

# STA 108 Project 1

Noe Velasquez, \_\_\_\_\_, and \_\_\_\_\_

2023-01-27

## Introduction

### 1

```
# 1
# a
cdi = read.table("CDI.txt")
colnames(cdi) <- c("id_num", "county",
                  "state", "land_area",
                  "pop_total", "pop_18_34",
                  "pop_65_old", "active_physicians",
                  "hospital_beds", "serious_crimes",
                  "pct_hsgrad", "pct_bachelors",
                  "pct_poverty", "pct_unemp",
                  "income_percap", "income_total",
                  "region")
model_1 = lm(active_physicians ~ pop_total, data = cdi)
model_2 = lm(active_physicians ~ hospital_beds, data = cdi)
model_3 = lm(active_physicians ~ income_total, data = cdi)
```

### a

The estimated regression functions are: 1. The number of active physicians in relation to total population is estimated by  $\hat{Y} = -110.6347772 + 0.0027954X$ .

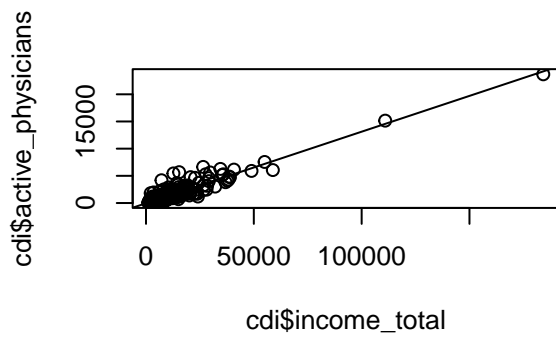
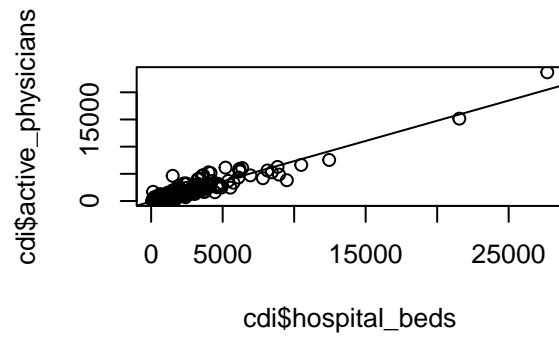
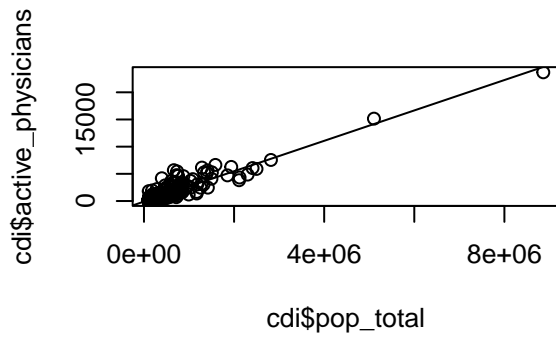
2. The number of active physicians in relation to number of hospital beds is estimated by  $\hat{Y} = -95.9321847 + 0.7431164X$ .

3. The number of active physicians in relation to total personal income is estimated by  $\hat{Y} = -48.3948489 + 0.1317012X$ .

```
#b
par(mfrow = c(2,2))
plot(cdi$pop_total, cdi$active_physicians)
abline(model_1)

plot(cdi$hospital_beds, cdi$active_physicians)
abline(model_2)
```

```
plot(cdi$income_total, cdi$active_physicians)
abline(model_3)
```



**b**

```
# c
```

**c**