



GERMAN SACK

JULIANA CAMPOS

LUCIANA LACERDA

MARCO MILANEZ

VÍTOR MENDONÇA

**MONITORAMENTO E ANÁLISE DE *TWEETS* RELACIONADOS À
ANSIEDADE E À DEPRESSÃO**

RIO DE JANEIRO

2020



GERMAN SACK

JULIANA CAMPOS

LUCIANA LACERDA

MARCO MILANEZ

VÍTOR MENDONÇA

**MONITORAMENTO E ANÁLISE DE *TWEETS* RELACIONADOS À
ANSIEDADE E À DEPRESSÃO**

Trabalho de conclusão de curso
apresentado à Faculdade Senac como
requisito para a obtenção de aprovação
no curso de Pós-graduação lato sensu
Especialização em Big Data

Docente orientador: Clayton
Escouper das Chagas

RIO DE JANEIRO

2020

S121 Sack, German.

Monitoramento e análise de tweets relacionados à ansiedade e à depressão
/ German Sack; Juliana Campos; Luciana Lacerda; Marco Milanez; Vítor
Mendonça – Rio de Janeiro, 2020.

53 f. : il. ; 30 cm.

Orientador: Clayton Escouper das Chagas.

Trabalho de conclusão de curso (Pós-graduação lato sensu em Big Data)
– Faculdade de Tecnologia Senac Rio, 2020.

Inclui bibliografia.

1. Big data. 2. Internet das coisas. 3. Recursos eletrônicos de informação.
4. Inovações tecnológicas. I. Sack, German. II. Campos, Juliana. III.
Lacerda, Luciana. IV. Milanez, Marco. V. Mendonça, Vítor. VI. Título.

CDD 005.75



Faculdade de Tecnologia Senac Rio

Curso: Pós-Graduação lato sensu Especialização em Big Data

Ano: 2020

Nome dos alunos:

German Sack

Juliana Campos

Luciana Lacerda

Marco Milanez

Vítor Mendonça

Título: Monitoramento e Análise de *Tweets* Relacionados à Ansiedade e à Depressão

Nome do docente orientador: Clayton Escouper das Chagas

Conceito: "O"

Sem pendências ou recomendações, trabalho de conclusão de curso aprovado. Liberado para os trâmites finais do curso e solicitação do certificado, após verificação de outros requisitos (aprovação nas outras disciplinas/módulos do curso, documentação e financeiro).

Rio de Janeiro, **10** de **OUTUBRO** de **2020**.

A handwritten signature in blue ink, reading "Escouper", with a stylized flourish above it.

CLAYTON ESCOUPER DAS CHAGAS - Orientador

RESUMO

Este estudo teve por objetivo analisar as publicações do Twitter entre o período de 27 de julho a 18 de agosto de 2020. Devido ao impacto da pandemia do novo coronavírus na saúde mental dos indivíduos, foram coletados, através do software R, os *tweets* que continham as palavras “ansiedade” e/ou “depressão”. Em primeiro momento, a partir da análise de sentimentos, a pesquisa se propôs a classificar as publicações quanto à polaridade. A partir dessa técnica, grande parte dos *tweets* foi classificada com polaridade negativa. Já com relação ao comportamento ao longo do dia, constatou-se pico das publicações entre meia noite e duas horas da manhã. Por último, a análise semântica mostrou que os usuários relatam estar ansiosos ou depressivos por motivos não necessariamente relacionados à pandemia, contudo termos relacionados aos sintomas, gatilhos, consequências e tratamento dessas doenças foram mencionados de forma recorrente.

Palavras-chave: análise de sentimentos; *tweets*; ansiedade; depressão.

ABSTRACT

This work aims to analyze the publications on Twitter from July 27 to August 18, 2020. Due to the new coronavirus's pandemic on individuals' mental health, the tweets contained the words "anxiety" and/or "depression". Based on sentiment analysis, the research proposed to classify the publications as to their polarity. Based on this technique, most of the tweets classified were negative polarity. There was a peak in publications between midnight and two in the morning regarding behavior throughout the day. Finally, the semantic analysis showed that users report being anxious or depressed for reasons not necessarily related to the pandemic, however, terms related to the symptoms, triggers, consequences and treatment of these diseases were mentioned repeatedly.

Keywords: sentiment analysis; *tweets*; anxiety; depression

LISTA DE FIGURAS

| | |
|---|----|
| Figura 1 - Diagrama de arquitetura do trabalho | 18 |
| Figura 2 - Etapas do Processo de Mineração de Textos | 19 |
| Figura 3 - Diagrama de correlação entre as palavras | 26 |
| Figura 4 – Total de <i>tweets</i> por polaridade | 29 |
| Figura 5 – Densidade de <i>tweets</i> por hora | 31 |
| Figura 6 – Evolução dos <i>tweets</i> por polaridade | 32 |
| Figura 7 – Nuvem de Comparação por Polaridade | 34 |
| Figura 8 – Mapa semântico do contexto de ansiedade e depressão no Twitter | 46 |

LISTA DE TABELAS

| | |
|---|----|
| Tabela 1 - Frequência de <i>tweets</i> por dia da semana e por polaridade | 23 |
| Tabela 2 - Média de <i>tweets</i> por dia da semana | 24 |
| Tabela 3 – Contagem de termos com frequência maior do que 20 (Top 5) | 25 |

LISTA DE QUADROS

Quadro 1 - Dicionário de variáveis coletadas

52

SUMÁRIO

| | |
|---|----|
| 1 INTRODUÇÃO | 11 |
| 2 REFERENCIAL TEÓRICO | 13 |
| 2.1 SAÚDE MENTAL EM TEMPOS DE PANDEMIA | 13 |
| 2.2 ANÁLISE DE SENTIMENTOS | 14 |
| 3 METODOLOGIA | 16 |
| 3.1 TWITTER | 16 |
| 3.2 COLETA DE DADOS | 16 |
| 3.3 INTEGRAÇÃO, LIMPEZA E SELEÇÃO DA BASE DE DADOS | 18 |
| 3.4 DEFINIÇÃO DE MÉTRICAS | 20 |
| 4 ANÁLISE E DISCUSSÃO DOS RESULTADOS | 22 |
| 4.1. MÉTRICAS | 22 |
| 4.2. CORRELAÇÃO ENTRE PALAVRAS NOS <i>TWEETS</i> | 25 |
| 4.3. ANÁLISE DE SENTIMENTOS | 27 |
| 4.3.1 ANÁLISE SEMÂNTICA | 35 |
| 4.3.1.1 PRINCIPAIS USOS DO TWITTER | 36 |
| 4.3.1.2 DIFICULDADES EM ACEITAR AS DOENÇAS | 37 |
| 4.3.1.3 GATILHOS | 37 |
| 4.3.1.4 SINTOMAS | 39 |
| 4.3.1.5 CONSEQUÊNCIAS | 40 |
| 4.3.1.6 UM OUTRO CONTEXTO DE ANSIEDADE E DEPRESSÃO NO TWITTER | 41 |
| 4.3.1.7 TRATAMENTOS E CURA | 44 |
| 4.3.1.8 MAPA SEMÂNTICO | 45 |
| 5 CONCLUSÃO | 47 |
| 6 REFERÊNCIAS | 49 |
| 7 ANEXOS | 52 |

1 INTRODUÇÃO

Atualmente, muitas pessoas e empresas utilizam as redes sociais para expressarem suas opiniões sobre pessoas, produtos, eventos, sentimentos, diversão, família e, inclusive sobre sua situação de saúde. Nesse último caso, muitas vezes o compartilhamento do conteúdo ocorre antes mesmo do usuário ter procurado um especialista. Com isso, a quantidade de dados textuais que existe na rede de computadores é enorme, e esses dados podem ser usados para aprofundamento de conhecimento sobre determinado tema, assunto ou empresa. A informação textual presente na Internet é muitas vezes um caminho pelo qual se pode ter informação valiosa e sem custos.

Apenas o Twitter é responsável por, em média, cerca de 9.000 *tweets* (termo designado para as publicações no Twitter) por segundo, mais de 550.000 *tweets* por minuto. Isso representa um imenso volume de informações com grande potencial para ser trabalhado (STATS, 2020).

Por esses motivos, vários estudos buscam monitorar o que acontece nas redes sociais, entre eles estão a análise de sentimentos dos conteúdos que são publicados, cujo objetivo é classificar a polaridade dos textos em positivo, neutro ou negativo. É um campo da ciência de dados que vem ganhando bastante importância e se tornando uma das grandes áreas de processamento de linguagem natural (CIÊNCIA E DADOS, 2019).

Text Mining, ou em português Mineração de Textos, é uma extensão da mineração de dados e se refere ao processo de extração de conhecimento de documentos de textos não estruturados. Em 1999, Tan sinalizou em seu estudo o grande potencial desse campo, pelo fato das empresas armazenarem cerca de 80% das suas informações em documentos de texto. A universalização das redes sociais, no período mais recente, ratifica a importância desse campo que é composto, além da análise de texto, por outras técnicas como, *clustering*, classificação, visualização, banco de dados, *machine learning*, etc.

Com o contexto de isolamento social no qual estamos inseridos, temos a hipótese de que esteja ocorrendo no Twitter, um grande número de compartilhamentos de conteúdos relacionados à ansiedade e depressão. Assim, o processo de *Text Mining* pode trazer benefícios por conta da grande quantidade de informação valiosa contida nos textos e que não está disponível nos dados estruturados clássicos.

Nesse contexto de análise em redes sociais, o presente trabalho tem como objetivo analisar os *tweets* relativos à ansiedade e à depressão e identificar, por meio do monitoramento dos termos citados, a polaridade dos *tweets*, bem como sua evolução ao longo dos dias e os termos correlacionados. E dessa forma, identificar padrões de publicações dos usuários quanto ao tema.

Sendo a dimensão da pandemia atual diferente de outras, assim como seus efeitos na população, acreditamos que a gravidade do problema precisa ser analisada, já que transtornos psicológicos como ansiedade e depressão representarão uma segunda onda de problemas na saúde no pós-pandemia.

A divisão do trabalho é feita em 7 seções, sendo esta a primeira seção. A segunda vai apresentar o referencial teórico. A terceira trata de metodologia. A quarta apresentará a análise e discussão dos resultados. Na quinta constarão as conclusões deste trabalho. A sexta trará as referências e a última seção contemplará os anexos.

A base utilizada neste trabalho, assim como o script utilizado para a geração das análises estão disponíveis em: <https://github.com/analise-twitter-ansiedade-depressao>. Essa disponibilização torna possível também, a quem interessar, a reprodução do estudo.

2 REFERENCIAL TEÓRICO

2.1 SAÚDE MENTAL EM TEMPOS DE PANDEMIA

Ansiedade, depressão ou transtornos psiquiátricos, como o Transtorno de Estresse Pós Traumático são respostas do corpo à situação de *stress*, que variam entre os indivíduos (MARGIS *et al.*, 2003),

Em um contexto de crise de saúde global como o vivenciado na crise do coronavírus, além dos impactos biológicos e econômicos, observam-se também impactos relacionados aos níveis de estresse e ansiedade que afetam pessoas, até então, consideradas saudáveis (SHIGEMURA, 2020). Pereira *et al.* (2020) afirmam que o número de pessoas com problemas ligados à saúde mental tende a ser maior do que, de fato, o de pessoas acometidas pela doença. Além disso, esses problemas podem durar muitos anos, como mostram crises do passado.

