This report explains the approach taken to forecast sales across stores. It is reproduble based on code available at …

The first stage is to load in the data. We will use the tidyverse package and load it with the library() function:

library(tidyverse)  
# load data  
source("code/load-data.R")

The task is to estimate sales for the Departments of each store based on the historical training data.

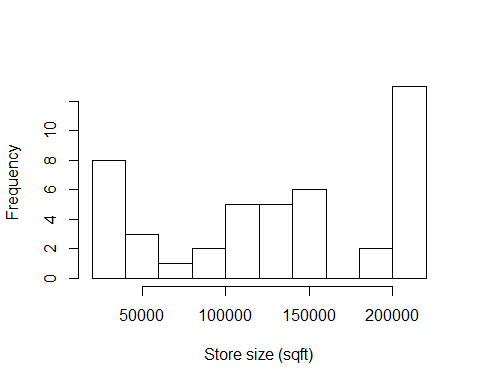
We can check that these files have been loaded into the R envionment with the following command:

ls()

## [1] "features" "stores" "test" "train"

## Describe data

hist(stores$`Size (sq ft)`, main = "", xlab = "Store size (sqft)")



## Joining the data

# join store-level data onto training dataset (so we know size)  
train\_joined = inner\_join(train, y = stores)

## Joining, by = "Store"