Lesson 4: Dive into Tidy: Data Wrangling and Manipulation Part 2 - Homework

Instructor: Caitlin Allen Akselrud (caitlin.allen_akselrud@noaa.gov)

February 04, 2021

Contents

Questions:

First, load your libraries!

```
library(tidyverse)
library(lubridate)
library(tidytuesdayR)
```

Use tree_dat for the following questions:

```
trees <- tidytuesdayR::tt_load('2020-01-28')
```

```
##
##   Downloading file 1 of 2: 'sf_trees.csv'
##   Downloading file 2 of 2: 'Street_Tree_Map.csv'
```

```
## Warning: 952155 parsing failures.
##  row       col              expected                                    actual         file
## 2670 XCoord    1/0/T/F/TRUE/FALSE 6005345.6576                                 <raw vector>
## 2670 YCoord    1/0/T/F/TRUE/FALSE 2111619.4154                                 <raw vector>
## 2670 Latitude  1/0/T/F/TRUE/FALSE 37.7784066780298                             <raw vector>
## 2670 Longitude 1/0/T/F/TRUE/FALSE -122.425067563729                            <raw vector>
## 2670 Location  1/0/T/F/TRUE/FALSE (37.7784066780298, -122.425067563729) <raw vector>
## .... ......... ................. ...................................... ............
## See problems(...) for more details.
```

```
trees
tree_dat <- trees$sf_trees
```

1   Are there any duplicate tree_id rows?

2   How many trees of each species are there?

3   How old is the oldest tree in the data set, in years?

4   How many trees are planted by year?

5   How many are planted by decade?

The below function might be helpful to you for determining the decade. In this case you'll use it in the dplyr::mutate() function.

```
floor_decade <- function(value){ return(value - value %% 10) } # I found this on stackoverflow
```