

# Concentrated Animal Feeding Operations: Impact on Incidence of Lung Cancer in Iowa



**ZACHARY BALES  
LINETTE DE LA ROSA  
VANESSA SALVARY**

**ISIB 2013  
THE UNIVERSITY OF IOWA  
JULY 18, 2013**

**MENTOR: MARY KATHRYN COWLES**



## Outline

Introduction

Variables

Data Analysis

Conclusions

References



# Introduction



- Background
- Definitions
  - AFO
  - CAFO
- Types of CAFOs
  - Open Feedlots
  - Confinements
- Facilities
  - 3055

# Pictures of CAFOs



# Open Feedlots



<http://iafarmwife.com/tag/cattle-confinement/>

<http://www.blueplanetgreenliving.com/2010/01/06/plains-justice-cafos-and-threats-to-human-health/>

# Concentrated Animal Feeding Operations: Impact on Incidence of Lung Cancer in Iowa



- Health Effects of Proximity to CAFOs
  - Common Emissions
    - Ammonia
    - Carbon Dioxide
    - Hydrogen Sulfide
    - Endotoxins
  - Several Diseases
    - Asthma
    - Chronic Bronchitis
    - Asphyxiation
    - Pneumonia
- Ecological Study to Assess the Impact of CAFOs on Incidence
- Incidence

# Research Question



**DOES THE PROXIMITY OF CAFOS HAVE AN  
IMPACT ON LUNG CANCER RATES?**



## Animal Units

- Independent variable
- Serves as our measure of C.A.F.O. presence
- EPA defined unit of measure
- Allows us to standardize and quantify different types of animals
- Calculated by multiplying number of each animal by an equivalency factor





**Known variables being controlled:**

❑ **Smoking**

**Smoking**

Alveoli that carry O<sub>2</sub> destroyed

Cilia obstructed and are no longer able to naturally clean lungs( tar/ pollen/ debris build up) with all of the toxins/chemicals

❑ **Radon**

**Radon**

2<sup>nd</sup> to smoking for leading causes of lung cancer

Because of location, IA has the highest levels of radon in the entire country

Dangerous because of radioactivity:

Uranium decomposes from rock and soil producing Radon as byproduct

❑ **Minority status**

**Minority status**

Strong correlation between socio- economic status and health care

Using census data, we compared 8 different ethnic groups

1. White 2. Black or African American 3. American Indian and Alaska Native 4. Asian alone 5. Native Hawaiian and other Pacific Islander 6. Two or more races 7. Hispanic or Latino 8. White not Hispanic or Latino



