EDE Group Project

Gaby Czarniak, Mara Michel, Sam Sedar

2023-11-14

Contents

Research Rationale

[RATIONALE HERE]

There are 100 counties in North Carolina. There are 843 power plants made up of 1190 generators.

##Data Set Information Data used in this analysis comes from a variety of sources. A brief overview of each source is outlined below and full citations can be found at the end of this report. More information on the metadata for each source can be found in the 'Metadata' folder.

*eGRID - 2020 and 2021 data from the U.S. EPA on all power plants in the U.S. Data includes information on plant locations, production, retirements,and emissions.The 2020 data set was only utilized for historical retirement data as that was removed from the 2021 data set.

*Income - 2021 data from the USDA on percentage of people in each US county living in poverty.

*Unemployment - 2000-2022 data from the USDA on unemployment rates in each US county.

*Social Vulnerability Index - 2020 data from the US CDC on key vulnerability criteria including socioeconomic status, household characteristics, racial & ethnic minority status, and house type & transportation.

##Data Wrangling The intial step in our research, once loaded, was data wrangling on the three distinct datasets: eGRID, unemployment, and income data. For the eGRID dataset, the code selects specific columns related to power plant information in North Carolina, converts certain columns to numeric format, and then combines the generator and plant data based on a common identifier. Duplicate columns are removed, resulting in the plant_gen_data dataframe. The code then proceeds to handle the unemployment dataset, selecting relevant columns covering the years 2002-2022 and merging it with the plant_gen_data dataframe using the FIPS code as a key. Similarly, the income dataset is processed by selecting specific columns, and the final step involves merging this income data with the previously created dataframe. The resulting dataset, named processed_data, contains information on power plants, unemployment rates, and income data, and the column names are displayed at the end of the code.

#CONFIRM WHETHER TABLE IS NEEDED - IT'S IN THE RUBRIC SO I THINK YES?

```
##  [1] "FIPS_Code"             "SEQGEN"               "YEAR"
##  [4] "PSTATABB_gen"          "PNAME"                "ORISPL"
##  [7] "GENID"                 "NUMBLR"               "GENSTAT"
## [10] "PRMVR"                 "FUELG1"               "NAMEPCAP_gen"
## [13] "CFACT"                 "GENNTAN"              "GENNTOZ"
## [16] "GENERSRC"              "GENYRONL"             "GENYRRET"
## [19] "PLFUELCT"              "NAMEPCAP_plant"       "PLNGENAN"
## [22] "PLSO2RTA"              "PLCO2RTA"             "PLCH4RTA"
## [25] "PLN2ORTA"              "PLCO2AN"              "PLN2OAN"
```

```
## [28] "PLNOXAN"                "PLSO2AN"                "PLCH4AN"
## [31] "CNTYNAME"               "LAT"                    "LON"
## [34] "COALFLAG"               "Unemployment_rate_2000" "Unemployment_rate_2001"
## [37] "Unemployment_rate_2002" "Unemployment_rate_2003" "Unemployment_rate_2004"
## [40] "Unemployment_rate_2005" "Unemployment_rate_2006" "Unemployment_rate_2007"
## [43] "Unemployment_rate_2008" "Unemployment_rate_2009" "Unemployment_rate_2010"
## [46] "Unemployment_rate_2011" "Unemployment_rate_2012" "Unemployment_rate_2013"
## [49] "Unemployment_rate_2014" "Unemployment_rate_2015" "Unemployment_rate_2016"
## [52] "Unemployment_rate_2017" "Unemployment_rate_2018" "Unemployment_rate_2019"
## [55] "Unemployment_rate_2020" "Unemployment_rate_2021" "Unemployment_rate_2022"
## [58] "PCTPOVALL_2021"         "MEDHHINC_2021"
```

#Exploratory Analysis *We then began data exploration and visualization for the processed_data dataframe, focusing on fuel types, income, and electricity generation in North Carolina. The summary(processed_data) provides a general overview of the dataset. Figure 1 illustrates the percentage makeup of primary fuel types in North Carolina based on total annual generation - the majority of which is natural gas. Subsequently, the scatter plot in figures 2 depicts the relationship between median household income and total plant annual generation, with points colored by fuel type. The analysis further drills down to scatter plots in figures 3 and 4 excluding nuclear data and for a selected set of fuel types (COAL, SOLAR, OIL, WIND, GAS). Additionally, stacked bar plots in figures 5, 6, and 7 visualize the total annual generation by fuel type and county, and for better clarity, separate plots are created for the top 10 and bottom 10 counties based on median household income. These visualizations help explore and understand the relationships between fuel types, income levels, and electricity generation in different counties of North Carolina.

```
##   FIPS_Code             SEQGEN           YEAR      PSTATABB_gen
##  Length:1190        Min.   :13538   Min.   :2021   Length:1190
##  Class :character   1st Qu.:13835   1st Qu.:2021   Class :character
##  Mode  :character   Median :14132   Median :2021   Mode  :character
##                     Mean   :14132   Mean   :2021
##                     3rd Qu.:14430   3rd Qu.:2021
##                     Max.   :14727   Max.   :2021
##
##     PNAME              ORISPL             GENID               NUMBLR
##  Length:1190        Length:1190        Length:1190        Min.   :1.000
##  Class :character   Class :character   Class :character   1st Qu.:1.000
##  Mode  :character   Mode  :character   Mode  :character   Median :1.000
##                                                           Mean   :2.642
##                                                           3rd Qu.:3.000
##                                                           Max.   :9.000
##                                                           NA's   :1137
##    GENSTAT             PRMVR              FUELG1           NAMEPCAP_gen
##  Length:1190        Length:1190        Length:1190        Min.   :  0.100
##  Class :character   Class :character   Class :character   1st Qu.:  2.000
##  Mode  :character   Mode  :character   Mode  :character   Median :  5.000
##                                                           Mean   : 27.218
##                                                           3rd Qu.:  6.425
##                                                           Max.   :950.900
##                                                           NA's   :6
##      CFACT           GENNTAN         GENNTOZ        GENERSRC
##  Min.   :0.0000   Min.   :  0.0   Min.   :  0.0   Length:1190
##  1st Qu.:0.1693   1st Qu.:  3.0   1st Qu.:  1.0   Class :character
##  Median :0.2050   Median : 70.5   Median :121.0   Mode  :character
##  Mean   :0.2367   Mean   :285.3   Mean   :273.3
##  3rd Qu.:0.2330   3rd Qu.:599.0   3rd Qu.:426.0
##  Max.   :1.0460   Max.   :915.0   Max.   :985.0
```

```
##  NA's   :44       NA's   :1016    NA's   :911
##     GENYRONL       GENYRRET      PLFUELCT         NAMEPCAP_plant
##  Min.   :1912   Min.   :2021   Length:1190        Min.   :   1.0
##  1st Qu.:2007   1st Qu.:2021   Class :character   1st Qu.:   4.0
##  Median :2014   Median :2021   Mode  :character   Median :   5.0
##  Mean   :2005   Mean   :2031                      Mean   : 143.1
##  3rd Qu.:2017   3rd Qu.:2049                      3rd Qu.:  20.0
##  Max.   :2021   Max.   :2049                      Max.   :2558.2
##  NA's   :15     NA's   :1161
##     PLNGENAN           PLSO2RTA          PLCO2RTA            PLCH4RTA
##  Min.   :   -1884   Min.   : 0.000   Min.   :    0.000   Min.   :0.0000
##  1st Qu.:    4556   1st Qu.: 0.006   1st Qu.:    9.291   1st Qu.:0.0010
##  Median :    9030   Median : 0.400   Median :  846.821   Median :0.0170
##  Mean   :  384657   Mean   : 1.977   Mean   : 1847.908   Mean   :0.0714
##  3rd Qu.:   27568   3rd Qu.: 2.768   3rd Qu.: 1688.995   3rd Qu.:0.0740
##  Max.   :19662114   Max.   :43.200   Max.   :24335.500   Max.   :1.1520
##  NA's   :22         NA's   :852      NA's   :852         NA's   :852
##     PLN2ORTA          PLCO2AN           PLN2OAN             PLNOXAN
##  Min.   :0.0000   Min.   :      0   Min.   :     0.00   Min.   :   0.000
##  1st Qu.:0.0000   1st Qu.:     70   1st Qu.:     1.29   1st Qu.:   0.000
##  Median :0.0020   Median :   1155   Median :    24.06   Median :   3.809
##  Mean   :0.0114   Mean   : 568291   Mean   :  8414.78   Mean   : 270.420
##  3rd Qu.:0.0158   3rd Qu.: 219009   3rd Qu.:   997.14   3rd Qu.: 106.072
##  Max.   :0.1970   Max.   :7610767   Max.   :218266.08   Max.   :7119.806
##  NA's   :852      NA's   :853       NA's   :853         NA's   :853
##     PLSO2AN            PLCH4AN            CNTYNAME              LAT
##  Min.   :   0.002   Min.   :      0.0   Length:1190        Min.   :33.94
##  1st Qu.:   0.353   1st Qu.:      6.4   Class :character   1st Qu.:35.14
##  Median :   2.924   Median :    120.3   Mode  :character   Median :35.44
##  Mean   : 137.718   Mean   :  60605.6                      Mean   :35.48
##  3rd Qu.:  12.102   3rd Qu.:   8159.5                      3rd Qu.:35.84
##  Max.   :2959.327   Max.   :1499963.2                      Max.   :36.53
##  NA's   :853        NA's   :853
##       LON           COALFLAG         Unemployment_rate_2000
##  Min.   :-84.18   Length:1190        Min.   :2.400
##  1st Qu.:-80.23   Class :character   1st Qu.:3.600
##  Median :-78.85   Mode  :character   Median :4.300
##  Mean   :-79.20                      Mean   :4.519
##  3rd Qu.:-77.98                      3rd Qu.:5.475
##  Max.   :-75.53                      Max.   :8.900
##
##  Unemployment_rate_2001 Unemployment_rate_2002 Unemployment_rate_2003
##  Min.   :3.000          Min.   : 3.100         Min.   : 3.400
##  1st Qu.:5.600          1st Qu.: 6.600         1st Qu.: 6.400
##  Median :6.400          Median : 7.600         Median : 7.500
##  Mean   :6.664          Mean   : 7.825         Mean   : 7.437
##  3rd Qu.:8.100          3rd Qu.: 9.300         3rd Qu.: 8.775
##  Max.   :9.800          Max.   :11.400         Max.   :11.900
##
##  Unemployment_rate_2004 Unemployment_rate_2005 Unemployment_rate_2006
##  Min.   : 3.10          Min.   : 3.300         Min.   :3.600
##  1st Qu.: 5.50          1st Qu.: 5.200         1st Qu.:4.800
##  Median : 6.40          Median : 6.200         Median :5.800
##  Mean   : 6.54          Mean   : 6.264         Mean   :5.777
```

```
##  3rd Qu.: 7.60          3rd Qu.: 7.300         3rd Qu.:6.600
##  Max.   :11.40          Max.   :10.300         Max.   :9.500
##
##  Unemployment_rate_2007 Unemployment_rate_2008 Unemployment_rate_2009
##  Min.   : 3.400         Min.   : 4.000         Min.   : 6.70
##  1st Qu.: 4.800         1st Qu.: 5.800         1st Qu.:10.40
##  Median : 5.700         Median : 6.800         Median :11.80
##  Mean   : 5.705         Mean   : 6.921         Mean   :11.99
##  3rd Qu.: 6.700         3rd Qu.: 7.800         3rd Qu.:13.93
##  Max.   :10.800         Max.   :11.000         Max.   :17.80
##
##  Unemployment_rate_2010 Unemployment_rate_2011 Unemployment_rate_2012
##  Min.   : 6.50          Min.   : 6.50          Min.   : 6.30
##  1st Qu.:10.00          1st Qu.: 9.80          1st Qu.: 9.30
##  Median :11.90          Median :11.30          Median :10.70
##  Mean   :11.77          Mean   :11.42          Mean   :10.71
##  3rd Qu.:13.40          3rd Qu.:12.40          3rd Qu.:11.80
##  Max.   :17.50          Max.   :18.10          Max.   :18.10
##
##  Unemployment_rate_2013 Unemployment_rate_2014 Unemployment_rate_2015
##  Min.   : 5.200         Min.   : 4.500         Min.   : 4.300
##  1st Qu.: 7.800         1st Qu.: 6.000         1st Qu.: 5.500
##  Median : 8.400         Median : 6.600         Median : 6.000
##  Mean   : 8.851         Mean   : 6.976         Mean   : 6.428
##  3rd Qu.: 9.875         3rd Qu.: 7.700         3rd Qu.: 7.200
##  Max.   :14.900         Max.   :13.000         Max.   :11.900
##
##  Unemployment_rate_2016 Unemployment_rate_2017 Unemployment_rate_2018
##  Min.   :3.900          Min.   :3.400          Min.   :3.000
##  1st Qu.:4.900          1st Qu.:4.300          1st Qu.:3.800
##  Median :5.500          Median :4.800          Median :4.300
##  Mean   :5.778          Mean   :5.052          Mean   :4.468
##  3rd Qu.:6.600          3rd Qu.:5.700          3rd Qu.:5.100
##  Max.   :9.500          Max.   :8.400          Max.   :8.300
##
##  Unemployment_rate_2019 Unemployment_rate_2020 Unemployment_rate_2021
##  Min.   :3.000          Min.   : 5.200         Min.   :3.600
##  1st Qu.:3.700          1st Qu.: 6.300         1st Qu.:4.500
##  Median :4.000          Median : 7.200         Median :5.100
##  Mean   :4.264          Mean   : 7.335         Mean   :5.429
##  3rd Qu.:4.800          3rd Qu.: 8.200         3rd Qu.:6.300
##  Max.   :8.500          Max.   :11.900         Max.   :9.500
##
##  Unemployment_rate_2022 PCTPOVALL_2021  MEDHHINC_2021
##  Min.   :3.000          Min.   : 7.70   Min.   :38613
##  1st Qu.:3.500          1st Qu.:13.10   1st Qu.:44477
##  Median :3.800          Median :17.00   Median :52463
##  Mean   :4.076          Mean   :16.89   Mean   :54735
##  3rd Qu.:4.500          3rd Qu.:20.70   3rd Qu.:59841
##  Max.   :6.500          Max.   :27.90   Max.   :91558
##

## # A tibble: 9 x 2
##   PLFUELCT total_generation
```

```
##   <chr>              <dbl>
## 1 BIOMASS          7461838.
## 2 COAL            79461837
## 3 GAS            253795203.
## 4 HYDRO           19343891.
## 5 NUCLEAR         78248681
## 6 OIL                46952.
## 7 OTHF              236895
## 8 SOLAR           10169755.
## 9 WIND              514703
```

