# Analysis 2: Joining Cholesterol with Crimes and Web Scraping Wikipedia

Sunghwu Song

March 18, 2021

## Instructions

**Overview:** For each question, show your R code that you used to answer each question in the provided chunks. When a written response is required, be sure to answer the entire question in complete sentences outside the code chunks. When figures are required, be sure to follow all requirements to receive full credit. Point values are assigned for every part of this analysis.

**Helpful:** Make sure you knit the document as you go through the assignment. Check all your results in the created PDF file.

**Submission:** Submit via an electronic document on Gradescope. Must be submitted as an PDF file generated in RStudio.

## Introduction

Does high cholesterol lead to high crime rates? Probably not, but web scraping will definitely lead to lower crime rates. This data analysis assignment is separated into three parts which cover material from the lectures on tidy data, joins, and webscraping. In Part 1, you will demonstrate the basic concept of joins by connecting relational data involving a cholesterol study. For this segment, `pivot_longer` and `pivot_wider` will be utilized to create a single tidy dataset ready for analysis. In Part 2, we will join all 5 datasets from the lecture series on web scraping. Part 3 will require an understanding of web scraping to import a table found on Wikipedia directly into R. The following R code reads in all datasets required for this assignment.

```
# Data for Part 1
CHOL1=read_csv("Cholesterol.csv")
CHOL2=read_csv("Cholesterol2.csv")

# Data for Part 2
VIOLENT=read_csv("FINAL_VIOLENT.csv")
ZIP=read_csv("FINAL_ZIP.csv")
STATE_ABBREV=read_csv("FINAL_STATE_ABBREV.csv")
CENSUS=read_csv("FINAL_CENSUS.csv")
S_VS_D=read_csv("FINAL_SAFE_VS_DANGEROUS.CSV")
```

## Assignment

## Part 1: Cholesterol Experiment

The data frame `CHOL1` contains experimental results from randomly assigning *18* people to one of two competing margarine brands "A" and "B". The cholesterol of these patients was measured once before using the margarine brand, once after *4* weeks with the margarine brand, and then again after *8* weeks with the margarine brand. Researchers want to see if there is benefit of these brands of margarine on reducing an individual's cholesterol and want to determine if there is a statistically significant difference between the two competing brands.

### Q1 *(3 Points)*

Start by examing the tables `CHOL1` and `CHOL2` and answering the following questions with *Yes* or *No* responses.

Is the variable `ID` in `CHOL1` a primary key?

Answer *(1 Point)*: Yes

Is the variable, `Margarine` in `CHOL1` a primary key?

Answer *(1 Point)*: No

Is the variable, `Brand` in `CHOL2` a primary key?

Answer *(1 Point)*: No

### Q2 *(2 Points)*

In a new data frame called `CHOL1a` based on `CHOL1`, rename the variables `After4weeks` and `After8weeks` to nonsynctactic variable names `4` and `8`, respectively. Use `names(CHOL1a)` to display this modification.

```
CHOL1a = rename(CHOL1, '4' = After4weeks, '8' = After8weeks)
names(CHOL1a)
```

```
## [1] "ID"        "Before"    "4"         "8"         "Margarine"
```

### Q3 *(4 Points)*

Use the `pivot_longer()` function or `gather()` function on `CHOL1a` to create a new numeric variable called `Week` that contains numeric values *4* or *8* and a new numeric variable called `Response` that contains the Cholesterol after the corresponding number of weeks. Create a new data frame called `CHOL1b` with these modifications and use `str(Chol1b)` to show that both variables have been created correctly and are indeed numeric (an integer variable is a specific type of numeric variable).

```
CHOL1b <- CHOL1a %>%
  gather("4","8", key = "Response", value = "Chloesterol", convert = TRUE)
str(CHOL1b)
```

```
## tibble [36 x 5] (S3: tbl_df/tbl/data.frame)
##  $ ID         : num [1:36] 1 2 3 4 5 6 7 8 9 10 ...
##  $ Before     : num [1:36] 6.42 6.76 6.56 4.8 8.43 7.49 8.05 5.05 5.77 3.91 ...
##  $ Margarine  : chr [1:36] "B" "A" "B" "A" ...
##  $ Response   : int [1:36] 4 4 4 4 4 4 4 4 4 4 ...
##  $ Chloesterol: num [1:36] 5.83 6.2 5.83 4.27 7.71 7.12 7.25 4.63 5.31 3.7 ...
```

### Q4 *(4 Points)*

Now working with `CHOL2`, we want to spread the variable `Statistic` across multiple columns. Do this in a new data frame called `CHOL2a` and use `print(CHOL2a)` to display the modified complete table.

```
CHOL2a <- CHOL2 %>%
  spread(key = Statistic, value = Value)

print(CHOL2a)
```

```
## # A tibble: 2 x 6
##   Brand Calories   Fat SatFat Serving Sodium
##   <chr>    <dbl> <dbl>  <dbl>   <dbl>  <dbl>
## 1 A           70     7    2.5      14    130
## 2 B           50     6    1.5      14     NA
```

### Q5 *(3 Points)*

Start by examing the tables `CHOL1b` and `CHOL2a` and answering the following questions with *Yes* or *No* responses.