O isolamento social se mostra relevante diante dessa situação, pois a mudança de rotina e a distância dos amigos e familiares remetem ao sentimento de solidão (SARTÓRIO *et al.*, 2020). A relação entre o isolamento social e o aumento do estresse foi observada em diversos estudos. Isso ocorre pois, de uma maneira geral, em situações como essa, o organismo passa a ficar em estado de exaustão prolongada, o que compromete em diferentes níveis a saúde mental, gerando sintomas de ansiedade, depressão e distúrbios de memória (CACCIPOPO *et al.*, 2011 *apud* SARTÓRIO *et al.*, 2020; WEBER *et al.*, 2019).

Em paralelo ao isolamento social, a alternativa para diminuir a distância física, o uso das mídias sociais, de uma maneira geral, ganha destaque. Por um lado, essa situação parece ser favorável. No entanto, Reed *et al.* (2018) listam várias consequências diretas relacionados ao tempo de uso da Internet, em que observou-se forte correlação com depressão, ansiedade, isolamento social, etc.

Além do isolamento social, outros fatores impactam a saúde mental dos indivíduos, podendo avançar para um quadro de depressão como, por exemplo, a ansiedade crônica, os conflitos conjugais e a instabilidade financeira. Em abril de 2020, somente um mês após o início da pandemia no Brasil, os conflitos conjugais

já haviam tido aumento de 431% (G1 *apud* Fórum Brasileiro de Segurança Pública, 2020). Já relacionado a instabilidade financeira, em setembro, a estimativa de contratos de trabalho suspensos ou redução da jornada de trabalho e salário, ultrapassava 16 milhões de indivíduos (MINISTÉRIO DA ECONOMIA *apud* EXTRA, 2020). Somado a isso, 11 estados do país registraram aumento do desemprego no segundo trimestre comparado ao primeiro trimestre de 2020 (IBGE *apud* UOL, 2020).

No Brasil, a alteração das rotinas das pessoas potencializou o nível de ansiedade na população. Segundo o Laboratório de Neuropsicologia Cognitiva e Esportiva (LaNCE/UERJ), o número de casos de depressão quase dobrou, enquanto os de ansiedade e estresse aumentaram 80% (UERJ, 2020). Outro ponto de alerta é que, muitas vezes, a depressão é erroneamente confundida pela população com tristeza, contudo é uma doença psiquiátrica grave, que afeta cerca de 300 milhões de pessoas em todo mundo e, pelo menos 11,5 milhões segundo a Organização Mundial da Saúde (MINHA VIDA *apud* OMS, 2020).

Outro fator que teve alteração na pandemia foi o comportamento de busca dos usuários no Google, pois um estudo mostrou que as buscas relacionadas à ansiedade tiveram um salto, ficando três vezes maior do que a média dos últimos 16 anos (GOOGLE *apud* G1).

As questões listadas ao longo dessa subseção possuem, direta ou indiretamente, impacto na saúde mental em um cenário de pandemia e isolamento social, pois diante da incerteza, os indivíduos estão naturalmente mais ansiosos, por exemplo. Diversos estudos apontam a relação entre a ocorrência de eventos estressores, como momentos de pandemia e crises financeiras, e o surgimento de sintomas depressivos. Reconhecer essa relação causal permite que estratégias de tratamento possam ser feitas (MARGIS *et al.*, 2003).

2.2 ANÁLISE DE SENTIMENTOS

Segundo Pang e Lee (2008), análise de sentimentos, também conhecida como mineração de opinião, é um ramo da mineração de texto (*Text Mining*) em

que a classificação dos textos é feita a partir do sentimento ou opinião. Em geral, a classificação é utilizada para atribuir sentimentos positivos, negativos ou neutros, que, como afirma Cambero (2016) também são definidos como polaridade.

A análise de sentimentos permite fornecer informações sobre diferentes universos. Contudo, esse é um ramo que está em constante aprimoramento pelo fato de os sentimentos serem algo subjetivo (CAMBERO, 2016).

A recente popularização desse ramo está relacionada ao crescimento da Internet e do conteúdo que é gerado por seus usuários e, principalmente, nas redes sociais, onde as pessoas publicam seus conteúdos em linguagem extremamente coloquial e, em muitas vezes, utilizando artifícios gráficos para tornar os diálogos ainda mais sucintos.

Nesse contexto, o foco deste conceito está atrelado a análise de sentimentos e opinião, expressa pelo público, sobre um determinado assunto, conceito ou serviço, e, no âmbito deste trabalho, classificando-os como positivo, negativo ou neutro.

3 METODOLOGIA

3.1 TWITTER

O Twitter é uma rede social, criada em 2006, com algumas características específicas como, por exemplo, a postagem (denominada *tweet*). Ela deve ter, no máximo, 280 caracteres (Twitter *apud* CORRÊA, 2017; TECMUNDO, 2020).

Apesar de criado em 2006, somente dois anos depois ganhou popularidade no Brasil e a versão em português foi criada somente em 2009. A rede social foi inicialmente pensada como relevante a ser utilizada no trabalho, pois é atualmente o *microblog* mais popular na Internet. O Twitter é bastante usado mundialmente, sendo apresentado em 37 idiomas e tendo aproximadamente 365 milhões de usuários ativos (STATS, 2020). Com o objetivo de manter os usuários ativos, o Twitter se adapta continuamente às necessidades do universo digital. Dentre algumas alterações, estão a inclusão de imagens nos *tweets* e a criação de enquetes (TWITTER *apud* CORRÊA, 2017).

A cada minuto, mais de 550 mil *tweets* são publicados¹ (STATS, 2020), o que reforça que a rede social possui quantidade de informações considerável para serem analisadas no que tange à opinião de seus usuários.

3.2 COLETA DE DADOS

Coletar e interpretar conjuntos massivos de dados só é possível graças ao grande poder computacional que monitora diversas fontes, como sensores e informações sociais, e as analisa usando algoritmos inteligentes (DAVENPORT, 2014).

¹ Crescimento de 36% do que foi observado por Corrêa (2017).

Utilizou-se a API do Twitter² integrada com o Software R (versão 4.0.2)³ para a extração dos dados. O software R contém pacotes e funções direcionados tanto para análises estatísticas, quanto para mineração de textos.

Diante do aumento de doenças relacionadas à ansiedade e à depressão no contexto da pandemia de Covid-19, como apresentado anteriormente, a coleta foi realizada tendo como base os *tweets* que contivessem pelo menos um dos termos a seguir:

- Ansiedade; e/ou
- Depressão

Foram extraídas postagens do Twitter, entre os dias 27 de julho a 18 de agosto de 2020, totalizando, após a exclusão de *tweets* duplicados, 567.382 postagens.

Os *tweets* coletados foram armazenados em formato JSON (JavaScript Object Notation)⁴, por ser um modelo para armazenamento e transmissão de informações no formato texto. Apesar de muito simples, tem sido bastante utilizado por aplicações Web, devido à sua capacidade de estruturar informações de uma forma bem mais compacta do que a conseguida pelo modelo XML, tornando mais rápida a análise dessas informações (DEV MEDIA, 2020).

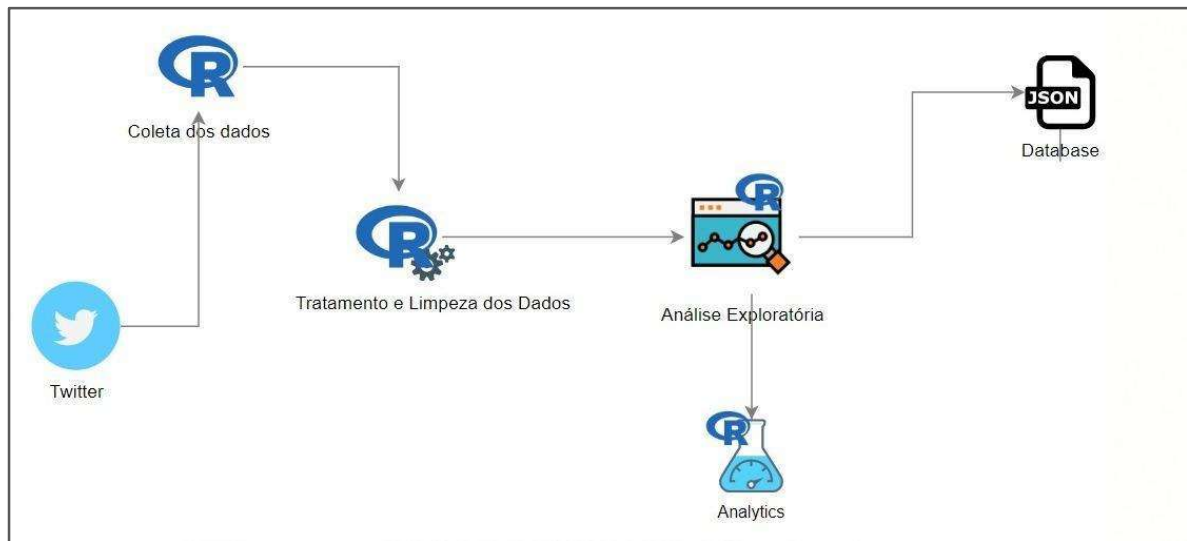
Sendo assim, o diagrama de arquitetura do trabalho está apresentado na Figura 1, sendo um passo fundamental para definição de padrões, de processos e de como as informações foram estruturadas e organizadas. No diagrama estão definidas as práticas de coleta, armazenamento e análise de dados.

² Para saber mais: <https://developer.twitter.com/en/docs>.

³ O software pode ser baixado diretamente na Internet pelo site '<https://www.r-project.org/>'.

⁴ Para saber mais: <http://json.org>

Figura 1 - Diagrama de arquitetura do trabalho



Fonte: Elaboração própria.

3.3 INTEGRAÇÃO, LIMPEZA E SELEÇÃO DA BASE DE DADOS

O processo de Mineração de Textos é iniciado com uma base pronta, e essa é considerada a etapa mais trabalhosa de todo processo. Nesse momento, a base passa por um pré-processamento que vai estruturar o texto não estruturado (ARANHA e PASSOS, 2006).

De uma forma geral, as etapas do processo de Mineração de Textos utilizadas foram: i) seleção dos *tweets*, que foi feita a partir da escolha dos termos monitorados; ii) análise estatística a partir das métricas que serão apresentadas; iii) análise semântica, que envolve avaliar a sequência e contexto dos termos; iv) indexação e normalização que têm o objetivo de facilitar as similaridades entre as palavras; e v) análise dos resultados. A ordem dessas etapas pode ser visualizada na Figura 2.

Figura 2 - Etapas do Processo de Mineração de Textos



Fonte: Adaptado de Aranha e Passos (2006).

A indexação e normalização envolveram a remoção de *stopwords*, que são palavras consideradas como não relevantes na análise de textos. Essa lista é composta por preposições, pronomes, artigos, conjunções e advérbios. Em complemento, essa etapa também contou com a remoção de:

- pontuação;
- menções a usuários;
- numeração;
- links;
- espaços desnecessários;
- emojis;
- caracteres especiais;
- quebra de linhas;
- tweets duplicados;
- tags HTML/Javascript/CSS, dentre outras; e
- além da transformação de letras maiúsculas para minúsculas.

Além disso, *tweets* com textos idênticos também foram removidos da análise e *retweets* não foram considerados na coleta.

