

```
#observations:

#Fields always present: ICULOS, Age, Gender, HospAdmTime, SepsisLabel
#Fields with < 15% missing values: HR, O2Sat, SBP, MAP, DBP
#Fields with 20% to 90% missing: Temp, Resp, Glucose, Unit1, Unit2
#All other fields have > 90% missing values.
#The median patient percent missing is 100% for 16 of the quantities.
#Demographics
#Only about 1.4% of raw records indicate sepsis. [2623 / 188,453]
#Only about 5.6% of the patients have sepsis. [279 / 5000]
#Sepsis reported in first hour for 20.1% of sepsis patients. [56 / 279]
#original data

orgdata <- read.csv("Dataset.csv")
summary(is.na(orgdata))
sum(is.na(orgdata))

install.packages(c("ggplot2", "reshape2"))

# Load libraries
library(ggplot2)
library(reshape2)

orgdatamatrix <- as.matrix(orgdata)
orgdatamatrix1 <- as.matrix(data)
data <- read.csv("data.csv")

heatmap(orgdatamatrix, scale = "none", Rowv = NA, Colv = NA, cexRow = 0.5, cexCol = 0.5)
heatmap(orgdatamatrix1, scale = "row", Rowv = NA, Colv = NA, cexRow = 0.7, cexCol = 1)

na_counts <- list()

for (col in names(orgdata)) {
  na_count <- sum(is.na(orgdata[[col]]))

  na_counts[[col]] <- na_count
}
print(na_counts)
```