# Latent Topic Modeling on Twitter Data

### Method: CA

Correspondence Analysis (CA) is a multivariate graphical technique designed to explore relationships among categorical variables. The outcome from correspondence analysis is a graphical display of the rows and columns of a contingency table that is designed to permit visualization of the salient relationships among the variable responses in a low-dimensional space. Such a representation reveals a more global picture of the relationships among row-column pairs which would otherwise not be detected through a pairwise analysis.

#### Calculate CA:

- Step 1: Compute row and column averages
- Step 2: Compute the expected values
- Step 3: Compute the residuals
- Step 4: Plotting labels with similar residuals close together
- Step 5: Interpreting the relationship between row and column labels

#### How to Interpret Correspondence Analysis Plots

Correspondence analysis does not show us which rows have the highest numbers, nor which columns have the highest numbers. It instead shows us the relativities.

- The further things are from the origin, the more discriminating they are.
- Look at the length of the line connecting the row label to the origin. Longer lines indicate that the row label is highly associated with some of the column labels (i.e., it has at least one high residual).
- Look at the length of the label connecting the column label to the origin. Longer lines again indicate a high association between the column label and one or more row labels.
- Look at the angle formed between these two lines. Really small angles indicate association. 90 degree angles indicate no relationship. Angles near 180 degrees indicate negative associations.

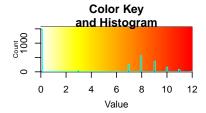
#### **Dataset**

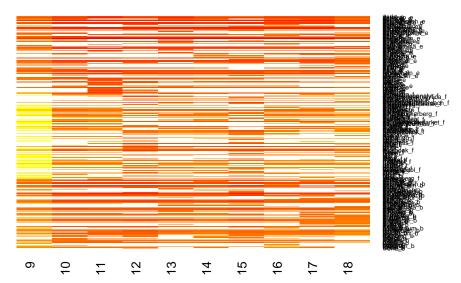
```
##
               9 10 11 12 13 14 15 16 17 18
## #ada e
                             8
## #airdrop_e
               9 10 11 10 11 10 10 10 10
## #altcoin e
               9 10 10
                       10
                          10
                             10
## #bch_e
                           9
                              8
                        0
## #binanc e
               8
                  9
                     9
                        9
                           9
                             9
## #bitcoin e 10 11 11 11 11 11 11 11 10
```

- Research Question
  - Do we see new words appearing on a particular week

# **Analysis**

### Heatmap



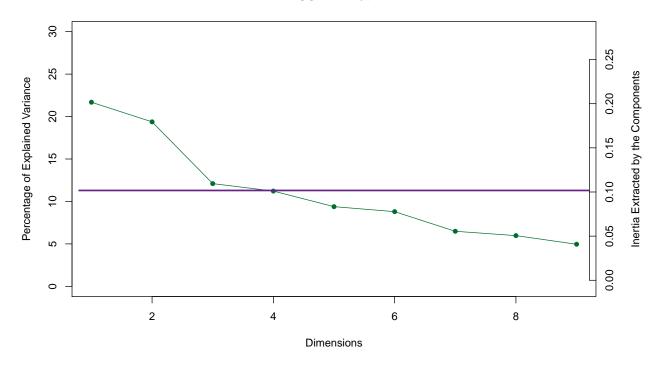


#### Scree Plot

Gives amount of information explained by corresponding component. Gives an intuition to decide which components best represent data in order to answer the research question.

P.S. The most contribution component may not always be most useful for a given research question.

