

Healthcare Cost and Utilization Project

Akashdeep Jaswal

10/31/2016

About Healthcare Cost and Utilization Project (HCUP):

HCUP is the Nation's most comprehensive source of hospital data, including information on in-patient care, ambulatory care, and emergency department visits. HCUP enables researchers, insurers, policymakers and others to study health care delivery and patient outcomes over time, and at the national, regional, State, and community levels. For details about this initiative see

```
# Importing Packages
library(tidyverse)
library(plm)
library(ggplot2)
library(lattice)
```

1. Importing Data

1.1 Reading Inpatient Hospitalization Data

Variable encoding scheme and definitions for HCUP Data: [here](#)

```
setwd("~/Desktop/AkashPersonal/hcup-population")

# Reading the Hospitalization Data
hosp.data <- NULL
for (year in seq(2004,2013,by=1)) {

  # path name on how to read the file
  file <- paste('data/HospData/WA_SIDC_',year,'_CORE','.rdat',sep='')
  print (c("Reading data from", file))

  # Loading RDat files and Appending to DF
  load(file)
  cols <- c("ayear","amonth","pstco2","zip","pl_ur_cat4","age","female","dx1","dxccs1")
  temp <- data.frame(temp)[cols]
  hosp.data <- rbind(hosp.data,temp)

  # Tabular Dataframe
  hosp.data <- tbl_df(hosp.data)
}
```

```
## [1] "Reading data from"
## [2] "data/HospData/WA_SIDC_2004_CORE.rdat"
## [1] "Reading data from"
## [2] "data/HospData/WA_SIDC_2005_CORE.rdat"
## [1] "Reading data from"
## [2] "data/HospData/WA_SIDC_2006_CORE.rdat"
## [1] "Reading data from"
```

```
## [2] "data/HospData/WA_SIDC_2007_CORE.rdat"
## [1] "Reading data from"
## [2] "data/HospData/WA_SIDC_2008_CORE.rdat"
## [1] "Reading data from"
## [2] "data/HospData/WA_SIDC_2009_CORE.rdat"
## [1] "Reading data from"
## [2] "data/HospData/WA_SIDC_2010_CORE.rdat"
## [1] "Reading data from"
## [2] "data/HospData/WA_SIDC_2011_CORE.rdat"
## [1] "Reading data from"
## [2] "data/HospData/WA_SIDC_2012_CORE.rdat"
## [1] "Reading data from"
## [2] "data/HospData/WA_SIDC_2013_CORE.rdat"
```

1.2 Importing Vaccination Data and Census Demographics

```
vaccine.data <- read.csv('~/.Desktop/AkashPersonal/hcup-population/data/Vaccine_clean_data.csv')
census.data <- read.csv('~/.Desktop/AkashPersonal/hcup-population/data/Final_Population_Interpolated.csv')
```

2. Data Wrangling

2.1 Data Type Casting

```
# Recasting Variables appropriately - Hospitalization
hosp.data$zip <- as.integer(hosp.data$zip)
hosp.data$dx1 <- as.integer(hosp.data$dx1)
hosp.data$pl_ur_cat4 <- as.integer(hosp.data$pl_ur_cat4)
hosp.data <- hosp.data %>%
  mutate(urban_rural=ifelse(pl_ur_cat4==1, "Large Metropolitan",
                            ifelse(pl_ur_cat4==2, "Small Metropolitan",
                            ifelse(pl_ur_cat4==3, "Micropolitan",
                            "Not Metro/Micro")))))
hosp.data$urban_rural <- as.factor(hosp.data$urban_rural)

# Demographic data from US Census (Monthly Interpolated by WA County)
census.data$year <- as.integer(census.data$year)
```

2.2 Filtering, Aggregating and Mapping Data

```
# Filtering, Aggregating and Summarizing Hosp Data
hosp.filtered <- hosp.data %>%
  filter(age >= 0 & age <= 4 & ayear >= 2004) %>%
  mutate(AgeGroup=1) %>%
  group_by(ayear, amonth, zip, pstco2, dx1, dxccs1, AgeGroup, urban_rural) %>%
  summarise(hospitalizations=n())

# Zip Code - County FIPS Mapping
zip.fips.mapping <- hosp.data %>%
  select(zip, pstco2) %>%
  group_by(zip, pstco2) %>%
  summarise(n())
```

```

# Mapping the FIPS County codes to Zip Codes in Vaccine Data
vaccine.data <- merge(vaccine.data, zip.fips.mapping[c('zip','pstco2')],
                      by.x=c("ADDRESS_ZIP"), by.y=c("zip"))
vaccine.data <- vaccine.data[ , !duplicated(colnames(vaccine.data))]

# Cleaning Census Data
census.clean <- census.data %>%
  select(year, month, County, FIPS, AgeGroup, Population) %>%
  group_by(year, month, County, FIPS, AgeGroup) %>%
  summarise(Population=sum(Population))

```

3. Building Fixed Effect Models

3.1 All Cause Hospitalization - Vaccination Fixed Effects

```

# All Cause Hospitalization
all.cause.hosp <- hosp.filtered %>%
  select(ayear, amonth, AgeGroup, hospitalizations, pstco2, urban_rural) %>%
  group_by(ayear, amonth, pstco2, AgeGroup, urban_rural) %>%
  summarise(patients=sum(hospitalizations))

```

Adding missing grouping variables: `zip`, `dx1`, `dxccs1`

```

all.cause.hosp <- merge(all.cause.hosp, census.clean, by.x=c("ayear","amonth","pstco2","AgeGroup"),
                      by.y=c("year","month","FIPS","AgeGroup"))
all.cause.hosp <- mutate(all.cause.hosp, hosp_norm = patients/Population)

names(all.cause.hosp)[1] <- 'year'
names(all.cause.hosp)[2] <- 'month'
names(all.cause.hosp)[3] <- 'FIPS'

vaccine.filtered <- vaccine.data %>%
  select(year, month, FIPS=pstco2, X4313314.SERIES, TOTAL.POPULATION) %>%
  group_by(year, month, FIPS) %>%
  summarise(X4313314_Norm=sum(X4313314.SERIES)/sum(TOTAL.POPULATION),
            X4313314=sum(X4313314.SERIES), TOTAL_POPULATION=sum(TOTAL.POPULATION))

all.cause.merge <- merge(all.cause.hosp, vaccine.filtered, by=c("year","month","FIPS"))
all.cause.merge$t_index <- interaction(all.cause.merge$year, all.cause.merge$month, sep=".")
summary(all.cause.merge)

