

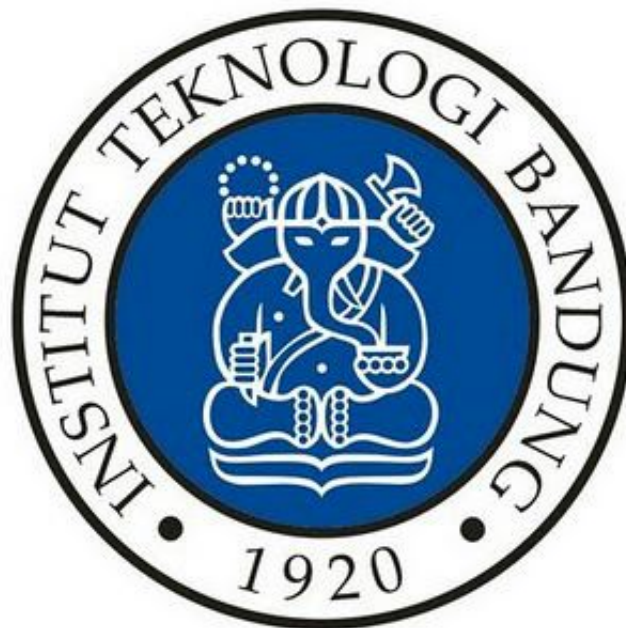
IF5170 Visualisasi Data

Taksonomi Tools Visualisasi Data

Adi Ferdian (23514030)

Hannif Kurniawan (23514072)

Apip Ramdlani (23514094)



**PROGRAM STUDI MAGISTER INFORMATIKA
SEKOLAH TEKNIK ELEKTRO DAN INFORMATIKA
INSTITUT TEKNOLOGI BANDUNG
2015**

Daftar Isi

[Daftar Isi](#)

[Deskripsi](#)

[Data dan Visualisasi](#)

[Word cloud](#)

[Network](#)

[TreeMap](#)

[Map](#)

[Bubble Hierarchy](#)

[Bubble Plot](#)

[Saran dan Perbaikan](#)

[Referensi](#)

1. Deskripsi

Visualisasi data riil menggunakan R, Weave atau tools lain yang mendukung dalam visualisasi data. Jenis visualisasi data yang digunakan adalah sebagai berikut.

- Word cloud
- Network
- TreeMap
- Map
- Bubble Hierarchy
- Bubble Plot

2. Data dan Visualisasi

2.1. Word cloud

2.1.1. Data

Data yang digunakan dalam visualisasi data dengan metode Word cloud adalah data Aspirasi dan Pengaduan Masyarakat Nasional. Data aspirasi dan pengaduan masyarakat nasional terkait program pembangunan nasional dan pelayanan publik pemerintah yang diperoleh melalui sistem LAPOR!. Mencakup aspirasi dan pengaduan terhadap 81 Kementerian/Lembaga, 5 Pemerintah Daerah, serta 44 BUMN di Indonesia.

2.1.2. Visualisasi Data

Visualisasi data dilakukan dengan R dengan menggunakan packages textmining dan wordcloud. Visualisasi dilakukan terhadap topik dari data laporan dengan tujuan melihat frekuensi kemunculan topik pelaporan. Berikut adalah proses yang dilakukan dengan bahasa pemrograman R.

```
1. //wordcloud.R
2. library(wordcloud);
3. library(tm);
4. setwd("/Users/adiferdian/Documents/Studi/Semester 3/Visualisasi
   Data/Tugas/visdat15-2");
5. streamlaporan <- read.csv('Data/streamlaporan.csv');
6. laporan <-streamlaporan$Kategori;
```

```

7. mycorpus <- Corpus(VectorSource(laporan));
8. #mycorpus <- tm_map(mycorpus, stripWhitespace);
9. mycorpus <- tm_map(mycorpus, removePunctuation)
10. mycorpus <- tm_map(mycorpus, removeNumbers)
11. mycorpus <- tm_map(mycorpus, tolower);
12. mycorpus <- tm_map(mycorpus, PlainTextDocument);
13. stopwordsID <- "stopwords.txt";
14. cStopwordID <- readLines(stopwordsID);
15. mycorpus <-
    tm_map(mycorpus, removeWords, c(stopwords("english"), cStopwordID));
16. wordcloud(mycorpus, scale=c(3,0.2), min.freq=4, max.words=300,
    random.order=FALSE, rot.per=0.35, use.r.layout=FALSE,
    colors=brewer.pal(8, "Dark2"));

```

Berikut adalah hasil visualisasi data yang diperoleh.