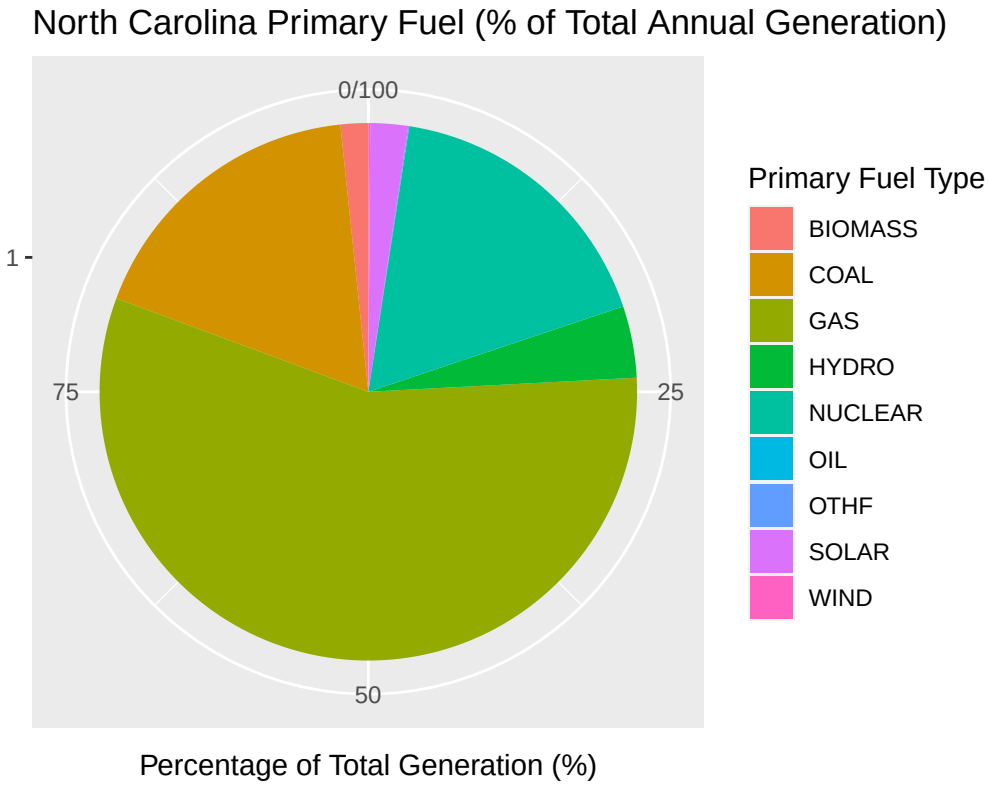
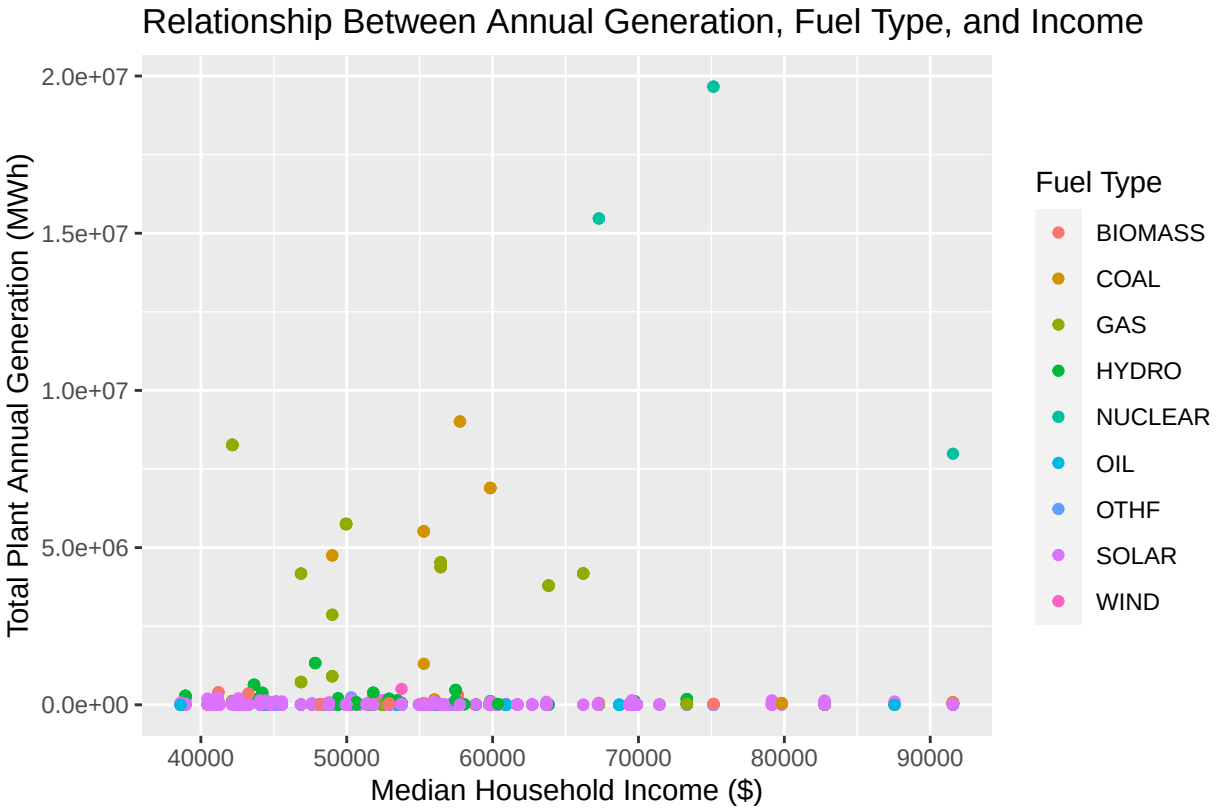
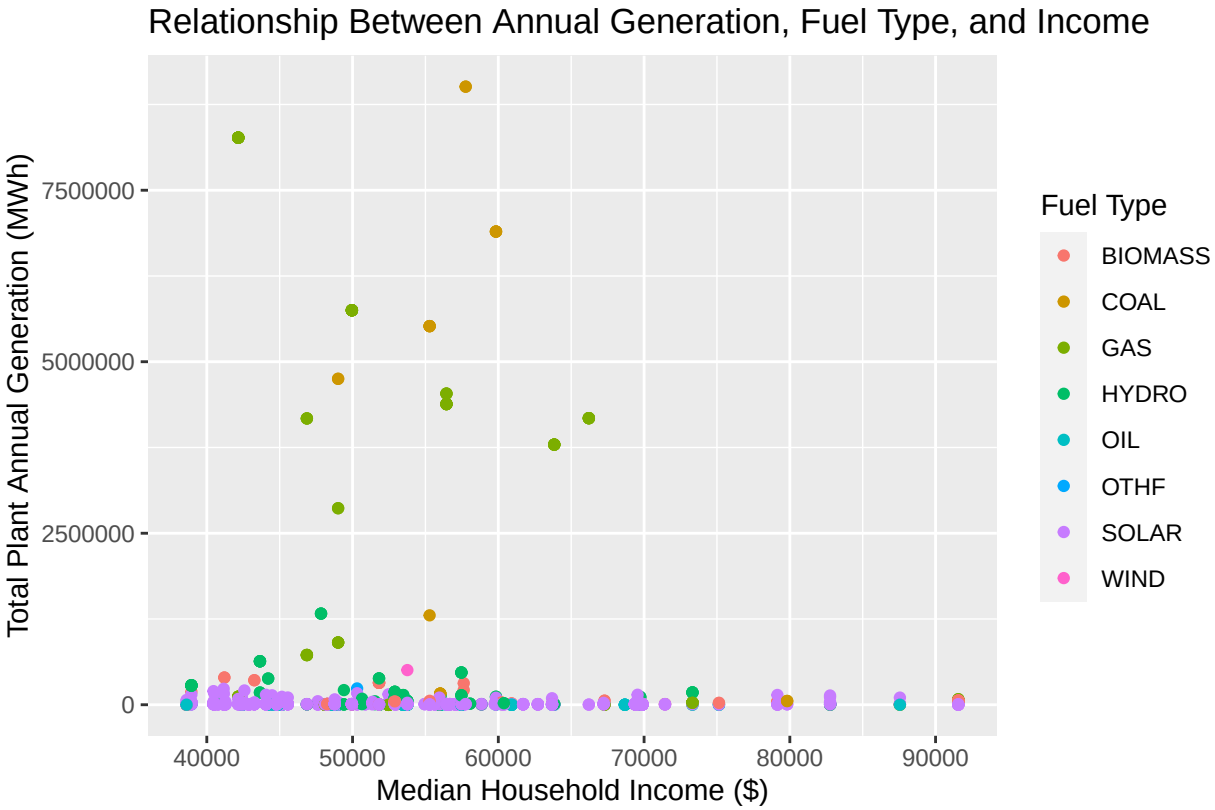


