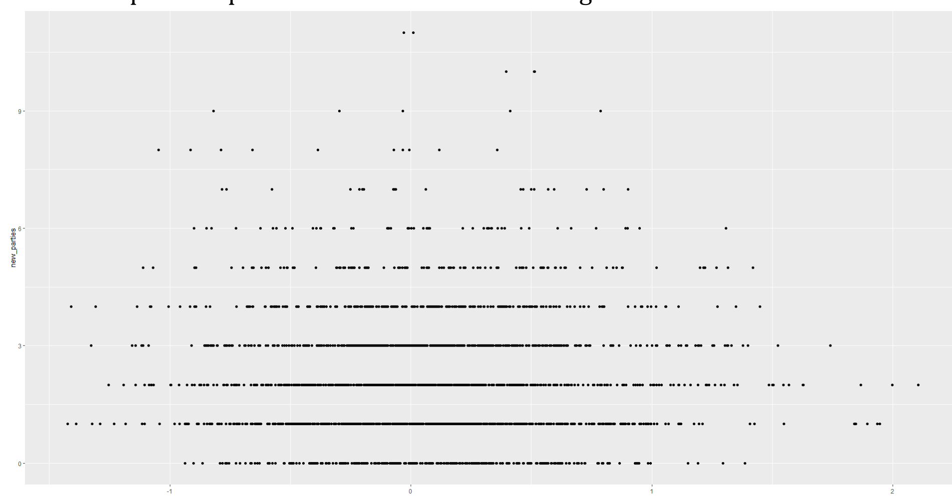
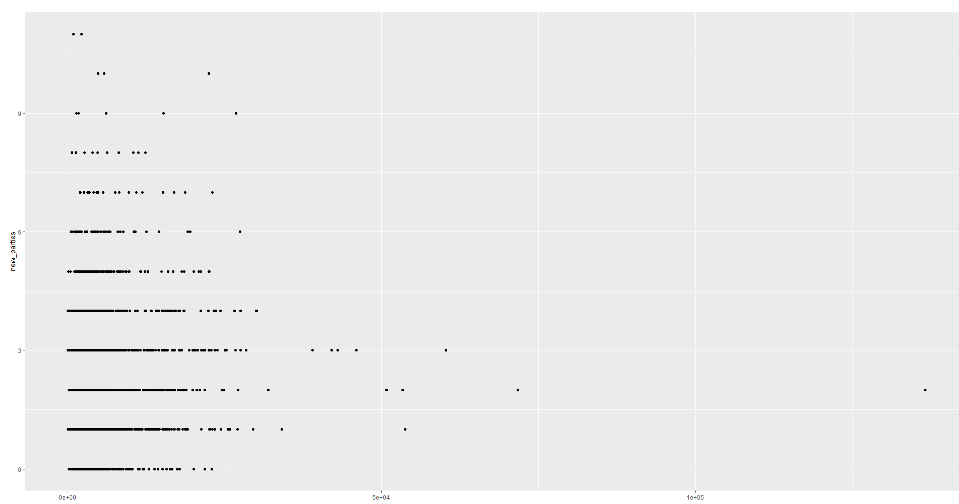
Please check the R script for detailed comment about the coding.

1. First, we will set up the relationship between rainfall and political party foundings, and then modify the rainfall measure to generate a statistically independent measure for droughts. This modification will allow us to isolate the effect of economic strain on political parties from other underlying features of a region that might influence its political structure.
    a. Create a figure, for example, a scatter plot, showing the visual relationship between the level of rainfall in a district in the period leading up to the current election, and the number of political parties that are founded in a region.





Insight:

Based on the scatter plot showing relationship b/w raw rainfall level and number of political parties, we can see when the rain level was around 0, many new parties were founded. When the rainfall level increased, the number of political parties founded decreased. So that we could conclude the doroughts condition is more likely to motivate people to found a new party than flood is.

    b. Using the election-period level rainfall measures created above, show that the raw level of rainfall, as well as the Standardized Precipitation Index, are not independent from one election period to the next within a district, as well as from neighboring districts from one election period to the next.

