

# sara-experimental

Sara Shao

10/7/2021

```
library(tidyverse)

## -- Attaching packages ----- tidyverse 1.3.0 --
## v tibble 3.0.6      v purrr 0.3.4
## v tidyr 1.1.2      v dplyr 1.0.4
## v readr 1.4.0      v forcats 0.5.1

## -- Conflicts ----- tidyverse_conflicts() --
## x lubridate::as.difftime() masks base::as.difftime()
## x lubridate::date() masks base::date()
## x dplyr::filter() masks stats::filter()
## x readr::guess_encoding() masks rvest::guess_encoding()
## x lubridate::intersect() masks base::intersect()
## x dplyr::lag() masks stats::lag()
## x purrr::pluck() masks rvest::pluck()
## x lubridate::setdiff() masks base::setdiff()
## x lubridate::union() masks base::union()

train <- read.csv('data-train.csv')

train

##      St Re   Fr R_moment_1 R_moment_2 R_moment_3 R_moment_4
## 1 0.10 224 0.052 0.00215700 1.3035e-01 1.4374e+01 1.5865e+03
## 2 3.00 224 0.052 0.00379030 4.7042e-01 6.9940e+01 1.0404e+04
## 3 0.70 224   Inf 0.00290540 4.3499e-02 8.2200e-01 1.5551e+01
## 4 0.05 90   Inf 0.06352800 9.0653e-02 4.6746e-01 3.2696e+00
## 5 0.70 398   Inf 0.00036945 6.2242e-03 1.2649e-01 2.5714e+00
## 6 2.00 90 0.300 0.14780000 2.0068e+00 3.6249e+01 6.7167e+02
## 7 0.20 90   Inf 0.08127300 3.2450e-01 3.0363e+00 3.2976e+01
## 8 3.00 224   Inf 0.00574730 1.1966e-01 2.7480e+00 6.3159e+01
## 9 0.90 224   Inf 0.00302150 4.5244e-02 8.4530e-01 1.5809e+01
## 10 0.60 398 0.052 0.00031431 4.4672e-03 8.2060e-02 1.5077e+00
## 11 0.90 90   Inf 0.09102700 5.9539e-01 7.2454e+00 9.5166e+01
## 12 0.30 398   Inf 0.00036022 6.2830e-03 1.3546e-01 2.9211e+00
## 13 2.00 224   Inf 0.00447250 8.0804e-02 1.6668e+00 3.4408e+01
## 14 1.00 224 0.052 0.00312380 3.6478e-01 5.3322e+01 7.7958e+03
## 15 0.50 90 0.052 0.12670000 6.8596e+02 5.4300e+06 4.2900e+10
## 16 0.60 224 0.300 0.00257400 3.6621e-02 6.7102e-01 1.2309e+01
## 17 0.10 90   Inf 0.07722700 2.2120e-01 1.8833e+00 2.0190e+01
## 18 1.00 90 0.300 0.11236000 1.1261e+00 1.7335e+01 2.8261e+02
## 19 0.70 224 0.052 0.00285610 3.1273e-01 4.4529e+01 6.3423e+03
## 20 0.20 90 0.052 0.11760000 5.1774e+02 3.8100e+06 2.8000e+10
## 21 0.10 90 0.300 0.06125200 6.9867e-02 2.4338e-01 1.1379e+00
```

