Analysis before fitting the CAR model

Alvin Sheng

6/28/2021

```
library(here)
```

```
## here() starts at /Users/Alvin/Documents/NCSU_Fall_2021/NIH_SIP/flood-risk-health-effects
```

```
library(ape)
library(GGally)
```

```
## Loading required package: ggplot2
```

```
## Registered S3 method overwritten by 'GGally':
##   method from
##   +.gg   ggplot2
```

```
library(usdm)
```

```
## Loading required package: sp
```

```
## Loading required package: raster
```

```
##
## Attaching package: 'raster'
```

```
## The following objects are masked from 'package:ape':
##
##     rotate, zoom
```

```
fls_model_df <- readRDS(here("intermediary_data/fls_model_df.rds"))
```

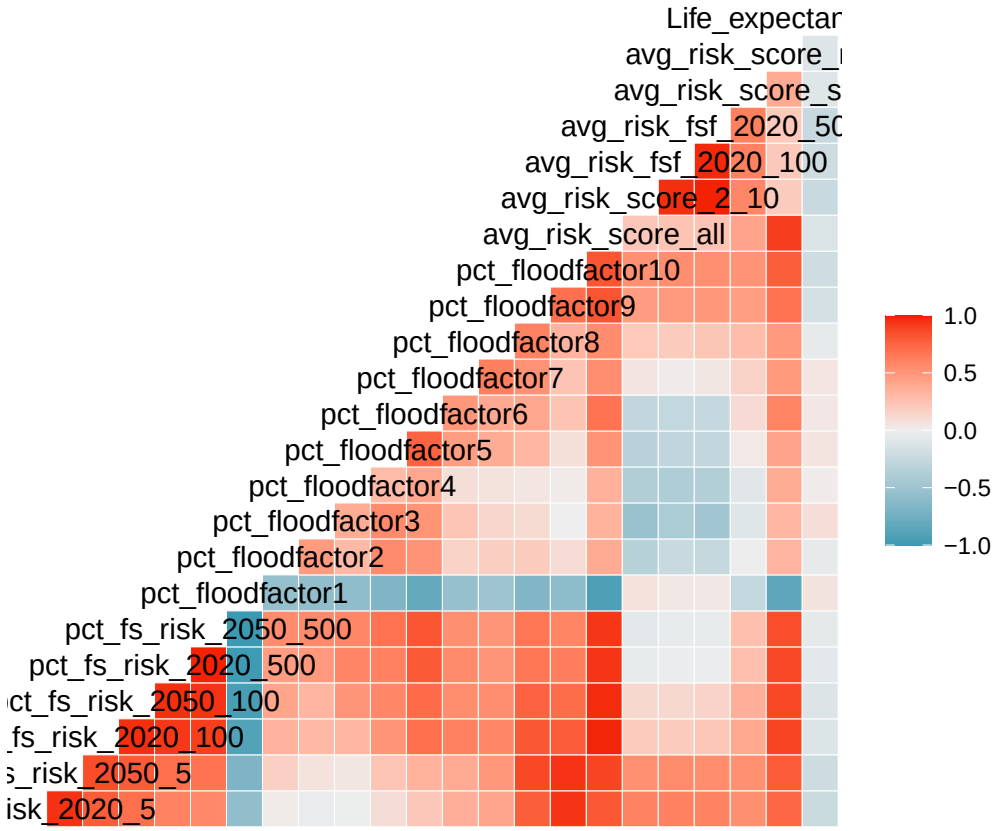
Checking for multicollinearity among the covariates

`S.CARleroux()` automatically puts a fixed ridge penalty on the beta coefficients. Therefore, the large number of covariates and multicollinearity would be accounted for.

Flood risk variables

```
ggcorr(data = fls_model_df[, c(12:33, ncol(fls_model_df))], progress = F)
```

```
## Warning: Ignoring unknown parameters: progress
```




```
flood_cor <- cor(fls_model_df[complete.cases(fls_model_df[, c(12:33, ncol(fls_model_df))]), c(12:33, nc

flood_cor[nrow(flood_cor), ]
```

```
##     pct_fs_risk_2020_5     pct_fs_risk_2050_5   pct_fs_risk_2020_100
##            -0.22403513            -0.20639134            -0.11964945
##   pct_fs_risk_2050_100   pct_fs_risk_2020_500   pct_fs_risk_2050_500
##            -0.11188615            -0.06844810            -0.06162909
##       pct_floodfactor1       pct_floodfactor2       pct_floodfactor3
##             0.06104602            -0.04091275             0.09312308
##       pct_floodfactor4       pct_floodfactor5       pct_floodfactor6
##             0.01773260             0.05967750             0.04447554
##       pct_floodfactor7       pct_floodfactor8       pct_floodfactor9
##             0.04774946            -0.05311612            -0.16650941
##      pct_floodfactor10     avg_risk_score_all    avg_risk_score_2_10
##            -0.20238523            -0.12602894            -0.23725350
##  avg_risk_fsf_2020_100  avg_risk_fsf_2020_500    avg_risk_score_sfha
##            -0.21519007            -0.24188687            -0.09702686
## avg_risk_score_no_sfha Life expectancy, 2014*
##            -0.10459813             1.00000000
```

For each variable, I take the summary of its correlations with other variables, not including itself.

```
diag(flood_cor) <- NA

summary(flood_cor)
```

```
##  pct_fs_risk_2020_5 pct_fs_risk_2050_5 pct_fs_risk_2020_100
##  Min.   :-0.5669    Min.   :-0.6817    Min.   :-0.9052
```

