project

```
knitr::opts_chunk$set(echo = TRUE)
library(tidyverse)
```

```
## -- Attaching packages ------------ tidyverse 1.3.0 --
```

```
## v ggplot2 3.3.0     v purrr   0.3.3
## v tibble  2.1.3     v dplyr   0.8.5
## v tidyr   1.0.2     v stringr 1.4.0
## v readr   1.3.1     v forcats 0.4.0
```

```
## -- Conflicts --------------- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
states <- read_csv("~/states.csv")
```

```
## Parsed with column specification:
## cols(
##   State = col_character(),
##   Abbrev = col_character(),
##   Code = col_character()
## )
```

```
data_1_ <- read_csv("~/data (1).csv")
```

```
## Parsed with column specification:
## cols(
##   State = col_character(),
##   LE = col_double()
## )
```

```
h08_2 <- read_csv("~/h08 2.csv")
```

```
## Parsed with column specification:
## cols(
##   State = col_character(),
##   `2018` = col_number()
## )
```

```
Heart_Disease_Mortality_by_State_3 <- read_csv("~/Heart Disease Mortality by State 3.csv")
```

```
## Parsed with column specification:
## cols(
##   Code = col_character(),
##   RATE = col_double()
## )
```

```
states1<-full_join(states,data_1_,by="State")
states2<-full_join(states1,h08_2,by="State")
states3<- full_join(states2, Heart_Disease_Mortality_by_State_3, by="Code")
```

```
states3 %>%
  rename(Income="2018")
```

```
## # A tibble: 51 x 6
##    State              Abbrev Code     LE Income  RATE
```

```
##    <chr>                 <chr>  <chr> <dbl>  <dbl> <dbl>
##  1 Alabama               Ala.   AL     75.4 49936  225.
##  2 Alaska                Alaska AK     78.3 68734  130.
##  3 Arizona               Ariz.  AZ     79.6 62283  136.
##  4 Arkansas              Ark.   AR     76   49781  217.
##  5 California            Calif. CA     80.8 70489  140.
##  6 Colorado              Colo.  CO     80   73034  124.
##  7 Connecticut           Conn.  CT     80.8 72812  142.
##  8 Delaware              Del.   DE     78.4 65012  159.
##  9 District of Columbia  D.C.   DC     NA      NA   NA
## 10 Florida               Fla.   FL     79.4 54644  143.
## # ... with 41 more rows
```
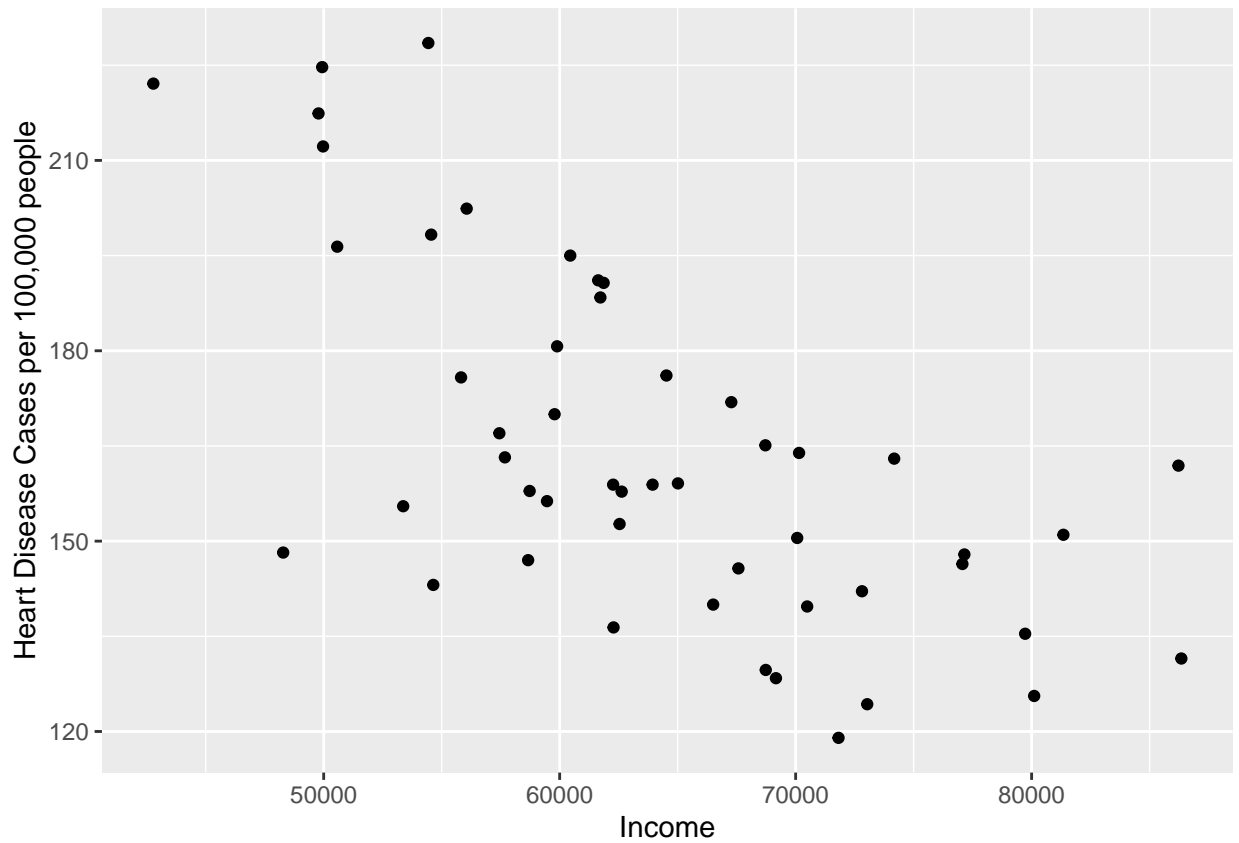
```
states3[,"Income"]<-states3[,5]
ggplot(data=states3, mapping = aes(x= Income, y= LE))+geom_point()+ylab("Life Expectancy")
```