```

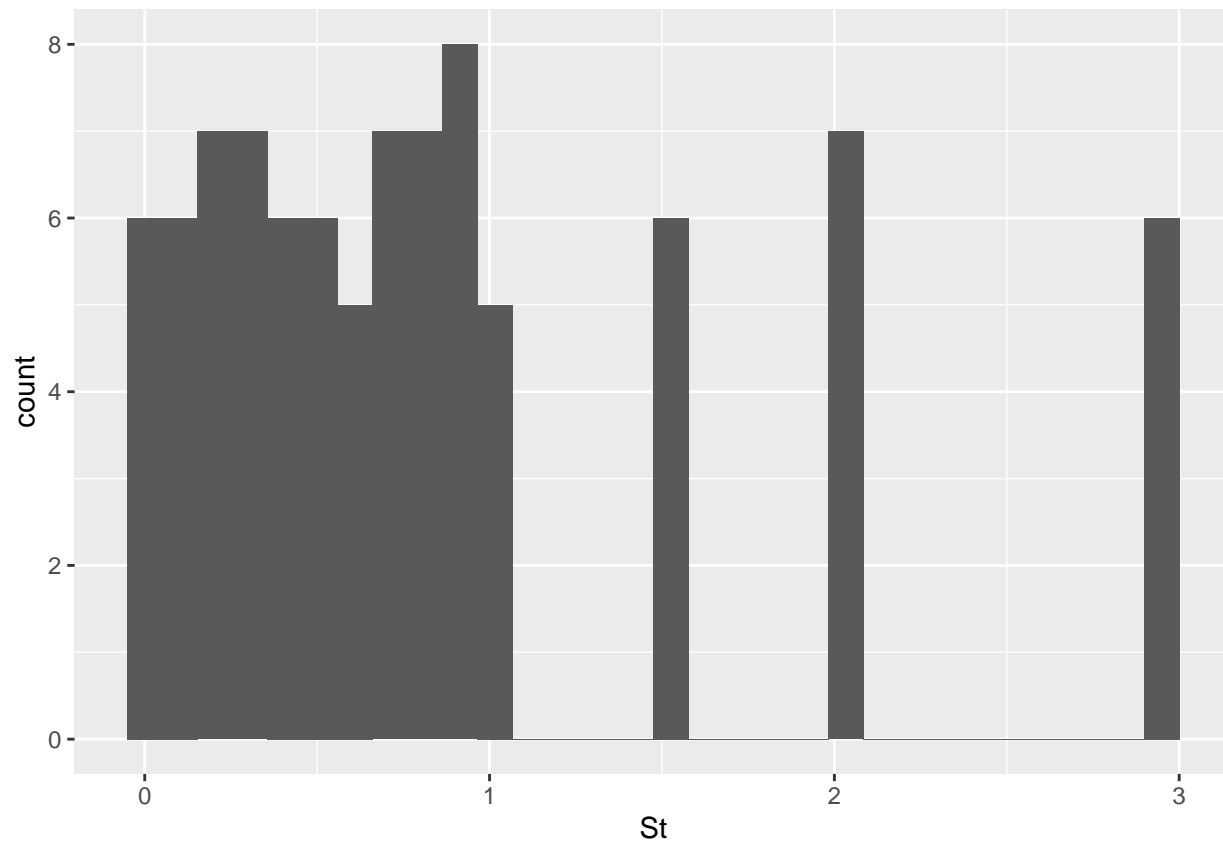
## 22 0.50 398   Inf 0.00036800 6.3559e-03 1.3341e-01 2.8013e+00
## 23 0.20 224   Inf 0.00269160 3.9016e-02 7.6384e-01 1.4978e+01
## 24 1.50 398 0.052 0.00038321 5.9338e-03 1.1156e-01 2.1004e+00
## 25 0.90 398   Inf 0.00038344 6.4432e-03 1.2925e-01 2.5935e+00
## 26 0.50 224 0.052 0.00274240 3.0355e-01 4.3911e+01 6.3530e+03
## 27 0.10 398 0.052 0.00027479 3.2549e-03 5.8006e-02 1.0344e+00
## 28 0.40 224 0.052 0.00268090 2.8897e-01 4.1585e+01 5.9861e+03
## 29 0.30  90 0.052 0.12261000 6.2727e+02 4.9100e+06 3.8500e+10
## 30 0.05 224 0.052 0.00173740 1.6633e-03 2.0228e-02 3.6438e-01
## 31 1.50 224 0.052 0.00341630 4.0300e-01 5.8417e+01 8.4710e+03
## 32 0.80  90   Inf 0.09107400 6.1825e-01 7.4973e+00 9.7048e+01
## 33 1.50  90 0.052 0.15181000 9.9690e+02 8.5500e+06 7.3300e+10
## 34 0.05 398   Inf 0.00022202 1.0055e-03 1.0857e-02 1.1782e-01
## 35 0.80 224   Inf 0.00298090 4.4580e-02 8.3764e-01 1.5759e+01
## 36 0.90 224 0.300 0.00280490 4.1143e-02 7.5132e-01 1.3729e+01
## 37 0.40 224   Inf 0.00292630 4.6261e-02 9.2914e-01 1.8681e+01
## 38 0.80 398 0.052 0.00033341 4.9036e-03 9.1143e-02 1.6948e+00
## 39 1.50 224 0.300 0.00341050 5.4101e-02 1.0222e+00 1.9340e+01
## 40 0.20 224 0.052 0.00257870 2.6830e-01 3.9080e+01 5.6959e+03
## 41 0.30 224   Inf 0.00283770 4.3589e-02 8.6962e-01 1.7373e+01
## 42 0.30 224 0.300 0.00250630 3.5881e-02 6.8596e-01 1.3132e+01
## 43 2.00 224 0.300 0.00381230 6.1927e-02 1.1844e+00 2.2705e+01
## 44 1.00  90   Inf 0.09691800 6.7696e-01 8.2384e+00 1.0602e+02
## 45 0.80 224 0.052 0.00295750 3.3361e-01 4.8161e+01 6.9539e+03
## 46 1.00 224 0.300 0.00289530 4.2300e-02 