FIT3152 assignment 1 code

Lingustic analysis using LIWC

Jason Ching Yuen Siu

knitr::opts\_chunk$set(echo = TRUE,include=F, message =FALSE)

Import needed library

Read the data

Clean the data

Change the font style

==============================================================================================================

* Q1
  1. How active are participants, are there periods where this increases or decreases?
* b Looking at the linguistic variables, 1 do these change over time?

Since Clout’s change has been most turbulent, it is worth know which variables define this variable. We will use linear regression model to know it ——————————————————————————————– c Is there a relationship between variables? ——————————————————————————————–

Regression model

After refiting the model 1

After refiting the model 2

After refiting the model 3

============================================================================================================== - Q2

Analyse the language used by groups. Some starting points:  
  
a Threads indicate groups of participants communicating on the same topic. Describe the  
threads present in your data.  
--------------------------------------------------------------------------------------------

1. Sentiments : are most the thread +ve ?
   1. Pronoun : What are the most used pronoun?
   2. Structure : What are the mean of WC and WPS?

* b By analysing the linguistic variables for all or some of the threads, is it possible to see a difference in the language used by different groups? ——————————————————————————————–
* Find out the languages used between the most postivie and negative threads. 1. find out the 10 most active threads
  1. In the most active threads, find out the most postivie and negative threads. The data set is called data\_nega\_pose

Based on the data collected from the most negative and positive, we can make a Radarchart to visualise the difference of language used

The comparison of pronouns used between the most positive and negative

c Does the language used within threads (or between threads) change over time? How  
 consistent or variable is the language used within threads?  
--------------------------------------------------------------------------------------------

The following graphs are to view the change and consistency  
**Language structure (LangStructure) = WPS + WC** The more the LangStructure, the more complex the structure is

==============================================================================================================  
Q3 Challenge: Social networks online

This is to create tabulate\_dt() for tabulising the adjacency of nodes

create function (network\_centrality) for centrality summary of node within the network

Lets have a look at the network of 2002 feb

network for 2002 march