Dari hasil visualisasi yang diperoleh dapat terlihat bahwa topik yang paling sering muncul dari data Aspirasi dan Laporan Masyarakat Nasional adalah kelola, birokrasi, reformasi dan tata.

2.2. Network

2.2.1. Data

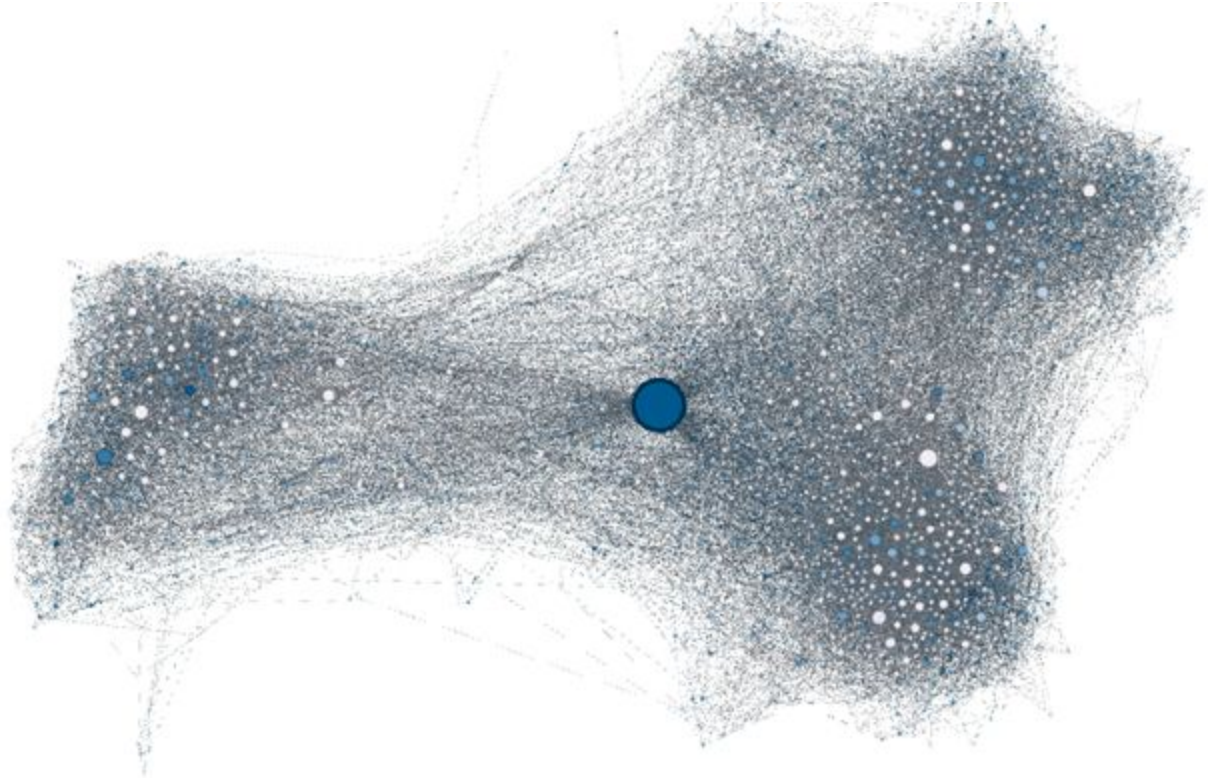
Data yang digunakan dalam visualisasi data dengan metode Network adalah data yang diambil dari API twitter tentang keterhubungan sebuah akun terhadap followers-nya secara random. Data terdiri atas dua set data yaitu nodes yang berisi identitas masing-masing akun serta edges yang berisi data akun follow akun yang lainnya.

