

Practice Text Network Analysis

Digital Business Ecosystem Research Center

WORDij



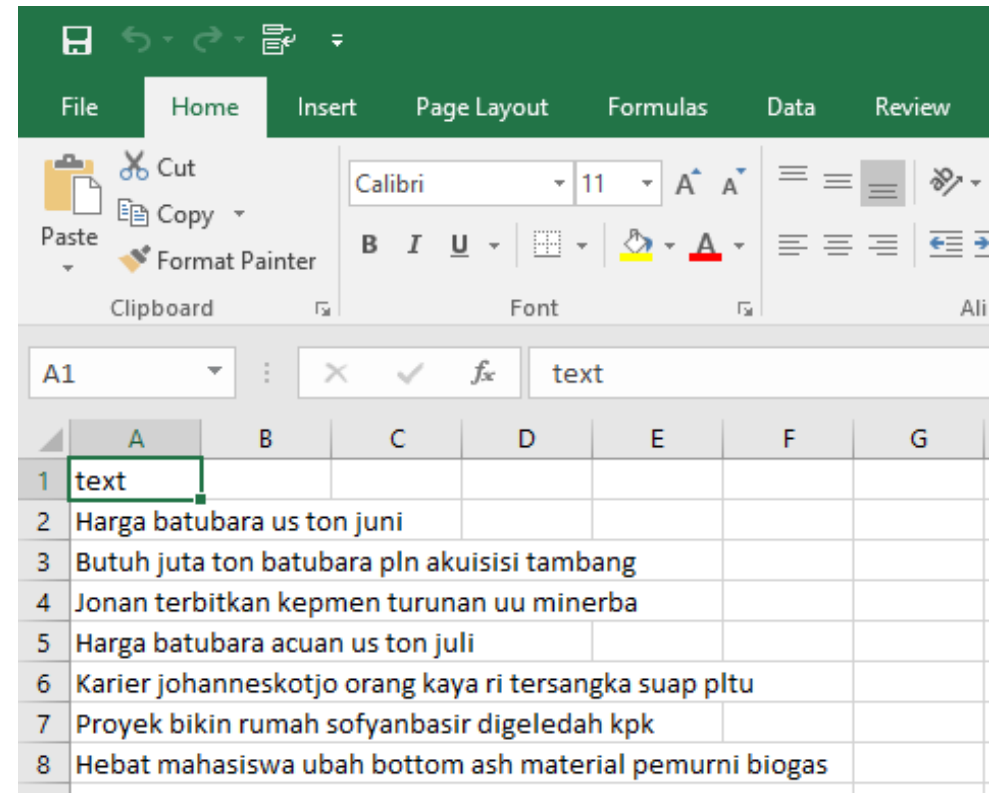
Text Network Analysis

- The analysis of text describing various kinds of "computer support solution" that allows the analyst to "extract networks of concepts" from text
- The assumption underlying the emergence text network is a relationship that is contained in every word so that it produces a pattern that can be analyzed

Source: Bratawisnu et al., (2017)