1.raw rainfall level

```
Twoways effects Within Model

Call:
plm(formula = rainfall ~ border_rainfall + year_betw + lagged_sum,
    data = combine, effect = "twoways", model = "within",
    index = "district")

Unbalanced Panel: n = 189, T = 1-12, N = 1959

Residuals:
     Min.   1st Qu.    Median   3rd Qu.      Max.
-7434.734  -552.905   -19.385   607.802  8360.059

Coefficients:
                  Estimate Std. Error t-value  Pr(>|t|)
border_rainfall 6.1338e-02 1.2565e-02  4.8817 1.147e-06 ***
year_betw       1.2440e+03 2.7255e+01 45.6416 < 2.2e-16 ***
lagged_sum      1.1443e-01 1.1674e-02  9.8022 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Total Sum of Squares:    8.11e+09
Residual Sum of Squares: 3693800000
R-Squared:      0.54453
Adj. R-Squared: 0.49185
F-statistic: 699.394 on 3 and 1755 DF, p-value: < 2.22e-16
>
```

2.spi

```
Twoways effects Within Model

Call:
plm(formula = spi ~ border_spi + year_betw + lagged_avg, data = combine,
    effect = "twoways", model = "within", index = "district")

Unbalanced Panel: n = 189, T = 1-12, N = 1959

Residuals:
     Min.   1st Qu.    Median   3rd Qu.      Max.
-1.527994 -0.325482 -0.010484  0.285332  1.819901

Coefficients:
             Estimate Std. Error t-value  Pr(>|t|)
border_spi -0.2706836  0.0562332 -4.8136 1.610e-06 ***
year_betw  -0.0155659  0.0093686 -1.6615   0.09679 .
lagged_avg  0.3591416  0.0495788  7.2439 6.501e-13 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Total Sum of Squares:    464.67
Residual Sum of Squares: 448.63
R-Squared:      0.034519
Adj. R-Squared: -0.077158
F-statistic: 20.9153 on 3 and 1755 DF, p-value: 2.5874e-13
>
```

Insight:

From the regression result, we can see both lagged value of the district and lagged value of its neighbours have small p-value in the models. So that the raw level of rainfall, as well as the Standardized Precipitation Index, are not independent from one election period to the next within a district, as well as from neighboring districts from one election period to the next.

c. Meteorological scientists consider moderate droughts to occur if the Standardized Precipitation Index falls below -1, and moderate floods to occur if it rises above 1. Create a measure that sums the number of years a district experiences either moderate droughts or floods during the interval starting from the year following the previous election up until the year of the current election. Perform the same test as in (B), using this new transformed measure. This measure will form the basis for the predictors used in the remainder of the regressions in Questions 2-5.

```
--------------------------------------------
Maximum Likelihood estimation
Newton-Raphson maximisation, 3 iterations
Return code 1: gradient close to zero
Log-Likelihood: -1882.122
3  free parameters
Estimates:
                  Estimate Std. error t value Pr(> t)
neighbour_extreme -0.039496   0.031262  -1.263   0.206
year_betw          0.243869   0.012132  20.101  <2e-16 ***
lagged_extreme    -0.001617   0.025939  -0.062   0.950
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
--------------------------------------------
```

Insight:

Difference: In QB we can see both lagged SPI and raw rainfall level are good predictors of current rainfall level. However, the lagged counts of droughts and floods of both the district itself and its neighbours are not good predictors of the weather condition of the district as the p-value of two predictors are very large. The reason may be that normal rainfall is dependent on the weather condition of the previous period, but extreme weathers like droughts and floods are more seasonal, which is not very related with the weather in the previous ime period.

2. Next, let's analyze whether there are more new political parties when droughts or floods occur. Run a regression predicting the number of new political parties that are formed as a function of the number of years a district experiences droughts or flooding in the interval starting from the year following the previous election up until the year of the current election.

a.

```
Panel Regression with AR(1) Prais-Winsten correction and panel heteroskedasticity-robust standard errors

Unbalanced Panel Design:
 Total obs.:        2502  Avg obs. per panel 11.0708
 Number of panels: 226   Max obs. per panel 14
 Number of times:  14    Min obs. per panel 1

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)     2.093564   0.037946  55.172   <2e-16 ***
number_extreme -0.016482   0.025015  -0.659   0.5100
year_betw       0.005347   0.002312   2.313   0.0208 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

R-squared: 0.1431
Wald statistic: 5.4373, Pr(>Chisq(2)): 0.066
```

Insight:

From the regression result we can see actually the foundings of new parties is not related with the number of extreme weathers, but dependent on the length of each election period.

b.

In addition to modeling the effect of extreme weather on the overall entry of new parties, do certain kinds of political parties seem to be more likely than other kinds to be formed when a district experiences extreme weather?

```
> reg_result
                     party        slope     p-value         r2
4      new_parties_secular -0.014605099 0.005750946 0.04846941
7    new_parties_religious  0.006552405 0.047460767 0.07829206
5  new_parties_nationalist -0.007732959 0.216641216 0.03508278
11     new_parties_farming -0.003647817 0.254240188 0.05868551
1        new_parties_caste -0.008706597 0.281558756 0.04436857
6      new_parties_liberal -0.016667242 0.286782201 0.05861551
10   new_parties_farright  0.004342296 0.304806927 0.04398934
9     new_parties_farleft -0.007706819 0.537915550 0.06299937
2    new_parties_socialist -0.005716679 0.549213973 0.05634759
8       new_parties_ethnic  0.002231855 0.682743013 0.03693184
3    new_parties_communist -0.002177684 0.741605876 0.03614548
> 
```

Insight: We can see that there are two parties having statistically significant relationship with extreme weathers, in which the founding of religious party is positively influenced by extreme weather and the founding of secular party is negatively influenced by extreme weathers.

