Knitted R Markdown Document for Final Project

Parker Johnson

2023-06-04

Loading in the dataset. Note that this is from the 2017 dataset retrieved from https://www.fractracker.org/2017/03/34-states-active-drilling-2016/. This dataset wasn't used for the final analysis as 2021 data was found that was more updated, but since it contains point-level data on the location of oil wells in Pennsylvania, the code for aggregating this data to the county-level could still prove to be useful to anyone hoping to do further analysis using frequency or proximity to oil wells.

```
PA_map <- ggmap(get_map(c(left = -80.95, bottom = 39.55,
                          right = -74.15, top = 42.52), source = "stamen"))
```

```
## i Map tiles by Stamen Design, under CC BY 3.0. Data by OpenStreetMap, under ODbL.
```

```
no_TX_wells <- read_csv("US_WellsFacilities2016/US_OGfacil_2016_noTX.csv")
```

```
## Rows: 794726 Columns: 8
```

```
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr (5): FacilityNa, Operator, API, State, Type
## dbl (3): Latitude, Longitude, Count
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
PA_wells <- no_TX_wells %>%
  filter(State == "PA") %>%
  filter(Longitude < 0 & Longitude > -100)

PaCounties = st_read("PACounty2023_04/PaCounty2023_04.shp") %>% as("Spatial")
```

```
## Reading layer `PaCounty2023_04' from data source
##   `C:\Users\pjohn\OneDrive\Documents\Stat310 Final Project\PaCounty2023_04\PaCounty2023_04.shp'
##   using driver `ESRI Shapefile'
## Simple feature collection with 67 features and 25 fields
## Geometry type: MULTIPOLYGON
## Dimension:     XY
## Bounding box:  xmin: -8963377 ymin: 4825316 xmax: -8314404 ymax: 5201413
## Projected CRS: WGS 84 / Pseudo-Mercator
```

```
PaCounties <- spTransform(PaCounties, "+proj=longlat +datum=NAD83 +no_defs")
```

```
## Warning: PROJ support is provided by the sf and terra packages among others
```

Pulling census data.

```
PA_county_tracts <- counties("PA")
```

```
## Retrieving data for the year 2021
```

```
##   |                                                                              |
```

```
PA_county_tracts <- as_Spatial(PA_county_tracts)

PA_census <- get_acs(geography = "county",
                         variables = c("B01003_001E", "B01001_002E",
                                       "B01001_026E", "B23025_003E",
                                       "B23025_004E", "B23025_005E",
                                       "B15003_022E", "B15003_017E",
                                       "B15003_023E", "B19013_001E",
                                       "B15003_001E", "B19058_001E",
                                       "B19058_002E", "B19058_003E",
                                       "B02001_001E", "B02001_002E",
                                       "B02001_003E", "B02001_004E",
                                       "B02001_005E","B02001_006E",
                                       "B02001_007E", "B02001_008E"), year = 2021,
                         state = "PA",
                         geometry = F)
```

```
## Getting data from the 2017-2021 5-year ACS
```

```
code_book_PA <- rbind(c("B01003_001", "total_pop"),
                      c("B01001_002", "total_male"),
                      c("B01001_026", "total_female"),
                      c("B23025_003", "total_labor_force"),
                      c("B23025_004", "total_employed"),
                      c("B23025_005", "total_unemployed"),
                      c("B15003_022", "ea_bachelors"),
                      c("B15003_017", "ea_hsdiplomma"),
                      c("B15003_023", "ea_masters"),
                      c("B19013_001", "med_hh_income"),
                      c("B15003_001", "total_ea"),
                      c("B19058_001", "total_foodstamp"),
                      c("B19058_002", "foodstamp_yes"),
                      c("B19058_003", "foodstamp_no"),
                      c("B02001_001", "total_race"),
                      c("B02001_002", "total_white"),
                      c("B02001_003", "total_black"),
                      c("B02001_004", "total_native"),
                      c("B02001_005", "total_asian"),
                      c("B02001_006", "total_islander"),
```

```
                  c("B02001_007", "total_otherrace"),
                  c("B02001_008", "total_twoormore"))

code_book_PA <- as.data.frame(code_book_PA)
colnames(code_book_PA) <- c("variable", "var_name")

PA_census <- left_join(PA_census, code_book_PA)
```

```
## Joining, by = "variable"
```

```
#format the data so there is a row for each census tract and column for every variable
PA_acs_data <- maditr::dcast(PA_census, GEOID ~ var_name,
                             value.var = "estimate", fun.aggregate = NULL)

#subset to only data that is in the shape file
PA_acs_data <- PA_acs_data[which(PA_acs_data$GEOID %in% PA_county_tracts$GEOID),]

#new variable for percent unemployed
PA_acs_data$perc_unemployed <- PA_acs_data$total_unemployed/
  PA_acs_data$total_labor_force
PA_acs_data$perc_employed <- PA_acs_data$total_employed/
  PA_acs_data$total_labor_force

PA_acs_data$perc_foodstamp <- PA_acs_data$foodstamp_yes/PA_acs_data$total_foodstamp
PA_acs_data$perc_nofoodstamp <- PA_acs_data$foodstamp_no/PA_acs_data$total_foodstamp

PA_acs_data$perc_female <- PA_acs_data$total_female/PA_acs_data$total_pop
PA_acs_data$perc_male <- PA_acs_data$total_male/PA_acs_data$total_pop

PA_acs_data$perc_white <- PA_acs_data$total_white/PA_acs_data$total_race


# Combine dataset with spatial dataset by GEOID
PA_county_tracts@data <- left_join(PA_county_tracts@data, PA_acs_data)
```

```
## Joining, by = "GEOID"
```

```
proj4string(PA_county_tracts)
```

```
## [1] "+proj=longlat +datum=NAD83 +no_defs"
```

```
proj4string(PaCounties) #check projection
```

```
## [1] "+proj=longlat +datum=NAD83 +no_defs"
```

```
PaCounties <- spTransform(PaCounties, proj4string(PA_county_tracts))

#overlay the two
over_data <- over(PA_county_tracts, PaCounties)
```

```
## Warning in RGEOSBinPredFunc(spgeom1, spgeom2, byid, func): spgeom1 and spgeom2
## have different proj4 strings
```

```
PA_county_tracts@data <- cbind(PA_county_tracts@data, over_data)
```

Aggregating the outdated dataset to the county level, and adding 1 to the number of oil wells in each county to allow for the log transformation to work later.

```
PA_wells2 <- PA_wells
coordinates(PA_wells2) <- ~Longitude + Latitude
proj4string(PA_wells2) <- "+proj=longlat +datum=NAD83 +no_defs"

over_dat <- over(PA_wells2, PaCounties)
length(which(is.na(over_dat$COUNTY_NAM))) #126 points outside of the counties
```

```
## [1] 126
```

```
agg_dat <- plyr::count(over_dat, c('COUNTY_NAM'))
agg_dat$COUNTY_NAM <- as.factor(agg_dat$COUNTY_NAM)
colnames(agg_dat) <- c("COUNTY_NAM", "num_wells")
agg_dat$COUNTY_NAM <- str_to_title(agg_dat$COUNTY_NAM)
agg_dat <- agg_dat %>%
  filter(is.na(COUNTY_NAM) == FALSE)

PA_county_tracts@data <- PA_county_tracts@data %>%
  select(-c(COUNTY_NAM))

PA_county_tracts@data <- left_join(PA_county_tracts@data, agg_dat,
                                   by = c("NAME" = "COUNTY_NAM"))

PA_county_tracts@data["num_wells"][is.na(PA_county_tracts@data["num_wells"])] <- 0
PA_county_tracts@data["num_wells"] <- PA_county_tracts@data["num_wells"] + 1
PA_county_tracts$well_rate <- PA_county_tracts$num_wells / PA_county_tracts$total_pop
```

Reading in the cancer, asthma, and updated data for the number of conventional and historic oil wells in Pennsylvania, all aggregated to the county level. Also adding the indicator variable for if a county is in eastern Pennsylvania or not, and standardizing the total population in each county to allow for the spatial Durbin models to work later.

```
Pennsylvania_Cancer <- read_csv("Pennsylvania_Cancer.csv")
```

```
## New names:
## Rows: 97 Columns: 13
## -- Column specification
## -------------------------------------------------------- Delimiter: "," chr
## (5): County, CI*Rank([rank note]), Lower CI (CI*Rank), Upper CI (CI*Rank... dbl
## (8): FIPS, Age-Adjusted Incidence Rate([rate note]) - cases per 100,000,...
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## * `Lower 95% Confidence Interval` -> `Lower 95% Confidence Interval...4`
## * `Upper 95% Confidence Interval` -> `Upper 95% Confidence Interval...5`
## * `Lower 95% Confidence Interval` -> `Lower 95% Confidence Interval...12`
## * `Upper 95% Confidence Interval` -> `Upper 95% Confidence Interval...13`
```

```r
Pennsylvania_Cancer <- Pennsylvania_Cancer[3:69, ]
Pennsylvania_Cancer <- Pennsylvania_Cancer %>%
  mutate(County = str_sub(County,1,-11)) %>%
  rename("cancer_rate" =
           "Age-Adjusted Incidence Rate([rate note]) - cases per 100,000") %>%
  select(County, cancer_rate)

PA_county_tracts@data <- left_join(PA_county_tracts@data, Pennsylvania_Cancer,
                                   by = c("NAME" = "County"))


counties <- c("Erie", "Crawford", "Mercer", "Lawrence", "Beaver", "Washington",
              "Greene", "Fayette", "Westmoreland", "Allegheny", "Butler",
              "Armstrong", "Clarion", "Venango", "Forest", "Warren", "McKean",
              "Elk", "Cameron", "Jefferson", "Clearfield", "Indiana", "Cambria",
              "Somerset", "Bedford", "Blair", "Centre", "Fulton", "Huntingdon",
              "Mifflin", "Juniata", "Snyder", "Union", "Clinton", "Potter",
              "Tioga", "Lycoming", "Montour", "Northumberland", "Columbia",
              "Sullivan", "Bradford", "Susquehanna", "Wyoming", "Luzerne",
              "Lackawanna", "Wayne", "Pike", "Monroe", "Carbon", "Schuylkill",
              "Northampton", "Lehigh", "Bucks", "Montgomery", "Philadelphia",
              "Delaware", "Chester", "Berks", "Lebanon", "Dauphin", "Perry",
              "Cumberland", "Franklin", "Adams", "York", "Lancaster")
asthma_ED_visits <- c(27.8, 15.9, 30, 22, 13.9, 28.2, 22.3, 24.4, 21.3, 56.1,
                      21.2, 19, 10.4, 19, 13.7, 23.6, 22.3, 15.8, 7.6, 13.1,
                      15.3, 15.3, 25.4, 6.3, 12.2, 29.4, 30.9, 17.7, 18.5, 22.1,
                      17.5, 19.5, 21, 13.2, 12.1, 24.5, 37.8, 26.7, 27.9, 26.4,
                      8.1, 12.7, 13.6, 11.3, 55.8, 35.2, 33.3, 93, 108, 32.1,
                      25.9, 57, 90.6, 34.1, 39.3, 111.9, 70.6, 41.4, 136.8,
                      83.6, 71.7, 37.3, 36.4, 25.2, 23.1, 31.3, 32.8)
number_of_wells <- c(2857, 3076, 3213, 174, 143, 1896, 2149, 3071, 5765, 1132,
                     1285, 8285, 3710, 8184, 5416, 12514, 11102, 2910, 49, 5371,
                     4156, 11285, 565, 72, 39, 0, 707, 0, 2, 0, 0, 0, 0, 503,
                     1187, 105, 4, 0, 0, 0, 0, 12, 6, 1, 1, 1, 2, 2, 0, 0, 2,
                     0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 1, 0)
asthma <- data.frame(counties, asthma_ED_visits, number_of_wells)

PA_county_tracts@data <- left_join(PA_county_tracts@data, asthma,
                                   by = c("NAME" = "counties"))
PA_county_tracts$total_asthma <- (PA_county_tracts$total_pop / 10000) *
  PA_county_tracts$asthma_ED_visits
PA_county_tracts$total_cancer <- (PA_county_tracts$total_pop / 10000) *
  PA_county_tracts$cancer_rate

PA_county_tracts@data["number_of_wells"][is.na(PA_county_tracts@data["number_of_wells"])] <- 0
PA_county_tracts@data["number_of_wells"] <- PA_county_tracts@data["number_of_wells"] + 1
PA_county_tracts$well_rate_v2 <- PA_county_tracts$number_of_wells /
  PA_county_tracts$total_pop

#Adding the indicator variable
PA_county_tracts$is_Eastern_PA <- ifelse(PA_county_tracts$number_of_wells <= 3, 1, 0)
PA_county_tracts@data["is_Eastern_PA"][PA_county_tracts@data["NAME"] == "Blair"] <- 0
PA_county_tracts@data["is_Eastern_PA"][PA_county_tracts@data["NAME"] == "Sullivan"] <- 0
```

```
#Standardizing population for the spatial Durbin model used later
PA_county_tracts$total_pop_standardized <- PA_county_tracts$total_pop / 1596865
```

Creating plots of the log number of oil wells, the median household income, cancer rate, and asthma ED visit rate based on Pennsylvania counties.