Is the variable `ID` in `CHOL1b` a primary key?

Answer *(1 Point)*: No

Is the variable, `Margarine` in `CHOL1b` a primary key?

Answer *(1 Point)*: No

Is the variable, `Brand` in `CHOL2a` a primary key?

Answer *(1 Point)*: Yes

### Q6 *(4 Points)*

Get the nutritional facts of the different margarine brands in `CHOL2a` into the experimental results found in `CHOL1b` using a join. Create a new data frame named `CHOL.COMBINED` and display the table using `head(CHOL.COMBINED)`. This final data frame should contain *36* observations and *10* variables.

```
CHOL.COMBINED <- CHOL1b %>%
  inner_join(CHOL2a, by = c("Margarine" = "Brand"))

head(CHOL.COMBINED)
```

```
## # A tibble: 6 x 10
##      ID Before Margarine Response Chloesterol Calories   Fat SatFat Serving
##   <dbl>  <dbl> <chr>        <int>       <dbl>    <dbl> <dbl>  <dbl>   <dbl>
## 1     1   6.42 B                4        5.83       50     6    1.5      14
## 2     2   6.76 A                4        6.2        70     7    2.5      14
## 3     3   6.56 B                4        5.83       50     6    1.5      14
## 4     4   4.8  A                4        4.27       70     7    2.5      14
## 5     5   8.43 B                4        7.71       50     6    1.5      14
## 6     6   7.49 A                4        7.12       70     7    2.5      14
## # ... with 1 more variable: Sodium <dbl>
```

## Part 2: Linking Important Information to 2017 Violent Crimes Data

In the zipped folder, there are *5* CSV files. In this section, we are going to merge all of that data into one object called `FINAL.VIOLENT`.

### Q1 *(2 Points)*

The dataset `S_VS_D` contains a variable `CLASS` where "S=Safe" and "D=Dangerous" according to the article *These Are the 2018 Safest and Most Dangerous States in the U.S* by Steve Karantzoulidis. We seek to compare the violent crime statistics for states not in this list. Use a filtering join to create a new data frame called `VIOLENT2` that only contains violent crime statistics from the states not represented in the data frame `S_VS_D`. Use `str(VIOLENT2)` to display the variables and the dimensions of `VIOLENT2`.

```
VIOLENT2 = anti_join(VIOLENT, S_VS_D, by = c("State" = "STATE"))

str(VIOLENT2)
```

```
## tibble [68 x 8] (S3: spec_tbl_df/tbl_df/tbl/data.frame)
##  $ State     : chr [1:68] "Arizona" "Arizona" "Arizona" "Arizona" ...
##  $ City      : chr [1:68] "Chandler" "Gilbert" "Glendale" "Mesa" ...
##  $ Population: num [1:68] 249355 242090 249273 492268 1644177 ...
##  $ Total     : num [1:68] 259.5 85.5 488.2 415.8 760.9 ...
##  $ Murder    : num [1:68] 2.01 2.07 4.81 4.67 9.55 ...
##  $ Rape      : num [1:68] 52.1 16.1 38.9 51.2 69.5 ...
##  $ Robbery   : num [1:68] 57 21.1 193 92.2 200.3 ...
##  $ Assault   : num [1:68] 148.4 46.3 251.5 267.7 481.6 ...
##  - attr(*, "spec")=
##   .. cols(
##   ..   State = col_character(),
##   ..   City = col_character(),
##   ..   Population = col_double(),
##   ..   Total = col_double(),
##   ..   Murder = col_double(),
##   ..   Rape = col_double(),
##   ..   Robbery = col_double(),
##   ..   Assault = col_double()
##   .. )
```

### Q2 *(4 Points)*

Start by creating a new data set called `VIOLENT3` based on `VIOLENT2` that fixes some problems in the variable `City`. Specifically, we would like to change "Louisville Metro" to "Louisville".

Next, create a new data frame named `VIOLENT4` that connects the population change and density measures from 2019 contained in `CENSUS` to the cities and states in `VIOLENT3`. Use `head(VIOLENT4)` to give a preview of the new merged dataset.

Finally, in a complete sentence, identify any location(s) (Cities and States) missing violent crime information.

