

one-stream wrangle

made with flipbookr and xaringan

Gina Reynolds

```
library(gapminder)
library(tidyverse)
```

```
## — Attaching packages ————— tidyverse 1.3.0 —
```

```
## ✓ ggplot2 3.3.2      ✓ purrr   0.3.3
## ✓ tibble  3.0.0      ✓ dplyr   0.8.5
## ✓ tidyr   1.0.2      ✓ stringr 1.4.0
## ✓ readr   1.3.1      ✓ forcats 0.5.0
```

```
## Warning: package 'ggplot2' was built under R version 3.6.2
```

```
## Warning: package 'tibble' was built under R version 3.6.2
```

```
## — Conflicts ————— tidyverse_conflicts() —
```

```
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()     masks stats::lag()
```

```
knitr::opts_chunk$set(cache = F, comment = "", warning = F)
```

```
gapminder
```

```
# A tibble: 1,704 x 6
  country    continent  year lifeExp      pop gdpPercap
  <fct>      <fct>    <int>  <dbl>    <int>    <dbl>
1 Afghanistan Asia      1952   28.8  8425333    779.
2 Afghanistan Asia      1957   30.3  9240934    821.
3 Afghanistan Asia      1962   32.0 10267083    853.
4 Afghanistan Asia      1967   34.0 11537966    836.
5 Afghanistan Asia      1972   36.1 13079460    740.
6 Afghanistan Asia      1977   38.4 14880372    786.
7 Afghanistan Asia      1982   39.9 12881816    978.
8 Afghanistan Asia      1987   40.8 13867957    852.
9 Afghanistan Asia      1992   41.7 16317921    649.
10 Afghanistan Asia      1997   41.8 22227415    635.
# ... with 1,694 more rows
```

```
gapminder %>%  
  filter(year == 2002)
```

```
# A tibble: 142 x 6  
  country    continent  year lifeExp      pop gdpPercap  
  <fct>      <fct>    <int> <dbl>    <int> <dbl>  
1 Afghanistan Asia      2002  42.1  25268405    727.  
2 Albania    Europe    2002  75.7   3508512   4604.  
3 Algeria    Africa    2002  71.0  31287142   5288.  
4 Angola     Africa    2002  41.0  10866106   2773.  
5 Argentina  Americas  2002  74.3  38331121   8798.  
6 Australia  Oceania   2002  80.4  19546792  30688.  
7 Austria    Europe    2002  79.0   8148312  32418.  
8 Bahrain    Asia      2002  74.8    656397  23404.  
9 Bangladesh Asia      2002  62.0  135656790  1136.  
10 Belgium   Europe    2002  78.3  10311970  30486.  
# ... with 132 more rows
```

```
gapminder %>%  
  filter(year == 2002) %>%  
  select(-lifeExp)
```

```
# A tibble: 142 x 5  
  country      continent  year      pop gdpPercap  
  <fct>        <fct>    <int>   <int>   <dbl>  
1 Afghanistan Asia      2002  25268405    727.  
2 Albania     Europe   2002   3508512   4604.  
3 Algeria     Africa   2002  31287142   5288.  
4 Angola      Africa   2002  10866106   2773.  
5 Argentina   Americas 2002  38331121   8798.  
6 Australia   Oceania   2002  19546792  30688.  
7 Austria     Europe   2002   8148312  32418.  
8 Bahrain     Asia     2002    656397  23404.  
9 Bangladesh  Asia     2002  135656790   1136.  
10 Belgium    Europe   2002   10311970  30486.  
# ... with 132 more rows
```

```
gapminder %>%
  filter(year == 2002) %>%
  select(-lifeExp) %>%
  rename(gdp_per_cap = gdpPerCap)
```

```
# A tibble: 142 x 5
  country    continent  year      pop gdp_per_cap
  <fct>      <fct>    <int>   <int>   <dbl>
1 Afghanistan Asia      2002  25268405    727.
2 Albania    Europe    2002   3508512   4604.
3 Algeria    Africa    2002  31287142   5288.
4 Angola     Africa    2002  10866106   2773.
5 Argentina  Americas  2002  38331121   8798.
6 Australia  Oceania   2002  19546792  30688.
7 Austria    Europe    2002   8148312  32418.
8 Bahrain    Asia      2002    656397  23404.
9 Bangladesh Asia      2002  135656790   1136.
10 Belgium   Europe    2002   10311970  30486.
# ... with 132 more rows
```

```
gapminder %>%
  filter(year == 2002) %>%
  select(-lifeExp) %>%
  rename(gdp_per_cap = gdpPerCap) %>%
  mutate(gdp = gdp_per_cap * pop)
```

```
# A tibble: 142 x 6
  country    continent  year      pop gdp_per_cap      gdp
  <fct>      <fct>    <int>   <int>    <dbl>    <dbl>
1 Afghanistan Asia      2002  25268405    727.  18363410424.
2 Albania    Europe    2002   3508512   4604.  16153932130.
3 Algeria    Africa    2002  31287142   5288.  165447670333.
4 Angola     Africa    2002  10866106   2773.  30134833901.
5 Argentina  Americas  2002   38331121   8798.  337223430800.
6 Australia  Oceania   2002   19546792  30688.  599847158654.
7 Austria    Europe    2002    8148312  32418.  264148781752.
8 Bahrain    Asia      2002    656397   23404.  15362026094.
9 Bangladesh Asia      2002  135656790   1136.  154159077921.
10 Belgium   Europe    2002   10311970  30486.  314369518653.
# ... with 132 more rows
```

```
gapminder %>%
  filter(year == 2002) %>%
  select(-lifeExp) %>%
  rename(gdp_per_cap = gdpPercap) %>%
  mutate(gdp = gdp_per_cap * pop) %>%
  mutate(percent_gdp = 100 * gdp / sum(gdp))
```

```
# A tibble: 142 x 7
  country    continent  year      pop gdp_per_cap      gdp percent_gdp
  <fct>      <fct>      <int>   <int>    <dbl>      <dbl>      <dbl>
1 Afghanistan Asia      2002  25268405      727.  18363410424.    0.0389
2 Albania    Europe    2002   3508512     4604.  16153932130.    0.0342
3 Algeria    Africa    2002  31287142     5288.  165447670333.    0.350
4 Angola     Africa    2002  10866106     2773.  30134833901.    0.0638
5 Argentina  Americas  2002  38331121     8798.  337223430800.    0.713