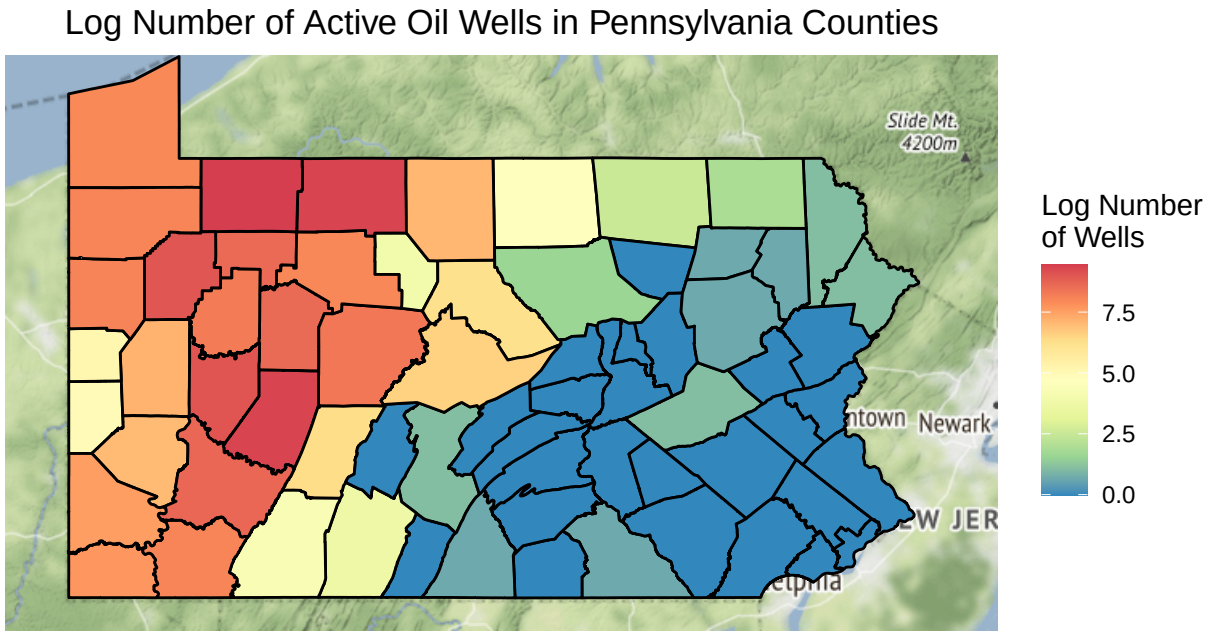
```
PA_tracts_tidy <- broom::tidy(PA_county_tracts)
```

```
## Regions defined for each Polygons
```

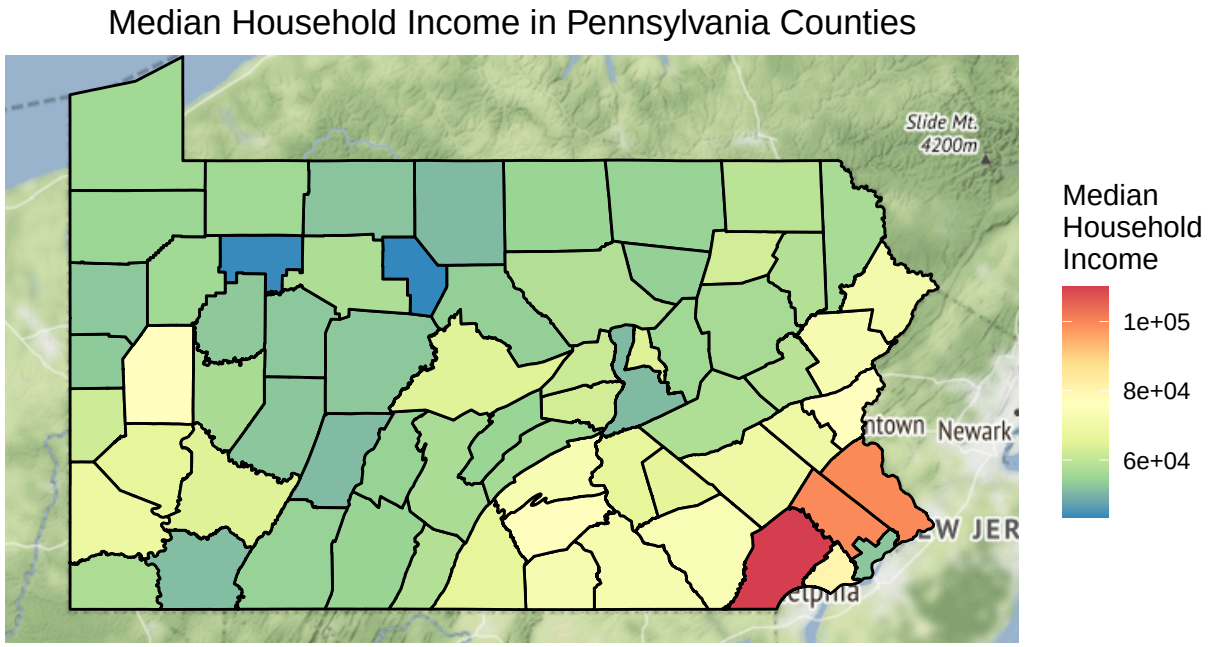
```
PA_county_tracts$id <- row.names(PA_county_tracts) #need to join data
PA_tracts_tidy <- left_join(PA_tracts_tidy, PA_county_tracts@data)
```

```
## Joining, by = "id"
```

```
PA_map + geom_polygon(data= PA_tracts_tidy, aes(x = long, y = lat,
                                group = group, fill = log(number_of_wells)),
                      col = "black") +
  labs(fill = "Log Number\nof Wells",
       title = "Log Number of Active Oil Wells in Pennsylvania Counties") +
  theme(plot.title = element_text(hjust = 0.5)) +
  scale_fill_distiller(palette = "Spectral") +
      theme(axis.title = element_blank(),
        axis.text = element_blank(),
        axis.ticks = element_blank())
```

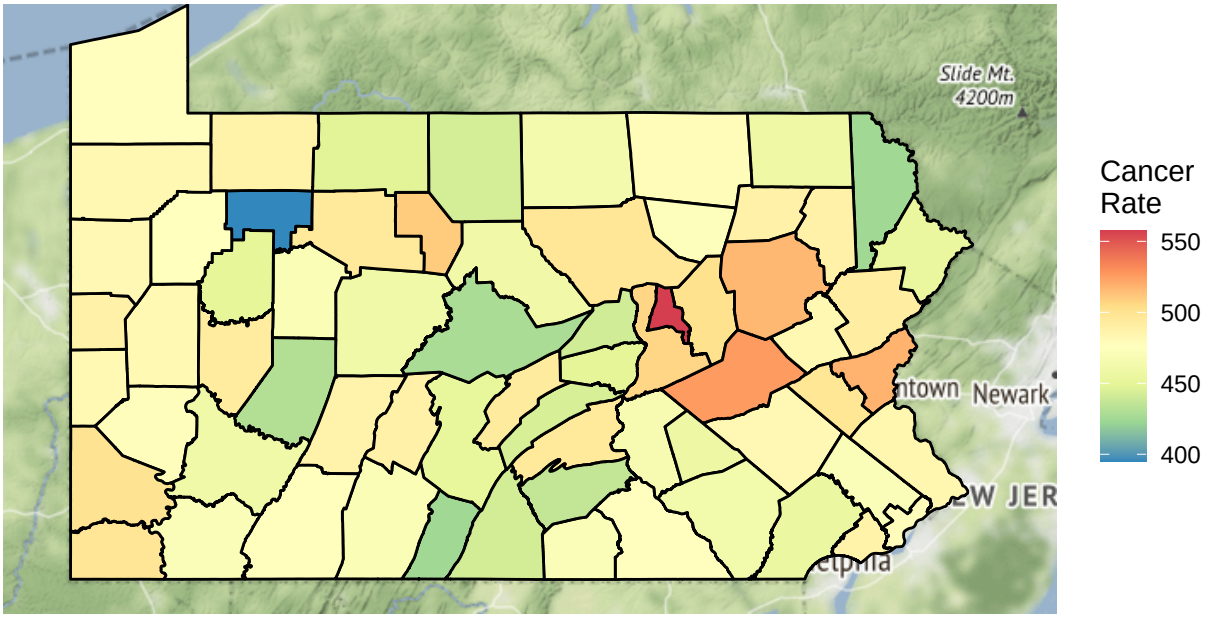


```
PA_map + geom_polygon(data= PA_tracts_tidy, aes(x = long, y = lat,
                                group = group, fill = med_hh_income),
                        col = "black") +
  labs(fill = "Median\nHousehold\nIncome",
       title = "Median Household Income in Pennsylvania Counties") +
  theme(plot.title = element_text(hjust = 0.5)) +
  scale_fill_distiller(palette = "Spectral") +
        theme(axis.title = element_blank(),
        axis.text = element_blank(),
        axis.ticks = element_blank())
```

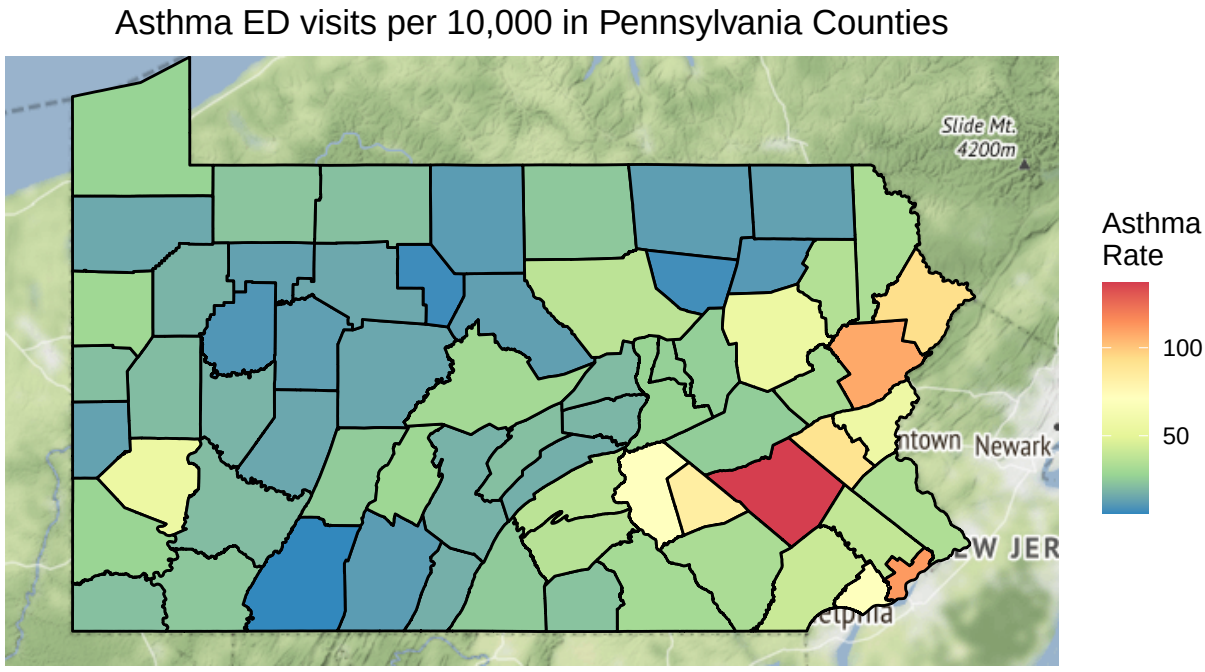



```
PA_map + geom_polygon(data= PA_tracts_tidy, aes(x = long, y = lat,
                                group = group, fill = cancer_rate), col = "black") +
  labs(fill = "Cancer\nRate",
       title = "Cancer Incidents per 10,000 in Pennsylvania Counties") +
  theme(plot.title = element_text(hjust = 0.5)) +
  scale_fill_distiller(palette = "Spectral") +
        theme(axis.title = element_blank(),
        axis.text = element_blank(),
        axis.ticks = element_blank())
```

Cancer Incidents per 10,000 in Pennsylvania Counties




```
PA_map + geom_polygon(data= PA_tracts_tidy, aes(x = long, y = lat,
                                group = group, fill = asthma_ED_visits),
                      col = "black") +
  labs(fill = "Asthma\nRate",
       title = "Asthma ED visits per 10,000 in Pennsylvania Counties") +
  theme(plot.title = element_text(hjust = 0.5)) +
  scale_fill_distiller(palette = "Spectral") +
        theme(axis.title = element_blank(),
        axis.text = element_blank(),
        axis.ticks = element_blank())
```



Specifying the initial 5 neighborhood matrices.

```
pennsylvania_nb <- poly2nb(PA_county_tracts)
B_list <- nb2listw(pennsylvania_nb, style="B") #binary
W_list <- nb2listw(pennsylvania_nb, style="W") #row standardized

areal_knn3 <- knearneigh(coordinates(PA_county_tracts), k=3)
#convert back to nb
areal_knn_nb3 <- knn2nb(areal_knn3)
#choose row-standardized or binary
W_list_knn3 <- nb2listw(areal_knn_nb3, style="W")
#average for row-standardized

areal_knn5 <- knearneigh(coordinates(PA_county_tracts), k=5)
areal_knn_nb5 <- knn2nb(areal_knn5)
W_list_knn5 <- nb2listw(areal_knn_nb5, style="W")

areal_knn10 <- knearneigh(coordinates(PA_county_tracts), k=10)
areal_knn_nb10 <- knn2nb(areal_knn10)
W_list_knn10 <- nb2listw(areal_knn_nb10, style="W")
```