```

```

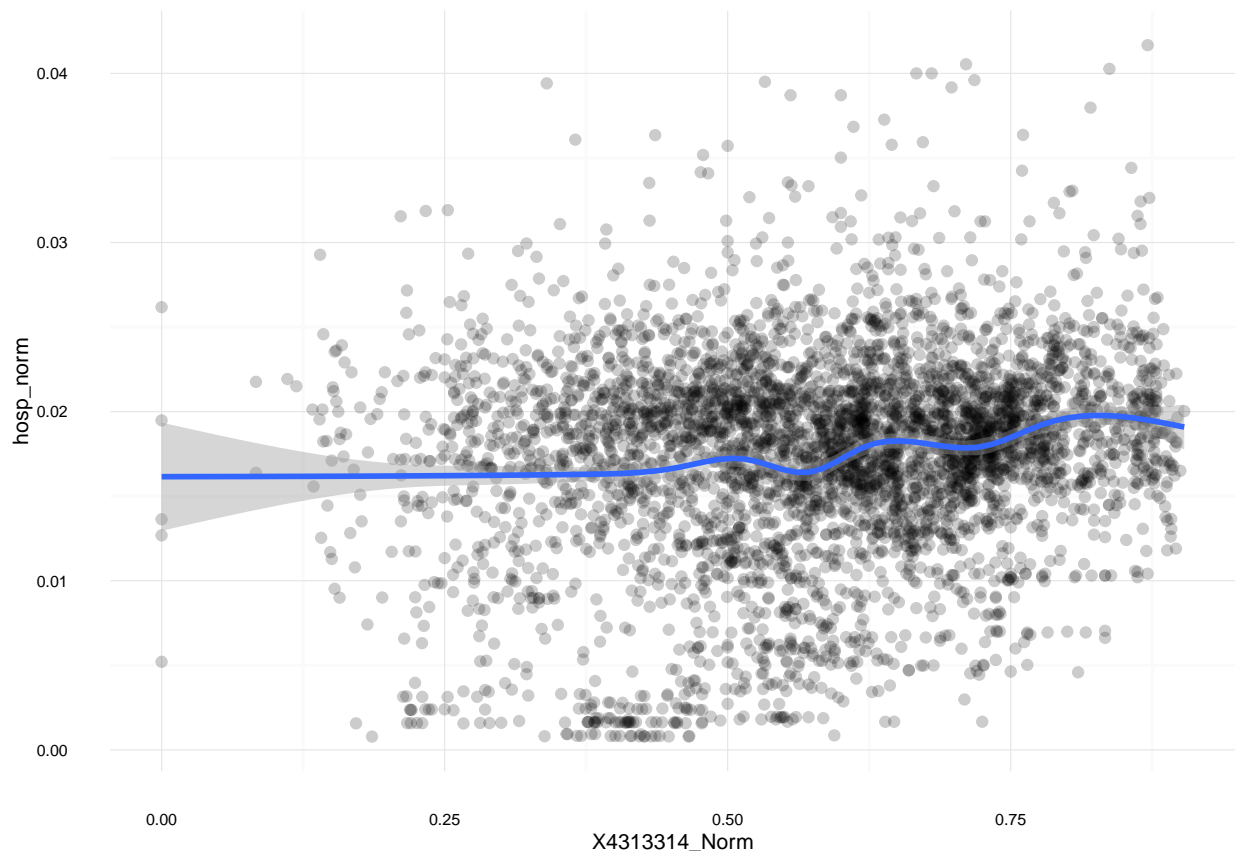
##      year      month      FIPS      AgeGroup
## Min.   :2004   Min.    : 1.000   Min.    :53001   Min.    :1
## 1st Qu.:2006   1st Qu.: 3.750   1st Qu.:53019   1st Qu.:1
## Median :2009   Median : 6.000   Median :53039   Median :1
## Mean   :2009   Mean    : 6.494   Mean    :53039   Mean    :1
## 3rd Qu.:2011   3rd Qu.: 9.000   3rd Qu.:53059   3rd Qu.:1
## Max.   :2013   Max.    :12.000   Max.    :53077   Max.    :1
##
##      urban_rural      patients      County
## Large Metropolitan: 595   Min.    : 1.0   Adams County : 120

```

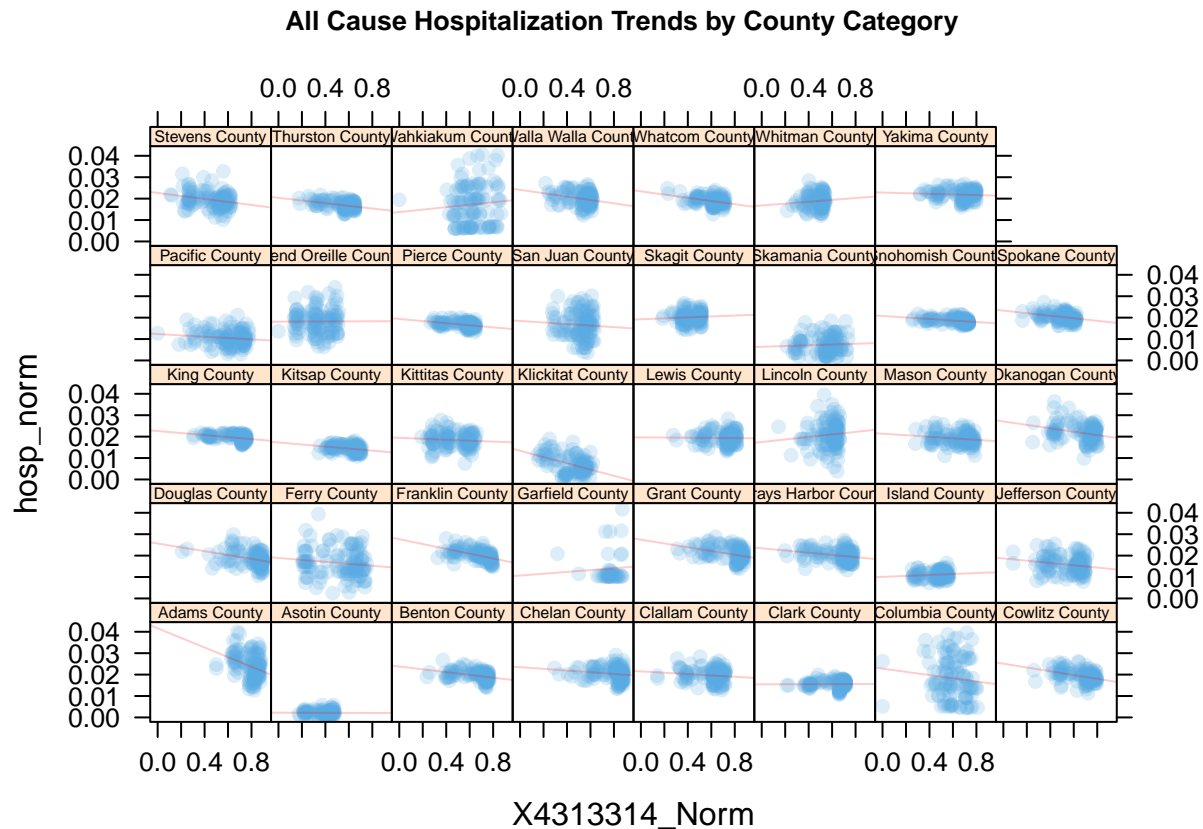
```
## Micropolitan      :1080   1st Qu.: 14.0   Benton County : 120
## Not Metro/Micro   :1455   Median  : 64.0   Chelan County : 120
## Small Metropolitan:1426   Mean    : 209.3   Clallam County: 120
##                  3rd Qu.: 195.0   Clark County  : 120
##                  Max.     :2574.0   Cowlitz County: 120
##                  (Other)      :3836
##
##      Population      hosp_norm      X4313314_Norm      X4313314
## Min.   :    76   Min.   :0.0007855   Min.   :0.0000   Min.   :    0.0
## 1st Qu.:  1113   1st Qu.:0.0149056   1st Qu.:0.4750   1st Qu.:  138.0
## Median :  3424   Median :0.0182642   Median :0.5983   Median :  583.5
## Mean   : 11298   Mean   :0.0175179   Mean   :0.5834   Mean   : 1902.4
## 3rd Qu.: 11169   3rd Qu.:0.0209590   3rd Qu.:0.7097   3rd Qu.: 1913.0
## Max.   :128161   Max.   :0.0416667   Max.   :0.9033   Max.   :26746.0
##
## TOTAL_POPULATION    t_index
## Min.   :    1.0   2010.1 : 39
## 1st Qu.:  284.8   2012.1 : 39
## Median :  882.5   2007.2 : 39
## Mean   : 3088.4   2008.3 : 39
## 3rd Qu.: 2998.2   2010.3 : 39
## Max.   :35370.0   2013.3 : 39
##                  (Other):4322
```

