TRABALHO FINAL

(revisado)

Aluno: Pedro Henrique Nogueira Silva

Matricula: 76435

Prof.: Cristiano Rodrigues de Carvalho

# Introdução

O surgimento das redes sociais trouxe um estrondoso crescimento no volume de dados e com cada vez mais recursos nas redes sociais a variedade e a veracidade dos dados disponibilizados podem se tornar maior e mais segura.

Algumas dessas redes sociais se destacam pela capacidade do seu usuário conseguir expressar suas opiniões e sentimentos a partir da elaboração de texto, podemos citar nessa categoria Facebook e Twitter. Podemos considerar hoje, ambas as redes, dentre as mais utilizadas no mundo. Com este número crescente de usuários cada dia mais, os dados oriundos das mesmas tornam-se consideravelmente relevantes e uteis em qualquer análise realizada.

A exploração de dados proporciona a descoberta de diversas informações, graças as tais técnicas podemos aplicar em contextos diversos. Para este estudo utilizaremos a análise léxica em dados obtidos do Twitter afim de inferir certas características sobre seus usuários.

As características encontradas serão baseadas sobre o Big Five, que se trata de um modelo com base em analise linguística, tendo como resultado o enquadramento da personalidade do indivíduo em cinco possíveis fatores.

# Objetivo

Tal estudo tem como finalidade a coleta de dados de um usuário específico da rede Twitter. Após, será realizado analise léxica em busca de fatores de personalidade daquele usuário.

Será utilizado para tal coleta, a ferramenta Knime Analytics e o modelo de personalidade o Big Five.

# Big Five e Personalidade

Diversos estudos tentam descrever com total certeza o que seria a personalidade de um indivíduo, quais características descrevem o comportamento exato. Desses estudos, aparecem diversas formas de origem das características e suas definições. Autores como D’Andrea em 1982 coloca como o meio em que o indivíduo esta como um dos possíveis fatores. Outros autores, como Allport em 1973 e Schultz & Schultz em 2002, trazem o conceito de traços de personalidade, onde as repostas de um indivíduo a estímulos dão resultados a personalidade.

Baseado neste novo conceito surgiram uma linha de pesquisa sobre os traços de personalidade, dessas novas linhas surgiu o modelo dos Cinco Grandes Fatores (CGF ou Big Five). Tendo como o objetivo descrever a personalidade do indivíduo de forma pratica e rápida, ganhou notoriedade devido a sua simplicidade e a possibilidade de ser utilizada de diversas maneiras. Atualmente o Big Five descreve os seguintes traços de personalidade:

* (E) Extroversão: intensidade de interação interpessoal
* (S) Socialização: interações que uma pessoa apresenta ao longo de um contínuo que se estende da compaixão ao antagonismo
* (C) Conscienciosidade/Realização: representa o grau de organização, persistência, controle e motivação para alcançar objetivos
* (N) Neutorismo: avalia o ajustamento afetivo versus instabilidade emocional
* (A) Abertura: indivíduos originais, independentes, indagadores, criativos, ousados, que deliberadamente procuram e apreciam experiências novas e os contrasta com aqueles mais convencionais.

Tais definições são encontrar e com maior explicação em (Braz, 2011)

# Metodologia

O trabalho consistiu das seguintes etapas:

* Coleta de tweets públicos de um usuário teste.
* As ferramentas utilizadas no estudo são voltadas para a língua inglesa. Foram utilizados classificadores léxicos e ferramentas de pré-processamento textual. Com essas ferramentas conseguimos classificar cada palavra de acordo com sua classificação gramatica e ainda, distinguir palavras positivas de negativas e termos de negação.
* Por fim, utilizamos essas palavras classificadas e enumeradas para determinar os traços de personalidade do usuário.

Com a coleta das palavras, foram utilizadas as seguintes expressões para a classificação das personalidades. Tais expressões podem ser encontradas no trabalho de Nunes et. Al. , 2013.