3. Now that we have established the baseline effect, we can look at how political activity is stimulated by droughts or floods in one district might affect political activity in another district. Use a similar regression to Question 2 to show that, even when taking into account a district's own droughts and floods, the level of entry of new political parties in a district will also depend on the number of years its neighboring districts experience years of droughts or flooding in the interval starting from the year following two elections ago, up until the year of the previous election—the election lead-up interval before the current one. Similar to Question 2, include a control in the regression for the number of years in the current election period, a control for the time-invariant features of a district as fixed effects, and a linear control for each election year.

```
Panel Regression with AR(1) Prais-Winsten correction and panel heteroskedasticity-robust standard errors

Unbalanced Panel Design:
 Total obs.:        1959 Avg obs. per panel 10.3651
 Number of panels: 189  Max obs. per panel 12
 Number of times:  13   Min obs. per panel 1

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)    11.496441   5.011215   2.294   0.0219 *
extreme         0.015795   0.035638   0.443   0.6577
border_extreme -0.050377   0.031630  -1.593   0.1114
year           -0.004612   0.002494  -1.849   0.0646 .
year_betw      -0.053179   0.025475  -2.087   0.0370 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

R-squared: 0.3547
Wald statistic: 6.4835, Pr(>Chisq(4)): 0.1658
> 
```

Insight:

The entry of new parties in a district is either not dependent on its own droughts and flooding or the neighbors' droughts and flooding.

4. Extreme weather events like droughts or floods can erode the stability of political systems and wear away at the entrenched power bases of large, national-scale parties that have difficulty responding to the needs of affected regions.

a. Does experiencing droughts or floods relate differently to the entry and diffusion of political parties depending on their scope

```
> summary(sub_national)

Panel Regression with AR(1) Prais-Winsten correction and panel heteroskedasticity-robust standard errors

Unbalanced Panel Design:
 Total obs.:       1959 Avg obs. per panel 10.3651
 Number of panels: 189  Max obs. per panel 12
 Number of times:  13   Min obs. per panel 1

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)    19.640255   3.509092   5.597 2.49e-08 ***
extreme        -0.005574   0.025706  -0.217 0.828367
border_extreme -0.102783   0.019453  -5.284 1.41e-07 ***
year           -0.009285   0.001757  -5.286 1.39e-07 ***
year_betw      -0.062555   0.016408  -3.813 0.000142 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

R-squared: 0.6714
Wald statistic: 48.351, Pr(>Chisq(4)): 0
```

```
> summary(national)

Panel Regression with AR(1) Prais-Winsten correction and panel heteroskedasticity-robust standard errors

Unbalanced Panel Design:
 Total obs.:       1959 Avg obs. per panel 10.3651
 Number of panels: 189  Max obs. per panel 12
 Number of times:  13   Min obs. per panel 1

Coefficients:
                 Estimate Std. Error t value Pr(>|t|)
(Intercept)     0.8512110  4.4946874   0.189    0.850
extreme         0.0049417  0.0279866   0.177    0.860
border_extreme  0.0353236  0.0254602   1.387    0.165
year            0.0002377  0.0022420   0.106    0.916
year_betw      -0.0213679  0.0192583  -1.110    0.267

R-squared: 0.1066
Wald statistic: 4.099, Pr(>Chisq(4)): 0.3928
```

```
> summary(regional)

Panel Regression with AR(1) Prais-Winsten correction and panel heteroskedasticity-robust standard errors

Unbalanced Panel Design:
 Total obs.:       1959 Avg obs. per panel 10.3651
 Number of panels: 189  Max obs. per panel 12
 Number of times:  13   Min obs. per panel 1

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)    11.590810   3.405421   3.404 0.000678 ***
extreme         0.001634   0.017596   0.093 0.926028
border_extreme -0.046945   0.016330  -2.875 0.004087 **
year           -0.005522   0.001703  -3.242 0.001206 **
year_betw      -0.033570   0.011922  -2.816 0.004915 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

R-squared: 0.0491
Wald statistic: 17.1746, Pr(>Chisq(4)): 0.0018
```

```
> summary(state)

Panel Regression with AR(1) Prais-Winsten correction and panel heteroskedasticity-robust standard errors

Unbalanced Panel Design:
 Total obs.:       1959 Avg obs. per panel 10.3651
 Number of panels: 189  Max obs. per panel 12
 Number of times:  13   Min obs. per panel 1

Coefficients:
                 Estimate Std. Error t value Pr(>|t|)
(Intercept)     4.5469684  1.9405175   2.343 0.019220 *
extreme        -0.0001820  0.0096739  -0.019 0.984992
border_extreme -0.0297884  0.0083448  -3.570 0.000366 ***
year           -0.0021632  0.0009729  -2.224 0.026296 *
year_betw      -0.0038907  0.0062955  -0.618 0.536634
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

R-squared: 0.0565
Wald statistic: 15.5773, Pr(>Chisq(4)): 0.0036
>
```