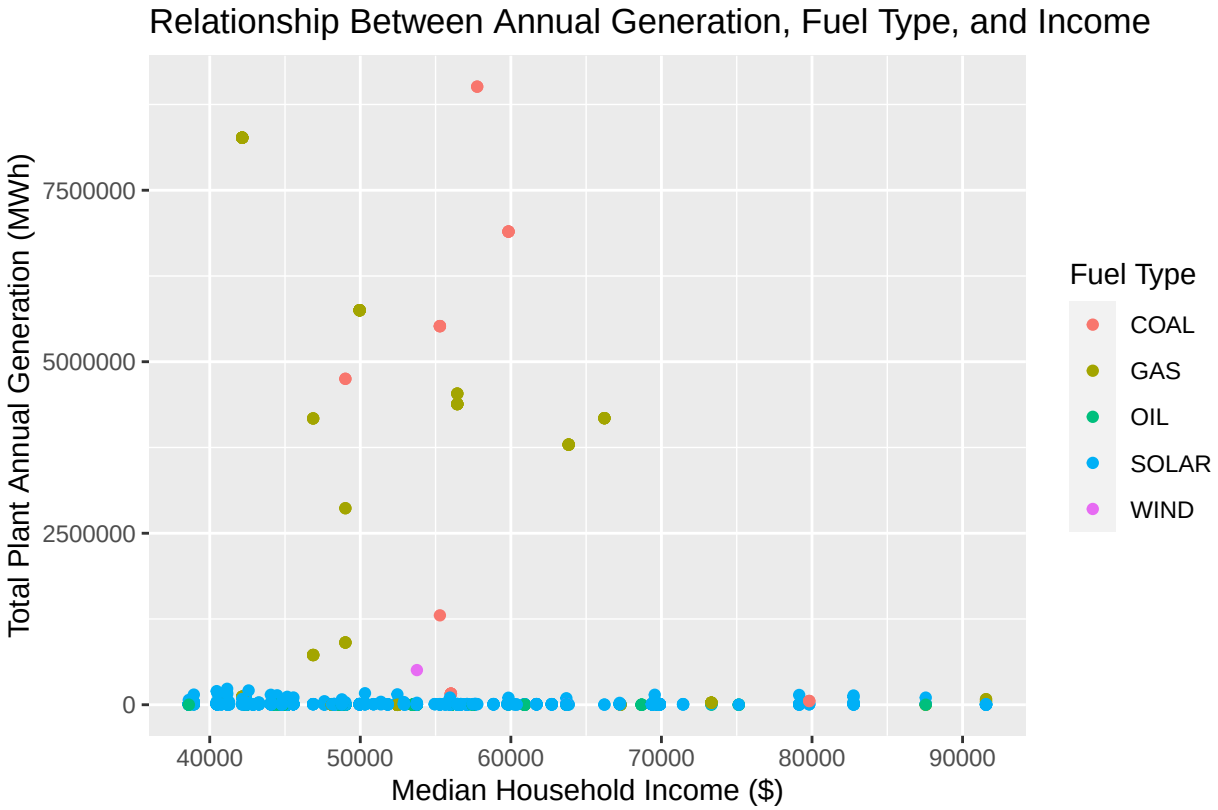
Figure 1: Pie chart illustrating the percentage distribution of primary fuel types in North Carolina based on total annual generation.



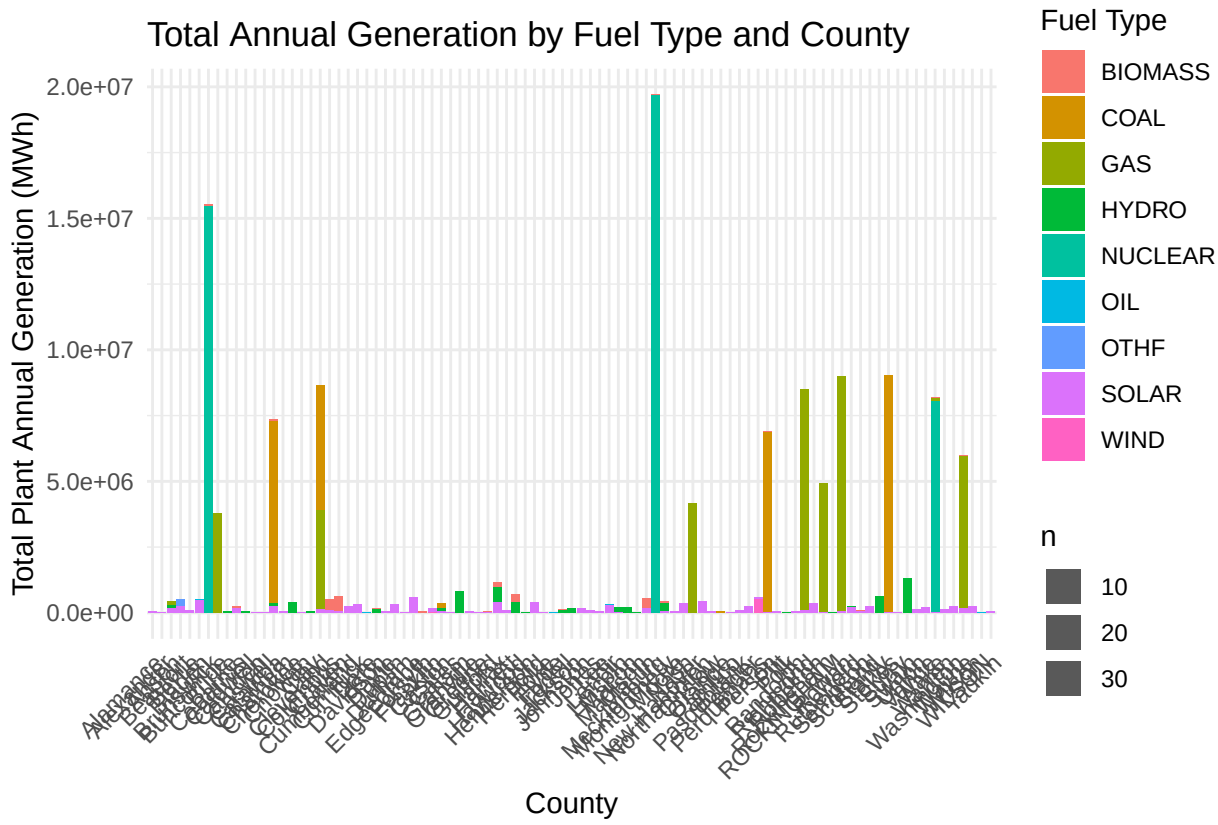
more energy, the plot fails to display any clear relationship, leading to necessary adjustments.



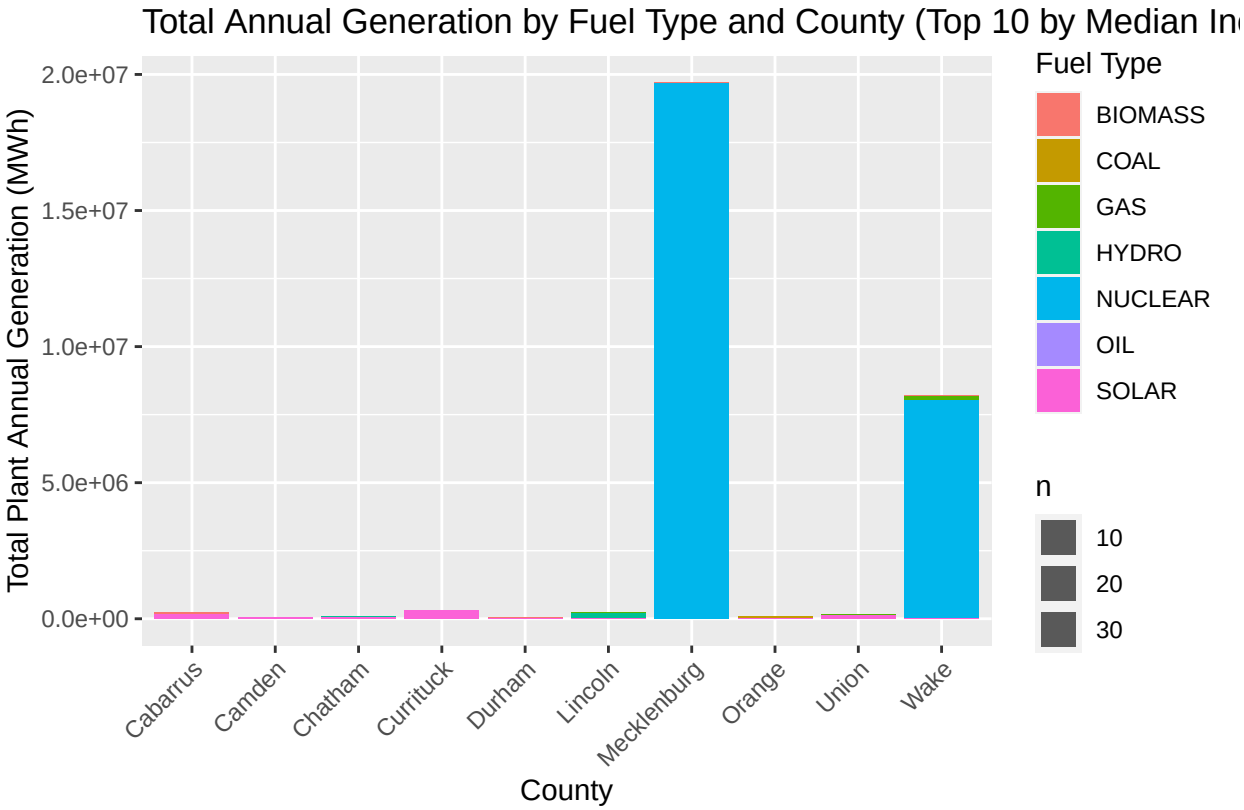
was regenerated and reflects a potential trend, particularly at the $70,000 median income mark.