6 Australia  Oceania   2002  19546792    30688.  599847158654.    1.27
7 Austria    Europe    2002   8148312    32418.  264148781752.    0.559
8 Bahrain    Asia      2002    656397    23404.  15362026094.    0.0325
9 Bangladesh Asia      2002  135656790     1136.  154159077921.    0.326
10 Belgium   Europe    2002   10311970    30486.  314369518653.    0.665
# ... with 132 more rows
```



```
gapminder %>%
  filter(year == 2002) %>%
  select(-lifeExp) %>%
  rename(gdp_per_cap = gdpPerCap) %>%
  mutate(gdp = gdp_per_cap * pop) %>%
  mutate(percent_gdp = 100 * gdp / sum(gdp)) %>%
  mutate(europe = continent == "Europe")
```

```
# A tibble: 142 x 8
  country    continent year      pop gdp_per_cap      gdp percent_gdp europe
  <fct>      <fct>    <int>   <int>    <dbl>    <dbl>    <dbl> <lgl>
1 Afghanist... Asia      2002  2.53e7    727.    1.84e10  0.0389 FALSE
2 Albania     Europe    2002  3.51e6   4604.    1.62e10  0.0342 TRUE
3 Algeria     Africa    2002  3.13e7   5288.    1.65e11  0.350  FALSE
4 Angola      Africa    2002  1.09e7   2773.    3.01e10  0.0638 FALSE
5 Argentina   Americas  2002  3.83e7   8798.    3.37e11  0.713  FALSE
6 Australia   Oceania   2002  1.95e7  30688.    6.00e11  1.27   FALSE
7 Austria     Europe    2002  8.15e6  32418.    2.64e11  0.559  TRUE
8 Bahrain     Asia      2002  6.56e5  23404.    1.54e10  0.0325 FALSE
9 Bangladesh  Asia      2002  1.36e8   1136.    1.54e11  0.326  FALSE
10 Belgium    Europe    2002  1.03e7  30486.    3.14e11  0.665  TRUE
# ... with 132 more rows
```

```
gapminder %>%
  filter(year == 2002) %>%
  select(-lifeExp) %>%
  rename(gdp_per_cap = gdpPercap) %>%
  mutate(gdp = gdp_per_cap * pop) %>%
  mutate(percent_gdp = 100 * gdp / sum(gdp)) %>%
  mutate(europe = continent == "Europe") %>%
  select(country, year, gdp, europe, pop)
```

```
# A tibble: 142 x 5
  country      year      gdp europe      pop
  <fct>      <int>    <dbl> <lgl>    <int>
1 Afghanistan 2002  18363410424. FALSE 25268405
2 Albania      2002  16153932130. TRUE  3508512
3 Algeria      2002  165447670333. FALSE 31287142
4 Angola       2002   30134833901. FALSE 10866106
5 Argentina    2002  337223430800. FALSE 38331121
6 Australia    2002  599847158654. FALSE 19546792
7 Austria      2002  264148781752. TRUE   8148312
8 Bahrain      2002   15362026094. FALSE   656397
9 Bangladesh   2002  154159077921. FALSE 135656790
10 Belgium     2002  314369518653. TRUE   10311970
# ... with 132 more rows
```

```
gapminder %>%
  filter(year == 2002) %>%
  select(-lifeExp) %>%
  rename(gdp_per_cap = gdpPerCap) %>%
  mutate(gdp = gdp_per_cap * pop) %>%
  mutate(percent_gdp = 100 * gdp / sum(gdp)) %>%
  mutate(europe = continent == "Europe") %>%
  select(country, year, gdp, europe, pop) %>%
  mutate(europe_category =
    case_when(europe == T ~ "Europe",
              europe == F ~ "Not Europe"))
```

```
# A tibble: 142 x 6
  country      year      gdp europe      pop europe_category
  <fct>      <int>      <dbl> <lgl>      <int> <chr>
1 Afghanistan 2002 18363410424. FALSE 25268405 Not Europe
2 Albania      2002 16153932130. TRUE 3508512 Europe
3 Algeria      2002 165447670333. FALSE 31287142 Not Europe
4 Angola       2002 30134833901. FALSE 10866106 Not Europe
5 Argentina    2002 337223430800. FALSE 38331121 Not Europe
6 Australia    2002 599847158654. FALSE 19546792 Not Europe
7 Austria      2002 264148781752. TRUE 8148312 Europe
8 Bahrain      2002 15362026094. FALSE 656397 Not Europe
9 Bangladesh   2002 154159077921. FALSE 135656790 Not Europe
10 Belgium     2002 314369518653. TRUE 10311970 Europe
# ... with 132 more rows
```

```
gapminder %>%
  filter(year == 2002) %>%
  select(-lifeExp) %>%
  rename(gdp_per_cap = gdpPerCap) %>%
  mutate(gdp = gdp_per_cap * pop) %>%
  mutate(percent_gdp = 100 * gdp / sum(gdp)) %>%
  mutate(europe = continent == "Europe") %>%
  select(country, year, gdp, europe, pop) %>%
  mutate(europe_category =
    case_when(europe == T ~ "Europe",
              europe == F ~ "Not Europe")) %>%
  arrange(-gdp)
```

```
# A tibble: 142 x 6
  country      year  gdp europe      pop europe_category
  <fct>      <int> <dbl> <lgl>      <int> <chr>
1 United States 2002 1.12e13 FALSE 287675526 Not Europe
2 China         2002 3.99e12 FALSE 1280400000 Not Europe
3 Japan         2002 3.63e12 FALSE 127065841 Not Europe
4 Germany       2002 2.47e12 TRUE  82350671 Europe
5 India         2002 1.81e12 FALSE 1034172547 Not Europe
6 United Kingdom 2002 1.77e12 TRUE  59912431 Europe
7 France        2002 1.73e12 TRUE  59925035 Europe
8 Italy         2002 1.62e12 TRUE  57926999 Europe
9 Brazil        2002 1.46e12 FALSE 179914212 Not Europe
10 Mexico       2002 1.10e12 FALSE 102479927 Not Europe
# ... with 132 more rows
```

```

gapminder %>%
  filter(year == 2002) %>%
  select(-lifeExp) %>%
  rename(gdp_per_cap = gdpPerCap) %>%
  mutate(gdp = gdp_per_cap * pop) %>%
  mutate(percent_gdp = 100 * gdp / sum(gdp)) %>%
  mutate(europe = continent == "Europe") %>%
  select(country, year, gdp, europe, pop) %>%
  mutate(europe_category =
    case_when(europe == T ~ "Europe",
              europe == F ~ "Not Europe")) %>%
  arrange(-gdp) %>%
  mutate(gdp_billions = gdp/1000000000)