3.4 DEFINIÇÃO DE MÉTRICAS

- a) **Contagem de *tweets*:** Considera o total de *tweets* coletados.

$$\sum n$$

Onde:

n = Quantidade de *tweets* coletados.

Percentual de *tweets* por polaridade: identifica a representatividade de cada polaridade sobre o total de *tweets* coletados.

$$\text{Percentual de } \textit{tweets} \text{ por polaridade: } Xi = \frac{\sum n_i}{\sum n}$$

Onde:

i varia de acordo com a polaridade do *tweet*.

n_i = Quantidade de *tweets* com polaridade i .

n = Quantidade de *tweets* coletados na amostra.

- b) **Contagem de *tweets* por polaridade e por dia:** identifica o dia da semana onde as polaridades são mais significativas.

$$\sum n_{i_d}$$

Onde:

i varia de acordo com a polaridade do *tweet*.

d indica o dia da semana

- c) **Média de tweets por dia de semana:** identifica se existe um dia da semana onde a média de *tweets* é mais significativa.

$$X_d = \frac{\sum N_d}{\sum Q_d}$$

Onde:

N_d = Quantidade de *tweets* coletados em um determinado dia da semana

Q_d = Quantidade de dias em que o referido dia aconteceu durante a coleta de dados

- d) **Contagem de palavras:** identifica a quantidade de cada palavra em todos os *tweets* coletados, para identificar as palavras mais significativas. Neste caso, serão consideradas apenas as palavras citadas com uma frequência maior do que 20.

$$\sum P$$

Onde:

P = Palavra

O cálculo das métricas, assim como os códigos utilizados para nestes cálculos serão apresentados, mais adiante, na seção 4.1.

4 ANÁLISE E DISCUSSÃO DOS RESULTADOS

Com as bases de dados unificadas e limpas, foi realizada uma análise das métricas definidas na Subseção 3.4, além de uma análise de *Text Mining*. Esta última, vem sendo amplamente utilizada; é conhecida como uma extensão do *Data Mining* e consiste, segundo Aranha e Passos (2006), na extração de padrões ou tendências de grandes volumes de textos.

Dessa forma, a Mineração de Textos emprega um conjunto de técnicas capazes de estruturar e reconhecer padrões em bases textuais. Nesta etapa, tenta-se descobrir alguma padronização dos dados, aplicando técnicas adequadas de análise para obter bons resultados.

Os *tweets* foram coletados utilizando a API do Twitter a partir da função *search_tweets*⁵. A imagem a seguir exemplifica o código utilizado para um dia de coleta de *posts* do Twitter.

```

95 # Definindo o volume de tweets e palavras que serão coletadas
96
97 tweets_tansiedade <- search_tweets("transtorno de ansiedade" , n = 40000,
98                                   include_rts = FALSE,
99                                   retryonratelimit = TRUE, lang="pt")
100 tweets_ansiedade  <- search_tweets("ansiedade" , n = 40000,
101                                   include_rts = FALSE,
102                                   retryonratelimit = TRUE, lang="pt")

```

O parâmetro `include_rts = FALSE` indica que os *retweets* não foram pegos na coleta e o parâmetro `retryonratelimit = TRUE` se faz necessário somente se a coleta tiver um valor superior a 18.000.

4.1. MÉTRICAS

A partir da base com 567.382 *tweets*, observou-se predominância dos *tweets* de polaridade negativa. A maneira como a classificação foi obtida será detalhada na Subseção 4.3. O total de *tweets* coletados com polaridade negativa equivale a

⁵ Para saber mais: https://www.rdocumentation.org/packages/rtweet/versions/0.7.0/topics/search_tweets.

81%, enquanto os de polaridade positiva e neutra equivalem a 7% e 12%, respectivamente. O percentual de cada polaridade pode ser obtido a partir do código a seguir:

```

167 # tweets por polaridade
168 freq <-stat %>%
169   select(polaridade) %>%
170   group_by(polaridade) %>%
171   count(polaridade)
172
173 # frequência
174 freq
175 tabela = data.frame(freq)
176 tabela
177
178 # percentual por polaridade
179 percent = tabela$n/sum(tabela$n)
180 as.data.frame(percent)

```

Já o total de *tweets* por dia da semana e por polaridade pode ser observada na Tabela 1. O objetivo da métrica foi observar se há algum dia em que a frequência absoluta dos *tweets* é mais relevante.

Tabela 1 - Frequência de *tweets* por dia da semana e por polaridade

| Dia da semana | Negativo | Neutro | Positivo | Total |
|---------------|----------|--------|----------|--------|
| Domingo | 54.260 | 12.974 | 5.133 | 72.367 |
| Segunda | 74.152 | 18.040 | 7.426 | 99.618 |
| Terça | 70.978 | 17.991 | 6.978 | 95.947 |
| Quarta | 59.056 | 14.638 | 5.797 | 79.491 |
| Quinta | 58.630 | 13.597 | 5.500 | 77.727 |
| Sexta | 52.567 | 12.688 | 5.176 | 70.431 |
| Sábado | 54.221 | 12.550 | 5.030 | 71.801 |

Fonte: Elaboração própria a partir dos dados do Twitter.

A coleta de *tweets* foi realizada em quatro segundas-feiras, em quatro terças-feiras e em três dos demais dias da semana. De acordo com a Tabela 1, observa-se que segunda-feira foi o dia com maior frequência de *tweets* de todas as polaridades, seguido de terça-feira, possivelmente, porque ambos foram coletados mais vezes que os demais dias.

No entanto, como a coleta não foi realizada o mesmo número de vezes em cada dia da semana, a análise apenas da frequência de *tweets* por dia da semana pode levar a uma análise equivocada sobre qual dia da semana, de fato, concentrou o maior volume de publicações. Desta forma, para melhor qualificar esta análise, foram calculados os valores médios de *tweets* por dia da semana, conforme aponta a Tabela 2.

Tabela 2 - Média de *Tweets* por dia da semana

| Dia da semana | Média de <i>Tweets</i> |
|----------------------|-------------------------------|
| Domingo | 24.122 |
| Segunda | 24.905 |
| Terça | 23.987 |
| Quarta | 26.497 |
| Quinta | 25.909 |
| Sexta | 23.477 |
| Sábado | 23.934 |

Fonte: Elaboração própria a partir dos dados do Twitter.

A Tabela 2 indica que de uma forma geral, as postagens relacionadas à ansiedade e à depressão por dia variam na ordem entre 23 mil e 24 mil, em média. A exceção ocorre nas quartas-feiras e quintas-feiras, quando as publicações são ainda mais volumosas, pois atingem a ordem de 26 mil e 25 mil, respectivamente. Parte desse volume ligeiramente maior de postagens às quartas e quintas se justifica pelo fato de serem dias com muitos jogos de futebol. Isso acaba por aumentar as postagens de ansiedade, mas não no contexto de doença, mas sim no contexto de expectativa pela competição esportiva.

Conforme indicado na Tabela 3, ao fazer a contagem dos termos com frequência maior que 20 nos tweets extraídos, os dois termos mais recorrentes foram “Ansiedade” (313.762) e “Depressão” (303.362), que foram as palavras-chave nas buscas. Termos como “Não” (136.507) e “Minha” (91.177) são de difícil interpretação fora de um contexto. Completando o ranking das cinco mais citadas aparece a palavra “Sai” (72.845): o termo “sai depressão” é utilizado com muita frequência representando um esforço dos doentes em sair da depressão ou diminuir a ansiedade, conforme será melhor detalhado na Seção 4.3.1.7.

Tabela 3 – Contagem de termos com frequência maior do que 20 (Top 5)

| Termo | Contagem |
|-----------|----------|
| Ansiedade | 313.762 |
| Depressão | 303.362 |
| Não | 136.507 |
| Minha | 91.177 |
| Sai | 72.845 |

Fonte: Elaboração própria a partir dos dados do Twitter.

4.2. CORRELAÇÃO ENTRE PALAVRAS NOS *TWEETS*

Após a sumarização da base de dados, foi aplicado o método de classificação (ou categorização) dos *tweets*, para identificar os principais temas, por meio da contagem de palavras que mais aparecem nos *tweets*. A partir disso, foi possível identificar os tópicos principais, os mais citados e se estão relacionados ao tema do trabalho.

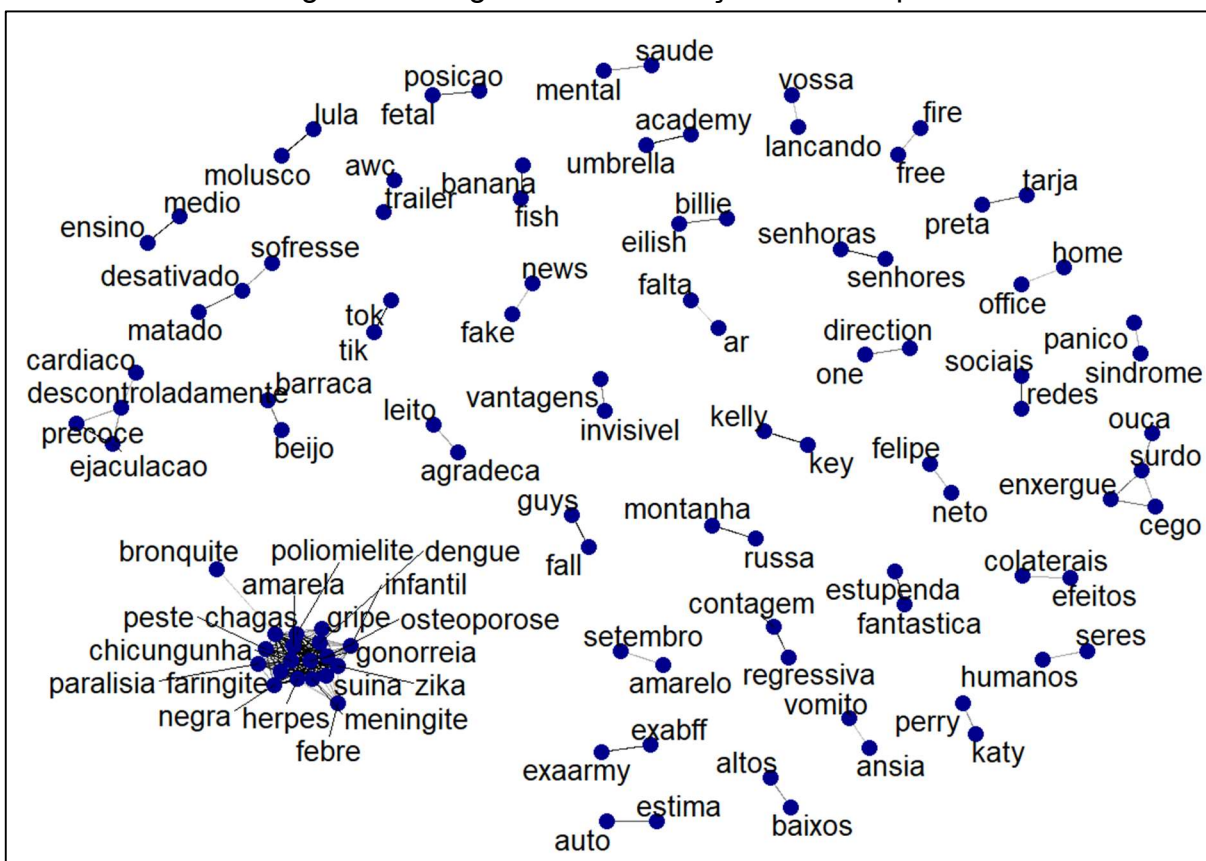
A contagem das palavras é importante pois obtém-se quantas vezes cada palavra apareceu em cada documento, por meio de uma matriz de termos.

