```
# Carregando pacotes necessários
install.packages(c("tidyverse", "lubridate", "ggplot2", "dplyr", "readr", "binom"))
Installing packages into '/cloud/lib/x86_64-pc-linux-gnu-library/4.2'
(as 'lib' is unspecified)
library(tidyverse)
                                                            — tidyverse 2.0.0 —
— Attaching core tidyverse packages ——
√ dplyr
            1.1.2
                      √ readr
                                   2.1.4
✓ forcats 1.0.0
                      ✓ stringr 1.5.0

√ ggplot2 3.4.2

                    ✓ tibble 3.2.1
✓ lubridate 1.9.2
                      √ tidyr
                                  1.3.0

√ purrr 1.0.1
— Conflicts -
                                                      - tidyverse_conflicts() --
X dplyr::filter() masks stats::filter()
★ dplyr::lag() masks stats::lag()
i Use the conflicted package (<a href="http://conflicted.r-lib.org/">http://conflicted.r-lib.org/</a>) to force all conflicts to become err
library(lubridate)
library(ggplot2)
library(dplyr)
library(readr)
library(binom)
# Carregando os dados
data <- read_csv("Artigo GPT Respostas.csv")</pre>
Rows: 103 Columns: 8
— Column specification —
Delimiter: ","
chr (7): Carimbo de data/hora, Qual curso você faz?, Qual nível você acredit...
lgl (1): Pontuação
i Use `spec()` to retrieve the full column specification for this data.
i Specify the column types or set `show_col_types = FALSE` to quiet this message.
# Convertendo a coluna de data e hora para o formato correto
data$`Carimbo de data/hora` <- mdy_hms(data$`Carimbo de data/hora`)</pre>
Warning: All formats failed to parse. No formats found.
# Estatísticas descritivas
summary(data)
                                     Qual curso você faz?
 Carimbo de data/hora Pontuação
 Min. :NA
                      Mode:logical Length:103
                                     Class :character
 1st Qu.:NA
                      NA's:103
 Median :NA
                                     Mode :character
 Mean :NaN
 3rd Qu.:NA
 Max.
       :NA
 NA's :103
 Qual nível você acredita que um chatbot pode ser útil como ferramenta de busca de conhecimento e
Length:103
 Class :character
 Mode :character
 Você atualmente usa mais o chatGPT ou outras ferramentas de pesquisa?
 Length:103
 Class :character
Mode :character
 Você já referenciou o ChatGPT depois de utilizá-lo como fonte de pesquisa em um trabalho acadêmi
 Length:103
 Class :character
Mode :character
 Como você avalia a eficácia do ChatGPT em relação a outras ferramentas ou métodos utilizados em
Length:103
 Class :character
Mode :character
 Eu concordo em participar voluntariamente do presente estudo como participante. Tenho ciência de
Length:103
 Class :character
 Mode :character
# Contagem de respostas por curso
 data %>%
  group_by(`Qual curso você faz?`) %>%
  summarise(count = n()) %>%
  ggplot(aes(x = `Qual curso você faz?`, y = count)) +
  geom_bar(stat = "identity") +
  theme(axis.text.x = element_text(angle = 45, hjust = 1)) +
  labs(title = "Número de respostas por curso", x = "Curso", y = "Número de respostas")
    Número de respostas por curso
  60 -
Número de respostas
                                         Curso
# Contagem de respostas por nível de utilidade percebida
 utilidade_respostas <- data %>%
  group_by(`Qual nível você acredita que um chatbot pode ser útil como ferramenta de busca de col
  summarise(count = n())
 utilidade_respostas %>%
  ggplot(aes(x = `Qual nível você acredita que um chatbot pode ser útil como ferramenta de busca
  geom_bar(stat = "identity") +
   theme(axis.text.x = element_text(angle = 45, hjust = 1)) +
   labs(title = "Percepção de utilidade do chatbot", x = "Nível de utilidade percebida", y = "Númo
     Percepção de utilidade do chatbot
Número de respostas
                                Nível de utilidade percebida
# Calculando o intervalo de confiança para as proporções de cada nível de utilidade
 utilidade_respostas <- utilidade_respostas %>%
  mutate(total = sum(count)) %>%
  rowwise() %>%
  mutate(conf_int = list(binom::binom.confint(count, total, methods = "wilson"))) %>%
  unnest(conf_int) %>%
  select(`Qual nível você acredita que um chatbot pode ser útil como ferramenta de busca de conhe
          count, total, lower, upper)
 print(utilidade_respostas)
# A tibble: 4 \times 5
 Qual nível você acredita que um chatbot pode ser …¹ count total lower upper
  <chr>>
                                                      <int> <int> <dbl> <dbl>
                                                             103 0.00534 0.0681
                                                              103 0.429 0.618
                                                              103 0.0152 0.0956
                                                             103 0.327 0.514
# Contagem de respostas por curso e nível de utilidade
 curso_utilidade <- data %>%
  summarise(count = n())
# Criando tabela de contingência
 contingency_table <- curso_utilidade %>%
  pivot_wider(names_from = `Qual nível você acredita que um chatbot pode ser útil como ferramenta
# Substituindo NA por 0
```

```
1 Indiferente
2 Muito útil
3 Pouco útil
4 Útil
# i abbreviated name:
  ¹`Qual nível você acredita que um chatbot pode ser útil como ferramenta de busca de conhecime
  group_by(`Qual curso você faz?`, `Qual nível você acredita que um chatbot pode ser útil como fo
`summarise()` has grouped output by 'Qual curso você faz?'. You can override
using the `.groups` argument.
```