```

```

# A tibble: 142 x 7
  country      year  gdp europe      pop europe_category gdp_billions
  <fct>      <int> <dbl> <lgl>      <int> <chr>          <dbl>
1 United States 2002 1.12e13 FALSE 287675526 Not Europe    11247.
2 China         2002 3.99e12 FALSE 1280400000 Not Europe    3994.
3 Japan         2002 3.63e12 FALSE 127065841 Not Europe    3635.
4 Germany       2002 2.47e12 TRUE  82350671 Europe      2473.
5 India         2002 1.81e12 FALSE 1034172547 Not Europe    1806.
6 United Kingdom 2002 1.77e12 TRUE  59912431 Europe      1766.
7 France        2002 1.73e12 TRUE  59925035 Europe      1733.
8 Italy         2002 1.62e12 TRUE  57926999 Europe      1620.
9 Brazil        2002 1.46e12 FALSE 179914212 Not Europe    1463.
10 Mexico       2002 1.10e12 FALSE 102479927 Not Europe    1101.
# ... with 132 more rows

```

```
gapminder %>%
  filter(year == 2002) %>%
  select(-lifeExp) %>%
  rename(gdp_per_cap = gdpPerCap) %>%
  mutate(gdp = gdp_per_cap * pop) %>%
  mutate(percent_gdp = 100 * gdp / sum(gdp)) %>%
  mutate(europe = continent == "Europe") %>%
  select(country, year, gdp, europe, pop) %>%
  mutate(europe_category =
    case_when(europe == T ~ "Europe",
              europe == F ~ "Not Europe")) %>%
  arrange(-gdp) %>%
  mutate(gdp_billions = gdp/1000000000) %>%
  slice(1:8)
```

```
# A tibble: 8 x 7
  country      year  gdp europe      pop europe_category gdp_billions
  <fct>      <int> <dbl> <lgl>      <int> <chr>          <dbl>
1 United States 2002 1.12e13 FALSE 287675526 Not Europe    11247.
2 China         2002 3.99e12 FALSE 1280400000 Not Europe     3994.
3 Japan         2002 3.63e12 FALSE 127065841 Not Europe     3635.
4 Germany       2002 2.47e12  TRUE  82350671  Europe       2473.
5 India         2002 1.81e12 FALSE 1034172547 Not Europe     1806.
6 United Kingdom 2002 1.77e12  TRUE  59912431  Europe       1766.
7 France        2002 1.73e12  TRUE  59925035  Europe       1733.
8 Italy         2002 1.62e12  TRUE  57926999  Europe       1620.
```

```
gapminder %>%
  filter(year == 2002) %>%
  select(-lifeExp) %>%
  rename(gdp_per_cap = gdpPercap) %>%
  mutate(gdp = gdp_per_cap * pop) %>%
  mutate(percent_gdp = 100 * gdp / sum(gdp)) %>%
  mutate(europe = continent == "Europe") %>%
  select(country, year, gdp, europe, pop) %>%
  mutate(europe_category =
    case_when(europe == T ~ "Europe",
              europe == F ~ "Not Europe")) %>%
  arrange(-gdp) %>%
  mutate(gdp_billions = gdp/1000000000) %>%
  slice(1:8) ->
europe_or_not_2002
```

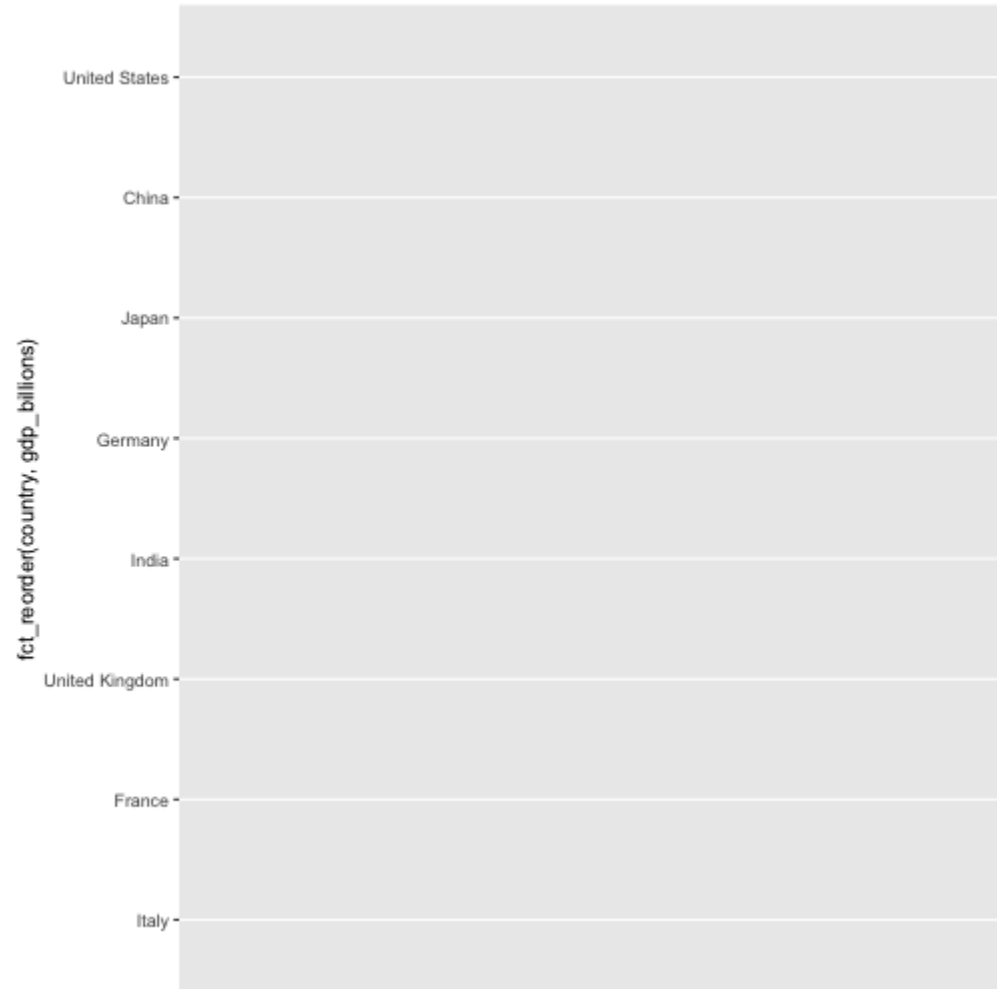
```
gapminder %>%
  filter(year == 2002) %>%
  select(-lifeExp) %>%
  rename(gdp_per_cap = gdpPerCap) %>%
  mutate(gdp = gdp_per_cap * pop) %>%
  mutate(percent_gdp = 100 * gdp / sum(gdp)) %>%
  mutate(europe = continent == "Europe") %>%
  select(country, year, gdp, europe, pop) %>%
  mutate(europe_category =
    case_when(europe == T ~ "Europe",
              europe == F ~ "Not Europe")) %>%
  arrange(-gdp) %>%
  mutate(gdp_billions = gdp/1000000000) %>%
  slice(1:8) ->
europe_or_not_2002

ggplot(data = europe_or_not_2002)
```