A seguir, foi realizada uma análise de correlação, para termos que foram citados mais de 100 vezes nos *tweets*. Foi fixado o nível de correlação de 0,50 para que seja possível identificar o contexto das palavras que mais aparecem juntas nos *tweets*.

A identificação de palavras mais correlacionadas nos *tweets* é uma forma de saber se a análise está seguindo pelo caminho correto, pois mostra as palavras que seriam supostamente esperadas no contexto estudado. Caso surgissem palavras que não tenham nenhum sentido com o trabalho, isso seria um indicativo de que algo estaria incorreto na análise.

A correlação de palavras é uma etapa fundamental para a contextualização de alguns pontos. Este método facilita o entendimento qualitativo das informações, pois identifica, de uma forma geral, as palavras que aparecem juntas nos *tweets*.

Figura 3 - Diagrama de correlação entre as palavras



Fonte: Elaboração própria a partir das informações do Twitter.

A Figura 3 demonstra correlação entre as doenças mencionadas nos *tweets* como, por exemplo, dengue, osteoporose, herpes, hpv, paralisia infantil, etc. No contexto analisado, algumas correlações chamam atenção e fazem sentido, são

elas: i) saúde e mental; ii) setembro e amarelo; iii) tarja e preta; iv) home e office; v) efeitos e colaterais; vi) síndrome e pânico e; vii) auto e estima. Algumas outras correlações se relacionam com eventos ou estreias relacionadas a entretenimento, como umbrella e academy, one e direction, awc e trailer. Por último, termos relacionados a famosos como felipe e neto e blie e elish também aparecem juntos.

Vale lembrar que essa correlação foi feita para os termos que apareceram mais de 100 vezes na base e o motivo desses termos estarem em evidência no período de coleta está detalhado na análise semântica.

4.3. ANÁLISE DE SENTIMENTOS

A análise de sentimentos é o processo de determinar o sentimento ou polaridade de opiniões expressas sobre determinado assunto. Esta análise pode ser aplicada em qualquer forma textual como, por exemplo, blogs, resenhas ou *microblogs*, em que este último são mensagens mais curtas com caracteres limitados, como os *tweets*. Tendo em vista que cada linha da base de dados corresponde a uma publicação, a avaliação de cada *tweet* é feita separadamente (YOUNIS, 2015).

A importância da análise de sentimentos justifica-se, porque as opiniões fazem parte da realidade humana e influenciam os nossos comportamentos. Se pensarmos, sempre que temos uma dúvida ou um problema, procuramos muitas vezes a opinião de outras pessoas, no intuito de tentar chegar a uma conclusão (ou solução) que seja a mais assertiva possível. Esta realidade se tornou ainda mais presente nas redes sociais, e por conta disso a análise de sentimentos ganha cada vez mais importância, e está sendo utilizada para analisar o contexto da ansiedade e da depressão neste trabalho.

Para a classificação dos sentimentos no Software R, em primeiro momento, é preciso ter um dicionário somente com as palavras que representam sentimento positivo e outro dicionário somente com as palavras que representam sentimento negativo. Assim, criou-se manualmente, um dicionário de 9.653 palavras positivas e outro dicionário com 15.958 palavras negativas, pertinentes ao contexto estudado e associadas ao idioma português. São necessários dois pacotes do R para o

desenvolvimento da análise: o *stringr* (para a manipulação de *strings*) e o *plyr* (que permite combinar várias fontes de dados). Cada palavra contida no dicionário de palavras positivas recebe pontuação +1 e as palavras contidas no dicionário negativo recebem a pontuação -1.

A função para a classificação dos *tweets* é utilizada de acordo com o seguinte código:

```

37 # Função de avaliação de tweets
38 score.sentiment <- function(sentences, pos.words, neg.words, .progress='none')
39 {
40   require(plyr)
41   require(stringr)
42   # Obtendo um vetor de frases dos tweets.
43   scores <- laply(sentences, function(sentence, pos.words, neg.words){
44     # Fazendo algumas limpezas ortográficas nos tweets
45     # como colocar tudo em minúsculo e tirar pontuações.
46     sentence <- gsub('[:punct:]', '', sentence)
47     sentence <- gsub('[:cntrl:]', '', sentence)
48     sentence <- gsub '\\d+', '', sentence)
49     sentence <- tolower(sentence)
50     # A função str_split é fornecida pelo pacote stringr que é usado
51     # para dividir as frases em palavras.
52     word.list <- str_split(sentence, '\\s+')
53     words <- unlist(word.list)
54     # A função Match é usada para comparar todas as palavras com o dicionário
55     # positivo e a negativo e retorna VERDADEIRO / FALSO.
56     pos.matches <- match(words, pos.words)
57     neg.matches <- match(words, neg.words)
58     pos.matches <- !is.na(pos.matches)
59     neg.matches <- !is.na(neg.matches)
60     score <- sum(pos.matches) - sum(neg.matches)
61     return(score)
62   }, pos.words, neg.words, .progress=.progress)
63   # Os três parâmetros (dicionário positivo, dicionário negativo e tweets são
64   # passados para essa função
65   # Os resultados serão as pontuações de cada tweet.
66   scores.df <- data.frame(score=scores, text=sentences)
67   return(scores.df)
68 }
```

A função de sentimentos calcula a pontuação de cada *tweet* de forma individual. Primeiro, calcula a pontuação positiva, comparando cada palavra do *tweet* com o dicionário de palavras positivas. Em seguida, calcula a pontuação negativa, comparando cada palavra do *tweet* com o léxico de palavras negativas. Dessa forma, a pontuação final de cada *tweet* é definida por:

$$Score_{total} = Score_{pos} - Score_{neg}$$

Esta classificação gerou um *score* que é interpretado da seguinte forma:

$Score_{total} > 0$; *tweet* com polaridade positiva;

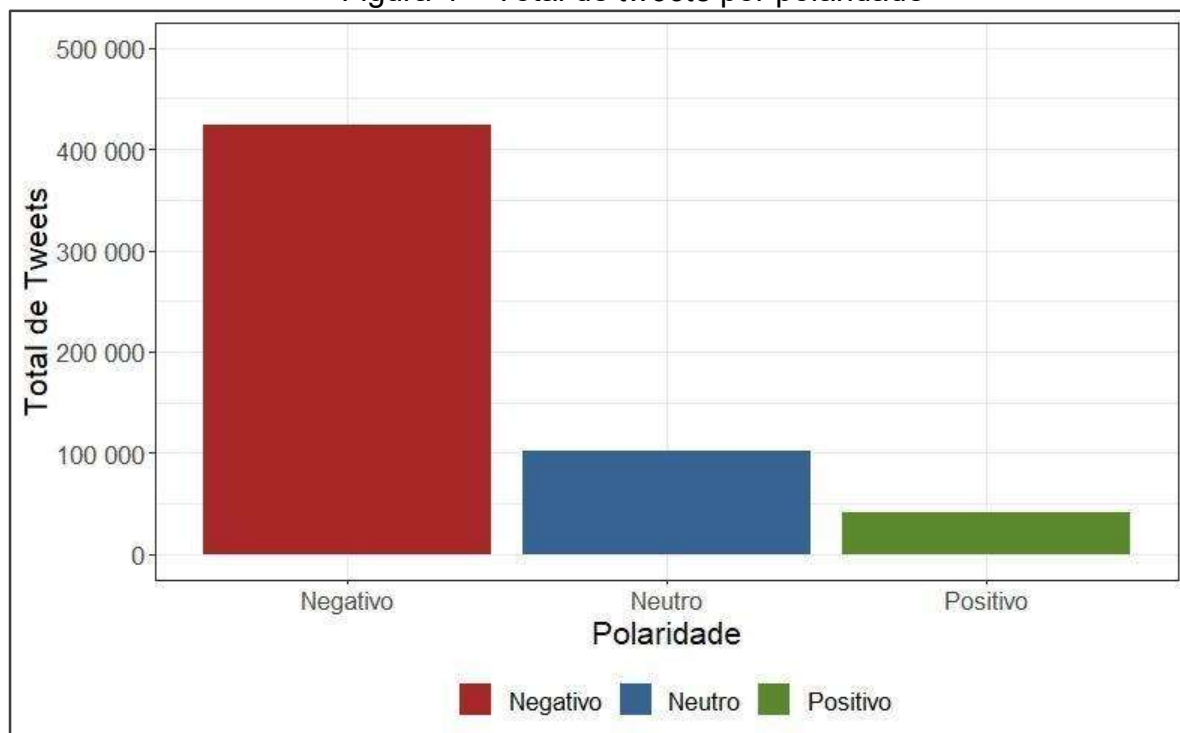
$Score_{total} = 0$; *tweet* com polaridade neutra;

$Score_{total} < 0$; *tweet* com polaridade negativa.

Toda a base foi classificada por meio dos *scores* citados anteriormente. Uma amostra aleatória (correspondente a 3% do total da base classificada) foi extraída com o objetivo de verificar a consistência das classificações. Essa consistência está associada aos dicionários de termos criados. É importante que os termos dos dicionários estejam de acordo com o tema que está sendo pesquisado. Caso contrário, haverá muitos *tweets* classificados de maneira equivocada.

A Figura 4 apresenta a distribuição de *tweets* em cada polaridade obtida por meio do *score* citado anteriormente. Este e todos os demais gráficos foram construídos através do `ggplot()` no Software R.

Figura 4 – Total de *tweets* por polaridade



Fonte: Elaboração própria a partir das informações do Twitter.

Como termos buscados estão relacionados à saúde mental, era esperado uma proporção bem superior de *tweets* classificados como negativos, o que indica que o resultado apontado na Figura 4 foi alinhado com as expectativas.

Com o objetivo de entender se há algum horário específico em que os usuários publicam mais sobre as palavras monitoradas, a Figura 5 demonstra a densidade ao longo das 24 horas do dia.

Nem sempre é possível ver a distribuição dos dados através de histograma com as frequências absolutas em cada intervalo. Uma alternativa quando se trabalha com um grande volume de dados, é utilizar a frequência relativa. Mas à medida que esses intervalos diminuem, fica cada vez mais difícil de visualizar no gráfico. Dessa forma, faz-se a curva de densidade para mostrar essas informações em percentuais, de maneira que a área sob a curva soma 100% (KHAN ACADEMY, 2020).

