

Homework 4: Two reasonably non-standard problems

Applied Stats

Due 11 Dec 2020

1 Donald Duck

Your task is to write a short report on the factors that influenced the voters in Wisconsin to vote for Trump in 2016. Specifically, what are the most important demographic factors which seem to be causing a strong spatial pattern in Trump support? Is Trumpism a primarily urban/rural phenomenon? Or is it rather a racial phenomenon with Trump appealing to White voters? Maybe there's some other spatial explanatory variable which are important but we haven't looked at? Or is there really very little spatial variation, with Trump voters being evenly distributed throughout Wisconsin?

The file `wisconsin.RData` on the quercus site is loaded into R below

```
(load("../data/wisconsin.RData"))  
[1] "wisconsinCsubm" "wisconsinCm"      "wisconsinTract"
```

The three objects contain election results at the sub-county, county and census tract respectively. The variables in these objects include

- `propWhite`, `propInd`: proportion of each region which is White and Indegenous, respectively
- `pop`, `area`, `pdens`, `logPdens`: total population, surface area (square km), ratio of the two, and log of the ratio
- `'trump'`, `'Total'`: number of votes for Trump and total number of votes, respectively

The objects are `SpatialPolygonsDataFrame`'s, which contain boundaries of regions as well as data.

I created the file `resWisconsin.RData` by running

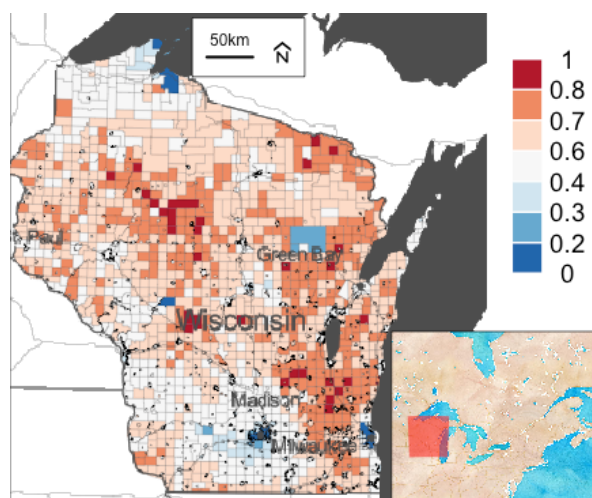
```
resTrump = diseasemapping::bym(trump ~ logPdens + propWhite +  
  propInd, data = wisconsinCsubm, prior = list(sd = c(log(2.5),  
  0.5), propSpatial = c(0.5, 0.5)), Ntrials = wisconsinCsubm$Total,  
  family = "binomial")  
save(resTrump, file = "../data/resWisconsin.RData", compress = "xz")
```

The prior distributions have a prior median of $\log(2.5)$ for σ and 0.5 for the spatial proportion ϕ . Loading the file from the quercus web site gives

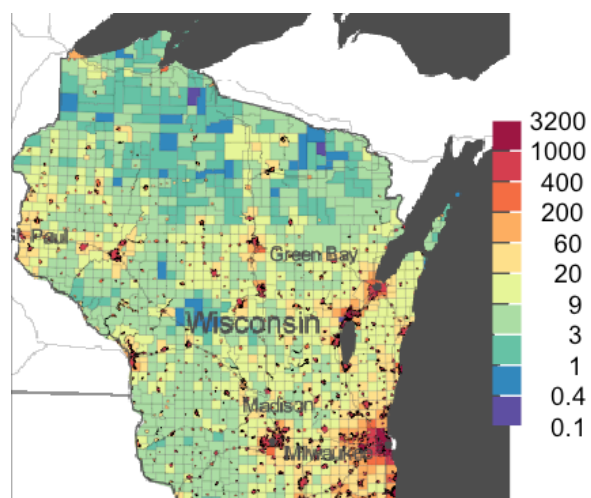
```
(load("../data/resWisconsin.RData"))  
[1] "resTrump"
```

Figure 1 was created with code in the appendix, the graphics files can be retrieved from quercus. Code for producing them is in the Appendix. Figure 2 was made by plotting the `random.mean` and `fitted.invlogit` objects from `resTrump`, which are the posterior means of the spatial random effect and predicted probability respectively.

