Project tutoré 2 - Murat Simsek et Yuquan Dai

# R pour la question 1

# Importer les données nécessaires  
dm1 <- read.csv2("dm1.csv", header = TRUE, sep = ",")  
dm2 <- read.csv2("dm2.csv", header = TRUE, sep = ",")  
ie1 <- read.csv2("ie1.csv", header = TRUE, sep = ",")  
ie2 <- read.csv2("ie2.csv", header = TRUE, sep = ",")  
lb1 <- read.csv2("lb1.csv", header = TRUE, sep = ",")  
lb2 <- read.csv2("lb2.csv", header = TRUE, sep = ",")  
  
# Jointure les données  
alldm <- full\_join(dm1,dm2)

## Joining, by = c("STUDYID", "DOMAIN", "USUBJID", "EPOCH", "VISIT", "VISITNUM",  
## "RFSTDTC", "RFENDTC", "SITEID", "BRTHDTC", "AGE", "AGEU", "SEX", "RACE",  
## "ETHNIC", "ARMCD", "ARM", "COUNTRY", "DMDTC", "DMDY")

allie <- full\_join(ie1,ie2)

## Joining, by = c("STUDYID", "DOMAIN", "USUBJID", "EPOCH", "IESEQ", "IETESTCD",  
## "IETEST", "IECAT", "IEORRES", "IESTRESC", "VISIT", "VISITNUM", "IEDTC", "IEDY")

alllb <- full\_join(lb1,lb2)

## Joining, by = c("STUDYID", "DOMAIN", "USUBJID", "EPOCH", "LBSEQ", "LBTESTCD",  
## "LBTEST", "LBCAT", "LBORRES", "LBORRESU", "LBORNRLO", "LBORNRHI", "LBSTRESC",  
## "LBSTRESN", "LBSTRESU", "LBSTNRLO", "LBSTNRHI", "LBNRIND", "LBSTAT",  
## "LBREASND", "LBSPEC", "LBMETHOD", "LBBLFL", "VISIT", "VISITNUM", "LBDTC",  
## "LBDY")

# On a choisit dm pour travailler  
n\_distinct(alldm$USUBJID)

## [1] 411

# Faire une exclusion spéciale  
allusrdm <- alldm$USUBJID[!(alldm$USUBJID %in% allie$USUBJID[allie$IECAT=="EXCLUSION"])]  
summary(allusrdm)

## Length Class Mode   
## 369 character character

# Séparer BUPRENORPHINE/NALOXONE et CLONIDINE  
allusrBN <- filter(alldm,   
 alldm$ARM=="BUPRENORPHINE/NALOXONE" &  
 alldm$USUBJID %in% allusrdm)  
allusrC <- filter(alldm,   
 alldm$ARM=="CLONIDINE" &   
 alldm$USUBJID %in% allusrdm)  
n\_distinct(allusrBN$USUBJID)

## [1] 233

n\_distinct(allusrC$USUBJID)

## [1] 110

# On considérera les patients qui ont des résultats au tests urinaires à J13 ou J14  
usrBNfin <- filter(alllb,   
 (alllb$VISITNUM==13 | alllb$VISITNUM==14)&  
 (alllb$LBSTAT!="NOT DONE")&  
 (alllb$USUBJID %in% allusrBN$USUBJID))  
usrCfin <- filter(alllb,   
 (alllb$VISITNUM==13 | alllb$VISITNUM==14)&  
 (alllb$LBSTAT!="NOT DONE")&  
 (alllb$USUBJID %in% allusrC$USUBJID))  
n\_distinct(usrBNfin$USUBJID)

## [1] 150

n\_distinct(usrCfin$USUBJID)

## [1] 28



# Python pour la question 2