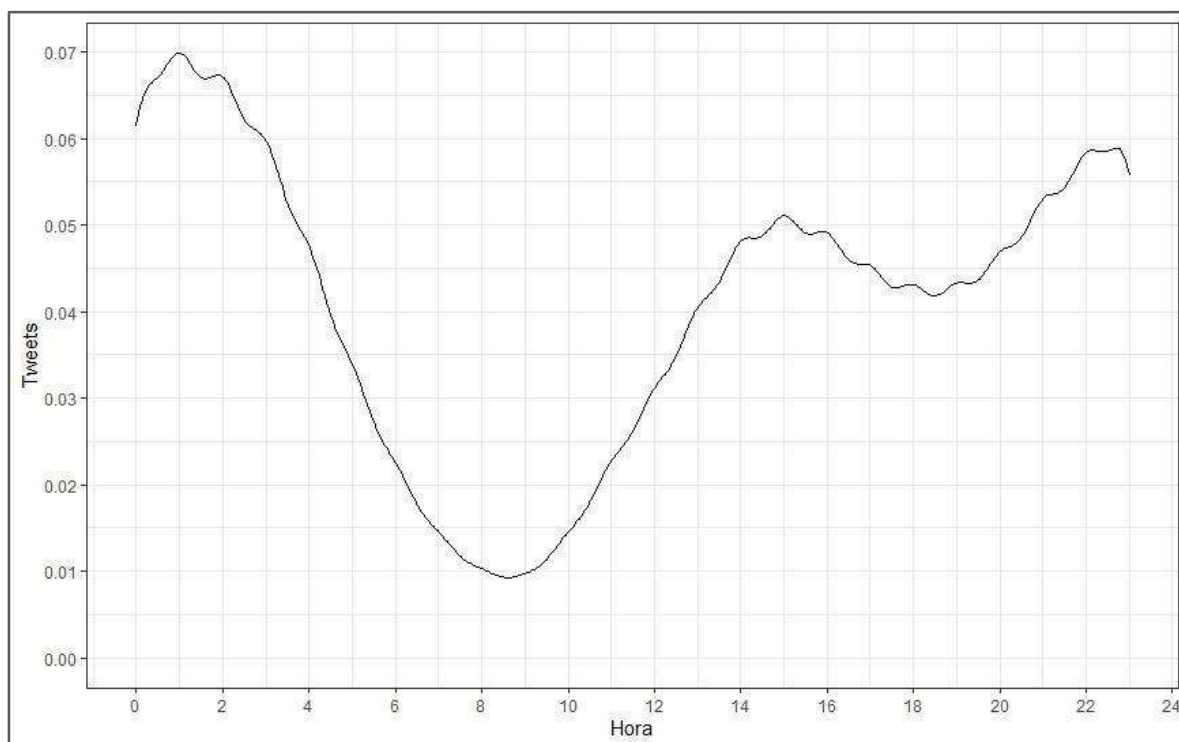
O Density plot (`geom_density`) é utilizado para visualizar a distribuição de uma variável numérica em intervalos contínuos. Ao trabalhar com uma variável contínua, uma alternativa a dividir os dados em intervalos e fazer um histograma, é calcular a densidade de Kernel estimada para a distribuição da variável, conforme o código a seguir.

```

134 library(lubridate)
135
136 tweets_textos3$date <- day(tweets_textos3$created_at)
137 tweets_textos3$hour <- hour(tweets_textos3$created_at)
138
139
140 gg <- ggplot(tweets_textos, aes(x = hour)) +
141   geom_density() +
142   theme_bw() +
143   scale_x_continuous(n.breaks = 13) +
144   scale_y_continuous(n.breaks = 8) +
145   labs(x = "Hora", y = "Tweets")
146
147 gg

```

Utilizou-se o método do núcleo (Kernel) para fazer uma análise de performance do processo de coleta, por meio do gráfico de densidade.

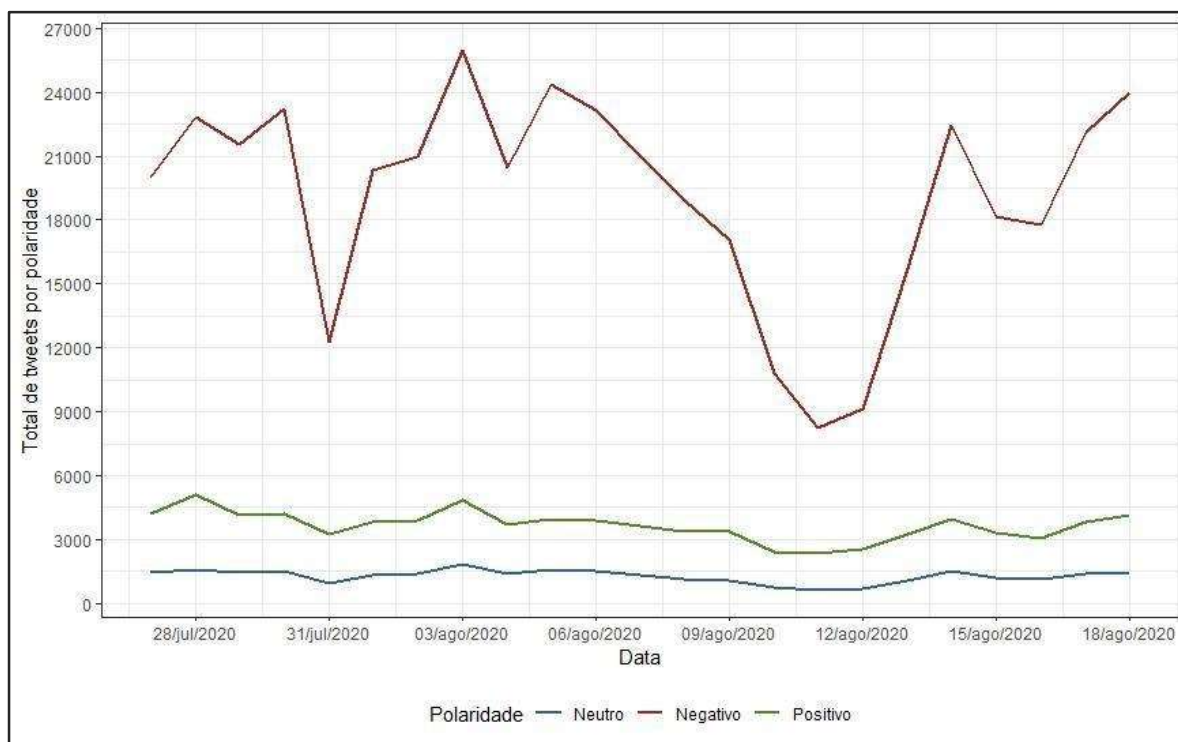
Figura 5 – Densidade de *tweets* por hora

Fonte: Fonte: Elaboração própria a partir das informações do Twitter.

Assim, é possível perceber na Figura 5 que nas três primeiras horas do dia é quando se tem o maior volume de postagens. A partir de então, há uma queda vertiginosa até chegar ao ponto de menor quantidade de *posts*, por volta de 8 horas e 30 minutos da manhã; é perfeitamente justificável por ser horário em que a grande maioria das pessoas está dormindo. Daí em diante, o número de *tweets* aumenta novamente, com leve queda entre 14 horas e 18 horas.

Esse resultado lança luz sobre uma questão que está diretamente relacionada à saúde mental: a insônia (CUIDAÍ, 2020). Apesar da queda de publicações após às 2 horas da manhã, o volume de *tweets* às 22 horas é praticamente o mesmo do que o que foi observado às 3 horas da manhã.

Já a Figura 6 representa a evolução de *tweets* por polaridade, a fim de verificar se houve alguma sazonalidade dos dados no período analisado.

Figura 6 – Evolução dos *tweets* por polaridade

Fonte: Elaboração própria a partir das informações do Twitter.

Analisando a Figura 6, percebe-se que ao longo dos 23 dias de coleta, os *tweets* de polaridade negativa predominaram. Há dois picos de queda quanto aos *posts* com essa polaridade, que representa apenas um volume menor de publicações, em 31 de julho e entre os dias 10 e 12 de agosto. Isso porque não houve redução significativa na proporção dos *tweets* quanto à sua polaridade. Dessa forma, não é possível afirmar que algum acontecimento possa ter motivado essa redução. O dia 5 de agosto foi o dia com maior proporção de *tweets* negativos (81,6%) e o dia 26 de julho foi o dia com menor proporção (72,1%).

Outra forma de análise de termos é a nuvem de palavras, que é uma representação gráfica que permite visualizar os termos mais relevantes (que aparecem mais vezes) em um conjunto de dados. Quanto maior a fonte utilizada, mais relevante é o termo. A Figura 7 mostra quais as palavras foram as mais frequentes, de acordo com as três polaridades identificadas através do algoritmo de classificação.

As palavras que aparecem na nuvem obedecem a uma lógica, onde a diferença entre a maior proporção de um termo dentre as polaridades e sua média

vai definir o tamanho da palavra, e em qual polaridade ela vai aparecer no gráfico. E ainda, quanto maior o tamanho da palavra, e mais ao centro da nuvem, maior será a sua relevância.

A nuvem de comparação vai além da nuvem de palavras, pois ela permite a divisão da nuvem de palavras por alguma outra variável. No R, a função `comparison.cloud` permite gerar a nuvem. Segundo o próprio Software R, a `comparison.cloud` é definida por:

$$(max_j(p_{ij} - p_i))$$

Em que:

$$p_j = \frac{\sum_i p_{ij}}{ndocs} = \text{média entre os documentos};$$

p_{ij} = frequência da palavra i no documento j , onde j varia de 1 a 3.

Tem-se, assim, três documentos, um correspondente a cada polaridade. A mesma palavra se repete nos três documentos, se diferenciando apenas pela frequência em cada um deles. Ou seja, o que determina se uma palavra aparece na polaridade positiva, por exemplo, é o fato dela possuir a maior diferença entre a sua frequência no documento (correspondente a polaridade positiva) e a média total dessa palavra nos três documentos.

Conforme pode ser observado pela linha de código a seguir, a nuvem de palavras é composta pelas 200 palavras mais relevantes de cada documento.

```
123 comparison.cloud(tdm_mx, colors = c("#A92626", "#5B882C", "#366392"), max.words = 200,
124                      scale = c(4,1), random.order = FALSE, title.size = 1.5)
125
```


“terapia”, “curar”, “psicólogo” e “psicóloga” aparecem, em sua maioria, associadas a *tweets* classificados como positivos.

Já as palavras associadas aos *tweets* de polaridade neutras como “lindo”, “linda”, “bom” apesar de caráter positivo, foram mais frequentes em publicações classificadas como neutras. Isso ocorre, pois apesar dessas palavras terem caráter positivo, no texto do *tweet*, elas estão associadas as outras palavras de cunho negativo e que podem configurar ironia, por exemplo.

Apesar de observarmos na Figura 4, que o total de *tweets* positivos é bem menor do que os demais, aqui, observa-se que, em termos de volume de palavras, a polaridade neutra é minoria. Isso se justifica porque as metodologias de análise de polaridade e da nuvem de comparação são diferentes e também porque uma palavra analisada isoladamente tem um contexto, e quando está contextualizada em um *post*, pode ter um contexto totalmente diferente.

4.3.1 ANÁLISE SEMÂNTICA

A cognição alcança todas as formas do conhecimento, como percepção, raciocínio e julgamento (GUIMARÃES, 2007). Os mapas semânticos apresentam, graficamente, conjuntos de representações discursivas feitas por um sujeito (o ator), com vistas a um objeto (o problema), em contextos de interações particulares, segundo Cossette e Audet (1992). Essa representação gráfica é o resultado da interpretação mental que o analista (facilitador) faz a partir da representação discursiva feita pelo sujeito (ator) sobre um problema.

Os mapas semânticos podem ser usados como produtos (mantendo estáveis) ou como ferramentas (com caráter dinâmico e passível de modificação) e podem ser:

- de identidade (designam as chaves físicas do problema: atores, eventos e processos) ou
- de categorização (criam escalas e convenções que oferecem informações sobre o relacionamento entre as entidades do problema),
e

- causais ou de argumentação (com vias alternativas para passar de uma posição a outra no mapa - as ligações potenciais entre a entidade de importância para a organização) (REFATTI *et al*, 2013).

