

# lab3\_estes

andrew.estes

9/11/2021

Install Data and Packages

```
#install.packages("AmesHousing")
#tinytex::install_tinytex()
library(AmesHousing)

## Warning: package 'AmesHousing' was built under R version 4.1.1

library(tidyverse)

## -- Attaching packages ----- tidyverse 1.3.1 --

## v ggplot2 3.3.5     v purrr    0.3.4
## v tibble   3.1.3     v dplyr    1.0.7
## v tidyr    1.1.3     v stringr  1.4.0
## v readr    2.0.0     vforcats  0.5.1

## Warning: package 'tidyverse' was built under R version 4.1.1

## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()   masks stats::lag()

library(tinytex)
AmesData <- make_ames()
#view(AmesData)
#str(AmesData)
```

Create a new column in the AmesData data frame which has a value of 1 if the house is built in year 2000 or later or 0 if it is built before year 2000 using the following method: 1) IF-ELSE 2) FOR 3) SPLY 4) Question 3 - Question 1

1) ifelse(name this column – w3ifelse). What is the first 10 elements of w3ifelse?

```
AmesData2 <- AmesData %>%
  select(Year_Built) %>%
  mutate(w3ifelse = ifelse(Year_Built > 1999, 1, 0))

head(AmesData2, 10)

## # A tibble: 10 x 2
##   Year_Built w3ifelse
##       <int>     <dbl>
## 1     1960      0
## 2     1961      0
## 3     1958      0
## 4     1968      0
## 5     1997      0
## 6     1998      0
## 7     2001      1
## 8     1992      0
## 9     1995      0
## 10    1999      0

view(AmesData2)
```

2) if-Else and a For loop (name this column – w3for). What is the first 10 elements of w3for?

```
AmesData3 <- AmesData2

w3for = for (i in 1:nrow(AmesData3)){
  if(AmesData3$Year_Built[i] < 2000){
    AmesData3$Year_Built[i] <- 0
  } else {
    AmesData3$Year_Built[i] <- 1
  }
}

head(AmesData3, 10)

## # A tibble: 10 x 2
##   Year_Built w3ifelse
##       <dbl>     <dbl>
## 1         0         0
## 2         0         0
## 3         0         0
## 4         0         0
## 5         0         0
## 6         0         0
## 7         1         1
## 8         0         0
## 9         0         0
## 10        0         0
```

3) Build your own function and use sapply (name this column - w3apply). What is the first 10 elements of w3apply?

```
woo <- function(x){  
  if(is.na(x)) return(x)  
  else if (x <= 1999) return (0)  
  else return (1)  
}  
  
AmesData4 <- AmesData2 %>%  
  select(Year_Built, w3ifelse) %>%  
  mutate(w3apply = sapply(AmesData2$Year_Built, woo))  
  
head(AmesData4)  
  
## # A tibble: 6 x 3  
##   Year_Built w3ifelse w3apply  
##       <int>     <dbl>     <dbl>  
## 1      1960      0.000      0  
## 2      1961      0.000      0  
## 3      1958      0.000      0  
## 4      1968      0.000      0  
## 5      1997      0.000      0  
## 6      1998      0.000      0
```

**4) Create a column w3diff which is the difference of the two columns w3ifelse and w3apply.**

DO NOT use a for loop or the apply set of functions for this step. Simple subtraction will work. What is the total sum of this column?

```
AmesDataFinal <- AmesData4 %>%
  mutate(w3diff = w3apply - w3ifelse)

head(AmesDataFinal, 10)

## # A tibble: 10 x 4
##   Year_Built w3ifelse w3apply w3diff
##       <int>     <dbl>    <dbl>    <dbl>
## 1      1960      0        0        0
## 2      1961      0        0        0
## 3      1958      0        0        0
## 4      1968      0        0        0
## 5      1997      0        0        0
## 6      1998      0        0        0
## 7      2001      1        1        0
## 8      1992      0        0        0
## 9      1995      0        0        0
## 10     1999      0        0        0
```