E=(substantivo+adjetivo+preposição+artigo-pronomes-verbos-adverbios-interjeição+100)/2

S = (palavras positivas-palavras negativas-artigos+100)/2

C = (palavras positivas-negações-palavras negativas+100)/2

N = (palavras negativa-palavras positivas+primeira pessoa-terceira pessoa+100)/2

A = (palavras longas-palavras curtas/2+terceira pessoa-primeira pessoa+100)/2

# Resultado

Para o desenvolvimento do trabalho, foi selecionado uma das personalidades mais influentes da atualidade e com traços de personalidade bem definidos.

Foi realizada a coleta dos tweets do atual presidente dos Estados Unidos, Donald Trump.

Para o experimento foi utilizado como exemplo o twiiter do atual presidente dos Estados Unidos – Donald Trump

Para o experimento foi coletado cerca de 200 tweets, sendo os mais recentes.

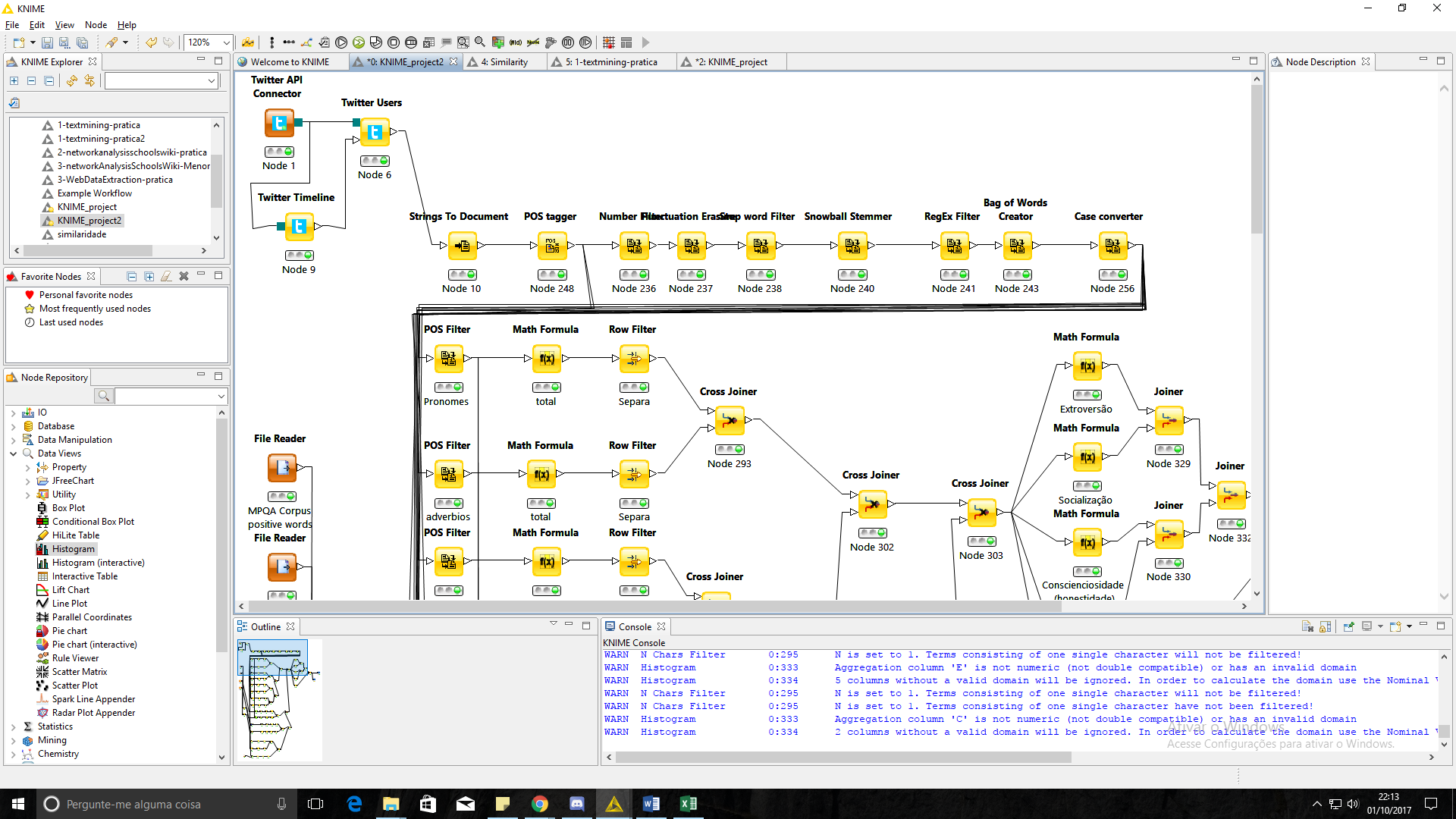


Figura 1: Preprocessamento e limpeza das palavras

Nesta etapa do experimento, realizamos a coleta dos tweets e a limpeza das palavras.

