STS 122 - Lab 1

Roella Louise Pineda

Lab 1: Working with Data Frames

Due 1.8.21

Scenario

You are the proprietor of a very small bookstore. You currently have the following books:

- *The Waste Land* by T.S. Eliot (published 1922; 10 copies in stock; $8 each)
- *Lord of the Flies* by William Golding (published 1954; 16 copies in stock; $8 each)
- *Grapes of Wrath* by John Steinbeck (published 1939; 5 copies in stock; $10 each)
- *Pride and Prejudice* by Jane Austen (published 1813; 7 copies in stock; $15 each)
- *Frankenstein* by Mary Shelley (published 1818; 15 copies in stock; $5 each)
- *Little Women* by Louisa May Alcott (published 1868; 9 copies in stock; $20 each)
- *Jane Eyre* by Charlotte Bronte (published 1847; 12 copies in stock; $25 each)
- *The Things They Carried* by Tim O'Brien (published 1990; 2 copies in stock; $17 each)
- *Huckleberry Finn* by Mark Twain (published 1884; 21 copies in stock; $12 each)
- *Siddhartha* by Hermann Hesse (published 1922; 4 copies in stock; $22 each)

Inventory

In the code block below, fill in the blanks to create a data frame with a row for each book and a column for each characteristic of the book (title, author, year, etc.). We will do this by creating a vector for each characteristic, then combining the vectors into a data frame. Print your result.

```
title <- c("The Waste Land", "Lord of the Flies", "Grapes of Wrath", "Pride and Prejudice", "Frankenstei
author <- c("T.S. Eliot", "William Golding", "John Steinbeck", "Jane Austen", "Mary Shelley", "Louisa Ma
pyear <- c(1922, 1954, 1939, 1813, 1818, 1868, 1847, 1990, 1884, 1922)
nstock <- c(10, 16, 5, 7, 15, 9, 12, 2, 21, 4)
price <- c(8, 8, 10, 15, 5, 20, 25, 17, 12, 22)

bookstore <- data.frame(title, author, pyear, nstock, price)
```

```
bookstore
```

```
##                       title            author pyear nstock price
## 1           The Waste Land        T.S. Eliot  1922     10     8
## 2         Lord of the Flies   William Golding  1954     16     8
## 3           Grapes of Wrath    John Steinbeck  1939      5    10
## 4       Pride and Prejudice       Jane Austen  1813      7    15
## 5              Frankenstein      Mary Shelley  1818     15     5
## 6              Little Women Louisa May Alcott  1868      9    20
## 7                 Jane Eyre  Charlotte Bronte  1847     12    25
## 8  The Things They Carried       Tim O'Brien  1990      2    17
## 9          Huckleberry Finn        Mark Twain  1884     21    12
## 10              Siddhartha     Hermann Hesse  1922      4    22
```

Organization

For the remainder of the lab, we will be using functions from `dplyr`. Load the `dplyr` package in the code block below.

```
library(dplyr)
```

```
## Warning: package 'dplyr' was built under R version 4.0.5
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

Saving as a new data frame called `books_sorted`, fill in the blanks below to sort your inventory by year published (most recent first, so descending order) and, within year, by price (least expensive first, so ascending order). This can be done in either one step or two, but pay attention to the order. Print your result.

```
books_sorted <- bookstore %>% arrange(-pyear, price)
books_sorted
```

```
##                       title            author pyear nstock price
## 1  The Things They Carried       Tim O'Brien  1990      2    17
## 2         Lord of the Flies   William Golding  1954     16     8
## 3           Grapes of Wrath    John Steinbeck  1939      5    10
## 4            The Waste Land        T.S. Eliot  1922     10     8
## 5                Siddhartha     Hermann Hesse  1922      4    22
## 6          Huckleberry Finn        Mark Twain  1884     21    12
## 7              Little Women Louisa May Alcott  1868      9    20
## 8                 Jane Eyre  Charlotte Bronte  1847     12    25
## 9              Frankenstein      Mary Shelley  1818     15     5
## 10      Pride and Prejudice       Jane Austen  1813      7    15
```

Sale

Your bookstore is having a sale!

- Books published before 1900 are 25% off
- Books with more than 10 copies in stock are 40% off
- Books published before 1900 with more than 10 copies in stock are 50% off

In the code block below, fill in the blanks to create a new data frame with the sale price for each book, in addition to the original price. You will need three ifelse statements, which can be either sequential or nested. Print your result.