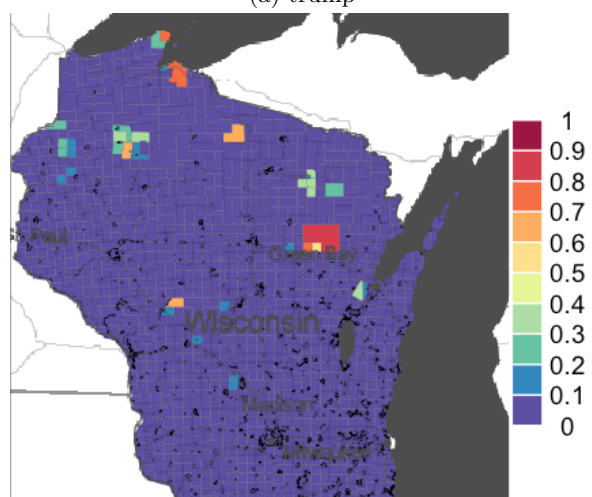
Here's a table of results which doesn't look very nice. The variables aren't well explained, they're on the log scale, and they refer to a one-unit change which isn't very interpretable for `logPdens`.



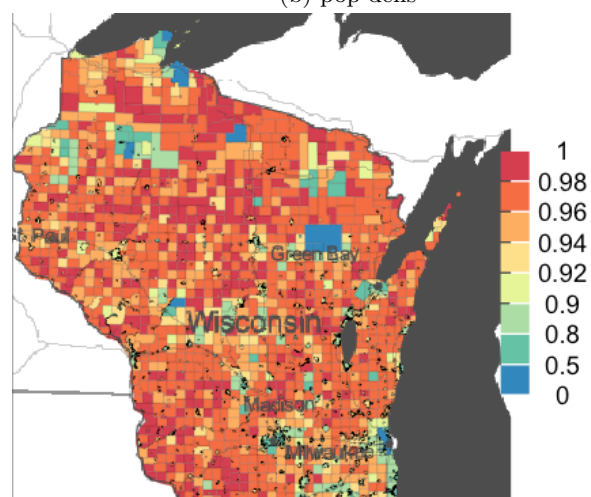
(a) trump



(b) pop dens



(c) ind



(d) white

Figure 1: maps

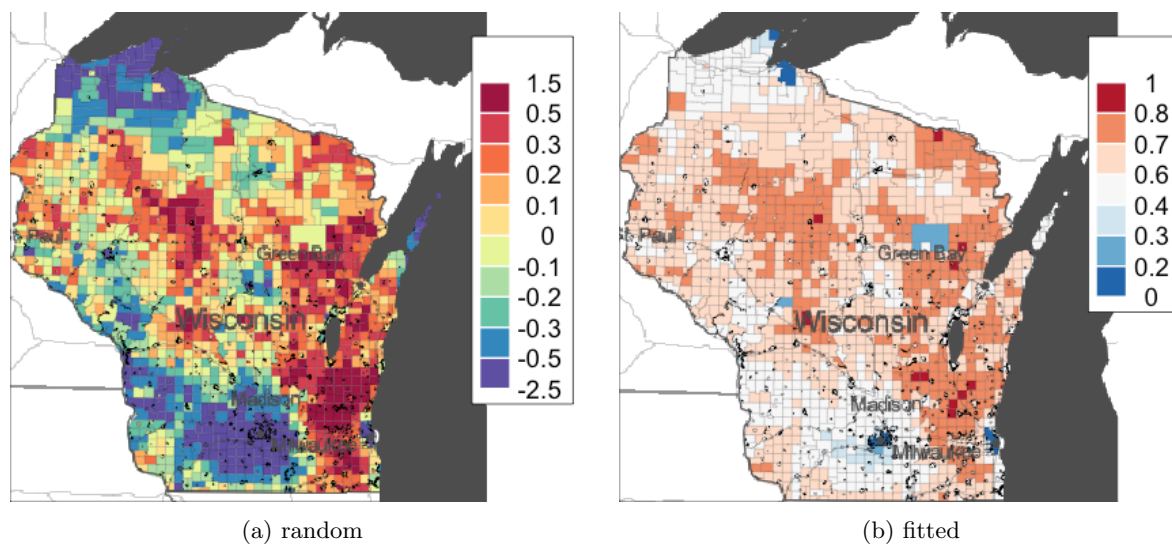


Figure 2: results

```
knitr::kable(resTrump$parameters$summary[, paste0(c(0.5,
0.025, 0.975), "quant")], digits = 5)
```

	0.5quant	0.025quant	0.975quant
(Intercept)	-0.56276	-0.82716	-0.29674
logPdens	-0.08105	-0.08979	-0.07232
propWhite	1.41879	1.15241	1.68307
propInd	-0.78943	-1.13430	-0.44628
sd	0.31830	0.30419	0.33446
propSpatial	0.96016	0.91715	0.98591

Some notes:

- Write a nice self-contained report clearly explaining the model you've used, about 2 pages of writing (plus tables and figures)
- You don't need to re-run the models, but certainly do so if you'd like. If you use this model you should justify why this model is sensible (or say why its wrong but you had no choice but use it anyways if you prefer).
- If you don't want to re-run the model, you can transform the numbers in the parameter table instead of scaling the covariates.
- Bonus points for finding and using the 2020 data instead of 2016!