Plotting the data

```
ggplot(all.cause.merge, aes(X4313314_Norm,hosp_norm)) + geom_point(alpha=0.2) +
  geom_smooth() + theme_minimal(base_size = 8)
```



```
xyplot(hosp_norm ~ X4313314_Norm | factor(County), data=all.cause.merge,
       pch=19, type=c("p","r"), alpha = 0.2,
       main=list(label="All Cause Hospitalization Trends by County Category",cex=0.8),
       sub=list(label="11",cex=0.1),
       par.settings=simpleTheme(col="#59ABE3", col.line="red"),
       par.strip.text=list(cex=0.5))
```



```
# Fixed Effects All Cause Hospitalization
fe.all.cause <- plm(hosp_norm ~ X4313314_Norm, data = all.cause.merge,
                    index = c("FIPS", "t_index"), model = "within", effect = "individual")
```

```
## series AgeGroup is constant and has been removed
```

```
summary(fe.all.cause)
```

```
## Oneway (individual) effect Within Model
##
## Call:
## plm(formula = hosp_norm ~ X4313314_Norm, data = all.cause.merge,
##      effect = "individual", model = "within", index = c("FIPS",
##      "t_index"))
##
## Unbalanced Panel: n=39, T=42-120, N=4556
##
```

```
## Residuals :
##      Min.    1st Qu.      Median    3rd Qu.      Max.
## -0.016700 -0.002040 -0.000122  0.001790  0.028200
##
## Coefficients :
##              Estimate Std. Error t-value Pr(>|t|)
## X4313314_Norm -0.00453738  0.00048982 -9.2634 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Total Sum of Squares:    0.073823
## Residual Sum of Squares: 0.072446
## R-Squared:    0.018647
## Adj. R-Squared: 0.018483
## F-statistic: 85.8109 on 1 and 4516 DF, p-value: < 2.22e-16
```

```
summary(fixef(fe.all.cause))
```

```
##              Estimate Std. Error t-value Pr(>|t|)
## 53001 0.02755305 0.00053104 51.8851 < 2.2e-16 ***
## 53003 0.00386665 0.00043053  8.9811 < 2.2e-16 ***
## 53005 0.02248882 0.00048481 46.3872 < 2.2e-16 ***
## 53007 0.02391527 0.00052789 45.3039 < 2.2e-16 ***
## 53009 0.02231946 0.00046473 48.0269 < 2.2e-16 ***
## 53011 0.01832089 0.00047177 38.8346 < 2.2e-16 ***
## 53013 0.02102689 0.00047709 44.0735 < 2.2e-16 ***
## 53015 0.02224810 0.00048866 45.5292 < 2.2e-16 ***
## 53017 0.02211800 0.00053453 41.3782 < 2.2e-16 ***
## 53019 0.01890095 0.00044484 42.4896 < 2.2e-16 ***
## 53021 0.02318781 0.00049082 47.2432 < 2.2e-16 ***
## 53023 0.01740024 0.00072242 24.0861 < 2.2e-16 ***
## 53025 0.02421015 0.00052111 46.4590 < 2.2e-16 ***
## 53027 0.02300644 0.00048313 47.6198 < 2.2e-16 ***
## 53029 0.01319425 0.00043006 30.6801 < 2.2e-16 ***
## 53031 0.01824523 0.00045205 40.3614 < 2.2e-16 ***
## 53033 0.02248442 0.00048383 46.4720 < 2.2e-16 ***
## 53035 0.01713577 0.00046763 36.6439 < 2.2e-16 ***
## 53037 0.02053019 0.00043191 47.5330 < 2.2e-16 ***
## 53039 0.00897091 0.00043383 20.6784 < 2.2e-16 ***
## 53041 0.02252287 0.00050134 44.9253 < 2.2e-16 ***
## 53043 0.02346053 0.00046070 50.9237 < 2.2e-16 ***
## 53045 0.02196008 0.00047979 45.7706 < 2.2e-16 ***
## 53047 0.02481185 0.00048891 50.7493 < 2.2e-16 ***
## 53049 0.01316703 0.00046623 28.2415 < 2.2e-16 ***
## 53051 0.01969226 0.00040167 49.0254 < 2.2e-16 ***
## 53053 0.01918815 0.00044988 42.6517 < 2.2e-16 ***
## 53055 0.01891713 0.00044289 42.7132 < 2.2e-16 ***
## 53057 0.02212873 0.00042202 52.4350 < 2.2e-16 ***
## 53059 0.00972746 0.00045543 21.3587 < 2.2e-16 ***
## 53061 0.02139546 0.00047483 45.0589 < 2.2e-16 ***
## 53063 0.02257590 0.00043868 51.4630 < 2.2e-16 ***
## 53065 0.02153604 0.00042775 50.3469 < 2.2e-16 ***
## 53067 0.01945070 0.00045074 43.1532 < 2.2e-16 ***
## 53069 0.02014257 0.00049253 40.8965 < 2.2e-16 ***
```

```
## 53071 0.02235203 0.00044646 50.0648 < 2.2e-16 ***
## 53073 0.02163777 0.00046829 46.2059 < 2.2e-16 ***
## 53075 0.02101266 0.00043296 48.5327 < 2.2e-16 ***
## 53077 0.02488112 0.00048931 50.8493 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
fixef.all.cause <- matrix(fixef(fe.all.cause))
fixef.df <- as.data.frame(cbind(sort(unique(all.cause.merge$FIPS)), fixef.all.cause))
colnames(fixef.df) <- c("FIPS", "coeff")
fixef.df$FIPS <- as.character(fixef.df$FIPS)
fixef.df <- mutate(fixef.df, id=substr(FIPS, 3,5))
fixef.df <- fixef.df[c("coeff", 'id')]
```

Setting up Mapping Libraries

```
library(GGally)
library(rgdal)
library(ggmap)
library(rgeos)
library(broom)
library(Cairo)
library(maptools)
library(maps)
library(choroplethr)
```

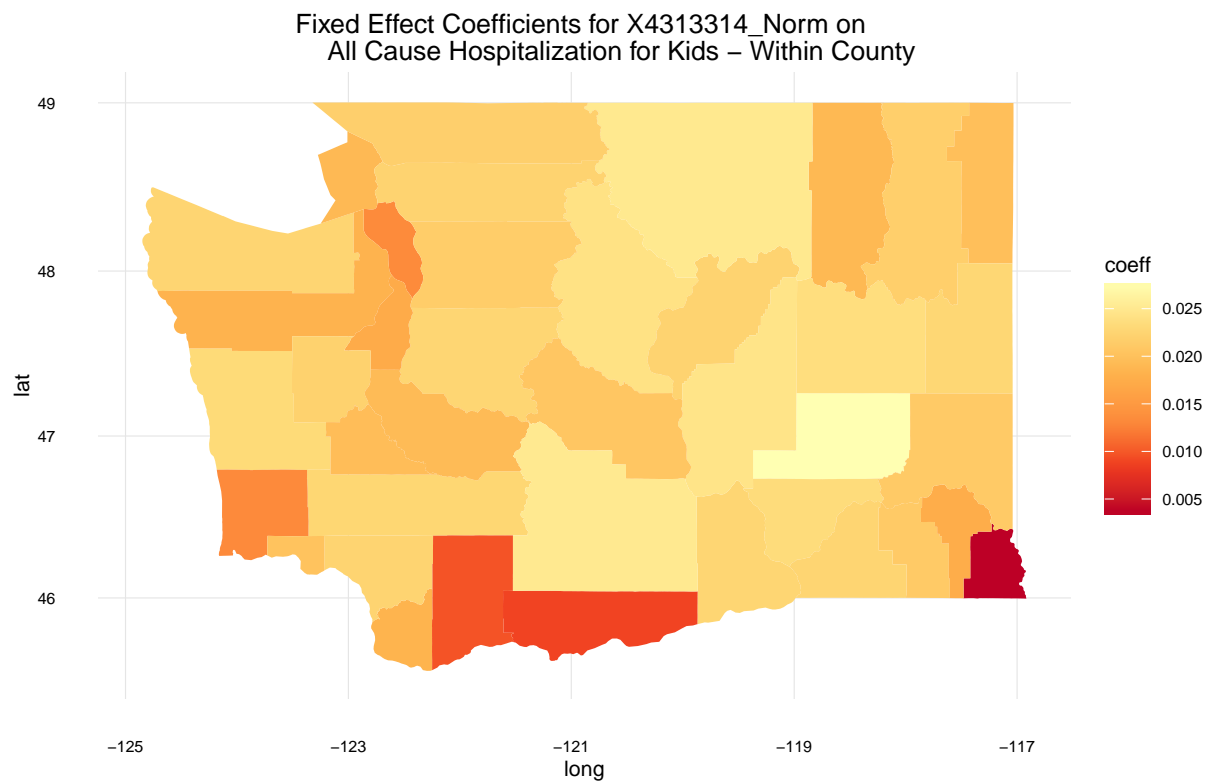
Loading Shapefiles and boundaries

```
map.shapefile <- readOGR(dsn = "/Users/ajaswal/Desktop/hcup-population/Shapefiles/tl_2013_53_cousub",
                        layer = "tl_2013_53_cousub")
```

