Coastal Analysis

Read in data

```
library("readxl")
library("lme4")
```

```
## Loading required package: Matrix
```

```
# Read in dataset with coastal coding. Read in summary sheet (sheet 13)
coastal <- read_excel("FIPS-based datasets_05232021.xlsx", sheet = 13)
```

```
## New names:
## * `` -> ...12
## * `` -> ...22
## * `` -> ...25
## * `` -> ...39
```

```
# summary(coastal)

# Read in PM25 data from our 2020 study, created with: PM25 = data.frame(fips =
# aggregate_pm_census_cdc_test_beds$fips, mean_pm25 =
# aggregate_pm_census_cdc_test_beds$mean_pm25) save(PM25, file = 'PM25.Rda')
load("PM25.Rda")
```

Create smaller dataset from previous dataset, dataclean, merge with PM25 dataset.

```
coastal.new = data.frame(coastal$`FIPS as Text`, coastal$state, coastal$cases, coastal$deaths,
    coastal$`Country REGION`, coastal$`Coastal Distance`, coastal$`Population 2019 Estimate`,
    coastal$`Population Density`, coastal$`All Ages in Poverty (%)`, coastal$`Under 18s in Poverty`,
    coastal$`Median Income`, coastal$`percent adult obesity`, coastal$`diff/total`,
    coastal$`Politcal alignment 2020 election`, coastal$`median age 2019`, coastal$Humid)
colnames(coastal.new) = c("fips", "state", "cases", "deaths", "region", "coastal.distance",
    "population2019", "popdensity", "poverty", "under18poverty", "median_income",
    "pct_obesity", "voter_margin_2020", "party", "median_age", "humidity")

# change NAs in coastal.distance to level 4, and save as factor with reference
# level 4.
coastal.new$coastal.distance[is.na(coastal.new$coastal.distance)] <- 4
coastal.new$coastal.distance = as.factor(coastal.new$coastal.distance)
coastal.new <- within(coastal.new, coastal.distance <- relevel(coastal.distance,
    ref = 4))

# change NAs in 'region' to 'Inland', convert all characters to lowercase
coastal.new$region[is.na(coastal.new$region)] <- "Inland"
coastal.new$region = tolower(coastal.new$region)

# Merge with PM25 dataset
```

```r
coastal.new = merge(coastal.new, PM25, by = "fips")
summary(coastal.new)
```

```
##      fips               state                cases              deaths
##  Length:3088        Length:3088        Min.   :      1   Min.   :    0.0
##  Class :character   Class :character   1st Qu.:   1025   1st Qu.:   18.0
##  Mode  :character   Mode  :character   Median :   2456   Median :   47.0
##                                        Mean   :   9416   Mean   :  165.9
##                                        3rd Qu.:   6160   3rd Qu.:  110.0
##                                        Max.   :1219237   Max.   :23101.0
##     region          coastal.distance population2019       popdensity
##  Length:3088        4:2417           Min.   :     169   Min.   :    0.1
##  Class :character   1: 300           1st Qu.:   11137   1st Qu.:   17.5
##  Mode  :character   2: 200           Median :   26163   Median :   45.3
##                     3: 171           Mean   :  102696   Mean   :  202.6
##                                      3rd Qu.:   68022   3rd Qu.:  112.7
##                                      Max.   :10039107   Max.   :17179.1
##     poverty       under18poverty   median_income     pct_obesity
##  Min.   :0.0270   Min.   :0.0240   Min.   : 24732   Min.   :13.6
##  1st Qu.:0.1050   1st Qu.:0.1370   1st Qu.: 46212   1st Qu.:29.4
##  Median :0.1340   Median :0.1870   Median : 53242   Median :32.4
##  Mean   :0.1447   Mean   :0.1999   Mean   : 55573   Mean   :32.1
##  3rd Qu.:0.1750   3rd Qu.:0.2490   3rd Qu.: 61767   3rd Qu.:35.1
##  Max.   :0.4770   Max.   :0.6340   Max.   :151806   Max.   :49.5
##  voter_margin_2020    party             median_age     humidity
##  Min.   :-0.8675   Length:3088        Min.   :23.4   Length:3088
##  1st Qu.: 0.1375   Class :character   1st Qu.:38.2   Class :character
##  Median : 0.3859   Mode  :character   Median :41.4   Mode  :character
##  Mean   : 0.3203                      Mean   :41.5
##  3rd Qu.: 0.5666                      3rd Qu.:44.6
##  Max.   : 0.9309                      Max.   :67.4
##    mean_pm25
##  Min.   : 2.060
##  1st Qu.: 6.335
##  Median : 8.789
##  Mean   : 8.398
##  3rd Qu.:10.483
##  Max.   :15.786
```

PRELIMINARY ANALYSIS on only coastal counties

```
# Subset coastal counties only
coastal.only = coastal.new[coastal.new$coastal.distance != 4, ]
nrow(coastal.only)
```

```
## [1] 671
```

```
nrow(na.omit(coastal.only))
```

```
## [1] 633
```

```
# Model cases
model.initial.cases = glmer(cases ~ (1 | state) + coastal.distance + offset(log(population2019)) +
    scale(popdensity) + scale(poverty) + scale(log(median_income)) + scale(pct_obesity) +
    scale(voter_margin_2020) + scale(median_age) + factor(party) + factor(humidity) +
    mean_pm25, family = poisson(link = "log"), data = coastal.only)
```

```
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, : Model is nearly unide
##  - Rescale variables?
```

```
summary(model.initial.cases)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: poisson  ( log )
## Formula:
## cases ~ (1 | state) + coastal.distance + offset(log(population2019)) +
##     scale(popdensity) + scale(poverty) + scale(log(median_income)) +
##     scale(pct_obesity) + scale(voter_margin_2020) + scale(median_age) +
##     factor(party) + factor(humidity) + mean_pm25
##    Data: coastal.only
##
##       AIC       BIC    logLik  deviance  df.resid
##  489510.4  489572.7 -244741.2  489482.4       619
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -137.952  -12.341   -1.138    9.761  302.294
##
## Random effects:
##  Groups Name        Variance Std.Dev.
##  state  (Intercept) 0.04082  0.202
## Number of obs: 633, groups:  state, 29
##
## Fixed effects:
##                              Estimate Std. Error  z value Pr(>|z|)
## (Intercept)                -2.8945317  0.0904450  -32.003  < 2e-16 ***
## coastal.distance2          -0.0083465  0.0007805  -10.694  < 2e-16 ***
## coastal.distance3           0.0032708  0.0010161    3.219  0.00129 **
## scale(popdensity)          -0.0157230  0.0002362  -66.553  < 2e-16 ***
## scale(poverty)              0.0453622  0.0009874   45.939  < 2e-16 ***
## scale(log(median_income)) -0.0333713  0.0009787  -34.098  < 2e-16 ***
## scale(pct_obesity)         -0.0667857  0.0005562 -120.075  < 2e-16 ***
## scale(voter_margin_2020)    0.0539723  0.0006596   81.828  < 2e-16 ***
## scale(median_age)          -0.1103006  0.0005729 -192.536  < 2e-16 ***
```

```
## factor(party)Republican   -0.0055143  0.0010650   -5.178 2.25e-07 ***
## factor(humidity)Marine    -0.3254818  0.0018670 -174.337  < 2e-16 ***
## factor(humidity)Moist      0.1085693  0.0993722    1.093  0.27459
## mean_pm25                  0.0326665  0.0002082  156.909  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Correlation matrix not shown by default, as p = 13 > 12.