7.6755e-01 1.3941e+01
## 47 0.70  90 0.300 0.09471100 6.9751e-01 9.1793e+00 1.3187e+02
## 48 0.30 224 0.052 0.00256750 2.6547e-01 3.7665e+01 5.3451e+03
## 49 0.40 224 0.300 0.00262070 3.9502e-02 7.6851e-01 1.4966e+01
## 50 0.10 224 0.300 0.00221530 2.4475e-02 4.2167e-01 7.2842e+00
## 51 3.00  90   Inf 0.17234000 2.2386e+00 4.0454e+01 7.6198e+02
## 52 1.00 224   Inf 0.00309680 4.6454e-02 8.6381e-01 1.6077e+01
## 53 2.00 398   Inf 0.00053647 1.0022e-02 2.1023e-01 4.4109e+00
## 54 0.80  90 0.052 0.13793000 8.2524e+02 6.8000e+06 5.6100e+10
## 55 0.40 398 0.052 0.00029691 4.1375e-03 7.6124e-02 1.4014e+00
## 56 0.50 398 0.052 0.00030716 4.3494e-03 8.0143e-02 1.4770e+00
## 57 0.70  90   Inf 0.09217600 5.6482e-01 6.7191e+00 8.8723e+01
## 58 2.00  90 0.052 0.15433000 1.0269e+03 8.8700e+06 7.6700e+10
## 59 0.90  90 0.300 0.10962000 1.0319e+00 1.5797e+01 2.6136e+02
## 60 0.30  90   Inf 0.07694500 3.2652e-01 3.4052e+00 4.1042e+01
## 61 0.50 224 0.300 0.00250710 3.5152e-02 6.4378e-01 1.1801e+01
## 62 0.50  90 0.300 0.08477300 4.9728e-01 6.0317e+00 8.3287e+01
## 63 0.80 398   Inf 0.00037399 6.2457e-03 1.2542e-01 2.5193e+00
## 64 0.20 398   Inf 0.00033521 5.4505e-03 1.1408e-01 2.3884e+00
## 65 0.70  90 0.052 0.13173000 7.3694e+02 5.8700e+06 4.6700e+10
## 66 2.00 398 0.052 0.00039644 6.1040e-03 1.1209e-01 2.0593e+00
## 67 0.70 224 0.300 0.00260870 3.6438e-02 6.5445e-01 1.1765e+01
## 68 0.60 224 0.052 0.00279390 3.0594e-01 4.3745e+01 6.2554e+03
## 69 0.30 398 0.052 0.00030066 4.3488e-03 8.3446e-02 1.6023e+00
## 70 0.90 224 0.052 0.00305410 3.5419e-01 5.1795e+01 7.5758e+03
## 71 0.80 224 0.300 0.00268160 3.7714e-02 6.7549e-01 1.2112e+01
## 72 0.20 224 0.300 0.00246950 3.4818e-02 6.7088e-01 1.2939e+01
## 73 3.00 398 0.052 0.00040188 5.4492e-03 9.1871e-02 1.5565e+00
## 74 0.90  90 0.052 0.14184000 8.7019e+02 7.2500e+06 6.0400e+10
## 75 0.40 398   Inf 0.00036977 6.4986e-03 1.3933e-01 2.9880e+00

