

```
In [1]: library(tidyverse)

# Attaching packages: tidyverse 1.3.0
ggplot2 3.3.3    purrr  0.3.4
tibble  3.0.6    dplyr  1.0.4
tidyr   1.1.2    stringr 1.4.0
readr   1.4.0    forcats 0.5.1

# Conflicts: tidyverse_conflicts()
* dplyr::filter() masks stats::filter()
* dplyr::lag()    masks stats::lag()
```

Warm-up 1

- Use `pivot_longer()` and `pivot_wider()` to convert one table into the other.

```
In [2]: phone_nos1 <- tribble(
#-----/-----/-----
  ~name,      ~"home_phone",  ~"cell_phone",
  ~"John Smith", ~"123-124-1111", ~"121-121-3212",
  ~"Jane Doe",   ~"321-421-7777", ~"431-456-1290",
  ~"Hadley Wickman", ~"111-222-8888", ~"121-111-9876",
)

phone_nos2 <- tribble(
#-----/-----/-----
  ~name,      ~"line_type",  ~"phone_nos",
  ~"John Smith", ~"home_phone", ~"123-124-1111",
  ~"John Smith", ~"cell_phone", ~"121-121-3212",
  ~"Jane Doe",   ~"home_phone", ~"321-421-7777",
  ~"Jane Doe",   ~"cell_phone", ~"431-456-1290",
  ~"Hadley Wickman", ~"home_phone", ~"111-222-8888",
  ~"Hadley Wickman", ~"cell_phone", ~"121-111-9876",
)

phone_nos1 %>%
  pivot_longer(c("home_phone", "cell_phone"), names_to = "line_type", values_to = "phone_nos")
```

A tibble: 6 × 3

name	line_type	phone_nos
<chr>	<chr>	<chr>
John Smith	home_phone	123-124-1111
John Smith	cell_phone	121-121-3212
Jane Doe	home_phone	321-421-7777
Jane Doe	cell_phone	431-456-1290
Hadley Wickman	home_phone	111-222-8888
Hadley Wickman	cell_phone	121-111-9876

```
In [5]: phone_nos2 %>%
  pivot_wider(names_from = line_type, values_from = phone_nos)
```

A tibble: 3 × 3

name	home_phone	cell_phone
<chr>	<chr>	<chr>
John Smith	123-124-1111	121-121-3212
Jane Doe	321-421-7777	431-456-1290
Hadley Wickman	111-222-8888	121-111-9876

Warm-up 2

- The following is a 2-way frequency table describing patient data
- Use `ggplot` to make a bar graph of how many patients are in each medication group, each bar filled in by age status.
 - Hint: Remember to use `stat = 'identity'` to override the y-value of a `geom_bar` layer.

```
In [7]: (patient_data <- tribble(
#-----/-----/-----
  ~medication, ~"child", ~"adult",
  ~"placebo",   ~25,      ~28,
  ~"medication1", ~14,      ~13,
  ~"medication2", ~15,      ~16
))
```

