

Examen Parcial 1

ByteMiners

23/02/2020

Análisis exploratorio de datos

1. Exploración y visualización de variables

1.1 Explorar la base de datos, realizar los cambios necesarios y convenientes para poder trabajar con ella.

```
#Se lee la base de datos
setwd("/home/jose/Escritorio/Semestre 6/Mineria de datos/portafolioJairo/parcial1/examen")
df <- read.csv("Ejercicio 1.csv", sep = ",")

#Trasponer las columnas del df para que las características de
#cada observacion estén ordenadas por columnas
df_transposed <- as.data.frame(t(as.matrix(df)))

# Eliminar observaciones con valores NA
df_transposed_omitted<-na.omit(df_transposed)

# Colocar nombre a las columnas y eliminar la primera fila
names(df_transposed_omitted) <- c("Peso","Estatura",
                                "Edad","Complexión",
                                "Papás_separados","Cantidad_hermanos",
                                "Trabaja", "Horas_trab_dia",
                                "Horas_trab_mes", "Salario_mes")

updated_data <- df_transposed_omitted[-c(1),]
```

2. Metodo de reducción de dimensión (PCA)

2.1 PCA base de datos PCA1

2.2 PCA base de datos PCA2

```
url<-"http://web.stanford.edu/~hastie/ElemStatLearn//datasets/zip.digits/train.2"
data<-read.csv(url)
```

4. Metodo de reducción de dimensión (CCA)