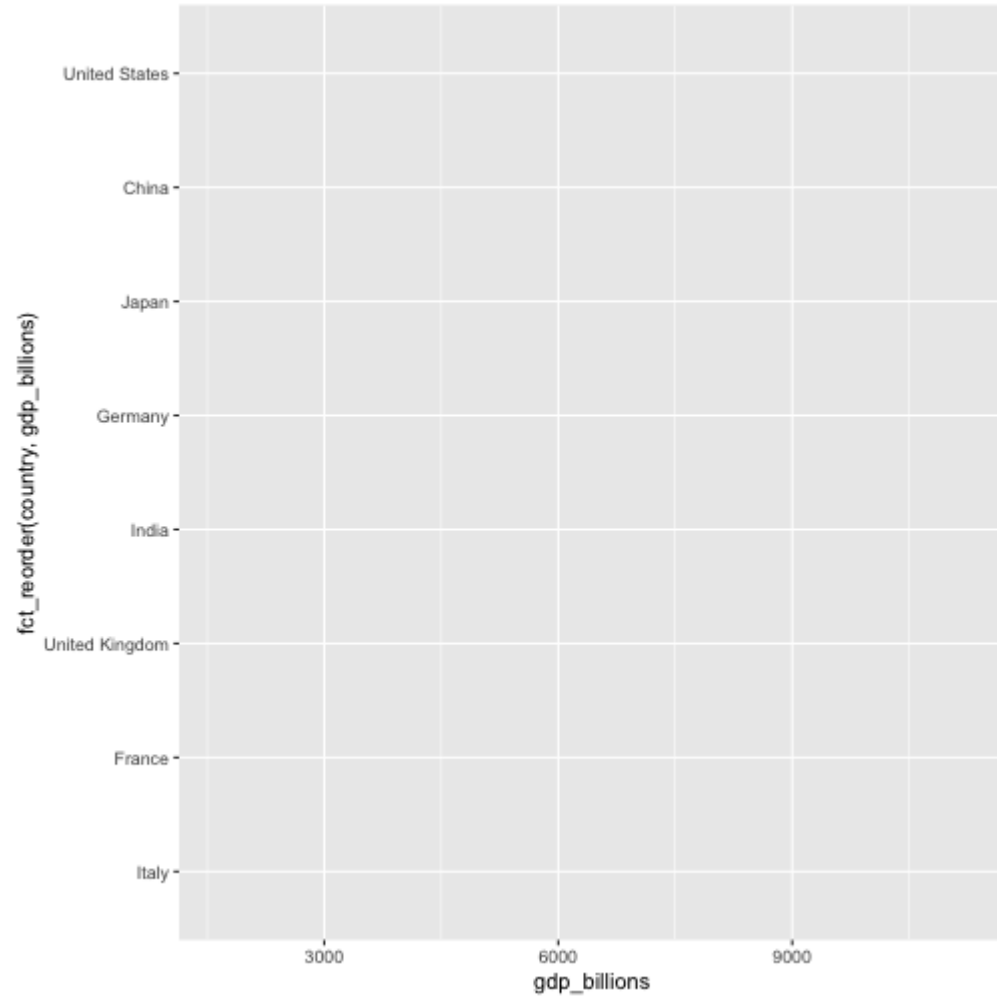
```
gapminder %>%
  filter(year == 2002) %>%
  select(-lifeExp) %>%
  rename(gdp_per_cap = gdpPercap) %>%
  mutate(gdp = gdp_per_cap * pop) %>%
  mutate(percent_gdp = 100 * gdp / sum(gdp)) %>%
  mutate(europe = continent == "Europe") %>%
  select(country, year, gdp, europe, pop) %>%
  mutate(europe_category =
    case_when(europe == T ~ "Europe",
              europe == F ~ "Not Europe")) %>%
  arrange(-gdp) %>%
  mutate(gdp_billions = gdp/1000000000) %>%
  slice(1:8) ->
europe_or_not_2002

ggplot(data = europe_or_not_2002) +
  aes(y = fct_reorder(country, gdp_billions))
```



```
gapminder %>%
  filter(year == 2002) %>%
  select(-lifeExp) %>%
  rename(gdp_per_cap = gdpPerCap) %>%
  mutate(gdp = gdp_per_cap * pop) %>%
  mutate(percent_gdp = 100 * gdp / sum(gdp)) %>%
  mutate(europe = continent == "Europe") %>%
  select(country, year, gdp, europe, pop) %>%
  mutate(europe_category =
    case_when(europe == T ~ "Europe",
              europe == F ~ "Not Europe")) %>%
  arrange(-gdp) %>%
  mutate(gdp_billions = gdp/1000000000) %>%
  slice(1:8) ->
europe_or_not_2002

ggplot(data = europe_or_not_2002) +
  aes(y = fct_reorder(country, gdp_billions)) +
  aes(x = gdp_billions)
```

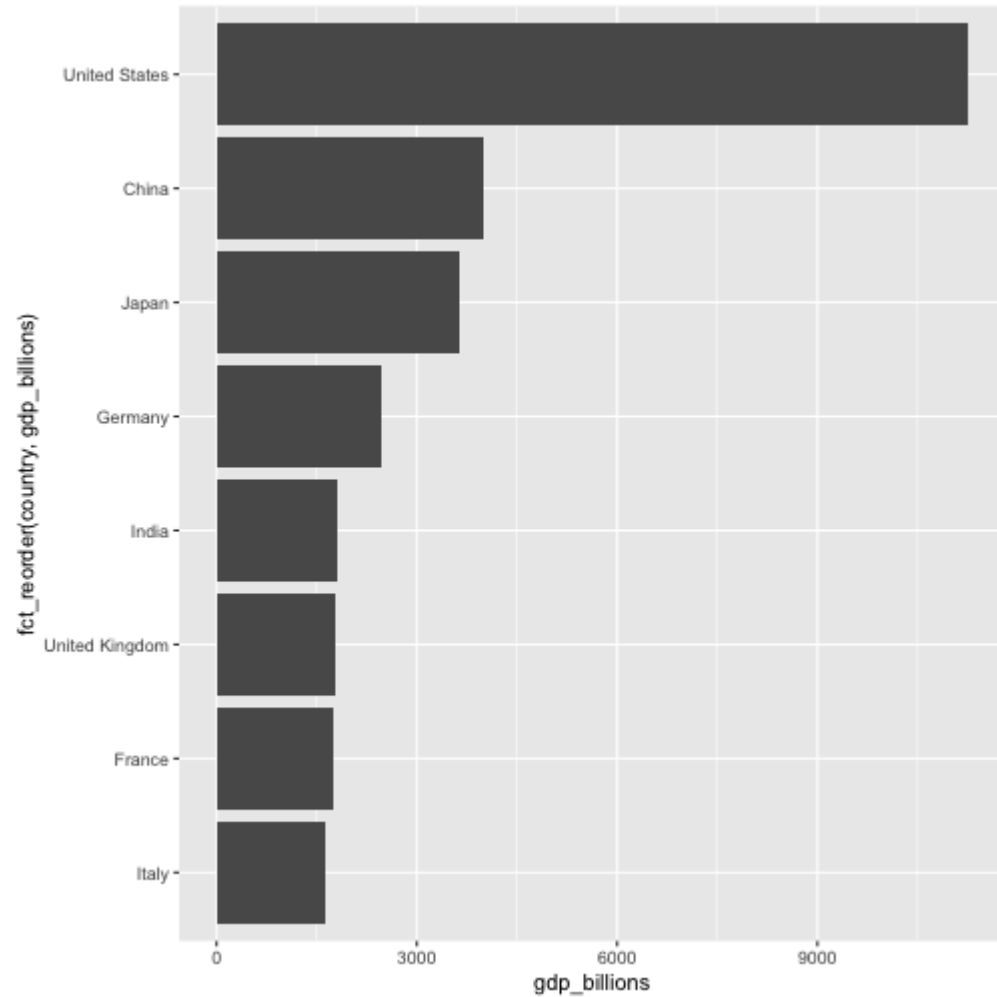


```

gapminder %>%
  filter(year == 2002) %>%
  select(-lifeExp) %>%
  rename(gdp_per_cap = gdpPerCap) %>%
  mutate(gdp = gdp_per_cap * pop) %>%
  mutate(percent_gdp = 100 * gdp / sum(gdp)) %>%
  mutate(europe = continent == "Europe") %>%
  select(country, year, gdp, europe, pop) %>%
  mutate(europe_category =
    case_when(europe == T ~ "Europe",
              europe == F ~ "Not Europe")) %>%
  arrange(-gdp) %>%
  mutate(gdp_billions = gdp/1000000000) %>%
  slice(1:8) ->
europe_or_not_2002

ggplot(data = europe_or_not_2002) +
  aes(y = fct_reorder(country, gdp_billions)) +
  aes(x = gdp_billions) +
  geom_col()

