

```
> library(tidyverse)
— Attaching packages — tidyverse 1.2.1 —
✓ ggplot2 3.2.1   ✓ purrr 0.3.2
✓ tibble 2.1.3    ✓ dplyr 0.8.3
✓ tidyr 0.8.3     ✓ stringr 1.4.0
✓ readr 1.3.1     ✓ forcats 0.4.0
— Conflicts — tidyverse_conflicts() —
✖ dplyr::filter() masks stats::filter()
✖ dplyr::lag() masks stats::lag()
> library(caret)
Loading required package: lattice
```

Attaching package: ‘caret’

The following object is masked from ‘package:purrr’:

lift

```
> set.seed(1, sample.kind="Rounding")
Warning message:
In set.seed(1, sample.kind = "Rounding") :
  non-uniform 'Rounding' sampler used
> n <- 100
> Sigma <- 9*matrix(c(1.0, 0.5, 0.5, 1.0), 2, 2)
> dat <- MASS::mvrnorm(n = 100, c(69, 69), Sigma) %>%
+   data.frame() %>% setNames(c("x", "y"))
> dim(dat)
[1] 100 2
> head(dat)
      x      y
1 68.30298 66.44188
2 69.41395 69.54029
3 68.19536 65.46259
4 72.90762 73.38170
5 70.83796 68.87421
6 64.21743 69.51929
>
>
> set.seed(1, sample.kind="Rounding")
Warning message:
In set.seed(1, sample.kind = "Rounding") :
  non-uniform 'Rounding' sampler used
> rmse <- replicate(100, {
+ test_index <- createDataPartition(dat$y, times = 1, p = 0.5, list = FALSE)
+ train_set <- dat %>% slice(-test_index)
+ test_set <- dat %>% slice(test_index)
+ fit <- lm(y ~ x, data = train_set)
+ y_hat <- predict(fit, newdata = test_set)
+ sqrt(mean((y_hat-test_set$y)^2))
+ })
> mean(rmse)
[1] 2.488661
> sd(rmse)
[1] 0.1243952
>
```