A tibble: 3 × 3

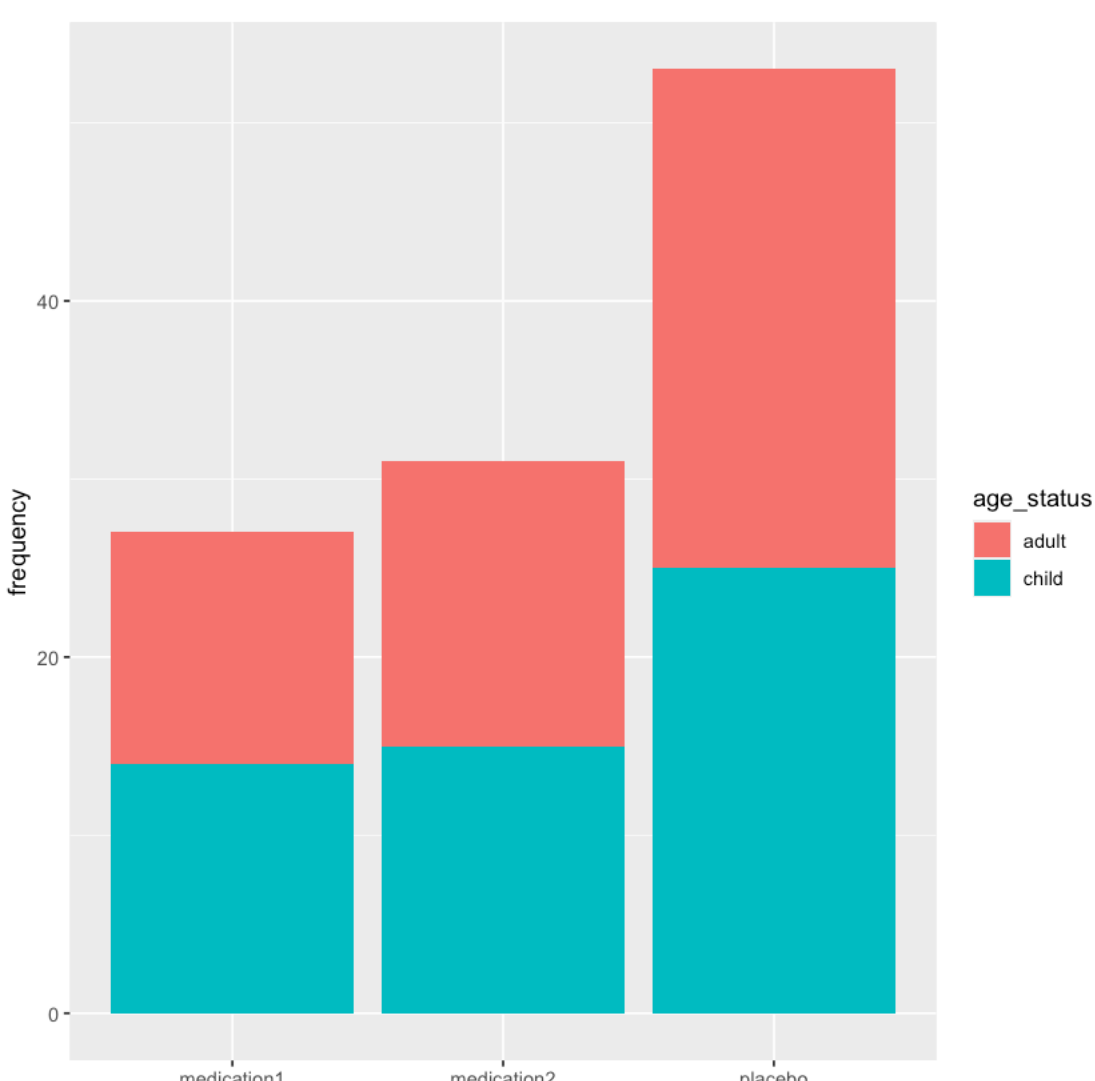
medication	child	adult
<chr>	<dbl>	<dbl>
placebo	25	28
medication1	14	13
medication2	15	16

```
In [10]: patient_data %>%
  pivot_longer(c("child", "adult"), names_to = "age_status", values_to = "frequency")
```

A tibble: 6 × 3

medication	age_status	frequency
<chr>	<chr>	<dbl>
placebo	child	25
placebo	adult	28
medication1	child	14
medication1	adult	13
medication2	child	15
medication2	adult	16

```
In [12]: patient_data %>%
  pivot_longer(c("child", "adult"), names_to = "age_status", values_to = "frequency") %>%
  ggplot(aes(x = medication, y = frequency)) +
  geom_bar(aes(fill = age_status), stat = "identity")
```



```
In [ ]: 
```

- Look at the `who` dataset
- Is this dataset tidy? Why or why not?
- Is there any redundant information?
- Hint: don't forget about using `?`

```
In [16]: who %>%
  select(-c("iso2", "iso3"))
```

