income_temperature

```r
# load data sets
income <- read_csv('IRSIncomeByZipCode.csv')
state_codes <- read_csv('statenamecode.csv')
state_temps <- read_csv('statetemps.csv')
```

```r
# tidy data to only include variables that will be used
income <- income %>% select(index,
                            ZIPCODE,
                            STATE,
                            `Adjusted gross income (AGI)`)

# rename some columns in <income>
income <- income %>%
  rename(
    zip = ZIPCODE,
    state_code = STATE,
    agi = `Adjusted gross income (AGI)`
  )

# create new variable that contains quartiles
income <- income %>%
  mutate(percentile = case_when(agi <= 33003 ~ "25%",
                                agi > 33003 & agi <= 67463 ~ "50%",
                                agi > 67463 & agi <= 122500 ~ "75%",
                                agi > 122500 ~ "0%"))
```

```r
# left merge
state_temps <- merge(x = state_temps, y = state_codes, by = "State")
```

```r
# further wrangle <state_temps>
state_temps <- state_temps %>%
  select(State,
         `Avg C`,
         Code) %>%
  rename(
    state_name = State,
    temp = `Avg C`,
    state_code = Code
  )
```

```r
state_temps_income <- merge(x = income, y = state_temps, by = "state_code")
```

```r
lower <- state_temps_income %>% filter(agi < 504420)
model1 <- lm(agi ~ temp, data = lower)
summary(model1)$coefficients
```

```
##              Estimate Std. Error  t value      Pr(>|t|)
## (Intercept) 66913.512  2435.3409 27.47603 2.069607e-163
## temp         3791.242   197.3764 19.20818  1.495110e-81
```

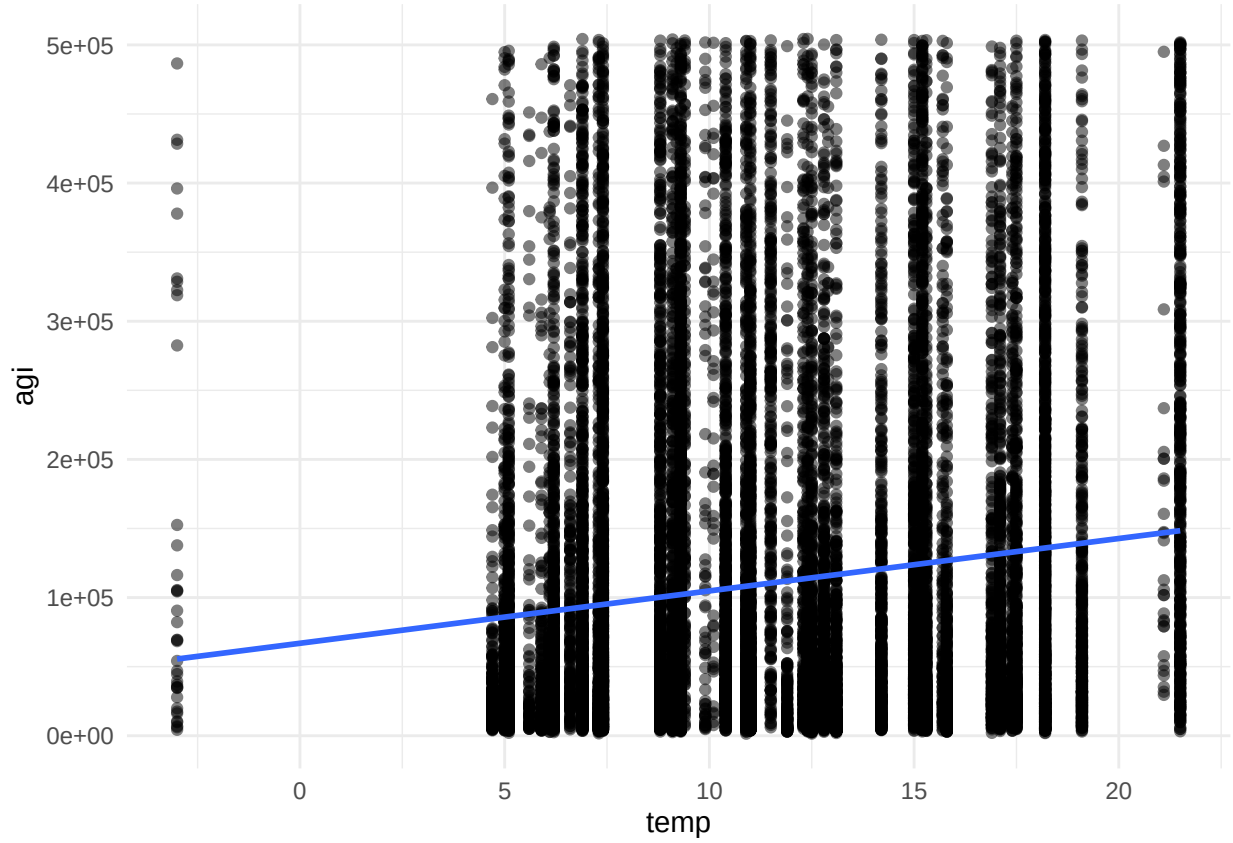
```
cor(x = lower$temp, y = lower$agi)
```

```
## [1] 0.129077
```

```
summary(model1)$r.squared
```

```
## [1] 0.01666087
```

```
state_temps_income %>%
  filter(agi < 504420) %>%
  ggplot(aes(x = temp, y = agi)) +
  geom_point(alpha = 0.5) +
  geom_smooth(method = 'lm', se = FALSE) +
  theme_minimal()
```



Bibliography

https://wid.world/data/ https://en.wikipedia.org/wiki/List_of_countries_by_average_yearly_temperature
https://dqydj.com/average-median-top-household-income-percentiles/ https://www.kaggle.com/datasets/

satoshidatamoto/irs-income-by-zip-code-a-deeper-looke?select=IRSIncomeByZipCode.csv https://worldpopulationreview.com/states/state-abbreviations https://www.currentresults.com/Weather/US/average-annual-state-temperatures.php