## Use print(x, correlation=TRUE)  or
##     vcov(x)        if you need it
```

```
## optimizer (Nelder_Mead) convergence code: 0 (OK)
## Model is nearly unidentifiable: very large eigenvalue
##  - Rescale variables?
```

```
# Model deaths
model.initial.deaths = glmer(deaths ~ (1 | state) + coastal.distance + offset(log(population2019)) +
    scale(popdensity) + scale(poverty) + scale(log(median_income)) + scale(pct_obesity) +
    scale(voter_margin_2020) + scale(median_age) + factor(party) + factor(humidity) +
    mean_pm25, family = poisson(link = "log"), data = coastal.only)
```

```
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, : Model is nearly unide
##  - Rescale variables?
```

```
summary(model.initial.deaths)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: poisson  ( log )
## Formula:
## deaths ~ (1 | state) + coastal.distance + offset(log(population2019)) +
##     scale(popdensity) + scale(poverty) + scale(log(median_income)) +
##     scale(pct_obesity) + scale(voter_margin_2020) + scale(median_age) +
##     factor(party) + factor(humidity) + mean_pm25
##    Data: coastal.only
##
##      AIC      BIC   logLik deviance df.resid
##  19261.6  19323.9  -9616.8  19233.6      619
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -21.5020  -2.6453  -0.4458   2.1176  25.4407
##
## Random effects:
##  Groups Name        Variance Std.Dev.
##  state  (Intercept) 0.09803  0.3131
## Number of obs: 633, groups:  state, 29
##
## Fixed effects:
##                            Estimate Std. Error z value Pr(>|z|)
## (Intercept)               -7.497283   0.146352 -51.228  < 2e-16 ***
## coastal.distance2          0.042181   0.005735   7.355 1.90e-13 ***
## coastal.distance3          0.023638   0.007407   3.191  0.00142 **
## scale(popdensity)         -0.019259   0.001658 -11.616  < 2e-16 ***
## scale(poverty)             0.217943   0.007131  30.563  < 2e-16 ***
```

```
## scale(log(median_income)) -0.021899   0.007304  -2.998  0.00271 **
## scale(pct_obesity)        -0.030917   0.004250  -7.275 3.46e-13 ***
## scale(voter_margin_2020)   0.023009   0.004848   4.746 2.08e-06 ***
## scale(median_age)          0.129801   0.004205  30.868  < 2e-16 ***
## factor(party)Republican    0.018457   0.007942   2.324  0.02012 *
## factor(humidity)Marine    -0.223662   0.015466 -14.462  < 2e-16 ***
## factor(humidity)Moist      0.280735   0.158825   1.768  0.07713 .
## mean_pm25                  0.085450   0.001641  52.064  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation matrix not shown by default, as p = 13 > 12.
## Use print(x, correlation=TRUE)  or
##     vcov(x)        if you need it
## optimizer (Nelder_Mead) convergence code: 0 (OK)
## Model is nearly unidentifiable: very large eigenvalue
##  - Rescale variables?
```

Redo: Create indicator for being a coast (levels 1,2,3) instead.

```
# Indicator Coastal or NonCoastal
coastal.new$indicatorcoast = ifelse(coastal.new$coastal.distance == "4", "Noncoastal",
    "Coastal")

# Model cases
model.indicator.cases = glmer(cases ~ (1 | state) + factor(indicatorcoast) + offset(log(population2019))
    scale(popdensity) + scale(poverty) + scale(log(median_income)) + scale(pct_obesity) +
    scale(voter_margin_2020) + scale(median_age) + factor(party) + factor(humidity) +
    mean_pm25, family = poisson(link = "log"), data = coastal.new)
```