Desta forma, as próximas seções - 4.3.1.1 a 4.3.1.7 - trarão uma representação discursiva para contextualizar os tweets relacionados à ansiedade e à depressão. Na seção 4.3.1.8, mais adiante, será apresentado um mapa semântico, que traz uma representação visual do contexto de ansiedade e depressão, a partir da análise discursiva.

4.3.1.1 PRINCIPAIS USOS DO TWITTER

Por meio da nuvem de palavras dos termos apontados nos *tweets* coletados, pela análise de correlação entre as palavras, e também pela análise qualitativa de uma amostra dos *tweets* coletados foi possível realizar uma análise semântica. Esta análise apontou que sobre ansiedade e depressão, o Twitter é utilizado com muita frequência para que seus usuários:

- Descrevam seus sofrimentos;
- Descrevam a falta de entendimento da doença entre os membros da família e amigos;
- Descrevam diversos sintomas da doença;
- Busquem outras pessoas na mesma situação;
- Busquem outras pessoas que se envolvam com o sofrimento;
- Façam recomendações e orientações a outras pessoas;
- Falem sobre a efetividade de seus tratamentos (cura).

Percebe-se que, por diversas vezes, os usuários buscam o Google como forma de entender a doença e seus sintomas. No entanto, esta ferramenta de busca e alguns testes que podem ser realizados pela Internet apavoram os doentes, pois apontam uma infinidade de sintomas durante as pesquisas e uma infinidade de transtornos como resultados dos testes da Internet.

4.3.1.2 DIFICULDADES EM ACEITAR AS DOENÇAS

Os *tweets* apontam que os usuários passam por muitas dificuldades em aceitar as doenças como ansiedade e depressão, principalmente, por conta da intensidade dos sintomas e da imprevisibilidade de sua ocorrência.

A fase de negação se prolonga por bastante tempo até que os doentes procurem ajuda de profissionais e aceitem os diagnósticos. Percebe-se uma dificuldade em se abordar o tema, principalmente, com amigos e familiares, pois traz uma sensação de vergonha ao doente e de solidão, pois a percepção é de estar doente e não poder pedir ajuda. Até mesmo porque muitos familiares acabam julgando como sendo frescura do doente. Observam-se relatos de que familiares afirmam que a cura para ansiedade é “lavar louça”, em tom sarcástico.

Uma citação bastante citada foi para ilustrar o sentido da hipocrisia e machismo no que se refere à ansiedade perante a sociedade foi:

“Tem ansiedade ??? Acalme

Tem depressão??? Sorria

É surdo?? Ouça

E cego ??? Enxergue

Sofre machismo?? Vire homem.”

Assim, alguns sentimentos apontam a forma como os doentes se sentem, tais como: fracassado, assustado, obcecado, sufocado, patético, paranoico, compulsivo, impotente, preguiçoso, angustiado, tenso, agoniado, melancólico, sozinho, louco, boiolinha, depressivo, inútil, ...

4.3.1.3 GATILHOS

Foram relatados uma infinidade de gatilhos para ansiedade e depressão, que estão atrelados a terminar um curso, relacionamentos mal sucedidos, corpo considerado como fora dos padrões “ideais”, consequência de estupro ou enfrentamento com figuras masculinas (violência ou abuso), dificuldade em enfrentar locais cheios, mudanças físicas nos últimos anos (“febre” nas redes

sociais nos últimos meses), obesidade ou magreza, preconceitos, autismo, machismo, homofobia, racismo, incompatibilidade religiosa, além de bullying. Neste contexto, a forma como as outras pessoas pensam tem grande interferência e relevância aos doentes: “pessoas acham que depressão é frescura, bullying é brincadeira e suicídio é para chamar atenção”.

A ansiedade advinda de provas para aprovação em Mestrados e Doutorados, além de entregas finais de dissertações e teses foram citadas como gatilhos para ansiedade também, já que existe uma cobrança de conteúdos e de prazos.

E ainda, dentro do atual contexto que estamos vivendo, alguns gatilhos foram citados, como o desemprego, o presidente Bolsonaro, a espera de resultados de Covid-19, a falta de convívio social (por conta de home office e por aulas em EAD, durante a pandemia), o uso de máscaras, a quarentena, o isolamento, as mudanças de pensamento e do modo de vida, em geral, por conta da pandemia. E ainda, pessoas que estão com depressão agradecem a Deus por não estarem em um leito de hospital, neste momento conturbado da saúde.

Especificamente sobre o *home office*, que é uma realidade em boa parte das empresas no momento, houve relatos de que é possível trabalhar, mesmo chorando de ansiedade na cama. A ausência dos colegas de trabalho, a percepção de pouca produtividade, por conta da divisão com tarefas domésticas, a percepção de que não se pára de trabalhar, o barulho, as constantes mensagens de WhatsApp, além da verificação constante se o microfone encontra-se no mudo pioram a ansiedade.

Em outro aspecto do *home office*, observou-se redução da ansiedade por não ser necessário o deslocamento no trânsito para o trabalho. E ainda, houve relatos de ansiedade quando as pessoas precisam ir presencialmente, ao trabalho, neste momento de pandemia (exposição ao risco).

Os *tweets* apontam ainda uma realidade bem difícil que se refere ao volume de pessoas muito jovens que sofrem de ansiedade. Os relatos apontam que jovens durante o ensino médio passam por uma série de mudanças para a adolescência, e que acabam por desencadear ansiedade por conta de gravidez precoce, bullying na escola, excesso de cobrança para realização de provas bem sucedidas para entrar em faculdade (Vestibular), etc.

Ansiedade e depressão também foram relacionadas a uma série de doenças, tais como: gripe, chikungunya, poliomelite, febre amarela, zika, doença de chagas, faringite, meningite, gonorréia, osteoporose, dengue, paralisia infantil, leptospirose, herpes, acne, bronquite, alergia, rinite, asma, gastrite, infarto, câncer, esquizofrenia, autismo, etc.

Um outro ponto interessante é o impacto das redes sociais como gatilho para a ansiedade - fobia social atrelada às redes sociais. Pessoas relatam a desativação de redes, como o Facebook, por conta de fotos e comentários que trazem depressão, tratando-se de um gatilho para a ansiedade. Outro exemplo, é o filtro de rosto invertido no Tik Tok (como as pessoas te vêem): pessoas achando terrível a forma como são vistas por outras e justificando tendências depressivas por isso. Por outro lado, usuários gostam e acham uma distração interessante para sair da depressão.

Foram observados diversos relatos de um grupo aberto do Facebook chamado Deep Web da Depressão. Segundo os usuários, este grupo faz conteúdos pesados, comentários debochados sobre depressão, além de comentários machistas, pornográficos, etc. Muitos usuários do Twitter utilizam a rede social para denunciar este grupo do Facebook.

4.3.1.4 SINTOMAS

Os sintomas relacionados à ansiedade e à depressão são extremamente citados nos *tweets*. Ou seja, a descrição do sofrimento nos *tweets* foi apurada de forma muito relevante.

Entre os sintomas da ansiedade e da depressão citados, destacam-se: dificuldade de respirar, medo de morrer, aceleração do coração, insegurança, medo, desespero, sentimento de insuficiência, entrar em pânico, sem um motivo aparente, transtornos de personalidade - bipolaridade, dificuldade em estar sozinho, humor depressivo, formigamento no corpo, síndrome de *burn out*, complexo de inferioridade, complexo de imagem, perturbações, dor, solidão, stress, choro, tristeza, culpa, sofrimento, dificuldade em levantar da cama, náuseas, alucinações, esgotamento, palpitações, calafrios, fadiga, exaustão, fraqueza, taquicardia,

tremedeira, enxaqueca, nervosismo, surtos, compulsão alimentar, além de deitar ou chorar em posição fetal e percepção de estar em uma montanha russa (sobre os altos e baixos da depressão). Houve ainda afirmações de que tomar Coca Cola piora a ansiedade, por conta da cafeína.

4.3.1.5 CONSEQUÊNCIAS

Entre as consequências da ansiedade e depressão citadas nos *tweets*, destacam-se morte, suicídio, mudança de comportamento, baixa autoestima, fumar muito, ejaculação precoce, automutilação, nostalgia, abstinência, espinhas, insônia, preocupação.

Desde 2014, a Associação Brasileira de Psiquiatria – ABP, em parceria com o Conselho Federal de Medicina – CFM, organiza nacionalmente o Setembro Amarelo. O dia 10 deste mês é, oficialmente, o Dia Mundial de Prevenção ao Suicídio, mas a campanha acontece durante todo o ano.

Segundo o site da campanha Setembro Amarelo, são registrados cerca de 12 mil suicídios todos os anos no Brasil e mais de 1 milhão no mundo. Trata-se de uma triste realidade, que registra cada vez mais casos, principalmente entre os jovens. Cerca de 96,8% dos casos de suicídio estavam relacionados aos transtornos mentais. Em primeiro lugar está a depressão, seguida do transtorno bipolar e abuso de substâncias (SETEMBRO AMARELO, 2020).

No entanto, chamou atenção que boa parte dos *tweets* que abordavam a campanha falam sobre a hipocrisia das pessoas no que se refere aos problemas com ansiedade e depressão: famílias que acham que o problema é frescura, pessoas que ao longo do ano apresentam comportamentos questionáveis e, que no mês de setembro discursam sobre a campanha Setembro Amarelo.

4.3.1.6 UM OUTRO CONTEXTO DE ANSIEDADE E DEPRESSÃO NO TWITTER

Observou-se um outro contexto de ansiedade no Twitter, pois seus usuários relatam sentirem ansiedade ao jogar um jogo eletrônico, ou na espera de um jogo esportivo, ou do lançamento de uma nova temporada de série, ou de um filme, ou de uma música ou trailer. Neste contexto, observa-se que a ansiedade é citada como forma de espera para jogar, ou para assistir um filme, série ou trailer. E ainda, há relatos também que estes entretenimentos fazem esquecer um pouco a ansiedade. Desta forma, gera-se uma contagem regressiva para alguma atividade, seja um filme, uma série, ou até mesmo alguma expectativa de viagem, etc

Além disso, observa-se também que alguns usuários do Twitter identificam o seu momento de sofrimento de ansiedade com algumas músicas, filmes e séries.

Os jogos *Free Fire* e *Fall Guys*, por exemplo, são muito citados neste contexto. Os usuários do Twitter relatam ansiedade para jogar, ou uma forma de afastar a depressão e, até mesmo, apontam depressão pós-jogo.

O jogo *Last of Us* também foi bastante destacado, com citações sobre sentimento de ansiedade para finalizar o jogo e depressão por assistir certas cenas, que retratam um contexto de sofrimento. O sucesso do *game* será adaptado para uma série de TV.

O *Selca Day* foi citado neste contexto também, visto que é uma ação, dentre os *fandoms* de *kpop*. Os fãs postam fotos de pares ou combinando com seus artistas favoritos. Cada fã clube tem seu próprio dia. Alguns são definidos pelos dias da semana, outros pelos dias do mês. O objetivo dessa ação é fazer com que os fãs conheçam seus “irmãos de *fandoms*”, estimule o amor próprio e espalhe o amor pelo artista. No entanto, a ação é encarada em alguns momentos como positiva e outros como negativa. Positiva, por se tratar de um momento engraçado e que tira as pessoas da depressão, além de deixar pessoas ansiosas para ver *selcas* de outras pessoas. E negativa, porque as pessoas julgam as *selcas* pela aparência e pelas minorias de representatividade.