```

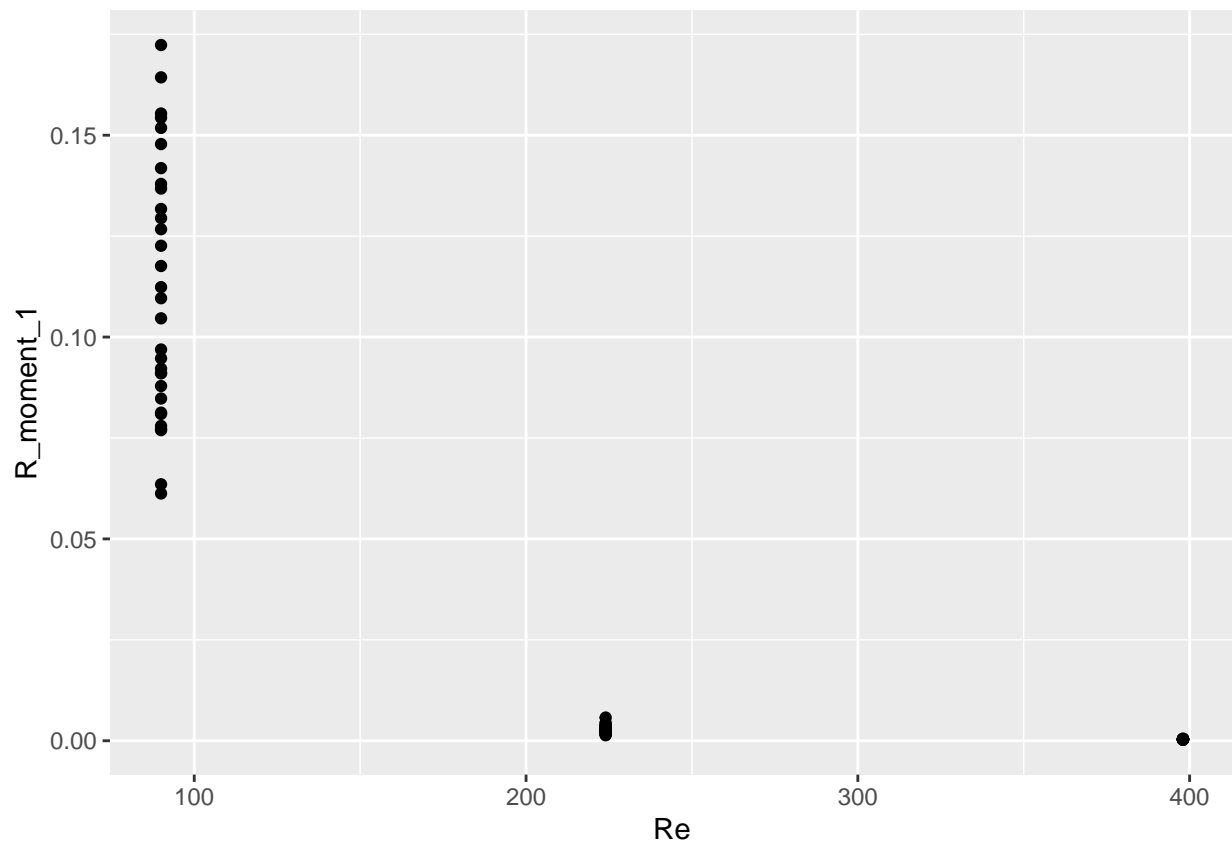
```
## 76 0.20 90 0.300 0.07798500 2.5598e-01 2.0965e+00 2.0849e+01
## 77 0.60 90 0.052 0.12946000 7.1816e+02 5.7200e+06 4.5600e+10
## 78 1.50 90 0.300 0.13678000 1.8254e+00 3.2833e+01 6.0903e+02
## 79 0.10 90 0.052 0.10464000 1.6015e+02 6.9900e+05 3.0700e+09
## 80 3.00 90 0.052 0.15538000 1.0443e+03 9.1400e+06 8.0000e+10
## 81 2.00 224 0.052 0.00363470 4.4512e-01 6.5387e+01 9.6105e+03
## 82 0.05 90 0.052 0.08786800 5.3449e-01 2.2205e+01 1.5679e+03
## 83 0.40 90 0.300 0.08095700 3.9996e-01 4.3303e+00 5.3618e+01
## 84 0.90 398 0.052 0.00034145 5.0555e-03 9.4083e-02 1.7522e+00
## 85 3.00 90 0.300 0.16433000 2.3317e+00 4.4516e+01 8.8779e+02
## 86 0.05 224 Inf 0.00153380 2.5653e-04 3.0407e-04 5.4466e-04
## 87 0.05 224 0.300 0.00135380 1.0303e-04 5.1400e-05 4.1600e-05
## 88 0.60 224 Inf 0.00291710 4.4317e-02 8.5282e-01 1.6431e+01
## 89 1.50 224 Inf 0.00370310 6.0910e-02 1.1829e+00 2.2990e+01
```

```
ggplot(data = train, mapping = aes(x = St)) +
  geom_histogram()
```