Insight:

We can see that at different scope the extreme weather has different influence on the entry and diffusion of new parties. Extreme weather at regional and sub-national has very strong influence on founding of new parties. Extreme weather at state level's influence is slightly weaker than that at regional level. At the national level, extreme weathers have very little influence on the entry and diffusion of new parties.

b. Does experiencing droughts or floods relate to political concentration?

```
Panel Regression with AR(1) Prais-Winsten correction and panel heteroskedasticity-robust standard errors

Unbalanced Panel Design:
 Total obs.:       1953 Avg obs. per panel 10.3333
 Number of panels: 189  Max obs. per panel 12
 Number of times:  13   Min obs. per panel 1

Coefficients:
                 Estimate Std. Error t value Pr(>|t|)
(Intercept)     89477.766   4508.614  19.846  < 2e-16 ***
extreme           -51.608     20.471  -2.521  0.01178 *
border_extreme    -78.766     17.415  -4.523 6.47e-06 ***
year              -44.300      2.263 -19.576  < 2e-16 ***
year_betw          33.012     12.428   2.656  0.00796 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

R-squared: 0.4845
Wald statistic: 496.3969, Pr(>Chisq(4)): 0
```

Insight:

Yes. As counts of extreme weather in the district and lagged counts of extreme weather in the neighbours both have very samll p-value, it does relate to political concentration. Based on the regression result, count of extreme weather is negatively related to the degree of concentration of political power at a district, especially the lagged count of extreme weather of neighbours. So the more the extreme weathers are, the less concentrated the political power is at the district level.

5. Political parties are formed to accomplish a variety of goals. Individual parties can also exist in the context of larger social and cultural trends, especially when regions influence each other as political organizing activity diffuses across regions over time. To understand the diffusion process more, we want to analyze whether the new parties that appear in a district are the same parties that have appeared in neighboring districts in the past, or if it the process of political organization, rather than the content of a specific political party, that is diffusing

```
Panel Regression with AR(1) Prais-Winsten correction and panel heteroskedasticity-robust standard errors

Unbalanced Panel Design:
 Total obs.:       1959 Avg obs. per panel 9.0694
 Number of panels: 216  Max obs. per panel 12
 Number of times:  14   Min obs. per panel 1

Coefficients:
                              Estimate Std. Error t value Pr(>|t|)
(Intercept)                 -26.201709   4.380256  -5.982 2.62e-09 ***
number_extreme               -0.076129   0.020276  -3.755 0.000179 ***
regress_lag2$border_extreme   0.001010   0.023828   0.042 0.966201
year                          0.013981   0.002208   6.333 2.98e-10 ***
year_betw                     0.023398   0.002021  11.577  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

R-squared: 0.349
Wald statistic: 138.2612, Pr(>Chisq(4)): 0
```

```
Panel Regression with AR(1) Prais-Winsten correction and panel heteroskedasticity-robust standard errors

Unbalanced Panel Design:
 Total obs.:       1959 Avg obs. per panel 9.0694
 Number of panels: 216  Max obs. per panel 12
 Number of times:  14   Min obs. per panel 1

Coefficients:
                              Estimate Std. Error t value Pr(>|t|)
(Intercept)                 -25.898242   3.729187  -6.945 5.14e-12 ***
number_extreme                0.016707   0.017303   0.966    0.334
regress_lag2$border_extreme   0.008446   0.019540   0.432    0.666
year                          0.013302   0.001881   7.072 2.12e-12 ***
year_betw                     0.022071   0.001806  12.220  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

R-squared: 0.1517
Wald statistic: 167.3681, Pr(>Chisq(4)): 0
```

Insight:

Based on the regression analysis result, we can see as the number of extreme weathers in a district is very related to the number of parties founded that have contested in neighbouring districts. The negative coefficient suggests that as the number of extreme weathers increases in a district, new parties that have contested an election in the neighbouring district are less likely to be founded in the district. The number of parties founded that have not contested an election in the neighbouring districts will increase if the number of extreme weathers increases in the district as the coefficient of regression model 2 is positive. But this relationship is very weak as the p-value is big.