Algumas séries retratam este contexto de ansiedade para serem finalizadas pelos usuários do Twitter, tais como *Good Girls*, *She Ra*, *One Piece*, *Akame ga Kill!*,

Vis a Vis, *Umbrella Academy*, *Banana Fish*. Há citações ainda de ansiedade pelo lançamento dos trailers da Liga da Justiça e do trailer do filme *After We Collided* (AWC) 2, além do Filme Barraca do Beijo 2.

Os personagens *Edward* e *Bella* de “O Crepúsculo” foram bastante citados. O filme Lua Nova mostra *Bella* em depressão, depois de se afastar de *Edward*. Durante o período de coleta foi anunciado o release de *Midnight Sun*, mais um livro sequencial. Pessoas relatam pontos depressivos dos personagens e dos filmes e livros, e que, possivelmente, terão depressão quando finalizarem a leitura do livro.

O *Big Brother* 2020, também denominado BB 2020, foi a quinta edição do *reality show Big Brother* em Portugal, transmitido pela TVI. Foi iniciado em 26 de abril de 2020 e terminou em 2 de agosto de 2020. Pessoas relatam ansiedade pelos episódios e pela final do programa foram bastante relevantes. Relatos semelhantes foram observados para o *Masterchef*, programa de disputa culinária na Band e para a novela Totalmente Demais, da TV Globo.

Este mesmo contexto de ansiedade pelos episódios é percebido em *Doramas*, que são histórias com começo, meio e fim, interpretadas por atores, contadas em formato episódico e exibidos semanalmente na televisão japonesa. Assim como séries, também é possível encontrar *Doramas* de diversos gêneros e para diversos públicos. Por serem exibidos em episódios semanais, os *Doramas* costumam ter temporadas curtas e histórias mais estruturadas, com uma quantidade mais enxuta de personagens.

O programa Fantástico, da TV Globo, vem sendo apontado como um programa com conteúdo depressivo e os usuários do Twitter relatam sentirem uma depressão pós-Fantástico.

Durante o período de coleta de *tweets*, um ilustrador fez a internet chorar e ampliar postagens sobre a depressão com o adeus de Lula Molusco a Bob Esponja - filme de pouco mais de um minuto, que criou uma história em que o Lula Molusco escreve uma carta póstuma ao Bob Esponja.

Sobre o *Youtuber* Felipe Neto, houve relatos de pessoas que assistem seus vídeos para saírem da depressão. E ainda, o *Youtuber* assumiu que enfrenta problemas com depressão. Foi entrevistado no Programa Roda Viva próximo ao

período de coleta dos tweets e sua frase “A gente precisa tirar a depressão do armário” foi muito bem vista entre os doentes. Muitos se identificaram.

Na música, este contexto também aparece, pois os usuários de Twitter relatam ansiedade por lançamentos de música, por premiações de artistas e novos integrantes de bandas.

Um exemplo é a banda BTS. Os usuários do Twitter estavam ansiosos pelo lançamento da música *Dynamite*, que ocorreu após o período da coleta. No entanto, durante a coleta, ocorreu um *teaser*, que foi o lançamento das primeiras fotos, que gerou ansiedade entre os fãs, inclusive utilizando *hashtags* associadas à banda (#Exabff e #ExaArmy). Além disso, a banda também estava concorrendo ao MTV’s *Hottest Summer SuperStar 2020*, gerando ansiedade entre os fãs para conhecerem os vencedores.

Fãs da *Beyoncé* também estavam ansiosos pelo lançamento do filme *Black is King*, com a cantora, que ocorreu durante o período da coleta. A *hashtag* associada ao filme (#BeyonceisComing) foi bastante utilizada.

Fãs da banda *Treasure* também estavam ansiosos com a estreia do novo integrante da banda, que ocorreu durante o período da coleta (*Debut do Treasure*). Além disso, para a *live* do grupo *One Direction*, os fãs relatam ansiedade pelo início da *live* e também depressão *pós-live*, por ter terminado. Ansiedade relatada também pelo lançamento da música *Cherry Flavoured*, da banda *Neighbourhood*: considerada também uma oportunidade para sair da depressão.

No caso de *Katty Perry*, os relatos de ansiedade estavam atrelados ao adiamento do novo álbum, além de momentos de depressão que ela viveu. Ela revelou lutar contra depressão durante o período de coleta.

Anitta estava concorrendo a premiação Meus Prêmios *Nick* (MPN) gerando ansiedade entre seus fãs no Twitter.

Ainda no contexto musical, algumas músicas são citadas por usuários em momentos de depressão e ansiedade, tais como músicas de *Lana Del Rey* e música Anjo (de Kelly Key). Neste último caso, ocorrem citações à letra da música “Mais uma noite chega, E com ela a depressão”. E ainda, com algumas adaptações: “Mais

uma segunda chega, E com ela a depressão"... "Mais um semestre chega, E com ele a depressão"...

Houve relatos também de ansiedade para a criação de *Lomotif*, que é um editor que permite criar pequenos filmes usando fotos e vídeos do celular. O *app* possui uma ferramenta de edição de fácil manuseio, em que é possível adicionar músicas famosas como trilha sonora e compartilhar o resultado com os amigos. No entanto, percebe-se um outro contexto: pessoas tão bonitas editando vídeos bonitos, que acabam gerando depressão em outras.

No caso do futebol, os relatos dos usuários do Twitter apontam que os jogos do Flamengo têm dado depressão, por conta da atuação do time, sem o ex-técnico Jorge Jesus. Para o Corinthians, também há relatos de depressão quando o time perde. No caso do Palmeiras, *Bayern* e Barcelona, as citações são relacionadas à ansiedade pelo início dos jogos. Já citações relacionadas ao Neymar referem-se à ansiedade para que ele faça gols.

Observa-se ainda citações de ansiedade ao Cartola FC, que é um jogo eletrônico de futebol fictício, no qual as pessoas montam seus times com jogadores de futebol da vida real. Foi criado e é mantido pelo site Globo.com e promovido pelo canal de TV por assinatura SporTV e também pela Globo.

O momento de pandemia apontou um aumento nas compras online e, com isso, houve relatos sobre ansiedade na espera por este tipo de compras durante a greve dos correios, e até mesmo pela impossibilidade de comprar online, já que a greve dificultaria a entrega dos produtos comprados.

4.3.1.7 TRATAMENTOS E CURA

O termo "sai depressão" é utilizado com muita frequência representando um esforço dos doentes em sair da depressão ou diminuir a ansiedade. É utilizado quando se busca alguma atividade para entretenimento, seja assistir a um filme, a uma série, ouvir uma música, jogar um jogo eletrônico, assistir a alguma competição esportiva, etc.

De acordo com os *tweets*, o tratamento para ansiedade é baseado em terapia (com psicólogos e psiquiatras), medicamentos e meditação. É considerado um tratamento longo. E, por conta disso, muitos doentes abandonam e apresentam retorno dos sintomas. O chá de camomila é um calmante natural que é bastante utilizado quando se percebe um aumento na ansiedade. Ansiolíticos, antidepressivos, anfetaminas, anti-histamínicos e calmantes foram citados como formas de tratamento envolvendo alguma sedação, relaxamento muscular, efeito tranquilizante e regulador de sono, apetite e ritmo cardíaco, tais como ritalina, dramin, clonazepam (rivotril), sertralina, serotonina.

Os usuários que relatam sobre a cura da ansiedade, relatam uma percepção de melhor bem-estar, de maior coragem para enfrentar os problemas e de maior controle de sono, apetite e ritmo cardíaco.

4.3.1.8 MAPA SEMÂNTICO

O mapa semântico traz uma representação visual do contexto de ansiedade e depressão no Twitter, a partir:

- da análise discursiva para contextualizar os tweets relacionados à ansiedade e à depressão (realizada entre as Seções 4.3.1.1 a 4.3.1.7);
- da computação cognitiva, por meio da correlação entre palavras, nuvem de palavras, etc;
- da análise qualitativa das informações.

Desta forma, está apresentado na Figura 8, o mapa semântico que contextualiza as análises qualitativas do trabalho.

Figura 8 – Mapa semântico do contexto de ansiedade e depressão no Twitter



Fonte: Elaboração própria, a partir da análise qualitativa das informações do Twitter.

5 CONCLUSÃO

O presente estudo se propôs a analisar os *tweets* que continham as palavras “ansiedade” e “depressão” em um período de 23 dias entre julho e agosto de 2020. A escolha desses termos ocorreu devido aos desdobramentos que a pandemia do novo Coronavírus acarreta diretamente a saúde mental, conforme estudos mostram (OLIVEIRA *et al*, 2020; SARTÓRIO *et al*, 2020; SHIGEMURA, 2020).

Devido à universalização das redes sociais e, conseqüentemente, da Mineração de Textos, optou-se pela análise de sentimentos. Observou-se a importância de se trabalhar com um dicionário mais alinhado possível ao tema pesquisado, para que a classificação das polaridades esteja adequada à realidade do problema. Como a análise de textos é um tema relativamente novo e em disseminação no Brasil, não existem muitos dicionários disponíveis em português. Com o objetivo de as análises refletirem os sentimentos mais próximos dos reais, foram criados dois dicionários com termos pertinentes ao contexto de ansiedade e de depressão. De acordo com o universo analisado, cerca de 81% dos *tweets* foram classificados com polaridade negativa.

Durante o período analisado, quarta e quinta-feira foram os dias da semana com as maiores médias de publicações. Parte desse volume ligeiramente maior de postagens pode estar relacionada com os jogos de futebol. Isso acaba por aumentar as postagens de ansiedade, mas não no contexto de doença, mas sim no contexto de expectativa pela competição esportiva. Já no que se refere ao horário, observou-se que o pico das publicações ocorreu entre meia noite e 2 horas da manhã. Era esperado que o pico não ocorresse durante o dia, por conta da rotina diária, e também por entender que as pessoas se sentiriam mais solitárias aos finais de semana e, conseqüentemente, publicariam mais. Esta última hipótese não se confirmou. Contudo, o pico das publicações dá indícios para uma das consequências da ansiedade e da depressão, que conforme visto na nuvem de comparação e na análise semântica, foi bastante mencionada entre os usuários: a insônia.

A partir do contexto geral dos *tweets* apresentado na nuvem de comparação já se torna possível observar que alguns aspectos relacionados ao tratamento são mencionados de maneira recorrente. Já algumas outras palavras, como “lindo”, “linda” e “bom” foram classificadas com polaridade neutra pois observou-se que, na maioria, estão inseridas em um contexto de ironia.

Por fim, a análise semântica evidenciou que os usuários também demonstram estarem ansiosos por estímulos não diretamente relacionados ao coronavírus como, por exemplo, volta do futebol, reexibição de novela, lançamento de jogos ou séries, espera por compras *on line*, etc. No entanto, esse tipo de publicação não foi a predominante. Os sintomas de ansiedade e de depressão estiveram muito presentes nos relatos do Twitter. Outro ponto que chama atenção são os gatilhos mencionados e que estão relacionados a questões atuais, tais como machismo, abusos sexuais, homofobia, racismo, etc.