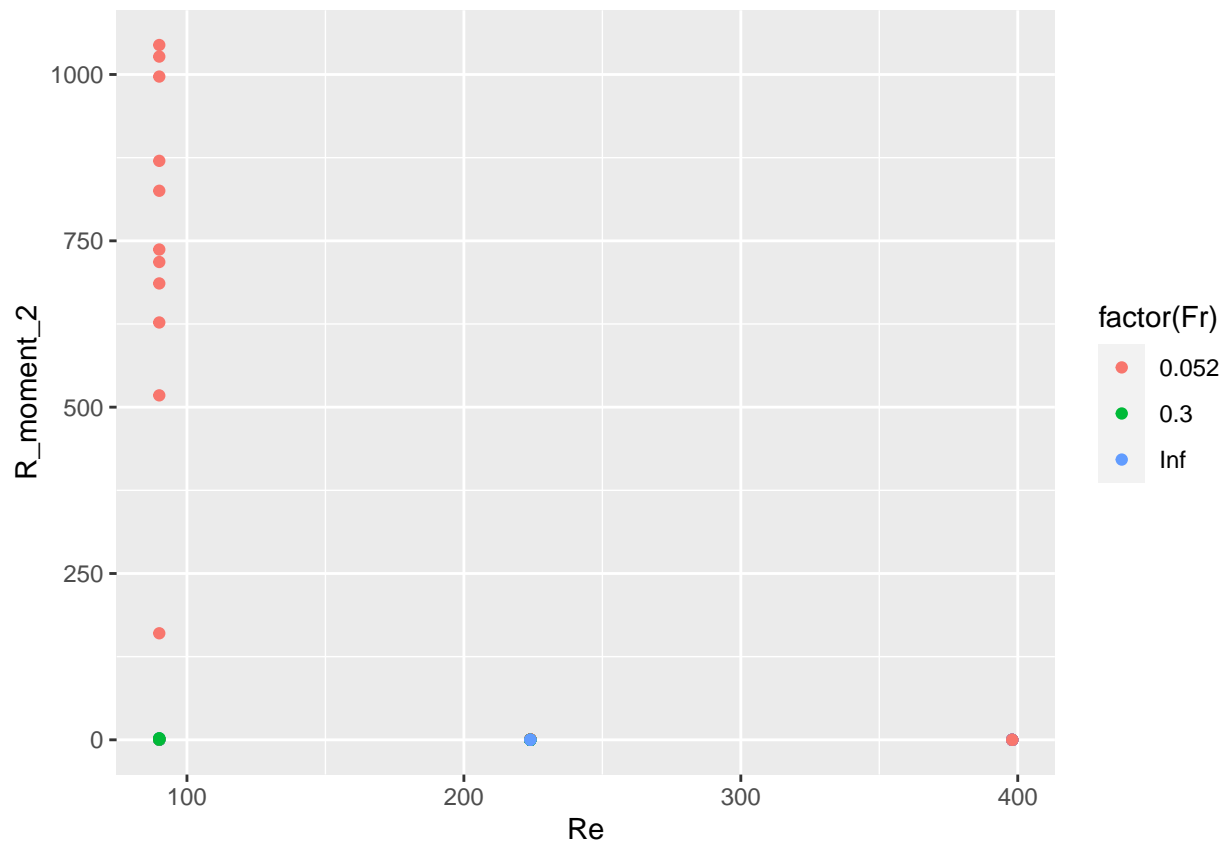
```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



