Topic ideas
STA 210 - Project

Ginger and Stats - Rakshita Ramakrishna, Aimi Wen, Nathan Nguyen, Bryan Pan

```r
library(tidyverse)
```

```
Warning in system("timedatectl", intern = TRUE): running command 'timedatectl'
had status 1


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


v ggplot2 3.3.5      v purrr   0.3.4
v tibble  3.1.6      v dplyr   1.0.7
v tidyr   1.1.4      v stringr 1.4.0
v readr   2.1.1      v forcats 0.5.1


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

Project idea 1

Introduction and data

- State the source of the data set.

- Describe when and how it was originally collected (by the original data curator, not necessarily how you found the data)

- Describe the observations and the general characteristics being measured in the data

Research question

- Describe a research question you're interested in answering using this data.

Glimpse of data

- Use the `glimpse` function to provide an overview of the dataset

```
details <- readr::read_csv('https://raw.githubusercontent.com/rfordatascience/tidytuesday/mas
```

```
detail_2016 <- details %>%
  filter(yearpublished == 2016)
```

```
glimpse(detail_2016)
```

```
Rows: 1,291
Columns: 23
$ num                     <dbl> 7, 12, 47, 54, 55, 57, 62, 78, 94, 97, 135, 14~
$ id                      <dbl> 167791, 169786, 204583, 205637, 193738, 201808~
$ primary                 <chr> "Terraforming Mars", "Scythe", "Kingdomino", "~
$ description             <chr> "In the 2400s, mankind begins to terraform the~
$ yearpublished           <dbl> 2016, 2016, 2016, 2016, 2016, 2016, 2016, 2016~
$ minplayers              <dbl> 1, 1, 2, 1, 2, 2, 2, 1, 2, 2, 5, 1, 2, 2, 1, 2~
$ maxplayers              <dbl> 5, 5, 4, 2, 4, 4, 4, 5, 8, 4, 10, 4, 8, 8, 4, ~
$ playingtime             <dbl> 120, 115, 15, 120, 150, 60, 20, 180, 20, 240, ~
$ minplaytime             <dbl> 120, 90, 15, 60, 75, 30, 20, 120, 20, 180, 45,~
$ maxplaytime             <dbl> 120, 115, 15, 120, 150, 60, 20, 180, 20, 240, ~
$ minage                  <dbl> 12, 14, 8, 14, 12, 12, 8, 14, 8, 14, 13, 12, 1~
$ boardgamecategory       <chr> "['Economic', 'Environmental', 'Industry / Man~
$ boardgamemechanic       <chr> "['Drafting', 'End Game Bonuses', 'Hand Manage~
$ boardgamefamily         <chr> "['Components: Map (Global Scale)', 'Component~
$ boardgameexpansion      <chr> "['French Championship Promo Cards (fan expans~
$ boardgameimplementation <chr> "['Terraforming Mars: Ares Expedition', 'Terra~
$ boardgamedesigner       <chr> "['Jacob Fryxelius']", "['Jamey Stegmaier']", ~
$ boardgameartist         <chr> "['Isaac Fryxelius']", "['Jakub Rozalski']", "~
$ boardgamepublisher      <chr> "['FryxGames', 'Arclight', 'Brädspel.se', 'Fan~
$ owned                   <dbl> 101872, 86371, 61775, 59731, 39730, 42118, 510~
$ trading                 <dbl> 538, 623, 735, 806, 361, 364, 846, 404, 336, 2~
$ wanting                 <dbl> 2011, 1896, 381, 780, 1568, 1217, 643, 1157, 4~
$ wishing                 <dbl> 19227, 19325, 3788, 8566, 10618, 9444, 7194, 1~
```

Project idea 2

Introduction and data

- State the source of the data set.

- Describe when and how it was originally collected (by the original data curator, not necessarily how you found the data)

- Describe the observations and the general characteristics being measured in the data

Research question

- Describe a research question you're interested in answering using this data.

Glimpse of data

- Use the `glimpse` function to provide an overview of the dataset

```
chocolate <- read_csv('https://raw.githubusercontent.com/rfordatascience/tidytuesday/master/
```

```
Rows: 2530 Columns: 10
```

```
-- Column specification -------------------------------------------------------
Delimiter: ","
chr (7): company_manufacturer, company_location, country_of_bean_origin, spe...
dbl (3): ref, review_date, rating

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.
```

```
glimpse(chocolate)
```

```
Rows: 2,530
Columns: 10
$ ref                     <dbl> 2454, 2458, 2454, 2542, 2546, 2546, 2~
$ company_manufacturer    <chr> "5150", "5150", "5150", "5150", "5150~
$ company_location        <chr> "U.S.A.", "U.S.A.", "U.S.A.", "U.S.A.~
$ review_date             <dbl> 2019, 2019, 2019, 2021, 2021, 2021, 2~
$ country_of_bean_origin  <chr> "Tanzania", "Dominican Republic", "Ma~
```

```
$ specific_bean_origin_or_bar_name <chr> "Kokoa Kamili, batch 1", "Zorzal, bat~
$ cocoa_percent                    <chr> "76%", "76%", "76%", "68%", "72%", "8~
$ ingredients                      <chr> "3- B,S,C", "3- B,S,C", "3- B,S,C", "~
$ most_memorable_characteristics   <chr> "rich cocoa, fatty, bready", "cocoa, ~
$ rating                           <dbl> 3.25, 3.50, 3.75, 3.00, 3.00, 3.25, 3~
```

Project idea 3

Introduction and data

- State the source of the data set.

- Describe when and how it was originally collected (by the original data curator, not necessarily how you found the data)

- Describe the observations and the general characteristics being measured in the data

Research question

- Describe a research question you're interested in answering using this data.

Glimpse of data

- Use the `glimpse` function to provide an overview of the dataset

```
# add code to load and glimpse data here
```