2 Question 2

The file `England_shp.RData` on quercus is loaded as follows

```
(load("../data/England_shp.RData"))
```

```
[1] "UK_shp"
```

Each row in the dataset is a public health region in England. Variables in this data set include

- `pm25modelled`: concentrations of fine particulate matter (PM 2.5) in the health authority
- `cases, E`: number of COVID-19 cases up to 15 October 2020 and expected number (computed from population data and known incidence rates)
- `Unemployment, Ethnicity`: percent of individuals who are unemployed and ethnic ethnic minorities respectively

The file `englandResults` on quercus are computed as follows

```
UK_shp$logExpected = log(UK_shp$E)
# remove an island
UK2 = UK_shp[grep("Wight", UK_shp$Name, invert = TRUE),
]
englandRes = diseasemapping::bym(cases ~ offset(logExpected) +
  Ethnicity + modelledpm25 + Unemployment, prior = list(sd = c(0.5,
  0.5), propSpatial = c(0.5, 0.5)), family = "poisson",
  data = UK2)
```

Warning in `inla.model.properties.generic(inla.trim.family(model), mm[names(mm) == : Model 'bym2' in sec`
Use this model with extra care!!! Further warnings are disabled.

```
save(englandRes, file = "../data/englandRes.RData")
```

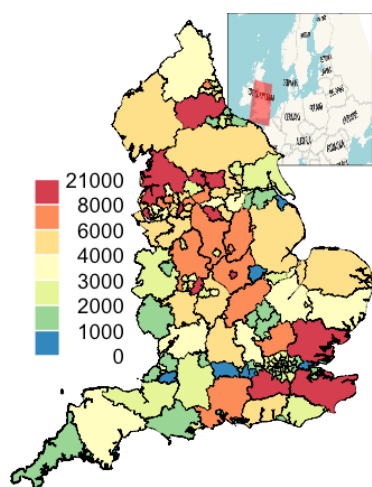
A bad table is

```
englandRes$parameters$summary[, c(1, 3:5)]
```