```
ggplot(data = train, mapping = aes(x = Re, y = R_moment_1)) +
  geom_point()
```

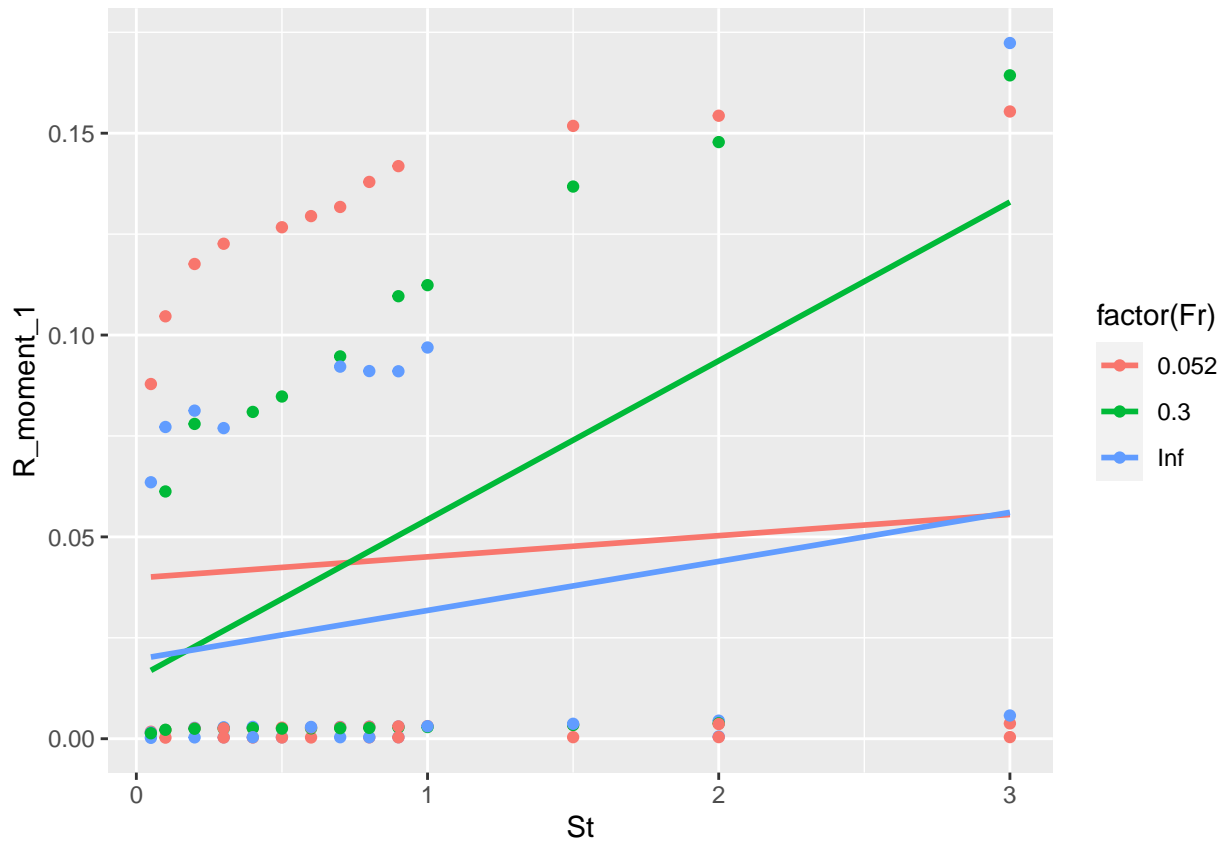


