**análises textuais: (21/5) 19h**

**- frequência ok**

**- gráfico de centralidade e co-ocorrência ok**

**- dispersão lexical (plot) – fazer tb na linha temporal ok**

**- análise LDA ok**

**- sentence similarity não sei se faz sentido (eu preciso comparar os corpus do Twitter com os dos jornais,**

**- sentiment analysis (você acha que a gente não supervisiona ou supervisiona? tô em dúvida por conta do Twitter. mas tb como a gente ia comparar os resultados usando um método para analisar mídia e outro para analisar Twitter?) – fazer tb na linha temporal**

[**https://spacy.io/usage/vectors-similarity**](https://spacy.io/usage/vectors-similarity)

**análises longitudinais - palavras? (Twitter, jornais) atores mais influentes (relevância, curtidas)?(Twitter)**

**impeachment**

**corrupção**

**análises de rede:**

**- clusterização em torno de governo\_Dilma + nós com maior indegree**

**- atores mais influentes (histograma e gráfico)**

**- tweets mais influentes (histograma e gráfico)**

**- histograma dos atores mais influentes**

**Depois vai ter a parte de comparar as análises de sentimento e tentar achar associação ou correlação entre mídia, Twitter e polls (eeee)**

**Eventos (análises Twitter e mídia)**

**1) Protestos de junho de 2013 – junho de 2013 (maio-julho de 2013)**

**2) Início Lava Jato – março de 2014 (fev-abril de 2014)**

**3) Eleições 2014 – outubro de 2014 (set-nov de 2014)**

**4) Protestos de março de 2015 – março de 2015 (fev-abril de 2015)**

**5) Agosto de 2015 (verificar o que foi - foram protestos ver Arnaudo p.18. Comprop) – (julho-set 2015)**

**6) Aprovação do impeachment de Dilma – maio de 2015 (abr- jun 2016)**

* **atenção na mídia + atenção no Twitter (já tem o histograma) na linha temporal -> tipo o gráfico que peguei num artigo, *trends in attention***

Checks:

Foi utilizado o seguinte subconjunto: **Protestos de junho de 2013 – junho de 2013 (maio-julho de 2013)**