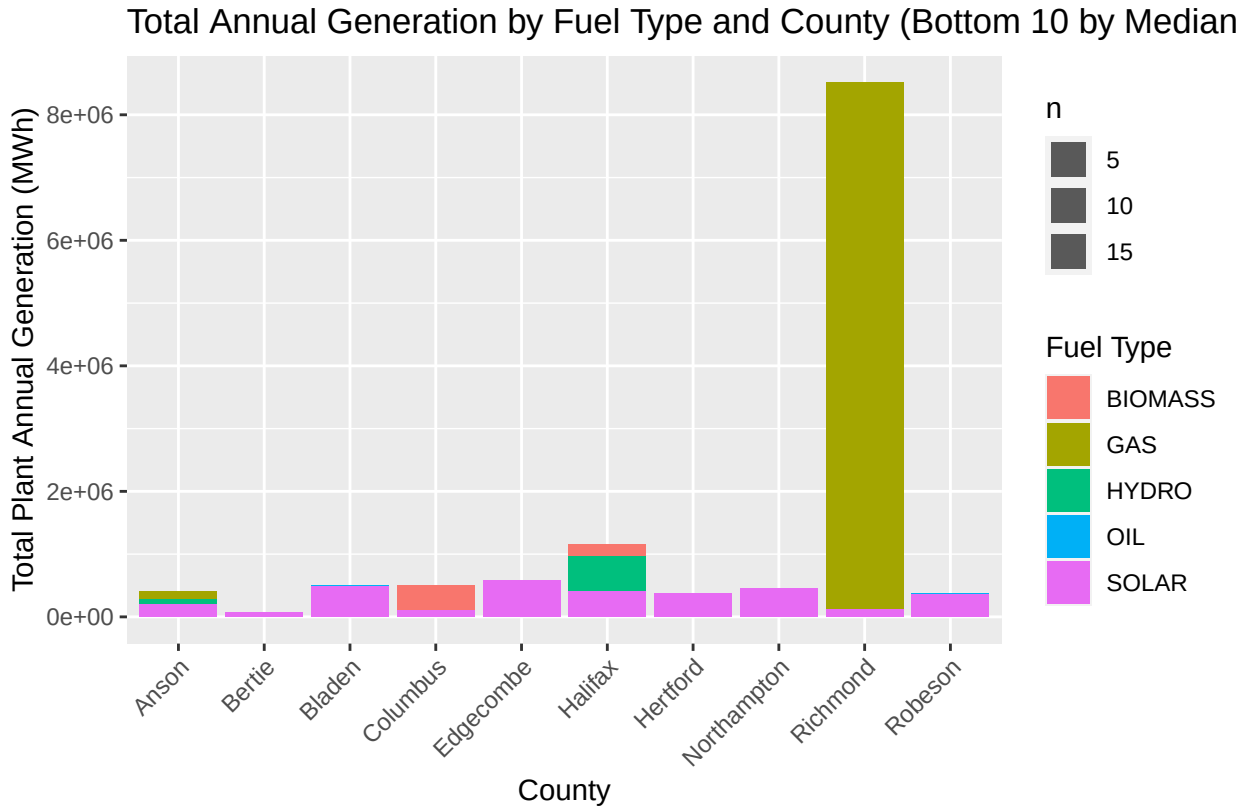
To narrow our exploration further, we selected key energy sources: coal, gas, oil, wind, and solar.



mber of counties in NC, this visualization did not prove fruitful in depicting any further analysis.



ıre digestable manner, we first show the top 10 counties (by median income) in North Carolina.



igestable manner, we then show the bottom 10 counties (by median income) in North Carolina.

*We then created initial maps to better understand the landscape of North Carolina in terms of socioeconomic status and geographic distribution of power plants. Figure 8 below shows the mapping of power plants across the state and helped us to visualize the concentration of power plants across the state.Figure 9 maps the nameplate capacity of those power plants by county. We noticed that multiple counties, especially Richmond county, stood out as having a high concentration of capacity. Figure 10 demonstrates a heat map of the percentage of each county's population living in poverty as of 2021. Richmond county is once again a clear visual outlier on this map.

```
## Reading layer `cb_2018_us_county_20m' from data source
##   `C:\Users\iharm\OneDrive\Documents\CzarniakMichelSedar\Data\Spatial\cb_2018_us_county_20m.shp'
##   using driver `ESRI Shapefile'
## Simple feature collection with 3220 features and 9 fields
## Geometry type: MULTIPOLYGON
## Dimension:     XY
## Bounding box:  xmin: -179.1743 ymin: 17.91377 xmax: 179.7739 ymax: 71.35256
## Geodetic CRS:  NAD83
```

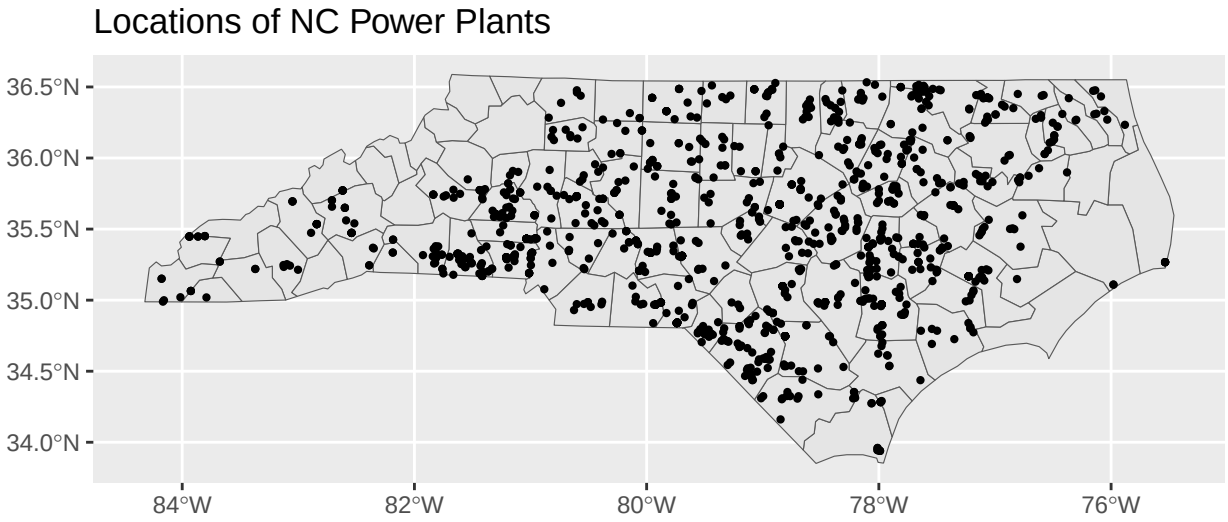
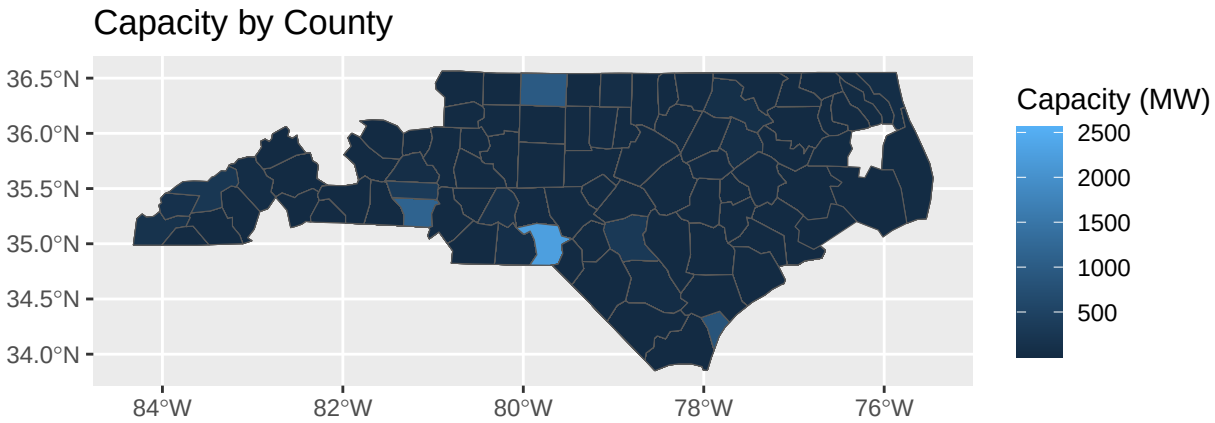
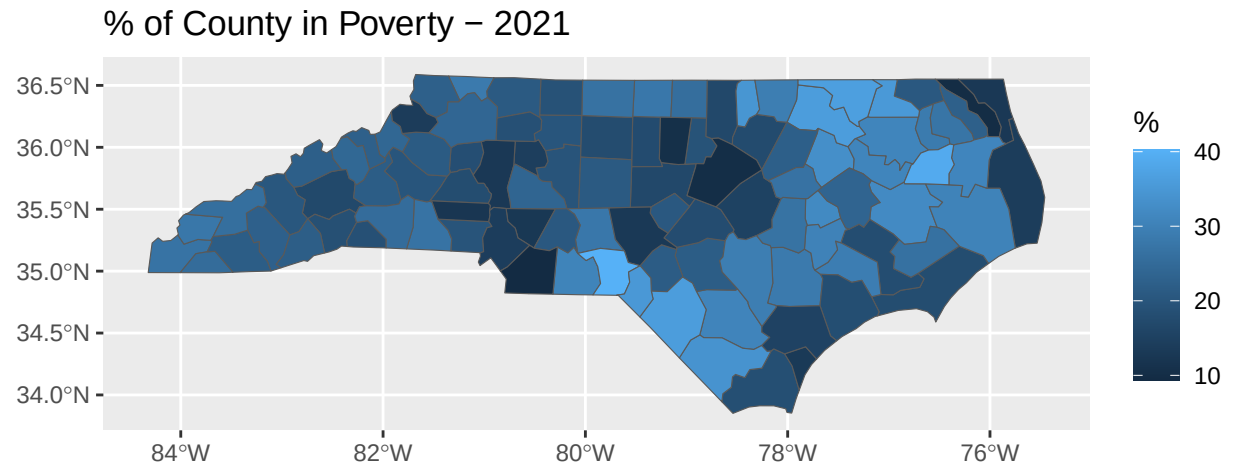


