Practice Problems on Prediction

1. When can loops make our code easier and/or simpler?

2. Take a look at the code below.

```
n <- 100
for (i in 1:n) {
  print(i^2)
}
```

   What values does i take on? What will be the first value printed? Second? Last?

3. Take a look at the code below.

```
x <- c(4, 7, 2, 1)
for (i in 1:length(x)) {
  print(x[i]*2)
}
```

   What values does i take on? What will be the first value printed? Second? Last?

4. Take a look at the code below.

```
x <- c(4, 7, 2, 1)
z <- rep(NA, length(x))
for (i in 1:length(x)) {
  z[i] <- x[i]*2)
}
```

   What values does i take on? What will be the first value printed (trick question)? What does the vector z look like before running the loop? After?

5. Calculate the mean ideology for Republicans from the 100th to the 113th Congress.

   (a) Load the `nominate.csv` data set.

   (b) Use the `subset` function to subset out Republicans (i.e., `party == "Republican"`) into a differently named data set (maybe `nominate_r`).

   (c) Create a vector `congresses` that contains each congress in the data set (i.e., 100 to 113).

   (d) Create a holder `means_r` that is the same length as the vector `congresses`. We will later fill this vector with the means for the Republicans in each Congress.

(e) Create a loop that takes `i` from 1 to the number of congresses (i.e., maybe `length(congresses)`?).

(f) Inside the loop, replace the `i`th value of `means_r` with the mean of `ideology_score` for the `congresses[i]`th Congress.

Repeat parts (b) through (f) for Democrats.

6. Go to the Huffington Post's Pollster and look at the current and past polling averages for the Republican candidates. This is comparable to the figure at the top of page 121 in *QSS*. Who seems most likely to win? Least likely?

7. Now go to the PredictWise page for the GOP nomination. Scroll down to the bottom and look at the betting data from Betfair, PredictIt, and Bookie. The amounts between $0.00 and $1.00 and represent the market price for a contract that pays $1 if the candidate wins and $0 if the candidate loses. This is informative, because if a candidate's contract is worth $0.33, we'd expect them to win about 33% of the time. This happens because of supply and demand–in equilibrium $1 × (chance of winning) = (market price). Why do you think the polls and betting markets diverge? Which do you think is more accurate?