## Other variables being controlled:

- ❑ **Urbanicity**

### Urbanicity

Medium metro (pop.) 250,000 - 999,9999

Small metro (pop.) < 250,000

Micropolitan (pop.) 10,000 - 50,000

Noncore (pop.) 2,500 - 10,000

- ❑ **Air Quality**

### Air quality and Gender

Considered, but insufficient data

Good air days far out weighed medium, poor and unhealthy

- ❑ **Gender**



## Maps:

Lung Cancer Incidence

Urbanicity

Race

Smoking

Radon

C.A.F.O.s

## Age Adjusted Lung Cancer Incidence Rates Per 100,000



Response variable



## Maps:

Lung Cancer Incidence

Urbanicity

Race

Smoking

Radon

C.A.F.O.s

## Urbanicity by County



**West:** Counties bordering Omaha, NE.

**Center:** Des Moines.

**East:** Iowa City, Cedar Rapids, and Quad Cities.



## Maps:

Lung Cancer Incidence

Urbanicity

**Race**

Smoking

Radon

C.A.F.O.s

### Percentage of Black or African Americans by County





## Maps:

Lung Cancer Incidence

Urbanicity

Race

**Smoking**

Radon

C.A.F.O.s

## Smoking Rates by County



Strong spatial correlation



## Maps:

Lung Cancer Incidence

Urbanicity

Race

Smoking

**Radon**

C.A.F.O.s

## Uranium Hotspots by County



Strong spatial clustering

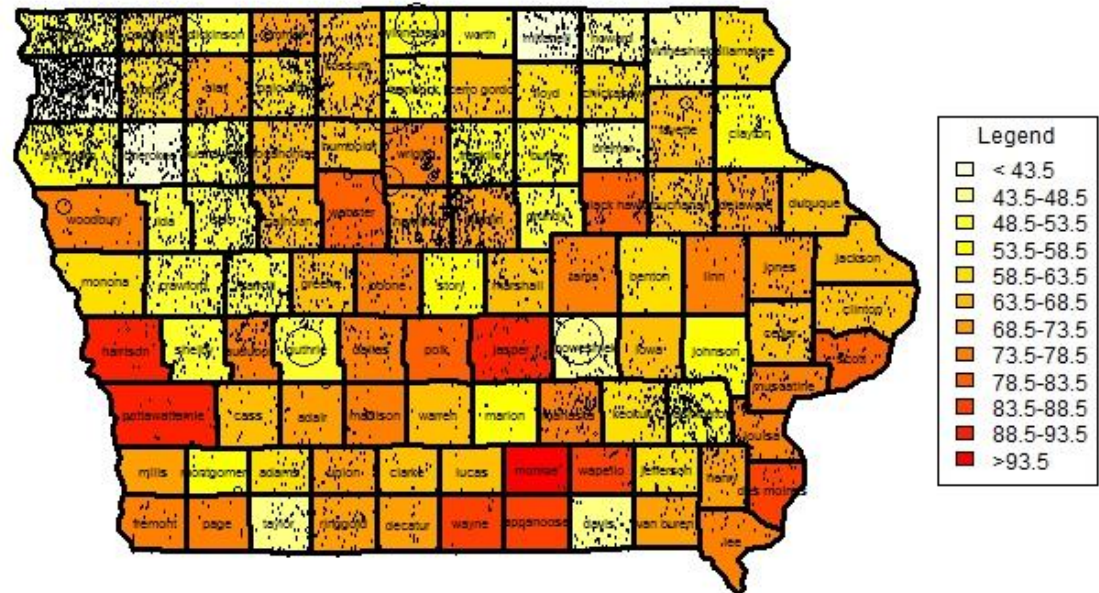
Counties that are close together have similar coloring and characteristics



## Maps:

- Lung Cancer Incidence
- Urbanicity
- Race
- Smoking
- Radon
- C.A.F.O.s**

### Lung Cancer Incidence Rates per 100,000 with Relative Size of C.A.F.O.s by County



This represents what we are **MOST** interested in





Expected to see a strong positive relationship between the predictors and Incidence

Relationship between predictors should be minimal to avoid explaining the same information in the model

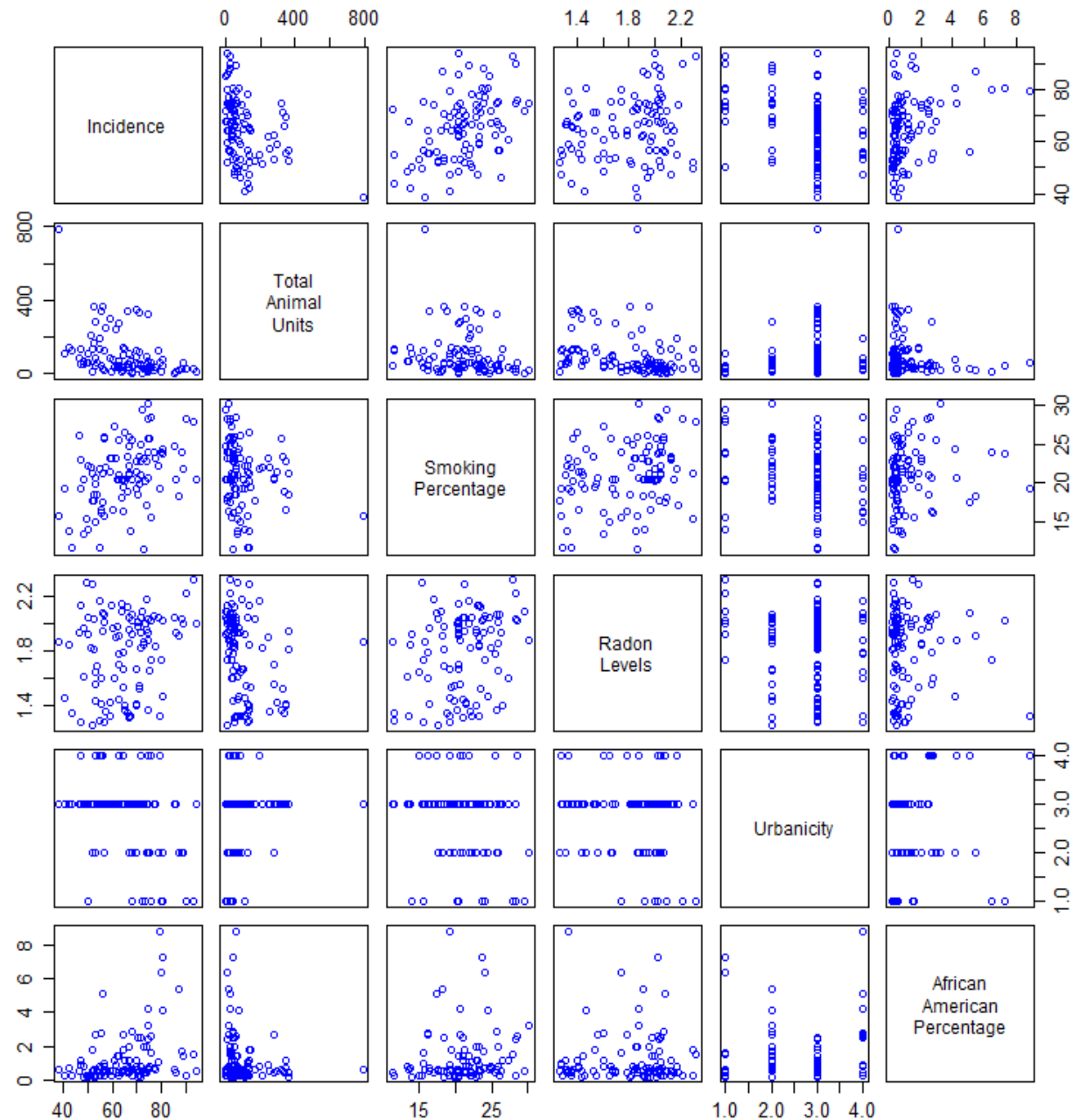
Relationship for Smoking, Radon, and Race as expected