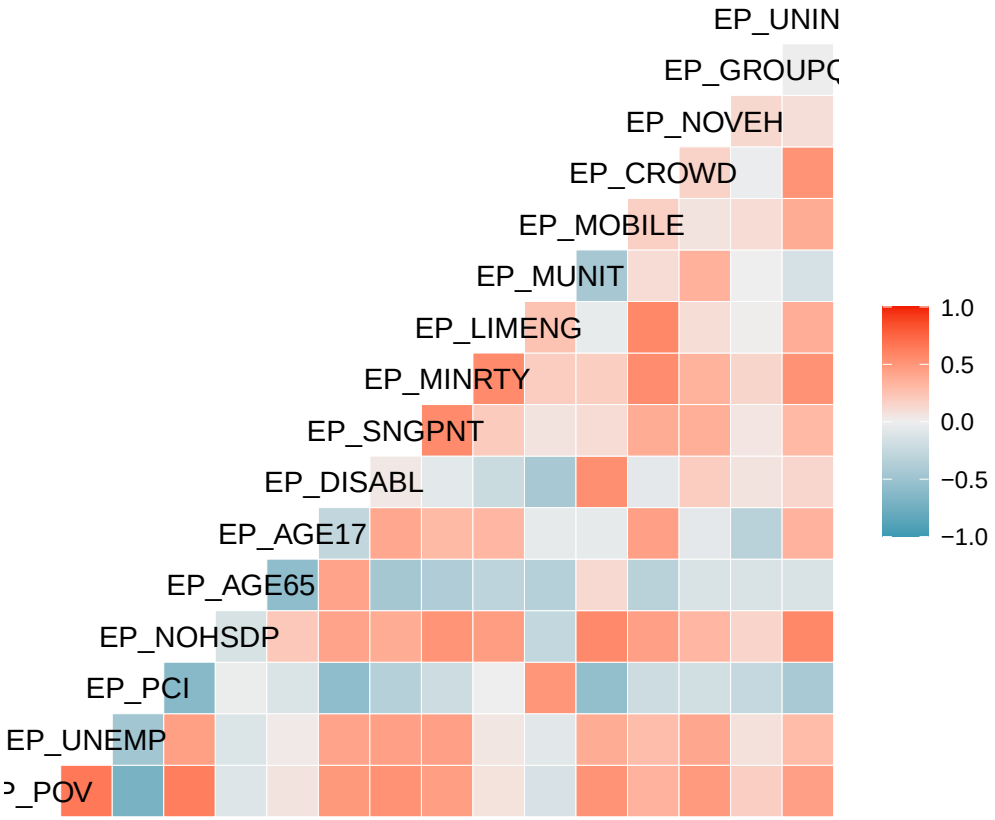
```
##  1st Qu.: 0.1282    1st Qu.: 0.2640    1st Qu.: 0.3083
##  Median : 0.5842    Median : 0.5462    Median : 0.6225
##  Mean   : 0.4349    Mean   : 0.4806    Mean   : 0.5151
##  3rd Qu.: 0.7316    3rd Qu.: 0.7986    3rd Qu.: 0.8482
##  Max.   : 0.9499    Max.   : 0.9499    Max.   : 0.9788
##  NA's   :1          NA's   :1          NA's   :1
##  pct_fs_risk_2050_100 pct_fs_risk_2020_500 pct_fs_risk_2050_500
##  Min.   :-0.9470      Min.   :-0.9852      Min.   :-0.9999
##  1st Qu.: 0.3352      1st Qu.: 0.3195      1st Qu.: 0.3371
##  Median : 0.5894      Median : 0.5973      Median : 0.5742
##  Mean   : 0.5029      Mean   : 0.4765      Mean   : 0.4758
##  3rd Qu.: 0.7864      3rd Qu.: 0.7746      3rd Qu.: 0.7825
##  Max.   : 0.9621      Max.   : 0.9850      Max.   : 0.9850
##  NA's   :1            NA's   :1            NA's   :1
##  pct_floodfactor1   pct_floodfactor2    pct_floodfactor3   pct_floodfactor4
##  Min.   :-0.99993   Min.   :-0.554100   Min.   :-0.56458   Min.   :-0.581415
##  1st Qu.:-0.84646   1st Qu.: 0.007492   1st Qu.:-0.02187   1st Qu.:-0.004162
##  Median :-0.58692   Median : 0.207284   Median : 0.19554   Median : 0.086429
##  Mean   :-0.56661   Mean   : 0.171865   Mean   : 0.12839   Mean   : 0.117315
##  3rd Qu.:-0.54250   3rd Qu.: 0.408212   3rd Qu.: 0.36865   3rd Qu.: 0.371173
##  Max.   : 0.07776   Max.   : 0.562637   Max.   : 0.56405   Max.   : 0.588232
##  NA's   :1          NA's   :1           NA's   :1          NA's   :1
##  pct_floodfactor5   pct_floodfactor6  pct_floodfactor7  pct_floodfactor8
##  Min.   :-0.67623   Min.   :-0.8160   Min.   :-0.5386   Min.   :-0.5557
##  1st Qu.: 0.06699   1st Qu.: 0.1405   1st Qu.: 0.1093   1st Qu.: 0.2232
##  Median : 0.36991   Median : 0.4254   Median : 0.3797   Median : 0.4314
##  Mean   : 0.25561   Mean   : 0.3266   Mean   : 0.3068   Mean   : 0.3587
##  3rd Qu.: 0.54810   3rd Qu.: 0.6622   3rd Qu.: 0.5355   3rd Qu.: 0.5556
##  Max.   : 0.75447   Max.   : 0.8161   Max.   : 0.6856   Max.   : 0.6856
##  NA's   :1          NA's   :1         NA's   :1         NA's   :1
##  pct_floodfactor9  pct_floodfactor10 avg_risk_score_all avg_risk_score_2_10
##  Min.   :-0.6663   Min.   :-0.5924   Min.   :-0.9302    Min.   :-0.5240
##  1st Qu.: 0.3417   1st Qu.: 0.1381   1st Qu.: 0.3432    1st Qu.:-0.1972
##  Median : 0.5078   Median : 0.5324   Median : 0.5803    Median : 0.1638
##  Mean   : 0.4598   Mean   : 0.4240   Mean   : 0.5265    Mean   : 0.1610
##  3rd Qu.: 0.6943   3rd Qu.: 0.7088   3rd Qu.: 0.8735    3rd Qu.: 0.5103
##  Max.   : 0.8749   Max.   : 0.9421   Max.   : 0.9788    Max.   : 0.9849
##  NA's   :1         NA's   :1         NA's   :1          NA's   :1
##  avg_risk_fsf_2020_100 avg_risk_fsf_2020_500 avg_risk_score_sfha
##  Min.   :-0.4151       Min.   :-0.4742       Min.   :-0.26923
##  1st Qu.:-0.1715       1st Qu.:-0.1917       1st Qu.: 0.05147
##  Median : 0.1634       Median : 0.1829       Median : 0.32088
##  Mean   : 0.1776       Mean   : 0.1805       Mean   : 0.27294
##  3rd Qu.: 0.5305       3rd Qu.: 0.5247       3rd Qu.: 0.50061
##  Max.   : 0.9710       Max.   : 0.9849       Max.   : 0.61792
##  NA's   :1             NA's   :1             NA's   :1
##  avg_risk_score_no_sfha Life expectancy, 2014*
##  Min.   :-0.8566        Min.   :-0.241887
##  1st Qu.: 0.3200        1st Qu.:-0.193416
##  Median : 0.5062        Median :-0.100812
##  Mean   : 0.4796        Mean   :-0.088779
##  3rd Qu.: 0.7968        3rd Qu.: 0.003071
##  Max.   : 0.9171        Max.   : 0.093123
##  NA's   :1              NA's   :1
```