Making the lm model for cancer rate, and checking the VIF for the linear model.

```
cancer_lm <- lm(cancer_rate ~ perc_white + perc_unemployed + med_hh_income +
                  total_pop + log(number_of_wells) + is_Eastern_PA +
                  I(log(number_of_wells) * is_Eastern_PA),
```

```
               data = PA_county_tracts@data)
summary(cancer_lm)
```

```
##
## Call:
## lm(formula = cancer_rate ~ perc_white + perc_unemployed + med_hh_income +
##     total_pop + log(number_of_wells) + is_Eastern_PA + I(log(number_of_wells) *
##     is_Eastern_PA), data = PA_county_tracts@data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -60.946 -17.574   1.694  16.594  87.218
##
## Coefficients:
##                                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)                           3.528e+02  7.614e+01   4.633 2.04e-05
## perc_white                            1.154e+02  6.568e+01   1.757   0.0841
## perc_unemployed                       4.937e+02  3.488e+02   1.415   0.1622
## med_hh_income                        -8.534e-05  3.659e-04  -0.233   0.8164
## total_pop                             2.921e-05  2.069e-05   1.412   0.1633
## log(number_of_wells)                 -2.344e+00  1.736e+00  -1.350   0.1822
## is_Eastern_PA                         3.169e+00  1.523e+01   0.208   0.8359
## I(log(number_of_wells) * is_Eastern_PA) -1.271e+01  1.327e+01  -0.958   0.3420
##
## (Intercept)                          ***
## perc_white                           .
## perc_unemployed
## med_hh_income
## total_pop
## log(number_of_wells)
## is_Eastern_PA
## I(log(number_of_wells) * is_Eastern_PA)
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 26.97 on 59 degrees of freedom
## Multiple R-squared:  0.09648,    Adjusted R-squared:  -0.01071
## F-statistic: 0.9001 on 7 and 59 DF,  p-value: 0.5126
```

```
library(car)
```

```
## Warning: package 'car' was built under R version 4.2.3
```

```
## Loading required package: carData
```

```
## Warning: package 'carData' was built under R version 4.2.3
```

```
##
## Attaching package: 'car'
```

```
## The following object is masked from 'package:dplyr':
##
##     recode
```

```
## The following object is masked from 'package:purrr':
##
##     some
```

```
vif(cancer_lm)
```

```
##                              perc_white                          perc_unemployed
##                                3.687980                                1.860221
##                           med_hh_income                               total_pop
##                                1.802330                                3.048814
##                     log(number_of_wells)                           is_Eastern_PA
##                                3.650169                                5.342124
## I(log(number_of_wells) * is_Eastern_PA)
##                                1.528325
```

```
#store the residuals to see if there is any spatial structure to the residuals
resid <- cancer_lm$residuals

#can join this to the SpatialPolygons
PA_county_tracts@data$resid <- resid
```

Moran's I tests for each neighborhood matrix on the linear model.

```
#set.seed(597)
#binary, spatial adjacency
mc <- moran(PA_county_tracts$resid, B_list, n = length(pennsylvania_nb),
            S0 = Szero(B_list))
mc$I
```

```
## [1] 0.1379875
```

```
moran.mc(x=PA_county_tracts$resid, listw=B_list, nsim=5000)
```

```
##
##  Monte-Carlo simulation of Moran I
##
## data:  PA_county_tracts$resid
## weights: B_list
## number of simulations + 1: 5001
##
## statistic = 0.13799, observed rank = 4887, p-value = 0.0228
## alternative hypothesis: greater
```

```
#row-standardized, spatial adjacency
mc <- moran(PA_county_tracts$resid, W_list, n = length(pennsylvania_nb),
            S0 = Szero(W_list))
mc$I
```

```
## [1] 0.1556063
```

```
moran.mc(x=PA_county_tracts$resid, listw=W_list, nsim=5000)
```

```
##
##  Monte-Carlo simulation of Moran I
##
## data:  PA_county_tracts$resid
## weights: W_list
## number of simulations + 1: 5001
##
## statistic = 0.15561, observed rank = 4926, p-value = 0.015
## alternative hypothesis: greater
```

```
#k-nn 3, row-standardized
mc <- moran(PA_county_tracts$resid, W_list_knn3, n = length(pennsylvania_nb),
            S0 = Szero(W_list_knn3))
mc$I
```

```
## [1] 0.1051032
```

```
moran.mc(x=PA_county_tracts$resid, listw=W_list_knn3, nsim=5000)
```

```
##
##  Monte-Carlo simulation of Moran I
##
## data:  PA_county_tracts$resid
## weights: W_list_knn3
## number of simulations + 1: 5001
##
## statistic = 0.1051, observed rank = 4512, p-value = 0.09778
## alternative hypothesis: greater
```

```
#k-nn 5, row-standardized
mc <- moran(PA_county_tracts$resid, W_list_knn5, n = length(pennsylvania_nb),
            S0 = Szero(W_list_knn5))
mc$I
```

```
## [1] 0.05201744
```

```
moran.mc(x=PA_county_tracts$resid, listw=W_list_knn5, nsim=5000)
```

```
##
##  Monte-Carlo simulation of Moran I
##
## data:  PA_county_tracts$resid
## weights: W_list_knn5
## number of simulations + 1: 5001
##
## statistic = 0.052017, observed rank = 4139, p-value = 0.1724
## alternative hypothesis: greater
```

```
#k-nn 10, row-standardized
mc <- moran(PA_county_tracts$resid, W_list_knn10, n = length(pennsylvania_nb),
            S0 = Szero(W_list_knn10))
mc$I
```

```
## [1] 0.08213889
```

```
moran.mc(x=PA_county_tracts$resid, listw=W_list_knn10, nsim=5000)
```

```
##
##  Monte-Carlo simulation of Moran I
##
## data:  PA_county_tracts$resid
## weights: W_list_knn10
## number of simulations + 1: 5001
##
## statistic = 0.082139, observed rank = 4855, p-value = 0.02919
## alternative hypothesis: greater
```

Creating the spatial lag models.

```
pennsylvania_lag_B <- lagsarlm(cancer_rate ~ perc_white + perc_unemployed +
                                 med_hh_income + total_pop +
                                 log(number_of_wells) + is_Eastern_PA +
                                 I(log(number_of_wells) * is_Eastern_PA),
                               data = PA_county_tracts, listw = B_list)
summary(pennsylvania_lag_B)
```

```
##
## Call:lagsarlm(formula = cancer_rate ~ perc_white + perc_unemployed +
##     med_hh_income + total_pop + log(number_of_wells) + is_Eastern_PA +
##     I(log(number_of_wells) * is_Eastern_PA), data = PA_county_tracts,
##     listw = B_list)
##
## Residuals:
##         Min          1Q      Median          3Q         Max
## -56.9365188 -16.9464643   0.0036152  15.2476669  91.4396547
##
## Type: lag
## Coefficients: (asymptotic standard errors)
##                                         Estimate  Std. Error z value
## (Intercept)                           3.3840e+02  7.3075e+01  4.6309
## perc_white                            1.1741e+02  6.1340e+01  1.9140
## perc_unemployed                       5.3190e+02  3.2922e+02  1.6156
## med_hh_income                        -5.2268e-05  3.4334e-04 -0.1522
## total_pop                             3.0004e-05  1.9331e-05  1.5522
## log(number_of_wells)                 -2.3979e+00  1.6206e+00 -1.4796
## is_Eastern_PA                         2.8840e+00  1.4209e+01  0.2030
## I(log(number_of_wells) * is_Eastern_PA) -1.3267e+01  1.2412e+01 -1.0689
##                                         Pr(>|z|)
## (Intercept)                           3.641e-06
## perc_white                              0.05562
```

```
## perc_unemployed                            0.10618
## med_hh_income                              0.87900
## total_pop                                  0.12062
## log(number_of_wells)                       0.13897
## is_Eastern_PA                              0.83916
## I(log(number_of_wells) * is_Eastern_PA)    0.28512
##
## Rho: 0.0036117, LR test value: 0.78556, p-value: 0.37545
## Asymptotic standard error: 0.004099
##     z-value: 0.88113, p-value: 0.37825
## Wald statistic: 0.77638, p-value: 0.37825
##
## Log likelihood: -311.1524 for lag model
## ML residual variance (sigma squared): 632.82, (sigma: 25.156)
## Number of observations: 67
## Number of parameters estimated: 10
## AIC: 642.3, (AIC for lm: 641.09)
## LM test for residual autocorrelation
## test value: 2.4922, p-value: 0.11441
```

```
pennsylvania_lag_W <- lagsarlm(cancer_rate ~ perc_white + perc_unemployed +
                                 med_hh_income + total_pop +
                                 log(number_of_wells) + is_Eastern_PA +
                                 I(log(number_of_wells) * is_Eastern_PA),
                               data = PA_county_tracts, listw = W_list)
summary(pennsylvania_lag_W)
```

```
##
## Call:lagsarlm(formula = cancer_rate ~ perc_white + perc_unemployed +
##     med_hh_income + total_pop + log(number_of_wells) + is_Eastern_PA +
##     I(log(number_of_wells) * is_Eastern_PA), data = PA_county_tracts,
##     listw = W_list)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -51.9417 -16.8188   2.3809  17.2226  75.9742
##
## Type: lag
## Coefficients: (asymptotic standard errors)
##                                             Estimate  Std. Error z value
## (Intercept)                                1.6631e+02  9.9224e+01  1.6761
## perc_white                                 1.1581e+02  5.8199e+01  1.9898
## perc_unemployed                            4.2533e+02  3.0894e+02  1.3768
## med_hh_income                             -8.0137e-05  3.2410e-04 -0.2473
## total_pop                                  2.7673e-05  1.8326e-05  1.5100
## log(number_of_wells)                      -2.0175e+00  1.5520e+00 -1.2999
## is_Eastern_PA                              2.5396e+00  1.3510e+01  0.1880
## I(log(number_of_wells) * is_Eastern_PA) -9.2978e+00  1.1759e+01 -0.7907
##                                            Pr(>|z|)
## (Intercept)                                 0.09372
## perc_white                                  0.04661
## perc_unemployed                             0.16859
## med_hh_income                               0.80471
## total_pop                                   0.13103
```

```
## log(number_of_wells)                     0.19362
## is_Eastern_PA                            0.85089
## I(log(number_of_wells) * is_Eastern_PA)  0.42913
##
## Rho: 0.39732, LR test value: 5.3856, p-value: 0.020303
## Asymptotic standard error: 0.14846
##     z-value: 2.6763, p-value: 0.0074448
## Wald statistic: 7.1624, p-value: 0.0074448
##
## Log likelihood: -308.8523 for lag model
## ML residual variance (sigma squared): 570.33, (sigma: 23.882)
## Number of observations: 67
## Number of parameters estimated: 10
## AIC: 637.7, (AIC for lm: 641.09)
## LM test for residual autocorrelation
## test value: 0.029533, p-value: 0.86355
```

```
pennsylvania_lag_W_3 <- lagsarlm(cancer_rate ~ perc_white + perc_unemployed +
                                   med_hh_income + total_pop +
                                   log(number_of_wells) + is_Eastern_PA +
                                   I(log(number_of_wells) * is_Eastern_PA),
                                 data = PA_county_tracts, listw = W_list_knn3)
summary(pennsylvania_lag_W_3)
```

```
##
## Call:lagsarlm(formula = cancer_rate ~ perc_white + perc_unemployed +
##     med_hh_income + total_pop + log(number_of_wells) + is_Eastern_PA +
##     I(log(number_of_wells) * is_Eastern_PA), data = PA_county_tracts,
##     listw = W_list_knn3)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -55.679398 -17.585746   0.012419  17.494361  84.595107
##
## Type: lag
## Coefficients: (asymptotic standard errors)
##                                          Estimate  Std. Error z value
## (Intercept)                             2.5895e+02  9.8873e+01  2.6190
## perc_white                              1.1235e+02  6.0416e+01  1.8596
## perc_unemployed                         4.1195e+02  3.2076e+02  1.2843
## med_hh_income                          -9.9555e-05  3.3652e-04 -0.2958
## total_pop                               2.8122e-05  1.9027e-05  1.4780
## log(number_of_wells)                   -1.9902e+00  1.6012e+00 -1.2430
## is_Eastern_PA                           3.5995e+00  1.4008e+01  0.2570
## I(log(number_of_wells) * is_Eastern_PA) -1.0151e+01  1.2213e+01 -0.8312
##                                        Pr(>|z|)
## (Intercept)                            0.008818
## perc_white                             0.062941
## perc_unemployed                        0.199043
## med_hh_income                          0.767358
## total_pop                              0.139421
## log(number_of_wells)                   0.213883
## is_Eastern_PA                          0.797207
## I(log(number_of_wells) * is_Eastern_PA) 0.405850
```

```
##
## Rho: 0.21076, LR test value: 1.8967, p-value: 0.16845
## Asymptotic standard error: 0.14361
##     z-value: 1.4676, p-value: 0.1422
## Wald statistic: 2.1539, p-value: 0.1422
##
## Log likelihood: -310.5968 for lag model
## ML residual variance (sigma squared): 614.81, (sigma: 24.795)
## Number of observations: 67
## Number of parameters estimated: 10
## AIC: 641.19, (AIC for lm: 641.09)
## LM test for residual autocorrelation
## test value: 0.54752, p-value: 0.45933
```

```
pennsylvania_lag_W_10 <- lagsarlm(cancer_rate ~ perc_white + perc_unemployed +
                                    med_hh_income + total_pop +
                                    log(number_of_wells) + is_Eastern_PA +
                                    I(log(number_of_wells) * is_Eastern_PA),
                                  data = PA_county_tracts, listw = W_list_knn10)
summary(pennsylvania_lag_W_10)
```

```
##
## Call:lagsarlm(formula = cancer_rate ~ perc_white + perc_unemployed +
##     med_hh_income + total_pop + log(number_of_wells) + is_Eastern_PA +
##     I(log(number_of_wells) * is_Eastern_PA), data = PA_county_tracts,
##     listw = W_list_knn10)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -57.40701 -18.14611   0.48947  16.92293  83.80198
##
## Type: lag
## Coefficients: (asymptotic standard errors)
##                                         Estimate  Std. Error z value
## (Intercept)                            1.9264e+02  1.2645e+02  1.5234
## perc_white                             1.2087e+02  6.0400e+01  2.0011
## perc_unemployed                        4.2335e+02  3.2068e+02  1.3202
## med_hh_income                         -1.4309e-04  3.3656e-04 -0.4252
## total_pop                              2.9423e-05  1.9023e-05  1.5467
## log(number_of_wells)                  -1.8933e+00  1.6114e+00 -1.1750
## is_Eastern_PA                          3.7917e+00  1.4023e+01  0.2704
## I(log(number_of_wells) * is_Eastern_PA) -1.3668e+01  1.2217e+01 -1.1187
##                                        Pr(>|z|)
## (Intercept)                             0.12766
## perc_white                              0.04538
## perc_unemployed                         0.18678
## med_hh_income                           0.67072
## total_pop                               0.12194
## log(number_of_wells)                    0.24000
## is_Eastern_PA                           0.78686
## I(log(number_of_wells) * is_Eastern_PA)  0.26326
##
## Rho: 0.33795, LR test value: 2.0269, p-value: 0.15454
## Asymptotic standard error: 0.21913
```

```
##     z-value: 1.5422, p-value: 0.12302
## Wald statistic: 2.3785, p-value: 0.12302
##
## Log likelihood: -310.5317 for lag model
## ML residual variance (sigma squared): 614.44, (sigma: 24.788)
## Number of observations: 67
## Number of parameters estimated: 10
## AIC: 641.06, (AIC for lm: 641.09)
## LM test for residual autocorrelation
## test value: 0.57189, p-value: 0.44951
```