Relationship for Total Animal Units was opposite

All predictor/Incidence relationships are significant

(p-value < 0.05)

## Bivariate Relationships Among CAFO Variables



# Models



**STANDARD LINEAR MODEL**  
**BAYESIAN SPATIAL MODEL**



## Linear Model

Which one

Why use it

How it was used

- The Linear Model used was the **lm ( )** function in R
- Used for the easy diagnostic tools
  - Plots
  - Charts
  - P – Value (significance)
- The method of use was as a **starting point** to check the known relationships
- Continued to use to see the spatial correlations impact



# Bayesian Spatial Model

Which one

Why use it

How it was used

Duncan Lee (2013). CARBayes:  
Spatial areal unit modelling. R  
package version 1.4. [http://CRAN.R-  
project.org/package=CARBayes](http://CRAN.R-project.org/package=CARBayes)

Reinhard Furrer, Stephan R. Sain  
(2010). spam: A Sparse Matrix R  
Package with Emphasis on MCMC  
Methods for Gaussian Markov  
Random Fields. Journal of Statistical  
Software, 36(10), 1-25. URL  
<http://www.jstatsoft.org/v36/i10/>.

- The model used was **gaussian.iarCAR()** in the “CARBayes” package in R
- It was used to control for **spatial correlation** in the residuals of the variables being tested
  - This was done with a **correlation matrix** of 0's and 1's
    - ✦ 0 if the counties are not adjacent
    - ✦ 1 if they share a border
      - The spam package was used to get this
- Note: The model works by drawing 1000 random samples
  - The seed for every random process was 9113