2.2.2. Visualisasi Data

Visualisasi data dilakukan dengan tools Gephi yang mempunyai kemampuan untuk membuat visualisasi network hierarchy. Visualisasi dilakukan untuk mendapatkan keterkaitan antara akun dan akun mana yang yang paling berpengaruh. Langkah-langkahnya adalah

1. Install gephi
2. Import data nodes
3. Import data edges
4. Rubah ukuran nodes pada menu rangking dengan ukuran (10-150)
5. Pilih metode Fruchterman Reingold pada menu layout lalu rubah nilainya menjadi (10000;10;10), klik run
6. Setelah stabil pilih metode Force Atlas 2 lalu ceklis "prevent overlap" dan rubah nilai "scalling" menjadi 10, klik run
7. Setelah stabil, pada panel rangking rubah "sex" atribut lalu pilih warna yang diinginkan.
8. Pada panel label, check nodes

Berikut adalah hasil visualisai data yang diperoleh.



Dari hasil visualisasi yang diperoleh dapat terlihat bahwa ada satu akun yang menjadi poros, kemudian terbagi menjadi tiga kelompok akun besar yang saling berhubungan.

2.3. TreeMap

2.3.1. Data

Data yang digunakan dalam visualisasi data dengan metode Word cloud adalah data Aspirasi dan Pengaduan Masyarakat Nasional. Data aspirasi dan pengaduan masyarakat nasional terkait program pembangunan nasional dan pelayanan publik pemerintah yang diperoleh melalui sistem LAPOR!. Mencakup aspirasi dan pengaduan terhadap 81 Kementerian/Lembaga, 5 Pemerintah Daerah, serta 44 BUMN di Indonesia.

2.3.2. Visualisasi Data

Visualisasi data dilakukan dengan R dengan menggunakan packages portfolio dan treemap. Visualisasi dilakukan terhadap topik, area dan kategori dari data

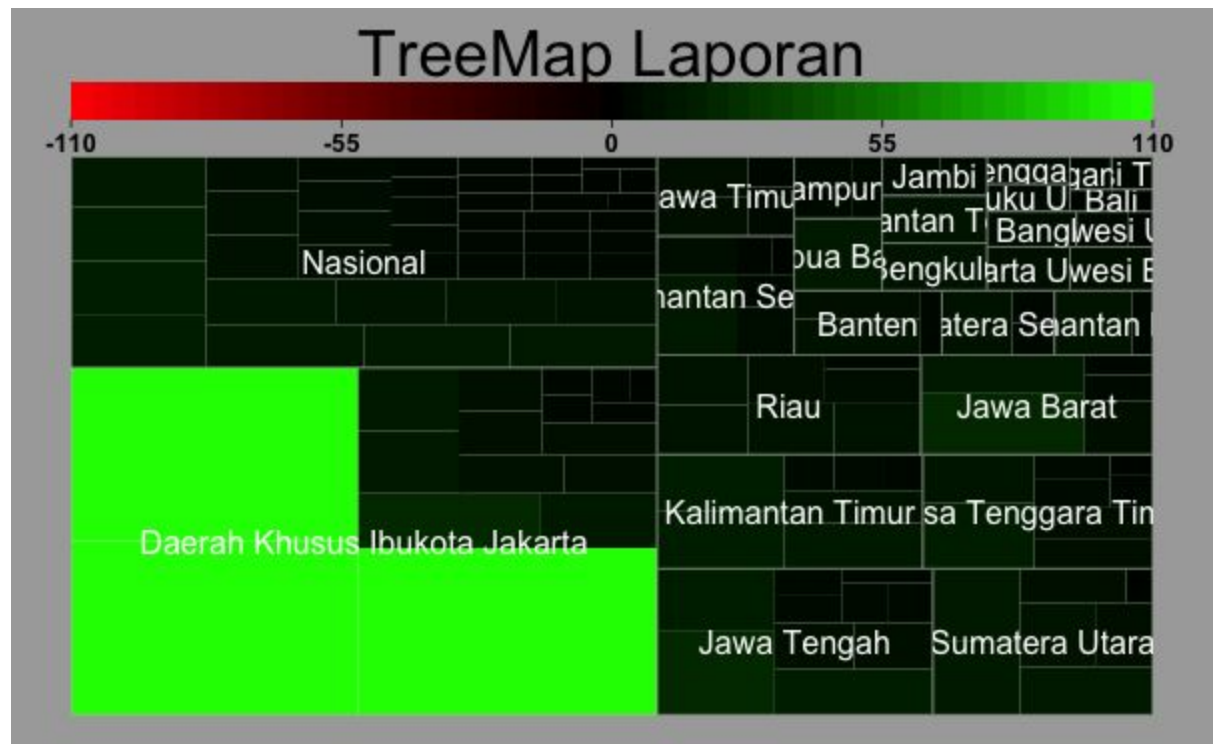
laporan dengan tujuan melihat pemetaan data. Berikut adalah proses yang dilakukan dengan bahasa pemrograman R.

```

1. //TreeMap.R
2. setwd("/Users/adiferdian/Documents/Studi/Semester 3/Visualisasi
   Data/Tugas/visdat15-2");
3. library(RColorBrewer);
4. library(portfolio);
5. library(rjson);
6. source("treemapbrewer.R");
7. streamlaporan <- read.csv('Data/streamlaporan.csv');
8. map.market(id      = streamlaporan$TrackingID,
9.           area     = streamlaporan$KategoriID,
10.          group    = streamlaporan$Area,
11.          color     = streamlaporan$KategoriID,
12.          lab       = c(TRUE, FALSE),
13.          main      = "TreeMap Laporan",
14.          );