```
## Warning: Removed 1 rows containing missing values (geom_point).
```
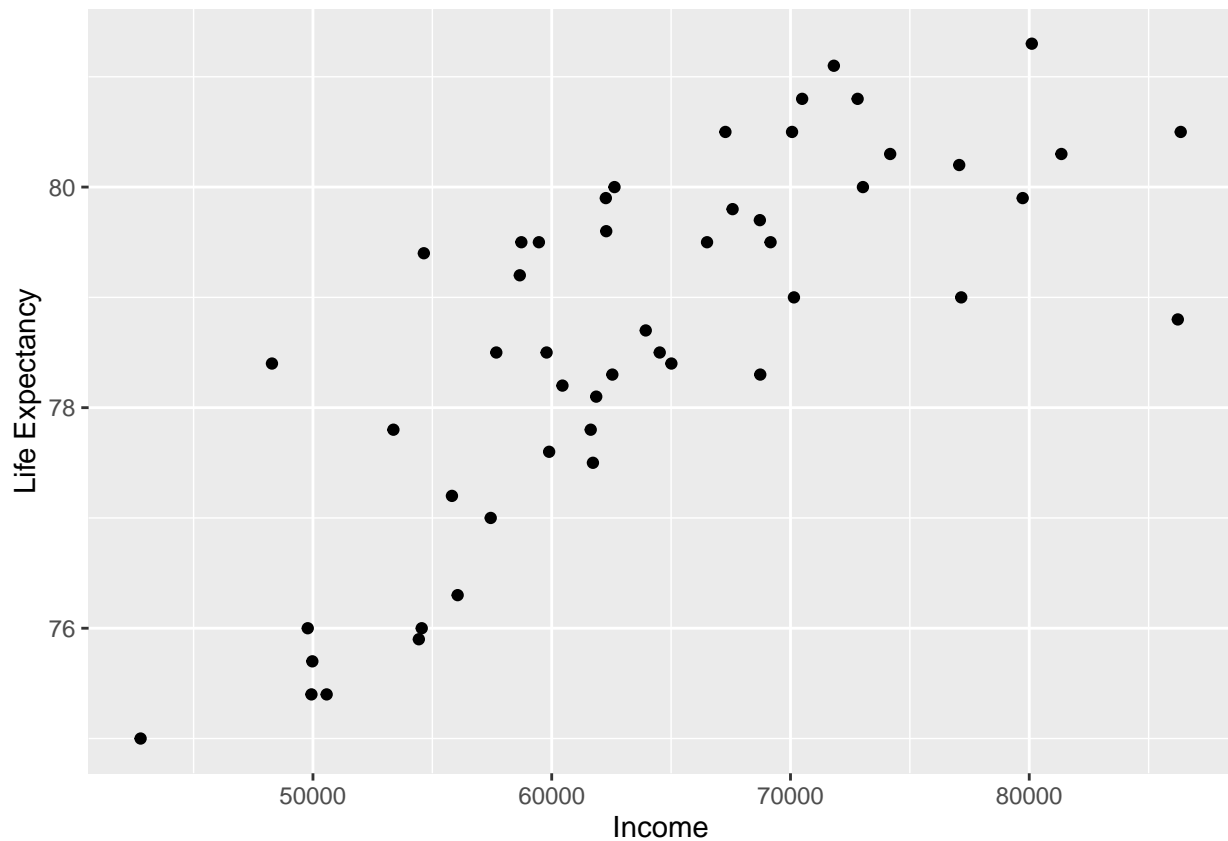


```
ggplot(data = states3, mapping= aes(x=Income, y= RATE))+geom_point()+ylab("Heart Disease Cases per 100,0
```

```
## Warning: Removed 1 rows containing missing values (geom_point).
```

There is a very strong positive correlation between income and life expectancy. There is a very strong negative correlation between income and rates of heart disease. The higher prevalence of heart disease in lower income communities is likely a factor in life expectancy being lower for poorer communities.