```
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, : Model is nearly unide
##  - Rescale variables?
```

```
summary(model.indicator.cases)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: poisson  ( log )
## Formula:
## cases ~ (1 | state) + factor(indicatorcoast) + offset(log(population2019)) +
##     scale(popdensity) + scale(poverty) + scale(log(median_income)) +
##     scale(pct_obesity) + scale(voter_margin_2020) + scale(median_age) +
##     factor(party) + factor(humidity) + mean_pm25
##    Data: coastal.new
##
##       AIC       BIC    logLik  deviance  df.resid
##  965452.9  965530.9 -482713.5  965426.9      2973
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -147.77   -6.66   -0.44    6.57  351.65
##
## Random effects:
##  Groups Name        Variance Std.Dev.
##  state  (Intercept) 0.07194  0.2682
## Number of obs: 2986, groups:  state, 49
##
## Fixed effects:
##                                   Estimate Std. Error  z value Pr(>|z|)
## (Intercept)                     -2.4439948  0.0383588  -63.714  < 2e-16 ***
## factor(indicatorcoast)Noncoastal  0.0017808  0.0006072    2.933  0.00336 **
## scale(popdensity)               -0.0014500  0.0001089  -13.310  < 2e-16 ***
## scale(poverty)                   0.0156984  0.0006444   24.359  < 2e-16 ***
## scale(log(median_income))       -0.0332435  0.0005610  -59.253  < 2e-16 ***
## scale(pct_obesity)              -0.0133383  0.0003348  -39.842  < 2e-16 ***
## scale(voter_margin_2020)         0.0830422  0.0004115  201.787  < 2e-16 ***
## scale(median_age)               -0.0824599  0.0003313 -248.867  < 2e-16 ***
## factor(party)Republican         -0.0222363  0.0007204  -30.866  < 2e-16 ***
## factor(humidity)Marine          -0.2696846  0.0015653 -172.285  < 2e-16 ***
## factor(humidity)Moist           -0.3756649  0.0016911 -222.147  < 2e-16 ***
## mean_pm25                        0.0428397  0.0001566  273.594  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Correlation of Fixed Effects:
##               (Intr) fct()N scl(pp) scl(pv) s((_)) scl(p_) s(__20 scl(m_)
## fctr(ndct)N   -0.017
## scl(ppdnst)    0.011  0.045
## scal(pvrty)    0.000  0.103 -0.158
## scl(lg(m_))    0.002  0.100 -0.081   0.869
## scl(pct_bs)    0.003  0.030  0.140   0.081   0.344
## scl(__2020)    0.007  0.021  0.252   0.152   0.118 -0.153
## scal(mdn_g)    0.004  0.063 -0.042   0.306   0.266  0.157  -0.173
## fctr(prty)R   -0.018 -0.051 -0.110   0.001  -0.024 -0.077  -0.662 -0.082
## fctr(hmdty)Mr -0.014  0.077 -0.101  -0.094  -0.194 -0.079   0.168 -0.112
## fctr(hmdty)Ms -0.030  0.117  0.007   0.036  -0.002 -0.021   0.014 -0.082
## mean_pm25     -0.031  0.083 -0.282  -0.026  -0.086 -0.035   0.155  0.062
##               fct()R fctr(hmdty)Mr fctr(hmdty)Ms
## fctr(ndct)N
## scl(ppdnst)
## scal(pvrty)
## scl(lg(m_))
## scl(pct_bs)
## scl(__2020)
## scal(mdn_g)
## fctr(prty)R
## fctr(hmdty)Mr -0.022
## fctr(hmdty)Ms  0.010 -0.015
## mean_pm25      0.055  0.395        -0.142
## optimizer (Nelder_Mead) convergence code: 0 (OK)
## Model is nearly unidentifiable: very large eigenvalue
##  - Rescale variables?
```

```
# Model deaths
model.indicator.deaths = glmer(deaths ~ (1 | state) + factor(indicatorcoast) + offset(log(population201
    scale(popdensity) + scale(poverty) + scale(log(median_income)) + scale(pct_obesity) +
    scale(voter_margin_2020) + scale(median_age) + factor(party) + factor(humidity) +
    mean_pm25, family = poisson(link = "log"), data = coastal.new)
```

