Data Wrangling for filtering rows and selecting columns

This lab is preconfigured to include all dependencies (libraries, packages, and datasets) you'll need to complete your work in RStudio. You can practice, run test cases, and work on assignments from your browser.

Practice filter() and select() with GOES-R FDC Data Product

- Data for the lab is the GOES-R fire detection and characterization (FDC) data product, which uses both visible and infrared (IR) ABI spectral channels (or bands) to locate fires and retrieve fire characteristics. It was retrieved and stored in a CSV file. You can find the data Crop_Range_GOES0901_R.csv in your "Files" tab of RStudio.

In the data, some variables are described below:

- year: data collected year.
- jday: the Julian date. The Julian day is the continuous count of days since the beginning of the Julian period. Here the beginning of the Julian period is the first day of the year (Jan 1th). - gmt: Greenwich Mean Time.
- frp: Fire Radiative Power.
- Bintime: round-up time for gmt. The data is expected with a temporal resolution of 0.5 minutes.] For example, 10.2 is to be 10.5, and 11.7 is to be 12.

Please practice the operations below.

- Read the data into R/RStudio.
- Selects the columns "year, jday,gmt,frp, Bintime"
- Filter the row by the jday >227

Try executing this chunk by clicking the *Run* button within the chunk or by placing your cursor inside it and pressing *Ctrl+Shift+Enter*.

```
#I need this package to read the CSV file
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
## v dplyr     1.1.2     v readr     2.1.4
## v forcats   1.0.0     v stringr   1.5.0
## v ggplot2   3.4.2     v tibble    3.2.1
## v lubridate 1.9.2     v tidyr     1.3.0
## v purrr     1.0.1
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
# set your own working directory - remember to modify this command to match your local or lab environme

# For example:
# - If you are using the In-Browser RStudio lab option for this activity, your working directory will s
# - If your work is stored in a sub-folder within your lab, let's say a folder titled "Module" you'll c
# For this activity your working directory should be set to "/home/rstudio".
```

```r
# setwd("C:/Users/R_Week2")


# Read the data into RStudio. I'm going to call the file fire_data
fire_data <- read_csv("Crop_Range_GOES0901_R.csv")
```

```
## Rows: 23652 Columns: 16
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## dbl (16): FID, year, jday, gmt, icount, lat, lon, code, frp, zen, 5kmIGBP, 5...
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```r
# Filter the row by the jday >227. I used filter and selected the data(fire_data) and told it to give m

filter (fire_data, jday > 227)
```

```
## # A tibble: 16,546 x 16
##      FID  year  jday   gmt icount   lat   lon  code   frp   zen `5kmIGBP`
##    <dbl> <dbl> <dbl> <dbl>  <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>     <dbl>
##  1     1  2019   229  11.7  22273  43.5 -94.5    15  12.1  36.6        12
##  2     2  2019   232  18.2  33351  43.5 -95.1    15  23.1  67.9        12
##  3     4  2019   229  13.3  22827  43.5 -95.9    15  10.1  29.7        12
##  4     5  2019   228  18.1  20748  43.4 -96.4    15  19.3  65.2        12
##  5     6  2019   232  17.2  33130  43.4 -95.2    15   7.4  57.0        12
##  6     7  2019   241  14.7  96282  43.4 -95.4    15  12.6  37.6        12
##  7     8  2019   242  13.7  97622  43.4 -95.1    15  24.9  34.2        12
##  8     9  2019   232  13.5  32355  43.2 -91.7    15  21.4  30.6        12
##  9    10  2019   239  15.5  94264  43.3 -94.1    15   6.2  43.5        12
## 10    11  2019   241  14.7  96283  43.3 -95.0    15  19.2  37.6        12
## # i 16,536 more rows
## # i 5 more variables: `500mIGBP` <dbl>, `30mFCCS` <dbl>, `30mCDL` <dbl>,
## #   `1kmGLCC` <dbl>, BinTime <dbl>
```

```r
# Use pipe `%>%` to connect filter() and select()
fire_data %>% select(year, jday, gmt, frp, BinTime) %>% filter (jday > 227)
```

```
## # A tibble: 16,546 x 5
##     year  jday   gmt   frp BinTime
##    <dbl> <dbl> <dbl> <dbl>   <dbl>
##  1  2019   229  11.7  12.1    12
##  2  2019   232  18.2  23.1    18.5
##  3  2019   229  13.3  10.1    13.5
##  4  2019   228  18.1  19.3    18.5
##  5  2019   232  17.2   7.4    17.5
##  6  2019   241  14.7  12.6    15
##  7  2019   242  13.7  24.9    14
##  8  2019   232  13.5  21.4    14
##  9  2019   239  15.5   6.2    16
## 10  2019   241  14.7  19.2    15
## # i 16,536 more rows
```