Creating the spatial Durbin models.

```
pennsylvania_Durbin_B <- lagsarlm(cancer_rate ~ perc_white + perc_unemployed +
                                    med_hh_income + total_pop_standardized +
                                    log(number_of_wells) + is_Eastern_PA +
                                    I(log(number_of_wells) * is_Eastern_PA),
                                data = PA_county_tracts, Durbin = T, listw = B_list)
summary(pennsylvania_Durbin_B)
```

```
##
## Call:lagsarlm(formula = cancer_rate ~ perc_white + perc_unemployed +
##     med_hh_income + total_pop_standardized + log(number_of_wells) +
##     is_Eastern_PA + I(log(number_of_wells) * is_Eastern_PA),
##     data = PA_county_tracts, listw = B_list, Durbin = T)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -54.6713 -15.9503   5.1352  15.6378  66.5044
##
## Type: mixed
## Coefficients: (asymptotic standard errors)
##                                              Estimate  Std. Error z value
## (Intercept)                                 3.4101e+02  8.0856e+01  4.2175
## perc_white                                  1.6116e+02  6.5236e+01  2.4705
## perc_unemployed                             4.1127e+02  3.6047e+02  1.1409
## med_hh_income                              -4.1508e-04  4.4094e-04 -0.9414
## total_pop_standardized                      5.9134e+01  3.3147e+01  1.7840
## log(number_of_wells)                       -3.2209e+00  2.2334e+00 -1.4421
## is_Eastern_PA                               2.8998e+00  1.5355e+01  0.1888
## I(log(number_of_wells) * is_Eastern_PA)    -1.3494e+01  1.2519e+01 -1.0779
## lag.(Intercept)                            -9.1421e+01  4.6503e+01 -1.9659
## lag.perc_white                              5.9734e+01  3.8857e+01  1.5373
## lag.perc_unemployed                         3.3565e+02  1.5450e+02  2.1725
## lag.med_hh_income                          -1.8062e-04  2.1832e-04 -0.8273
## lag.total_pop_standardized                  2.4633e+01  1.8607e+01  1.3239
## lag.log(number_of_wells)                    2.9083e-01  6.7220e-01  0.4327
## lag.is_Eastern_PA                           1.3268e+01  7.0276e+00  1.8880
## lag.I(log(number_of_wells) * is_Eastern_PA) -1.4747e+01  5.9700e+00 -2.4703
##                                              Pr(>|z|)
## (Intercept)                                 2.471e-05
## perc_white                                    0.01349
## perc_unemployed                               0.25390
```

```
## med_hh_income                                  0.34652
## total_pop_standardized                         0.07443
## log(number_of_wells)                           0.14926
## is_Eastern_PA                                  0.85021
## I(log(number_of_wells) * is_Eastern_PA)        0.28108
## lag.(Intercept)                                0.04931
## lag.perc_white                                 0.12422
## lag.perc_unemployed                            0.02982
## lag.med_hh_income                              0.40805
## lag.total_pop_standardized                     0.18555
## lag.log(number_of_wells)                       0.66526
## lag.is_Eastern_PA                              0.05902
## lag.I(log(number_of_wells) * is_Eastern_PA)    0.01350
##
## Rho: 0.047415, LR test value: 1.5992, p-value: 0.20601
## Asymptotic standard error: 0.033088
##     z-value: 1.433, p-value: 0.15186
## Wald statistic: 2.0535, p-value: 0.15186
##
## Log likelihood: -305.0128 for mixed model
## ML residual variance (sigma squared): 520.4, (sigma: 22.812)
## Number of observations: 67
## Number of parameters estimated: 18
## AIC: 646.03, (AIC for lm: 645.62)
## LM test for residual autocorrelation
## test value: 0.0056144, p-value: 0.94027
```

```
pennsylvania_Durbin_W <- lagsarlm(cancer_rate ~ perc_white + perc_unemployed +
                                    med_hh_income + total_pop_standardized +
                                    log(number_of_wells) + is_Eastern_PA +
                                    I(log(number_of_wells) * is_Eastern_PA),
                                  data = PA_county_tracts, Durbin = T, listw = W_list)
summary(pennsylvania_Durbin_W)
```

```
##
## Call:lagsarlm(formula = cancer_rate ~ perc_white + perc_unemployed +
##     med_hh_income + total_pop_standardized + log(number_of_wells) +
##     is_Eastern_PA + I(log(number_of_wells) * is_Eastern_PA),
##     data = PA_county_tracts, listw = W_list, Durbin = T)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -52.9601 -15.4705   3.2853  15.3087  55.6120
##
## Type: mixed
## Coefficients: (asymptotic standard errors)
##                                              Estimate  Std. Error z value
## (Intercept)                                -2.2863e+02  2.5590e+02 -0.8934
## perc_white                                  1.7136e+02  6.1580e+01  2.7828
## perc_unemployed                             4.9919e+02  3.6638e+02  1.3625
## med_hh_income                              -2.0674e-04  4.2165e-04 -0.4903
## total_pop_standardized                      5.5000e+01  3.1369e+01  1.7533
## log(number_of_wells)                       -6.0291e+00  2.2585e+00 -2.6695
## is_Eastern_PA                               9.6630e+00  1.5178e+01  0.6367
```

```
## I(log(number_of_wells) * is_Eastern_PA)     -7.7302e+00  1.1851e+01 -0.6523
## lag.perc_white                               3.4937e+02  1.8086e+02  1.9317
## lag.perc_unemployed                          1.7964e+03  7.3885e+02  2.4313
## lag.med_hh_income                           -5.0486e-04  9.7146e-04 -0.5197
## lag.total_pop_standardized                   1.1856e+02  8.3730e+01  1.4160
## lag.log(number_of_wells)                     4.4184e+00  3.2942e+00  1.3413
## lag.is_Eastern_PA                            5.7613e+01  3.3363e+01  1.7268
## lag.I(log(number_of_wells) * is_Eastern_PA) -7.7537e+01  2.4547e+01 -3.1588
##                                             Pr(>|z|)
## (Intercept)                                 0.371618
## perc_white                                  0.005390
## perc_unemployed                             0.173045
## med_hh_income                               0.623903
## total_pop_standardized                      0.079550
## log(number_of_wells)                        0.007597
## is_Eastern_PA                               0.524349
## I(log(number_of_wells) * is_Eastern_PA)     0.514204
## lag.perc_white                              0.053397
## lag.perc_unemployed                         0.015046
## lag.med_hh_income                           0.603278
## lag.total_pop_standardized                  0.156785
## lag.log(number_of_wells)                    0.179838
## lag.is_Eastern_PA                           0.084199
## lag.I(log(number_of_wells) * is_Eastern_PA) 0.001584
##
## Rho: 0.26101, LR test value: 2.076, p-value: 0.14964
## Asymptotic standard error: 0.16418
##     z-value: 1.5898, p-value: 0.11187
## Wald statistic: 2.5275, p-value: 0.11187
##
## Log likelihood: -301.7553 for mixed model
## ML residual variance (sigma squared): 471.19, (sigma: 21.707)
## Number of observations: 67
## Number of parameters estimated: 17
## AIC: 637.51, (AIC for lm: 637.59)
## LM test for residual autocorrelation
## test value: 0.019107, p-value: 0.89006
```

```
pennsylvania_Durbin_W_3 <- lagsarlm(cancer_rate ~ perc_white + perc_unemployed +
                                      med_hh_income + total_pop_standardized +
                                      log(number_of_wells) + is_Eastern_PA +
                                      I(log(number_of_wells) * is_Eastern_PA),
                      data = PA_county_tracts, Durbin = T, listw = W_list_knn3)
summary(pennsylvania_Durbin_W_3)
```

```
##
## Call:lagsarlm(formula = cancer_rate ~ perc_white + perc_unemployed +
##     med_hh_income + total_pop_standardized + log(number_of_wells) +
##     is_Eastern_PA + I(log(number_of_wells) * is_Eastern_PA),
##     data = PA_county_tracts, listw = W_list_knn3, Durbin = T)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -53.1166 -18.9635   2.4641  14.3378  72.4744
```

```
##
## Type: mixed
## Coefficients: (asymptotic standard errors)
##                                            Estimate  Std. Error z value
## (Intercept)                                1.6474e+02  1.9222e+02  0.8571
## perc_white                                 1.3536e+02  5.9166e+01  2.2878
## perc_unemployed                            1.8709e+02  3.8568e+02  0.4851
## med_hh_income                             -5.5702e-04  4.4970e-04 -1.2387
## total_pop_standardized                     4.6907e+01  3.1163e+01  1.5052
## log(number_of_wells)                      -6.3207e+00  2.4473e+00 -2.5827
## is_Eastern_PA                             -9.2511e+00  1.5450e+01 -0.5988
## I(log(number_of_wells) * is_Eastern_PA)   -1.2418e+00  1.2361e+01 -0.1005
## lag.perc_white                             9.2193e+01  1.2721e+02  0.7247
## lag.perc_unemployed                        1.0379e+03  5.5467e+02  1.8712
## lag.med_hh_income                         -4.1999e-04  8.8702e-04 -0.4735
## lag.total_pop_standardized                 6.3672e+01  5.9985e+01  1.0615
## lag.log(number_of_wells)                   5.6896e+00  3.0778e+00  1.8486
## lag.is_Eastern_PA                          5.0734e+01  2.7466e+01  1.8472
## lag.I(log(number_of_wells) * is_Eastern_PA) -5.2260e+01  1.7985e+01 -2.9057
##                                           Pr(>|z|)
## (Intercept)                               0.391404
## perc_white                                0.022147
## perc_unemployed                           0.627622
## med_hh_income                             0.215471
## total_pop_standardized                    0.132267
## log(number_of_wells)                      0.009803
## is_Eastern_PA                             0.549323
## I(log(number_of_wells) * is_Eastern_PA)   0.919981
## lag.perc_white                            0.468624
## lag.perc_unemployed                       0.061322
## lag.med_hh_income                         0.635866
## lag.total_pop_standardized                0.288477
## lag.log(number_of_wells)                  0.064519
## lag.is_Eastern_PA                         0.064719
## lag.I(log(number_of_wells) * is_Eastern_PA) 0.003664
##
## Rho: 0.15995, LR test value: 1.0397, p-value: 0.3079
## Asymptotic standard error: 0.14609
##     z-value: 1.0949, p-value: 0.27357
## Wald statistic: 1.1987, p-value: 0.27357
##
## Log likelihood: -304.5469 for mixed model
## ML residual variance (sigma squared): 515.98, (sigma: 22.715)
## Number of observations: 67
## Number of parameters estimated: 17
## AIC: 643.09, (AIC for lm: 642.13)
## LM test for residual autocorrelation
## test value: 0.4542, p-value: 0.50035
```

```
pennsylvania_Durbin_W_10 <- lagsarlm(cancer_rate ~ perc_white + perc_unemployed +
                                       med_hh_income + total_pop_standardized +
                                       log(number_of_wells) + is_Eastern_PA +
                                       I(log(number_of_wells) * is_Eastern_PA),
                     data = PA_county_tracts, Durbin = T, listw = W_list_knn10)
```

```
summary(pennsylvania_Durbin_W_10)
```

```
##
## Call:lagsarlm(formula = cancer_rate ~ perc_white + perc_unemployed +
##     med_hh_income + total_pop_standardized + log(number_of_wells) +
##     is_Eastern_PA + I(log(number_of_wells) * is_Eastern_PA),
##     data = PA_county_tracts, listw = W_list_knn10, Durbin = T)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -57.82868 -15.05188   0.24235  14.73290  60.74380
##
## Type: mixed
## Coefficients: (asymptotic standard errors)
##                                            Estimate  Std. Error z value
## (Intercept)                                1.1162e+03  4.9201e+02  2.2687
## perc_white                                 3.7490e+01  6.7839e+01  0.5526
## perc_unemployed                            1.0254e+02  3.4755e+02  0.2951
## med_hh_income                             -7.8195e-04  4.3409e-04 -1.8014
## total_pop_standardized                     3.9198e+01  3.1062e+01  1.2619
## log(number_of_wells)                      -5.4648e+00  2.4685e+00 -2.2138
## is_Eastern_PA                              8.1692e+00  1.6621e+01  0.4915
## I(log(number_of_wells) * is_Eastern_PA)   -5.3192e+00  1.2942e+01 -0.4110
## lag.perc_white                            -4.4541e+02  3.6104e+02 -1.2337
## lag.perc_unemployed                       -7.9060e+02  1.3962e+03 -0.5663
## lag.med_hh_income                         -5.1551e-03  1.9055e-03 -2.7054
## lag.total_pop_standardized                 2.1100e+02  1.3279e+02  1.5890
## lag.log(number_of_wells)                   2.6123e+00  5.0445e+00  0.5178
## lag.is_Eastern_PA                         -2.0772e+01  5.9975e+01 -0.3464
## lag.I(log(number_of_wells) * is_Eastern_PA)  2.7944e+01  6.0179e+01  0.4643
##                                           Pr(>|z|)
## (Intercept)                               0.023288
## perc_white                                0.580513
## perc_unemployed                           0.767955
## med_hh_income                             0.071644
## total_pop_standardized                    0.206970
## log(number_of_wells)                      0.026841
## is_Eastern_PA                             0.623082
## I(log(number_of_wells) * is_Eastern_PA)   0.681064
## lag.perc_white                            0.217323
## lag.perc_unemployed                       0.571223
## lag.med_hh_income                         0.006823
## lag.total_pop_standardized                0.112070
## lag.log(number_of_wells)                  0.604564
## lag.is_Eastern_PA                         0.729077
## lag.I(log(number_of_wells) * is_Eastern_PA) 0.642399
##
## Rho: 0.23469, LR test value: 0.90291, p-value: 0.342
## Asymptotic standard error: 0.23368
##     z-value: 1.0043, p-value: 0.31523
## Wald statistic: 1.0086, p-value: 0.31523
##
## Log likelihood: -303.6124 for mixed model
```

```
## ML residual variance (sigma squared): 502.78, (sigma: 22.423)
## Number of observations: 67
## Number of parameters estimated: 17
## AIC: 641.22, (AIC for lm: 640.13)
## LM test for residual autocorrelation
## test value: 2.668, p-value: 0.10238
```