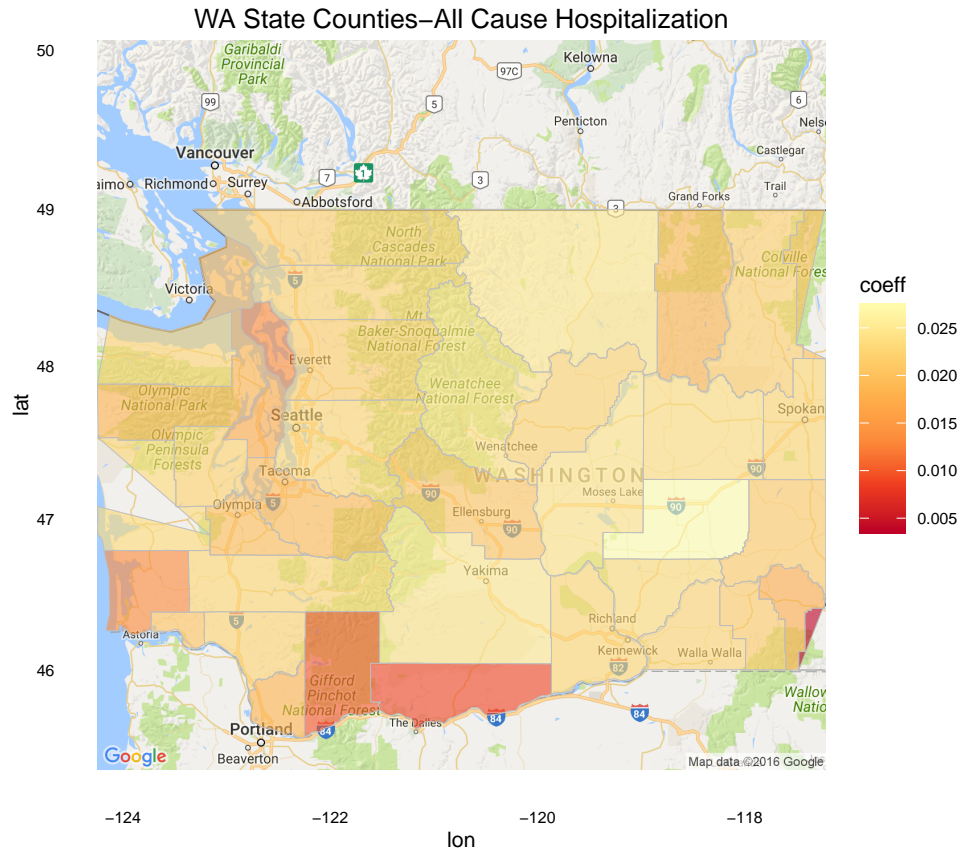
```
## OGR data source with driver: ESRI Shapefile
## Source: "/Users/ajaswal/Desktop/hcup-population/Shapefiles/tl_2013_53_cousub", layer: "tl_2013_53_cousub"
## with 242 features
## It has 18 fields
```

```
map.df <- fortify(map.shapefile, region="COUNTYFP")
```

```
map.df.merge <- merge(map.df, fixef.df, by="id")
ggplot() +
  geom_polygon(data = map.df.merge, aes(x = long, y = lat, group = group,
    fill = coeff), color = NA) +
  theme_minimal(base_size = 8) + coord_map() +
  scale_fill_distiller(palette = "YlOrRd") +
  ggtitle("Fixed Effect Coefficients for X4313314_Norm on
    All Cause Hospitalization for Kids - Within County")
```



```
# Trying Google Maps Overlay
map <- get_map("WA State", zoom = 7, maptype = "roadmap")
p <- ggmap(map)
p + geom_polygon(data = map.df.merge, aes(x = long, y = lat, group = group,
  fill = coeff), color = "grey", size=0.1, alpha=0.6) +
  scale_fill_distiller(palette = "YlOrRd") +
  ggtitle("WA State Counties-All Cause Hospitalization") +
  theme_minimal(base_size = 8)
```

Now Running Fixed Effects for Specific Type of Diagnosis

Single-Level CCS - Diagnoses: Codes

Diagnoses related to Respiratory Diseases

122 Pneumonia (except that caused by tuberculosis or sexually transmitted disease) 123 Influenza 124 Acute and chronic tonsillitis 125 Acute bronchitis 126 Other upper respiratory infections 127 Chronic obstructive pulmonary disease and bronchiectasis 128 Asthma 129 Aspiration pneumonitis; food/vomitus 130 Pleurisy; pneumothorax; pulmonary collapse 131 Respiratory failure; insufficiency; arrest (adult) 132 Lung disease due to external agents 133 Other lower respiratory disease 134 Other upper respiratory disease

```
# Filtering only Diagnoses related to Respiratory Diseases
hosp.resp <- hosp.filtered %>%
  select(ayear, amonth, zip, pstco2, AgeGroup, urban_rural, hospitalizations, dxccs1) %>%
  filter(dxccs1 >= 122 & dxccs1 <= 134)

# Aggregating
hosp.resp <- hosp.resp[c("ayear", "amonth", "pstco2", "hospitalizations", "AgeGroup")]
hosp.resp <- aggregate(hospitalizations~ayear+amonth+pstco2+AgeGroup, data=hosp.resp, sum)

# Joining population for normalization
hosp.resp <- merge(hosp.resp, census.clean, by.x=c("ayear", "amonth", "pstco2", "AgeGroup"),
  by.y=c("year", "month", "FIPS", "AgeGroup"))
hosp.resp <- mutate(hosp.resp, resp_hosp_norm = hospitalizations/Population)
```