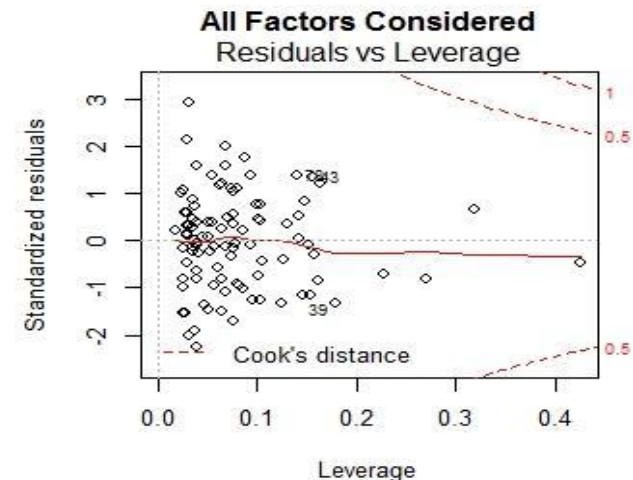
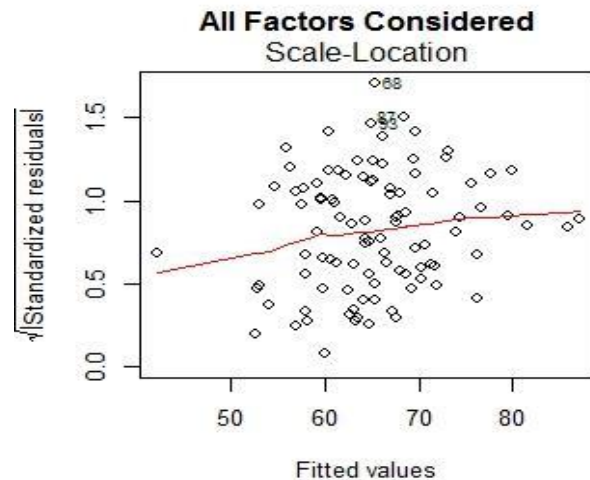
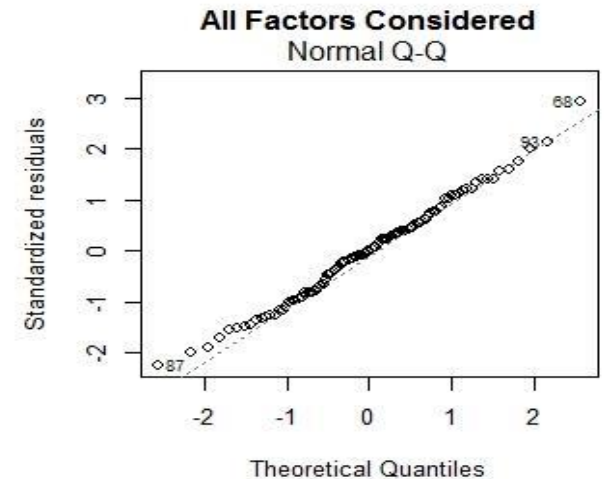
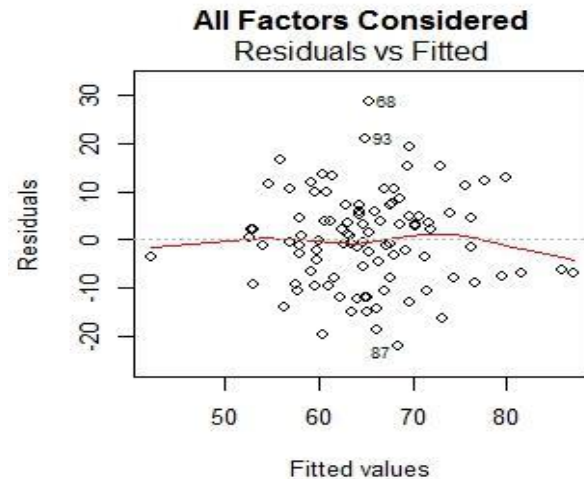
All Factors considers **all** of the variables we account for in the same model

No large curves on the left plots -- **linearity assumption is met**

Normal Q-Q – **residuals are normally distributed**

Residuals vs. Leverage – **No points drastically impacting the data**

Real data never matches perfectly so we check for **influential data** to correct for