Creating the spatial error models.

```
pennsylvania_sar_B <- errorsarlm(cancer_rate ~ perc_white + perc_unemployed +
                                   med_hh_income + total_pop +
                                   log(number_of_wells) + is_Eastern_PA +
                                   I(log(number_of_wells) * is_Eastern_PA),
                                 data = PA_county_tracts, listw = B_list)
summary(pennsylvania_sar_B)
```

```
##
## Call:errorsarlm(formula = cancer_rate ~ perc_white + perc_unemployed +
##     med_hh_income + total_pop + log(number_of_wells) + is_Eastern_PA +
##     I(log(number_of_wells) * is_Eastern_PA), data = PA_county_tracts,
##     listw = B_list)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -53.3677 -17.1711   1.4563  16.2915  80.9895
##
## Type: error
## Coefficients: (asymptotic standard errors)
##                                       Estimate  Std. Error z value
## (Intercept)                         3.8118e+02  6.7025e+01  5.6872
## perc_white                          9.7297e+01  5.5714e+01  1.7464
## perc_unemployed                     2.8054e+02  3.3310e+02  0.8422
## med_hh_income                      -6.2038e-05  3.5050e-04 -0.1770
## total_pop                           2.5333e-05  1.8492e-05  1.3699
## log(number_of_wells)               -2.4444e+00  1.8247e+00 -1.3397
## is_Eastern_PA                      -2.8296e+00  1.4620e+01 -0.1935
## I(log(number_of_wells) * is_Eastern_PA) -4.1695e+00  1.1837e+01 -0.3522
##                                       Pr(>|z|)
## (Intercept)                         1.292e-08
## perc_white                            0.08075
## perc_unemployed                       0.39967
## med_hh_income                         0.85951
## total_pop                             0.17071
## log(number_of_wells)                  0.18035
## is_Eastern_PA                         0.84653
## I(log(number_of_wells) * is_Eastern_PA)   0.72466
##
## Lambda: 0.072581, LR test value: 3.7268, p-value: 0.053545
## Asymptotic standard error: 0.030111
##     z-value: 2.4105, p-value: 0.015932
## Wald statistic: 5.8104, p-value: 0.015932
##
## Log likelihood: -309.6818 for error model
```

```
## ML residual variance (sigma squared): 587.38, (sigma: 24.236)
## Number of observations: 67
## Number of parameters estimated: 10
## AIC: 639.36, (AIC for lm: 641.09)
```

```
pennsylvania_sar_W <- errorsarlm(cancer_rate ~ perc_white + perc_unemployed +
                                   med_hh_income + total_pop +
                                   log(number_of_wells) + is_Eastern_PA +
                                   I(log(number_of_wells) * is_Eastern_PA),
                                 data = PA_county_tracts, listw = W_list)
summary(pennsylvania_sar_W)
```

```
##
## Call:errorsarlm(formula = cancer_rate ~ perc_white + perc_unemployed +
##     med_hh_income + total_pop + log(number_of_wells) + is_Eastern_PA +
##     I(log(number_of_wells) * is_Eastern_PA), data = PA_county_tracts,
##     listw = W_list)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -52.7792 -17.9384   3.2981  17.0488  74.8116
##
## Type: error
## Coefficients: (asymptotic standard errors)
##                                        Estimate  Std. Error z value
## (Intercept)                           3.8686e+02  6.4832e+01  5.9672
## perc_white                            9.4476e+01  5.4506e+01  1.7333
## perc_unemployed                       2.7375e+02  3.3207e+02  0.8244
## med_hh_income                        -7.3103e-05  3.4924e-04 -0.2093
## total_pop                             2.2520e-05  1.8361e-05  1.2265
## log(number_of_wells)                 -2.9638e+00  1.8447e+00 -1.6066
## is_Eastern_PA                        -2.8361e+00  1.4598e+01 -0.1943
## I(log(number_of_wells) * is_Eastern_PA) -6.9428e-01  1.1717e+01 -0.0593
##                                        Pr(>|z|)
## (Intercept)                           2.414e-09
## perc_white                              0.08304
## perc_unemployed                         0.40972
## med_hh_income                           0.83420
## total_pop                               0.22000
## log(number_of_wells)                    0.10813
## is_Eastern_PA                           0.84596
## I(log(number_of_wells) * is_Eastern_PA)   0.95275
##
## Lambda: 0.44111, LR test value: 5.0338, p-value: 0.024857
## Asymptotic standard error: 0.14489
##     z-value: 3.0444, p-value: 0.0023314
## Wald statistic: 9.2684, p-value: 0.0023314
##
## Log likelihood: -309.0283 for error model
## ML residual variance (sigma squared): 568.13, (sigma: 23.836)
## Number of observations: 67
## Number of parameters estimated: 10
## AIC: 638.06, (AIC for lm: 641.09)
```

```
pennsylvania_sar_W_3 <- errorsarlm(cancer_rate ~ perc_white + perc_unemployed +
                                     med_hh_income + total_pop +
                                     log(number_of_wells) + is_Eastern_PA +
                                     I(log(number_of_wells) * is_Eastern_PA),
                                   data = PA_county_tracts, listw = W_list_knn3)
summary(pennsylvania_sar_W_3)
```

```
##
## Call:errorsarlm(formula = cancer_rate ~ perc_white + perc_unemployed +
##     med_hh_income + total_pop + log(number_of_wells) + is_Eastern_PA +
##     I(log(number_of_wells) * is_Eastern_PA), data = PA_county_tracts,
##     listw = W_list_knn3)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -53.876722 -18.463500  -0.010814  17.994732  83.899825
##
## Type: error
## Coefficients: (asymptotic standard errors)
##                                       Estimate  Std. Error z value
## (Intercept)                          3.7517e+02  6.9099e+01  5.4295
## perc_white                           1.0274e+02  5.8610e+01  1.7530
## perc_unemployed                      3.1370e+02  3.3846e+02  0.9268
## med_hh_income                       -8.3977e-05  3.4764e-04 -0.2416
## total_pop                            2.5975e-05  1.9013e-05  1.3662
## log(number_of_wells)                -2.4645e+00  1.7255e+00 -1.4283
## is_Eastern_PA                       -7.0518e-01  1.4738e+01 -0.0478
## I(log(number_of_wells) * is_Eastern_PA) -4.0878e+00  1.2256e+01 -0.3335
##                                      Pr(>|z|)
## (Intercept)                          5.65e-08
## perc_white                            0.07961
## perc_unemployed                       0.35401
## med_hh_income                         0.80912
## total_pop                             0.17189
## log(number_of_wells)                  0.15321
## is_Eastern_PA                         0.96184
## I(log(number_of_wells) * is_Eastern_PA)  0.73873
##
## Lambda: 0.23724, LR test value: 1.7249, p-value: 0.18907
## Asymptotic standard error: 0.14405
##     z-value: 1.6469, p-value: 0.099577
## Wald statistic: 2.7123, p-value: 0.099577
##
## Log likelihood: -310.6827 for error model
## ML residual variance (sigma squared): 614.29, (sigma: 24.785)
## Number of observations: 67
## Number of parameters estimated: 10
## AIC: 641.37, (AIC for lm: 641.09)
```

```
pennsylvania_sar_W_10 <- errorsarlm(cancer_rate ~ perc_white + perc_unemployed +
                                      med_hh_income + total_pop +
                                      log(number_of_wells) + is_Eastern_PA +
                                      I(log(number_of_wells) * is_Eastern_PA),
```

```
                                  data = PA_county_tracts, listw = W_list_knn10)
summary(pennsylvania_sar_W_10)
```

```
##
## Call:errorsarlm(formula = cancer_rate ~ perc_white + perc_unemployed +
##     med_hh_income + total_pop + log(number_of_wells) + is_Eastern_PA +
##     I(log(number_of_wells) * is_Eastern_PA), data = PA_county_tracts,
##     listw = W_list_knn10)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -57.53586 -18.59095   0.70016  16.19593  81.13880
##
## Type: error
## Coefficients: (asymptotic standard errors)
##                                         Estimate  Std. Error z value
## (Intercept)                             3.6286e+02  6.9578e+01  5.2152
## perc_white                              1.1820e+02  5.8294e+01  2.0276
## perc_unemployed                         3.4981e+02  3.4413e+02  1.0165
## med_hh_income                          -1.6020e-04  3.5374e-04 -0.4529
## total_pop                               2.6649e-05  1.8619e-05  1.4312
## log(number_of_wells)                   -2.4715e+00  1.8173e+00 -1.3600
## is_Eastern_PA                           2.3332e+00  1.4430e+01  0.1617
## I(log(number_of_wells) * is_Eastern_PA) -1.2047e+01  1.2132e+01 -0.9930
##                                          Pr(>|z|)
## (Intercept)                             1.836e-07
## perc_white                                 0.0426
## perc_unemployed                            0.3094
## med_hh_income                              0.6506
## total_pop                                  0.1524
## log(number_of_wells)                       0.1738
## is_Eastern_PA                              0.8715
## I(log(number_of_wells) * is_Eastern_PA)    0.3207
##
## Lambda: 0.38601, LR test value: 2.2805, p-value: 0.13101
## Asymptotic standard error: 0.21059
##     z-value: 1.833, p-value: 0.066797
## Wald statistic: 3.36, p-value: 0.066797
##
## Log likelihood: -310.4049 for error model
## ML residual variance (sigma squared): 609.8, (sigma: 24.694)
## Number of observations: 67
## Number of parameters estimated: 10
## AIC: 640.81, (AIC for lm: 641.09)
```