```
ggplot(data = train, mapping = aes(x = Re, y = R_moment_2, color = factor(Fr))) +  
  geom_point()
```



```
ggplot(data = train, mapping = aes(x = St, y = R_moment_1, color = factor(Fr))) +  
  geom_point() +  
  geom_smooth(method = lm, se = F)
```

```
## `geom_smooth()` using formula 'y ~ x'
```



```
train_data <- train %>%
  mutate(Fr = as.factor(Fr)) %>%
  mutate(Re = as.factor(Re))

lm1 <- lm(log(R_moment_1) ~ log(St) + Re + Fr + St*Fr + Re*Fr, data = train_data)
summary(lm1)
```

```
##
## Call:
## lm(formula = log(R_moment_1) ~ log(St) + Re + Fr + St * Fr +
##     Re * Fr, data = train_data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.22072 -0.04215 -0.01259  0.04746  0.16243
##
## Coefficients: (1 not defined because of singularities)
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -1.957525   0.035198  -55.614 < 2e-16 ***
## log(St)      0.146804   0.014517   10.112 8.83e-16 ***
## Re224        -3.826548   0.030071 -127.250 < 2e-16 ***
## Re398        -5.998459   0.032830 -182.715 < 2e-16 ***
## Fr0.3        -0.421373   0.042544  -9.904 2.20e-15 ***
## FrInf        -0.434221   0.039192 -11.079 < 2e-16 ***
## St            0.001924   0.022326   0.086  0.932
## Fr0.3:St      0.149371   0.027594   5.413 6.79e-07 ***
## FrInf:St      0.139218   0.024067   5.785 1.48e-07 ***
```

```
## Re224:Fr0.3 0.253548 0.044660 5.677 2.31e-07 ***
## Re398:Fr0.3 NA NA NA NA
## Re224:FrInf 0.375558 0.045292 8.292 2.79e-12 ***
## Re398:FrInf 0.492235 0.048761 10.095 9.53e-16 ***
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0764 on 77 degrees of freedom
## Multiple R-squared: 0.999, Adjusted R-squared: 0.9988
## F-statistic: 6860 on 11 and 77 DF, p-value: < 2.2e-16

lm2 <- lm(log(R_moment_2) ~ log(St) + Re + Fr + St*Fr + Re*Fr, data = train_data)
summary(lm2)
```

```
##
## Call:
## lm(formula = log(R_moment_2) ~ log(St) + Re + Fr + St * Fr +
##     Re * Fr, data = train_data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -3.8715 -0.2898 -0.0746  0.4546  1.8660
##
## Coefficients: (1 not defined because of singularities)
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   7.77314    0.39607  19.626 < 2e-16 ***
## log(St)        1.49507    0.16336   9.152 6.10e-14 ***
## Re224         -7.52845    0.33838 -22.249 < 2e-16 ***
## Re398        -11.57137    0.36942 -31.323 < 2e-16 ***
## Fr0.3         -6.82133    0.47873 -14.249 < 2e-16 ***
## FrInf         -6.58705    0.44101 -14.936 < 2e-16 ***
## St            -0.98457    0.25122  -3.919 0.000191 ***
## Fr0.3:St       0.25047    0.31050   0.807 0.422354
## FrInf:St       0.06622    0.27082   0.245 0.807492
## Re224:Fr0.3    4.51652    0.50254   8.987 1.27e-13 ***
## Re398:Fr0.3    NA         NA         NA      NA
## Re224:FrInf    4.54159    0.50966   8.911 1.78e-13 ***
## Re398:FrInf    6.98196    0.54868  12.725 < 2e-16 ***
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.8597 on 77 degrees of freedom
## Multiple R-squared: 0.9531, Adjusted R-squared: 0.9464
## F-statistic: 142.1 on 11 and 77 DF, p-value: < 2.2e-16
```