O Node POS tagger tem a função de classificar as palavras quanto sua colocação gramatical.

Após essa etapa começaremos o processo de filtragem das palavras e contagem das mesmas.

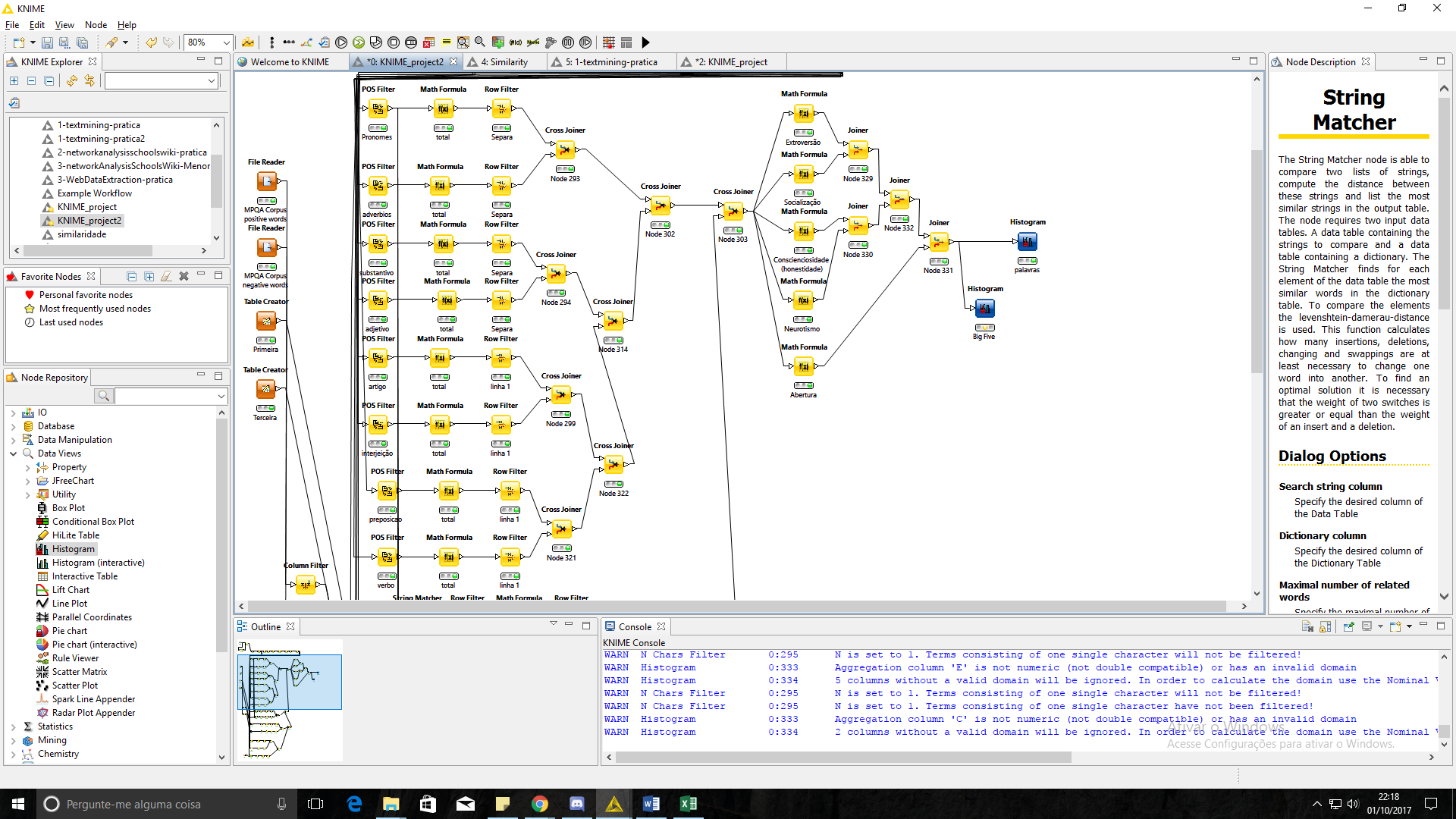
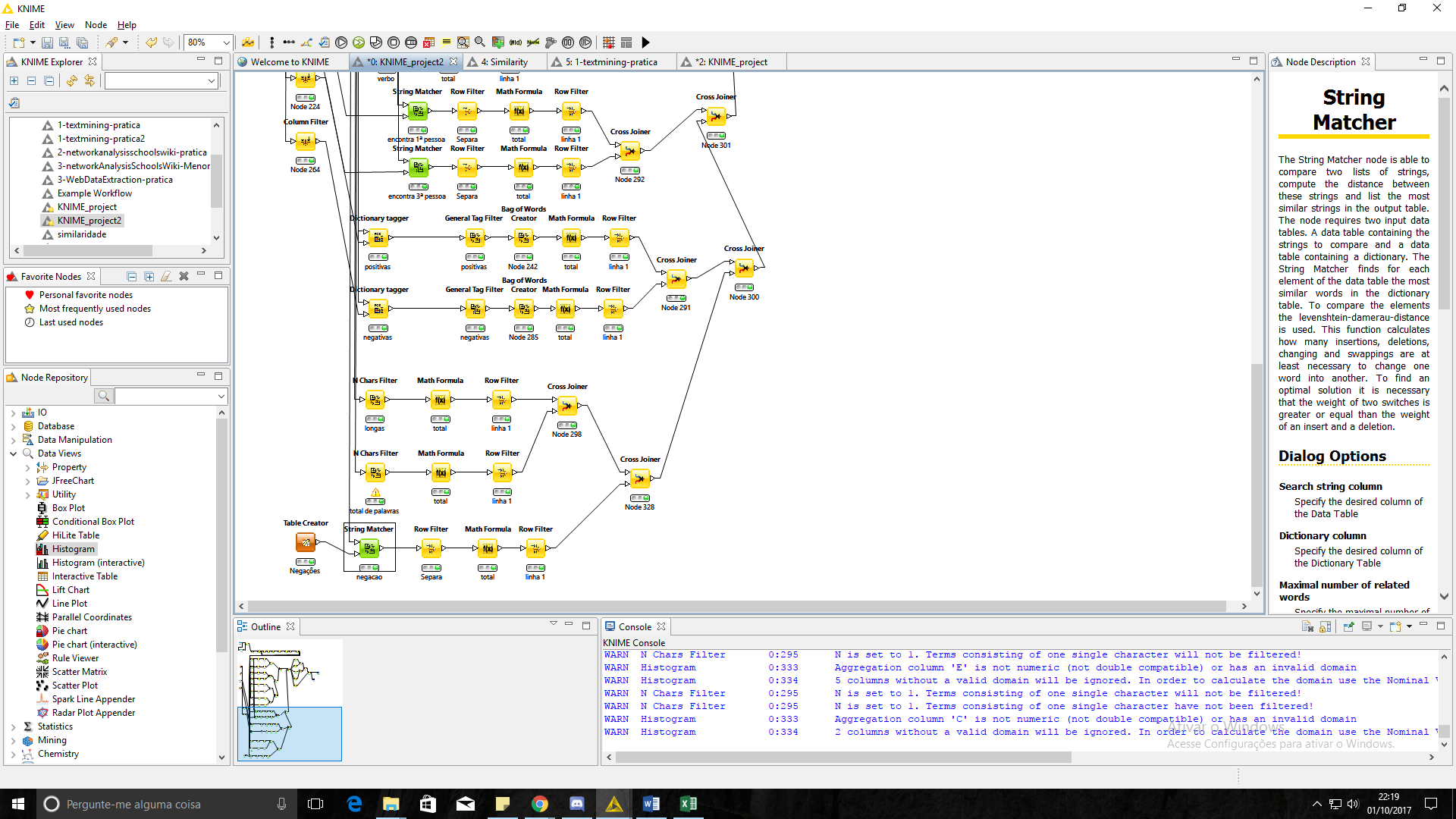
 

Figura 2: Classificação das palavras. Figura 3: Busca de palavras especificas no conjunto.