Creating the CAR models.

```
pennsylvania_car_B <- spautolm(cancer_rate ~ perc_white + perc_unemployed +
                                 med_hh_income + total_pop +
                                 log(number_of_wells) + is_Eastern_PA +
                                 I(log(number_of_wells) * is_Eastern_PA),
                         data = PA_county_tracts, family = "CAR", listw = B_list)
summary(pennsylvania_car_B)
```

```
##
## Call: spautolm(formula = cancer_rate ~ perc_white + perc_unemployed +
##     med_hh_income + total_pop + log(number_of_wells) + is_Eastern_PA +
##     I(log(number_of_wells) * is_Eastern_PA), data = PA_county_tracts,
##     listw = B_list, family = "CAR")
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -54.0454 -16.5400   1.7241  14.7077  78.6157
##
## Coefficients:
##                                       Estimate  Std. Error z value
## (Intercept)                          3.7732e+02  6.7734e+01  5.5707
## perc_white                           1.0000e+02  5.6516e+01  1.7695
## perc_unemployed                      3.0165e+02  3.3360e+02  0.9042
## med_hh_income                       -7.0129e-05  3.5006e-04 -0.2003
## total_pop                            2.5797e-05  1.8585e-05  1.3880
## log(number_of_wells)                -2.3444e+00  1.8112e+00 -1.2944
## is_Eastern_PA                       -1.7844e+00  1.4581e+01 -0.1224
## I(log(number_of_wells) * is_Eastern_PA) -5.7913e+00  1.1939e+01 -0.4851
##                                       Pr(>|z|)
## (Intercept)                          2.537e-08
## perc_white                             0.07681
## perc_unemployed                        0.36588
## med_hh_income                          0.84122
## total_pop                              0.16514
## log(number_of_wells)                   0.19554
## is_Eastern_PA                          0.90260
## I(log(number_of_wells) * is_Eastern_PA)   0.62762
##
## Lambda: 0.11447 LR test value: 3.3226 p-value: 0.068335
## Numerical Hessian standard error of lambda: 0.046971
##
## Log likelihood: -309.8839
## ML residual variance (sigma squared): 583.33, (sigma: 24.152)
## Number of observations: 67
## Number of parameters estimated: 10
## AIC: 639.77
```

```
pennsylvania_car_W <- spautolm(cancer_rate ~ perc_white + perc_unemployed +
                                 med_hh_income + total_pop +
                                 log(number_of_wells) + is_Eastern_PA +
                                 I(log(number_of_wells) * is_Eastern_PA),
                       data = PA_county_tracts, family = "CAR", listw = W_list)
```

```
## Warning in spautolm(cancer_rate ~ perc_white + perc_unemployed + med_hh_income
## + : Non-symmetric spatial weights in CAR model
```

```
summary(pennsylvania_car_W)
```

```
##
## Call: spautolm(formula = cancer_rate ~ perc_white + perc_unemployed +
##     med_hh_income + total_pop + log(number_of_wells) + is_Eastern_PA +
```

```
##     I(log(number_of_wells) * is_Eastern_PA), data = PA_county_tracts,
##     listw = W_list, family = "CAR")
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -56.77789 -18.24277   0.60317  16.26521  86.06579
##
## Coefficients:
##                                         Estimate  Std. Error z value
## (Intercept)                            3.6736e+02  7.1099e+01  5.1669
## perc_white                             9.9212e+01  6.1280e+01  1.6190
## perc_unemployed                        3.9985e+02  3.2795e+02  1.2193
## med_hh_income                         -4.4376e-06  3.4356e-04 -0.0129
## total_pop                              2.6100e-05  1.9376e-05  1.3470
## log(number_of_wells)                  -2.0698e+00  1.6422e+00 -1.2604
## is_Eastern_PA                          7.4927e-01  1.4335e+01  0.0523
## I(log(number_of_wells) * is_Eastern_PA) -7.0412e+00  1.2435e+01 -0.5663
##                                        Pr(>|z|)
## (Intercept)                            2.38e-07
## perc_white                               0.1054
## perc_unemployed                          0.2227
## med_hh_income                            0.9897
## total_pop                                0.1780
## log(number_of_wells)                     0.2075
## is_Eastern_PA                            0.9583
## I(log(number_of_wells) * is_Eastern_PA)  0.5712
##
## Lambda: 0.053583 LR test value: 0.27982 p-value: 0.59682
## Numerical Hessian standard error of lambda: 0.59093
##
## Log likelihood: -311.4053
## ML residual variance (sigma squared): 637.47, (sigma: 25.248)
## Number of observations: 67
## Number of parameters estimated: 10
## AIC: 642.81
```

```
pennsylvania_car_W_3 <- spautolm(cancer_rate ~ perc_white + perc_unemployed +
                                   med_hh_income + total_pop +
                                   log(number_of_wells) + is_Eastern_PA +
                                   I(log(number_of_wells) * is_Eastern_PA),
                 data = PA_county_tracts, family = "CAR", listw = W_list_knn3)
```

```
## Warning in spautolm(cancer_rate ~ perc_white + perc_unemployed + med_hh_income
## + : Non-symmetric spatial weights in CAR model
```

```
summary(pennsylvania_car_W_3)
```

```
##
## Call: spautolm(formula = cancer_rate ~ perc_white + perc_unemployed +
##     med_hh_income + total_pop + log(number_of_wells) + is_Eastern_PA +
##     I(log(number_of_wells) * is_Eastern_PA), data = PA_county_tracts,
##     listw = W_list_knn3, family = "CAR")
##
```

```
## Residuals:
##      Min       1Q   Median       3Q      Max
## -61.8921 -17.7640   1.6453  16.6023  86.9649
##
## Coefficients:
##                                          Estimate  Std. Error z value
## (Intercept)                             3.4743e+02  7.1420e+01  4.8647
## perc_white                              1.2151e+02  6.1599e+01  1.9726
## perc_unemployed                         5.1205e+02  3.2755e+02  1.5633
## med_hh_income                          -1.1763e-04  3.4342e-04 -0.3425
## total_pop                               3.0458e-05  1.9415e-05  1.5688
## log(number_of_wells)                   -2.3310e+00  1.6307e+00 -1.4295
## is_Eastern_PA                           4.7150e+00  1.4302e+01  0.3297
## I(log(number_of_wells) * is_Eastern_PA) -1.3740e+01  1.2456e+01 -1.1031
##                                          Pr(>|z|)
## (Intercept)                             1.146e-06
## perc_white                                0.04854
## perc_unemployed                           0.11799
## med_hh_income                             0.73195
## total_pop                                 0.11670
## log(number_of_wells)                      0.15287
## is_Eastern_PA                             0.74165
## I(log(number_of_wells) * is_Eastern_PA)   0.26997
##
## Lambda: 0.0078511 LR test value: 0.02766 p-value: 0.86791
## Numerical Hessian standard error of lambda: 0.21453
##
## Log likelihood: -311.5313
## ML residual variance (sigma squared): 640.06, (sigma: 25.299)
## Number of observations: 67
## Number of parameters estimated: 10
## AIC: 643.06
```

```
pennsylvania_car_W_10 <- spautolm(cancer_rate ~ perc_white + perc_unemployed +
                                    med_hh_income + total_pop +
                                    log(number_of_wells) + is_Eastern_PA +
                                    I(log(number_of_wells) * is_Eastern_PA),
                  data = PA_county_tracts, family = "CAR", listw = W_list_knn10)
```

```
## Warning in spautolm(cancer_rate ~ perc_white + perc_unemployed + med_hh_income
## + : Non-symmetric spatial weights in CAR model
```

```
## Warning in sqrt(fdHess[1, 1]): NaNs produced
```

```
summary(pennsylvania_car_W_10)
```

```
##
## Call: spautolm(formula = cancer_rate ~ perc_white + perc_unemployed +
##     med_hh_income + total_pop + log(number_of_wells) + is_Eastern_PA +
##     I(log(number_of_wells) * is_Eastern_PA), data = PA_county_tracts,
##     listw = W_list_knn10, family = "CAR")
##
```

```
## Residuals:
##      Min       1Q   Median       3Q      Max
## -61.5155 -17.8402   1.9362  16.4124  86.4832
##
## Coefficients:
##                                          Estimate  Std. Error z value
## (Intercept)                             3.5235e+02  7.1439e+01  4.9322
## perc_white                              1.1653e+02  6.1619e+01  1.8911
## perc_unemployed                         4.8461e+02  3.2746e+02  1.4799
## med_hh_income                          -9.9257e-05  3.4343e-04 -0.2890
## total_pop                               2.8953e-05  1.9414e-05  1.4914
## log(number_of_wells)                   -2.2585e+00  1.6300e+00 -1.3856
## is_Eastern_PA                           4.3991e+00  1.4294e+01  0.3078
## I(log(number_of_wells) * is_Eastern_PA) -1.3396e+01  1.2454e+01 -1.0756
##                                          Pr(>|z|)
## (Intercept)                             8.131e-07
## perc_white                                 0.0586
## perc_unemployed                            0.1389
## med_hh_income                              0.7726
## total_pop                                  0.1359
## log(number_of_wells)                       0.1659
## is_Eastern_PA                              0.7583
## I(log(number_of_wells) * is_Eastern_PA)    0.2821
##
## Lambda: 0.0051889 LR test value: 0.014289 p-value: 0.90485
## Numerical Hessian standard error of lambda: NaN
##
## Log likelihood: -311.538
## ML residual variance (sigma squared): 640.19, (sigma: 25.302)
## Number of observations: 67
## Number of parameters estimated: 10
## AIC: 643.08
```

The best model for cancer rate.

```
pennsylvania_Durbin_W <- lagsarlm(cancer_rate ~ perc_white + perc_unemployed +
                                    med_hh_income + total_pop_standardized +
                                    log(number_of_wells) + is_Eastern_PA +
                                    I(log(number_of_wells) * is_Eastern_PA),
                            data = PA_county_tracts, Durbin = T, listw = W_list)
summary(pennsylvania_Durbin_W)
```

```
##
## Call:lagsarlm(formula = cancer_rate ~ perc_white + perc_unemployed +
##     med_hh_income + total_pop_standardized + log(number_of_wells) +
##     is_Eastern_PA + I(log(number_of_wells) * is_Eastern_PA),
##     data = PA_county_tracts, listw = W_list, Durbin = T)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -52.9601 -15.4705   3.2853  15.3087  55.6120
##
## Type: mixed
```

```
## Coefficients: (asymptotic standard errors)
##                                           Estimate  Std. Error z value
## (Intercept)                             -2.2863e+02  2.5590e+02 -0.8934
## perc_white                               1.7136e+02  6.1580e+01  2.7828
## perc_unemployed                          4.9919e+02  3.6638e+02  1.3625
## med_hh_income                           -2.0674e-04  4.2165e-04 -0.4903
## total_pop_standardized                   5.5000e+01  3.1369e+01  1.7533
## log(number_of_wells)                    -6.0291e+00  2.2585e+00 -2.6695
## is_Eastern_PA                            9.6630e+00  1.5178e+01  0.6367
## I(log(number_of_wells) * is_Eastern_PA) -7.7302e+00  1.1851e+01 -0.6523
## lag.perc_white                           3.4937e+02  1.8086e+02  1.9317
## lag.perc_unemployed                      1.7964e+03  7.3885e+02  2.4313
## lag.med_hh_income                       -5.0486e-04  9.7146e-04 -0.5197
## lag.total_pop_standardized               1.1856e+02  8.3730e+01  1.4160
## lag.log(number_of_wells)                 4.4184e+00  3.2942e+00  1.3413
## lag.is_Eastern_PA                        5.7613e+01  3.3363e+01  1.7268
## lag.I(log(number_of_wells) * is_Eastern_PA) -7.7537e+01  2.4547e+01 -3.1588
##                                         Pr(>|z|)
## (Intercept)                             0.371618
## perc_white                              0.005390
## perc_unemployed                         0.173045
## med_hh_income                           0.623903
## total_pop_standardized                  0.079550
## log(number_of_wells)                    0.007597
## is_Eastern_PA                           0.524349
## I(log(number_of_wells) * is_Eastern_PA) 0.514204
## lag.perc_white                          0.053397
## lag.perc_unemployed                     0.015046
## lag.med_hh_income                       0.603278
## lag.total_pop_standardized              0.156785
## lag.log(number_of_wells)                0.179838
## lag.is_Eastern_PA                       0.084199
## lag.I(log(number_of_wells) * is_Eastern_PA) 0.001584
##
## Rho: 0.26101, LR test value: 2.076, p-value: 0.14964
## Asymptotic standard error: 0.16418
##     z-value: 1.5898, p-value: 0.11187
## Wald statistic: 2.5275, p-value: 0.11187
##
## Log likelihood: -301.7553 for mixed model
## ML residual variance (sigma squared): 471.19, (sigma: 21.707)
## Number of observations: 67
## Number of parameters estimated: 17
## AIC: 637.51, (AIC for lm: 637.59)
## LM test for residual autocorrelation
## test value: 0.019107, p-value: 0.89006
```

Making the lm for asthma rate, and checking the VIF of the linear model.

```
asthma_lm <- lm(asthma_ED_visits ~ perc_white + perc_unemployed + med_hh_income +
                  total_pop + log(number_of_wells) + is_Eastern_PA +
                  I(log(number_of_wells) * is_Eastern_PA),
                data = PA_county_tracts@data)
summary(asthma_lm)
```

```
##
## Call:
## lm(formula = asthma_ED_visits ~ perc_white + perc_unemployed +
##     med_hh_income + total_pop + log(number_of_wells) + is_Eastern_PA +
##     I(log(number_of_wells) * is_Eastern_PA), data = PA_county_tracts@data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -37.342  -7.422  -1.890   6.095  74.103
##
## Coefficients:
##                                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)                          1.442e+02  4.701e+01   3.068  0.00325
## perc_white                          -1.760e+02  4.055e+01  -4.341 5.64e-05
## perc_unemployed                      4.501e+02  2.154e+02   2.090  0.04092
## med_hh_income                        3.005e-04  2.259e-04   1.330  0.18856
## total_pop                           -1.197e-05  1.278e-05  -0.937  0.35276
## log(number_of_wells)                -3.282e-01  1.072e+00  -0.306  0.76048
## is_Eastern_PA                        1.386e+01  9.403e+00   1.474  0.14569
## I(log(number_of_wells) * is_Eastern_PA) -5.894e+00  8.195e+00  -0.719  0.47484
##
## (Intercept)                         **
## perc_white                          ***
## perc_unemployed                     *
## med_hh_income
## total_pop
## log(number_of_wells)
## is_Eastern_PA
## I(log(number_of_wells) * is_Eastern_PA)
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 16.65 on 59 degrees of freedom
## Multiple R-squared:  0.6521, Adjusted R-squared:  0.6109
## F-statistic:  15.8 on 7 and 59 DF,  p-value: 1.745e-11
```

```
library(car)
vif(asthma_lm)
```

```
##                              perc_white                          perc_unemployed
##                                3.687980                                1.860221
##                           med_hh_income                               total_pop
##                                1.802330                                3.048814
##                     log(number_of_wells)                           is_Eastern_PA
##                                3.650169                                5.342124
## I(log(number_of_wells) * is_Eastern_PA)
##                                1.528325
```

```
#store the residuals to see if there is any spatial structure to the residuals
resid <- asthma_lm$residuals

#can join this to the SpatialPolygons
PA_county_tracts@data$resid2 <- resid
```