```
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, : Model is nearly unide
##  - Rescale variables?
```

```
summary(model.indicator.deaths)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: poisson  ( log )
## Formula:
## deaths ~ (1 | state) + factor(indicatorcoast) + offset(log(population2019)) +
##     scale(popdensity) + scale(poverty) + scale(log(median_income)) +
##     scale(pct_obesity) + scale(voter_margin_2020) + scale(median_age) +
##     factor(party) + factor(humidity) + mean_pm25
##    Data: coastal.new
##
##      AIC      BIC   logLik deviance df.resid
##  56181.9  56259.9 -28077.9  56155.9     2973
##
## Scaled residuals:
```

```
##      Min       1Q   Median       3Q      Max
## -25.4192  -1.8723  -0.1705   1.8628  28.5901
##
## Random effects:
##  Groups Name        Variance Std.Dev.
##  state  (Intercept) 0.2004   0.4476
## Number of obs: 2986, groups:  state, 49
##
## Fixed effects:
##                                    Estimate Std. Error z value Pr(>|z|)
## (Intercept)                       -6.4523005  0.0655898 -98.374  < 2e-16 ***
## factor(indicatorcoast)Noncoastal -0.0390304  0.0045180  -8.639  < 2e-16 ***
## scale(popdensity)                  0.0003630  0.0007515   0.483   0.6290
## scale(poverty)                     0.1428396  0.0046432  30.763  < 2e-16 ***
## scale(log(median_income))         -0.0837036  0.0041728 -20.059  < 2e-16 ***
## scale(pct_obesity)                 0.0062329  0.0025400   2.454   0.0141 *
## scale(voter_margin_2020)           0.0926073  0.0030716  30.150  < 2e-16 ***
## scale(median_age)                  0.1103579  0.0024355  45.312  < 2e-16 ***
## factor(party)Republican           -0.0290070  0.0054692  -5.304 1.13e-07 ***
## factor(humidity)Marine            -0.1816046  0.0128505 -14.132  < 2e-16 ***
## factor(humidity)Moist             -0.6519447  0.0119238 -54.676  < 2e-16 ***
## mean_pm25                          0.0792389  0.0011936  66.386  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##               (Intr) fct()N scl(pp) scl(pv) s((_)) scl(p_) s(__20 scl(m_)
## fctr(ndct)N   -0.073
## scl(ppdnst)    0.050  0.044
## scal(pvrty)   -0.003  0.105 -0.171
## scl(lg(m_))    0.009  0.109 -0.069   0.863
## scl(pct_bs)    0.016  0.026  0.183   0.055   0.336
## scl(__2020)    0.034 -0.001  0.264   0.149   0.106 -0.136
## scal(mdn_g)    0.009  0.076  0.005   0.294   0.276  0.181  -0.164
## fctr(prty)R   -0.080 -0.046 -0.141   0.009  -0.010 -0.084  -0.663 -0.105
## fctr(hmdty)Mr -0.058  0.057 -0.090  -0.084  -0.182 -0.075   0.155 -0.100
## fctr(hmdty)Ms -0.123  0.135 -0.004   0.031  -0.011 -0.035  -0.006 -0.082
## mean_pm25     -0.137  0.064 -0.257  -0.027  -0.097 -0.038   0.162  0.080
##               fct()R fctr(hmdty)Mr fctr(hmdty)Ms
## fctr(ndct)N
## scl(ppdnst)
## scal(pvrty)
## scl(lg(m_))
## scl(pct_bs)
## scl(__2020)
## scal(mdn_g)
## fctr(prty)R
## fctr(hmdty)Mr -0.014
## fctr(hmdty)Ms  0.019 -0.021
## mean_pm25      0.052  0.379        -0.152
## optimizer (Nelder_Mead) convergence code: 0 (OK)
## Model is nearly unidentifiable: very large eigenvalue
##  - Rescale variables?
```