Data Tidying

Monika Sikand

1/16/2019

```r
library(dplyr)
library(tidyr)
#library(stats)
```

to call a function from a specific package 'package_name::function_name(. . .)'

Data Cleaning

Read in data file

```r
catch_df <- read.csv(url( "https://knb.ecoinformatics.org/knb/d1/mn/v2/object/df35b.302.1",method = "lil
head(catch_df)
```

```
##   Region Year Chinook Sockeye Coho Pink Chum All notesRegCode
## 1    SSE 1886       0       5    0    0   0   5
## 2    SSE 1887       0     155    0    0   0 155
## 3    SSE 1888       0     224   16    0   0 240
## 4    SSE 1889       0     182   11   92   0 285
## 5    SSE 1890       0     251   42    0   0 292
## 6    SSE 1891       0     274   24    0   0 298
```

control +SHift + m gives %>% pipe operator or

command + shift + m gives %>% pipe operator

catch_cleaned <- catch_df %>% select(-All, -notesRegCode)

- Remove marginal some and notes column
- Move from wide to long format

```r
catch_long <- catch_df %>%
  select(-All, -notesRegCode)%>%  # negative values remove columns
  gather(key = "species", value = "catch",-Year, -Region)

head(catch_long)
```

```
##   Region Year species catch
## 1    SSE 1886 Chinook     0
## 2    SSE 1887 Chinook     0
## 3    SSE 1888 Chinook     0
## 4    SSE 1889 Chinook     0
## 5    SSE 1890 Chinook     0
## 6    SSE 1891 Chinook     0
```

%>%

Split- Apply-Combine

Joins