#### **SCREE Plot**

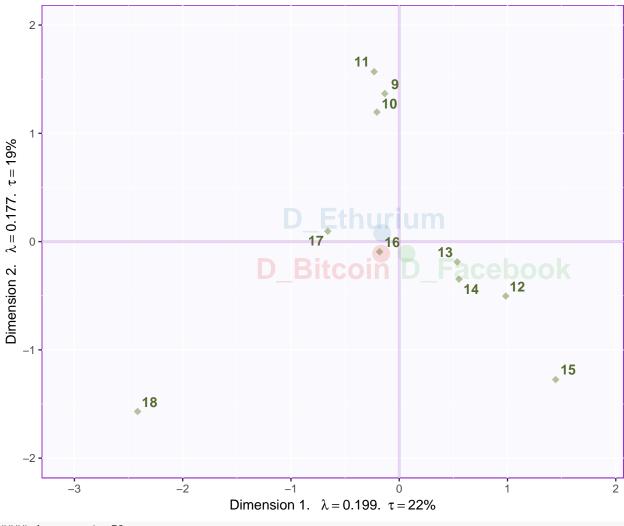


### **Factor Scores**

### Asymmetric Plot

```
#### Asymmetric Plot

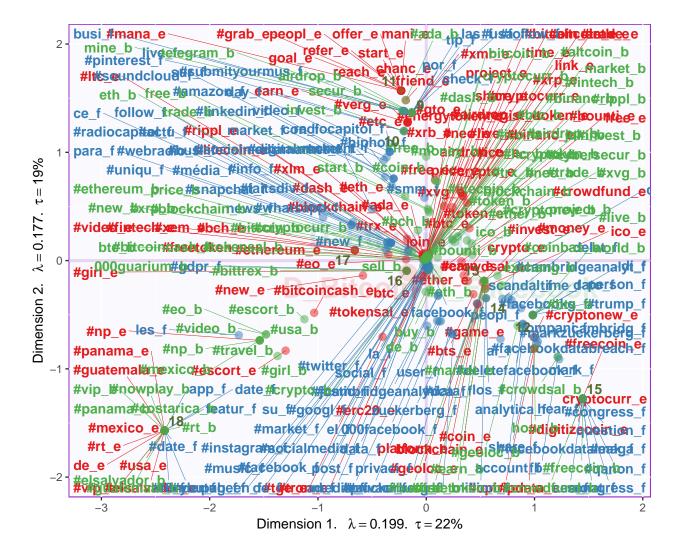
map.IJ.asym <- asymMap$baseMap + asymMap$J_labels +
   asymMap$J_points + labels4CA + legend$zeMap_dots + legend$zeMap_text
print(map.IJ.asym)</pre>
```