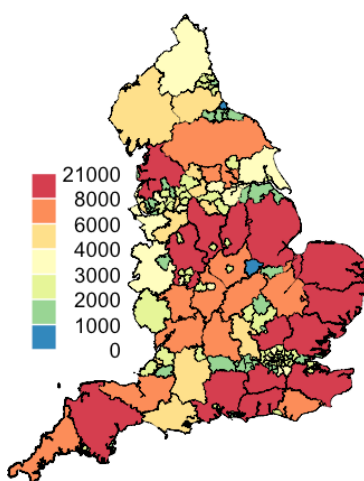
	mean	0.025quant	0.5quant	0.975quant
(Intercept)	-1.00751925	-1.523288113	-1.00723306	-0.49379891
Ethnicity	0.01204782	0.008097690	0.01204645	0.01600188
modelledpm25	0.05578866	-0.004375153	0.05574960	0.11611600
Unemployment	0.11321387	0.057647691	0.11321298	0.16873306
sd	0.29402227	0.258725327	0.29304915	0.33554463
propSpatial	0.89801415	0.767852225	0.90705357	0.97547907

Write a report investigating whether exposure to ambient air pollution makes individuals more susceptible to COVID-19 (about 2 pages of writing).

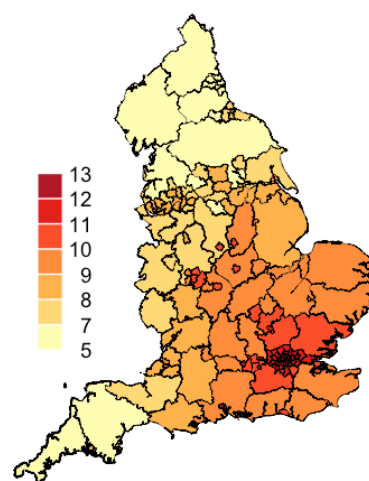
- Figures are available on the quercus site
- Someone with some knowledge of COVID-19 says that air pollution puts stress on the lungs and respiratory tract, so it should be expected that we'd see more covid where air pollution is high
- We would expected to see more COVID-19 where there is high unemployment, as such areas tend to have high deprivation and low access to health care.
- There are a number of hypotheses as to why areas with many ethnic minorities have more COVID-19, which has been documented in many countries. These include ethnic minorities being more likely to live in large, multi-generational households, work in high-risk occupations (i.e. the transport industry), and structural racism making access to health care harder.