Many of the flood risk variables are very correlated.

SVI Variables

```
ggcorr(data = fls_model_df[, 34:49], progress = F)
```

```
## Warning: Ignoring unknown parameters: progress
```




```
(svi_cor <- cor(fls_model_df[complete.cases(fls_model_df[, 34:49]), 34:49]))
```

```
##                 EP_POV    EP_UNEMP        EP_PCI  EP_NOHSDP    EP_AGE65
## EP_POV      1.00000000  0.65154857 -0.7103056275  0.6254853 -0.10657946
## EP_UNEMP    0.65154857  1.00000000 -0.4612103295  0.4532279 -0.12022992
## EP_PCI     -0.71030563 -0.46121033  1.0000000000 -0.6262797 -0.01683822
## EP_NOHSDP   0.62548531  0.45322790 -0.6262796652  1.0000000 -0.14111559
## EP_AGE65   -0.10657946 -0.12022992 -0.0168382185 -0.1411156  1.00000000
## EP_AGE17    0.06475125  0.02827912 -0.1243383212  0.2177170 -0.57476412
## EP_DISABL   0.48586544  0.42395670 -0.5794913164  0.4271945  0.42412972
## EP_SNGPNT   0.51887260  0.44944620 -0.3470227931  0.3785708 -0.45743733
## EP_MINRTY   0.45604006  0.44889574 -0.2090773909  0.5090794 -0.38753061
## EP_LIMENG   0.07009639  0.04580458 -0.0009591033  0.4598988 -0.29887556
## EP_MUNIT   -0.13836697 -0.07605523  0.4995965155 -0.2657774 -0.35262468
## EP_MOBILE   0.51545109  0.37718262 -0.5600628496  0.5625601  0.12101238
## EP_CROWD    0.34431163  0.28356947 -0.2148883334  0.4496746 -0.32904339
## EP_NOVEH    0.48055115  0.40918645 -0.1791498409  0.3147644 -0.12910911
## EP_GROUPQ   0.18226204  0.07750866 -0.2538691690  0.1511509 -0.13308534
## EP_UNINSUR  0.44883447  0.28895719 -0.4289763742  0.5736897 -0.12938690
```

```
##               EP_AGE17   EP_DISABL   EP_SNGPNT   EP_MINRTY     EP_LIMENG
## EP_POV      0.06475125  0.48586544  0.51887260  0.45604006  0.0700963895
## EP_UNEMP    0.02827912  0.42395670  0.44944620  0.44889574  0.0458045846
## EP_PCI     -0.12433832 -0.57949132 -0.34702279 -0.20907739 -0.0009591033
## EP_NOHSDP   0.21771701  0.42719446  0.37857077  0.50907944  0.4598988190
## EP_AGE65   -0.57476412  0.42412972 -0.45743733 -0.38753061 -0.2988755623
## EP_AGE17    1.00000000 -0.27658786  0.40354000  0.30149642  0.3149737539
## EP_DISABL  -0.27658786  1.00000000  0.04096641 -0.07177311 -0.2240600940
## EP_SNGPNT   0.40354000  0.04096641  1.00000000  0.56026164  0.2001874832
## EP_MINRTY   0.30149642 -0.07177311  0.56026164  1.00000000  0.5624017490
## EP_LIMENG   0.31497375 -0.22406009  0.20018748  0.56240175  1.0000000000
## EP_MUNIT   -0.04236691 -0.43118805  0.06426726  0.19472114  0.2498708116
## EP_MOBILE  -0.05200253  0.53993350  0.10384622  0.18382861 -0.0380200909
## EP_CROWD    0.44438111 -0.06539029  0.37813629  0.55290358  0.5775159407
## EP_NOVEH   -0.06444933  0.19360660  0.36659333  0.34205431  0.1000213363
## EP_GROUPQ  -0.32948858  0.06484295  0.05079929  0.14754534  0.0065066513
## EP_UNINSUR  0.33667300  0.13672110  0.30741216  0.52056679  0.3744399807
##                EP_MUNIT   EP_MOBILE    EP_CROWD    EP_NOVEH    EP_GROUPQ
## EP_POV     -0.138366969  0.51545109  0.34431163  0.48055115  0.182262045
## EP_UNEMP   -0.076055227  0.37718262  0.28356947  0.40918645  0.077508662
## EP_PCI      0.499596515 -0.56006285 -0.21488833 -0.17914984 -0.253869169
## EP_NOHSDP  -0.265777410  0.56256010  0.44967465  0.31476438  0.151150853
## EP_AGE65   -0.352624675  0.12101238 -0.32904339 -0.12910911 -0.133085338
## EP_AGE17   -0.042366909 -0.05200253  0.44438111 -0.06444933 -0.329488580
## EP_DISABL  -0.431188049  0.53993350 -0.06539029  0.19360660  0.064842953
## EP_SNGPNT   0.064267262  0.10384622  0.37813629  0.36659333  0.050799287