A tibble: 7240 × 58

country	year	new_sp_m014	new_sp_m1524	new_sp_m2534	new_sp_m3544	new_sp_m4554	new_sp_m5564	new_sp_m65	new_sp_f014	...	newrel_m4554	newrel_m5564	newrel_m65
<chr>	<int>	<int>	<int>	<int>	<int>	<int>	<int>	<int>	<int>	<int>	<int>	<int>	<int>
Afghanistan	1980	NA	NA	NA	NA	NA	NA	NA	NA	...	NA	NA	NA
Afghanistan	1981	NA	NA	NA	NA	NA	NA	NA	NA	...	NA	NA	NA
Afghanistan	1982	NA	NA	NA	NA	NA	NA	NA	NA	...	NA	NA	NA
Afghanistan	1983	NA	NA	NA	NA	NA	NA	NA	NA	...	NA	NA	NA
Afghanistan	1984	NA	NA	NA	NA	NA	NA	NA	NA	...	NA	NA	NA
Afghanistan	1985	NA	NA	NA	NA	NA	NA	NA	NA	...	NA	NA	NA
Afghanistan	1986	NA	NA	NA	NA	NA	NA	NA	NA	...	NA	NA	NA
Afghanistan	1987	NA	NA	NA	NA	NA	NA	NA	NA	...	NA	NA	NA
Afghanistan	1988	NA	NA	NA	NA	NA	NA	NA	NA	...	NA	NA	NA
Afghanistan	1989	NA	NA	NA	NA	NA	NA	NA	NA	...	NA	NA	NA
Afghanistan	1990	NA	NA	NA	NA	NA	NA	NA	NA	...	NA	NA	NA
Afghanistan	1991	NA	NA	NA	NA	NA	NA	NA	NA	...	NA	NA	NA
Afghanistan	1992	NA	NA	NA	NA	NA	NA	NA	NA	...	NA	NA	NA
Afghanistan	1993	NA	NA	NA	NA	NA	NA	NA	NA	...	NA	NA	NA
Afghanistan	1994	NA	NA	NA	NA	NA	NA	NA	NA	...	NA	NA	NA
Afghanistan	1995	NA	NA	NA	NA	NA	NA	NA	NA	...	NA	NA	NA
Afghanistan	1996	NA	NA	NA	NA	NA	NA	NA	NA	...	NA	NA	NA
Afghanistan	1997	0	10	6	3	5	2	0	5	...	NA	NA	NA
Afghanistan	1998	30	129	128	90	89	64	41	45	...	NA	NA	NA
Afghanistan	1999	8	55	55	47	34	21	8	25	...	NA	NA	NA
Afghanistan	2000	52	228	183	149	129	94	80	93	...	NA	NA	NA
Afghanistan	2001	129	379	349	274	204	139	103	146	...	NA	NA	NA
Afghanistan	2002	90	476	481	368	246	241	189	192	...	NA	NA	NA
Afghanistan	2003	127	511	436	284	256	288	203	245	...	NA	NA	NA
Afghanistan	2004	139	537	568	360	358	386	310	256	...	NA	NA	NA
Afghanistan	2005	151	606	560	472	453	470	419	320	...	NA	NA	NA
Afghanistan	2006	193	837	791	574	572	572	410	442	...	NA	NA	NA
Afghanistan	2007	186	856	840	597	566	630	507	475	...	NA	NA	NA
