Holden Herrell IST687 HW5

```
#Step 1
url<-"http://opendata.maryland.gov/resource/pdvh-tf2u.json"
df<-read.socrata(url)

#Step 2
namesOfColumns <- c("ACC_DATE","ACC_TIME","ACC_TIME_CODE","BARRACK","CASE_NUMBER","CITY_NAME",
                    "COLLISION_WITH_1","COLLISION_WITH_2","COUNTY_CODE","COUNTY_NAME","DAY_OF_WEEK",
                    "DIST_DIRECTION","DIST_FROM_INTERSECT","INJURY","INTERSECT_ROAD","PROP_DEST",
                    "ROAD","VEHICLE_COUNT")
cleandf<-function(df, namesOfColumns){
  colnames(df)<-namesOfColumns
  return(df)
}

df<-cleandf(df, namesOfColumns)

#Step 3
df$DAY_OF_WEEK <-gsub(" ","",df$DAY_OF_WEEK)
SundayAccidents<-sqldf("select count(DAY_OF_WEEK) as 'Accidents', DAY_OF_WEEK as 'Day'
                       from df where DAY_OF_WEEK = 'SUNDAY'")
SundayAccidents
```

```
##   Accidents    Day
## 1      2373 SUNDAY
```

```
AccidentsWithInjury<-sqldf("select count(*) as 'Accidents with Injuries' from df where INJURY='YES' ")
AccidentsWithInjury
```

```
##   Accidents with Injuries
## 1                    6433
```

```
DailyInjuries<- sqldf(" select  DAY_OF_WEEK as 'Day', count(DAY_OF_WEEK) as 'Accidents'
                      from df where INJURY='YES' group by DAY_OF_WEEK")
DailyInjuries
```

```
##         Day Accidents
## 1    FRIDAY      1043
## 2    MONDAY       915
## 3  SATURDAY       950
```

```
## 4    SUNDAY       818
## 5  THURSDAY       968
## 6   TUESDAY       843
## 7 WEDNESDAY       896
```

```
#Step 4
tapply(df$CASE_NUMBER, df$DAY_OF_WEEK=='SUNDAY',length)
```

```
## FALSE  TRUE
## 16265  2373
```

```
tapply(df$CASE_NUMBER, df$INJURY == 'YES', length)
```

```
## FALSE  TRUE
## 12204  6433
```

```
injuries<-df[which(df$INJURY=='YES'),]
tapply(injuries$CASE_NUMBER, injuries$DAY_OF_WEEK, length)
```

```
##    FRIDAY    MONDAY  SATURDAY    SUNDAY  THURSDAY   TUESDAY WEDNESDAY
##      1043       915       950       818       968       843       896
```