## EP_MINRTY   0.194721142  0.18382861  0.55290358  0.34205431  0.147545344
## EP_LIMENG   0.249870812 -0.03802009  0.57751594  0.10002134  0.006506651
## EP_MUNIT    1.000000000 -0.44223776  0.10540992  0.35186391 -0.000369879
## EP_MOBILE  -0.442237763  1.00000000  0.17663229  0.06226905  0.105594681
## EP_CROWD    0.105409916  0.17663229  1.00000000  0.15812169 -0.023615080
## EP_NOVEH    0.351863906  0.06226905  0.15812169  1.00000000  0.129367093
## EP_GROUPQ  -0.000369879  0.10559468 -0.02361508  0.12936709  1.000000000
## EP_UNINSUR -0.155772925  0.37796514  0.51676600  0.09040975  0.003735602
##              EP_UNINSUR
## EP_POV      0.448834467
## EP_UNEMP    0.288957193
## EP_PCI     -0.428976374
## EP_NOHSDP   0.573689707
## EP_AGE65   -0.129386901
## EP_AGE17    0.336673001
## EP_DISABL   0.136721103
## EP_SNGPNT   0.307412160
## EP_MINRTY   0.520566789
## EP_LIMENG   0.374439981
## EP_MUNIT   -0.155772925
## EP_MOBILE   0.377965140
## EP_CROWD    0.516766001
## EP_NOVEH    0.090409750
## EP_GROUPQ   0.003735602
## EP_UNINSUR  1.000000000
```

```
diag(svi_cor) <- NA
```

```
summary(svi_cor)
```

```
##      EP_POV            EP_UNEMP           EP_PCI          EP_NOHSDP
##  Min.   :-0.71031   Min.   :-0.46121   Min.   :-0.7103   Min.   :-0.6263
##  1st Qu.: 0.06742   1st Qu.: 0.03704   1st Qu.:-0.5106   1st Qu.: 0.1844
##  Median : 0.44883   Median : 0.28896   Median :-0.2539   Median : 0.4272
##  Mean   : 0.25925   Mean   : 0.21867   Mean   :-0.2809   Mean   : 0.2727
##  3rd Qu.: 0.50066   3rd Qu.: 0.43643   3rd Qu.:-0.1517   3rd Qu.: 0.4845
##  Max.   : 0.65155   Max.   : 0.65155   Max.   : 0.4996   Max.   : 0.6255
##  NA's   :1          NA's   :1          NA's   :1         NA's   :1
##     EP_AGE65          EP_AGE17          EP_DISABL         EP_SNGPNT
##  Min.   :-0.5748   Min.   :-0.57476   Min.   :-0.57949   Min.   :-0.45744
##  1st Qu.:-0.3408   1st Qu.:-0.09439   1st Qu.:-0.14792   1st Qu.: 0.05753
##  Median :-0.1331   Median : 0.02828   Median : 0.06484   Median : 0.30741
##  Mean   :-0.1754   Mean   : 0.04319   Mean   : 0.07258   Mean   : 0.20123
##  3rd Qu.:-0.1134   3rd Qu.: 0.30824   3rd Qu.: 0.42404   3rd Qu.: 0.39106
##  Max.   : 0.4241   Max.   : 0.44438   Max.   : 0.53993   Max.   : 0.56026
##  NA's   :1         NA's   :1          NA's   :1          NA's   :1
##    EP_MINRTY         EP_LIMENG           EP_MUNIT          EP_MOBILE
##  Min.   :-0.3875   Min.   :-0.298876   Min.   :-0.44224   Min.   :-0.56006
##  1st Qu.: 0.1657   1st Qu.: 0.002774   1st Qu.:-0.21078   1st Qu.: 0.01212
##  Median : 0.3421   Median : 0.100021   Median :-0.04237   Median : 0.12101
##  Mean   : 0.2741   Mean   : 0.159987   Mean   :-0.02927   Mean   : 0.13560
##  3rd Qu.: 0.5148   3rd Qu.: 0.344707   3rd Qu.: 0.15007   3rd Qu.: 0.37757
##  Max.   : 0.5624   Max.   : 0.577516   Max.   : 0.49960   Max.   : 0.56256
##  NA's   :1         NA's   :1           NA's   :1          NA's   :1
##     EP_CROWD          EP_NOVEH          EP_GROUPQ          EP_UNINSUR
##  Min.   :-0.3290   Min.   :-0.17915   Min.   :-0.32949   Min.   :-0.42898
##  1st Qu.: 0.0409   1st Qu.: 0.07634   1st Qu.:-0.01199   1st Qu.: 0.04707
##  Median : 0.2836   Median : 0.15812   Median : 0.05080   Median : 0.30741
##  Mean   : 0.2236   Mean   : 0.17507   Mean   : 0.01193   Mean   : 0.21747
##  3rd Qu.: 0.4470   3rd Qu.: 0.34696   3rd Qu.: 0.11748   3rd Qu.: 0.41340
##  Max.   : 0.5775   Max.   : 0.48055   Max.   : 0.18226   Max.   : 0.57369
##  NA's   :1         NA's   :1          NA's   :1          NA's   :1
```