**Plot of Linear Model With All 5 Variables Considered**



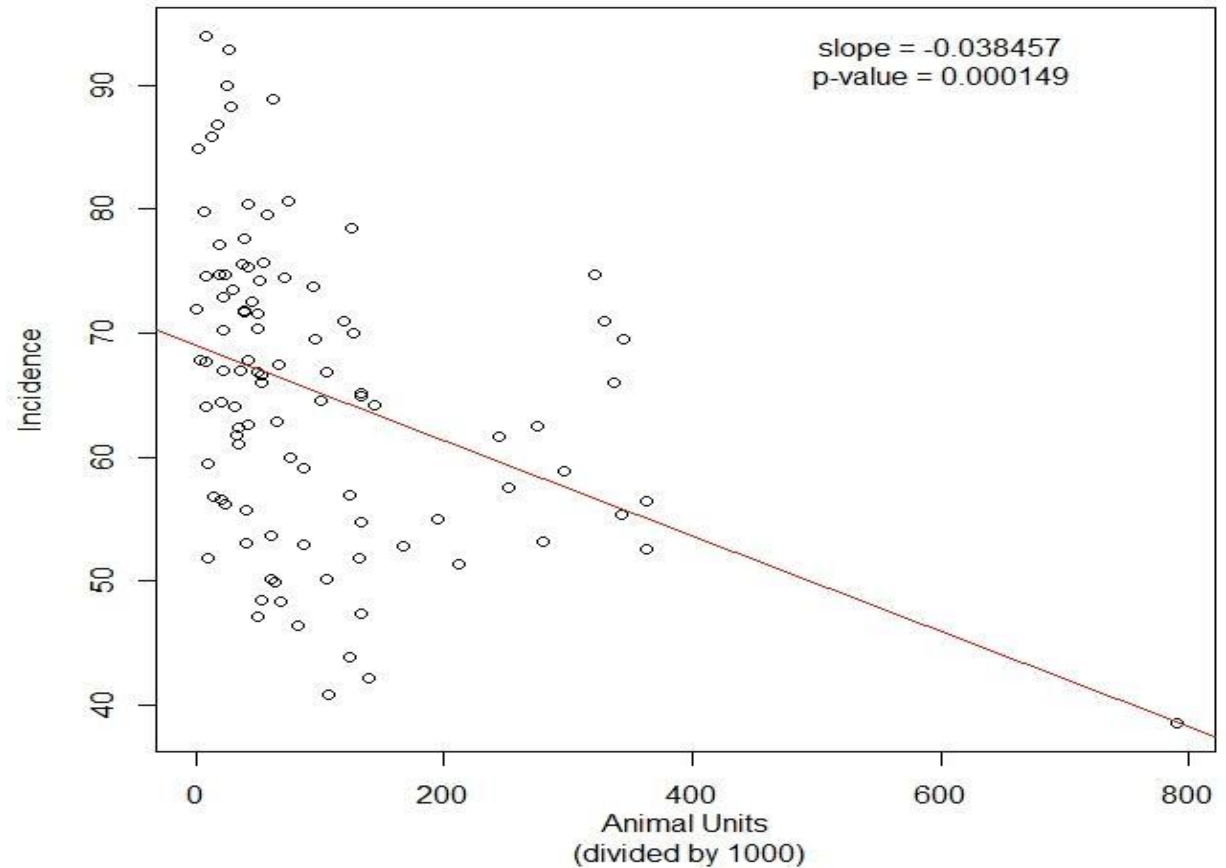
Sioux County –

Extremely **high** number  
of Animal Units

Extremely **low** incidence  
rate

Thus very **influential**

**Animal Units vs. Lung Cancer**



**Finding Influential Points**

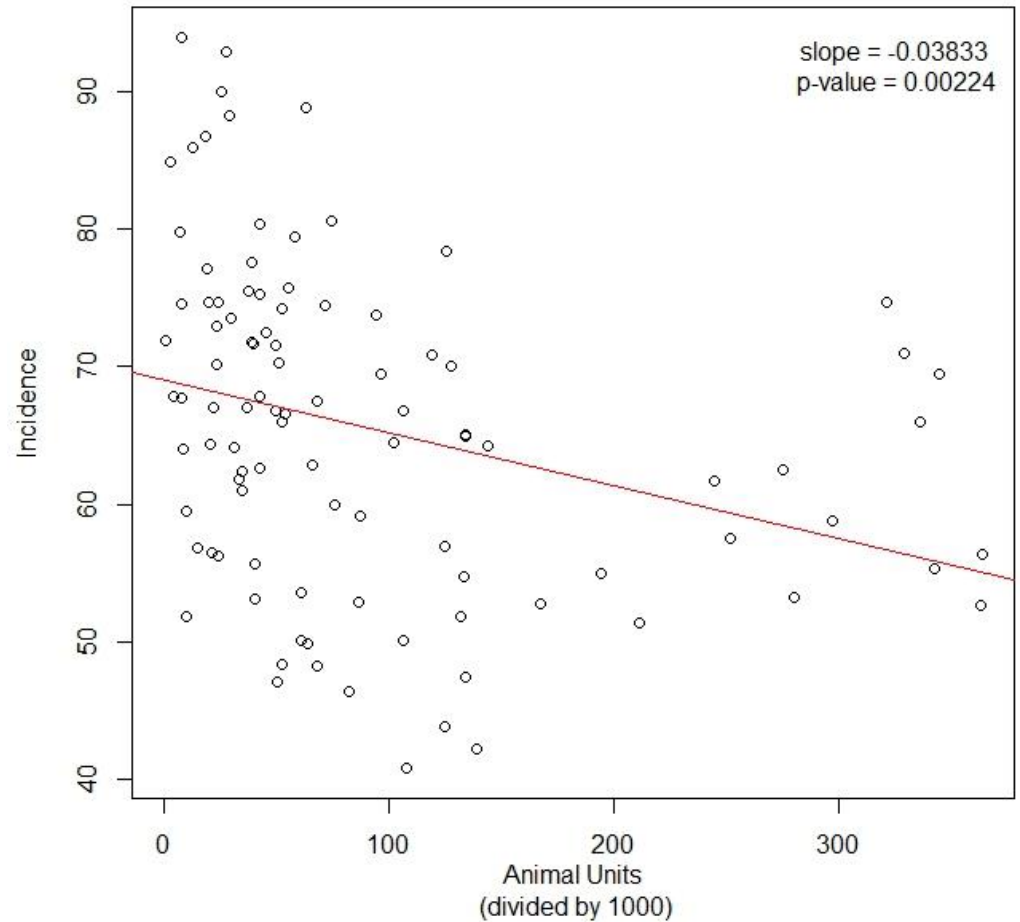


Remove Sioux County -

