

# PROC COMPARE in R

Hamza Oukassou

2023-09-08

```
# Load necessary packages
library(dplyr)
library(arsenal)

# Create the first dataset with 3 observations
dataset1 <- data.frame(
  Patient_ID = 1:3,
  Treatment = c("Drug A", "Drug B", "Drug A"),
  Age = c(45, 55, 40),
  Aval = c(78.5, 88.7, 92.3)
)

# Create the second dataset with 3 observations
dataset2 <- data.frame(
  Patient_ID = 1:3,
  Treatment = c("Drug A", "Drug B", "Drug A"),
  Age = c(45, 55, 38), # Differences in Age
  Aval = c(78.5, 91.2, 92.3) # Differences in Aval (Efficacy Score)
)

#To compare these datasets, simply pass them to the comparedf() function
#Use summary() to get a more detailed summary

compare01 <- summary(comparedf(dataset1, dataset2))
print(compare01)

##
##
## Table: Summary of data.frames
##
## version  arg          ncol  nrow
## -----  -
## x        dataset1      4     3
## y        dataset2      4     3
##
##
##
## Table: Summary of overall comparison
##
## statistic                                     value
## -----
## Number of by-variables                               0
```

```

## Number of non-by variables in common      4
## Number of variables compared              4
## Number of variables in x but not y        0
## Number of variables in y but not x        0
## Number of variables compared with some values unequal  2
## Number of variables compared with all values equal  2
## Number of observations in common          3
## Number of observations in x but not y     0
## Number of observations in y but not x     0
## Number of observations with some compared variables unequal  2
## Number of observations with all compared variables equal  1
## Number of values unequal                  2
##
##
##
## Table: Variables not shared
##
##
## -----
## No variables not shared
## -----
##
##
##
## Table: Other variables not compared
##
##
## -----
## No other variables not compared
## -----
##
##
##
## Table: Observations not shared
##
##
## -----
## No observations not shared
## -----
##
##
##
## Table: Differences detected by variable
##
## var.x      var.y      n    NAs
## -----  -----  ---  ----
## Patient_ID Patient_ID    0    0
## Treatment  Treatment    0    0
## Age        Age         1    0
## Aval       Aval         1    0

```

```

##
##
##
## Table: Differences detected
##
## var.x    var.y    ..row.names.. values.x  values.y  row.x  row.y
## -----  -----  -----  -----  -----  -----  -----
## Age     Age         3  40      38      3      3
## Aval    Aval        2  88.7    91.2    2      2
##
##
##
## Table: Non-identical attributes
##
##
## -----
## No non-identical attributes
## -----

# Capture the summary text by using capture.output() function
summary_compare <- capture.output(print(compare01))

# Specify the file path where you want to save the summary
file_path <- "summary_compare.txt"

# Save the summary compare to a text file
writeLines(summary_compare, file_path)

```