Moran's I tests for each neighborhood matrix on the linear model.

```
#set.seed(597)
#binary, spatial adjacency
mc <- moran(PA_county_tracts$resid2, B_list, n = length(pennsylvania_nb),
            S0 = Szero(B_list))
mc$I
```

```
## [1] 0.04404035
```

```
moran.mc(x=PA_county_tracts$resid2, listw=B_list, nsim=5000)
```

```
##
##  Monte-Carlo simulation of Moran I
##
## data:  PA_county_tracts$resid2
## weights: B_list
## number of simulations + 1: 5001
##
## statistic = 0.04404, observed rank = 4077, p-value = 0.1848
## alternative hypothesis: greater
```

```
#row-standardized, spatial adjacency
mc <- moran(PA_county_tracts$resid2, W_list, n = length(pennsylvania_nb),
            S0 = Szero(W_list))
mc$I
```

```
## [1] 0.09274497
```

```
moran.mc(x=PA_county_tracts$resid2, listw=W_list, nsim=5000)
```

```
##
##  Monte-Carlo simulation of Moran I
##
## data:  PA_county_tracts$resid2
## weights: W_list
## number of simulations + 1: 5001
##
## statistic = 0.092745, observed rank = 4663, p-value = 0.06759
## alternative hypothesis: greater
```

```
#k-nn 3, row-standardized
mc <- moran(PA_county_tracts$resid2, W_list_knn3, n = length(pennsylvania_nb),
            S0 = Szero(W_list_knn3))
mc$I
```

```
## [1] 0.05982705
```

```
moran.mc(x=PA_county_tracts$resid2, listw=W_list_knn3, nsim=5000)
```

```
##
##  Monte-Carlo simulation of Moran I
##
## data:  PA_county_tracts$resid2
## weights: W_list_knn3
## number of simulations + 1: 5001
##
## statistic = 0.059827, observed rank = 4027, p-value = 0.1948
## alternative hypothesis: greater
```

```
#k-nn 5, row-standardized
mc <- moran(PA_county_tracts$resid2, W_list_knn5, n = length(pennsylvania_nb),
            S0 = Szero(W_list_knn5))
mc$I
```

```
## [1] 0.08734474
```

```
moran.mc(x=PA_county_tracts$resid2, listw=W_list_knn5, nsim=5000)
```

```
##
##  Monte-Carlo simulation of Moran I
##
## data:  PA_county_tracts$resid2
## weights: W_list_knn5
## number of simulations + 1: 5001
##
## statistic = 0.087345, observed rank = 4658, p-value = 0.06859
## alternative hypothesis: greater
```

```
#k-nn 10, row-standardized
mc <- moran(PA_county_tracts$resid2, W_list_knn10, n = length(pennsylvania_nb),
            S0 = Szero(W_list_knn10))
mc$I
```

```
## [1] -0.04438311
```

```
moran.mc(x=PA_county_tracts$resid2, listw=W_list_knn10, nsim=5000)
```

```
##
##  Monte-Carlo simulation of Moran I
##
## data:  PA_county_tracts$resid2
## weights: W_list_knn10
## number of simulations + 1: 5001
##
## statistic = -0.044383, observed rank = 1365, p-value = 0.7271
## alternative hypothesis: greater
```

Creating the spatial lag models.

```r
asthma_lag_W <- lagsarlm(asthma_ED_visits ~ perc_white + perc_unemployed +
                           med_hh_income + total_pop + log(number_of_wells) +
                           is_Eastern_PA + I(log(number_of_wells) * is_Eastern_PA),
                         data = PA_county_tracts, listw = W_list)
summary(asthma_lag_W)
```

```
##
## Call:lagsarlm(formula = asthma_ED_visits ~ perc_white + perc_unemployed +
##     med_hh_income + total_pop + log(number_of_wells) + is_Eastern_PA +
##     I(log(number_of_wells) * is_Eastern_PA), data = PA_county_tracts,
##     listw = W_list)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -38.3886  -7.8410  -2.1800   5.4135  73.8442
##
## Type: lag
## Coefficients: (asymptotic standard errors)
##                                       Estimate  Std. Error z value
## (Intercept)                          1.5141e+02  4.3974e+01  3.4430
## perc_white                          -1.7551e+02  3.7684e+01 -4.6573
## perc_unemployed                      3.7656e+02  2.1137e+02  1.7815
## med_hh_income                        1.6716e-04  2.3304e-04  0.7173
## total_pop                           -1.2264e-05  1.1921e-05 -1.0287
## log(number_of_wells)                -2.0972e-01  9.9974e-01 -0.2098
## is_Eastern_PA                        1.1691e+01  9.0196e+00  1.2961
## I(log(number_of_wells) * is_Eastern_PA) -6.3947e+00  7.6333e+00 -0.8377
##                                       Pr(>|z|)
## (Intercept)                          0.0005752
## perc_white                           3.204e-06
## perc_unemployed                      0.0748295
## med_hh_income                        0.4732036
## total_pop                            0.3036160
## log(number_of_wells)                 0.8338453
## is_Eastern_PA                        0.1949307
## I(log(number_of_wells) * is_Eastern_PA) 0.4021803
##
## Rho: 0.16023, LR test value: 1.0009, p-value: 0.31709
## Asymptotic standard error: 0.14697
##     z-value: 1.0902, p-value: 0.27563
## Wald statistic: 1.1885, p-value: 0.27563
##
## Log likelihood: -278.7325 for lag model
## ML residual variance (sigma squared): 239.18, (sigma: 15.465)
## Number of observations: 67
## Number of parameters estimated: 10
## AIC: 577.47, (AIC for lm: 576.47)
## LM test for residual autocorrelation
## test value: 1.1328, p-value: 0.28717
```

```r
asthma_lag_W_5 <- lagsarlm(asthma_ED_visits ~ perc_white + perc_unemployed +
                             med_hh_income + total_pop + log(number_of_wells) +
                             is_Eastern_PA + I(log(number_of_wells) * is_Eastern_PA),
```

```
                          data = PA_county_tracts, listw = W_list_knn5)
summary(asthma_lag_W_5)
```

```
##
## Call:lagsarlm(formula = asthma_ED_visits ~ perc_white + perc_unemployed +
##     med_hh_income + total_pop + log(number_of_wells) + is_Eastern_PA +
##     I(log(number_of_wells) * is_Eastern_PA), data = PA_county_tracts,
##     listw = W_list_knn5)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -37.9069  -8.6996  -2.0096   5.1838  73.8245
##
## Type: lag
## Coefficients: (asymptotic standard errors)
##                                       Estimate  Std. Error z value
## (Intercept)                          1.4969e+02  4.4239e+01  3.3836
## perc_white                          -1.7470e+02  3.7796e+01 -4.6222
## perc_unemployed                      3.8074e+02  2.1352e+02  1.7832
## med_hh_income                        1.8893e-04  2.4201e-04  0.7807
## total_pop                           -1.2299e-05  1.1947e-05 -1.0295
## log(number_of_wells)                -2.2516e-01  1.0015e+00 -0.2248
## is_Eastern_PA                        1.1603e+01  9.0858e+00  1.2770
## I(log(number_of_wells) * is_Eastern_PA) -5.5386e+00  7.7086e+00 -0.7185
##                                       Pr(>|z|)
## (Intercept)                          0.0007154
## perc_white                           3.797e-06
## perc_unemployed                      0.0745561
## med_hh_income                        0.4349920
## total_pop                            0.3032618
## log(number_of_wells)                 0.8221137
## is_Eastern_PA                        0.2015974
## I(log(number_of_wells) * is_Eastern_PA) 0.4724481
##
## Rho: 0.14299, LR test value: 0.71863, p-value: 0.39659
## Asymptotic standard error: 0.15594
##     z-value: 0.91691, p-value: 0.35919
## Wald statistic: 0.84073, p-value: 0.35919
##
## Log likelihood: -278.8737 for lag model
## ML residual variance (sigma squared): 240.6, (sigma: 15.511)
## Number of observations: 67
## Number of parameters estimated: 10
## AIC: 577.75, (AIC for lm: 576.47)
## LM test for residual autocorrelation
## test value: 1.0659, p-value: 0.30187
```

Creating the spatial Durbin models.

```
asthma_Durbin_W <- lagsarlm(asthma_ED_visits ~ perc_white + perc_unemployed +
                              med_hh_income + total_pop_standardized +
                              log(number_of_wells) + is_Eastern_PA +
                              I(log(number_of_wells) * is_Eastern_PA),
```

```
                           data = PA_county_tracts, Durbin = T, listw = W_list)
summary(asthma_Durbin_W)
```

```
##
## Call:lagsarlm(formula = asthma_ED_visits ~ perc_white + perc_unemployed +
##     med_hh_income + total_pop_standardized + log(number_of_wells) +
##     is_Eastern_PA + I(log(number_of_wells) * is_Eastern_PA),
##     data = PA_county_tracts, listw = W_list, Durbin = T)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -36.2474  -6.4505  -2.9177   6.4123  70.7856
##
## Type: mixed
## Coefficients: (asymptotic standard errors)
##                                             Estimate  Std. Error z value
## (Intercept)                                7.9621e+01  1.6937e+02  0.4701
## perc_white                                -1.7165e+02  4.0756e+01 -4.2116
## perc_unemployed                            4.1431e+02  2.4109e+02  1.7185
## med_hh_income                              3.8255e-04  2.8022e-04  1.3652
## total_pop_standardized                    -1.0805e+01  2.0801e+01 -0.5195
## log(number_of_wells)                       1.2856e+00  1.4989e+00  0.8577
## is_Eastern_PA                              8.6081e+00  1.0029e+01  0.8583
## I(log(number_of_wells) * is_Eastern_PA)   -1.1505e+01  7.7124e+00 -1.4918
## lag.perc_white                             1.9408e+01  1.2353e+02  0.1571
## lag.perc_unemployed                       -4.9889e+01  4.8657e+02 -0.1025
## lag.med_hh_income                          6.5140e-04  6.4618e-04  1.0081
## lag.total_pop_standardized                -6.4281e+01  5.5348e+01 -1.1614
## lag.log(number_of_wells)                  -9.9548e-01  2.1490e+00 -0.4632
## lag.is_Eastern_PA                         -1.7247e-01  2.2213e+01 -0.0078
## lag.I(log(number_of_wells) * is_Eastern_PA)  2.1585e+01  1.6023e+01  1.3472
##                                            Pr(>|z|)
## (Intercept)                                 0.63827
## perc_white                                2.536e-05
## perc_unemployed                             0.08571
## med_hh_income                               0.17219
## total_pop_standardized                      0.60344
## log(number_of_wells)                        0.39104
## is_Eastern_PA                               0.39070
## I(log(number_of_wells) * is_Eastern_PA)     0.13575
## lag.perc_white                              0.87516
## lag.perc_unemployed                         0.91833
## lag.med_hh_income                           0.31342
## lag.total_pop_standardized                  0.24548
## lag.log(number_of_wells)                    0.64320
## lag.is_Eastern_PA                           0.99381
## lag.I(log(number_of_wells) * is_Eastern_PA)  0.17792
##
## Rho: 0.23599, LR test value: 1.4311, p-value: 0.23158
## Asymptotic standard error: 0.16786
##     z-value: 1.4058, p-value: 0.15978
## Wald statistic: 1.9764, p-value: 0.15978
##
```

```
## Log likelihood: -274.2266 for mixed model
## ML residual variance (sigma squared): 207.73, (sigma: 14.413)
## Number of observations: 67
## Number of parameters estimated: 17
## AIC: 582.45, (AIC for lm: 581.88)
## LM test for residual autocorrelation
## test value: 1.4681, p-value: 0.22565
```

```
asthma_Durbin_W_5 <- lagsarlm(asthma_ED_visits ~ perc_white + perc_unemployed +
                                med_hh_income + total_pop_standardized +
                                log(number_of_wells) + is_Eastern_PA +
                                I(log(number_of_wells) * is_Eastern_PA),
                          data = PA_county_tracts, Durbin = T, listw = W_list_knn5)
summary(asthma_Durbin_W_5)
```

```
##
## Call:lagsarlm(formula = asthma_ED_visits ~ perc_white + perc_unemployed +
##     med_hh_income + total_pop_standardized + log(number_of_wells) +
##     is_Eastern_PA + I(log(number_of_wells) * is_Eastern_PA),
##     data = PA_county_tracts, listw = W_list_knn5, Durbin = T)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -34.1688  -8.1430  -2.7057   6.4495  66.1966
##
## Type: mixed
## Coefficients: (asymptotic standard errors)
##                                            Estimate  Std. Error z value
## (Intercept)                              -3.4135e+01  1.6491e+02 -0.2070
## perc_white                               -1.5384e+02  3.9512e+01 -3.8935
## perc_unemployed                           5.6252e+02  2.3759e+02  2.3676
## med_hh_income                             6.4860e-04  3.0424e-04  2.1319
## total_pop_standardized                   -9.0481e+00  1.9448e+01 -0.4652
## log(number_of_wells)                      1.7680e+00  1.5070e+00  1.1732
## is_Eastern_PA                             1.0719e+01  9.5896e+00  1.1178
## I(log(number_of_wells) * is_Eastern_PA)  -1.4625e+01  7.8371e+00 -1.8662
## lag.perc_white                            8.6635e+01  1.1901e+02  0.7280
## lag.perc_unemployed                      -5.2146e+01  4.8499e+02 -0.1075
## lag.med_hh_income                         7.3994e-04  6.1892e-04  1.1955
## lag.total_pop_standardized               -6.2338e+01  4.7919e+01 -1.3009
## lag.log(number_of_wells)                 -7.0094e-01  2.0397e+00 -0.3436
## lag.is_Eastern_PA                         3.7144e+00  2.0548e+01  0.1808
## lag.I(log(number_of_wells) * is_Eastern_PA)  2.7131e+01  1.8574e+01  1.4607
##                                           Pr(>|z|)
## (Intercept)                                0.83602
## perc_white                                9.882e-05
## perc_unemployed                            0.01790
## med_hh_income                              0.03302
## total_pop_standardized                     0.64176
## log(number_of_wells)                       0.24073
## is_Eastern_PA                              0.26367
## I(log(number_of_wells) * is_Eastern_PA)    0.06202
## lag.perc_white                             0.46663
## lag.perc_unemployed                        0.91438
```

```
## lag.med_hh_income                               0.23188
## lag.total_pop_standardized                      0.19329
## lag.log(number_of_wells)                        0.73111
## lag.is_Eastern_PA                               0.85655
## lag.I(log(number_of_wells) * is_Eastern_PA)     0.14410
##
## Rho: 0.31667, LR test value: 2.5163, p-value: 0.11267
## Asymptotic standard error: 0.16481
##     z-value: 1.9214, p-value: 0.054681
## Wald statistic: 3.6918, p-value: 0.054681
##
## Log likelihood: -272.9367 for mixed model
## ML residual variance (sigma squared): 198.49, (sigma: 14.089)
## Number of observations: 67
## Number of parameters estimated: 17
## AIC: 579.87, (AIC for lm: 580.39)
## LM test for residual autocorrelation
## test value: 4.4822, p-value: 0.03425
```