p - value increases

Still highly significant  
negative slope

**Animal Units vs. Incidence of Lung Cancer  
(without Sioux County)**



**Controlling for Influential Points**

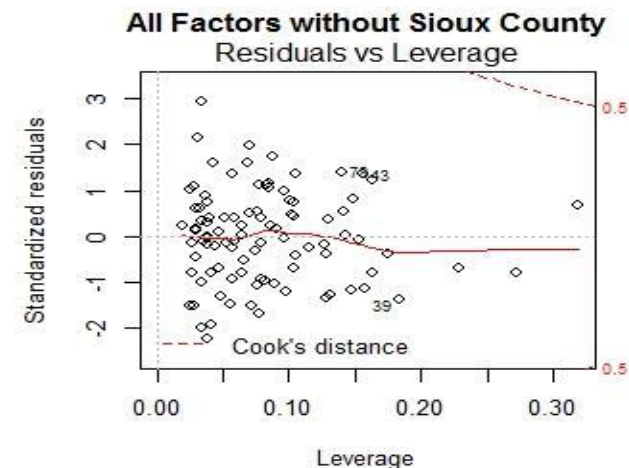
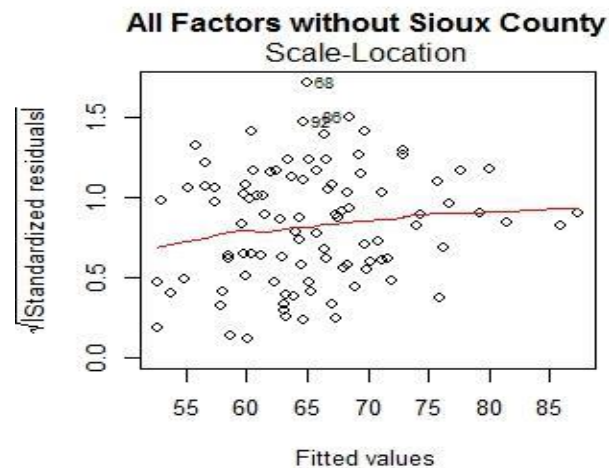
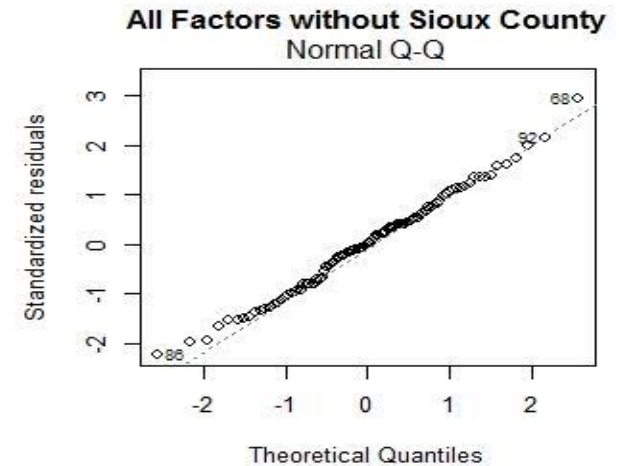
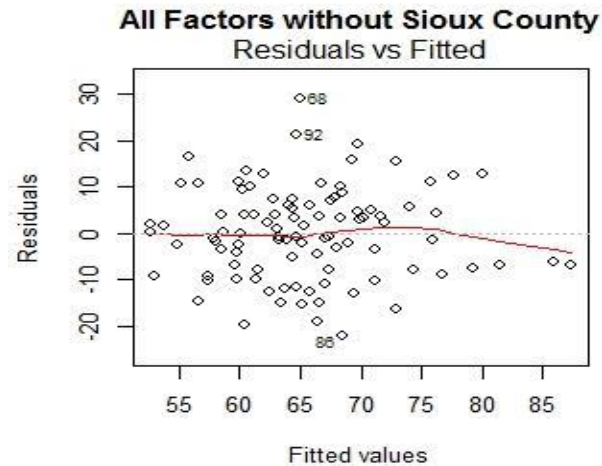