Figure 8:Black dots on this map represent the locations of power plants across North Carolina.



lank. Richmond County is an interesting outlier which the highest capacity of all the counties.



f total county population living in poverty in 2021, as defined by the USDA Economic Research Service

Our final stage of exploratory analysis was to visualize the power plant emissions geographically. Figures 11-15 demonstrate the aggregated power plant emissions per county in 2021 for CO2, NOx, SO2, N20, and CH4 respectively. Unfortunately eGrid did not provide data for Mercury (Hg) in 2021 so that was not analyzed. N20 and CH4 are provided in lbs while CO2, NOx, and SO2 are in short tons. For CO2 emissions, Richmond county stood out. For NOx emissions, Catawba county stood out. For SO2 emissions, Catawba, Haywood, and Person counties stood out.For N20 and CH4 emissions, Catawba and Person counties stood out.

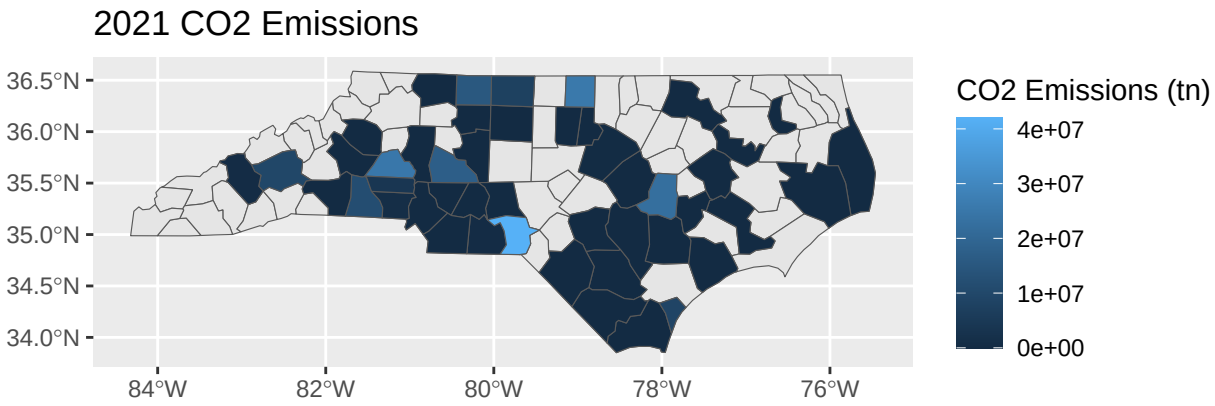


Figure 11: Total CO2 Emissions in 2021 across North Carolina Power Plants

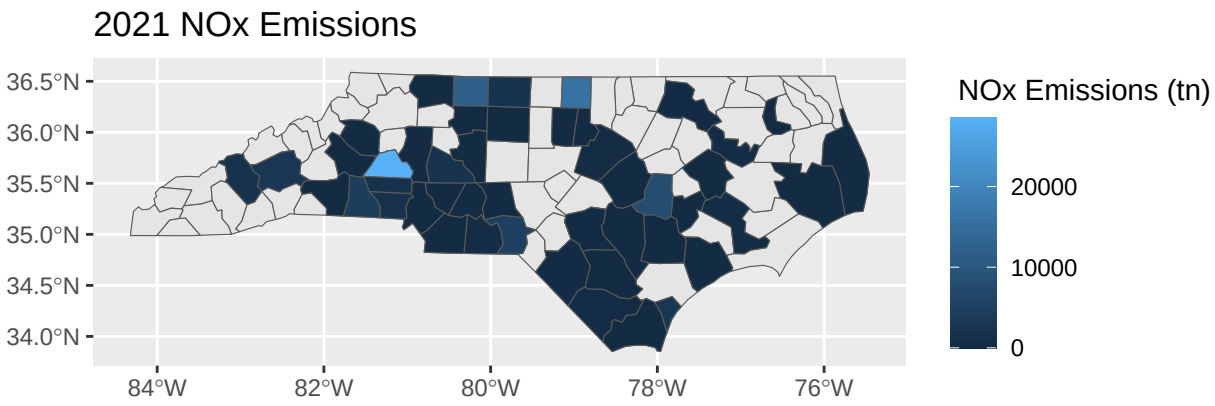


Figure 12: Total NOx Emissions in 2021 across North Carolina Power Plants

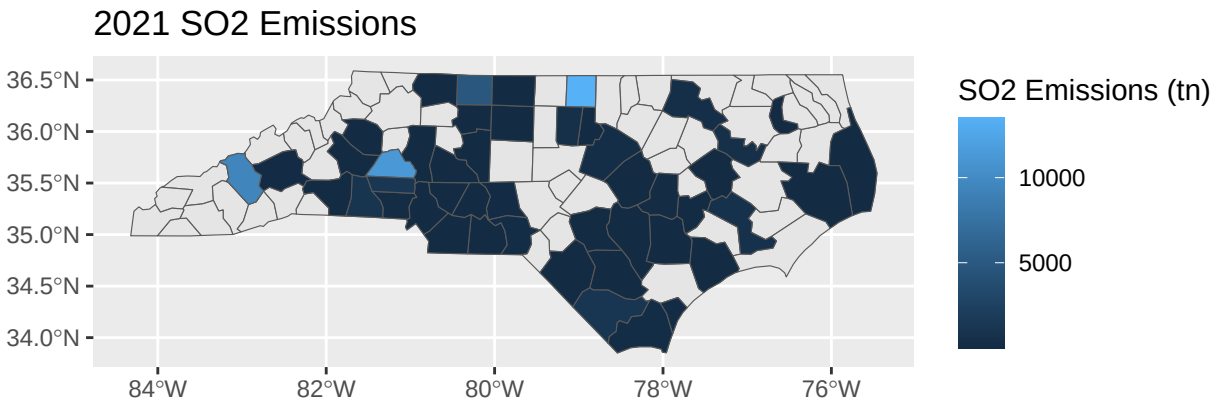


Figure 13: Total SO2 Emissions in 2021 across North Carolina Power Plants

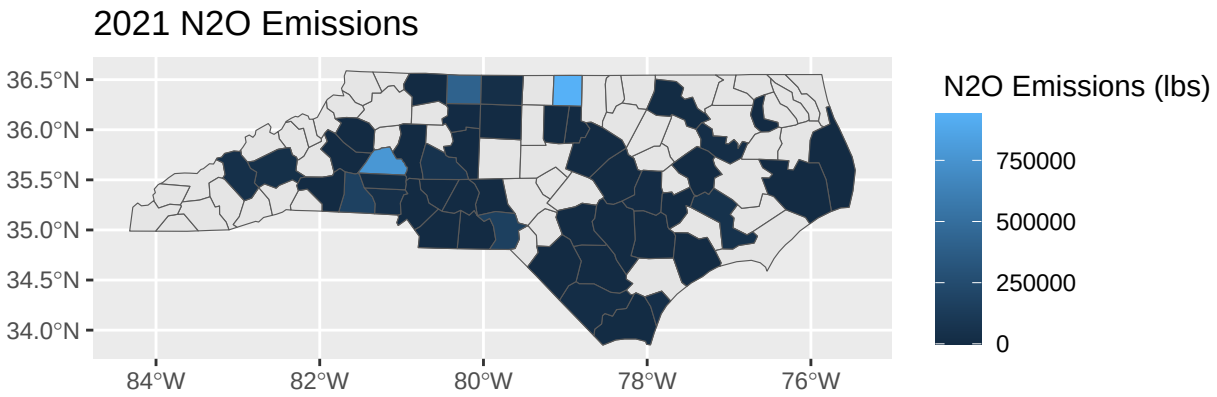


Figure 14: Total N20 Emissions in 2021 across North Carolina Power Plants

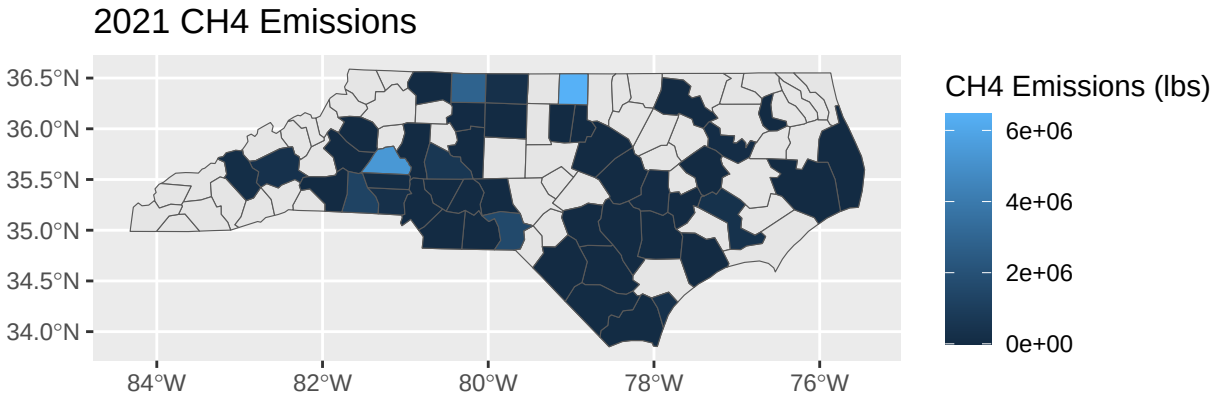
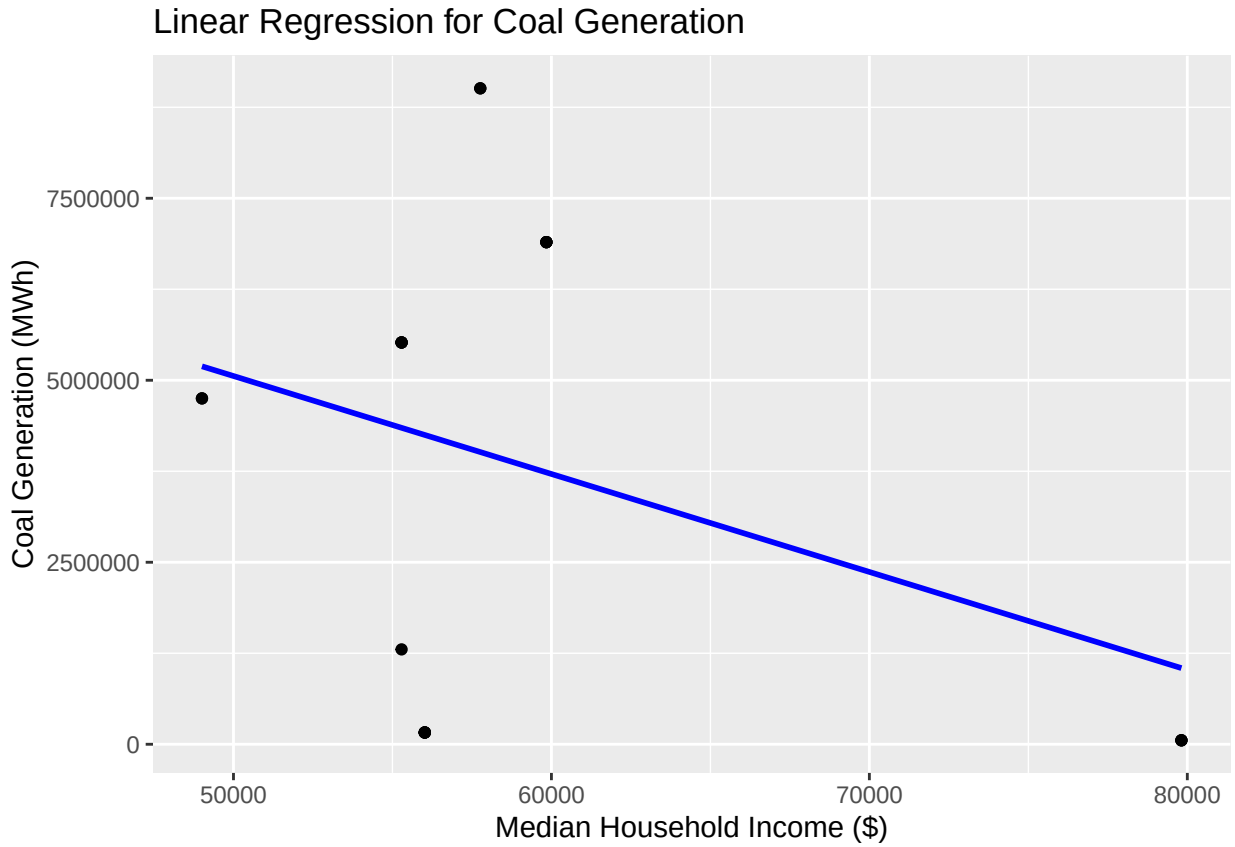


Figure 15: Total CH4 Emissions in 2021 across North Carolina Power Plants

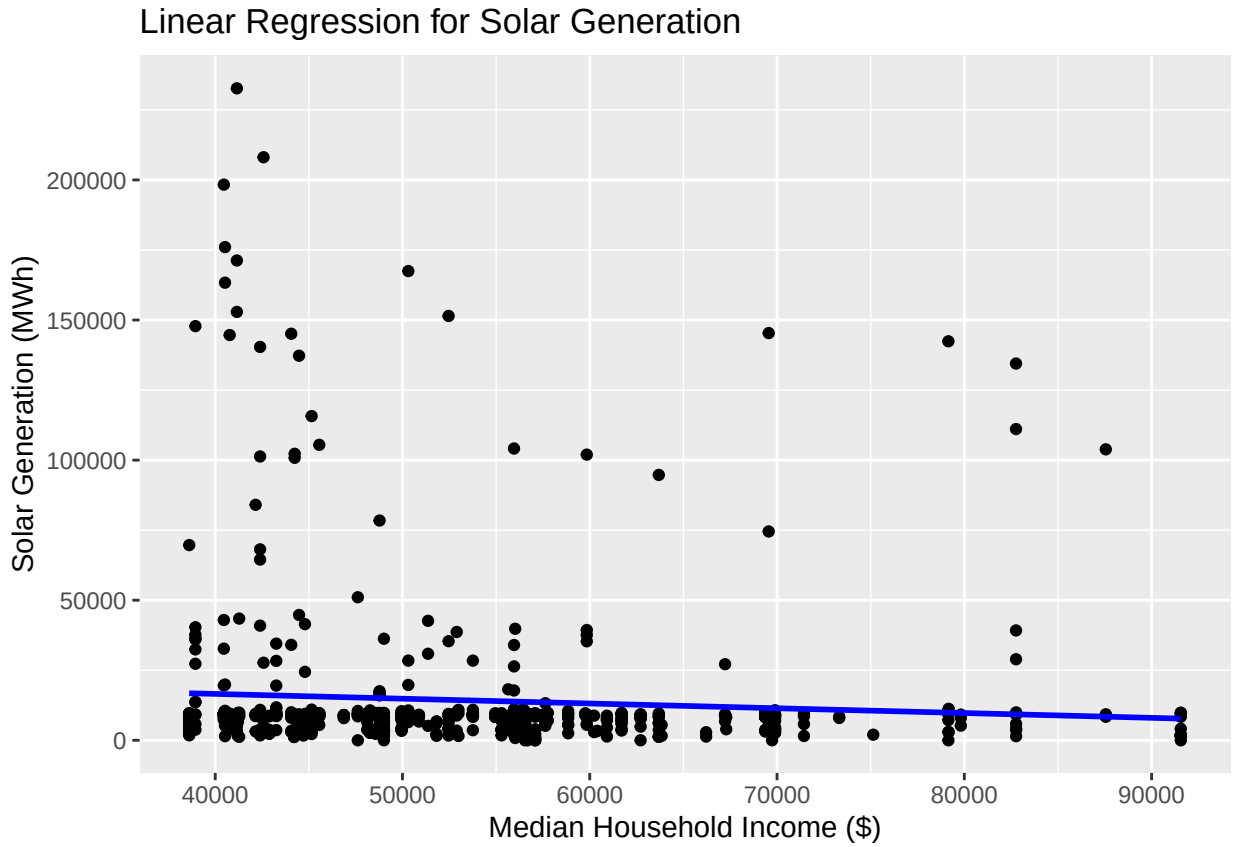
```
##
## Call:
## lm(formula = PLNGENAN ~ MEDHHINC_2021, data = coal_data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -4087126 -3043657  -440501  3163443  4996399
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.179e+07  4.751e+06   2.481   0.0226 *
## MEDHHINC_2021 -1.346e+02  7.904e+01  -1.703   0.1049
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3175000 on 19 degrees of freedom
## Multiple R-squared:  0.1324, Adjusted R-squared:  0.08671
## F-statistic: 2.899 on 1 and 19 DF,  p-value: 0.1049

## `geom_smooth()` using formula = 'y ~ x'
```




```
##
## Call:
## lm(formula = PLNGENAN ~ MEDHHINC_2021, data = solar_data)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -15257  -8368  -6603  -4095 216336
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    2.339e+04  4.662e+03   5.016 6.68e-07 ***
## MEDHHINC_2021 -1.707e-01  8.583e-02  -1.989   0.0471 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 27570 on 707 degrees of freedom
##   (16 observations deleted due to missingness)
## Multiple R-squared:  0.005565,   Adjusted R-squared:  0.004159
## F-statistic: 3.957 on 1 and 707 DF,  p-value: 0.04707
```

```
## `geom_smooth()` using formula = 'y ~ x'
```




```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##    2021    2021    2021    2031    2049    2049    1161

##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    2021    2021    2021    2031    2049    2049

##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##    1912    2007    2014    2005    2017    2021      15
```

```
##   FIPS_Code           CNTYNAME             SEQGEN           YEAR
##  Length:1175        Length:1175        Min.   :13538   Min.   :2021
##  Class :character   Class :character   1st Qu.:13836   1st Qu.:2021
##  Mode  :character   Mode  :character   Median :14133   Median :2021
##                                        Mean   :14133   Mean   :2021
##                                        3rd Qu.:14434   3rd Qu.:2021
##                                        Max.   :14727   Max.   :2021
##
##     PNAME              GENID             GENSTAT             GENYRONL
##  Length:1175        Length:1175        Length:1175        Min.   :1912
##  Class :character   Class :character   Class :character   1st Qu.:2007
##  Mode  :character   Mode  :character   Mode  :character   Median :2014
##                                                           Mean   :2005
##                                                           3rd Qu.:2017
##                                                           Max.   :2021
##
##  NAMEPCAP_plant      PLNGENAN           PLCO2AN          COALFLAG
##  Min.   :   1.0   Min.   :  -1884   Min.   :      0   Length:1175
##  1st Qu.:   4.0   1st Qu.:   4427   1st Qu.:     78   Class :character
```

```
##  Median :   5.0   Median :    9001   Median :   1155   Mode  :character
##  Mean   : 141.6   Mean   :  388933   Mean   : 588841
##  3rd Qu.:  19.7   3rd Qu.:   27526   3rd Qu.: 247693
##  Max.   :2558.2   Max.   :19662114   Max.   :7610767
##                   NA's   :22         NA's   :851
##  MEDHHINC_2021   Unemployment_rate_2000 Unemployment_rate_2001
##  Min.   :38613   Min.   :2.400          Min.   :3.000
##  1st Qu.:44477   1st Qu.:3.600          1st Qu.:5.600
##  Median :52463   Median :4.300          Median :6.400
##  Mean   :54738   Mean   :4.526          Mean   :6.666
##  3rd Qu.:60228   3rd Qu.:5.500          3rd Qu.:8.100
##  Max.   :91558   Max.   :8.900          Max.   :9.800
##
##  Unemployment_rate_2002 Unemployment_rate_2003 Unemployment_rate_2004
##  Min.   : 3.100         Min.   : 3.400         Min.   : 3.100
##  1st Qu.: 6.600         1st Qu.: 6.400         1st Qu.: 5.500
##  Median : 7.600         Median : 7.500         Median : 6.400
##  Mean   : 7.829         Mean   : 7.444         Mean   : 6.548
##  3rd Qu.: 9.300         3rd Qu.: 8.800         3rd Qu.: 7.600
##  Max.   :11.400         Max.   :11.900         Max.   :11.400
##
##  Unemployment_rate_2005 Unemployment_rate_2006 Unemployment_rate_2007
##  Min.   : 3.30          Min.   :3.600          Min.   : 3.400
##  1st Qu.: 5.20          1st Qu.:4.800          1st Qu.: 4.800
##  Median : 6.20          Median :5.800          Median : 5.700
##  Mean   : 6.27          Mean   :5.783          Mean   : 5.712
##  3rd Qu.: 7.30          3rd Qu.:6.600          3rd Qu.: 6.700
##  Max.   :10.30          Max.   :9.500          Max.   :10.800
##
##  Unemployment_rate_2008 Unemployment_rate_2009 Unemployment_rate_2010
##  Min.   : 4.000         Min.   : 6.7           Min.   : 6.50
##  1st Qu.: 5.850         1st Qu.:10.4           1st Qu.:10.00
##  Median : 6.800         Median :11.8           Median :11.90
##  Mean   : 6.929         Mean   :12.0           Mean   :11.78
##  3rd Qu.: 7.800         3rd Qu.:13.7           3rd Qu.:13.40
##  Max.   :11.000         Max.   :17.8           Max.   :17.50
##
##  Unemployment_rate_2011 Unemployment_rate_2012 Unemployment_rate_2013
##  Min.   : 6.50          Min.   : 6.30          Min.   : 5.200
##  1st Qu.: 9.80          1st Qu.: 9.30          1st Qu.: 7.800
##  Median :11.30          Median :10.70          Median : 8.400
##  Mean   :11.42          Mean   :10.72          Mean   : 8.857
##  3rd Qu.:12.45          3rd Qu.:11.80          3rd Qu.: 9.900
##  Max.   :18.10          Max.   :18.10          Max.   :14.900
##
##  Unemployment_rate_2014 Unemployment_rate_2015 Unemployment_rate_2016
##  Min.   : 4.50          Min.   : 4.300         Min.   :3.900
##  1st Qu.: 6.00          1st Qu.: 5.500         1st Qu.:4.900
##  Median : 6.70          Median : 6.000         Median :5.500
##  Mean   : 6.98          Mean   : 6.433         Mean   :5.781
##  3rd Qu.: 7.80          3rd Qu.: 7.200         3rd Qu.:6.600
##  Max.   :13.00          Max.   :11.900         Max.   :9.500
##
##  Unemployment_rate_2017 Unemployment_rate_2018 Unemployment_rate_2019
```

```
##  Min.   :3.400          Min.   :3.000          Min.   :3.000
##  1st Qu.:4.300          1st Qu.:3.800          1st Qu.:3.700
##  Median :4.800          Median :4.300          Median :4.000
##  Mean   :5.055          Mean   :4.471          Mean   :4.267
##  3rd Qu.:5.800          3rd Qu.:5.100          3rd Qu.:4.800
##  Max.   :8.400          Max.   :8.300          Max.   :8.500
##
##  Unemployment_rate_2020 Unemployment_rate_2021 Unemployment_rate_2022
##  Min.   : 5.20          Min.   :3.600          Min.   :3.000
##  1st Qu.: 6.30          1st Qu.:4.500          1st Qu.:3.500
##  Median : 7.20          Median :5.100          Median :3.800
##  Mean   : 7.34          Mean   :5.434          Mean   :4.079
##  3rd Qu.: 8.20          3rd Qu.:6.300          3rd Qu.:4.500
##  Max.   :11.90          Max.   :9.500          Max.   :6.500
##
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    1912    2007    2014    2005    2017    2021
```

```
##    Length     Class      Mode
##      1175 character character
```

```
## [1] "01-01-2017" "01-01-1982" "01-01-2016" "01-01-2016" "01-01-2015"
## [6] "01-01-2016"
```

```
## [1] FALSE
```

```
## [1] TRUE
```

```
## [1] TRUE
```

```
## [1] 48
```

```
##  [1] "Shannon Farm"                     "Turkey Branch Solar LLC"
##  [3] "Stone Solar"                      "Ajax Solar"
##  [5] "ZV Solar 1"                       "Progress Solar III, LLC"
##  [7] "South Robeson Farm"               "Railroad Farm"
##  [9] "Innovative Solar 55"              "Panda Solar NC 1, LLC"
## [11] "Railroad Farm 2"                  "Walters Solar (FLS 260)"
## [13] "Bo Biggs Solar"                   "Maxton Solar, LLC"
## [15] "Deep Branch Farm"                 "Panda Solar NC 3, LLC"
## [17] "Innovative Solar 47"              "Baltimore Church Solar, LLC"
## [19] "Lumberton"                        "McGoogan Farm, LLC"
## [21] "Pate Farm"                        "W H Weatherspoon"
## [23] "Railroad Solar Farm, LLC"         "Barker Solar, LLC"
## [25] "Innovative Solar 48"              "FLS Solar 170, LLC"
## [27] "St. Pauls Solar 1, LLC"           "Floyd Solar, LLC"
## [29] "Innovative Solar 44"              "Jersey Holdings"
## [31] "Broadridge Solar, LLC"            "Robeson County LFG to Energy"
## [33] "Watts Farm"                       "FLS Solar 200, LLC"
## [35] "Holly Swamp Solar, LLC"           "ZV Solar 2, LLC"
## [37] "Three Bridge"                     "Progress Solar II, LLC"
## [39] "Fairmont-FLS 100"                 "McCallum Farm"
## [41] "Laurinburg Solar, LLC (Heelstone)" "Panda Solar NC 2, LLC"
## [43] "St. Pauls Solar 2"                "ZV Solar 3, LLC"
```

```
##  [1] "1"     "TURKY" "PV1"   "PROG3" "SROB"  "IS044" "20002" "FLS1"  "BB08"
## [10] "MSPV"  "20011" "IS047" "GEN1"  "GT3"   "NCRAL" "BARK"  "IS048" "GT2"
```

```
## [19] "BROAD" "GEN01" "13502" "ZV204" "PGR01" "PROG2" "GEN 1" "LAUR"  "20003"
## [28] "GT1"   "STPAU" "GEN02" "ZV3"   "GT4"
```

```
## [1] TRUE
```