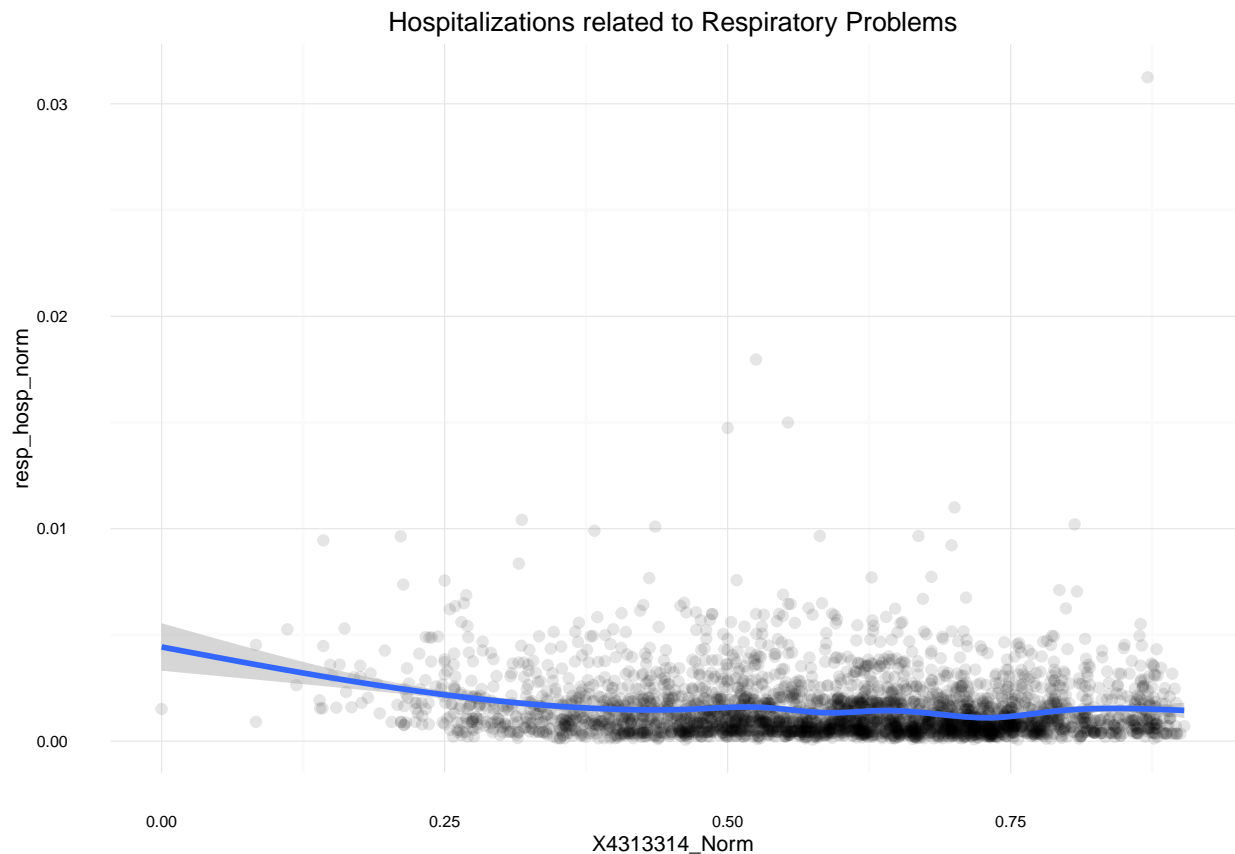
```

names(hosp.resp)[1] <- 'year'
names(hosp.resp)[2] <- 'month'
names(hosp.resp)[3] <- 'FIPS'

# Respiratory Hospitalizations
resp.merge <- merge(hosp.resp, vaccine.filtered, by=c("year","month","FIPS"))
resp.merge$t_index <- interaction(resp.merge$year, resp.merge$month, sep=".")

# Plotting the data
ggplot(resp.merge , aes(X4313314_Norm, resp_hosp_norm)) +
  geom_point(alpha=0.1) + geom_smooth() + theme_minimal(base_size = 8) +
  ggtitle("Hospitalizations related to Respiratory Problems")

```

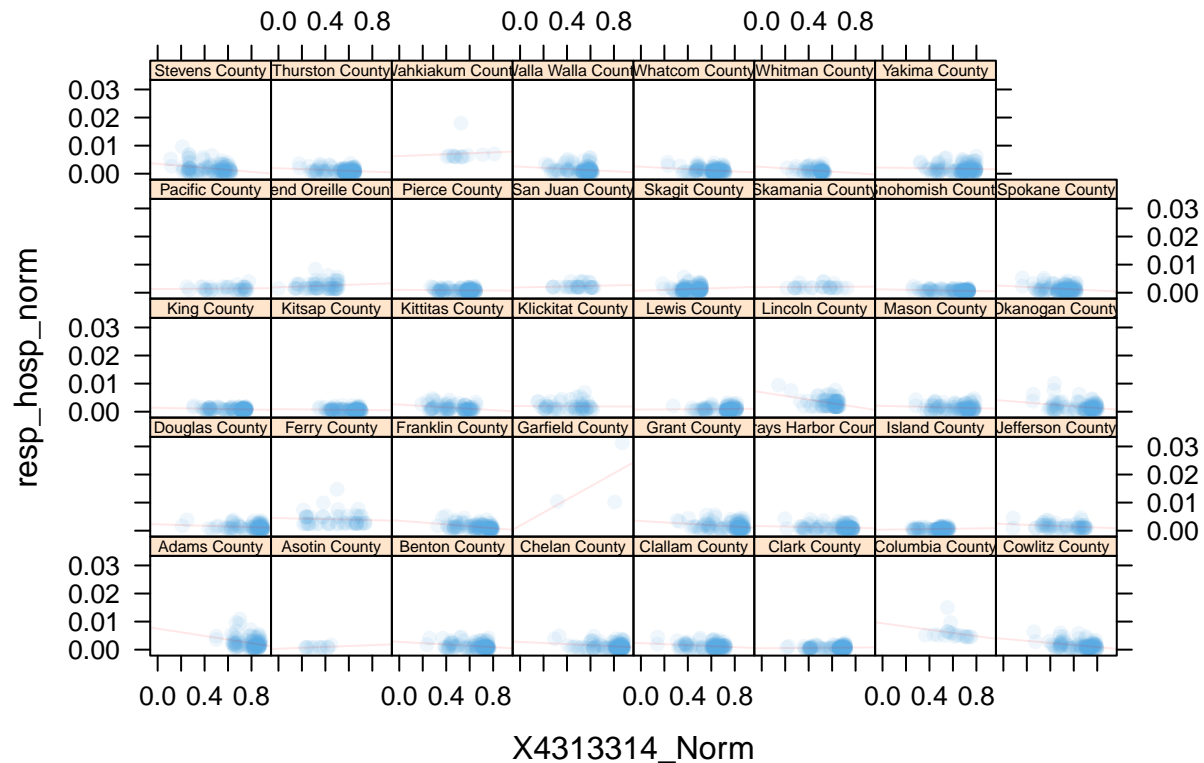


```

xyplot(resp_hosp_norm ~ X4313314_Norm | factor(County), data=resp.merge,
  pch=19, type=c("p","r"), alpha = 0.1,
  main=list(label="Hospitalizations related to Respiratory Problems By County",cex=0.8),
  sub=list(label="11",cex=0.1),
  par.settings=simpleTheme(col="#59ABE3", col.line="red"),
  par.strip.text=list(cex=0.5))