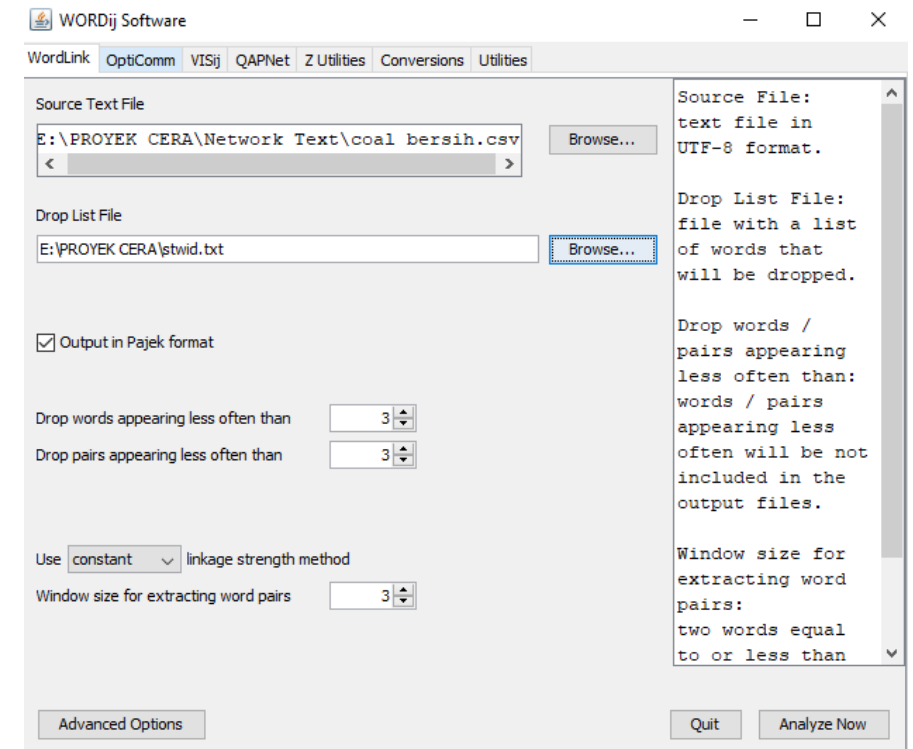
Data Preparation

- Prepare the data which consist textual content
- Place the textual content in one column
- Named the column with “text”, then save file as .csv
- *don't forget to preprocess the data












Data Preparation

- Open WORDij
- On the “Source Text File” click browse > choose the file
- On the “Drop List File” click browse > choose file stopwords
- Then, click “Analyze Now”