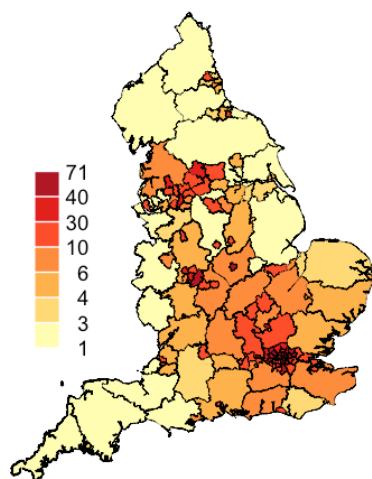
(a) cases



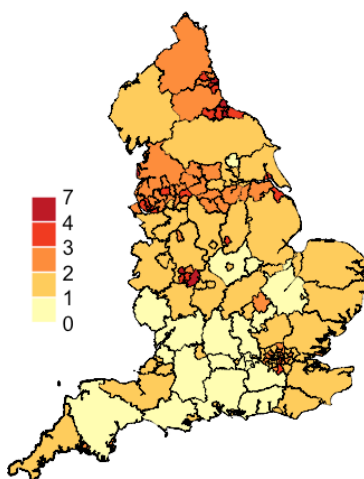
(b) E



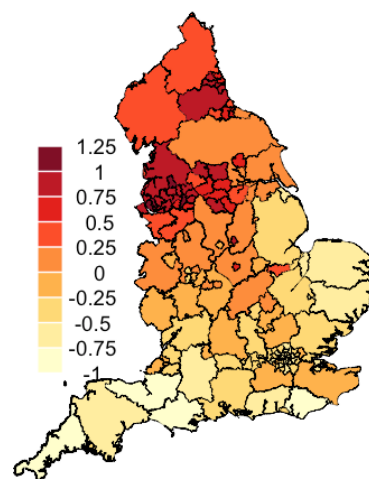
(c) pm



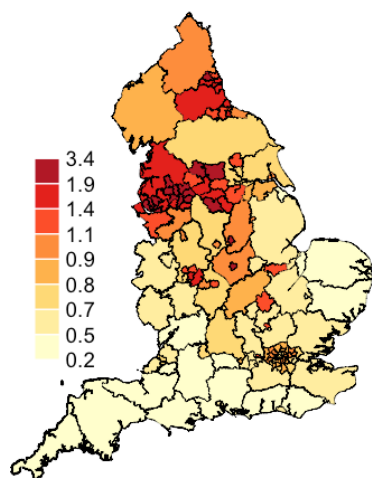
(d) eth



(e) unemp



(f) r



(g) lambda

Figure 3: map

3 Appendix

3.1 Donald

```
theColTrump = mapmisc::colourScale(wisconsinCsubm$propTrump,
  col = "RdBu", breaks = sort(unique(setdiff(c(0, 1, seq(0.2,
    0.8, by = 0.1))), 0.5))), style = "fixed", rev = TRUE)
theColPop = mapmisc::colourScale(wisconsinCsubm$pdens, col = "Spectral",
  breaks = 11, style = "equal", transform = "log", digits = 1,
  rev = TRUE)
theColWhite = mapmisc::colourScale(wisconsinCsubm$propWhite,
  col = "Spectral", breaks = c(0, 0.5, 0.8, 0.9, seq(0.9,
    1, by = 0.02)), style = "fixed", rev = TRUE)
theColInd = mapmisc::colourScale(wisconsinCsubm$propInd,
  col = "Spectral", breaks = seq(0, 1, by = 0.1), style = "fixed",
  rev = TRUE)
theBg = mapmisc::tonerToTrans(mapmisc::openmap(wisconsinCm,
  fact = 2, path = "stamen-toner"), col = "grey30")
theInset = mapmisc::openmap(wisconsinCm, zoom = 6, path = "stamen-watercolor",
  crs = mapmisc::crsMerc, buffer = c(0, 1500, 100, 700) *
    1000)

library("sp")
mapmisc::map.new(wisconsinCsubm, 0.85)
sp::plot(wisconsinCsubm, col = theColTrump$plot, add = TRUE,
  lwd = 0.2)
raster::plot(theBg, add = TRUE, maxpixels = 10^7)
mapmisc::insetMap(wisconsinCsubm, "bottomright", theInset,
  outer = TRUE, width = 0.35)
mapmisc::scaleBar(wisconsinCsubm, "top", cex = 0.8)
mapmisc::legendBreaks("topright", theColTrump, bty = "n",
  inset = 0)
mapmisc::map.new(wisconsinCsubm, 0.85)
plot(wisconsinCsubm, col = theColPop$plot, add = TRUE, lwd = 0.2)
plot(theBg, add = TRUE, maxpixels = 10^7)
mapmisc::legendBreaks("right", theColPop, bty = "n", inset = 0)
mapmisc::map.new(wisconsinCsubm, 0.85)
plot(wisconsinCsubm, col = theColInd$plot, add = TRUE, lwd = 0.2)
plot(theBg, add = TRUE, maxpixels = 10^7)
mapmisc::legendBreaks("right", theColInd, bty = "n", inset = 0)
mapmisc::map.new(wisconsinCsubm, 0.85)
plot(wisconsinCsubm, col = theColWhite$plot, add = TRUE,
  lwd = 0.2)
plot(theBg, add = TRUE, maxpixels = 10^7)
mapmisc::legendBreaks("right", theColWhite, bty = "n", inset = 0)

theColRandom = mapmisc::colourScale(resTrump$data$random.mean,
  col = "Spectral", breaks = 11, style = "quantile", rev = TRUE,
  dec = 1)
theColFit = mapmisc::colourScale(resTrump$data$fitted.invlogit,
  col = "RdBu", rev = TRUE, breaks = sort(unique(setdiff(c(0,
    1, seq(0.2, 0.8, by = 0.1))), 0.5))), style = "fixed")
mapmisc::map.new(wisconsinCsubm, 0.85)
plot(resTrump$data, col = theColRandom$plot, add = TRUE,
  lwd = 0.2)
```



```

plot(theBg, add = TRUE, maxpixels = 10^7)
mapmisc::legendBreaks("topright", theColRandom)
mapmisc::map.new(wisconsinCsubm, 0.85)
plot(resTrump$data, col = theColFit$plot, add = TRUE, lwd = 0.2)
plot(theBg, add = TRUE, maxpixels = 10^7)
mapmisc::legendBreaks("topright", theColFit)

