First_Year_Project

2023-02-07

Question one

Begin with describing and fitting a full model in which the intercepts and slopes of the extinction times versus numbers of pairs may be different in all four combinations of size and migratory status.
We start by importing the dataset directly from the csv file and saving it to the data variable:

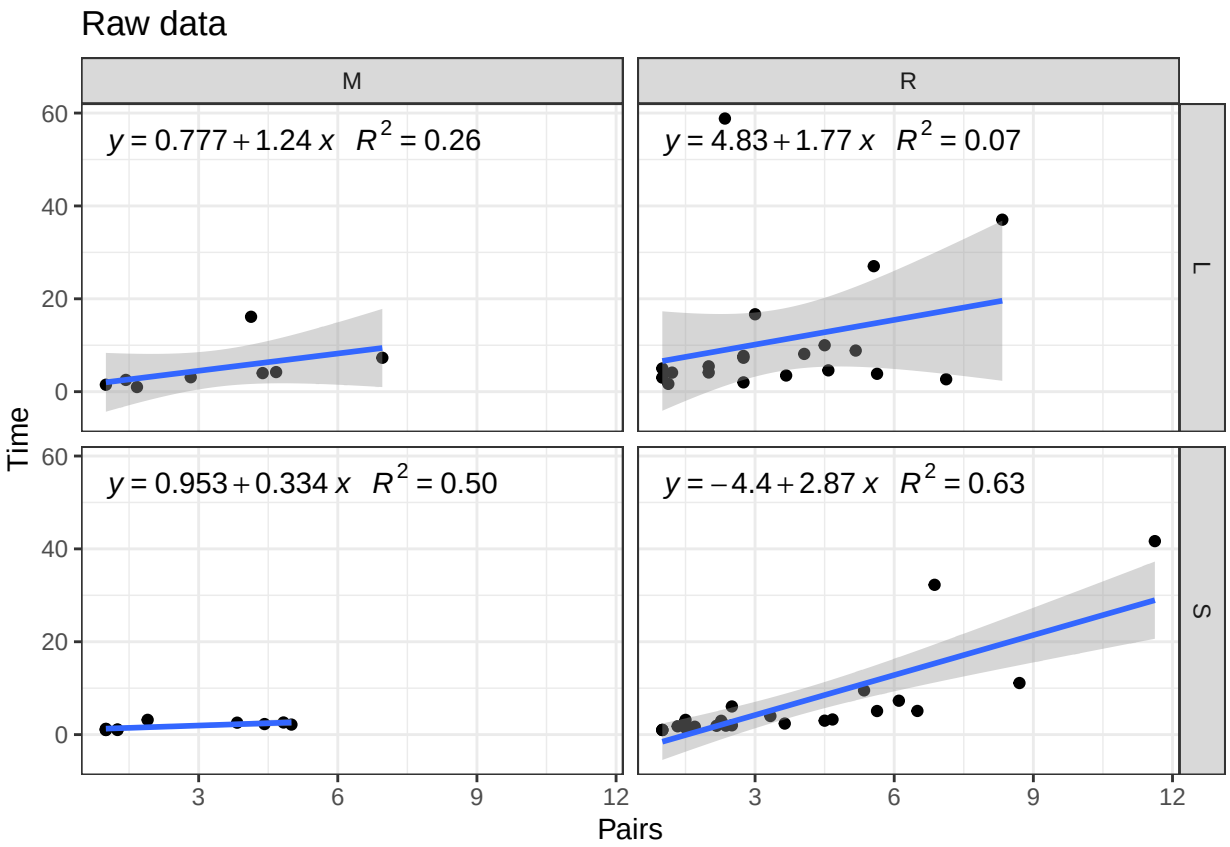
```
head(data <- read.csv('Factors Affecting Extinction.csv', header=T))
```

```
##          Species Time Pairs Size Status
## 1    Sparrowhawk 3.03  1.00    L      R
## 2        Buzzard 5.46  2.00    L      R
## 3        Kestrel 4.10  1.21    L      R
## 4      Peregrine 1.68  1.13    L      R
## 5 Grey_partridge 8.85  5.17    L      R
## 6          Quail 1.49  1.00    L      M
```

