

# Visualisasi SNA

(Pelatihan data sains menggunakan R dan Gephi)

Ujang Fahmi

Pelajaran ke-12



Salam kenal dan selamat datang.

Semoga kita semua bisa saling berbagi pengalaman dan pengetahuan. Saya adalah Ujang Fahmi, Co-founder dan mentor Sadasa Academy.

Jika anda berada dan sedang membaca tutorial ini, maka kemungkinan anda adalah orang yang sedang ingin belajar data sains, atau mungkin ditugaskan untuk mempelajari R oleh institusi atau organisasi anda. Sama seperti saya dulu, dimana tanpa latar belakang engineering saya didiharuskan untuk belajar R, demi menyelesaikan tugas akhir dan akhirnya jadilah seperti saya sekarang ini.

Satu hal yang pasti, ini adalah langkah pertama dari banyak langkah yang harus dilalui, entah melalui lembaga resmi atau belajar secara mandiri. Jadi selamat belajar!!!

Ujang Fahmi,  
Yogyakarta, 2021-10-14

*Materi yang disampaikan disimpan dan dokumentasikan **disini***

# Persiapan

Materi sebelumnya sudah membahas tentang pre-processing untuk mempersiapkan data network (materi 10) serta mengukur centrality dan modularity (materi 9) dengan package `igraph`. Untuk di sini kita akan mencoba untuk:

1. Menggunakan data network yang sudah dibuat sebelumnya
2. Menambahkan centrality pada network
3. Menambahkan modularity pada network
4. Membuat visualisasi network

# Data Network

Untuk dapat menggunakan data network yang sudah dibuat sebelumnya, kita bisa mengimpornya terlebih dahulu menggunakan skrip berikut:

```
library(igraph)

data_net = read_graph(file = "data/tes_net.graphml", format = "graphml")
class(data_net)
gsize(data_net)

# tes 1
# plot(data_net, edge.arrow.size = 0.2)
```

## Menambahkan centrality dan modularity

Sebelum mengukur centrality dan modularity, untuk mempermudah proses visualisasi kita perlu mengekstrak nodes dan edges dari data network terlebih dahulu.

```
# extract nodes
nodes = data_frame(id = V(data_net)$id, nodes = V(data_net)$name)
nodes = nodes %>%
  filter(!duplicated(nodes))
glimpse(nodes)
```

```
# extract edges
edges = ends(data_net, E(data_net))
edges = as.data.frame(edges)
edges$V1 = as.character(edges$V1)
edges$V2 = as.character(edges$V2)
edges = edges %>%
  group_by(V1) %>%
  filter(!duplicated(V2))

glimpse(edges)
```

## Degree Centrality

```
degree.cent <- centr_degree(data_net, mode = "all")  
# setNames(degree.cent$res, V(data_net)$name)  
  
degree.cent = data_frame(degree.cent = degree.cent$res)  
degree.cent %>%  
  head(10)  
  
hasil = bind_cols(nodes, degree.cent)  
glimpse(hasil)
```

## Betweenness Centrality

```
betweenness.cent =  
  betweenness(data_net)  
betweenness.cent =  
  data_frame(betweenness_cent = betweenness.cent)  
betweenness.cent$betweenness_cent =  
  round(betweenness.cent$betweenness_cent, 2)  
hasil = bind_cols(hasil, betweenness.cent)
```

## Eigenvector Centrality

```
eigen_centrality =  
  eigen centrality(data_net)  
eigen_centrality =  
  data_frame(eigen centrality = eigen centrality[["vector"]])  
eigen_centrality$eigen centrality =  
  round(eigen centrality$eigen centrality, 2)  
hasil = bind_cols(hasil, eigen centrality)
```



# Modularity Centrality

```
wtc <- cluster_walktrap(data_net)
modularity(wtc)
modularity(data_net, membership(wtc))

member = data_frame(modularity_class = wtc$membership)
hasil = bind_cols(hasil, member)
```

# Visualisasi SNA

```
library(igraph)

##
## Attaching package: 'igraph'

## The following objects are masked from 'package:stats':
##
##      decompose, spectrum

## The following object is masked from 'package:base':
##
##      union

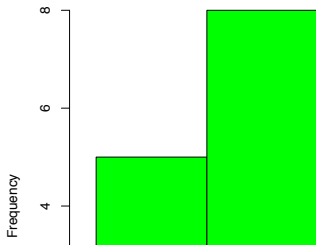
nodes <- read.csv("https://kateto.net/workshops/data/Dataset1-Media-Exa
links <- read.csv("https://kateto.net/workshops/data/Dataset1-Media-Exa
# head(nodes)
# head(links)
net <- graph_from_data_frame(d=links, vertices=nodes, directed=T)
```