```

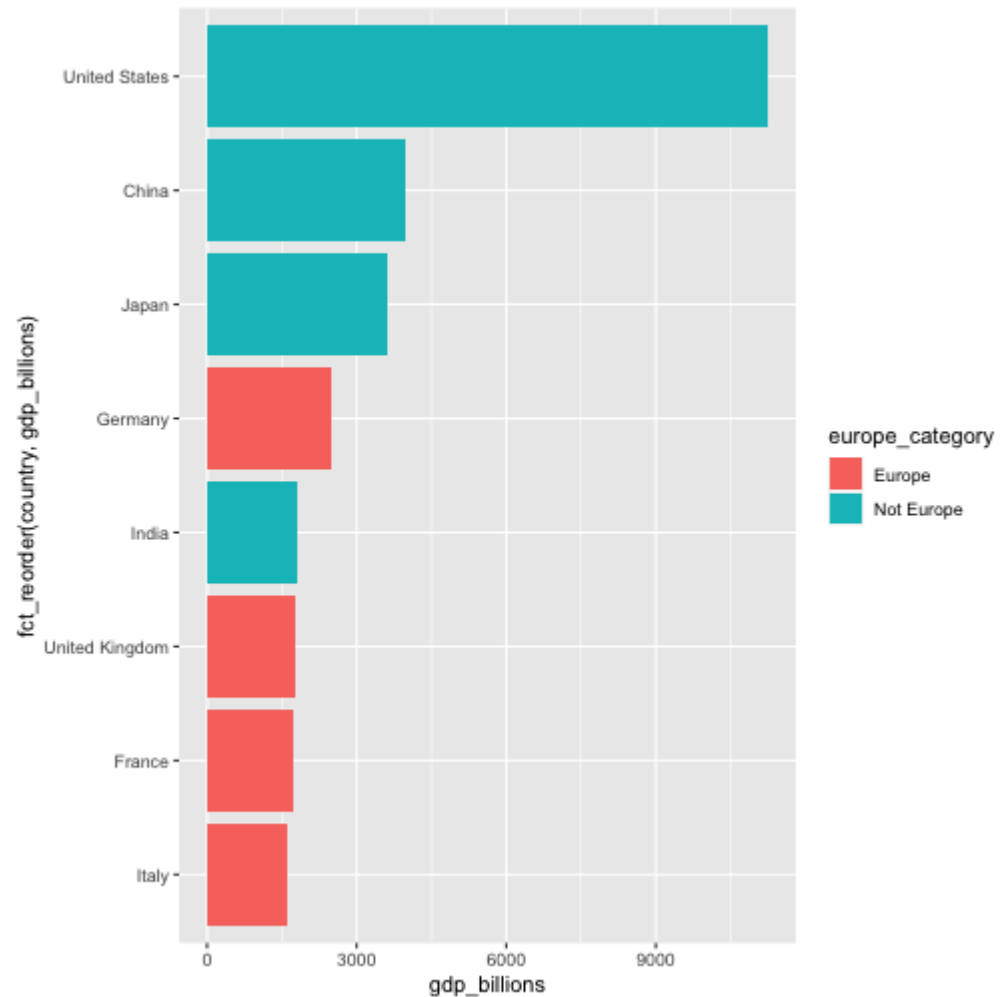


```

gapminder %>%
  filter(year == 2002) %>%
  select(-lifeExp) %>%
  rename(gdp_per_cap = gdpPerCap) %>%
  mutate(gdp = gdp_per_cap * pop) %>%
  mutate(percent_gdp = 100 * gdp / sum(gdp)) %>%
  mutate(europe = continent == "Europe") %>%
  select(country, year, gdp, europe, pop) %>%
  mutate(europe_category =
    case_when(europe == T ~ "Europe",
              europe == F ~ "Not Europe")) %>%
  arrange(-gdp) %>%
  mutate(gdp_billions = gdp/1000000000) %>%
  slice(1:8) ->
europe_or_not_2002

ggplot(data = europe_or_not_2002) +
  aes(y = fct_reorder(country, gdp_billions)) +
  aes(x = gdp_billions) +
  geom_col() +
  aes(fill = europe_category)

```

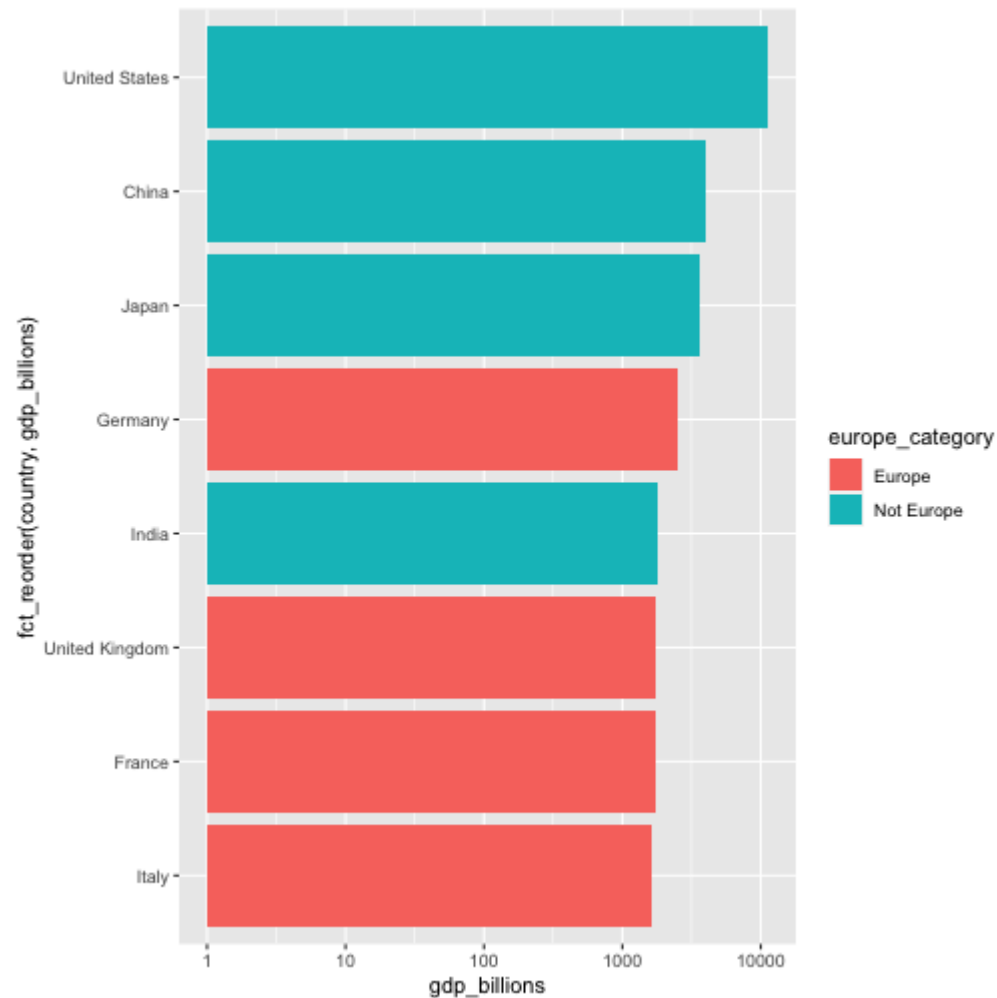


```

gapminder %>%
  filter(year == 2002) %>%
  select(-lifeExp) %>%
  rename(gdp_per_cap = gdpPerCap) %>%
  mutate(gdp = gdp_per_cap * pop) %>%
  mutate(percent_gdp = 100 * gdp / sum(gdp)) %>%
  mutate(europe = continent == "Europe") %>%
  select(country, year, gdp, europe, pop) %>%
  mutate(europe_category =
    case_when(europe == T ~ "Europe",
              europe == F ~ "Not Europe")) %>%
  arrange(-gdp) %>%
  mutate(gdp_billions = gdp/1000000000) %>%
  slice(1:8) ->
europe_or_not_2002

ggplot(data = europe_or_not_2002) +
  aes(y = fct_reorder(country, gdp_billions)) +
  aes(x = gdp_billions) +
  geom_col() +
  aes(fill = europe_category) +
  scale_x_log10()

