**Tugas Besar Data Mining Kelompok 10**

**videogamesales**

**THEOFILUS.S (3311901070)**

**Muhammad Putra Caesar Zellya (3311901084)**

buat pengaturan lokasi directory yang foldernya dibuat adalah **data** dan nama file excel sesuai dengan nama nya

nama file yang dibuat adalah

buatlah variable **lokasi\_kerja** sesuaikan nilainya dengan nama file excel

**lokasi\_kerja <- "D:/data"**

**getwd()**

perintah untuk pengaturan lokasi directory

**setwd(lokasi\_kerja)**

**getwd()**

perintah untuk Jalankan perintah untuk mengatur working directory

**dataset <- read.csv("videogamesales.csv", sep = ";")**

**head(dataset)**

Gunakan perintah library yang sudah didownload  
**library(arules)**

**library(arulesViz)**

**library(mlbench)**

**library(C50)**

**library(printr)**

Ketika belum ada maka menggunakan perintah **install.packages()** contoh **install.packages(arules)** dan seterusnya Ketika sudah maka tinggal di library

Import dataset yang sudah dibuat

**dataset <- read.transactions("videogamesales.csv", sep = ",")**

**inspect(dataset)**

Perintah rules

**rules<- apriori(dataset, parameter = list(support = 0.2, confidence = 0.6))**

**inspect(rules)**

pertama run dulu ulang perintah **dataset <- read.csv("videogamesales.csv", sep = ";")** dan **head(dataset)**

Dan gunakan perintah

**arulesViz::plotly\_arules(rules)**

Untuk melihat berbagai korelasi antara Na\_Sales dan Other\_sales di dataset videogame sales maka menggunakan

**cor(dataset$NA\_Sales, dataset$Other\_Sales)**

atau

**cor(dataset$NA\_Sales, dataset$Year)**

**cor(dataset$Other\_Sales, dataset$Year)**

membuat dataset baru antara NA\_Sales dan Other\_Sales

**databaru.NA\_Sales <-data.frame( NA\_Sales = c(1))**

**databaru.Other\_Sales <-data.frame( Other\_Sales = c(12))**

Gunakan perintah untuk kelas dataset untuk NA\_Sales dan Other\_Sales seperti berikut

**class(dataset)**

**class(dataset$NA\_Sales)**

**class(dataset$Other\_Sales)**

untuk merubah data kedalam faktor

dan ketik

**dataset$NA\_Sales <- as.factor(dataset$NA\_Sales)**

**dataset$Other\_Sales <- as.factor(dataset$Other\_Sales)**

untuk ingin melihat tipe datanya yang sudah datanya dalam factor  
Setelah itu baru mengetik peritah yang diatas dengan perintah

**model <- C5.0(NA\_Sales ~., data=dataset)**

dan untuk melihat model

**model**

**summary(model)**

atau bisa saja secara berulang seperti

**model <- C5.0(Other\_Sales ~., data=dataset)**

**model**

**summary(model)**

ketika data sudah dibuat model maka tampilkan pohon modelnya atau plot dari data tersebut.  
ketika ingin menampilkan pohon model maka mengetik

**plot(model)**

ketika ingin menjadikan dataset menjadi data testing maka sebagai kolom saja dan tanpa label

**datatesting <- dataset[,1:4]**

ketika ingin membuat prediksi data set maka menggunakan perintah

**predictions <- predict(model, datatesting)**

Untuk membandingkan hasil prediksi dengan dataset NA\_Sales dan Order\_Sales maka menggunakan perintah

**table(predictions, dataset$NA\_Sales)**

**table(predictions, dataset$Other\_Sales)**