1. **OBJETIVO**

O trabalho tem por objetivo a análise de sentimento em relação ao aborto para os lusófonos e hispanófonos, mais especificamente a mudança deste frente ao evento da descriminalização pela Corte Constitucional da Colômbia do aborto até a 24ª semana de gestação, decisão que ocorreu no dia 21 de fevereiro de 2022. Assim, pretende-se investigar a reação desse coorte frente a uma movimentação jurídica relevante em países culturalmente fronteiriços.

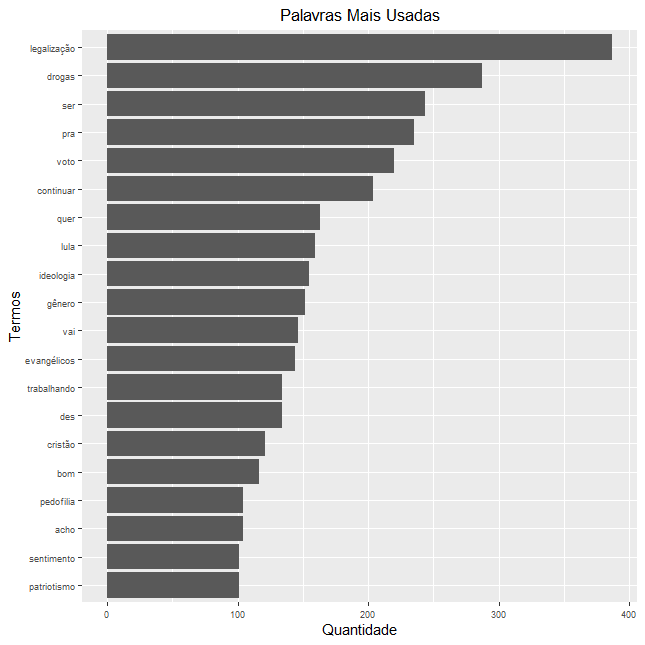
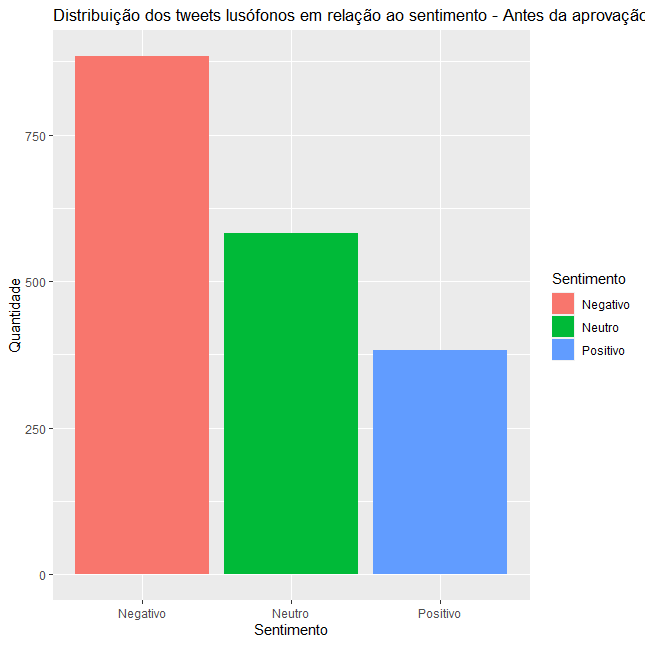
1. **METODOLOGIA**

Para a realização desse estudo foram coletados tweets que continham a palavra-chave “aborto”. Adicionalmente, foram aplicados dois filtros nessa busca: o primeiro se relaciona a língua utilizada pelo texto, selecionando apenas aqueles em espanhol e em português; e o segundo em relação ao período da postagem, selecionando apenas aqueles postados entre o dia 18 de fevereiro, três dias antes da descriminalização, até o dia 25 de fevereiro, data da raspagem dos dados. A base de dados, assim, foi limpa e manejada em quatros blocos: tweets portugueses e espanhóis, nos dois períodos supracitados.