```

Hospitalizations related to Respiratory Problems By County



```
# Fixed Effects Respiratory related Hospitalization
```

```
fe.resp <- plm(resp_hosp_norm ~ X4313314_Norm, data = resp.merge,
               index = c("FIPS", "t_index"), model = "within", effect = "individual")
```

```
## series AgeGroup is constant and has been removed
```

```
summary(fe.resp)
```

```
## Oneway (individual) effect Within Model
```

```
##
```

```
## Call:
```

```
## plm(formula = resp_hosp_norm ~ X4313314_Norm, data = resp.merge,
##      effect = "individual", model = "within", index = c("FIPS",
##      "t_index"))
```

```
##
```

```
## Unbalanced Panel: n=39, T=3-120, N=3435
```

```
##
```

```
## Residuals :
```

```
##      Min.    1st Qu.      Median    3rd Qu.      Max.
## -0.007390 -0.000692 -0.000242  0.000400  0.014300
```

```
##
```

```
## Coefficients :
```

```
##              Estimate Std. Error t-value Pr(>|t|)
## X4313314_Norm -0.00149682  0.00016132 -9.2786 < 2.2e-16 ***
```

```
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Total Sum of Squares:    0.0045778
## Residual Sum of Squares: 0.0044646
## R-Squared:              0.024732
## Adj. R-Squared: 0.024444
## F-statistic: 86.0928 on 1 and 3395 DF, p-value: < 2.22e-16
```

```
summary(fixef(fe.resp))
```

```
##           Estimate Std. Error t-value Pr(>|t|)
## 53001 0.00358212 0.00017031 21.0325 < 2.2e-16 ***
## 53003 0.00137844 0.00034995  3.9390 8.183e-05 ***
## 53005 0.00221893 0.00014838 14.9539 < 2.2e-16 ***
## 53007 0.00239587 0.00016432 14.5804 < 2.2e-16 ***
## 53009 0.00215156 0.00014530 14.8082 < 2.2e-16 ***
## 53011 0.00160516 0.00014406 11.1425 < 2.2e-16 ***
## 53013 0.00684908 0.00028047 24.4201 < 2.2e-16 ***
## 53015 0.00243729 0.00015132 16.1065 < 2.2e-16 ***
## 53017 0.00244547 0.00017058 14.3362 < 2.2e-16 ***
## 53019 0.00476445 0.00018490 25.7672 < 2.2e-16 ***
## 53021 0.00232691 0.00015049 15.4621 < 2.2e-16 ***
## 53023 0.01828593 0.00067072 27.2630 < 2.2e-16 ***
## 53025 0.00271368 0.00016139 16.8147 < 2.2e-16 ***
## 53027 0.00208581 0.00014973 13.9305 < 2.2e-16 ***
## 53029 0.00130575 0.00013499  9.6729 < 2.2e-16 ***
## 53031 0.00235121 0.00016883 13.9268 < 2.2e-16 ***
## 53033 0.00185495 0.00014782 12.5490 < 2.2e-16 ***
## 53035 0.00153943 0.00014232 10.8168 < 2.2e-16 ***
## 53037 0.00207127 0.00014575 14.2108 < 2.2e-16 ***
## 53039 0.00255868 0.00016650 15.3676 < 2.2e-16 ***
## 53041 0.00199950 0.00015932 12.5501 < 2.2e-16 ***
## 53043 0.00403663 0.00016088 25.0912 < 2.2e-16 ***
## 53045 0.00237110 0.00015291 15.5066 < 2.2e-16 ***
## 53047 0.00272292 0.00015519 17.5463 < 2.2e-16 ***
## 53049 0.00238381 0.00020963 11.3714 < 2.2e-16 ***
## 53051 0.00300132 0.00016151 18.5824 < 2.2e-16 ***
## 53053 0.00166899 0.00013569 12.3001 < 2.2e-16 ***
## 53055 0.00306123 0.00022077 13.8663 < 2.2e-16 ***
## 53057 0.00196344 0.00012561 15.6318 < 2.2e-16 ***
## 53059 0.00279610 0.00024230 11.5399 < 2.2e-16 ***
## 53061 0.00166843 0.00014462 11.5368 < 2.2e-16 ***
## 53063 0.00215261 0.00013165 16.3507 < 2.2e-16 ***
## 53065 0.00257968 0.00013318 19.3700 < 2.2e-16 ***
## 53067 0.00192159 0.00013600 14.1297 < 2.2e-16 ***
## 53069 0.00796069 0.00032951 24.1589 < 2.2e-16 ***
## 53071 0.00220345 0.00013735 16.0431 < 2.2e-16 ***
## 53073 0.00211463 0.00014259 14.8296 < 2.2e-16 ***
## 53075 0.00183058 0.00015314 11.9533 < 2.2e-16 ***
## 53077 0.00276942 0.00014976 18.4923 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```

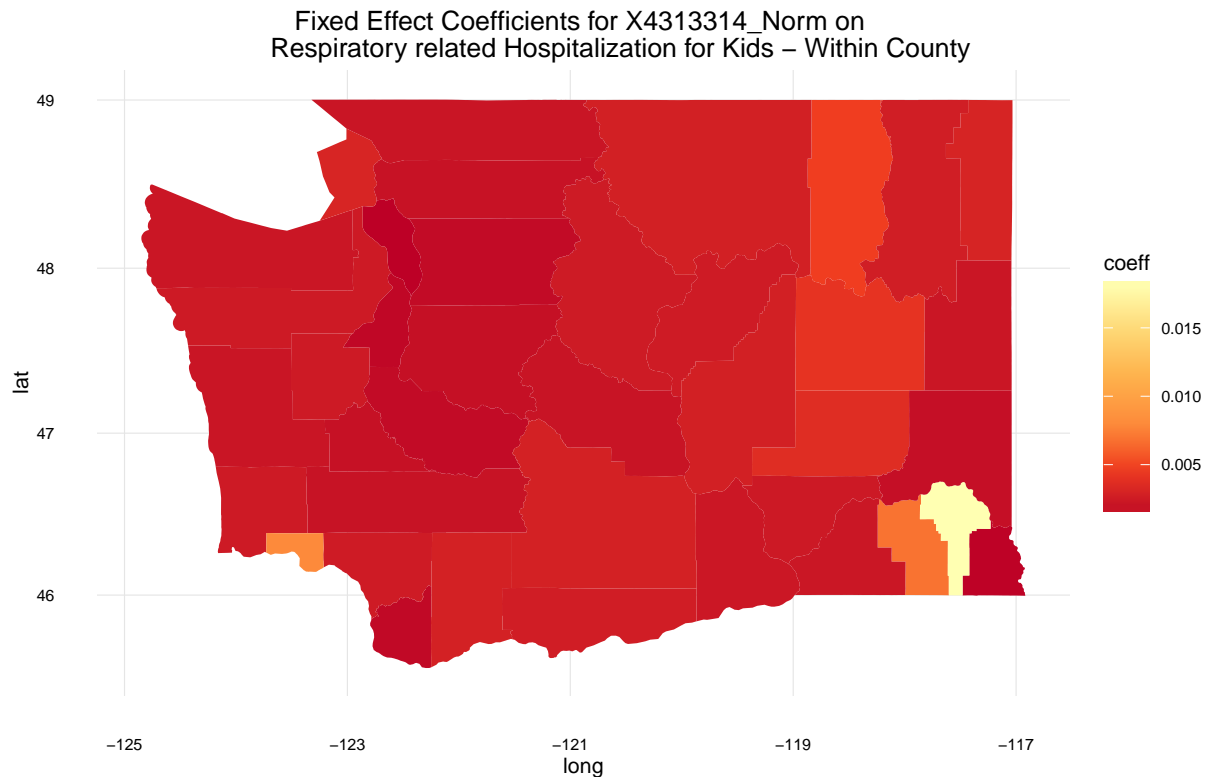
fixef.resp <- matrix(fixef(fe.resp))
fixef.resp.df <- as.data.frame(cbind(sort(unique(resp.merge$FIPS)), fixef.resp))
colnames(fixef.resp.df) <- c("FIPS", "coeff")
fixef.resp.df$FIPS <- as.character(fixef.resp.df$FIPS)
fixef.resp.df <- mutate(fixef.resp.df, id=substr(FIPS, 3,5))
fixef.resp.df <- fixef.resp.df[c("coeff", 'id')]

```

```

# Choropleth for Respiratory related diagnoses
map.df.merge <- merge(map.df, fixef.resp.df, by="id")
ggplot() +
  geom_polygon(data = map.df.merge, aes(x = long, y = lat, group = group,
    fill = coeff), color = NA) +
  theme_minimal(base_size = 8) + coord_map() +
  scale_fill_distiller(palette = "YlOrRd") +
  ggtitle("Fixed Effect Coefficients for X4313314_Norm on
    Respiratory related Hospitalization for Kids - Within County")

