Topic ideas

Pandas: Jessica Tang, Victoria Midkiff, Alan Wang, June Lee

October 13, 2022

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

```
fastfood <- read_csv("data/fastfood_calories.csv")
steak_survey <- read_csv("data/steak-risk-survey.csv")
```

Data Set 1

Introduction and Data

This data set is called Fast Food and was taken from the tidytuesday repository. The data was taken from https://fastfoodnutrition.org/fast-food-restaurants. The original data curator found the data from the nutrition contents of the various fast food items, taken by the fast food restaurants. The observations are individual menu items at a particular fast food restaurant. The characteristics include various nutritional measurements such as calories, sodium, cholesterol, and vitamin A.

Research questions

How can total carbohydrates, fiber, sugar, and protein predict the total calories of a fast food item?

Glimpse of data

```
glimpse(fastfood)
```

```
Rows: 515
Columns: 18
$ ...1        <dbl> 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17,~
$ restaurant  <chr> "Mcdonalds", "Mcdonalds", "Mcdonalds", "Mcdonalds", "Mcdon~
$ item        <chr> "Artisan Grilled Chicken Sandwich", "Single Bacon Smokehou~
$ calories    <dbl> 380, 840, 1130, 750, 920, 540, 300, 510, 430, 770, 380, 62~
$ cal_fat     <dbl> 60, 410, 600, 280, 410, 250, 100, 210, 190, 400, 170, 300,~
$ total_fat   <dbl> 7, 45, 67, 31, 45, 28, 12, 24, 21, 45, 18, 34, 20, 34, 8, ~
$ sat_fat     <dbl> 2.0, 17.0, 27.0, 10.0, 12.0, 10.0, 5.0, 4.0, 11.0, 21.0, 4~
$ trans_fat   <dbl> 0.0, 1.5, 3.0, 0.5, 0.5, 1.0, 0.5, 0.0, 1.0, 2.5, 0.0, 1.5~
$ cholesterol <dbl> 95, 130, 220, 155, 120, 80, 40, 65, 85, 175, 40, 95, 125, ~
$ sodium      <dbl> 1110, 1580, 1920, 1940, 1980, 950, 680, 1040, 1040, 1290, ~
$ total_carb  <dbl> 44, 62, 63, 62, 81, 46, 33, 49, 35, 42, 38, 48, 48, 67, 31~
$ fiber       <dbl> 3, 2, 3, 2, 4, 3, 2, 3, 2, 3, 2, 3, 3, 5, 2, 2, 3, 3, 5, 2~
$ sugar       <dbl> 11, 18, 18, 18, 18, 9, 7, 6, 7, 10, 5, 11, 11, 11, 6, 3, 1~
$ protein     <dbl> 37, 46, 70, 55, 46, 25, 15, 25, 25, 51, 15, 32, 42, 33, 13~
$ vit_a       <dbl> 4, 6, 10, 6, 6, 10, 10, 0, 20, 20, 2, 10, 10, 10, 2, 4, 6,~
$ vit_c       <dbl> 20, 20, 20, 25, 20, 2, 2, 4, 4, 6, 0, 10, 20, 15, 2, 6, 15~
$ calcium     <dbl> 20, 20, 50, 20, 20, 15, 10, 2, 15, 20, 15, 35, 35, 35, 4, ~
$ salad       <chr> "Other", "Other", "Other", "Other", "Other", "Other", "Oth~
```

Data Set 2

Introduction and Data

This data set is called Steak Survey and is from FiveThirty Eight. The data were collected through a SurveyMonkey Audience poll that DataLab sent out to 550 people in May 2014. The observations are individuals' responses to the survey. The questions asked about whether or not they would engage in risky activities (i.e. smoking, skydiving, gambling), if they eat steak, how they like their steak cooked, as well as demographic data such as the individual's age, gender, household income, education, and the region in the US that they live in.

Research questions

How can an individual's desire to engage in certain risky activities and how they like/order steak predict their household income?

Glimpse of data

```
glimpse(steak_survey)
```

```
Rows: 551
Columns: 15
$ RespondentID
$ `Consider the following hypothetical situations: <br>In Lottery A, you have a 50% chance o
$ `Do you ever smoke cigarettes?`
$ `Do you ever drink alcohol?`
$ `Do you ever gamble?`
$ `Have you ever been skydiving?`
$ `Do you ever drive above the speed limit?`
$ `Have you ever cheated on your significant other?`
$ `Do you eat steak?`
$ `How do you like your steak prepared?`
$ Gender
$ Age
$ `Household Income`
$ Education
$ `Location (Census Region)`
```

Data Set 3

Introduction and Data

(Introduce and discuss data here)

Research questions

(Discuss research questions here)

Glimpse of data

```
# glimpse data set 3
```