Afghanistan	2008	187	941	773	545	570	630	575	428	...	NA	NA	NA
Afghanistan	2009	200	906	705	499	491	596	570	439	...	NA	NA	NA
:	:	:	:	:	:	:	:	:	:	...	:	:	:
Zimbabwe	1984	NA	NA	NA	NA	NA	NA	NA	NA	...	NA	NA	NA
Zimbabwe	1985	NA	NA	NA	NA	NA	NA	NA	NA	...	NA	NA	NA
Zimbabwe	1986	NA	NA	NA	NA	NA	NA	NA	NA	...	NA	NA	NA
Zimbabwe	1987	NA	NA	NA	NA	NA	NA	NA	NA	...	NA	NA	NA
Zimbabwe	1988	NA	NA	NA	NA	NA	NA	NA	NA	...	NA	NA	NA
Zimbabwe	1989	NA	NA	NA	NA	NA	NA	NA	NA	...	NA	NA	NA
Zimbabwe	1990	NA	NA	NA	NA	NA	NA	NA	NA	...	NA	NA	NA
Zimbabwe	1991	NA	NA	NA	NA	NA	NA	NA	NA	...	NA	NA	NA
Zimbabwe	1992	NA	NA	NA	NA	NA	NA	NA	NA	...	NA	NA	NA
Zimbabwe	1993	NA	NA	NA	NA	NA	NA	NA	NA	...	NA	NA	NA
Zimbabwe	1994	NA	NA	NA	NA	NA	NA	NA	NA	...	NA	NA	NA
Zimbabwe	1995	NA	NA	NA	NA	NA	NA	NA	NA	...	NA	NA	NA
Zimbabwe	1996	NA	NA	NA	NA	NA	NA	NA	NA	...	NA	NA	NA
Zimbabwe	1997	NA	NA	NA	NA	NA	NA	NA	NA	...	NA	NA	NA
Zimbabwe	1998	NA	NA	NA	NA	NA	NA	NA	NA	...	NA	NA	NA
Zimbabwe	1999	NA	NA	NA	NA	NA	NA	NA	NA	...	NA	NA	NA
Zimbabwe	2000	NA	NA	NA	NA	NA	NA	NA	NA	...	NA	NA	NA
Zimbabwe	2001	NA	NA	NA	NA	NA	NA	NA	NA	...	NA	NA	NA
Zimbabwe	2002	191	600	2548	1662	744	315	159	222	...	NA	NA	NA
Zimbabwe	2003	133	874	3048	2228	981	367	205	180	...	NA	NA	NA
Zimbabwe	2004	187	833	2908	2298	1056	366	198	225	...	NA	NA	NA
Zimbabwe	2005	210	837	2264	1855	762	295	656	269	...	NA	NA	NA
Zimbabwe	2006	215	736	2391	1939	896	348	199	237	...	NA	NA	NA
Zimbabwe	2007	138	500	3693	0	716	292	153	185	...	NA	NA	NA
Zimbabwe	2008	127	614	0	3316	704	263	185	145	...	NA	NA	NA
Zimbabwe	2009	125	578	NA	3471	681	293	192	180	...	NA	NA	NA
Zimbabwe	2010	150	710	2208	1682	761	350	252	173	...	NA	NA	NA
Zimbabwe	2011	152	784	2467	2071	780	377	278	174	...	NA	NA	NA
Zimbabwe	2012	120	783	2421	2086	796	360	271	173	...	NA	NA	NA
Zimbabwe	2013	NA	NA	NA	NA	NA	NA	NA	NA	...	2349	1206	1206

```
In [14]: ?who
```

