# SUPPLY CHAIN MANAGEMENT: PREDICTING THE FUTURE PRICE OF PRODUCT

PRESENTED BY YEN-CHING, LIN

# DATA DESCRIPTION

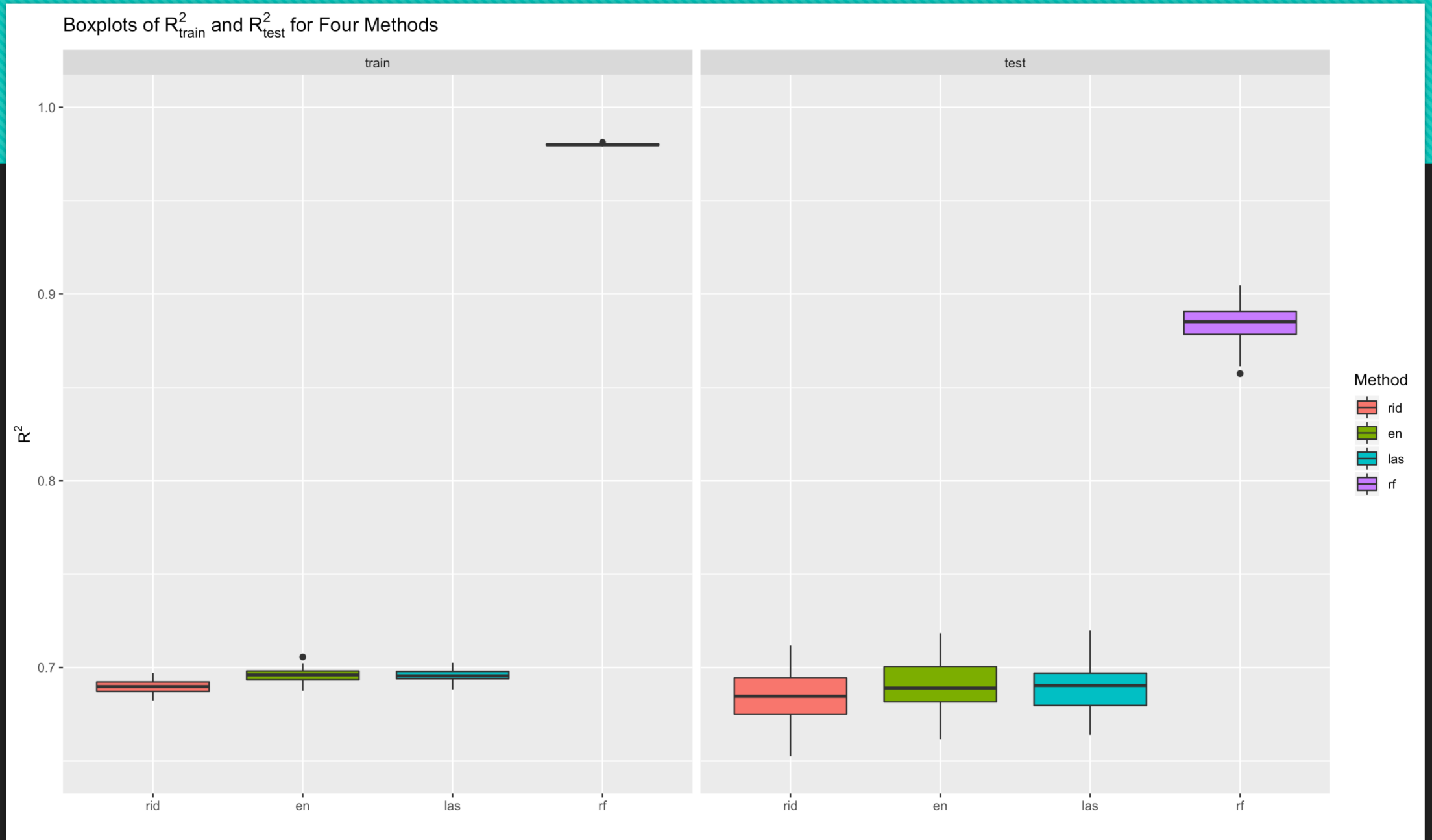
- Data source: scm20d (https://www.openml.org/d/41486)
- N = 8,966, P = 61 (includes response)
- All variables are numeric, no missing values
- Response:

  Mean price of product(computer) for 20-days in the future
- Features:

  Observed **prices** and historical price in specific day for product and it's component;

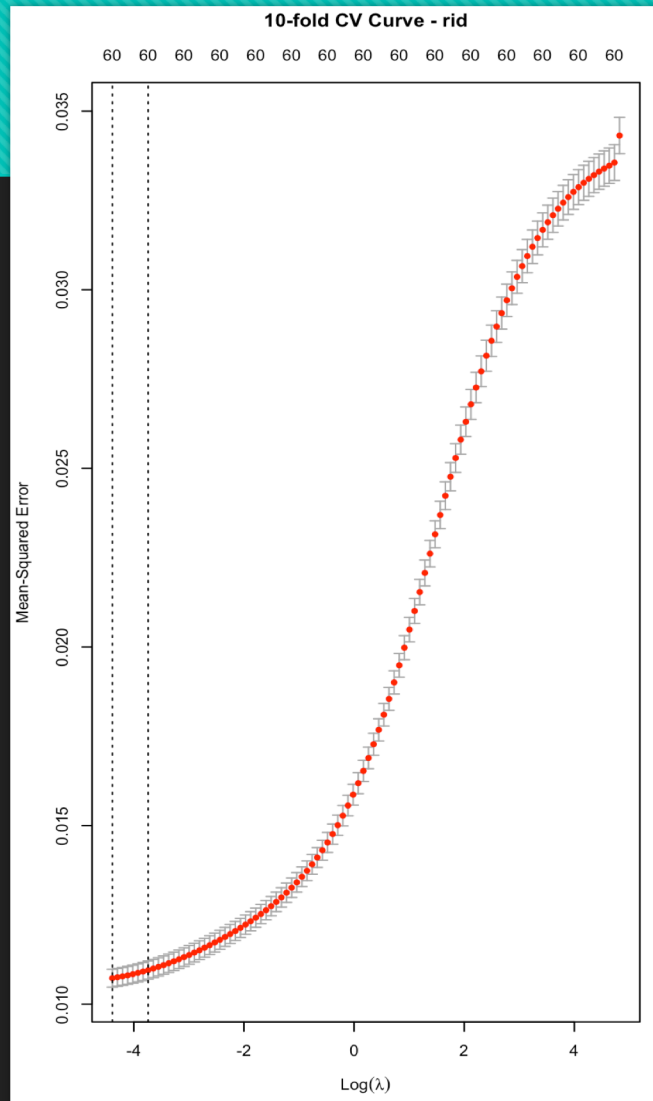
  Stock keeping units, Storage cost, Interest rates, Demand, etc.