Creating the spatial error models.

```
asthma_sar_W <- errorsarlm(asthma_ED_visits ~ perc_white + perc_unemployed +
                             med_hh_income + total_pop + log(number_of_wells) +
                             is_Eastern_PA + I(log(number_of_wells) * is_Eastern_PA),
                           data = PA_county_tracts, listw = W_list)
summary(asthma_sar_W)
```

```
##
## Call:errorsarlm(formula = asthma_ED_visits ~ perc_white + perc_unemployed +
##     med_hh_income + total_pop + log(number_of_wells) + is_Eastern_PA +
##     I(log(number_of_wells) * is_Eastern_PA), data = PA_county_tracts,
##     listw = W_list)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -38.1543  -7.4879  -1.2932   5.6584  73.4206
##
## Type: error
## Coefficients: (asymptotic standard errors)
##                                         Estimate  Std. Error z value
## (Intercept)                             1.4427e+02  4.1835e+01  3.4485
## perc_white                             -1.7336e+02  3.5559e+01 -4.8754
## perc_unemployed                         4.4616e+02  2.0711e+02  2.1543
## med_hh_income                           2.4705e-04  2.1624e-04  1.1425
## total_pop                              -7.6434e-06  1.1714e-05 -0.6525
## log(number_of_wells)                   -2.0873e-01  1.1088e+00 -0.1883
## is_Eastern_PA                           1.4218e+01  9.1124e+00  1.5603
## I(log(number_of_wells) * is_Eastern_PA) -9.6385e+00  7.5097e+00 -1.2835
##                                         Pr(>|z|)
## (Intercept)                            0.0005636
## perc_white                             1.086e-06
## perc_unemployed                        0.0312179
## med_hh_income                          0.2532443
```

```
## total_pop                                0.5140737
## log(number_of_wells)                     0.8506751
## is_Eastern_PA                            0.1187000
## I(log(number_of_wells) * is_Eastern_PA) 0.1993274
##
## Lambda: 0.3047, LR test value: 1.9686, p-value: 0.1606
## Asymptotic standard error: 0.16249
##     z-value: 1.8751, p-value: 0.060775
## Wald statistic: 3.5161, p-value: 0.060775
##
## Log likelihood: -278.2487 for error model
## ML residual variance (sigma squared): 232.3, (sigma: 15.241)
## Number of observations: 67
## Number of parameters estimated: 10
## AIC: 576.5, (AIC for lm: 576.47)
```

```
asthma_sar_W_5 <- errorsarlm(asthma_ED_visits ~ perc_white + perc_unemployed +
                               med_hh_income + total_pop + log(number_of_wells) +
                               is_Eastern_PA + I(log(number_of_wells) * is_Eastern_PA),
                             data = PA_county_tracts, listw = W_list_knn5)
summary(asthma_sar_W_5)
```

```
##
## Call:errorsarlm(formula = asthma_ED_visits ~ perc_white + perc_unemployed +
##     med_hh_income + total_pop + log(number_of_wells) + is_Eastern_PA +
##     I(log(number_of_wells) * is_Eastern_PA), data = PA_county_tracts,
##     listw = W_list_knn5)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -37.87071  -8.37811  -0.85753   5.36308  72.39648
##
## Type: error
## Coefficients: (asymptotic standard errors)
##                                          Estimate  Std. Error z value
## (Intercept)                              1.4254e+02  4.2337e+01  3.3668
## perc_white                              -1.7325e+02  3.5821e+01 -4.8365
## perc_unemployed                          4.3286e+02  2.1064e+02  2.0550
## med_hh_income                            2.9351e-04  2.1539e-04  1.3627
## total_pop                               -7.8360e-06  1.1646e-05 -0.6729
## log(number_of_wells)                    -1.6317e-01  1.1061e+00 -0.1475
## is_Eastern_PA                            1.3000e+01  9.0658e+00  1.4340
## I(log(number_of_wells) * is_Eastern_PA) -8.5326e+00  7.6512e+00 -1.1152
##                                          Pr(>|z|)
## (Intercept)                             0.0007605
## perc_white                              1.321e-06
## perc_unemployed                         0.0398825
## med_hh_income                           0.1729824
## total_pop                               0.5010325
## log(number_of_wells)                    0.8827166
## is_Eastern_PA                           0.1515714
## I(log(number_of_wells) * is_Eastern_PA) 0.2647635
##
## Lambda: 0.31488, LR test value: 1.8781, p-value: 0.17055
```

```
## Asymptotic standard error: 0.16836
##     z-value: 1.8702, p-value: 0.061451
## Wald statistic: 3.4978, p-value: 0.061451
##
## Log likelihood: -278.2939 for error model
## ML residual variance (sigma squared): 232.96, (sigma: 15.263)
## Number of observations: 67
## Number of parameters estimated: 10
## AIC: 576.59, (AIC for lm: 576.47)
```

Creating the CAR models.

```
asthma_car_W <- spautolm(asthma_ED_visits ~ perc_white + perc_unemployed +
                           med_hh_income + total_pop + log(number_of_wells) +
                           is_Eastern_PA + I(log(number_of_wells) * is_Eastern_PA),
                         data = PA_county_tracts, family = "CAR", listw = W_list)
```

```
## Warning in spautolm(asthma_ED_visits ~ perc_white + perc_unemployed +
## med_hh_income + : Non-symmetric spatial weights in CAR model
```

```
summary(asthma_car_W)
```

```
##
## Call: spautolm(formula = asthma_ED_visits ~ perc_white + perc_unemployed +
##     med_hh_income + total_pop + log(number_of_wells) + is_Eastern_PA +
##     I(log(number_of_wells) * is_Eastern_PA), data = PA_county_tracts,
##     listw = W_list, family = "CAR")
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -39.3304  -7.9124  -1.6971   5.2959  74.1708
##
## Coefficients:
##                                       Estimate  Std. Error z value
## (Intercept)                          1.5059e+02  4.2570e+01  3.5374
## perc_white                          -1.8155e+02  3.6360e+01 -4.9931
## perc_unemployed                      3.9034e+02  2.0605e+02  1.8944
## med_hh_income                        3.0755e-04  2.1514e-04  1.4296
## total_pop                           -9.0793e-06  1.1810e-05 -0.7688
## log(number_of_wells)                -2.3361e-02  1.0811e+00 -0.0216
## is_Eastern_PA                        1.3269e+01  9.0483e+00  1.4664
## I(log(number_of_wells) * is_Eastern_PA) -7.1682e+00  7.5984e+00 -0.9434
##                                      Pr(>|z|)
## (Intercept)                          0.0004041
## perc_white                           5.941e-07
## perc_unemployed                      0.0581719
## med_hh_income                        0.1528442
## total_pop                            0.4420399
## log(number_of_wells)                 0.9827605
## is_Eastern_PA                        0.1425280
## I(log(number_of_wells) * is_Eastern_PA) 0.3454845
##
```

```
## Lambda: 0.40763 LR test value: 1.3713 p-value: 0.24158
## Numerical Hessian standard error of lambda: 0.3672
##
## Log likelihood: -278.5473
## ML residual variance (sigma squared): 234.69, (sigma: 15.32)
## Number of observations: 67
## Number of parameters estimated: 10
## AIC: 577.09
```

```
asthma_car_W_5 <- spautolm(asthma_ED_visits ~ perc_white + perc_unemployed +
                           med_hh_income + total_pop + log(number_of_wells) +
                           is_Eastern_PA + I(log(number_of_wells) * is_Eastern_PA),
                     data = PA_county_tracts, family = "CAR", listw = W_list_knn5)
```

```
## Warning in spautolm(asthma_ED_visits ~ perc_white + perc_unemployed +
## med_hh_income + : Non-symmetric spatial weights in CAR model
```

```
summary(asthma_car_W_5)
```

```
##
## Call: spautolm(formula = asthma_ED_visits ~ perc_white + perc_unemployed +
##     med_hh_income + total_pop + log(number_of_wells) + is_Eastern_PA +
##     I(log(number_of_wells) * is_Eastern_PA), data = PA_county_tracts,
##     listw = W_list_knn5, family = "CAR")
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -39.0973  -8.7373  -1.3667   6.0161  71.7084
##
## Coefficients:
##                                         Estimate  Std. Error z value
## (Intercept)                             1.5070e+02  4.3156e+01  3.4920
## perc_white                             -1.7903e+02  3.6889e+01 -4.8532
## perc_unemployed                         3.9735e+02  2.0730e+02  1.9168
## med_hh_income                           2.6012e-04  2.1365e-04  1.2175
## total_pop                              -1.1813e-05  1.1834e-05 -0.9982
## log(number_of_wells)                   -7.4438e-02  1.0632e+00 -0.0700
## is_Eastern_PA                           1.5989e+01  8.9915e+00  1.7782
## I(log(number_of_wells) * is_Eastern_PA) -1.0646e+01  7.6916e+00 -1.3842
##                                          Pr(>|z|)
## (Intercept)                             0.0004794
## perc_white                              1.215e-06
## perc_unemployed                         0.0552660
## med_hh_income                           0.2234076
## total_pop                               0.3181748
## log(number_of_wells)                    0.9441842
## is_Eastern_PA                           0.0753682
## I(log(number_of_wells) * is_Eastern_PA) 0.1663083
##
## Lambda: 0.34086 LR test value: 1.0617 p-value: 0.30284
## Numerical Hessian standard error of lambda: 0.39391
##
## Log likelihood: -278.7021
```

```
## ML residual variance (sigma squared): 237.6, (sigma: 15.414)
## Number of observations: 67
## Number of parameters estimated: 10
## AIC: 577.4
```

The best model for the asthma rate.

```
asthma_sar_W <- errorsarlm(asthma_ED_visits ~ perc_white + perc_unemployed +
                             med_hh_income + total_pop + log(number_of_wells) +
                             is_Eastern_PA + I(log(number_of_wells) * is_Eastern_PA),
                           data = PA_county_tracts, listw = W_list)
summary(asthma_sar_W)
```

```
##
## Call:errorsarlm(formula = asthma_ED_visits ~ perc_white + perc_unemployed +
##     med_hh_income + total_pop + log(number_of_wells) + is_Eastern_PA +
##     I(log(number_of_wells) * is_Eastern_PA), data = PA_county_tracts,
##     listw = W_list)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -38.1543  -7.4879  -1.2932   5.6584  73.4206
##
## Type: error
## Coefficients: (asymptotic standard errors)
##                                        Estimate  Std. Error z value
## (Intercept)                           1.4427e+02  4.1835e+01  3.4485
## perc_white                           -1.7336e+02  3.5559e+01 -4.8754
## perc_unemployed                       4.4616e+02  2.0711e+02  2.1543
## med_hh_income                         2.4705e-04  2.1624e-04  1.1425
## total_pop                            -7.6434e-06  1.1714e-05 -0.6525
## log(number_of_wells)                 -2.0873e-01  1.1088e+00 -0.1883
## is_Eastern_PA                         1.4218e+01  9.1124e+00  1.5603
## I(log(number_of_wells) * is_Eastern_PA) -9.6385e+00  7.5097e+00 -1.2835
##                                        Pr(>|z|)
## (Intercept)                           0.0005636
## perc_white                            1.086e-06
## perc_unemployed                       0.0312179
## med_hh_income                         0.2532443
## total_pop                             0.5140737
## log(number_of_wells)                  0.8506751
## is_Eastern_PA                         0.1187000
## I(log(number_of_wells) * is_Eastern_PA) 0.1993274
##
## Lambda: 0.3047, LR test value: 1.9686, p-value: 0.1606
## Asymptotic standard error: 0.16249
##     z-value: 1.8751, p-value: 0.060775
## Wald statistic: 3.5161, p-value: 0.060775
##
## Log likelihood: -278.2487 for error model
## ML residual variance (sigma squared): 232.3, (sigma: 15.241)
## Number of observations: 67
## Number of parameters estimated: 10
## AIC: 576.5, (AIC for lm: 576.47)
```