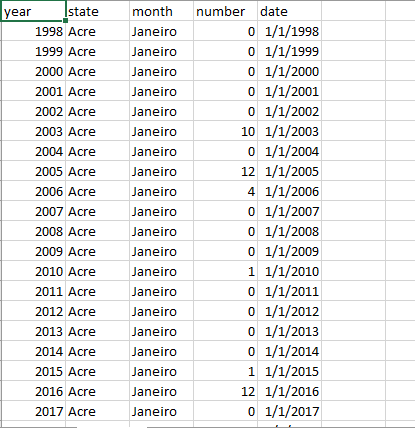
***Forest Fires in Brazil***

***Dataset :***

**

***Penjelasan :***

Laporan dataset ini berisi data tentang informasi kebakaran hutan di Brazil yang dibagi menjadi beberapa bagian. Dataset ini memuat informasi kebakaran pada periode awal sekitar 20 tahun lalu (1998 hingga 2017).

***Tujuan :***

Pada dataset ini kita ingin mengelola informasi yang lebih spesifik tentang mengenai kebakaran hutan di brasil yaitu :

* Menentukan peningkatan/pengurangan kebakaran pada tiap tahun
* Menentukan daerah yang paling rawan kebakaran

***Jumlah Attribute :***

Ada 5 yaitu :

* Year
* State
* Month
* Number
* Date

***Jumlah Dataset :***

Hanya ada 1 yaitu :

* amazon.csv

***Missing Value*** : -

***Link Github*** : Buat akun github kelompk dan pribadi :v

<https://github.com/TotoSugiarto/Tugas-Besar_067_063_048_074>

* 3311811067 : <https://github.com/TotoSugiarto>
* 3311811063 : //github.com/AndhikaPratama-PX1
* 3311811074 ://github.com/dimaspanjiperdana
* 3311811048 ://github.com/ErvylArwianda

***Metode dengan alasan :***

Metode Clustering alasannya ialah, untuk mempermudah membandingkan peningkatan kebakaran dengan melalukan pengelompokan.

***Pengolahan Data :***

* Menentukan peningkatan/pengurangan kebakaran pada tiap tahun

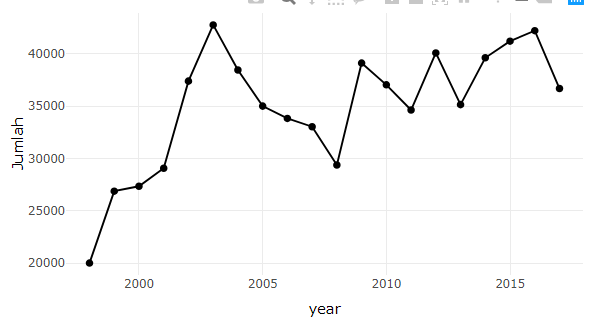
*Source Code :*

df <- read.csv("datamining.csv", sep = ",")

by\_years<- df %>% group\_by(year)

g <- ggplot(data=summarise(.data=by\_years, Jumlah = sum(number)),aes(x=year, y = Jumlah )) + geom\_line() + geom\_point() + theme\_minimal()

ggplotly(g,tooltip="Fire Cases Reported")



Dapat dilihat dari statistik bahwa tragedy kebakaran pada hutan amazon pertahun mengalami peningkatan lebih besar jika dibandingkan dengan pengurangan.

* Menentukan daerah yang paling rawan kebakaran

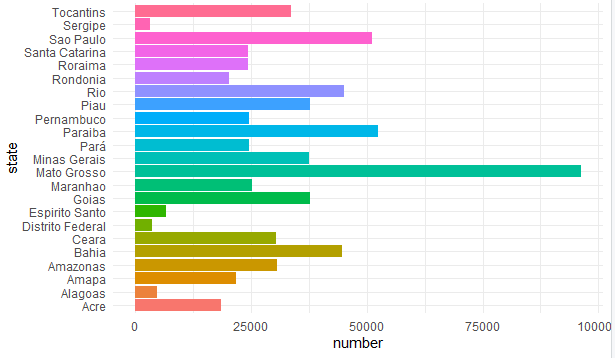
*Source Code :*

df <- read.csv("datamining.csv", sep = ",")

by\_states <- df %>% group\_by(state)

ggplot(data=by\_states,aes(x=state,y=number)) + geom\_col(aes(fill=state)) + coord\_flip(

) + theme\_minimal() + theme(legend.position = "none")



Daerah yang paling rawan kebakaran ialah daerah bagian Mato Grosso

* Clustering

Source Code:

data.numerik<-by\_states[2:4]

data.stds<-scale(data.numerik)

fviz\_nbclust(data.stds, kmeans, method = "silhouette")

Clustering=kmeans(data.stds,centers=3,nstart=25)

fviz\_cluster(Clustering, geom = "point", data = data.stds)+ggtitle("k=3")

final=data.frame(df, Clustering$cluster)

View(final)