# SIDE-BY-SIDE BOXPLOTS OF R-SQUARED



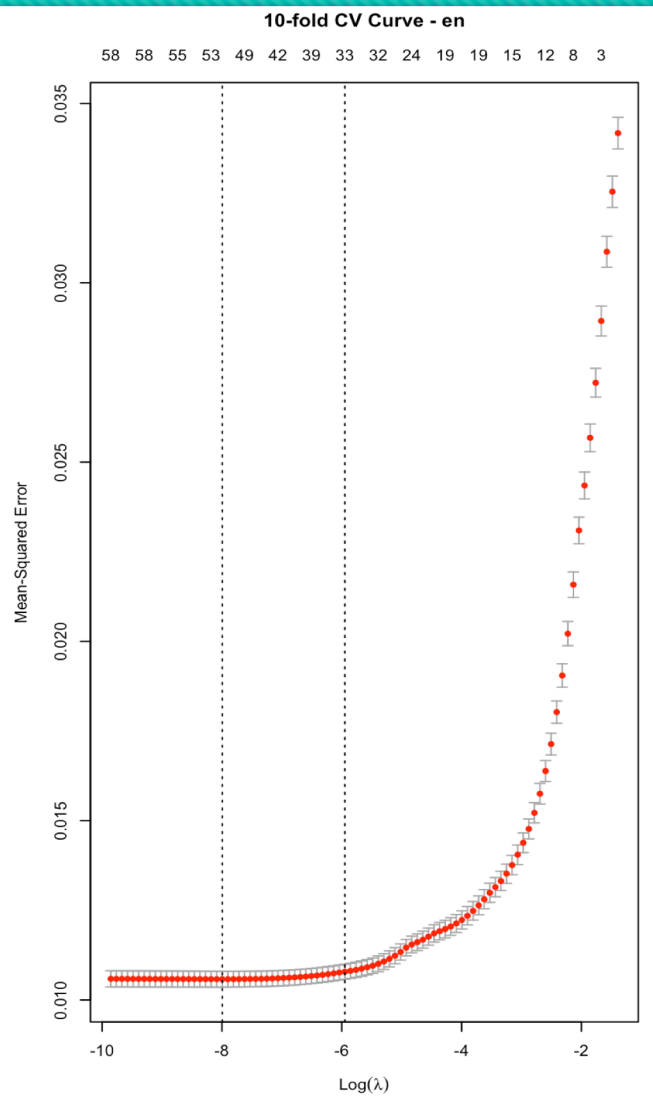
# 10-FOLD CV CURVES

0.543 sec



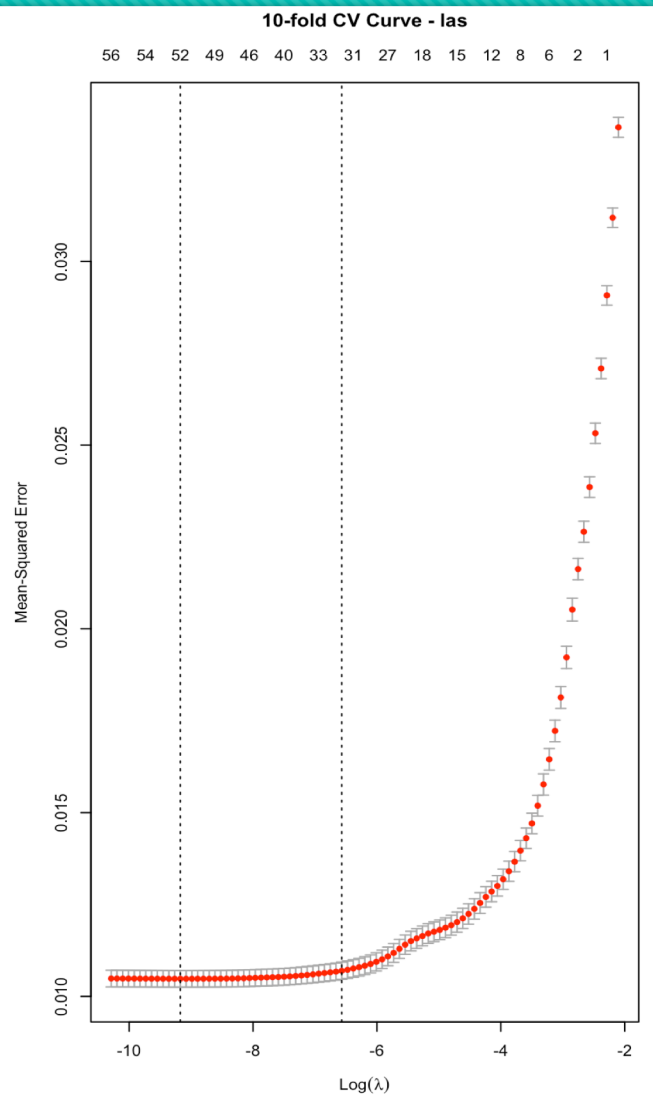

Ridge

0.872 sec



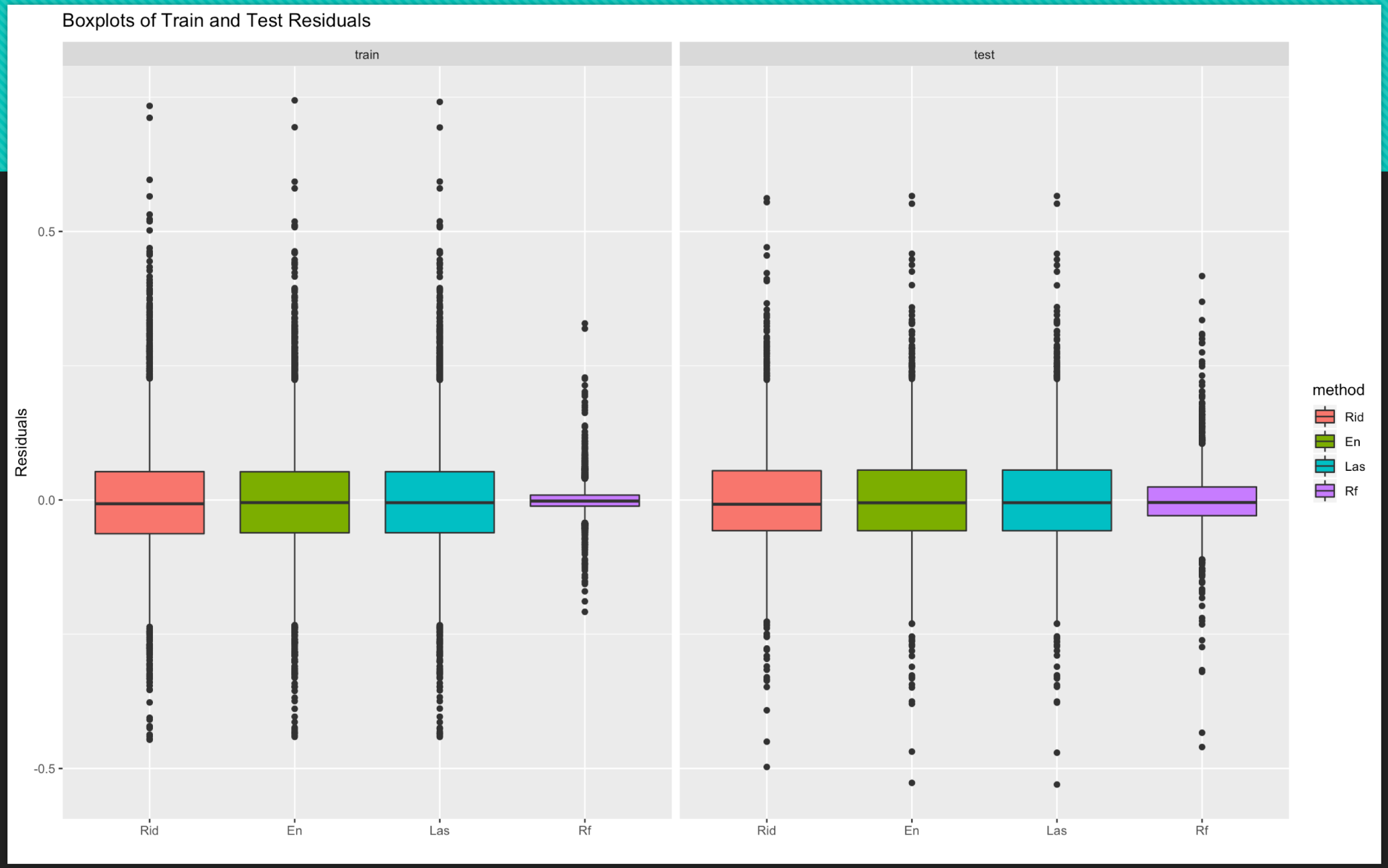

Elastic-Net

0.899 sec




Lasso

# SIDE-BY-SIDE BOXPLOTS OF RESIDUALS

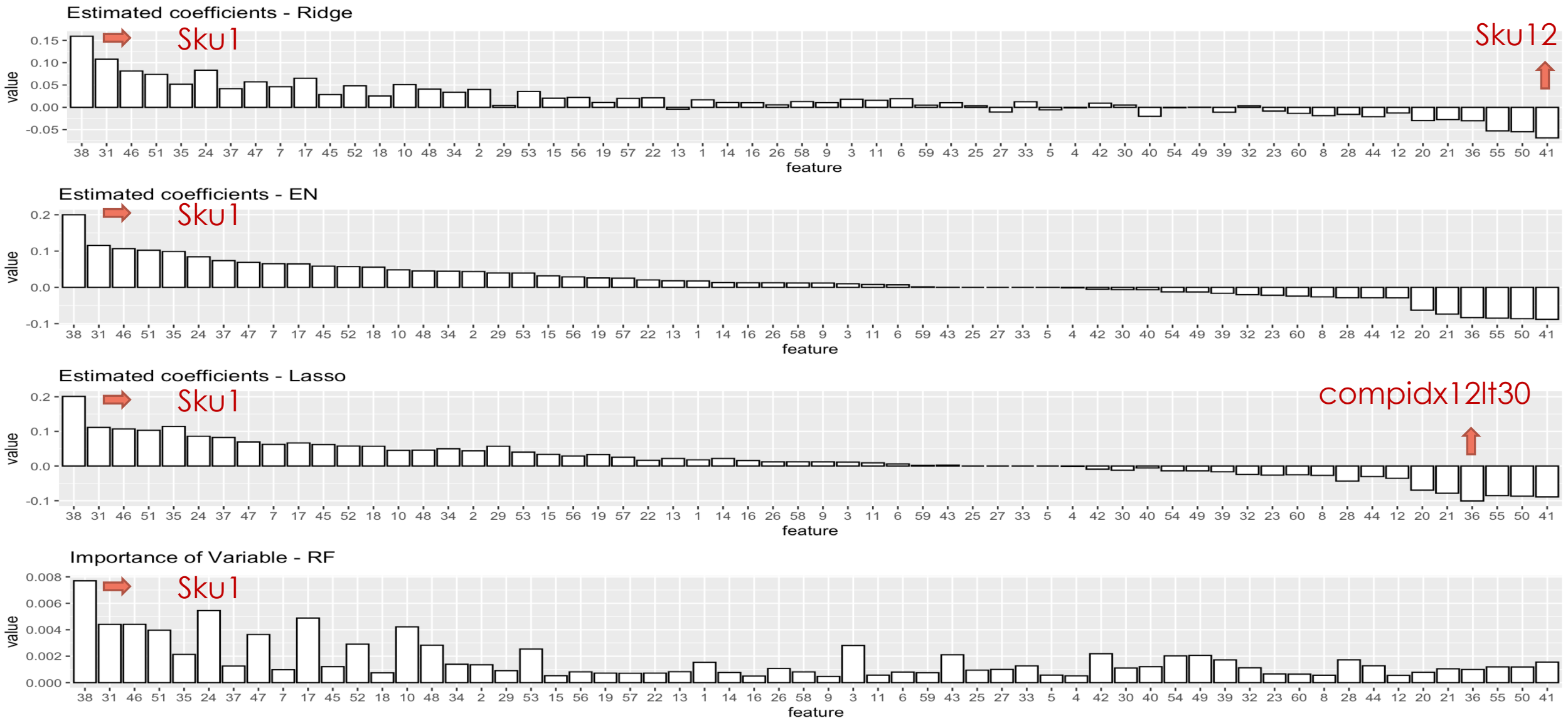


# PERFORMANCE AND RUNTIME

| Method | Test R-squared/ 90% interval | Time(sec) |
|---|---|---|
| Ridge | 0.6892 (0.6638, 0.7080) | 0.918 |
| Elastic-Net | 0.6954 (0.6700,0.7132) | 0.897 |
| Lasso | 0.6956 (0.6708,0.7062) | 0.887 |
| Random Forest | 0.9821 (0.8697,0.8958) | 99.186 |

- Trade-Off: Random Forest has the best performance, but also has the longest runtime

# BARPLOTS OF ESTIMATED COEFFICIENTS



# CONCLUSION

- Random forest performed better than Ridge, Lasso and Elastic-Net in both train and test datasets
- Trade-off between accuracy and fitting time

# THANK YOU!