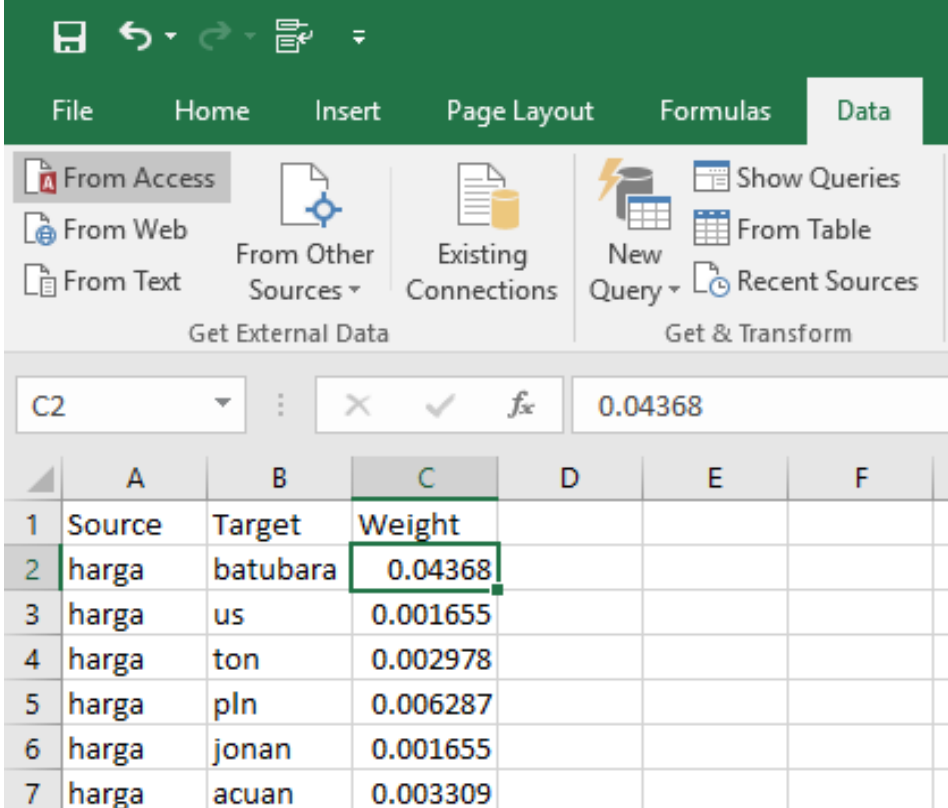


WORDij Output

 coal bersih.csv	20/08/2018 10:45	Microsoft Excel C...	28 KB
 coal bersih.csv.log	14/09/2018 07:32	Text Document	1 KB
 coal bersih.net	14/09/2018 07:32	NET File	8 KB
 coal bersih.pr	14/09/2018 07:32	PR File	22 KB
 coal bersih.ptg	14/09/2018 07:32	PTG File	22 KB
 coal bersih.stp.csv	14/09/2018 07:32	Microsoft Excel C...	24 KB
 coal bersih.stw.csv	14/09/2018 07:32	Microsoft Excel C...	8 KB
 coal bersih.wrd	14/09/2018 07:32	WRD File	8 KB
 coal bersih.wtg	14/09/2018 07:32	WTG File	12 KB

Data Preparation

- Create new excel file
- Place into column “pair” on the .stp file to the new excel file.
- Named the 1st column as “Source” and 2nd column as “Target”
- Place the data from “proportion” in .stp file to .xlsx which has column named source and target
- Place the data to the 3rd column and named its first row as “Weight”
- Save file as .CSV

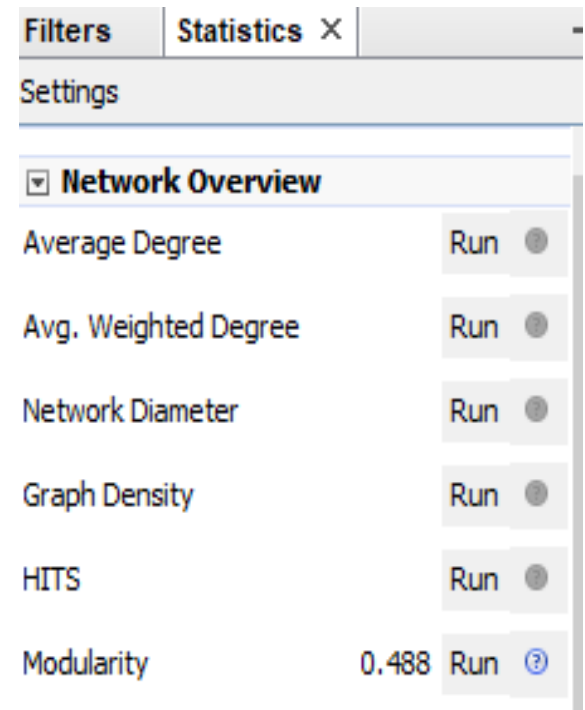
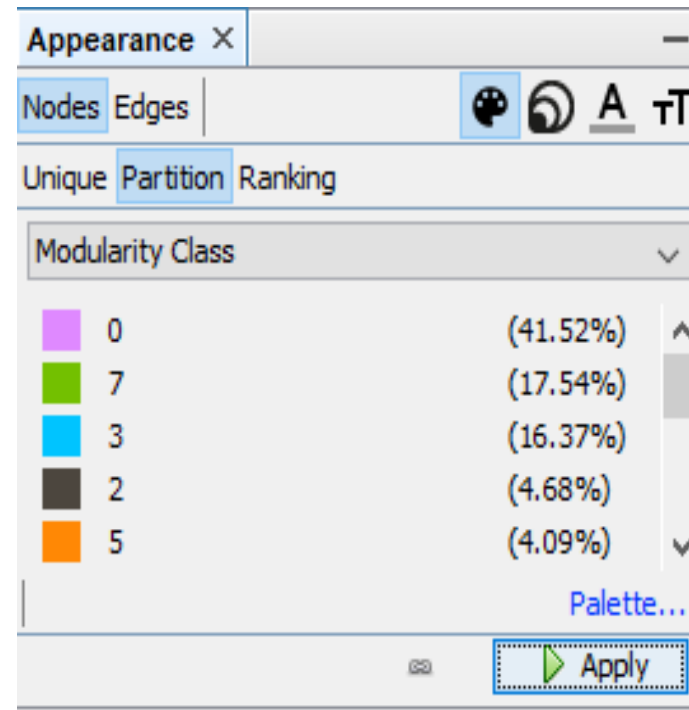