```



Diagnoses not related to Vaccinations

226 Fracture of neck of femur (hip) 227 Spinal cord injury 228 Skull and face fractures 229 Fracture of upper limb 230 Fracture of lower limb 231 Other fractures 232 Sprains and strains 233 Intracranial injury 234 Crushing injury or internal injury 235 Open wounds of head; neck; and trunk 236 Open wounds of extremities 237 Complication of device; implant or graft 238 Complications of surgical procedures or medical care 239 Superficial injury; contusion 240 Burns 241 Poisoning by psychotropic agents 242 Poisoning by other

medications and drugs 243 Poisoning by nonmedicinal substances 244 Other injuries and conditions due to external causes

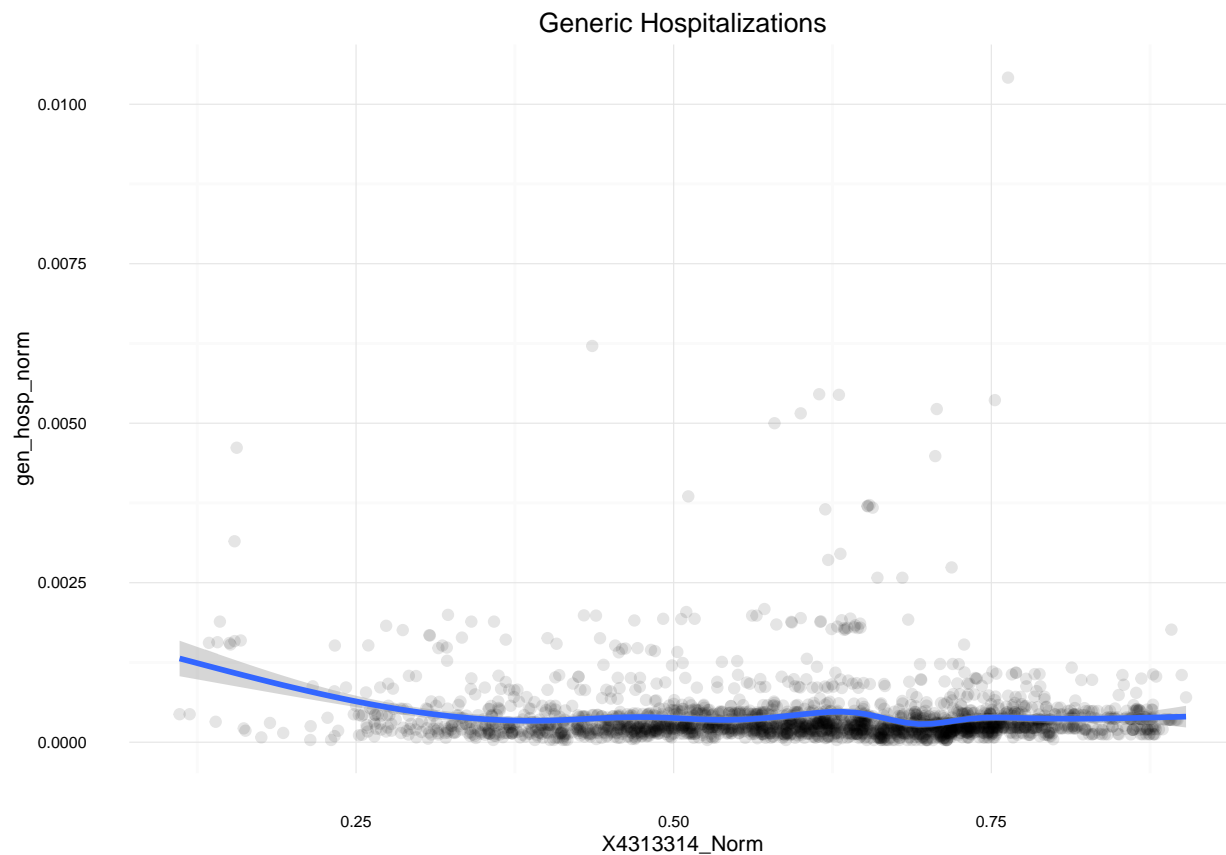
```
# Filtering only Diagnoses related to Non-Respiratory Generic Diseases
hosp.general <- hosp.filtered %>%
  select(ayear, amonth, zip, pstco2, AgeGroup, urban_rural, hospitalizations, dxccs1) %>%
  filter(dxccs1 >= 226 & dxccs1 <= 244)

# Aggregating
hosp.general <- hosp.general[c("ayear", "amonth", "pstco2", "hospitalizations", "AgeGroup")]
hosp.general <- aggregate(hospitalizations~ayear+amonth+pstco2+AgeGroup,
  data=hosp.general, sum)

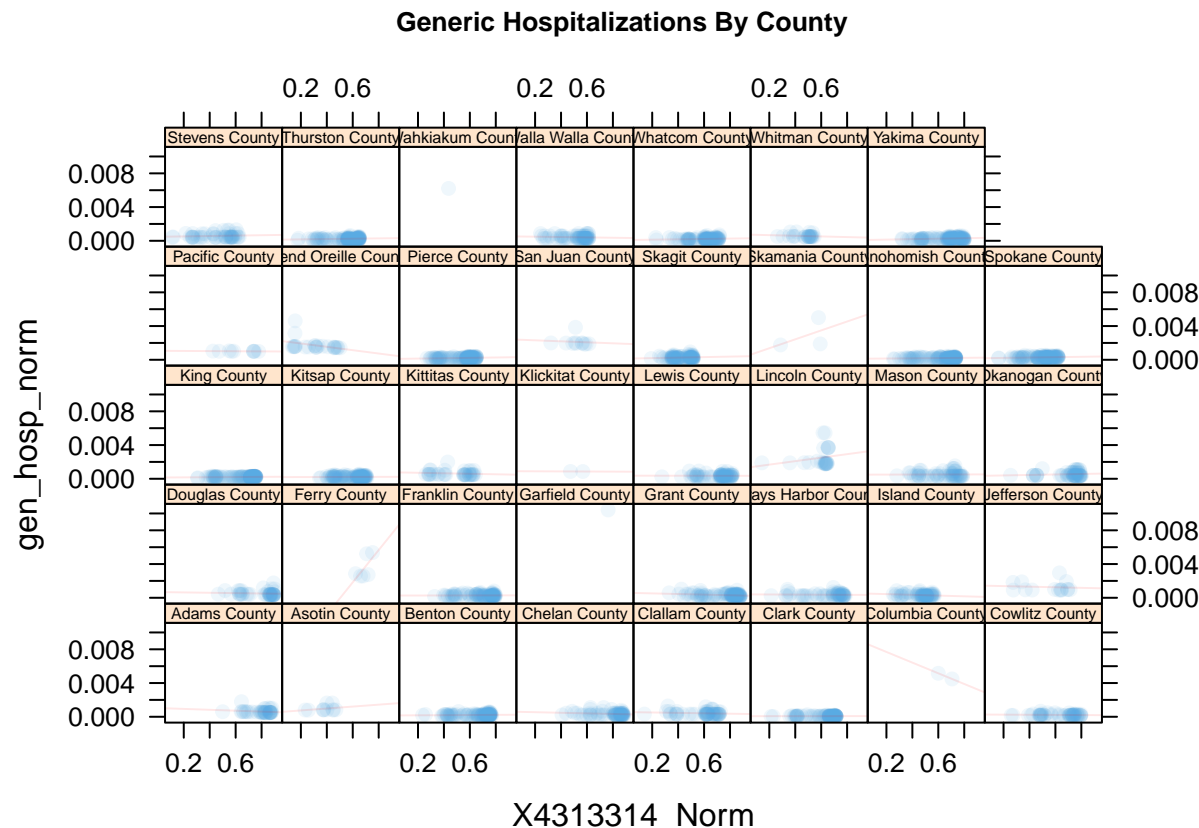
# Joining population for normalization
hosp.general <- merge(hosp.general, census.clean,
  by.x=c("ayear", "amonth", "pstco2", "AgeGroup"),
  by.y=c("year", "month", "FIPS", "AgeGroup"))
hosp.general <- mutate(hosp.general, gen_hosp_norm = hospitalizations/Population)
names(hosp.general)[1] <- 'year'
names(hosp.general)[2] <- 'month'
names(hosp.general)[3] <- 'FIPS'

# Respiratory Hospitalizations
gen.merge <- merge(hosp.general, vaccine.filtered, by=c("year", "month", "FIPS"))
gen.merge$t_index <- interaction(gen.merge$year, gen.merge$month, sep=".")

# Plotting the data
ggplot(gen.merge, aes(X4313314_Norm, gen_hosp_norm)) +
  geom_point(alpha=0.1) + geom_smooth() + theme_minimal(base_size = 8) +
  ggtitle("Generic Hospitalizations")
```



```
xyplot(gen_hosp_norm ~ X4313314_Norm | factor(County), data=gen.merge,
  pch=19, type=c("p","r"), alpha = 0.1,
  main=list(label="Generic Hospitalizations By County",cex=0.8),
  sub=list(label="11",cex=0.1),
  par.settings=simpleTheme(col="#59ABE3", col.line="red"),
  par.strip.text=list(cex=0.5))
```



```
# Fixed Effects - Generic Hospitalizations
fe.gen <- plm(gen_hosp_norm ~ X4313314_Norm, data = gen.merge,
              index = c("FIPS", "t_index"), model = "within", effect = "individual")
```

```
## series AgeGroup is constant and has been removed
```

```
summary(fe.gen)
```

```
## Oneway (individual) effect Within Model
##
## Call:
## plm(formula = gen_hosp_norm ~ X4313314_Norm, data = gen.merge,
##      effect = "individual", model = "within", index = c("FIPS",
##      "t_index"))
##
## Unbalanced Panel: n=39, T=1-120, N=2328
##
## Residuals :
##      Min.   1st Qu.   Median   3rd Qu.   Max.
## -1.12e-03 -1.22e-04 -4.07e-05  6.17e-05  2.96e-03
##
## Coefficients :
##              Estimate Std. Error t-value Pr(>|t|)
## X4313314_Norm 2.2289e-05 4.1563e-05  0.5363   0.5918
##
```



```