Teste qui-quadrado de independência chisq_test <- chisq.test(contingency_table[, -1])</pre> Warning in chisq.test(contingency_table[, -1]): Chi-squared approximation may be

Filtrando os dados para incluir apenas respostas 'útil' e 'muito útil'

Adicionando coluna de porcentagem e de etiqueta ao conjunto de dados

contingency_table[is.na(contingency_table)] <- 0</pre>

data_util <- data %>%

Contagem de respostas por curso

group_by(`Qual curso você faz?`) %>%

mutate(percent = count / sum(count) * 100,

curso_count <- data_util %>%

curso_count <- curso_count %>%

summarise(count = n())

```
incorrect
# Verificando o resultado do teste
 chisq_test
    Pearson's Chi-squared test
data: contingency_table[, -1]
X-squared = 142.04, df = 39, p-value = 1.236e-13
# O teste qui-quadrado indica que temos sim uma relação entre o curso e o nível de utilidade atr
```

filter(`Qual nível você acredita que um chatbot pode ser útil como ferramenta de busca de conh

```
label = paste(`Qual curso você faz?`, sprintf("(%0.1f%%)", percent)))
# Criando gráfico de pizza para útil e muito útil
 ggplot(curso_count, aes(x = "", y = count, fill = label)) +
   geom_bar(width = 1, stat = "identity") +
   coord_polar("y", start=0) +
   theme_void() +
   labs(title = "Distribuição de respostas 'útil' e 'muito útil' por curso",
        fill = "Curso")
Distribuição de respostas 'útil' e 'muito útil' por curso
                                                               Curso
                                                                   Arquitetura e Urbanismo (1.0%)
                                                                   Ciência da Computação (76.3%)
                                                                   Ciências Contábeis (1.0%)
                                                                   Direito (2.1%)
                                                                   Enfermagem (1.0%)
                                                                   História (1.0%)
                                                                   Letras (3.1%)
                                                                   Medicina (7.2%)
```

Outro (5.2%)

Outro (16.7%)

Psicologia (16.7%)

Pedagogia (1.0%)

```
Psicologia (1.0%)
# Filtrando os dados para incluir apenas respostas 'pouco útil ' e 'inútil'
data_inutil <- data %>%
 filter(`Qual nível você acredita que um chatbot pode ser útil como ferramenta de busca de conhe
# Contagem de respostas por curso
curso_count_inutil <- data_inutil %>%
  group_by(`Qual curso você faz?`) %>%
  summarise(count = n())
# Adicionando coluna de porcentagem e de etiqueta ao conjunto de dados
curso_count_inutil <- curso_count_inutil %>%
 mutate(percent = count / sum(count) * 100,
         label = paste(`Qual curso você faz?`, sprintf("(%0.1f%%)", percent)))
# Criando gráfico de pizza
ggplot(curso_count_inutil, aes(x = "", y = count, fill = label)) +
```

```
geom_bar(width = 1, stat = "identity") +
   coord_polar("y", start=0) +
   theme_void() +
   labs(title = "Distribuição de respostas 'pouco útil', 'Indiferente' e 'inútil' por curso",
        fill = "Curso")
Distribuição de respostas 'pouco útil', 'Indiferente' e 'inútil' por curso
                                                                Curso
                                                                    Ciência da Computação (16.7%)
                                                                    Ciências Sociais (16.7%)
                                                                    Comunicação Social (16.7%)
                                                                    Desenho Industrial (16.7%)
```