All Factors is still all the variables but applied to the 98 counties **excluding Sioux**

Slightly better than before

Not a huge difference so we will **continue to check both models**

i.e . with and without Sioux County



**Plot of Linear Model With All 5 Variables Considered**



| <i>Multiple Linear Model Outputs</i> |         |          |          |
|--------------------------------------|---------|----------|----------|
| All Factors                          |         |          |          |
| Variable                             | Slope   | Lower CI | Upper CI |
| Total Animal Units                   | -0.0245 | -0.0431  | -0.0058  |
| Smoking Percentage                   | 0.89023 | 0.35785  | 1.42262  |
| African American Percentage          | 2.0622  | 0.54063  | 3.58377  |
| Radon                                | 1.71563 | -5.8818  | 9.31303  |
| Noncore                              | -5.8028 | -13.485  | 1.87993  |
| Micropolitan                         | -3.2588 | -11.644  | 5.12608  |
| Small Metro                          | -10.771 | -19.941  | -1.602   |
| Without Sioux                        |         |          |          |
| Variable                             | Slope   | Lower CI | Upper CI |
| Total Animal Units                   | -0.0209 | -0.0452  | 0.00342  |
| Smoking Percentage                   | 0.8755  | 0.33697  | 1.41403  |
| African American Percentage          | 2.0841  | 0.55278  | 3.61543  |
| Radon                                | 2.23032 | -5.7164  | 10.177   |
| Noncore                              | -5.9224 | -13.657  | 1.81205  |
| Micropolitan                         | -3.2424 | -11.528  | 5.18055  |
| Small Metro                          | -10.804 | -20.016  | -1.5921  |

| <i>Bayesian Spatial Model Outputs</i> |         |          |          |
|---------------------------------------|---------|----------|----------|
| All Factors                           |         |          |          |
| Variable                              | Slope   | Lower CI | Upper CI |
| Total Animal Units                    | -0.0218 | -0.0426  | 0.0009   |
| Smoking Percentage                    | 0.8941  | 0.3564   | 1.4339   |
| African American Percentage           | 2.0073  | 0.5234   | 3.569    |
| Radon                                 | -0.8018 | -13.425  | 9.8912   |
| Noncore                               | -4.7763 | -12.474  | 3.2218   |
| Micropolitan                          | -2.8467 | -11.018  | 5.4981   |
| Small Metro                           | -9.3549 | -18.925  | -0.3632  |
| Without Sioux                         |         |          |          |
| Variable                              | Slope   | Lower CI | Upper CI |
| Total Animal Units                    | -0.0189 | -0.0466  | 0.0094   |
| Smoking Percentage                    | 0.8895  | 0.3492   | 1.4105   |
| African American Percentage           | 2.0408  | 0.4034   | 3.6268   |
| Radon                                 | 0.7463  | -11.241  | 11.582   |
| Noncore                               | -5.1676 | -13.011  | 2.819    |
| Micropolitan                          | -2.9924 | -12.023  | 5.2894   |
| Small Metro                           | -10.062 | -19.623  | 0.3198   |



## Making the Best Fit Model

Radon is **always insignificant** – with the highest p - value

Radon is **highly spatially correlated**

Controlling for spatial correlation then **may cause Radon to become less significant** in the model