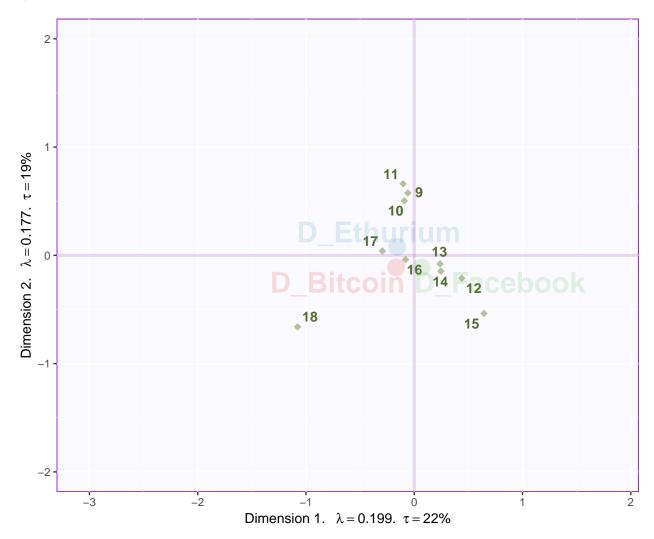
```
#### Asymmetric Plot

map.IJ.asym <- asymMap$baseMap + asymMap$I_labels +
   asymMap$I_points + labels4CA + legend$zeMap_dots + legend$zeMap_text
print(map.IJ.asym)</pre>
```

```
chlariethola alaschies hare glistryettad teoin
                 #subrityeurnase mani e offer_e start_e
                                                                                         #grab refer_e goal_e por_ffgto fico #affdoirkchain
               #pinterestmine b airdrop b
                                                                                                                           e#followigaitchgproedinkoe token_e
               #linkedin_radiocapitol_f busi_ftelegrampeop!reach
                                                                                                                                                     ndre#bec##lite#binabc e
               eth b #ada_e #amazon_tlay f #etc_e secur b
                                                                                                                              earn teade time eeth #tokensal b
               #actu_#trx_e fre#sbundclouddeo_f sur_finvest b
                                                                                                                              verg#grb##ve/##ico/market_lifree_e
                 #faits#webradi#dash market fron #hiphop f
                      para#média ce f #sippla lenediamankent f
             #litecoin_#info#whatsapportow #neodigitalmarkia#irecoin_technican was released by
 \lambda = 0.177. \tau = 19\%
                         #new_f news#snapchat#busi #xlm #xvg
                #xem#ethereutholockch#imjetradiocapheth etbch b
               #video_e #finterty peoculibits oin b
               sell_b #etbiecoincash_b #new_b
                                       su f 000guarium #freetoke#b@trex
               #adpr f
 Dimension 2.
                                                         #bitcoincash e
                                                                                                                                crowdsale #live b#earn#busd b
                                                                                                      #socialmedia #btsseandal To
                        #video b #eo_b
                                                                      #escort b
                                                                                                                                                                      ico excharcallbf
                              #mexico b les f
                                                                               #escort
                                                                                                                                                                #bts_btele#busi b
                                                                                                            noofacebook
#cambingeama
                       #usa_b #travel_b #np e
                          de e #np_b #girl b #f8 f
                                                                                        #justic/foe
                                                                                                                platfor trade exerace bookmark f world b
                                                                             #youtub
                                                                                               #tge eser #data / al f #pdata efb #freecoin_e
               #usa_e#panama_efeatur_f
               #mexico_e date_f#costaric#girl_e el_f cryptoclerbengeop#cfyptonewoevdsyptocureas_f
                                                                          #cambridgeanalpiockthain on pattopirium reecompanon_f
                #nowplay b
               #elsalvaldorate maildate_f en_f post f post 
               #ctuateanalevatildeterittela fattetc20/parbianitation
                      -3
                                                                          Dimension 1. \lambda = 0.199. \tau = 22\%
#### Asymmetric Plot
map.IJ.asym <- asymMap$baseMap + asymMap$I_labels +</pre>
    asymMap$I points + asymMap$J labels +
    asymMap$J_points + labels4CA + legend$zeMap_dots + legend$zeMap_text
print(map.IJ.asym)
```