```
sale <- bookstore %>% mutate(saleprice =
ifelse(pyear < 1900 & nstock > 10, price * 0.50,
ifelse(pyear < 1900, price * 0.75,
ifelse(nstock > 10, price * 0.60, price))))
sale
```

```
##                       title            author pyear nstock price saleprice
## 1           The Waste Land         T.S. Eliot  1922     10     8      8.00
## 2        Lord of the Flies    William Golding  1954     16     8      4.80
## 3          Grapes of Wrath     John Steinbeck  1939      5    10     10.00
## 4      Pride and Prejudice        Jane Austen  1813      7    15     11.25
## 5              Frankenstein       Mary Shelley  1818     15     5      2.50
## 6              Little Women Louisa May Alcott  1868      9    20     15.00
## 7                 Jane Eyre   Charlotte Bronte  1847     12    25     12.50
## 8  The Things They Carried        Tim O'Brien  1990      2    17     17.00
## 9         Huckleberry Finn         Mark Twain  1884     21    12      6.00
## 10              Siddhartha      Hermann Hesse  1922      4    22     22.00
```

Sale price check

Add a column to the data frame with the sale prices that indicates the percentage discount for each book. Do this by calculating the difference between sale price and original price. Print your result.

```
sale <- sale %>% mutate(percentdiscount =
(price - saleprice)/price)
sale
```

```
##                       title            author pyear nstock price saleprice
## 1           The Waste Land         T.S. Eliot  1922     10     8      8.00
## 2        Lord of the Flies    William Golding  1954     16     8      4.80
## 3          Grapes of Wrath     John Steinbeck  1939      5    10     10.00
## 4      Pride and Prejudice        Jane Austen  1813      7    15     11.25
## 5              Frankenstein       Mary Shelley  1818     15     5      2.50
## 6              Little Women Louisa May Alcott  1868      9    20     15.00
## 7                 Jane Eyre   Charlotte Bronte  1847     12    25     12.50
## 8  The Things They Carried        Tim O'Brien  1990      2    17     17.00
## 9         Huckleberry Finn         Mark Twain  1884     21    12      6.00
## 10              Siddhartha      Hermann Hesse  1922      4    22     22.00
```

```
##    percentdiscount
## 1             0.00
## 2             0.40
## 3             0.00
## 4             0.25
## 5             0.50
## 6             0.25
## 7             0.50
## 8             0.00
## 9             0.50
## 10            0.00
```

Sale list

Fill in the blanks below to create a new data frame that includes only books on sale and only author, title, and sale price. You will need to use the `filter()` and `select()` functions, and the order you use them in matters. Print your result.

```
#This is my first answer.
salebooks <- sale %>%
  filter(percentdiscount >= 0.01) %>%
  select(author, title, saleprice)
salebooks
```

```
##              author               title saleprice
## 1   William Golding   Lord of the Flies      4.80
## 2       Jane Austen Pride and Prejudice     11.25
## 3      Mary Shelley        Frankenstein      2.50
## 4 Louisa May Alcott        Little Women     15.00
## 5  Charlotte Bronte           Jane Eyre     12.50
## 6        Mark Twain    Huckleberry Finn      6.00
```

```
#Or is this the answer?
salebooks <- sale %>%
  filter(percentdiscount >= 0.01) %>%
  select(percentdiscount,author, title, saleprice)
salebooks
```

```
##   percentdiscount            author               title saleprice
## 1            0.40   William Golding   Lord of the Flies      4.80
## 2            0.25       Jane Austen Pride and Prejudice     11.25
## 3            0.50      Mary Shelley        Frankenstein      2.50
## 4            0.25 Louisa May Alcott        Little Women     15.00
## 5            0.50  Charlotte Bronte           Jane Eyre     12.50
## 6            0.50        Mark Twain    Huckleberry Finn      6.00
```

Turning in this assignment

Save and preview this file, then upload the `Lab1.nb.html` file to Canvas under Lab 1.