

WS #7 - Factor variables

Wednesday, September 25, 2025

DS 002R - Jo Hardin

Name: _____

Names of people you worked with: _____

Do you go to the mentor sessions? How can you get the most out of the mentor sessions?

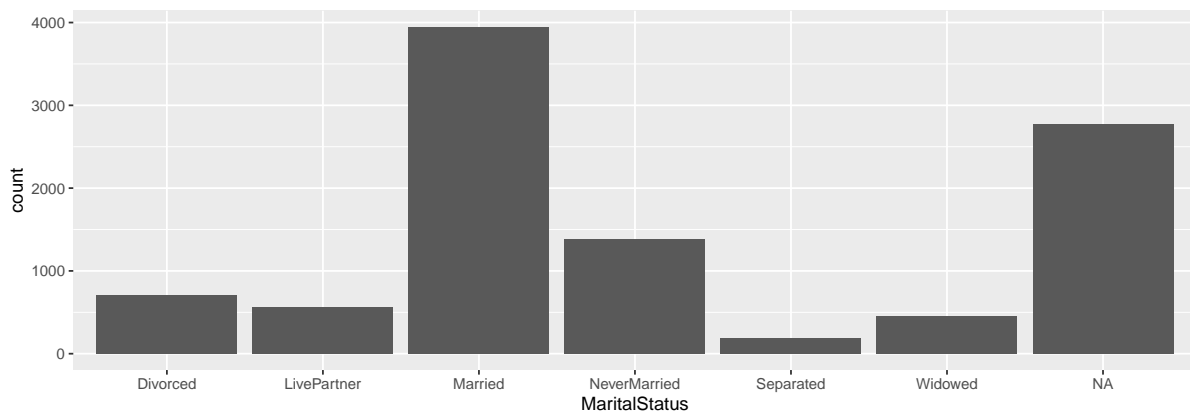
Task:

1. How would you order `MaritalStatus` in a more meaningful way?
2. Fill in the last 4 columns of the table below as if for new additional rows with `MaritalStatus` that contain the values `Divorced`, `LivePartner`, `Separated`, and `Widowed`?

```
select(NHANES, MaritalStatus) |> pull() |> levels()
```

```
[1] "Divorced"      "LivePartner"  "Married"      "NeverMarried"  
[5] "Separated"    "Widowed"
```

```
ggplot(NHANES, aes(x = MaritalStatus)) +  
  geom_bar()
```



```
NHANES |>
  select(MaritalStatus) |>
  sample_n(10) |>
  mutate(Mar_num = as.numeric(MaritalStatus),
         Mar_num_1 = as.numeric(MaritalStatus) + 1,
         Mar_char = as.character(MaritalStatus),
         Mar_change = ifelse(MaritalStatus == "Married", "Lol", MaritalStatus))
```

```
# A tibble: 10 x 5
  MaritalStatus Mar_num Mar_num_1 Mar_char    Mar_change
  <fct>         <dbl>    <dbl> <chr>      <chr>
1 NeverMarried     4        5 NeverMarried 4
2 Married          3        4 Married     Lol
3 Married          3        4 Married     Lol
4 NeverMarried     4        5 NeverMarried 4
5 <NA>             NA       NA <NA>      <NA>
6 Married          3        4 Married     Lol
7 Married          3        4 Married     Lol
8 <NA>             NA       NA <NA>      <NA>
9 Married          3        4 Married     Lol
10 Married         3        4 Married     Lol
```

Solution:

1. I'm not sure that there is a single right answer, but one potential ordering is:

NeverMarried, LivePartner, Married, Separated, Divorced, Widowed

2. Here are two tables with the relevant information on the 4 levels: Divorced, LivePartner, Separated, and Widowed.

```
set.seed(5)
NHANES |> select(MaritalStatus) |> sample_n(3) |>
  mutate(Mar_num = as.numeric(MaritalStatus),
         Mar_num_1 = as.numeric(MaritalStatus) + 1,
         Mar_char = as.character(MaritalStatus),
         Mar_change = ifelse(MaritalStatus == "Married", "Lol", MaritalStatus))
```

A tibble: 3 x 5

	MaritalStatus	Mar_num	Mar_num_1	Mar_char	Mar_change
	<fct>	<dbl>	<dbl>	<chr>	<int>
1	<NA>	NA	NA	<NA>	NA
2	NeverMarried	4	5	NeverMarried	4
3	LivePartner	2	3	LivePartner	2

```
set.seed(14)
NHANES |> select(MaritalStatus) |> sample_n(10) |>
  mutate(Mar_num = as.numeric(MaritalStatus),
         Mar_num_1 = as.numeric(MaritalStatus) + 1,
         Mar_char = as.character(MaritalStatus),
         Mar_change = ifelse(MaritalStatus == "Married", "Lol", MaritalStatus))
```

A tibble: 10 x 5

	MaritalStatus	Mar_num	Mar_num_1	Mar_char	Mar_change
	<fct>	<dbl>	<dbl>	<chr>	<chr>
1	Married	3	4	Married	Lol
2	Married	3	4	Married	Lol
3	Married	3	4	Married	Lol
4	Separated	5	6	Separated	5
5	Divorced	1	2	Divorced	1
6	Married	3	4	Married	Lol
7	Married	3	4	Married	Lol
8	Married	3	4	Married	Lol
9	<NA>	NA	NA	<NA>	<NA>
10	Widowed	6	7	Widowed	6