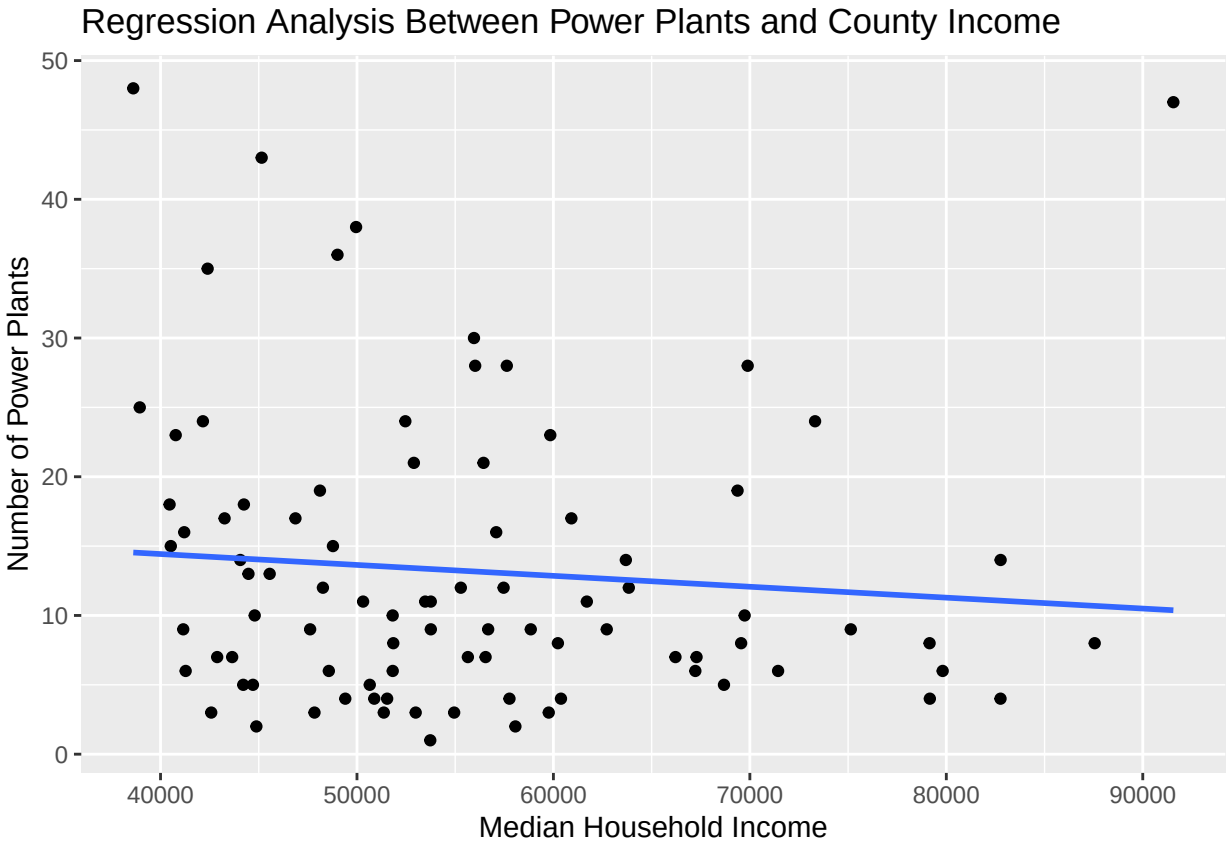
```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   38613   44477   52463   54735   59841   91558
```

#Question 1: Are there more/larger power plants in counties with lower average socioeconomic status?