Em síntese, apesar da riqueza de informações sobre o contexto de saúde mental como um todo, obtida a partir da análise semântica, fatores diretamente relacionados ao contexto mundial de pandemia foram mencionados significativamente. Gatilhos como desemprego, instabilidade política e isolamento social, assim como vistos em outros estudos, também aparecem quando partimos da ótica de Mineração de Textos.

6 REFERÊNCIAS

- 1 Ano após Novo Limite de Caracteres no Twitter, Tweets ficaram mais curtos. **Tecmundo**. Disponível em: < <https://www.tecmundo.com.br/redes-sociais/135727-1-ano-novo-limite-caracteres-twitter-tweets-curtos.htm> >. Acesso em 15 set. 2020.
- A Campanha do Setembro Amarelo Salva Vidas. **Setembro Amarelo**. Disponível em: <<https://www.setembroamarelo.com/>>. Acesso em: 10 set. 2020.
- Análise de Sentimentos e Machine Learning. **Ciência e Dados. Data Science for Professionals**. Disponível em: <<http://www.cienciaedados.com/analise-de-sentimentos-e-machine-learning/#:~:text=An%C3%A1lise%20de%20sentimentos%20%C3%A9%20a,texto%2C%20%C3%A9%20positiva%20ou%20negativa>>. Acesso em: 21 set. 2020.
- ARANHA, C. PASSOS. **A Tecnologia de Mineração de Textos**. Revista Eletrônica de Sistemas de Informação, n. 2, 2006.
- Brasileiros Buscaram Três Vezes mais pelo Termo “Ansiedade” no Google do que a Média dos últimos 16 anos. **G1**. Disponível em: <<https://g1.globo.com/bemestar/viva-voce/noticia/2020/08/25/brasileiros-buscaram-tres-vezes-mais-pelo-termo-ansiedade-no-google-do-que-a-media-dos-ultimos-16-anos.ghtml>>. Acesso em: 18 set. 2020.
- CAMBERO, A. **A Comparative Study of Twitter Sentiment Analysis Methods for Live Applications**. Orientador: Joe Geigel. 2016. 46 f. Tese (Mestrado em Ciência da Computação) – Instituto de Tecnologia de Rochester, Rochester – 2016.
- CORRÊA, I, T. **Análise dos Sentimentos expressos na Rede Social Twitter em relação aos Filmes Indicados ao Oscar 2017**. 2017, 72 f. Trabalho de Conclusão de Curso – Universidade Federal de Uberlândia, Uberlândia, 2017.
- COSSETTE, P.; AUDET, M. **Mapping of an idiosyncratic schema**. *Journal of Management Studies*, p. 325 – 47, 1992.
- COMO A QUALIDADE DO SONO AFETA A SAÚDE MENTAL. **CUIDAÍ**. Disponível em: < <https://cuidai.com.br/qualidade-sono-saude-mental/> >. Acesso em: 18 set. 2020.
- Curvas de Densidade. **Khan Academy**. Disponível em: <<https://pt.khanacademy.org/math/ap-statistics/density-curves-normal-distribution-ap/density-curves/v/density-curves>>. Acesso em: 23 set. 2020.
- DAVENPORT, T. H. **How Strategists use “big data” to Support Internal Business Decisions, Discovery and Production**. *Strategy & Leadership*, v. 42, n. 4, p. 45 – 50, jul./2014.

Desemprego Aumenta em 11 Estados no 2º Trimestre; Só 2 Registram Queda.

Uol. Disponível em: < <https://economia.uol.com.br/empregos-e-carreiras/noticias/redacao/2020/08/28/desemprego-pnad-continua-ibge.htm#:~:text=A%20taxa%20de%20desemprego%20aumentou,%2C%20de%2013%2C3%25>>. Acesso em: 19 set. 2020.

Docs. **Twitter Developer**. Disponível em: <<https://developer.twitter.com/en/docs>>. Acesso em 07 ago. 2020.

Getting Started. The R Project for Statistical Computing Disponível em: <<https://www.r-project.org/>>. Acesso em: 30 jul. 2020.

GUIMARÃES, F. M. **Aplicação do mapeamento cognitivo como apoio à implementação de estratégias empresariais: o caso de uma organização hospitalar**. 181 p. Dissertação de mestrado em Administração e Negócios, da Faculdade de Administração, Contabilidade e Economia- Universidade Católica do Rio Grande do Sul, 2007.

Introducing JSON. **JSON Org**. Disponível em: <<https://www.json.org/json-en.html>>. Acesso em: 10 ago. 2020

MARGIS, R.; PICON, P.; COSNER, A. F.; SILVEIRA, R. OLIVEIRA de. **Relação entre Estressores**. Revista de Psiquiatria do Rio Grande do Sul, v. 25, n. 1, abr./2003.

PANG, B; LEE, L. **Opinion Mining and Sentiment Analysis**. Foundations and Trends in Information Retrieval, v. 2, p. 1 – 135, jun./2008.

PEREIRA, M. D.; de OLIVEIRA, L. C.; COSTA, C, F, T.; BEZERRA, C. M. O. de; PEREIRA, M. D. **A Pandemia de COVID-19, o Isolamento Social, Consequências da Saúde Mental e Estratégias de Enfrentamento: Uma Revisão Integrativa**. Research, Society and Development, v. 9, n. 7, p. 1 – 35, 2020.

Pesquisa da UERJ Indica Aumento de Casos de Depressão entre Brasileiros Durante a Quarentena. **UERJ**. Disponível em: <<https://www.uerj.br/noticia/11028/>>. Acesso em: 12 de set. 2020.

REED, P.; VILE, R.; OSBORNE, L. A.; ROMANO, M.; TRUZZOLI, R. **Problematic Internet Usage and Immune Function**. Plos One, v. 10, n. 8, ago./2015.

REFATTI, M; GOULART, M. C. F.; STEIN M., MERINO, E. A. D. **Uso de mapas semânticos no processo de naming – uma abordagem estratégica em um estudo de caso aplicado**. Revista Brasileira de Design da Informação, v. 11, n. 1, 2014.

Relatos de Brigas de Casais Aumentam 431% desde o Início do Isolamento Provocado pelo Coronavírus, diz estudo. **G1**. Disponível em: < <https://g1.globo.com/bemestar/coronavirus/noticia/2020/04/20/relatos-de-briga-de->

[casais-aumentam-431percent-desde-o-inicio-do-isolamento-provocado-pelo-coronavirus-diz-estudo.ghml](#)>. Acesso em: 20 set. 2020.

Rtweet. **R Documentation**. Disponível em: <
<https://www.rdocumentation.org/packages/rtweet/versions/0.7.0>>. Acesso em: 10 ago. 2020.

SARTÓRIO, C. L.; JUIZ, P. J. L.; de MELO RODRIGUES, L. C.; ÁLVARES-DA-SILVA, A., M. **Paradoxos de Retroalimentação da Pandemia da COVID-19: Quebrando o Ciclo**. Cadernos de Prospecção, v. 13, n. 2, 2020.

SHIGEMURA, J.; MORGANSTEN, J. C.; BENEDEK, D. M. **Public responses to the novel 2019 coronavirus (2019-nCoV) in Japan: Mental health consequences and target populations**. Psychiatry and Clinical Neurosciences, v. 74, p. 277 – 283, 2020.

TAN, A. **Text Mining: The state of the art and challenges**. In *Proceedings PAKDD'99 Workshop on Knowledge Discovery from Advanced Databases*, p. 71–76, 1999.

Twitter Usage Statistics. **Internet Live Stats**. Disponível em:
<<https://www.internetlivestats.com/>>. Acesso em: 13 set. 2020.

YOUNIS, E. M. G. **Sentiment Analysis and Text Mining for Social Media Microblogs using Open Source Tools: Na Empirical Study**. International Journal of Computer Applications, v. 112, p. 44 – 48, fev./2015.

Uma Introdução ao JSON. **Devmedia**. Disponível em:
<<https://www.devmedia.com.br/json-tutorial/25275>>. Acesso em: 01 set. 2020.

Veja Como Aderir à Nova Prorrogação de Acordos de Suspensão de Contrato e Redução de Salário de Doméstica. **Extra**. Disponível em: <
<https://extra.globo.com/noticias/economia/veja-como-aderir-nova-prorrogacao-de-acordos-de-suspensao-de-contrato-reducao-de-salario-de-domestica-24639412.html>>. Acesso em: 19 set. 2020.

WEBER, J.; JAVELLE, F.; KLEIN, T.; FOITSCHIK, T.; CRUCIAN, B.; SCHNEIDER, S.; ALBELN. **Neurophysiological, Neuropsychological, and Cognitive Effects of 30 days of Isolation**. Experimental Brain Research, v. 237, p. 1563 – 1573, mar./2019.

7 ANEXOS

Quadro 1 – Dicionário de variáveis coletadas

(continua)

| Variável | Tipo | Descrição |
|----------------------|----------|---|
| user_id | string | Número de identificação do usuário |
| status_id | string | Número de identificação do status |
| created_at | string | Data e hora em que o tweet foi criado |
| screen_name | string | Arroba do autor do tweet |
| Text | string | Texto da publicação |
| Source | string | Dispositivo usado para postar. Se for postado pelo site do twitter aparecerá como origem da web, se for mobile aparecerá se foi Iphone, Ipad, Android, TweetDeck.. |
| display_text_width | inteiro | Quantidade de caracteres na publicação |
| is_quote | booleano | Indica se a publicação contém algum retweet como comentário. |
| favorite_count | inteiro | Indica aproximadamente quantas vezes este Tweet foi curtido pelos usuários do Twitter. |
| retweet_count | inteiro | Número de vezes que o tweet foi retweetado. |
| hashtags | string | Hashtags usadas no tweet. |
| mentions_user_id | string | Id do usuário mencionado na publicação. |
| mentions_screen_name | string | Arroba do usuário mencionado na publicação. |
| bbox_coords | array | É uma série de pontos de longitude e latitude, que define uma caixa que conterá a localização do usuário à qual esta caixa delimitadora está relacionada. Cada ponto é uma matriz na forma de [longitude, latitude]. Os pontos são agrupados em uma matriz por caixa delimitadora. Matrizes de caixas delimitadoras são agrupadas em uma matriz adicional para serem compatíveis com a notação de polígono. |
| name | string | Nome do usuário |
| location | string | Localização definida pelo usuário. É preenchido manualmente pelo usuário e pode ser editado a qualquer momento. |
| protected | booleano | Indica se o usuário optou por proteger seus tweets. |
| followers_count | inteiro | Número de seguidores |
| friends_count | inteiro | Número de perfis seguidos pelo usuário |
| listed_count | inteiro | Número de listas públicas das quais o usuário é membro. |
| statuses_count | inteiro | Número de tweets (incluindo retweets) que a conta publicou. |

(conclusão)

| | | |
|--------------------|----------|--|
| favorites_count | inteiro | Número de tweets que o usuário curtiu durante a vida útil da conta. |
| account_created_at | string | Data e hora em que a conta foi criada. |
| verified | booleano | Indica se o usuário tem uma conta verificada. Por exemplo, pessoas de interesse público. |
| place_type | string | Tipo de local. Exemplo: cidade, estado |
| place_name | string | Nome do local |
| place_fullname | string | Nome da cidade e do estado |

Fonte: Twitter