Após a classificação e contagem é feito o processo de execução das equações.

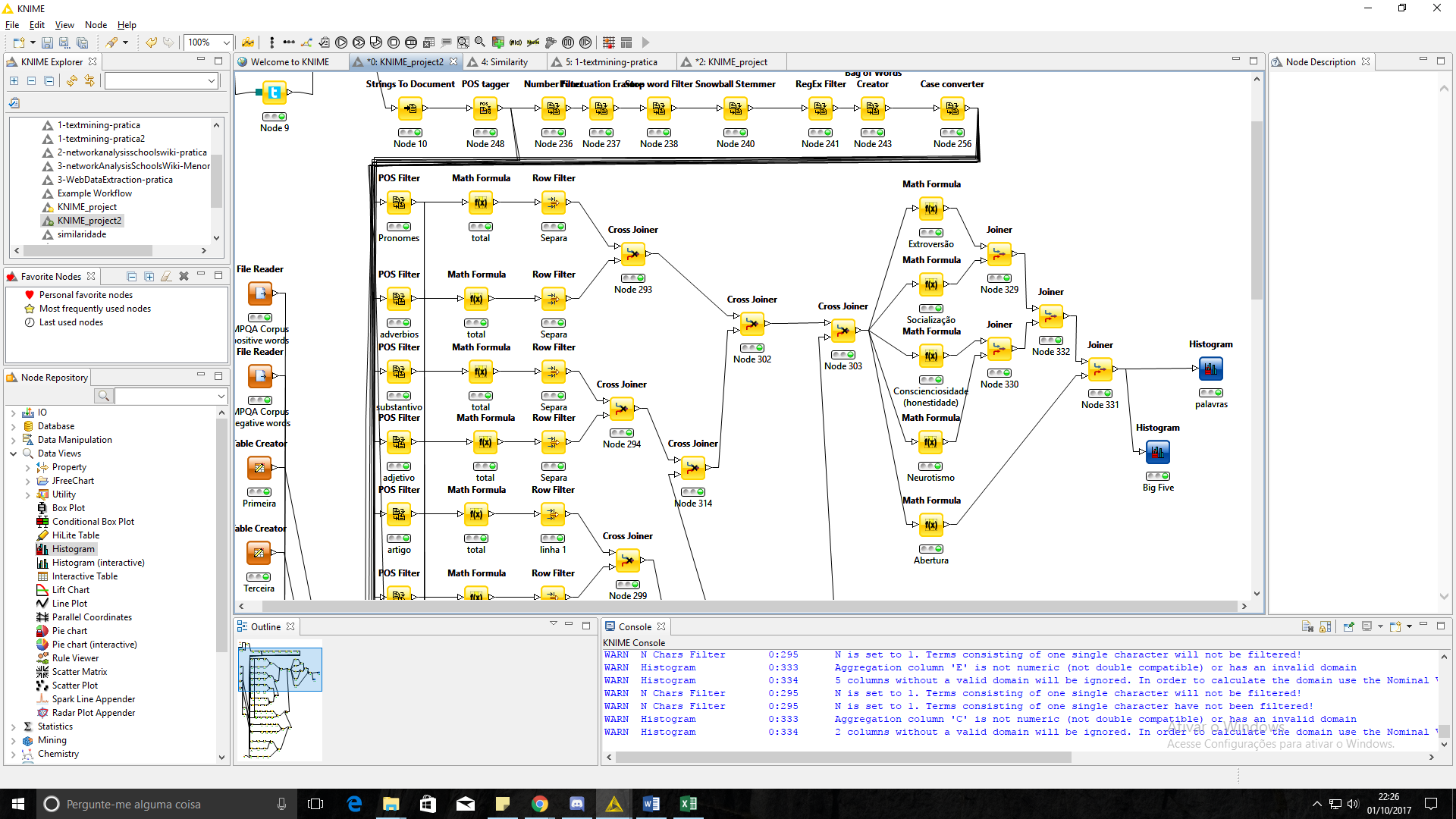


Figura 4: Execução das expressões.