Code and Output *(2 Points)*:

```
VIOLENT3 <- VIOLENT2 %>%
  mutate(City=ifelse(City=="Louisville Metro", "Louisville", City))
```

```r
VIOLENT4 <- VIOLENT3 %>%
  left_join(CENSUS, by = c("City" = "Name", "State" = "State"))
head(VIOLENT4)
```

```
## # A tibble: 6 x 10
##   State   City      Population Total Murder  Rape Robbery Assault Change Density
##   <chr>   <chr>          <dbl> <dbl>  <dbl> <dbl>   <dbl>   <dbl>  <dbl>   <dbl>
## 1 Arizona Chandler      249355 259.    2.01  52.1    57.0   148.    5.55    1554
## 2 Arizona Gilbert       242090  85.5   2.07  16.1    21.1    46.3   7.21    1443
## 3 Arizona Glendale      249273 488.    4.81  38.9   193.    252.    2.58    1648
## 4 Arizona Mesa          492268 416.    4.67  51.2    92.2   268.    6.87    1450
## 5 Arizona Phoenix      1644177 761.    9.55  69.5   200.    482.    4.49    1254
## 6 Arizona Scottsda~     251840 157.    1.99  40.9    39.7    74.6   4.6      542
```

```r
which(is.na(VIOLENT4$Total))
```

```
## [1] 52 53 54 58
```

```r
VIOLENT4
```

```
## # A tibble: 68 x 10
##    State    City    Population Total Murder  Rape Robbery Assault Change Density
##    <chr>    <chr>        <dbl> <dbl>  <dbl> <dbl>   <dbl>   <dbl>  <dbl>   <dbl>
##  1 Arizona  Chandl~     249355 259.    2.01  52.1    57.0   148.    5.55    1554
##  2 Arizona  Gilbert     242090  85.5   2.07  16.1    21.1    46.3   7.21    1443
##  3 Arizona  Glenda~     249273 488.    4.81  38.9   193.    252.    2.58    1648
##  4 Arizona  Mesa        492268 416.    4.67  51.2    92.2   268.    6.87    1450
##  5 Arizona  Phoenix    1644177 761.    9.55  69.5   200.    482.    4.49    1254
##  6 Arizona  Scotts~     251840 157.    1.99  40.9    39.7    74.6   4.6      542
##  7 Arizona  Tucson      532323 802.    8.64  93.6   269.    431.    3.25     917
##  8 Califor~ Anaheim     353400 355.    2.83  32.5   136.    183.   -0.2     2706
##  9 Califor~ Bakers~     381154 479.   10.8   24.1   198.    247.    2.13     997
## 10 Califor~ Chula ~     271109 298.    0.74  22.9   112.    162.    2.7     2136
## # ... with 58 more rows
```

Answer *(2 Points)*: The cities missing violent crime information are those that match rows 52, 53, 54, and 58, which are Charlotte, Durham, Greensboro, and Toledo.

### Q3 *(6 Points)*

Either ambitiously using one step or less-ambitiously using multiple steps add the longitude and latitude information provided in `ZIP` to the cities and states in `VIOLENT4`. You will need to use `STATE_ABBREV` data to link these two data frames. Your final data frame named `FINAL.VIOLENT` should contain all of the information in `VIOLENT4` along with the variables `lat` and `lon` from `ZIP`. There should be **no** state abbreviations in `FINAL.VIOLENT` since this information is redundant. Use `str(FINAL.VIOLENT)` to demonstrate that everything worked as planned.

In `FINAL.VIOLENT` identify what cities are missing latitude and longitude. Closely, inspect both the `ZIP` and `VIOLENT4` data frames. Report the location(s) missing geographical information and explain in complete sentences why this happened.

Finally, challenge yourself and attempt to fix this problem in a new data frame called `FINAL.VIOLENT.FIX`. Use a combination of `str()` and `filter()` to only display the data in `FINAL.VIOLENT.FIX` for the location(s) that `FINAL.VIOLENT` was missing latitude and longitude. Do this in the second code chunk below.

