisfeito (sim)	1.4400722	0.0178	322.10%
Desfrute da gastronomia retratada na tela (sim)	1.5610513	0.0114	376.38%

Os valores dos coeficientes não são muito diferentes se comparado ao modelo proposto. Logo, não será retirada a observação e será considerado o modelo inicial.

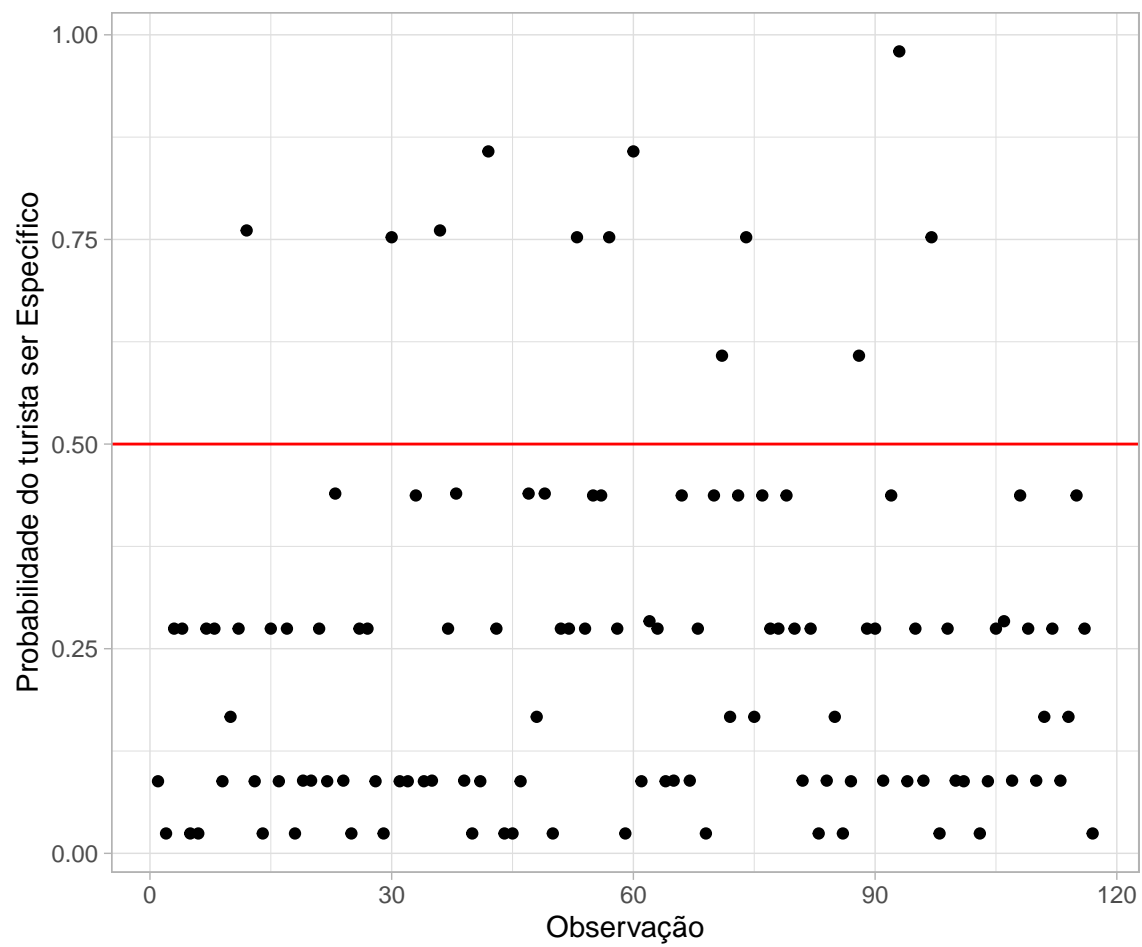
**Gráfico 49: Resíduos contra os valores ajustados do modelo**



A curva Lowess mostra proximidade ao patamar 0. Contudo, ela apresenta formato de parábola, o que indica a necessidade de ajustar o MLG logístico com efeito quadrático de alguma covariável. Contudo, como as variáveis são todas binárias, essa transformação não teria efeito. Assim, o modelo possível obtido é este.

Abaixo podemos ver o gráfico que evidencia a probabilidade de um indivíduo ser classificado como pertencente ao grupo “específico”, conforme o modelo prevê.

Gráfico 50: Probabilidade do turista ser classificado como “Específico” pelo modelo



## 4. Conclusão

A análise descritiva e estatística dos dados revelou, as variáveis mais influentes no contexto da motivação de turistas por produções audiovisuais. Destacaram-se, entre elas, as perguntas relacionadas às *atividades de lazer durante a viagem*, *preferências em produções audiovisuais* e o *nível de satisfação com a experiência audiovisual-induzida*. Esses insights forneceram uma compreensão mais profunda das dinâmicas envolvidas e também direcionaram a construção do modelo final de regressão logística. Ao apresentar esse modelo, buscamos não apenas descrever, mas também prever a propensão de um indivíduo integrar o grupo de turistas motivados pelo audiovisual, fazendo assim para uma abordagem extensiva e preditiva nas pesquisas sobre o tema.