```

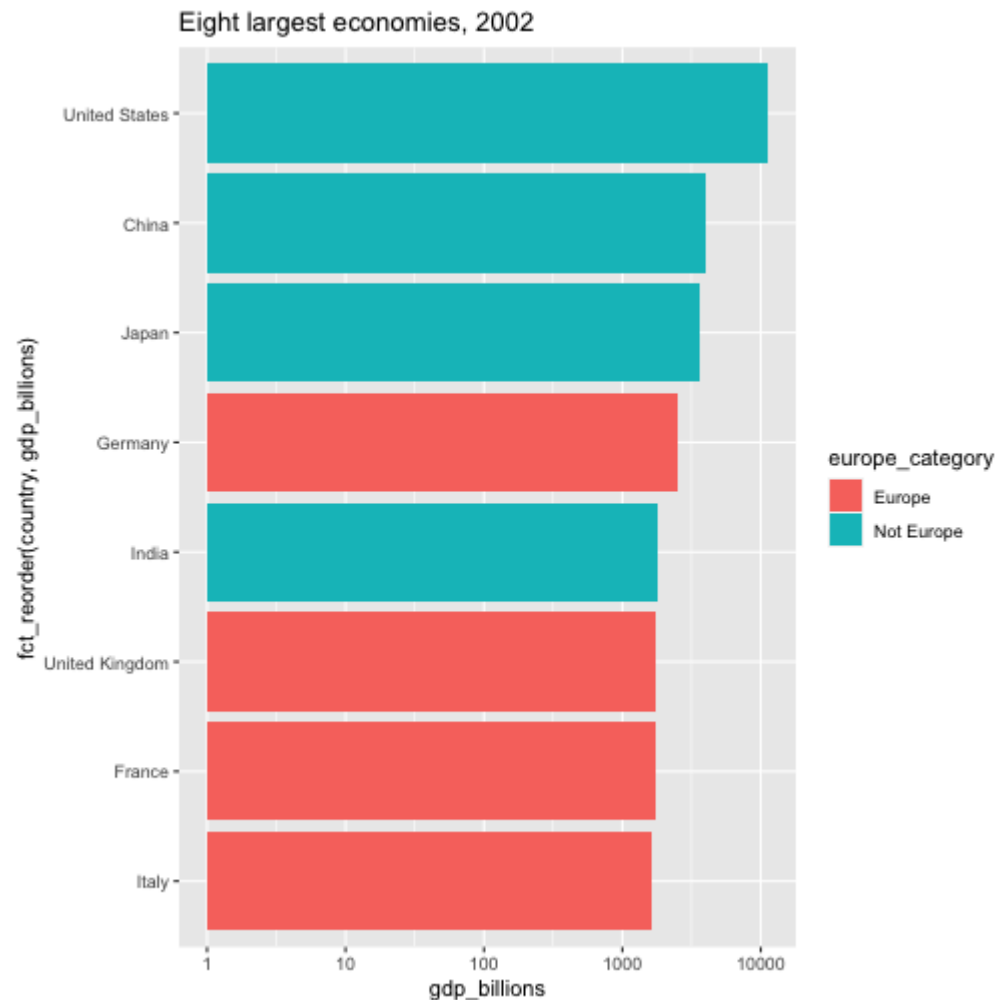


```

gapminder %>%
  filter(year == 2002) %>%
  select(-lifeExp) %>%
  rename(gdp_per_cap = gdpPerCap) %>%
  mutate(gdp = gdp_per_cap * pop) %>%
  mutate(percent_gdp = 100 * gdp / sum(gdp)) %>%
  mutate(europe = continent == "Europe") %>%
  select(country, year, gdp, europe, pop) %>%
  mutate(europe_category =
    case_when(europe == T ~ "Europe",
              europe == F ~ "Not Europe")) %>%
  arrange(-gdp) %>%
  mutate(gdp_billions = gdp/1000000000) %>%
  slice(1:8) ->
europe_or_not_2002

ggplot(data = europe_or_not_2002) +
  aes(y = fct_reorder(country, gdp_billions)) +
  aes(x = gdp_billions) +
  geom_col() +
  aes(fill = europe_category) +
  scale_x_log10() +
  labs(title = "Eight largest economies, 2002")

```



```
my_countries <- c("Brazil", "France", "Russia",  
                  "United Kingdom", "United States",  
                  "China", "India", "Japan")
```

```
my_countries <- c("Brazil", "France", "Russia",  
  "United Kingdom", "United States",  
  "China", "India", "Japan")
```

```
gapminder
```

```
# A tibble: 1,704 x 6  
  country    continent  year lifeExp      pop gdpPercap  
  <fct>      <fct>    <int> <dbl>    <int>    <dbl>  
1 Afghanistan Asia      1952  28.8  8425333    779.  
2 Afghanistan Asia      1957  30.3  9240934    821.  
3 Afghanistan Asia      1962  32.0 10267083    853.  
4 Afghanistan Asia      1967  34.0 11537966    836.  
5 Afghanistan Asia      1972  36.1 13079460    740.  
6 Afghanistan Asia      1977  38.4 14880372    786.  
7 Afghanistan Asia      1982  39.9 12881816    978.  
8 Afghanistan Asia      1987  40.8 13867957    852.  
9 Afghanistan Asia      1992  41.7 16317921    649.  
10 Afghanistan Asia      1997  41.8 22227415    635.  
# ... with 1,694 more rows
```



```
my_countries <- c("Brazil", "France", "Russia",
  "United Kingdom", "United States",
  "China", "India", "Japan")
```

```
gapminder %>%
```

```
  filter(year == 2002)
```

```
# A tibble: 142 x 6
```

	country	continent	year	lifeExp	pop	gdpPercap
	<fct>	<fct>	<int>	<dbl>	<int>	<dbl>
1	Afghanistan	Asia	2002	42.1	25268405	727.
2	Albania	Europe	2002	75.7	3508512	4604.
3	Algeria	Africa	2002	71.0	31287142	5288.
4	Angola	Africa	2002	41.0	10866106	2773.
5	Argentina	Americas	2002	74.3	38331121	8798.
6	Australia	Oceania	2002	80.4	19546792	30688.
7	Austria	Europe	2002	79.0	8148312	32418.
8	Bahrain	Asia	2002	74.8	656397	23404.
9	Bangladesh	Asia	2002	62.0	135656790	1136.
10	Belgium	Europe	2002	78.3	10311970	30486.

```
# ... with 132 more rows
```

```
my_countries <- c("Brazil", "France", "Russia",
  "United Kingdom", "United States",
  "China", "India", "Japan")
```

```
gapminder %>%
```

```
  filter(year == 2002) %>%
```

```
  filter(country %in% my_countries)
```

```
# A tibble: 7 x 6
```

	country	continent	year	lifeExp	pop	gdpPercap
	<fct>	<fct>	<int>	<dbl>	<int>	<dbl>
1	Brazil	Americas	2002	71.0	179914212	8131.
2	China	Asia	2002	72.0	1280400000	3119.
3	France	Europe	2002	79.6	59925035	28926.
4	India	Asia	2002	62.9	1034172547	1747.
5	Japan	Asia	2002	82	127065841	28605.
6	United Kingdom	Europe	2002	78.5	59912431	29479.
7	United States	Americas	2002	77.3	287675526	39097.

```

my_countries <- c("Brazil", "France", "Russia",
                 "United Kingdom", "United States",
                 "China", "India", "Japan")

gapminder %>%
  filter(year == 2002) %>%
  filter(country %in% my_countries) %>%
  mutate(pop_category =
    case_when(pop >= 1000000000 ~ "billions",
              pop >= 100000000 ~ "hundred millions",
              pop < 100000000 ~ "fewer than hundred m...)