Code and Output *(4 Points)*:

```
x = left_join(VIOLENT4, STATE_ABBREV, by = "State")
FINAL.VIOLENT = left_join(x, ZIP, by = c("City" = "city", "state" = "state")) %>%
  select(-state)

str(FINAL.VIOLENT)
```

```
## tibble [68 x 12] (S3: tbl_df/tbl/data.frame)
##  $ State     : chr [1:68] "Arizona" "Arizona" "Arizona" "Arizona" ...
##  $ City      : chr [1:68] "Chandler" "Gilbert" "Glendale" "Mesa" ...
##  $ Population: num [1:68] 249355 242090 249273 492268 1644177 ...
##  $ Total     : num [1:68] 259.5 85.5 488.2 415.8 760.9 ...
##  $ Murder    : num [1:68] 2.01 2.07 4.81 4.67 9.55 ...
##  $ Rape      : num [1:68] 52.1 16.1 38.9 51.2 69.5 ...
##  $ Robbery   : num [1:68] 57 21.1 193 92.2 200.3 ...
##  $ Assault   : num [1:68] 148.4 46.3 251.5 267.7 481.6 ...
##  $ Change    : num [1:68] 5.55 7.21 2.58 6.87 4.49 4.6 3.25 -0.2 2.13 2.7 ...
##  $ Density   : num [1:68] 1554 1443 1648 1450 1254 ...
##  $ lat       : num [1:68] 33.3 33.3 33.5 33.4 33.4 ...
##  $ lon       : num [1:68] -112 -112 -112 -112 -112 ...
```

```
FINAL.VIOLENT
```

```
## # A tibble: 68 x 12
##    State City  Population Total Murder  Rape Robbery Assault Change Density
##    <chr> <chr>      <dbl> <dbl>  <dbl> <dbl>   <dbl>   <dbl>  <dbl>   <dbl>
##  1 Ariz~ Chan~     249355 259.    2.01  52.1    57.0   148.    5.55    1554
##  2 Ariz~ Gilb~     242090  85.5   2.07  16.1    21.1    46.3   7.21    1443
##  3 Ariz~ Glen~     249273 488.    4.81  38.9   193.    252.    2.58    1648
##  4 Ariz~ Mesa      492268 416.    4.67  51.2    92.2   268.    6.87    1450
##  5 Ariz~ Phoe~    1644177 761.    9.55  69.5   200.    482.    4.49    1254
##  6 Ariz~ Scot~     251840 157.    1.99  40.9    39.7    74.6   4.6      542
##  7 Ariz~ Tucs~     532323 802.    8.64  93.6   269.    431.    3.25     917
##  8 Cali~ Anah~     353400 355.    2.83  32.5   136.    183.   -0.2     2706
##  9 Cali~ Bake~     381154 479.   10.8   24.1   198.    247.    2.13     997
## 10 Cali~ Chul~     271109 298.    0.74  22.9   112.    162.    2.7     2136
## # ... with 58 more rows, and 2 more variables: lat <dbl>, lon <dbl>
```

Answer *(1 Points)*: Washington D.C. lacks the values for latitude and longitude; I think it is because its state is the District of Columbia but D.C. part in the name created an error when using the ZIP data.

Code and Output *(1 Point)*:

```
y = x %>%
  mutate(City = ifelse(City == "Washington DC", "Washington", City)) %>%
  mutate(state = ifelse(is.na(state), "DC", state))

z = left_join(y, ZIP, by = c("City" = "city", "state"))
```

```
FINAL.VIOLENT.FIX = filter(z, City %in% "Washington")

str(FINAL.VIOLENT.FIX)
```

```
## tibble [1 x 13] (S3: spec_tbl_df/tbl_df/tbl/data.frame)
##  $ State     : chr "District of Columbia"
##  $ City      : chr "Washington"
##  $ Population: num 693972
##  $ Total     : num 949
##  $ Murder    : num 16.7
##  $ Rape      : num 63.8
##  $ Robbery   : num 339
##  $ Assault   : num 529
##  $ Change    : num 3.65
##  $ Density   : num 4461
##  $ state     : chr "DC"
##  $ lat       : num 38.9
##  $ lon       : num -77
##  - attr(*, "spec")=
##   .. cols(
##   ..   State = col_character(),
##   ..   City = col_character(),
##   ..   Population = col_double(),
##   ..   Total = col_double(),
##   ..   Murder = col_double(),
##   ..   Rape = col_double(),
##   ..   Robbery = col_double(),
##   ..   Assault = col_double()
##   .. )
```

## Part 3: Web Scraping a Table From Wikipedia

Wikipedia contains a rough estimate of a billion tables. Search through Wikipedia pages and identify an article, completely unrelated to crimes data, that contains an HTML table that has at least *5* rows and *3* columns. You will be required to web scrape the table into a data frame or tibble into R. This portion will require a minor knowledge of the `rvest` package. Utilize information from the web scraping lectures and tutorials to assist you with this.

### Q1 *(4 Points)*

What is the URL of the Wikipedia page you plan on webscraping (Knit the Document and Check the Hyperlink)?

Answer *(2 Points)*: https://en.wikipedia.org/wiki/List_of_Crayola_crayon_colors

In *2* to *5* sentences, Identify and describe the specific table you plan on web scraping. State the variables in *1* of the sentences.

Answer *(2 Points)*: The table I plan to web scrape is the standard list of colors for crayola crayon colors. It shows a table with the following columns: photo showing the color, name of the color, its hexdecimal code, years in production, any notes, and whether they are in the 16-box, 24-box, and/or the 64-box.

**Q2** ***(4 Points)***

Utilize the functions `read_html()` and `html_table()` to web scrape the specific table you described above. Internet access will be required for these functions to work. Create an R data frame named `DATA` which contains the information from the Wikipedia table. All code should be contained in the R code chunk below. Finally, use the `print()` function to display the table to demonstrate that everything worked as planned. The variable names and the content should match the table on the Wikipedia page you chose exactly. You are not required to perform any cleaning of this data. As long as the content of the table you describe matches `DATA`, then you are good. Don't worry if the table bleeds over multiple pages.

```
URL.DATA = "https://en.wikipedia.org/wiki/List_of_Crayola_crayon_colors"
DATA = URL.DATA %>%
  read_html() %>%
  html_table(fill=T) %>%
  .[[2]]

print(DATA)
```

```
##     Color                     Name Hexadecimal in their website depiction[b]
## 1      NA                      Red                                 #ED0A3F[1]
## 2      NA                   Maroon                                 #C32148[1]
## 3      NA                  Scarlet                                 #FD0E35[1]
## 4      NA                Brick Red                                 #C62D42[1]
## 5      NA        English Vermilion
## 6      NA              Madder Lake
## 7      NA  Permanent Geranium Lake
## 8      NA              Maximum Red
## 9      NA                 Chestnut                                 #B94E48[1]
## 10     NA               Orange-Red                                 #FF5349[1]
## 11     NA            Sunset Orange                                 #FE4C40[1]
## 12     NA              Bittersweet                                 #FE6F5E[1]
## 13     NA        Dark Venetian Red
## 14     NA             Venetian Red
## 15     NA       Light Venetian Red
## 16     NA          Vivid Tangerine                                 #FF9980[1]
## 17     NA               Middle Red
## 18     NA             Burnt Orange                                 #FF7034[1]
## 19     NA               Red-Orange                                 #FF681F[1]
## 20     NA                   Orange                                 #FF8833[1]
## 21     NA      Macaroni and Cheese                                 #FFB97B[1]
## 22     NA        Middle Yellow Red
## 23     NA              Mango Tango                                 #E77200[1]
## 24     NA            Yellow-Orange                                 #FFAE42[1]
## 25     NA       Maximum Yellow Red
## 26     NA             Banana Mania                                 #FBE7B2[1]
## 27     NA                    Maize
## 28     NA            Orange-Yellow                                 #F8D568[1]
## 29     NA                Goldenrod                                 #FCD667[1]
## 30     NA                Dandelion                                 #FED85D[1]
## 31     NA                   Yellow                                 #FBE870[1]
## 32     NA             Green-Yellow                                 #F1E788[1]
## 33     NA            Middle Yellow
## 34     NA              Olive Green                                 #B5B35C[1]
## 35     NA             Spring Green                                 #ECEBBD[1]
```

```
## 36     NA          Maximum Yellow
## 37     NA                  Canary                                 #FFFF99[1]
## 38     NA            Lemon Yellow
## 39     NA    Maximum Green Yellow
## 40     NA     Middle Green Yellow
## 41     NA                Inchworm                                 #AFE313[1]
## 42     NA      Light Chrome Green
## 43     NA            Yellow-Green                                 #C5E17A[1]
## 44     NA           Maximum Green
## 45     NA               Asparagus                                 #7BA05B[1]
## 46     NA      Granny Smith Apple                                 #9DE093[1]
## 47     NA                    Fern                                 #63B76C[1]
## 48     NA            Middle Green
## 49     NA                   Green                                 #01A638[1]
## 50     NA     Medium Chrome Green
## 51     NA            Forest Green                                 #5FA777[1]
## 52     NA               Sea Green                                 #93DFB8[1]
## 53     NA                Shamrock                                 #33CC99[1]
## 54     NA          Mountain Meadow                                #1AB385[1]
## 55     NA            Jungle Green                                 #29AB87[1]
## 56     NA          Caribbean Green                                #00CC99[1]
## 57     NA    Tropical Rain Forest                                 #00755E[1]
## 58     NA       Middle Blue Green
## 59     NA              Pine Green                                 #01786F[1]
## 60     NA      Maximum Blue Green
## 61     NA        Robin's Egg Blue                                 #00CCCC[1]
## 62     NA               Teal Blue                                 #008080[1]
## 63     NA              Light Blue                                 #8FD8D8[1]
## 64     NA              Aquamarine
## 65     NA          Turquoise Blue                                 #6CDAE7[1]
## 66     NA             Outer Space                                 #2D383A[1]
## 67     NA                Sky Blue                                 #76D7EA[1]
## 68     NA             Middle Blue
## 69     NA              Blue-Green                                 #0095B7[1]
## 70     NA            Pacific Blue                                 #009DC4[1]
## 71     NA                Cerulean                                 #02A4D3[1]
## 72     NA            Maximum Blue
## 73     NA                Blue (I)
## 74     NA           Cerulean Blue
## 75     NA              Cornflower                                 #93CCEA[1]
## 76     NA              Green-Blue
## 77     NA           Midnight Blue                                 #003366[1]
## 78     NA               Navy Blue                                 #0066CC[1]
## 79     NA                   Denim                                 #1560BD[1]
## 80     NA              Blue (III)                                 #0066FF[1]
## 81     NA              Cadet Blue                                 #A9B2C3[1]
## 82     NA              Periwinkle                                 #C3CDE6[1]
## 83     NA               Blue (II)
## 84     NA               Bluetiful
## 85     NA        Wild Blue Yonder                                 #7A89B8[1]
## 86     NA                  Indigo                                 #4F69C6[1]
## 87     NA                 Manatee                                 #8D90A1[1]
## 88     NA             Cobalt Blue
## 89     NA          Celestial Blue
```

```
## 90     NA                Blue Bell                              #9999CC[1]
## 91     NA      Maximum Blue Purple
## 92     NA              Violet-Blue
## 93     NA              Blue-Violet                              #6456B7[1]
## 94     NA         Ultramarine Blue
## 95     NA       Middle Blue Purple
## 96     NA             Purple Heart                              #652DC1[1]
## 97     NA             Royal Purple                              #6B3FA0[1]
## 98     NA              Violet (II)                              #8359A3[1]
## 99     NA            Medium Violet
## 100    NA                 Wisteria                              #C9A0DC[1]
## 101    NA             Lavender (I)
## 102    NA             Vivid Violet                              #803790[1]
## 103    NA           Maximum Purple
## 104    NA Purple Mountains' Majesty                             #D6AEDD[1]
## 105    NA                  Fuchsia                              #C154C1[1]
## 106    NA            Pink Flamingo                              #FC74FD[1]
## 107    NA               Violet (I)
## 108    NA           Brilliant Rose
## 109    NA                   Orchid                              #E29CD2[1]
## 110    NA                     Plum                              #8E3179[1]
## 111    NA              Medium Rose
## 112    NA                  Thistle                              #D8BFD8[1]
## 113    NA                 Mulberry
## 114    NA               Red-Violet                              #BB3385[1]
## 115    NA            Middle Purple
## 116    NA       Maximum Red Purple
## 117    NA            Jazzberry Jam                              #A50B5E[1]
## 118    NA                 Eggplant                              #614051[1]
## 119    NA                  Magenta                              #F653A6[1]
## 120    NA                   Cerise                              #DA3287[1]
## 121    NA          Wild Strawberry                              #FF3399[1]
## 122    NA            Lavender (II)                              #FBAED2[1]
## 123    NA             Cotton Candy                              #FFB7D5[1]
## 124    NA           Carnation Pink                              #FFA6C9[1]
## 125    NA               Violet-Red                              #F7468A[1]
## 126    NA               Razzmatazz                              #E30B5C[1]
## 127    NA               Piggy Pink                              #FDD7E4[1]
## 128    NA                  Carmine
## 129    NA                    Blush                              #DB5079[1]
## 130    NA           Tickle Me Pink                              #FC80A5[1]
## 131    NA                Mauvelous                              #F091A9[1]
## 132    NA                   Salmon                              #FF91A4[1]
## 133    NA        Middle Red Purple
## 134    NA                 Mahogany                              #CA3435[1]
## 135    NA                    Melon                              #FEBAAD[1]
## 136    NA            Pink Sherbert                              #F7A38E[1]
## 137    NA             Burnt Sienna                              #E97451[1]
## 138    NA                    Brown                              #AF593E[1]
## 139    NA                    Sepia                              #9E5B40[1]
## 140    NA              Fuzzy Wuzzy                              #87421F[1]
## 141    NA                   Beaver                              #926F5B[1]
## 142    NA               Tumbleweed                              #DEA681[1]
## 143    NA               Raw Sienna                              #D27D46[1]
```

```
## 144    NA            Van Dyke Brown
## 145    NA                       Tan
## 146    NA               Desert Sand                                  #EDC9AF[1]
## 147    NA                     Peach                                  #FFCBA4[1]
## 148    NA               Burnt Umber
## 149    NA                   Apricot                                  #FDD5B1[1]
## 150    NA                    Almond                                  #EED9C4[1]
## 151    NA                 Raw Umber
## 152    NA                    Shadow                                  #837050[1]
## 153    NA            Raw Sienna (I)
## 154    NA                Timberwolf                                  #D9D6CF[1]
## 155    NA                  Gold (I)
## 156    NA                 Gold (II)                                  #E6BE8A[1]
## 157    NA                    Silver                                  #C9C0BB[1]
## 158    NA                    Copper                                  #DA8A67[1]
## 159    NA             Antique Brass                                  #C88A65[1]
## 160    NA                     Black                                  #000000[1]
## 161    NA             Charcoal Gray
## 162    NA                      Gray                                  #8B8680[1]
## 163    NA                 Blue-Gray                                  #C8C8CD[1]
## 164    NA                     White                                  #FFFFFF[1]
##       Years in production[2]
## 1             1903-present
## 2             1949-present
## 3             1998-present
## 4             1958-present
## 5                1903-1935
## 6                1903-1935
## 7          1903-circa 1910
## 8                1926-1944
## 9             1903-present
## 10               1958-1990
## 11            1997-present
## 12            1958-present
## 13         1903-circa 1910
## 14               1903-1944
## 15         1903-circa 1910
## 16            1990-present
## 17               1926-1944
## 18            1958-present
## 19            1930-present
## 20            1903-present
## 21            1993-present
## 22     1926-1944, 1949-1958
## 23            2003-present
## 24            1930-present
## 25               1926-1944
## 26            1998-present
## 27               1903-1990
## 28               1958-1990
## 29            1903-present
## 30       1990-2017[2][3][4]
## 31            1903-present
## 32            1958-present
```

```
## 33                1926-1944
## 34             1903-present
## 35             1958-present
## 36                1926-1944
## 37             1998-present
## 38                1903-1990
## 39                1926-1944
## 40                1926-1944
## 41             2003-present
## 42                1903-1935
## 43             1930-present
## 44                1926-1944
## 45             1993-present
## 46             1993-present
## 47             1998-present
## 48                1926-1944
## 49             1903-present
## 50                1903-1939
## 51             1949-present
## 52             1949-present
## 53             1993-present
## 54             1998-present
## 55             1990-present
## 56             1997-present
## 57             1993-present
## 58                1926-1944
## 59  1903-1949, 1958-present
## 60                1926-1944
## 61             1993-present
## 62                1990-2003
## 63                     1958
## 64             1949-present
## 65             1935-present
## 66             1998-present
## 67             1958-present
## 68                1926-1944
## 69             1949-present
## 70             1993-present
## 71             1990-present
## 72                1926-1958
## 73                1903-1958
## 74                1949-1958
## 75             1958-present
## 76                1958-1990
## 77             1903-present
## 78             1958-present
## 79             1993-present
## 80             1949-present
## 81             1958-present
## 82             1958-present
## 83                1935-1958
## 84          2017-present[5]
## 85             2003-present
## 86             1999-present
```

```
## 87            1998-present
## 88               1903-1958
## 89         1903-circa 1910
## 90            1998-present
## 91               1926-1944
## 92         1903-circa 1910
## 93            1949-present
## 94               1903-1944
## 95               1926-1944
## 96            1998-present
## 97            1990-present
## 98  1930-1949, 1958-present
## 99               1949-1958
## 100           1993-present
## 101              1949-1958
## 102           1997-present
## 103              1926-1944
## 104           1993-present
## 105           1990-present
## 106           1997-present
## 107              1903-1930
## 108              1949-1958
## 109           1949-present
## 110           1958-present
## 111              1949-1958
## 112              1949-1999
## 113              1958-2003
## 114           1930-present
## 115              1926-1944
## 116              1926-1944
## 117           2003-present
## 118           1998-present
## 119           1903-present
## 120           1993-present
## 121           1990-present
## 122           1958-present
## 123           1998-present
## 124           1903-present
## 125           1958-present
## 126           1993-present
## 127           1998-present
## 128              1935-1958
## 129           1998-present
## 130           1993-present
## 131           1993-present
## 132           1949-present
## 133              1926-1944
## 134           1949-present
## 135           1958-present
## 136           1998-present
## 137           1903-present
## 138           1935-present
## 139 1935-1944, 1958-present
## 140           1998-present
```

```
## 141            1998-present
## 142            1993-present
## 143            1958-present
## 144               1903-1935
## 145            1958-present
## 146            1998-present
## 147            1903-present
## 148               1903-1944
## 149            1958-present
## 150            1998-present
## 151               1903-1990
## 152            1998-present
## 153         1903-circa 1910
## 154            1993-present
## 155               1903-1944
## 156            1953-present
## 157            1903-present
## 158 1903-1915, 1958-present
## 159            1998-present
## 160            1903-present
## 161               1903-1910
## 162            1926-present
## 163               1958-1990
## 164            1903-present
##
## 1
## 2
## 3
## 4
## 5
## 6
## 7
## 8
## 9
## 10
## 11
## 12
## 13
## 14
## 15
## 16
## 17
## 18
## 19
## 20
## 21                                                                      Also found as
## 22                                                   Part of the Munsell line, 1926-1944
## 23
## 24
## 25
## 26
## 27                                                                   Known as "Gold Ochre
## 28
## 29                                                             Known as "Medium Chrome Yello
```

```
## 30
## 31
## 32
## 33
## 34
## 35
## 36
## 37
## 38       Also known as "Light Chrome Yellow" (on labels "Chrome Yellow, Light") or "Light Yellow",
## 39
## 40
## 41
## 42                                                               "Chrome Green, Light" on labe
## 43
## 44
## 45
## 46
## 47
## 48
## 49
## 50                                         "Chrome Green, Medium" on labels. Produced 1903-1939
## 51
## 52
## 53
## 54
## 55
## 56
## 57
## 58
## 59                                               Known as "Dark Chrome Green" ("Chrome Green,
## 60
## 61
## 62
## 63
## 64
## 65
## 66
## 67
## 68
## 69
## 70
## 71
## 72                                                              Part of the Munsell line, 1926-
## 73                                                                     Known as "Celestial Bl
## 74
## 75
## 76
## 77
## 78
## 79
## 80
## 81
## 82
## 83
```

```
## 84
## 85
## 86
## 87
## 88
## 89
## 90
## 91
## 92
## 93
## 94
## 95
## 96
## 97
## 98
## 99
## 100
## 101
## 102
## 103
## 104                                                                 Also found as "Purple Mountai
## 105
## 106
## 107
## 108
## 109
## 110
## 111
## 112
## 113
## 114
## 115
## 116
## 117
## 118
## 119
## 120
## 121
## 122
## 123
## 124                                                                              Known as "Ro
## 125
## 126
## 127
## 128
## 129
## 130
## 131
## 132
## 133
## 134
## 135
## 136
## 137
```

```
## 138
## 139
## 140
## 141
## 142
## 143
## 144
## 145
## 146
## 147                                             Known as "Flesh Tint" (1903-1949), "Flesh" (1949-1956
## 148
## 149
## 150
## 151
## 152
## 153
## 154
## 155                                                             Metallic; swatch represents nominal 
## 156                                                             Metallic; swatch represents nominal 
## 157                                                             Metallic; swatch represents nominal 
## 158                                                                                                  M
## 159                                                                                                  M
## 160
## 161
## 162 As "Middle Grey", part of the Munsell line, 1926-1944. Spelled "Grey" on labels, but "Gray" on b
## 163
## 164
##     16-Box 24-Box 64-Box
## 1      yes    yes    yes
## 2       no     no     no
## 3       no    yes    yes
## 4       no     no    yes
## 5
## 6
## 7
## 8
## 9       no     no    yes
## 10
## 11      no     no     no
## 12      no     no    yes
## 13
## 14
## 15
## 16      no     no     no
## 17
## 18      no     no    yes
## 19     yes    yes    yes
## 20     yes    yes    yes
## 21      no     no    yes
## 22
## 23      no     no     no
## 24     yes    yes    yes
## 25
## 26      no     no     no
```

```
## 27
## 28
## 29      no     no    yes
## 30      no  * yes  * yes
## 31     yes    yes    yes
## 32      no    yes    yes
## 33
## 34      no     no    yes
## 35      no     no    yes
## 36
## 37      no     no     no
## 38
## 39
## 40
## 41      no     no     no
## 42
## 43     yes    yes    yes
## 44
## 45      no     no    yes
## 46      no     no    yes
## 47      no     no     no
## 48
## 49     yes    yes    yes
## 50
## 51      no     no    yes
## 52      no     no    yes
## 53      no     no     no
## 54      no     no     no
## 55      no     no     no
## 56      no     no     no
## 57      no     no     no
## 58
## 59      no     no     no
## 60
## 61      no     no    yes
## 62
## 63
## 64      no     no     no
## 65      no     no    yes
## 66      no     no     no
## 67      no     no    yes
## 68
## 69     yes    yes    yes
## 70      no     no    yes
## 71      no    yes    yes
## 72
## 73
## 74
## 75      no     no    yes
## 76
## 77      no     no     no
## 78      no     no     no
## 79      no     no     no
## 80     yes    yes    yes
```

```
## 81       no     no    yes
## 82       no     no    yes
## 83
## 84       no  yes *  yes *
## 85       no     no     no
## 86       no    yes    yes
## 87       no     no     no
## 88
## 89
## 90       no     no     no
## 91
## 92
## 93      yes    yes    yes
## 94
## 95
## 96       no     no     no
## 97       no     no     no
## 98      yes    yes    yes
## 99
## 100      no     no    yes
## 101
## 102      no     no     no
## 103
## 104      no     no    yes
## 105      no     no     no
## 106      no     no     no
## 107
## 108
## 109      no     no    yes
## 110      no     no    yes
## 111
## 112
## 113
## 114     yes    yes    yes
## 115
## 116
## 117      no     no     no
## 118      no     no     no
## 119      no     no    yes
## 120      no     no     no
## 121      no     no    yes
## 122      no     no    yes
## 123      no     no     no
## 124     yes    yes    yes
## 125      no    yes    yes
## 126      no     no     no
## 127      no     no     no
## 128
## 129      no     no     no
## 130      no     no    yes
## 131      no     no    yes
## 132      no     no    yes
## 133
## 134      no     no    yes
```

```
## 135    no    no   yes
## 136    no    no    no
## 137    no    no   yes
## 138   yes   yes   yes
## 139    no    no   yes
## 140    no    no    no
## 141    no    no    no
## 142    no    no   yes
## 143    no    no   yes
## 144
## 145    no    no   yes
## 146    no    no    no
## 147    no    no   yes
## 148
## 149    no   yes   yes
## 150    no    no    no
## 151
## 152    no    no    no
## 153
## 154    no    no   yes
## 155
## 156    no    no   yes
## 157    no    no   yes
## 158    no    no    no
## 159    no    no    no
## 160   yes   yes   yes
## 161
## 162    no   yes   yes
## 163
## 164   yes   yes   yes
```