```

3.2 Covid

```

casesCol = mapmisc::colourScale(UK2$cases, dec = -3, breaks = 12,
  col = "Spectral", style = "quantile", rev = TRUE)
Ecol = mapmisc::colourScale(UK2$E, breaks = casesCol$breaks,
  col = casesCol$col, style = "fixed")
pmCol = mapmisc::colourScale(UK2$modelledpm25, breaks = 9,
  dec = 0, style = "quantile")
ethCol = mapmisc::colourScale(UK2$Ethnicity, breaks = 9,
  digits = 1, style = "quantile")
uCol = mapmisc::colourScale(UK2$Unemployment, breaks = 12,
  dec = 0, style = "quantile")
rCol = mapmisc::colourScale(englandRes$data$random.mean,
  breaks = 12, dec = -log10(0.25), style = "quantile")
fCol = mapmisc::colourScale(englandRes$data$fitted.exp,
  breaks = 9, dec = 1, style = "quantile")
insetEngland1 = mapmisc::openmap(UK2, zoom = 3, fact = 4,
  path = "waze", crs = CRS("+init=epsg:3035"))
library("raster")
insetEngland = raster::crop(insetEngland1, extend(extent(insetEngland1),
  -c(25, 7, 4, 9.5) * 100 * 1000))

library("sp")
mapmisc::map.new(UK2)
mapmisc::insetMap(UK_shp, "topright", insetEngland, width = 0.4)
plot(UK2, col = casesCol$plot, add = TRUE, lwd = 0.2)
mapmisc::legendBreaks("left", casesCol, bty = "n")
mapmisc::map.new(UK2)
plot(UK2, col = Ecol$plot, add = TRUE, lwd = 0.2)
mapmisc::legendBreaks("left", casesCol, bty = "n")
mapmisc::map.new(UK2)
plot(UK2, col = pmCol$plot, add = TRUE, lwd = 0.2)
mapmisc::legendBreaks("left", pmCol, bty = "n")
mapmisc::map.new(UK2)
plot(UK2, col = ethCol$plot, add = TRUE, lwd = 0.2)
mapmisc::legendBreaks("left", ethCol, bty = "n")
mapmisc::map.new(UK2)
plot(UK2, col = uCol$plot, add = TRUE, lwd = 0.2)
mapmisc::legendBreaks("left", uCol, bty = "n")
mapmisc::map.new(UK2)
plot(UK2, col = rCol$plot, add = TRUE, lwd = 0.2)
mapmisc::legendBreaks("left", rCol, bty = "n")
mapmisc::map.new(UK2)
plot(UK2, col = fCol$plot, add = TRUE, lwd = 0.2)
mapmisc::legendBreaks("left", fCol, bty = "n")

```