## Total Sum of Squares:    0.00012804
## Residual Sum of Squares: 0.00012802
## R-Squared:              0.00012568
## Adj. R-Squared: 0.00012352
## F-statistic: 0.287583 on 1 and 2288 DF, p-value: 0.59183
```

```
summary(fixef(fe.gen))
```

```
##           Estimate Std. Error t-value Pr(>|t|)
## 53001 6.0063e-04 4.9238e-05 12.1985 < 2.2e-16 ***
## 53003 9.5147e-04 7.3085e-05 13.0187 < 2.2e-16 ***
## 53005 1.9807e-04 3.5188e-05  5.6287 1.815e-08 ***
## 53007 3.2246e-04 4.2769e-05  7.5397 4.707e-14 ***
## 53009 4.0253e-04 3.9674e-05 10.1460 < 2.2e-16 ***
## 53011 6.0568e-05 3.5723e-05  1.6955  0.08998 .
## 53013 4.8049e-03 1.6945e-04 28.3559 < 2.2e-16 ***
## 53015 2.0140e-04 4.0795e-05  4.9368 7.942e-07 ***
## 53017 4.9142e-04 4.7746e-05 10.2925 < 2.2e-16 ***
## 53019 3.5405e-03 1.0074e-04 35.1454 < 2.2e-16 ***
## 53021 2.6439e-04 3.7070e-05  7.1323 9.872e-13 ***
## 53023 1.0400e-02 2.3866e-04 43.5746 < 2.2e-16 ***
## 53025 2.9797e-04 3.9302e-05  7.5815 3.419e-14 ***
## 53027 3.4594e-04 4.0623e-05  8.5158 < 2.2e-16 ***
## 53029 3.0851e-04 3.4217e-05  9.0161 < 2.2e-16 ***
## 53031 1.2511e-03 6.5214e-05 19.1852 < 2.2e-16 ***
## 53033 1.9567e-04 3.4485e-05  5.6740 1.395e-08 ***
## 53035 2.0523e-04 3.3065e-05  6.2068 5.406e-10 ***
## 53037 6.1473e-04 4.3333e-05 14.1859 < 2.2e-16 ***
## 53039 8.5420e-04 1.6866e-04  5.0647 4.090e-07 ***
## 53041 3.3814e-04 3.9436e-05  8.5745 < 2.2e-16 ***
## 53043 2.4759e-03 5.2508e-05 47.1528 < 2.2e-16 ***
## 53045 4.9425e-04 3.9857e-05 12.4005 < 2.2e-16 ***
## 53047 5.2605e-04 4.2230e-05 12.4568 < 2.2e-16 ***
## 53049 9.9674e-04 7.9309e-05 12.5678 < 2.2e-16 ***
## 53051 1.7016e-03 4.8184e-05 35.3140 < 2.2e-16 ***
## 53053 2.0318e-04 3.1000e-05  6.5543 5.589e-11 ***
## 53055 2.1090e-03 7.4510e-05 28.3050 < 2.2e-16 ***
## 53057 2.5679e-04 3.0209e-05  8.5002 < 2.2e-16 ***
## 53059 2.8706e-03 1.3806e-04 20.7927 < 2.2e-16 ***
## 53061 1.9452e-04 3.3573e-05  5.7939 6.876e-09 ***
## 53063 2.7615e-04 2.9822e-05  9.2599 < 2.2e-16 ***
## 53065 5.6959e-04 3.7544e-05 15.1711 < 2.2e-16 ***
## 53067 2.2153e-04 3.1825e-05  6.9608 3.384e-12 ***
## 53069 6.2015e-03 2.3724e-04 26.1402 < 2.2e-16 ***
## 53071 4.0862e-04 3.5240e-05 11.5955 < 2.2e-16 ***
## 53073 2.0538e-04 3.4255e-05  5.9955 2.029e-09 ***
## 53075 5.5591e-04 4.8054e-05 11.5684 < 2.2e-16 ***
## 53077 2.5331e-04 3.5038e-05  7.2297 4.841e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
fixef.gen <- matrix(fixef(fe.gen))
fixef.gen.df <- as.data.frame(cbind(sort(unique(gen.merge$FIPS)), fixef.gen))
```

```

colnames(fixef.gen.df) <- c("FIPS","coeff")
fixef.gen.df$FIPS <- as.character(fixef.gen.df$FIPS)
fixef.gen.df <- mutate(fixef.gen.df, id=substr(FIPS,3,5))
fixef.gen.df <- fixef.gen.df[c("coeff",'id')]

# Choropleth for Respiratory related diagnoses
map.df.merge <- merge(map.df,fixef.gen.df, by="id")
ggplot() +
  geom_polygon(data = map.df.merge, aes(x = long, y = lat, group = group,
    fill = coeff), color = NA) +
  theme_minimal(base_size = 8) + coord_map() +
  scale_fill_distiller(palette = "YlOrRd") +
  ggtitle("Fixed Effect Coefficients for X4313314_Norm on
    Non-Respiratory related Generic Hospitalization for Kids - Within County")

```

