



# NEW YEAR'S SPEECHES

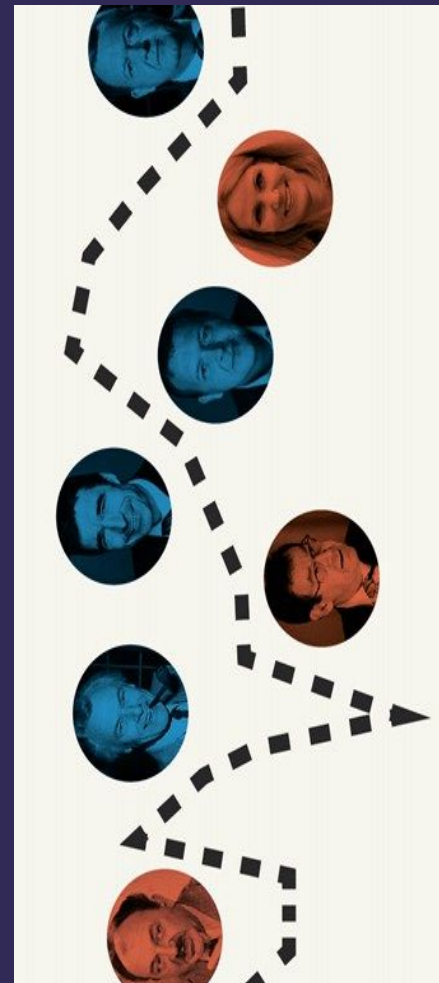
## DANISH PRIME MINISTERS

CHANGES OVER TIME / ACROSS PARTIES?

SENTIMENT, TOPICS, LIX, POLARISATION, WORDCLOUDS



df_transcripts					Python
...		Transcripts	Minister	Party	Wing
1998	godaften . skelsættende år danmark . år , dans...		Poul Nyrup	Socialdemokratiet	Red
1999	godaften ! mennesker nytåret naturlig anlednin...		Poul Nyrup	Socialdemokratiet	Red
2000	godaften – første dag nye årtusinde . sandelig...		Poul Nyrup	Socialdemokratiet	Red
2001	godaften – nytår.i oktober besøgte sydkorea , ...		Poul Nyrup	Socialdemokratiet	Red
2003	godaften – nytår ! danske familier netop fejre...		Anders Fogh	Venstre	Blue
2004	godaften – nytår ! måneder tid afslappede form...		Anders Fogh	Venstre	Blue
2005	godaften ! nytår præget tragedie , kender fuld...		Anders Fogh	Venstre	Blue
2006	godaften ! år stod verden ansigt ansigt ufatte...		Anders Fogh	Venstre	Blue
2007	godaften ! går tog afsked . år , længe huske d...		Anders Fogh	Venstre	Blue
2008	godaften ! dag taget hul år . friskt ubrugt li...		Anders Fogh	Venstre	Blue
2009	godaften ! første dag nye år . spændte , år br...		Anders Fogh	Venstre	Blue
2010	godaften.i går tog afsked årti , lærte kende n...		Lars Løkke	Venstre	Blue
2011	tror , : nytåret tid året , status . tænker ti...		Lars Løkke	Venstre	Blue
2012	godaften.for måneder danmark regering . regeri...	Helle Thorning Schmidt	Socialdemokratiet	Red	
2013	godaften.vi danskere grundlæggende optimister ...	Helle Thorning Schmidt	Socialdemokratiet	Red	
2014	godaften.for uger deltog mindehøjtideligheden ...	Helle Thorning Schmidt	Socialdemokratiet	Red	
2015	aften.en torsdag december sad gulvet stue valb...	Helle Thorning Schmidt	Socialdemokratiet	Red	
2016	aften.i december deltog arrangement avis , hve...	Lars Løkke	Venstre	Blue	
2017	godaften.de seneste dage , danskere verdensber...	Lars Løkke	Venstre	Blue	
2018	aften.mellem jul nytår gik tænkte mennesker , ...	Lars Løkke	Venstre	Blue	
2019	godaften.i dag taget fat år , umiddelbart tegn...	Lars Løkke	Venstre	Blue	
2020	godaften.i år hundrede år , sønderjylland genf...	Mette Frederiksen	Socialdemokratiet	Red	
2021	aften.vi taget hul år . bedre gamle . hårdt år...	Mette Frederiksen	Socialdemokratiet	Red	



# GET & CLEAN DATA USING PYTHON



```
def cleaning_transcripts(transcripts):  
    import string  
    transcripts = transcripts.lower()  
    transcripts = re.sub('\((.*?\))', '', transcripts)  
    transcripts = re.sub('\[.*?\]', '', transcripts)  
    # transcripts = re.sub('[%s]' % re.escape(string.punctuation), ' ', transcripts)  
    transcripts = re.sub('\w*\d\w*', '', transcripts)  
    return transcripts
```

[ ]

Python



```
def remove_stopwords(transcripts):  
    tokenized_transcripts=word_tokenize(transcripts)  
    transcripts_wo_stopwords= [w for w in tokenized_transcripts if not w in stop_words]  
    transcripts_wo_stopwords=' '.join(transcripts_wo_stopwords)  
    return transcripts_wo_stopwords  
  
#text = ' '.join([word for word in text.split() if word not in stopwordlist])
```

[ ]

Python

# "ANALYSIS" & VISUALISATION

## SENTIMENT USING SENTIDA

```
[ ]  
  
from sentida import Sentida  
  
sent_scores = []  
  
for text in df_transcripts['Transcripts']:  
    sentiment_score = Sentida().sentida(text,output = "total")  
    sent_scores.append(sentiment_score)  
  
[ ]
```

Python

## TØPIC MØDELLING

## WØRD CLOUDS



## LIX

### Lixtalsberegner

```
> def lix():  
    import re # regular expression modul importeres  
  
    text = raw_input("Text goes here:")  
  
    #Antal punktkummer beregnes  
    pktArray = text.split(".")  
    antalPkt = len(pktArray)-1  
  
    #Antal ord beregnes  
    textIdenTegn = re.sub('[\.,!'\',',text)  
    textArray = textIdenTegn.split()  
    antalOrd = len(textArray)  
  
    #Lange ord beregnes  
    langeOrd = 0  
    for i in textArray:  
        if len(i)>6:  
            langeOrd = langeOrd + 1  
  
    #LIX beregnes og udskrives  
    if antalPkt != 0 and antalOrd != 0:  
        lixTal = antalOrd/antalPkt + (langeOrd*100)/antalOrd #LIX = O/P + (L*100)/O  
        print "Lixtal:", lixTal  
        print "Antal ord i teksten:", antalOrd  
        print "Antal punktkummer i teksten:", antalPkt  
        print "Antal langeord i teksten:", langeOrd  
    else:  
        print "Lixtal kan ikke beregnes"  
  
    lix() #Tilslidst kaldes funktionen igen, så programmet ikke slutter  
  
lix()  
  
[ ]
```

Python

## BERT EMOTION CLASSIFICATION?