The screenshot shows the Microsoft Excel interface with the 'Data' tab selected. The ribbon includes options for 'Get External Data' (From Access, From Web, From Text, From Other Sources, Existing Connections) and 'Get & Transform' (Show Queries, From Table, New Query, Recent Sources). The formula bar shows 'C2' and the value '0.04368'. The table below has 7 rows and 6 columns (A-F). The first row is the header: Source, Target, Weight, D, E, F. The subsequent rows contain data: harga, batubara, 0.04368; harga, us, 0.001655; harga, ton, 0.002978; harga, pln, 0.006287; harga, jonan, 0.001655; and harga, acuan, 0.003309.

	A	B	C	D	E	F
1	Source	Target	Weight			
2	harga	batubara	0.04368			
3	harga	us	0.001655			
4	harga	ton	0.002978			
5	harga	pln	0.006287			
6	harga	jonan	0.001655			
7	harga	acuan	0.003309			

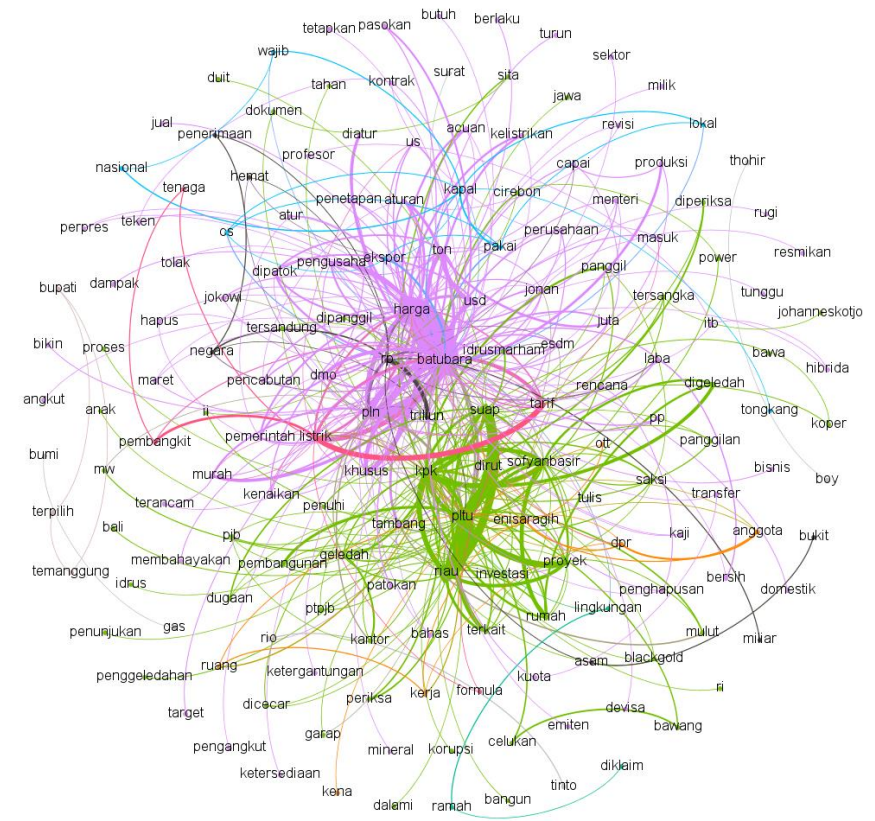
Analyze with Gephi

- On overview, choose layout and click run
- On statistics toolbar, run “modularity” > Click Ok > finish
- On the tab appearances toolbar, click “nodes” > “Partition” > “Modularity Class” > Apply
- The modularity network will appear as different color for each groups of network.



Visualization with Gephi

- Go to data laboratory, click “copy data to other column”
> choose “id” > copy to “label”.
- Go to preview
- Click “show labels”
- Modify the visualization by the font or the color in the left side.
- Click Refresh





THANK YOU