(1) Initial Plotting

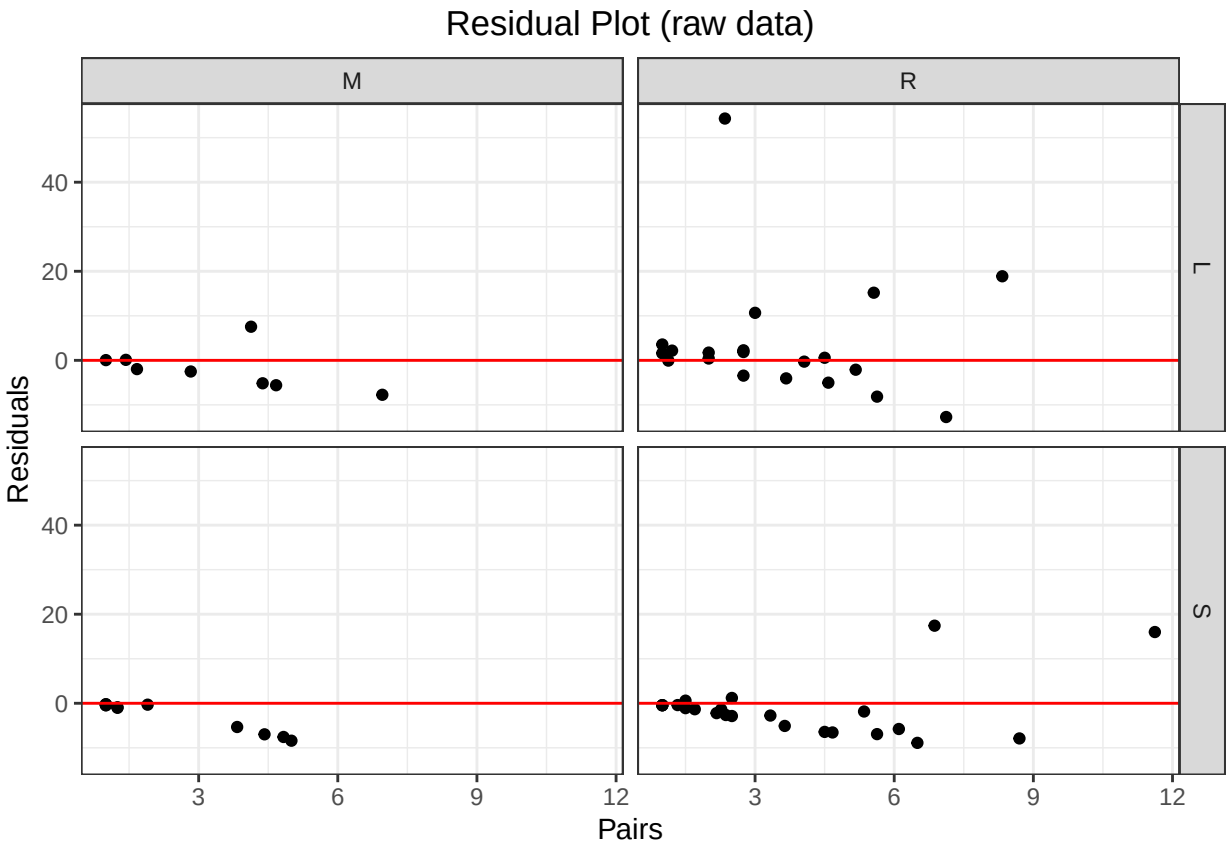
There are four different combinations of Size and Status, LR, LM, SR and SM. If we want to find the correlation between *extinction time* as a function of *pairs*, we can make a regression line with *pairs* as the predictor value and *extinction time* as the predicted value.

```
ggplot(data, aes(x = Pairs, y = Time)) +
  geom_point() +
  facet_grid(Size ~ Status) +
  theme(legend.position = "top") +
  geom_smooth(method = "lm", formula = y ~ x) +

  stat_poly_eq(formula = y ~ x,
  aes(label = paste(after_stat(eq.label), after_stat(rr.label), sep = "~~~")),
  parse = TRUE) +

  labs(title = "Raw data") +
  theme_bw()
```