# Resultados

Com o final da coleta temos os dados classificados e assim podemos começar a enquadrar os traços de personalidade expressos pelo atual presidente.

Temos a seguir a contagem de palavras, seguindo algumas características.

Palavras Longas: Palavras com mais de 7 letras.

Palavras Negação: Palavras que expressão negação e tornam as orações negativas.

|  |  |
| --- | --- |
| Palavras | Valores |
| Pronomes | 163 |
| Adverbio | 9 |
| Substantivo | 1350 |
| Adjetvio | 195 |
| Artigo | 215 |
| Interjeição | 1 |
| Preposição | 361 |
| Verbo | 551 |
| 1ª Pessoa | 69 |
| 3ª Pessoa | 63 |
| Positivas | 244 |
| Negativas | 94 |
| Longas | 397 |
| Curtas | 3045 |
| Negação | 21 |

Após a classificação das palavras e utilização das expressões, temos os dados abaixo e assim podemos mostrar os traços de personalidade encontrados, referente aos últimos 200 tweets.

|  |  |
| --- | --- |
| Personalidade | Valores |
| Extroversão | 749 |
| Socialização | 18 |
| Conscienciosidade | 115 |
| Neurotismo | -22 |
| Abertura | -516 |

# Análise dos resultados

## Abertura

Polo positivo: gostam de novidades e tendem a ser criativas.

Polo negativo: gostam da rotina e têm senso aguçado do certo e do errado

Valor obtido: - 516

## Neurotismo

Polo positivo: ansiosas, inibidas, melancólicas e dotadas de baixa autoestima

Polo negativo: otimistas e dotadas de boa estima consigo mesmas

Valor obtido: - 22

## Conscienciosidade

Polo positivo: grande motivação, disciplinado, comprometido e confiáveis.

Polo negativo: indisciplinado

Valor obtido: 115

## Socialização

Polo positivo: compassiva, amistosa e calorosa

Polo negativo: retraídos, críticos e egocêntricos

Valor obtido: 18

## Extroversão

Polo positivo: afabilidade, sociabilidade e capacidade de se impor

Polo negativo: introversão, reserva e submissão

Valor obtido: 749

# Conclusão

Ao final do estudo podemos perceber que alguns traços de personalidade encontrados correspondem claramente com o que Donald Trump demonstra em suas colocações, tanto nos seus tweets como também em seus discursos.

Podemos destacar a alta capacidade de se impor, polo positivo da extroversão, juntamente com sua grande motivação e comprometimento, polo positivo de conscienciosidade, onde ele permanece com sua posição desde o momento da sua pré-candidatura até o momento atual. Outro ponto que podemos destacar é sua abertura, por ser um candidato Republicano é normal o mesmo ter pensamentos estáticos em relação aos costumes do seu pais e aos ideais já difundidas.

# Referências Bibliográficas

(NUNES et al., 2013) NUNES, M.a.s.n. et al. INFERINDO PERSONALIDADE VIA TWEETS. **Revista Gestão, Inovação e Tecnologias,**[s.l.], v. 3, n. 3, p.045-057, 15 set. 2013. Associacao Academica de Propriedade Intelectual. http://dx.doi.org/10.7198/s2237-0722201300030004.

# (BRAZ, 2011) BRAZ, Lana Magna Sousa. ***Configuração dos Cinco Grandes Fatores de Personalidade em Estudantes de Psicologia.*** Trabalho apresentado XIX Seminário de Iniciação Científica da UFG – PIBIC, 2011.

(SCHULTZ & SCHULTZ , 2002) SCHULTZ, D. P, SCHULTZ, S. E. ***Teorias da Personalidade.*** Ed. Pioneira Thomson Learning, São Paulo- SP, 2002, citado em BRAZ, (2011).

(ALLPORT, 1973) ALLPORT, G.W.. ***Personalidade.***  São Paulo, Universidade de São Paulo – USP, 1973, citado em BRAZ, (2011).