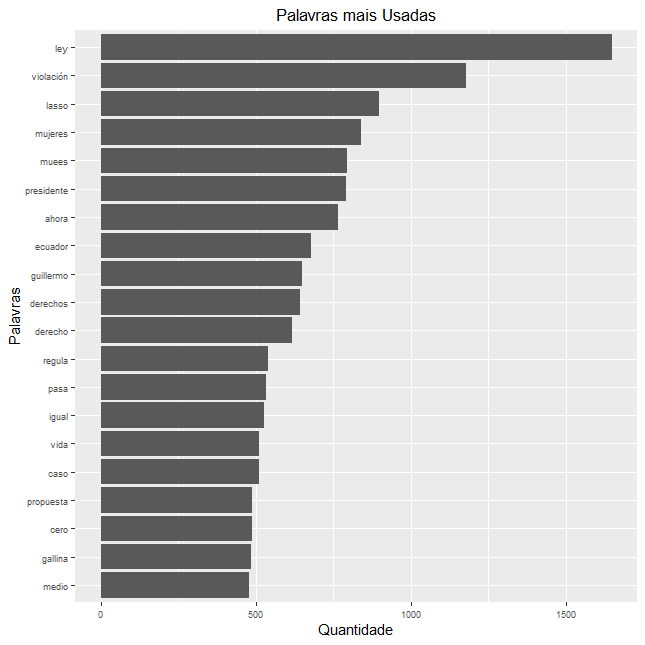
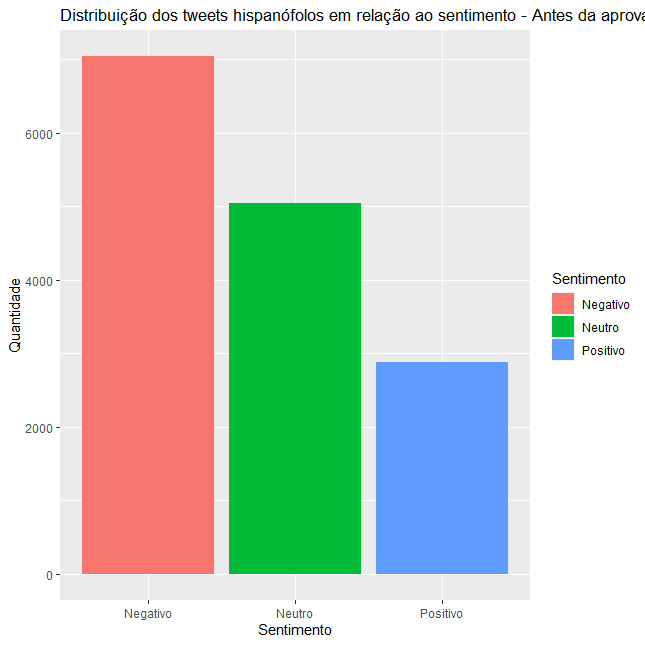
Na raspagem de dados foram coletados 1835 tweets em língua portuguesa antes do evento, 15000 tweets em língua espanhola antes do evento, 15000 tweets em língua portuguesa depois do evento e 15000 tweets em língua espanhola depois do evento.

Posteriormente, foi utilizado o software R para a confecção das visualizações do item 3 desse trabalho. Nele, foram construídos 2 ou 3 gráficos para cada bloco. O primeiro gráfico produzido foi a distribuição dos tweets em três classes de sentimento: positivo, negativo e neutro. Para classificá-los, foram necessários dicionários que informassem o “valor”, positivo ou negativo, de cada léxico. Tais dicionário são de produção externa e foram resgatados no Kaggle, uma comunidade online de cientistas de dados. Depois, foi calculado um score para cada tweet baseado na utilização de termos “positivos” e “negativos” no texto. Ou seja, cada tweet foi valorado de acordo com a conotação das palavras que o compõe. De acordo com o valor do score, o texto pode, assim, ser classificado em positivo, negativo ou neutro. Em seguida, foi construída a Nuvem de Palavras, uma lista hierarquizada visualmente que tem por fito a representação visual das frequências das palavras de uma dada base de dados. Para a sua produção, uma tabela de frequências foi necessária. Por fim, para os blocos de dados que se referem ao período anterior a aprovação, foi elaborado um gráfico de barras com a frequência das palavras mais utilizadas.

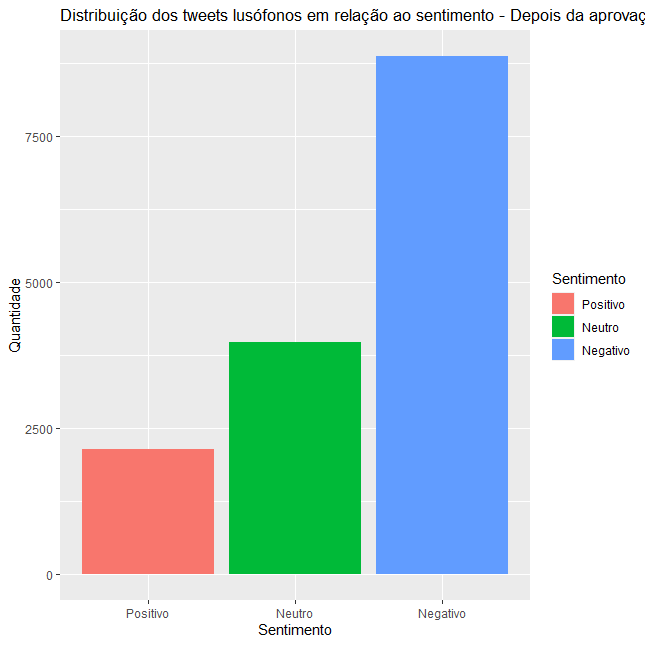
1. **VISUALIZAÇÕES**
   1. Tweets em Língua Portuguesa – Antes da Aprovação



* 1. Tweets em Língua Espanhola – Antes da Aprovação



3.3. Tweets em Língua Portuguesa – Depois da Aprovação





3.4. Tweets em Língua Espanhola – Após a Aprovação