```

Berikut adalah hasil visualisai data yang diperoleh.



Dari hasil visualisasi yang diperoleh dapat terlihat bahwa pelaporan dan aspirasi yang paling sering muncul adalah data dengan yang berada dari DKI Jakarta dan Nasional.

2.4. Map

2.4.1. Data

Data yang digunakan dalam visualisasi data dengan metode Map adalah data yang ada pada Open Flight. Data tersebut merupakan data lokasi bandara yang ada di eropa beserta jumlah keberangkatan dan pendaratannya dari masing-masing bandara. Data terdiri atas nama bandara, kota, negara tersebut berasal, dan lain-lain.

2.4.2. Visualisai Data

Visualisasi data dilakukan dengan R dengan menggunakan packages rworldmap dan plyr. Visualisasi dilakukan terhadap topik dari data laporan dengan tujuan melihat besarnya keberangkatan dan pendaratan dari tiap lokasi bandara. Dengan menggunakan API Gmaps, dapat dihasilkan peta yang sesuai dengan peta yang ada pada Gmaps. Berikut adalah proses yang dilakukan dengan bahasa pemrograman R

```
1. airports <- read.csv("airports.csv")
2. colnames(airports) <- c("ID", "name", "city", "country",
  "IATA_FAA", "ICAO", "lat", "lon", "altitude", "timezone", "DST")
3. library(rworldmap)
4. newmap <- getMap(resolution = "low")
5. plot(newmap, xlim = c(-20, 59), ylim = c(35, 71), asp = 1)
6. points(airports$lon, airports$lat, col = "red", cex = .6)
7. routes <- read.csv("routes.csv")
8. colnames(routes) <- c("airline", "airlineID", "sourceAirport",
  "sourceAirportID", "destinationAirport", "destinationAirportID",
  "codeshare", "stops", "equipment")
9. library(plyr)
10. departures <- ddply(routes, .(sourceAirportID), "nrow")
11. names(departures)[2] <- "flights"
12. arrivals <- ddply(routes, .(destinationAirportID), "nrow")
13. names(arrivals)[2] <- "flights"
```



```

14.airportD <- merge(airports, departures, by.x = "ID", by.y =
    "sourceAirportID")
15.airportA <- merge(airports, arrivals, by.x = "ID", by.y =
    "destinationAirportID")
16.
17.library(ggmap)
18.map <- get_map(location = 'Europe', zoom = 4)
19.mapPoints <- ggmap(map) +
20. +   geom_point(aes(x = lon, y = lat, size = sqrt(flights)), data =
    airportD, alpha = .5)
21.
22.mapPointsLegend <- mapPoints +
23. +   scale_area(breaks = sqrt(c(1, 5, 10, 50, 100, 500)), labels =
    c(1, 5, 10, 50, 100, 500), name = "departing routes")
24.mapPointsLegend

```

Berikut adalah hasil visualisai data yang diperoleh.



Dari hasil visualisasi yang diperoleh dapat terlihat bahwa pemberangkatan yang paling sering terjadi dari wilayah Jerman dan Perancis.

2.5. Bubble Hierarchy

2.5.1. Data

Data yang digunakan pada visualisasi adalah data produksi padi pada tahun 2013 yang dibagi berdasarkan jumlah produksi perprovinsi. Format data yang digunakan adalah json dengan contoh struktur sebagai berikut.

```
var data = {
  label: 'Total',
  amount: 1000,
  color: '#0066bb',
  children: [
```


2.6. Bubble Plot

2.6.1. Data

Data yang digunakan pada visualisasi Bubble Plot adalah data Jumlah Pemegang Kartu Askes di Kabupaten Bojonegoro serta data Rekapitulasi Penggunaan Kartu Askes di Kabupaten Bojonegoro.

2.6.2. Visualisasi Data

Visualisasi data dilakukan dengan R dengan menggunakan fungsi `symbols`. Visualisasi dilakukan untuk melihat perbandingan jumlah pemilik dan pengguna kartu askes pertahun. Berikut adalah proses yang dilakukan dengan bahasa pemrograman R.

```
1. //BubblePlot.R
2. rm(list = ls());
3. setwd("/Users/adiferdian/Documents/Studi/Semester 3/Visualisasi
   Data/Tugas/visdat15-2");
4. pemegang <- read.csv("~/Documents/Studi/Semester 3/Visualisasi
   Data/Tugas/visdat15-2/Data/jumlahpemegangkartuaskesdikabupatenbojo
   negoro.csv", sep=";")
5. pengguna <- read.csv("~/Documents/Studi/Semester 3/Visualisasi
   Data/Tugas/visdat15-2/Data/penggunaankartuaskesdikabupatenbojonego
   ro.csv", sep=";")
6. library(plyr);
7. pemegang$pemegang_kaartu <- NULL;
8. pengguna$Penggunaan <- NULL;
9. p1 <- ddply(pemegang, .(Tahun), summarise,
   jumPemegang=sum(Jumlah));
10. p2 <- ddply(pengguna, .(Tahun), summarise,
   jumPengguna=sum(Jumlah));
11. data <- merge(p1,p2,by="Tahun");
12. symbols(data$Tahun,
13.         data$jumPemegang,
14.         circles=data$jumPengguna,
15.         inches = 0.35,
16.         fg=data$Tahun,
17.         bg=data$Tahun,
18.         xlab="Tahun",
19.         ylab="Jumlah Pemegang",
```

```
20.         main = "  
21.         Jumlah Penggunaan Askes Kabupaten Bojonegoro ");
```

Berikut adalah hasil visualisai data yang diperoleh.



Dari visualisasi yang diperoleh terlihat bahwa populasi pengguna kartu askes yang paling besar adalah pada tahun 2013 di saat jumlah pemegangnya berbeda jauh pada tahun 2012.

3. Kesimpulan

Ada beberapa kesimpulan yang diperoleh dari percobaan yang telah dilakukan, yakni.

1. Pemilihan metode visualisasi data yang akan digunakan sangat dekat kaitannya dengan tujuan visualisasi serta data yang digunakan.
2. Preprocessing data merupakan proses yang penting untuk mendapatkan hasil visualisasi yang tepat

4. Referensi

<http://data.go.id/dataset/data-aspirasi-dan-pengaduan-masyarakat>

<http://data.go.id/dataset/jumlah-pemegang-kartu-askes-di-kabupaten-bojonegoro>

<http://data.go.id/dataset/penggunaan-kartu-askes-di-kabupaten-bojonegoro>

<http://www.bps.go.id/linkTabelStatis/view/id/1163>

<http://openflights.org/data.html>

<http://www.martingrandjean.ch>

<http://d3js.org/>