

## STAT 213: MODIFIED EXERCISE 10.8

COLIN REIMER DAWSON

The `CAFE` dataset (in the `Stat2Data`), also described in examples throughout Chapter 10 of the text, includes information about how various U.S. senators voted on an amendment that would hamper the proposed Corporate Average Fuel Economy (CAFE) bill. The bill would have tightened regulations on fuel economy standards, and so a Yes vote on the amendment acts in opposition to tightened regulations. Examine the documentation on the dataset for more information.

The quantitative variable `LogContr` records how much money each senator received from the auto industry, on a log scale. The `Dem` indicator is 1 if the senator caucused with Democrats, 0 otherwise. Fit and compare a set of logistic regression models to address the following questions. For each question, identify the model(s) you used to address the question, and interpret each coefficient.

- (a) Does the probability of a Yes vote increase with (log) campaign contributions?
- (b) Does the probability of a Yes vote differ between parties?
- (c) Is the relationship between log campaign contribution and Vote different for those that caucus with Democrats?
- (d) Does knowing whether a senator caucuses with Democrats improve predictive ability *after controlling for* campaign contributions?

Key differences in R code between linear and logistic regression (all caps indicates a placeholder).

```
## To fit a model
model <- glm(FORMULA, family = "binomial", data = DATA)
## To do an overall likelihood ratio test
anova(model, test = "LRT")
## To do a nested likelihood ratio test
anova(REduced, FULL)
## To do stepwise selection based on AIC
```

---

*Date:* Last Revised January 21, 2018.

```
step(NULLMODEL, scope = list(upper = FULLMODEL))
```