

Dataset :

ANZDATA is the Australia and New Zealand Dialysis and Transplant Registry. The Registry records the incidence, prevalence and outcome of dialysis and transplant treatment for patients with end stage renal failure. ANZDATA is located at The Royal Adelaide Hospital, South Australia.

- 1) <https://www.anzdata.org.au/>
- 2) I will upload later when the assignment upload

```
library(tidyverse)
# set your working directory to the same location as the data.
# setwd('/dataLocation/')

patientData <- read.csv("42264_AnzdataPatients.csv")
transplantData <- read.csv("42264_AnzdataTransplants.csv")
creatinineData <- read.csv("42264_AnzdataTransplantSerumCreatinine.csv")

# For simplicity let's ignore patient's multiple transplants. What does the !
# do?
transplantData2 <- transplantData[!duplicated(transplantData$id), ]

# Join patientData and transplantData
Data <- full_join(patientData, transplantData2, by = "id")

# Extract some interesting variables
use <- c("gendercode", "laterreferralcode", "creatinineatentry", "height", "weight",
        "smokingcode", "cancerever", "chroniclungcode", "coronaryarterycode", "peripheralvascularcode",
        "cerebrovascularcode", "diabetescode", "graftno", "transplantcentrestate", "recipientantibodycmvc
ode",
        "recipientantibodyebvcode", "donorsourcecode", "donorage", "donorgendercode",
        "ischaemia", "ageattransplant", "h1amismatchesa", "h1amismatchesb", "h1amismatchesdr",
        "h1amismatchesdq", "maxcytotoxicantibodies", "currentcytotoxicantibodies", "timeondialysis",
        "transplantstatus")

Data <- Data[, use]

# Recode some NA
Data[Data == ""] = NA
Data[Data == "-"] = NA
```