Goal: Graph the number of cases over time grouped by method of diagnosis for any given country

1. Move the variables from column headers to a new variable so that each row represents a case count and the type of case count
2. Separate the various variables into each string using the `separate` command

- Hint: at some point you'll need to use the line

```
mutate(variable_name = string::str_replace(variable_name, "newrel", "new_rel"))
```

which replaces each instance of "newrel" with "new_rel".

3. Create the desired graph

```
In [22]: who %>%
  pivot_longer(new_sp_m014:newrel_f65, names_to = "some_variables", values_to = "cases") %>%
  mutate(some_variables = string::str_replace(some_variables, "newrel", "new_rel")) %>%
  separate(some_variables, into = c("new", "method", "gender_age"), sep = "_")
```

A tibble: 405440 × 8							
country	iso2	iso3	year	new	method	gender_age	cases
<chr>	<chr>	<chr>	<int>	<chr>	<chr>	<chr>	<int>
Afghanistan	AF	AFG	1980	new	sp	m014	NA
Afghanistan	AF	AFG	1980	new	sp	m1524	NA
Afghanistan	AF	AFG	1980	new	sp	m2534	NA
Afghanistan	AF	AFG	1980	new	sp	m3544	NA
Afghanistan	AF	AFG	1980	new	sp	m4554	NA
Afghanistan	AF	AFG	1980	new	sp	m5564	NA
Afghanistan	AF	AFG	1980	new	sp	m65	NA
Afghanistan	AF	AFG	1980	new	sp	f014	NA
Afghanistan	AF	AFG	1980	new	sp	f1524	NA
Afghanistan	AF	AFG	1980	new	sp	f2534	NA
Afghanistan	AF	AFG	1980	new	sp	f3544	NA
Afghanistan	AF	AFG	1980	new	sp	f4554	NA
Afghanistan	AF	AFG	1980	new	sp	f5564	NA
Afghanistan	AF	AFG	1980	new	sp	f65	NA
Afghanistan	AF	AFG	1980	new	sn	m014	NA
Afghanistan	AF	AFG	1980	new	sn	m1524	NA
Afghanistan	AF	AFG	1980	new	sn	m2534	NA
Afghanistan	AF	AFG	1980	new	sn	m3544	NA
Afghanistan	AF	AFG	1980	new	sn	m4554	NA
Afghanistan	AF	AFG	1980	new	sn	m5564	NA
Afghanistan	AF	AFG	1980	new	sn	m65	NA
Afghanistan	AF	AFG	1980	new	sn	f014	NA
Afghanistan	AF	AFG	1980	new	sn	f1524	NA
Afghanistan	AF	AFG	1980	new	sn	f2534	NA
Afghanistan	AF	AFG	1980	new	sn	f3544	NA
Afghanistan	AF	AFG	1980	new	sn	f4554	NA
Afghanistan	AF	AFG	1980	new	sn	f5564	NA
Afghanistan	AF	AFG	1980	new	sn	f65	NA
Afghanistan	AF	AFG	1980	new	ep	m014	NA
Afghanistan	AF	AFG	1980	new	ep	m1524	NA
:	:	:	:	:	:	:	:
Zimbabwe	ZW	ZWE	2013	new	sn	f5564	NA
Zimbabwe	ZW	ZWE	2013	new	sn	f65	NA
Zimbabwe	ZW	ZWE	2013	new	ep	m014	NA
Zimbabwe	ZW	ZWE	2013	new	ep	m1524	NA
Zimbabwe	ZW	ZWE	2013	new	ep	m2534	NA
Zimbabwe	ZW	ZWE	2013	new	ep	m3544	NA
Zimbabwe	ZW	ZWE	2013	new	ep	m4554	NA
Zimbabwe	ZW	ZWE	2013	new	ep	m5564	NA
Zimbabwe	ZW	ZWE	2013	new	ep	m65	NA
Zimbabwe	ZW	ZWE	2013	new	ep	f014	NA
Zimbabwe	ZW	ZWE	2013	new	ep	f1524	NA
Zimbabwe	ZW	ZWE	2013	new	ep	f2534	NA
Zimbabwe	ZW	ZWE	2013	new	ep	f3544	NA
Zimbabwe	ZW	ZWE	2013	new	ep	f4554	NA
Zimbabwe	ZW	ZWE	2013	new	ep	f5564	NA
Zimbabwe	ZW	ZWE	2013	new	ep	f65	NA
Zimbabwe	ZW	ZWE	2013	new	rel	m014	1315
Zimbabwe	ZW	ZWE	2013	new	rel	m1524	1642
Zimbabwe	ZW	ZWE	2013	new	rel	m2534	5331
Zimbabwe	ZW	ZWE	2013	new	rel	m3544	5363
Zimbabwe	ZW	ZWE	2013	new	rel	m4554	2349
Zimbabwe	ZW	ZWE	2013	new	rel	m5564	1206
Zimbabwe	ZW	ZWE	2013	new	rel	m65	1208
Zimbabwe	ZW	ZWE	2013	new	rel	f014	1252
Zimbabwe	ZW	ZWE	2013	new	rel	f1524	2059
Zimbabwe	ZW	ZWE	2013	new	rel	f2534	4649
Zimbabwe	ZW	ZWE	2013	new	rel	f3544	3526
Zimbabwe	ZW	ZWE	2013	new	rel	f4554	1453
Zimbabwe	ZW	ZWE	2013	new	rel	f5564	811