HW_5

Benjamin Ramsey

3/23/2021

This is a visualization of solar energy project costs and their associated money incentives as reported to the New York State Energy Research and Development Authority. These projects were designated as commercial or industrial.

```
library(tidyverse)
```

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

## v ggplot2 3.3.3     v purrr   0.3.4
## v tibble  3.1.0     v dplyr   1.0.5
## v tidyr   1.1.3     v stringr 1.4.0
## v readr   1.4.0     v forcats 0.5.1

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

```
library(tinytex)
```

```
Solar_Electric_Programs <- read_csv("Solar_Electric_Programs_NYSERDA.csv")
```

```
##
## -- Column specification --------------------------------------------------------
## cols(
##   .default = col_character(),
##   `Zip Code` = col_double(),
##   `Total Inverter Quantity` = col_double(),
##   `Total PV Module Quantity` = col_double(),
##   `Project Cost` = col_double(),
##   `$Incentive` = col_double(),
##   `Total Nameplate kW DC` = col_double(),
##   `Expected KWh Annual Production` = col_double()
## )
## i Use `spec()` for the full column specifications.
```

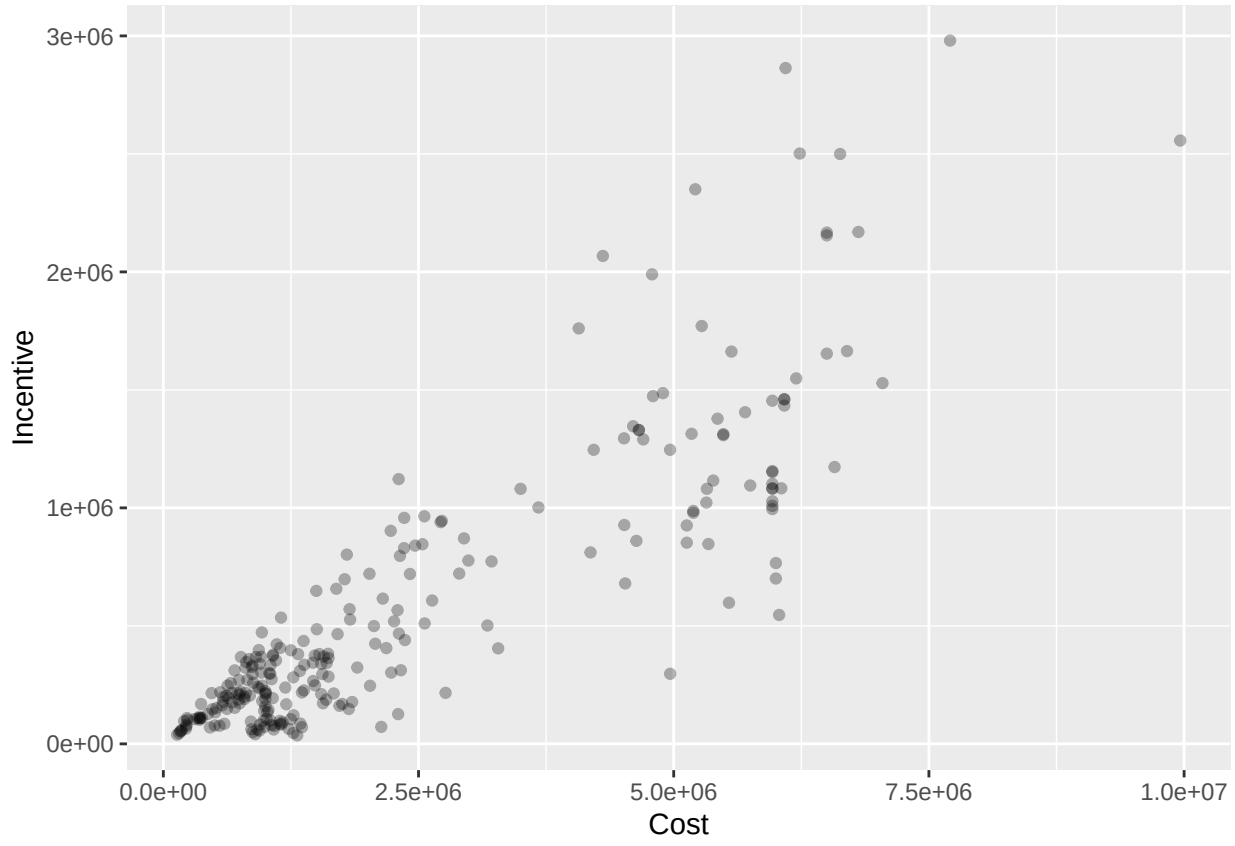
```
Cost_Incentive <- filter(Solar_Electric_Programs, `Program Type` == "Commercial/Industrial (Competitive)
```

```
summarize(Cost_Incentive, mean(`Project Cost`), median(`Project Cost`), mean(`$Incentive`), median(`$In
```

```
## # A tibble: 1 x 4
##   `mean(\`Project Co~ `median(\`Project C~ `mean(\`$Incentiv~ `median(\`$Incent~
##                 <dbl>                <dbl>              <dbl>              <dbl>
## 1            2378666.             1483835            576346.            334973.
```

My hypothesis is that there is positive relationship between project cost and the amount of incentive it received because bigger projects require more funding.

```
ggplot(data = Cost_Incentive) +
  geom_point(mapping = aes(x = `Project Cost`, y = `$Incentive`), alpha = .3) +
  labs(x = 'Cost', y = 'Incentive')
```




```
cor.test(Cost_Incentive$`Project Cost`, Cost_Incentive$`$Incentive`)
```

```
##
##  Pearson's product-moment correlation
##
## data:  Cost_Incentive$`Project Cost` and Cost_Incentive$`$Incentive`
## t = 28.047, df = 251, p-value < 2.2e-16
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.8372717 0.8976323
```

```
## sample estimates:
##       cor
## 0.8706928
```

The scatterplot of cost against incentive shows a strong positive realtionship between the two variables. The correlation coefficient is 0.87, also indicating a strong positive relationship between project cost and incentive amount.