```

```

# A tibble: 7 x 7
  country    continent  year lifeExp      pop gdpPercap pop_category
  <fct>      <fct>    <int>  <dbl>    <int>    <dbl> <chr>
1 Brazil    Americas   2002   71.0  1.80e8    8131. hundred millions
2 China      Asia       2002   72.0  1.28e9    3119. billions
3 France     Europe     2002   79.6  5.99e7   28926. fewer than hundred m...
4 India      Asia       2002   62.9  1.03e9    1747. billions
5 Japan      Asia       2002    82   1.27e8   28605. hundred millions
6 United King... Europe     2002   78.5  5.99e7   29479. fewer than hundred m...
7 United Stat... Americas   2002   77.3  2.88e8   39097. hundred millions

```

```

my_countries <- c("Brazil", "France", "Russia",
  "United Kingdom", "United States",
  "China", "India", "Japan")

gapminder %>%
  filter(year == 2002) %>%
  filter(country %in% my_countries) %>%
  mutate(pop_category =
    case_when(pop >= 1000000000 ~ "billions",
              pop >= 100000000 ~ "hundred millions",
              pop < 100000000 ~ "fewer than hundred m...)
  mutate(pop_category = fct_relevel(pop_category,
    level = c("fewer than hundred million",
              "hundred millions",
              "billions")))

```

```

# A tibble: 7 x 7
  country      continent  year lifeExp      pop gdpPercap pop_category
  <fct>        <fct>    <int>   <dbl>   <int>   <dbl> <fct>
1 Brazil      Americas    2002    71.0  1.80e8  8131. hundred millions
2 China        Asia      2002    72.0  1.28e9  3119. billions
3 France      Europe     2002    79.6  5.99e7 28926. fewer than hundred m...
4 India        Asia      2002    62.9  1.03e9  1747. billions
5 Japan        Asia      2002     82  1.27e8 28605. hundred millions
6 United King... Europe     2002    78.5  5.99e7 29479. fewer than hundred m...
7 United Stat... Americas    2002    77.3  2.88e8 39097. hundred millions

```

```
my_countries <- c("Brazil", "France", "Russia",  
  "United Kingdom", "United States",  
  "China", "India", "Japan")  
  
gapminder %>%  
  filter(year == 2002) %>%  
  filter(country %in% my_countries) %>%  
  mutate(pop_category =  
    case_when(pop >= 1000000000 ~ "billions",  
              pop >= 100000000 ~ "hundred millions",  
              pop < 100000000 ~ "fewer than hundred mi  
  mutate(pop_category = fct_relevel(pop_category,  
    level = c("fewer than hundred million",  
              "hundred millions",  
              "billions"))) %>%  
  ggplot()
```

```

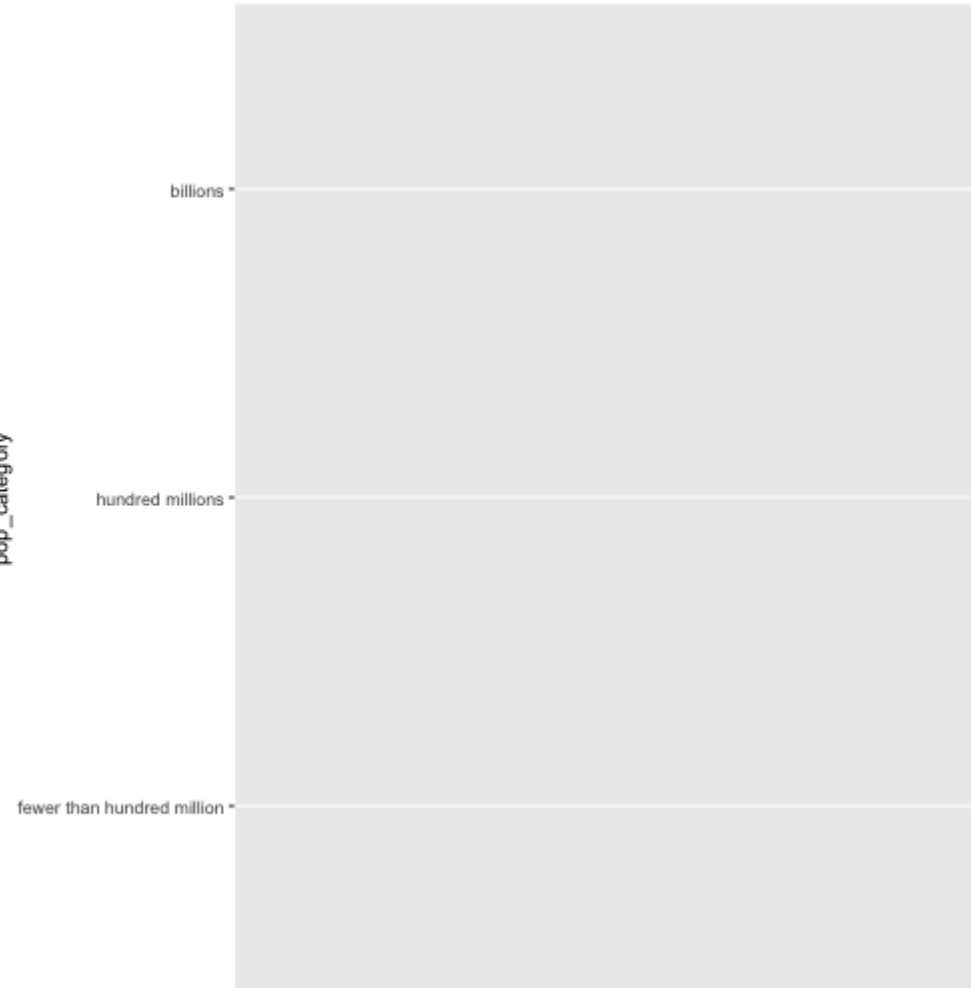
my_countries <- c("Brazil", "France", "Russia",
                 "United Kingdom", "United States",
                 "China", "India", "Japan")

gapminder %>%
  filter(year == 2002) %>%
  filter(country %in% my_countries) %>%
  mutate(pop_category =
    case_when(pop >= 1000000000 ~ "billions",
              pop >= 100000000 ~ "hundred millions",
              pop < 100000000 ~ "fewer than hundred mi
  mutate(pop_category = fct_relevel(pop_category,
                                    level = c("fewer than hundred million",
                                                "hundred millions",
                                                "billions"))) %>%

ggplot() +
  aes(y = pop_category)