```
lm3 <- lm(log(R_moment_3) ~ log(St) + Re + Fr + St*Fr + Re*Fr, data = train_data)
summary(lm3)
```

```
##
## Call:
## lm(formula = log(R_moment_3) ~ log(St) + Re + Fr + St * Fr +
##     Re * Fr, data = train_data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
```

```

## -7.1976 -0.4849 -0.1536 0.6956 3.0521
##
## Coefficients: (1 not defined because of singularities)
##           Estimate Std. Error t value Pr(>|t|)
## (Intercept) 17.419361 0.667233 26.107 < 2e-16 ***
## log(St)      2.349916 0.275193 8.539 9.29e-13 ***
## Re224       -11.303329 0.570040 -19.829 < 2e-16 ***
## Re398       -17.331199 0.622329 -27.849 < 2e-16 ***
## Fr0.3       -13.049707 0.806475 -16.181 < 2e-16 ***
## FrInf       -12.579479 0.742946 -16.932 < 2e-16 ***
## St          -1.634527 0.423216 -3.862 0.000233 ***
## Fr0.3:St     0.323641 0.523081 0.619 0.537927
## FrInf:St    -0.003946 0.456227 -0.009 0.993122
## Re224:Fr0.3 8.560027 0.846596 10.111 8.88e-16 ***
## Re398:Fr0.3      NA      NA      NA      NA
## Re224:FrInf 8.502945 0.858580 9.903 2.21e-15 ***
## Re398:FrInf 13.221542 0.924330 14.304 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.448 on 77 degrees of freedom
## Multiple R-squared: 0.9431, Adjusted R-squared: 0.9349
## F-statistic: 116 on 11 and 77 DF, p-value: < 2.2e-16
lm4 <- lm(log(R_moment_4) ~ log(St) + Re + Fr + St*Fr + Re*Fr, data = train_data)
summary(lm4)

##
## Call:
## lm(formula = log(R_moment_4) ~ log(St) + Re + Fr + St * Fr +
##     Re * Fr, data = train_data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -10.2229  -0.6354  -0.2095   0.8916   4.0676
##
## Coefficients: (1 not defined because of singularities)
##           Estimate Std. Error t value Pr(>|t|)
## (Intercept) 26.95317 0.90843 29.670 < 2e-16 ***
## log(St)      3.09216 0.37467 8.253 3.31e-12 ***
## Re224       -15.08892 0.77611 -19.442 < 2e-16 ***
## Re398       -23.11540 0.84730 -27.281 < 2e-16 ***
## Fr0.3       -19.16880 1.09801 -17.458 < 2e-16 ***
## FrInf       -18.49447 1.01152 -18.284 < 2e-16 ***
## St          -2.18993 0.57621 -3.801 0.000287 ***
## Fr0.3:St     0.36827 0.71217 0.517 0.606566
## FrInf:St    -0.08068 0.62115 -0.130 0.896995
## Re224:Fr0.3 12.52478 1.15263 10.866 < 2e-16 ***
## Re398:Fr0.3      NA      NA      NA      NA
## Re224:FrInf 12.40772 1.16895 10.614 < 2e-16 ***
## Re398:FrInf 19.37138 1.25847 15.393 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.972 on 77 degrees of freedom

```



```
## Multiple R-squared:  0.943, Adjusted R-squared:  0.9348
## F-statistic: 115.7 on 11 and 77 DF,  p-value: < 2.2e-16
```

```
plot(lm1)
```