(2) Residual plot of raw (not transformed) data

```
ggplot(data,
  mapping = aes(x = Pairs,
                y = resid(lm(Time ~ Pairs, data = data)))) +
  geom_point() +
  facet_grid(Size ~ Status) +
  geom_hline(yintercept = 0, color = "red") +
  xlab("Pairs") +
  ylab("Residuals") +
  labs(title = "Residual Plot (raw data)") +
  theme_bw() +
  theme(plot.title = element_text(hjust = 0.5)) +
  theme(legend.position = "bottom")
```

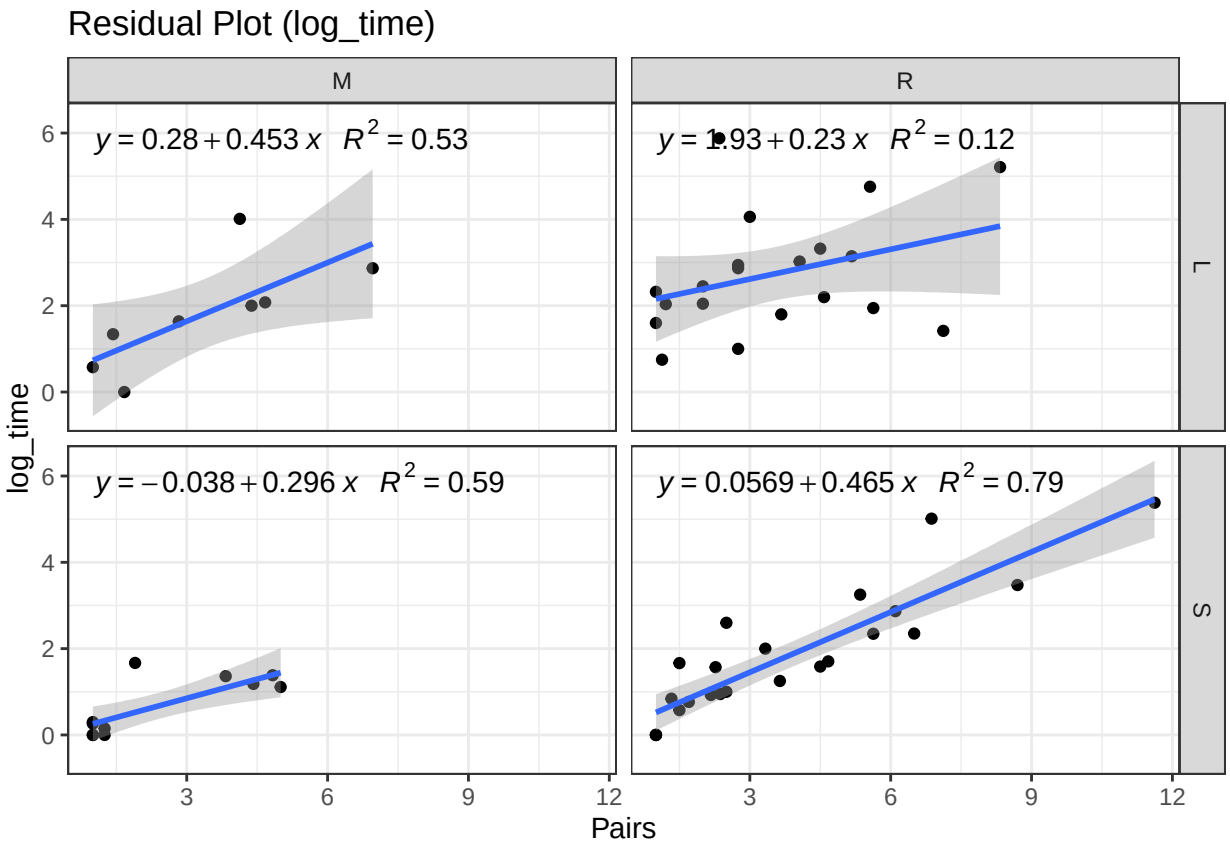


(2) Transformations of time to log_time, sqrt_time, inverse_time

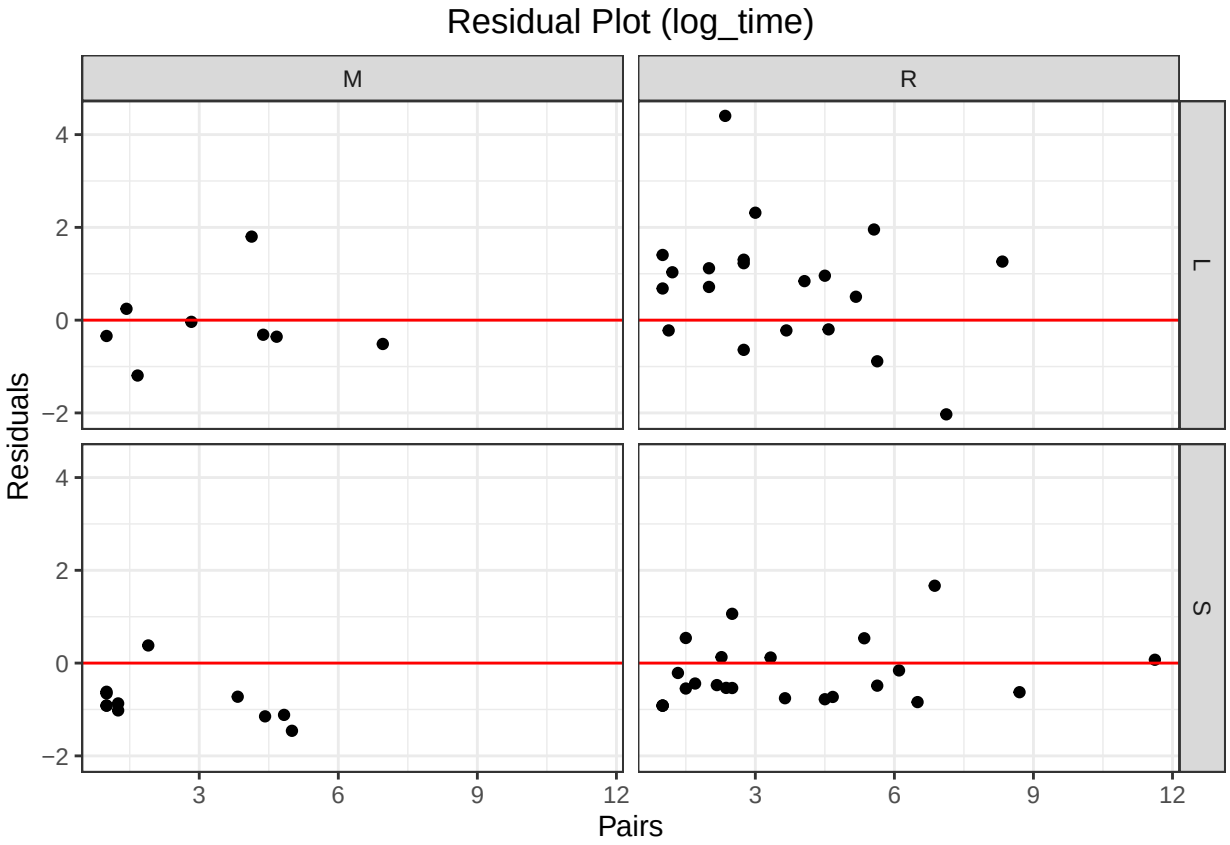
```
data$log_time <- log2(data$Time)
data$sqrt_time <- sqrt(data$Time)
data$inverse_time <- 1/data$Time
head(data)
```

```
##          Species Time Pairs Size Status  log_time sqrt_time inverse_time
## 1    Sparrowhawk 3.03  1.00    L      R 1.5993178  1.740690    0.3300330
## 2        Buzzard 5.46  2.00    L      R 2.4489010  2.336664    0.1831502
## 3        Kestrel 4.10  1.21    L      R 2.0356239  2.024846    0.2439024
## 4      Peregrine 1.68  1.13    L      R 0.7484612  1.296148    0.5952381
## 5 Grey_partridge 8.85  5.17    L      R 3.1456775  2.974895    0.1129944
## 6          Quail 1.49  1.00    L      M 0.5753123  1.220656    0.6711409
```