Air pollution variables

```
ggpairs(data = fls_model_df, columns = 50:55, progress = F)
```

```
## Warning: Removed 1 rows containing non-finite values (stat_density).
```

```
## Warning in ggally_statistic(data = data, mapping = mapping, na.rm = na.rm, :
## Removing 1 row that contained a missing value
```

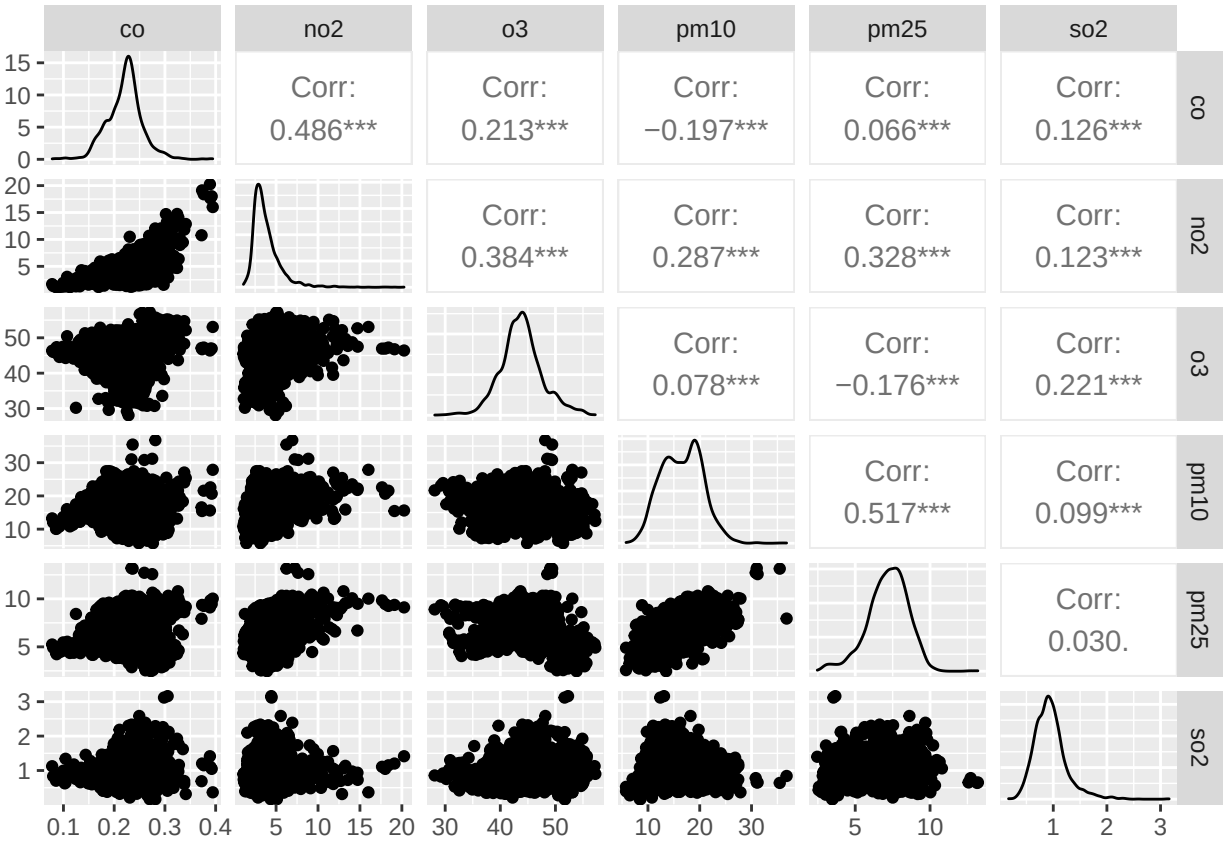
```
## Warning in ggally_statistic(data = data, mapping = mapping, na.rm = na.rm, :
## Removing 1 row that contained a missing value
```

```
## Warning in ggally_statistic(data = data, mapping = mapping, na.rm = na.rm, :
## Removing 1 row that contained a missing value
```

```
## Warning in ggally_statistic(data = data, mapping = mapping, na.rm = na.rm, :
## Removing 1 row that contained a missing value
```

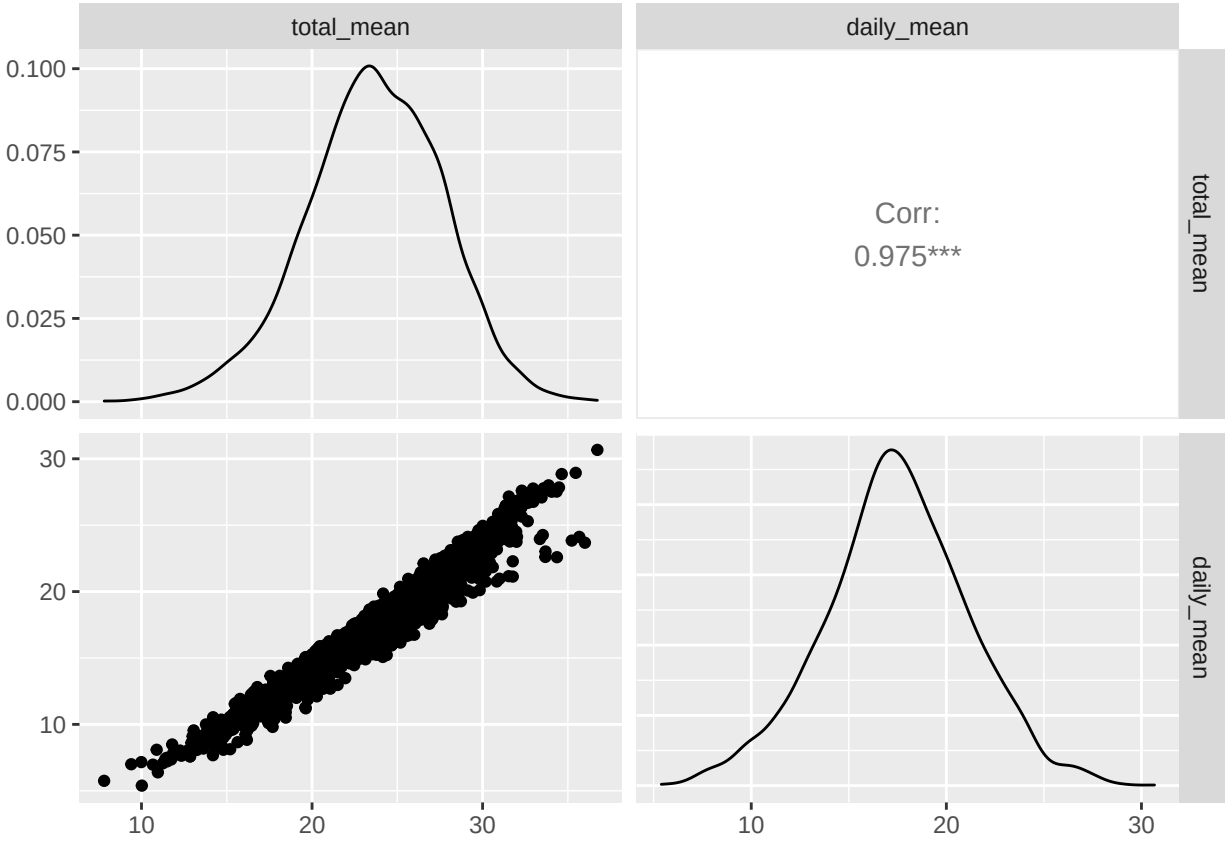
```
## Warning in ggally_statistic(data = data, mapping = mapping, na.rm = na.rm, :
## Removing 1 row that contained a missing value
## Warning: Removed 1 rows containing missing values (geom_point).
## Warning: Removed 1 rows containing non-finite values (stat_density).
## Warning in ggally_statistic(data = data, mapping = mapping, na.rm = na.rm, :
## Removing 1 row that contained a missing value
## Warning in ggally_statistic(data = data, mapping = mapping, na.rm = na.rm, :
## Removing 1 row that contained a missing value
## Warning in ggally_statistic(data = data, mapping = mapping, na.rm = na.rm, :
## Removing 1 row that contained a missing value
## Warning in ggally_statistic(data = data, mapping = mapping, na.rm = na.rm, :
## Removing 1 row that contained a missing value
## Warning: Removed 1 rows containing missing values (geom_point).
## Warning: Removed 1 rows containing missing values (geom_point).
## Warning: Removed 1 rows containing non-finite values (stat_density).
## Warning in ggally_statistic(data = data, mapping = mapping, na.rm = na.rm, :
## Removing 1 row that contained a missing value
## Warning in ggally_statistic(data = data, mapping = mapping, na.rm = na.rm, :
## Removing 1 row that contained a missing value
## Warning in ggally_statistic(data = data, mapping = mapping, na.rm = na.rm, :
## Removing 1 row that contained a missing value
## Warning: Removed 1 rows containing missing values (geom_point).
## Warning: Removed 1 rows containing missing values (geom_point).
## Warning: Removed 1 rows containing missing values (geom_point).
## Warning: Removed 1 rows containing non-finite values (stat_density).
## Warning in ggally_statistic(data = data, mapping = mapping, na.rm = na.rm, :
## Removing 1 row that contained a missing value
## Warning in ggally_statistic(data = data, mapping = mapping, na.rm = na.rm, :
## Removing 1 row that contained a missing value
## Warning: Removed 1 rows containing missing values (geom_point).
## Warning: Removed 1 rows containing missing values (geom_point).
## Warning: Removed 1 rows containing missing values (geom_point).
## Warning: Removed 1 rows containing missing values (geom_point).
## Warning: Removed 1 rows containing non-finite values (stat_density).
## Warning in ggally_statistic(data = data, mapping = mapping, na.rm = na.rm, :
## Removing 1 row that contained a missing value
```

```
## Warning: Removed 1 rows containing missing values (geom_point).
## Warning: Removed 1 rows containing missing values (geom_point).
## Warning: Removed 1 rows containing missing values (geom_point).
## Warning: Removed 1 rows containing missing values (geom_point).
## Warning: Removed 1 rows containing missing values (geom_point).
## Warning: Removed 1 rows containing non-finite values (stat_density).
```



Smoking prevalence variables

```
ggpairs(data = fls_model_df, columns = 56:57, progress = F)
```



The correlation between total_mean and daily_mean is almost one.

Non-spatial modeling

```
Y <- fls_model_df$`Life expectancy, 2014*`

# extract the covariates matrix

X <- fls_model_df[, 12:(ncol(fls_model_df) - 1)]

X           <- scale(X) # Scale covariates
X[is.na(X)] <- 0        # Fill in missing values with the mean
```

```
fls_lm <- lm(Y ~ X)
```

```
summary(fls_lm)
```

```
##
## Call:
## lm(formula = Y ~ X)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -7.6148 -0.5842  0.0020  0.5967  4.7797
##
## Coefficients: (1 not defined because of singularities)
```

```
##                          Estimate Std. Error  t value Pr(>|t|)
## (Intercept)              77.74540    0.01814 4286.284  < 2e-16 ***
## Xpct_fs_risk_2020_5      -0.12978    0.08373   -1.550 0.121230
## Xpct_fs_risk_2050_5      -0.27259    0.15827   -1.722 0.085121 .
## Xpct_fs_risk_2020_100     0.04546    0.18177    0.250 0.802542
## Xpct_fs_risk_2050_100    -0.14482    0.16201   -0.894 0.371435
## Xpct_fs_risk_2020_500    -0.17210    0.24889   -0.691 0.489310
## Xpct_fs_risk_2050_500     0.99660    1.59887    0.623 0.533126
## Xpct_floodfactor1         1.10693    7.39622    0.150 0.881041
## Xpct_floodfactor2        -0.02279    0.60871   -0.037 0.970132
## Xpct_floodfactor3        -0.03343    0.89880   -0.037 0.970337
## Xpct_floodfactor4         0.03795    1.56848    0.024 0.980699
## Xpct_floodfactor5         0.07052    0.40655    0.173 0.862302
## Xpct_floodfactor6        -0.04761    0.88780   -0.054 0.957232
## Xpct_floodfactor7         0.01979    0.21318    0.093 0.926062
## Xpct_floodfactor8        -0.03836    0.04656   -0.824 0.410098
## Xpct_floodfactor9         0.10969    0.13559    0.809 0.418581
## Xpct_floodfactor10             NA         NA       NA       NA
## Xavg_risk_score_all       0.43074    4.57588    0.094 0.925010
## Xavg_risk_score_2_10      0.35554    0.17976    1.978 0.048040 *
## Xavg_risk_fsf_2020_100    0.19057    0.10067    1.893 0.058445 .
## Xavg_risk_fsf_2020_500   -0.49068    0.20284   -2.419 0.015617 *
## Xavg_risk_score_sfha      0.05073    0.02829    1.793 0.073015 .
## Xavg_risk_score_no_sfha   0.14669    0.05255    2.791 0.005280 **
## XEP_POV                  -0.29373    0.03807   -7.715 1.63e-14 ***
## XEP_UNEMP                 0.02219    0.02751    0.807 0.419908
## XEP_PCI                   0.08419    0.03861    2.180 0.029304 *
## XEP_NOHSDP               -0.29450    0.04189   -7.030 2.54e-12 ***
## XEP_AGE65                 0.33244    0.03494    9.514  < 2e-16 ***
## XEP_AGE17                -0.19718    0.03703   -5.325 1.08e-07 ***
## XEP_DISABL               -0.44940    0.03297  -13.631  < 2e-16 ***
## XEP_SNGPNT               -0.07873    0.02899   -2.716 0.006653 **
## XEP_MINRTY               -0.44342    0.04876   -9.094  < 2e-16 ***
## XEP_LIMENG                0.60838    0.03503   17.369  < 2e-16 ***
## XEP_MUNIT                 0.10052    0.03304    3.042 0.002367 **
## XEP_MOBILE               -0.14341    0.03022   -4.745 2.18e-06 ***
## XEP_CROWD                 0.10004    0.02741    3.651 0.000266 ***
## XEP_NOVEH                -0.02927    0.02881   -1.016 0.309836
## XEP_GROUPQ                0.08741    0.02560    3.415 0.000646 ***
## XEP_UNINSUR              -0.19173    0.02774   -6.911 5.83e-12 ***
## Xco                      -0.20118    0.02709   -7.428 1.42e-13 ***
## Xno2                      0.09657    0.04059    2.379 0.017400 *
## Xo3                      -0.08766    0.02538   -3.455 0.000559 ***
## Xpm10                     0.05711    0.02846    2.007 0.044840 *
## Xpm25                    -0.28845    0.02902   -9.938  < 2e-16 ***
## Xso2                      0.04559    0.02105    2.166 0.030423 *
## Xtotal_mean              -1.00888    0.15038   -6.709 2.33e-11 ***
## Xdaily_mean               0.13921    0.15642    0.890 0.373529
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.011 on 3062 degrees of freedom
## Multiple R-squared:  0.8214, Adjusted R-squared:  0.8188
## F-statistic:   313 on 45 and 3062 DF,  p-value: < 2.2e-16
```

Checking for spatial autocorrelation

```
W <- readRDS(here("intermediary_data", "countyadj_reorganize.rds"))
```

Moran's I

```
(moran_results <- Moran.I(residuals(fls_lm), W))
```

```
## $observed
## [1] 0.2953613
##
## $expected
## [1] -0.0003218539
##
## $sd
## [1] 0.01057679
##
## $p.value
## [1] 0
```

The p-value is negligible, so we can reject the null hypothesis of zero spatial autocorrelation. Since the observed value of I is significantly greater then the expected value, the life expectancies are positively autocorrelated, in contrast to negatively autocorrelated. Thus, using a CAR model is justified.

Using VIF to exlude variables

```
X <- fls_model_df[, 12:(ncol(fls_model_df) - 1)]

X <- X[, names(X) != "pct_floodfactor1"]

X           <- scale(X) # Scale covariates

X <- as.data.frame(X)
```

```
vif(X)
```

```
##                 Variables          VIF
## 1      pct_fs_risk_2020_5    20.887236
## 2      pct_fs_risk_2050_5    75.169888
## 3    pct_fs_risk_2020_100    98.039137
## 4    pct_fs_risk_2050_100    78.060121
## 5    pct_fs_risk_2020_500   184.242932
## 6    pct_fs_risk_2050_500  7675.863781
## 7        pct_floodfactor2    87.124313
## 8        pct_floodfactor3   393.259773
## 9        pct_floodfactor4  2666.163478
## 10       pct_floodfactor5   392.270485
## 11       pct_floodfactor6  4226.245166
## 12       pct_floodfactor7   588.992691
## 13       pct_floodfactor8    48.880929
## 14       pct_floodfactor9  3286.079174
## 15      pct_floodfactor10 21259.891968
## 16     avg_risk_score_all 62899.973519
```

```
## 17    avg_risk_score_2_10    95.703864
## 18  avg_risk_fsf_2020_100    30.204905
## 19  avg_risk_fsf_2020_500   121.787408
## 20    avg_risk_score_sfha     2.417030
## 21 avg_risk_score_no_sfha     8.703305
## 22                 EP_POV     4.353438
## 23               EP_UNEMP     2.309368
## 24                 EP_PCI     4.597700
## 25              EP_NOHSDP     5.620278
## 26               EP_AGE65     3.809307
## 27               EP_AGE17     4.387969
## 28              EP_DISABL     3.379470
## 29              EP_SNGPNT     2.461728
## 30              EP_MINRTY     7.260267
## 31              EP_LIMENG     3.841383
## 32               EP_MUNIT     3.407034
## 33              EP_MOBILE     2.903874
## 34               EP_CROWD     2.467657
## 35               EP_NOVEH     2.513816
## 36              EP_GROUPQ     2.123159
## 37             EP_UNINSUR     2.355983
## 38                     co     2.294206
## 39                    no2     5.209123
## 40                     o3     1.956409
## 41                   pm10     2.473370
## 42                   pm25     2.499547
## 43                    so2     1.367940
## 44             total_mean    67.305006
## 45             daily_mean    73.049730
```

```
vifstep(X)
```

```
## 10 variables from the 45 input variables have collinearity problem:
##
## avg_risk_score_all pct_fs_risk_2050_500 pct_fs_risk_2020_500 avg_risk_fsf_2020_500 pct_fs_risk_2050_5
##
## After excluding the collinear variables, the linear correlation coefficients ranges between:
## min correlation ( EP_AGE17 ~ EP_UNEMP ):  0.000531153
## max correlation ( pct_floodfactor9 ~ pct_fs_risk_2020_5 ):  0.7815086
##
## ---------- VIFs of the remained variables --------
##                 Variables      VIF
## 1      pct_fs_risk_2020_5 6.244765
## 2        pct_floodfactor2 1.802137
## 3        pct_floodfactor3 1.855638
## 4        pct_floodfactor4 1.824333
## 5        pct_floodfactor5 3.075937
## 6        pct_floodfactor6 3.722657
## 7        pct_floodfactor7 2.581890
## 8        pct_floodfactor8 2.683663
## 9        pct_floodfactor9 4.562798
## 10  avg_risk_fsf_2020_100 5.308554
## 11    avg_risk_score_sfha 2.314375
## 12 avg_risk_score_no_sfha 5.796045
## 13                EP_POV 4.320384
```

```
## 14               EP_UNEMP 2.294512
## 15                 EP_PCI 4.489667
## 16              EP_NOHSDP 5.377951
## 17               EP_AGE65 3.742109
## 18               EP_AGE17 4.319653
## 19              EP_DISABL 3.320381
## 20              EP_SNGPNT 2.452418
## 21              EP_MINRTY 4.191313
## 22              EP_LIMENG 3.813787
## 23               EP_MUNIT 3.359209
## 24              EP_MOBILE 2.858549
## 25               EP_CROWD 2.433964
## 26               EP_NOVEH 2.470080
## 27              EP_GROUPQ 2.117524
## 28             EP_UNINSUR 2.277158
## 29                     co 2.245493
## 30                    no2 5.129353
## 31                     o3 1.902229
## 32                   pm10 2.303467
## 33                   pm25 2.437553
## 34                    so2 1.347527
## 35             total_mean 3.497329
```

This procedure detects that the following variables have collinearity problems. Let's exclude these variables and then rerun the analysis.

```
collin_var_names <- c("avg_risk_score_all", "pct_fs_risk_2050_500", "pct_fs_risk_2020_500", "avg_risk_fs
```