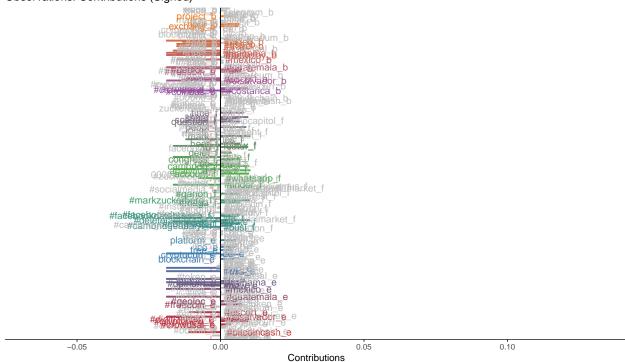
# Symmetric Plot



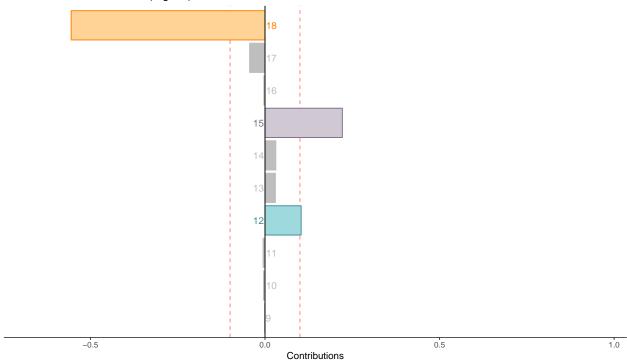
```
#mana_eofferstart_reachpeopl_e
                                                                                                                                            marfriendashedkulfallfultdein_e
                                                                                                                                  chance tip f #xmr #bitcoi#trade_e
               #sebmityourmus fefer e
                                                                                                                          #grab_#ada/b /#gto_e /#icø_e eth_#ltc_b
                                   mine_b
                                                               busitelegram
                     #linkedin_invest_b
                                               #hiphop_f#verg_e
                                                                                                                                               por airdrop Herypeochimo ebken e
                                                                                                                                            energytoken time the confintech_b
               #dash_etrade_b moment_f
                                                                                                                                                                          istusticom ibin hfree_e
               ce_f#rippl_e#xlm_e mark#anazoadiocapitol#etc
               #actu f #busi f #eth e fre#sbundcloudde
                    #faitsdi#whatsappollow #dashdighitalmar
\lambda = 0.177. \tau = 19\%
                                                                                                                               COI #freeireko
                                                                                                                                                                    pereuring_b
               eth wradiocapiwohedia #woxiamediamaskertib
                           #Itc e news#snap#wabradio#ada e
               #fint#thoekchair#rew_f#uniqu#info
               #blockchain_#bch_e #xem#vi
#cryptocurr_b
#new_b
               000quariun#bbttrex b
Dimension 2.
                 sell_b #gdpr_f
                                                                                                                                                                            ico e exchangousi b
                                                                                                                                       platfor e Gambadgealeatytf
                             #np b #new_e#escort_b
               #mexico_b les_f #video_b #eo_b#tokensal
                                                                                         #twitter/////b/w/_b/#earm/b#geølø¢
               #escort_teatur_f #f8_f #girl_#market #tge_e #bts
               #panama_b \ date_f #tinder frypto spciate c20user
                                                                                                                                              data to terroptone wife many procurre
                                                                                                                  le trucke betta de biook#chewdeath bir ecgina b
                                                                thereum cealmedia f
                                                                                                                        call #freetokerarbbridg
                                                                                                  el_f#marke#geolocaec#unark#ucker#peorggfes#mfaga f
                          elvació i peetati i na sito utia getto o lo éti pelva iditili vie teco in strare f
                                                                                                                                                                                #digitizecoinestion f
               #gsatemalade perpo o gle facttf coosta delicionalisticke qualivi le albanis o perpo o gle facttf coosta delicionalisticke qualivi le albanis o perpo o gle facttf coosta delicionalisticke qualivi le albanis o perpo o gle facttf coosta delicionalisticke qualivi le albanis o perpo o gle facttf coosta delicionalisticke quality i le albanis o perpo o gle facttf coosta delicionalisticke quality i le albanis o perpo o gle facttf coosta delicionalisticke quality i le albanis o perpo o gle facttf coosta delicionalisticke quality i le albanis o perpo o gle facttf coosta delicionalisticke quality i le albanis o perpo o gle facttf coosta delicionalisticke quality i le albanis o perpo o gle facttf coosta delicionalisticke quality i le albanis o perpo o gle facttf coosta delicionalisticke quality i le albanis o perpo o gle facttf coosta delicionalisticke quality i le albanis o perpo o gle facttf coosta delicionalisticke quality i le albanis o perpo o gle facttf coosta delicionalisticke quality i le albanis o positic delicionalisticke quality i le albanis o perpo o gle facttr coosta delicionalisticke quality i le albanis o perpo o gle facttr coosta delicionalisticke quality i le albanis o gle facttr coosta delicionalisticke quality i le albanis o gle facttr coosta delicionalisticke quality i le albanis o gle facttr coosta delicionalisticke quality i le albanis o gle facttr coosta delicionalisticke quality i le albanis o gle facttr coosta delicionalisticke quality i le albanis o gle facttr coosta delicionalisticke quality i le albanis o gle facttr coosta delicionalisticke quality i le albanis o gle facttr coosta delicionalisticke quality i le albanis o gle facttr coosta delicionalisticke quality i le albanis o gle facttr coosta delicionalisticke quality i le albanis o gle facttr coosta delicionalisticke quality i le albanis o gle facttr coosta delicionalisticke quality i le albanis o gle facttr coosta delicionalisticke quality delicionalisticke quality delicionalisticke quality delicionalisticke quality delicionalisticke quali
                       -3
                                                                                  Dimension 1. \lambda = 0.199. \tau = 22\%
```

### Most Contributing Variables









# Inference CA