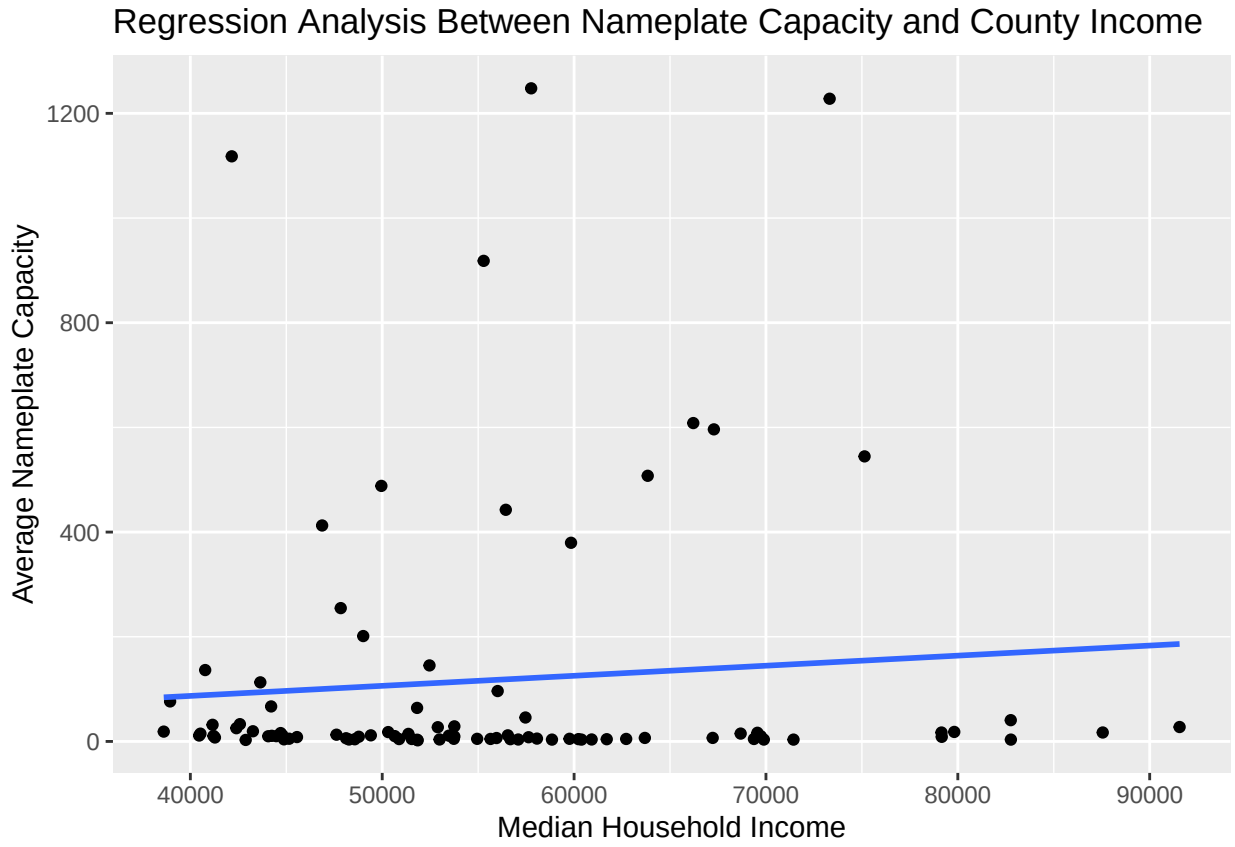
[Description here]

```
##
## Call:
## lm(formula = Num_Plants ~ MEDHHINC_2021, data = county_plants_df)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -12.348  -7.174  -2.691   3.609  36.624
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.757e+01  5.241e+00   3.352  0.00119 **
## MEDHHINC_2021 -7.859e-05  9.217e-05  -0.853  0.39619
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 10.42 on 87 degrees of freedom
## Multiple R-squared:  0.008287,   Adjusted R-squared:  -0.003112
## F-statistic: 0.727 on 1 and 87 DF,  p-value: 0.3962
```

```
## `geom_smooth()` using formula = 'y ~ x'
```

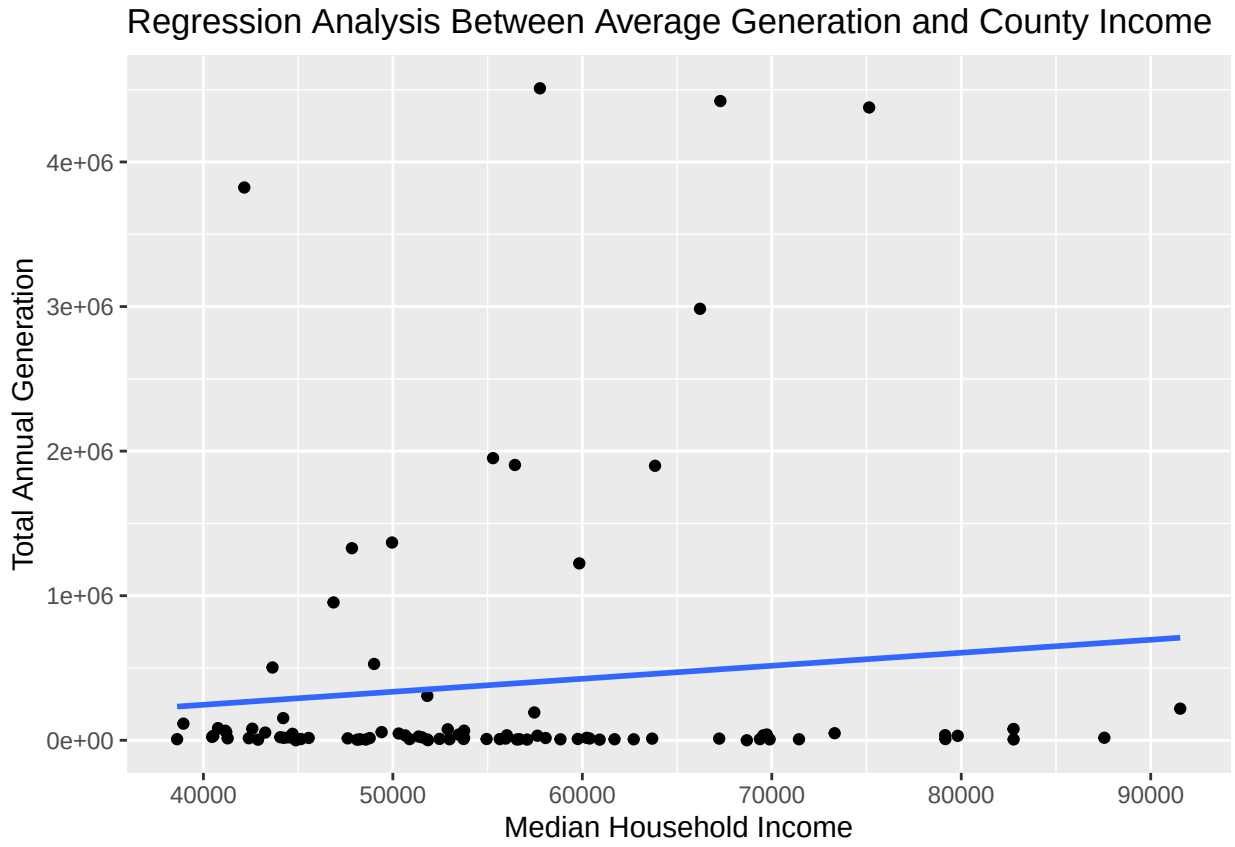



```
## `geom_smooth()` using formula = 'y ~ x'
```



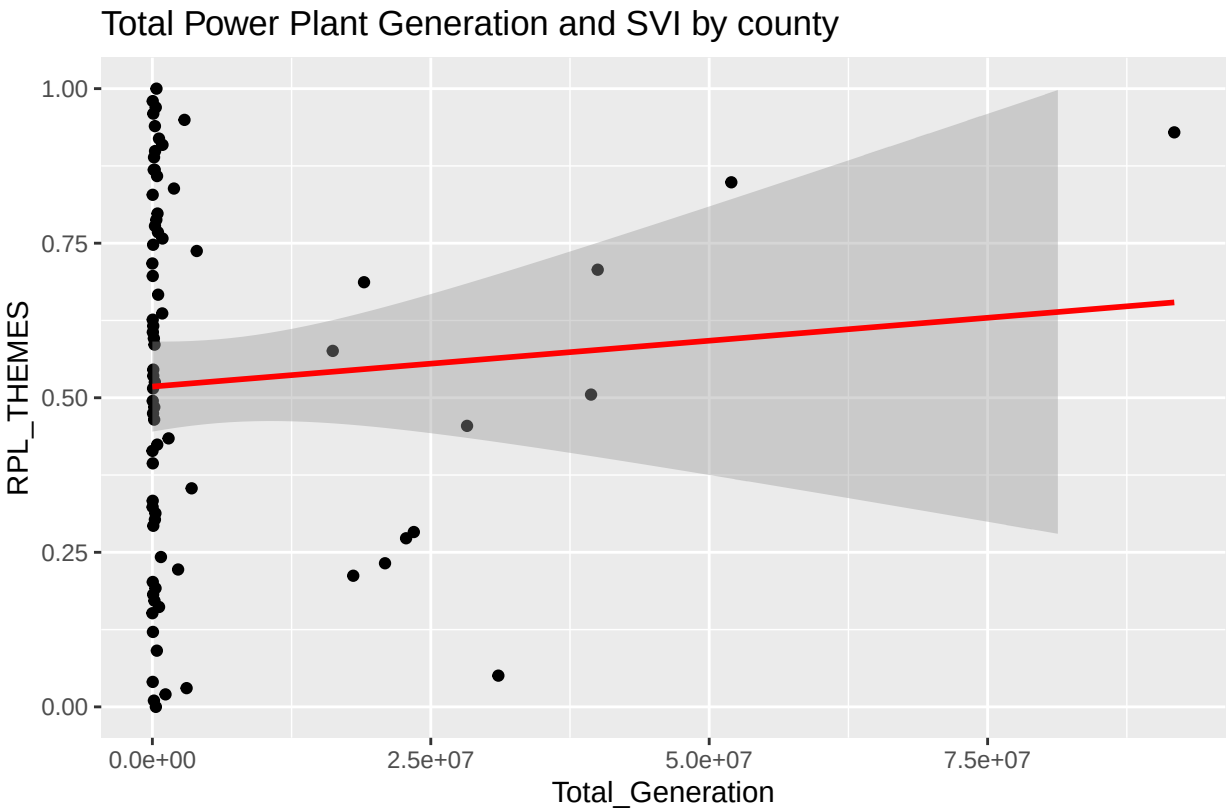

```
##
## Call:
## lm(formula = PLNGENAN ~ MEDHHINC_2021, data = merged_data_generation)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -655989 -408368 -312983 -219659 4103780
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -1.142e+05  5.008e+05  -0.228     0.82
## MEDHHINC_2021  8.997e+00  8.807e+00   1.022     0.31
##
## Residual standard error: 995200 on 87 degrees of freedom
## Multiple R-squared:  0.01185,    Adjusted R-squared:  0.0004951
## F-statistic: 1.044 on 1 and 87 DF,  p-value: 0.3098
```

```
## `geom_smooth()` using formula = 'y ~ x'
```




```
##
## Call:
## lm(formula = Total_Generation ~ RPL_THEMES, data = Generation_SVI)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -7359828 -6145039 -5111880 -3741737 84555556
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  3820133    3462723   1.103    0.274
## RPL_THEMES   3643958    5744791   0.634    0.528
##
## Residual standard error: 14690000 on 74 degrees of freedom
##   (13 observations deleted due to missingness)
## Multiple R-squared:  0.005408,   Adjusted R-squared:  -0.008033
## F-statistic: 0.4023 on 1 and 74 DF,  p-value: 0.5278
```

```
## `geom_smooth()` using formula = 'y ~ x'
```



ionship between the Social Vulnerability Index (SVI) and 2021 annual generation across all power plants in a county.