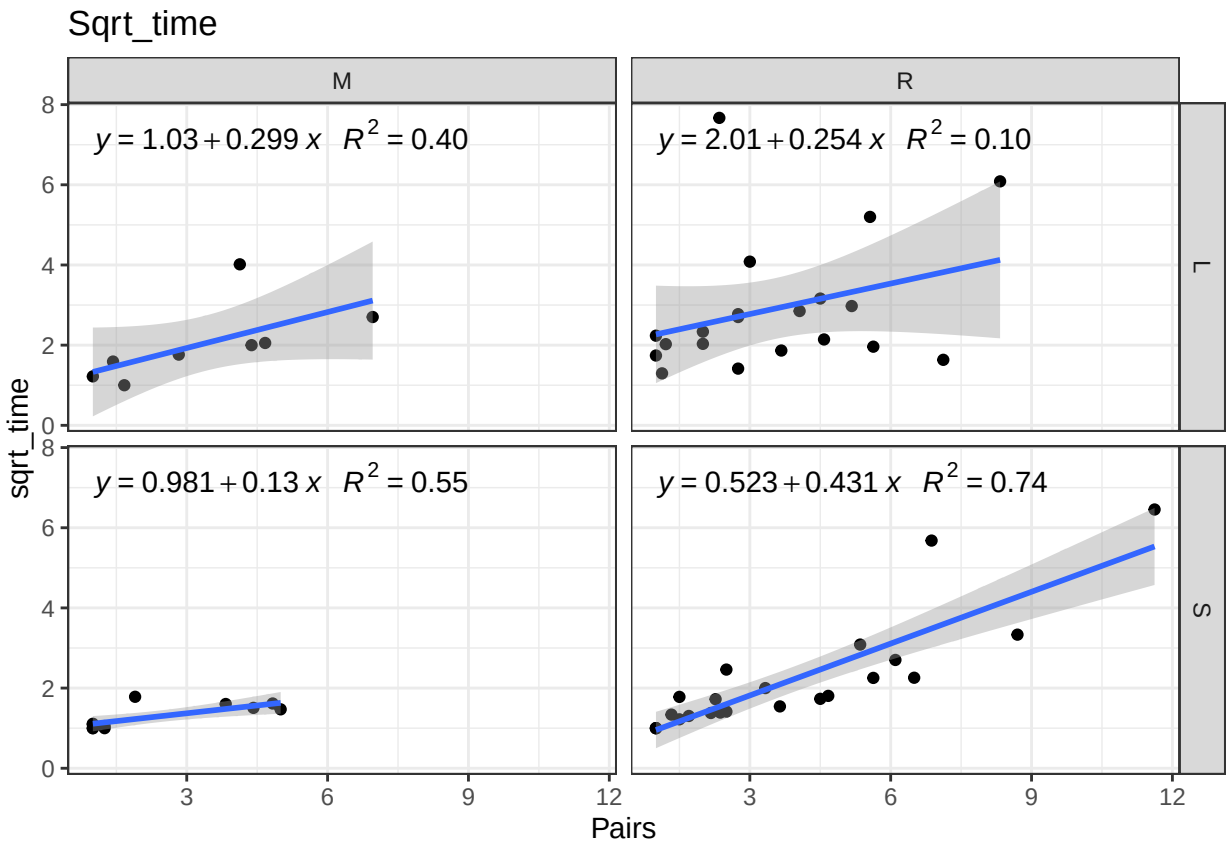
Plotting of log_time:



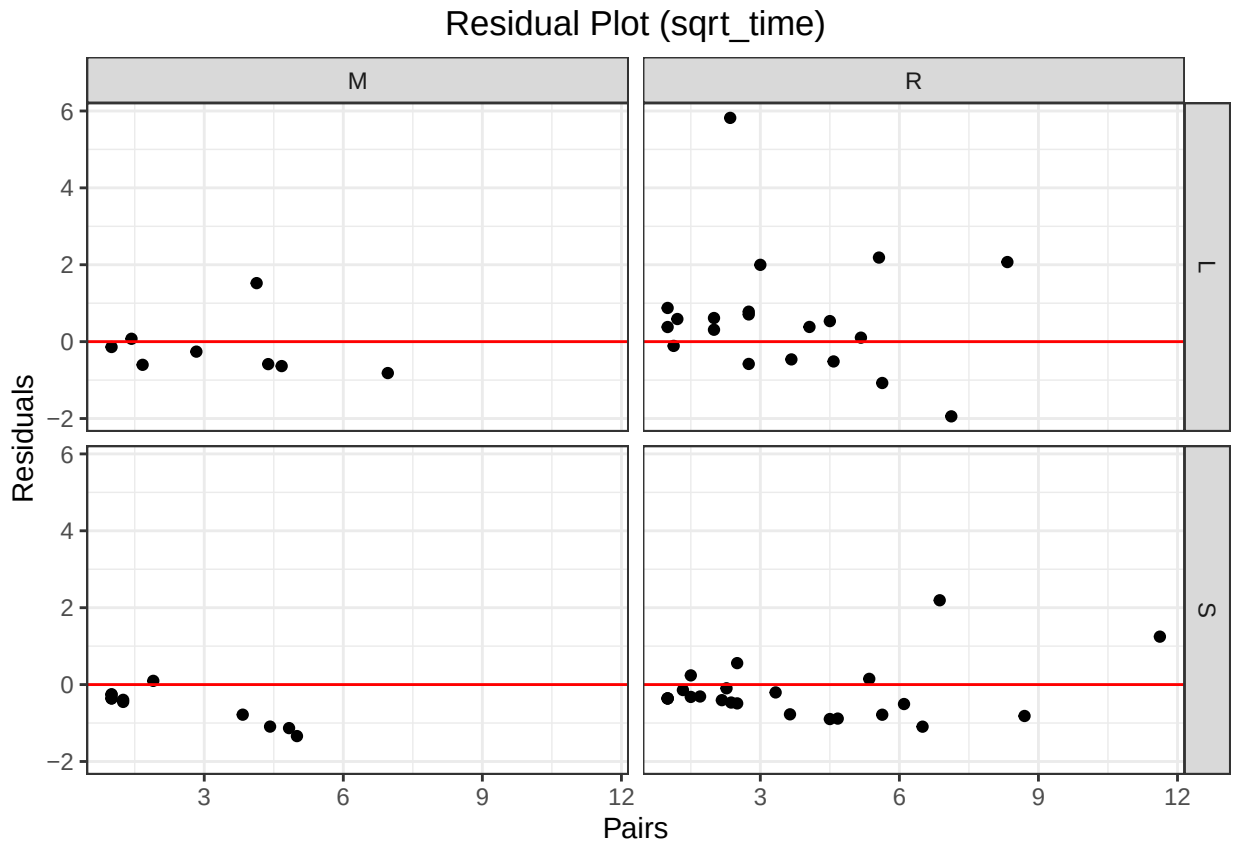
Residual plot of log_time:



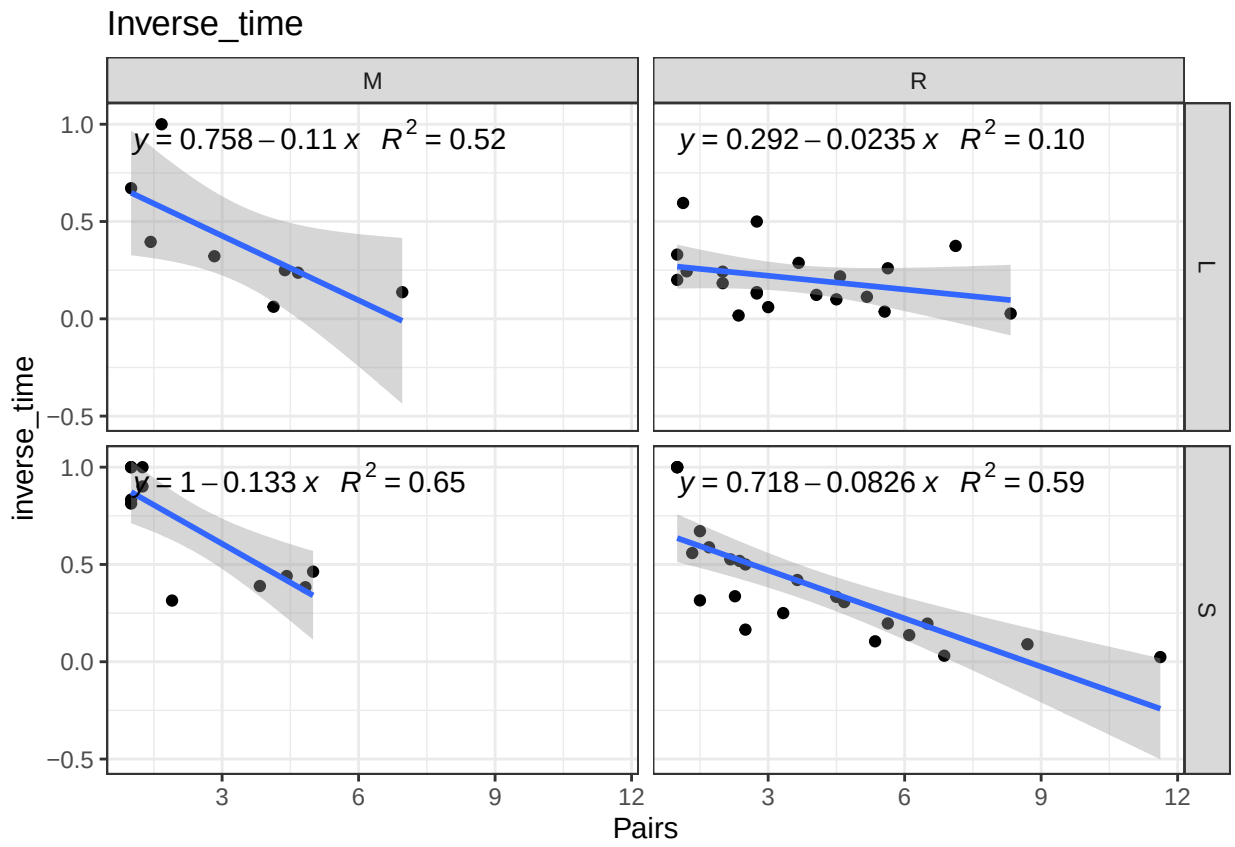
Plotting of sqrt_time:



Residual plot of sqrt_time:



Plotting of inverse_time:



Residual plot of inverse_time: