```r
library(readr) #importing csv files
library(dplyr) #general analysis
library(tidyr)


readmission <- read_csv("Hospital_Readmissions_Reduction_Program.csv") %>% filter(State=='MN')

infections <-  read_csv("Hospital-Acquired_Condition_Reduction_Program.csv") %>% filter(State=='MN')

survey <- read_csv("Patient_survey_HCAHPS_Hospital.csv") %>% filter(State=='MN')



readmission2 <- readmission %>% mutate(discharges=case_when(`Number of Discharges`=='Not Available'~0,
                                                            TRUE~as.numeric(`Number of Discharges`)))



#infections table has one record per hospital
#winnow down to just the fields we care about
infections_2 <-  infections %>% select(Hospital_Name, `Provider ID`,Total_HAC_Score, Total_HAC_Footnote) %>%
  rename(hospitalID=`Provider ID`, hospital=`Hospital_Name`)



#pull out readmission data for heart failure
readmission_hf<-  readmission %>% filter(`Measure Name`=='READM-30-HF-HRRP') %>% select(`Hospital Name`, `Provider Number`,
  `Excess Readmission Ratio`) %>%
  rename(HF_readmission=`Excess Readmission Ratio`, hospitalID=`Provider Number`, hospital=`Hospital Name`)

#readmission rates for pneumonia
readmission_pn<-  readmission %>% filter(`Measure Name`=='READM-30-PN-HRRP') %>% select(`Hospital Name`, `Provider Number`,
 `Excess Readmission Ratio`) %>%
  rename(PN_readmission=`Excess Readmission Ratio`, hospitalID=`Provider Number`, hospital=`Hospital Name`)


#pull out readmission data for hip/knee surgeries
readmission_hip_knee<-  readmission %>% filter(`Measure Name`=='READM-30-HIP-KNEE-HRRP') %>% select(`Hospital Name`, `Provider Number`,
  `Excess Readmission Ratio`) %>%
  rename(HIP_KNEE_readmission = `Excess Readmission Ratio`, hospitalID=`Provider Number`, hospital=`Hospital Name`)

```

```r
#pull out summary star rating from survey data


star_rating <-  survey %>% filter(`HCAHPS Measure ID`=='H_STAR_RATING') %>%
  select(`Provider ID`, `Hospital Name`, `Address`, `City`, `County Name`, `Patient Survey Star Rating` ) %>%
  rename(hospitalID=`Provider ID`, hospital=`Hospital Name`)




#join data frames together

hospitals <- inner_join(star_rating, infections_2 %>% select(hospitalID, Total_HAC_Score, Total_HAC_Footnote) , by=c("hospitalID"="hospitalID"))

hospitals <-  inner_join(hospitals, readmission_hf %>% select(hospitalID, HF_readmission), by=c("hospitalID"="hospitalID"))

hospitals <-  inner_join(hospitals, readmission_hip_knee %>% select(hospitalID, HIP_KNEE_readmission), by=c("hospitalID"="hospitalID"))

hospitals <-  inner_join(hospitals, readmission_pn %>% select(hospitalID, PN_readmission), by=c("hospitalID"="hospitalID"))


write.csv(hospitals, "hospitals_fordatawrapper.csv", row.names=FALSE)

```