EDS241: Assignment 1

Charles Hendrickson

01/19/2022

```
# Load the CalEnviroScreen 4.0 data from the California Office of Environmental Health Hazards Assessme

mydata <- read.xlsx("data/CES4.xlsx")

# Select the specific columns we will be using in our analysis.

mydata <- mydata %>%
  select("Census.Tract", "Total.Population", "California.County", "Low.Birth.Weight", "PM2.5", "Poverty

# omit all NA values from the dataset

mydata <- na.omit(mydata)
```

(a) What is the average concentration of PM2.5 across all census tracts in California?

```
pm2.5_avg <- mydata %>%
  summarise(pm2.5_avg = mean(PM2.5))

print(pm2.5_avg)
```

```
##   pm2.5_avg
## 1  10.19529
```

(b) What county has the highest level of poverty in California?

```
max_poverty <- mydata %>%
  group_by(California.County) %>%
  summarise(mean_poverty = mean(Poverty))
```