- Due to this, Radon was **removed**
- Radon is a **KNOWN** predictor of **lung cancer**
  - Thus we report two models
    - ✦ With Radon
    - ✦ Without Radon

| <i>Multiple Linear Model Outputs</i> |         |          |          |
|--------------------------------------|---------|----------|----------|
| Without Radon                        |         |          |          |
| Variable                             | Slope   | Lower CI | Upper CI |
| Total Animal Units                   | -0.0255 | -0.0435  | -0.0075  |
| Smoking Percentage                   | 0.91862 | 0.40362  | 1.43363  |
| African American Percentage          | 2.03462 | 0.52478  | 3.54446  |
| Noncore                              | -6.0832 | -13.631  | 1.46432  |
| Micropolitan                         | -3.6484 | -11.817  | 4.52016  |
| Small Metro                          | -11.035 | -20.089  | -1.9809  |
| Without Sioux and Radon              |         |          |          |
| Variable                             | Slope   | Lower CI | Upper CI |
| Total Animal Units                   | -0.0233 | -0.0459  | -0.0008  |
| Smoking Percentage                   | 0.9145  | 0.39625  | 1.43274  |
| African American Percentage          | 2.04366 | 0.52514  | 3.56218  |
| Noncore                              | -6.2184 | -13.85   | 1.41365  |
| Micropolitan                         | -3.715  | -11.935  | 4.50529  |
| Small Metro                          | -11.109 | -20.22   | -1.9975  |

| <i>Bayesian Spatial Model Outputs</i> |         |          |          |
|---------------------------------------|---------|----------|----------|
| Without Radon                         |         |          |          |
| Variable                              | Slope   | Lower CI | Upper CI |
| Total Animal Units                    | -0.021  | -0.0407  | -0.0004  |
| Smoking Percentage                    | 0.8833  | 0.3843   | 1.4463   |
| African American Percentage           | 2.023   | 0.5058   | 3.5311   |
| Noncore                               | -4.7598 | -12.372  | 2.8466   |
| Micropolitan                          | -2.2872 | -10.671  | 5.4675   |
| Small Metro                           | -9.3249 | -19.296  | -0.1964  |
| Without Sioux and Radon               |         |          |          |
| Variable                              | Slope   | Lower CI | Upper CI |
| Total Animal Units                    | -0.0245 | -0.0472  | 0.0001   |
| Smoking Percentage                    | 0.9212  | 0.3877   | 1.4546   |
| African American Percentage           | 2.0658  | 0.5768   | 3.5044   |
| Noncore                               | -4.9938 | -12.628  | 2.281    |
| Micropolitan                          | -2.2729 | -10.579  | 5.8319   |
| Small Metro                           | -9.8094 | -18.831  | -0.7901  |



## Final Model

Set.seed(9113)

Gaussian.iarCAR( )

- Bayesian Spatial Model
- Age Adjusted Incidence of Lung Cancer Rates per 100,000 people in Iowa by County as **Dependent Variable**
- Total Animal Units as **Independent Variable** being studied
- **Control** for Smoking, Radon, Race, Urbanicity, and Spatial Correlation



## Conclusion

The predictor variable of greatest interest, total animal units, was found to be insignificant after controlling for other variables

Why would a negative relationship be happening?

$R^2$  values still small  $\rightarrow$  not a large proportion of the variability in response variable is explained by the model

Spatial correlation captures effects of missing variables

Limitations: ecological study



# Looking Forward

Given access to individual data, it's possible that we may find a positive relationship between animal units and lung cancer.

# References



- Duncan Lee (2013). CARBayes: Spatial areal unit modelling. R package version 1.4. <http://CRAN.R-project.org/package=CARBayes>
- Original S code by Richard A. Becker and Allan R. Wilks. R version by Ray Brownrigg. Enhancements by Thomas P Minka <tpminka@media.mit.edu> (2013). maps: Draw Geographical Maps. R package version 2.3-2. <http://CRAN.R-project.org/package=maps>
- Reinhard Furrer, Stephan R. Sain (2010). spam: A Sparse Matrix R Package with Emphasis on MCMC Methods for Gaussian Markov Random Fields. Journal of Statistical Software, 36(10), 1-25. URL <http://www.jstatsoft.org/v36/i10/>.
- RStudio Team (2012). RStudio: Integrated Development for R. RStudio, Inc., Boston, MA URL <http://www.rstudio.com/>.