

Packages

```
library(naniar)
library(tidyverse)
```

Data

```
library(car)
```

```
## Loading required package: carData
```

```
##
```

```
## Attaching package: 'car'
```

```
## The following object is masked from 'package:dplyr':
```

```
##
```

```
##      recode
```

```
## The following object is masked from 'package:purrr':
```

```
##
```

```
##      some
```

```
data(prestige)
```

```
## Warning in data(prestige): data set 'prestige' not found
```

Training and test set

```
df <- Prestige %>% mutate(id = row_number())
## set the seed to make your partition reproducible
set.seed(123)
train <- df %>% sample_frac(.80)
dim(train)
```

```
## [1] 82  7
```

```
head(train)
```

```
##               education income women prestige census type id
## medical.technicians    12.79  5180 76.04    67.5   3156   wc  31
## welders                 7.92  6477  5.17    41.8   8335   bc  79
## commercial.travellers  11.13  8780  3.16    40.2   5133   wc  51
## economists             14.44  8049 57.31    62.2   2311  prof  14
## farmers                 6.84  3643  3.60    44.1   7112 <NA> 67
## receptionsts           11.04  2901 92.86    38.7   4171   wc  42
```

```
test <- anti_join(df, train, by = 'id')
dim(test)
```

```
## [1] 20  7
```