## Distribusi Degree

```
V(net)$label <- V(net)$media  
V(net)$degree <- degree(net)
```

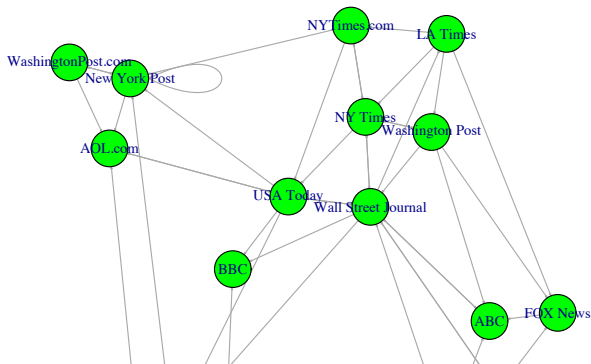
```
hist(V(net)$degree,  
     col = 'green',  
     main = 'Histogram of Node Degree',  
     ylab = 'Frequency',  
     xlab = 'Degree of Vertices')
```

Histogram of Node Degree



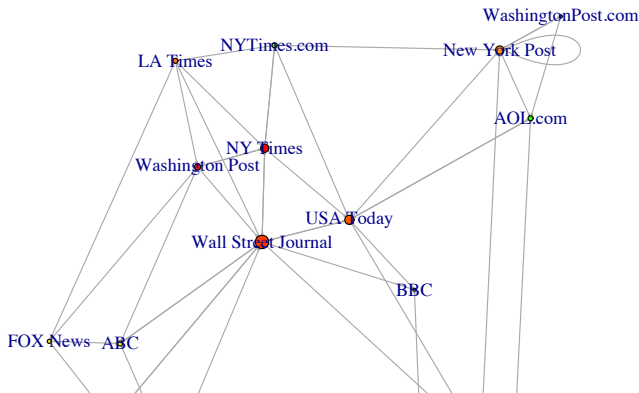
# Network diagram

```
set.seed(222)
plot(net,
      vertex.color = 'green',
      vertex.size = 2,
      edge.arrow.size = 0.1,
      vertex.label.cex = 0.8)
```



# Highlighting degrees & layouts

```
plot(net,  
  vertex.color = rainbow(52),  
  vertex.size = V(net)$degree*0.4,  
  edge.arrow.size = 0.1,  
  layout=layout.fruchterman.reingold)
```



# Hub and Authorities

Basically, the hub has many outgoing links, and Authorities have many incoming links.

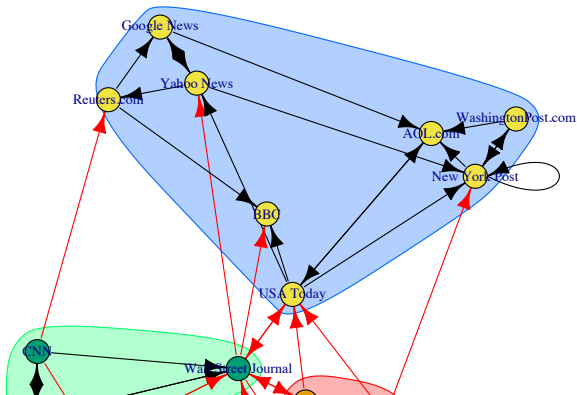
```
hs <- hub_score(net)$vector
as <- authority_score(net)$vector
set.seed(123)
plot(net,
      vertex.size=hs*30,
      main = 'Hubs',
      vertex.color = rainbow(52),
      edge.arrow.size=0.1,
      layout = layout.kamada.kawai)
```

**Hubs**



# Community Detection

```
cnet <- cluster_edge_betweenness(net)
plot(cnet,
     net,
     vertex.size = 10,
     vertex.label.cex = 0.8)
```



# Analisis dengan gephi

Untuk bisa melakukan analisis network di Gephi kita perlu tahu dulu apa persyaratan yang diperlukan: Dalam konteks ini file seperti apa yang dibutuhkan di gephi atau perangkat lunak sejenis.



# Table of Contents

## Persiapan

- Data Network

- Menambahkan centrality dan modularity

## Visualisasi SNA

- Distribusi Degree

- Network diagram

- Highlighting degrees & layouts

- Hub and Authorities

- Community Detection

## Analisis dengan gephi