```

pop_category



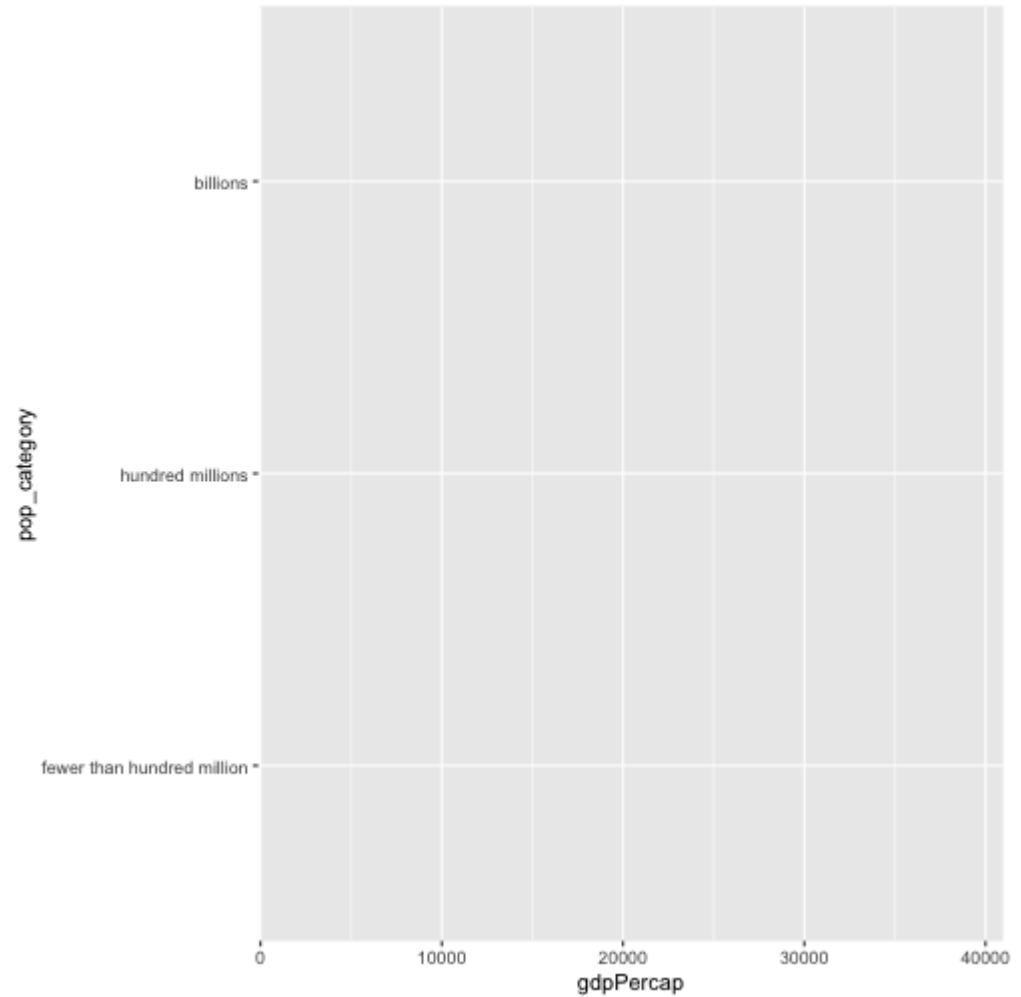
```

my_countries <- c("Brazil", "France", "Russia",
                  "United Kingdom", "United States",
                  "China", "India", "Japan")

gapminder %>%
  filter(year == 2002) %>%
  filter(country %in% my_countries) %>%
  mutate(pop_category =
    case_when(pop >= 1000000000 ~ "billions",
              pop >= 100000000 ~ "hundred millions",
              pop < 100000000 ~ "fewer than hundred mi
  mutate(pop_category = fct_relevel(pop_category,
                                    level = c("fewer than hundred million",
                                                "hundred millions",
                                                "billions"))) %>%

ggplot() +
  aes(y = pop_category) +
  aes(x = gdpPercap)

```



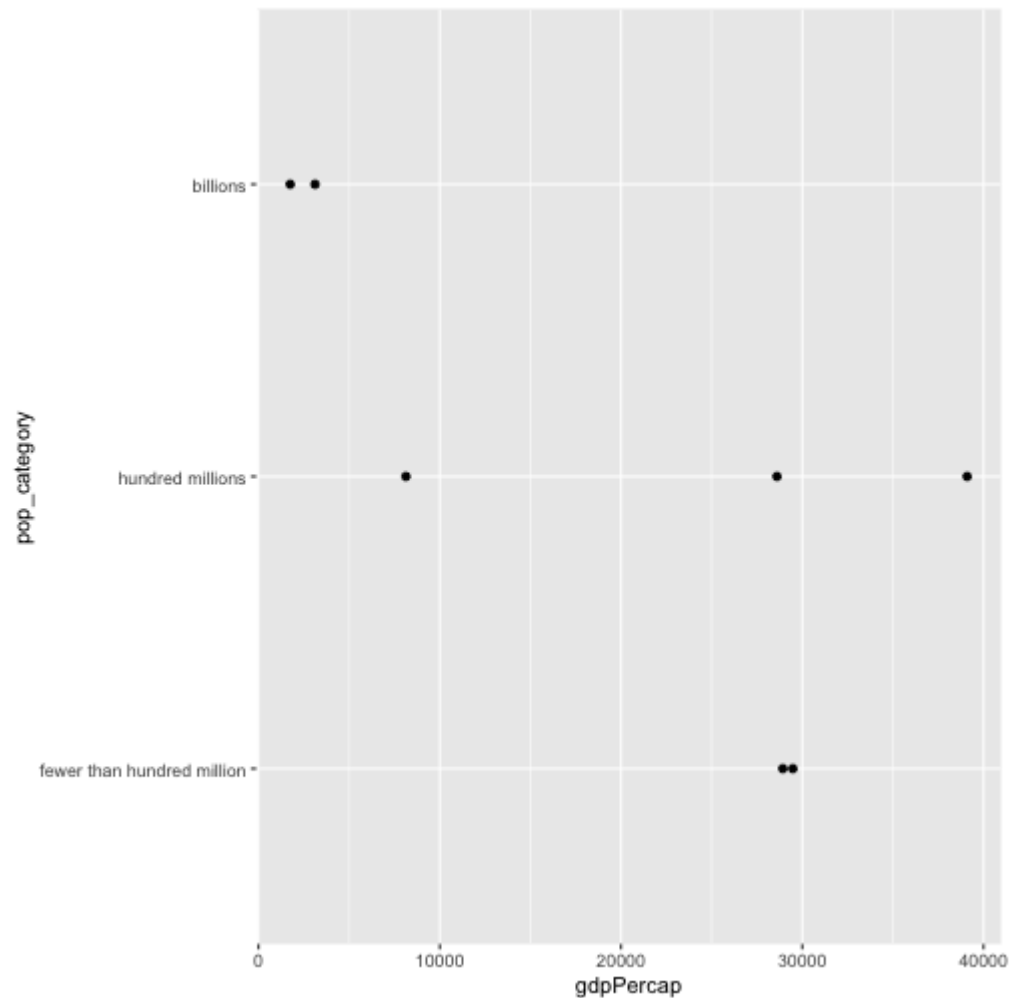
```

my_countries <- c("Brazil", "France", "Russia",
                 "United Kingdom", "United States",
                 "China", "India", "Japan")

gapminder %>%
  filter(year == 2002) %>%
  filter(country %in% my_countries) %>%
  mutate(pop_category =
    case_when(pop >= 1000000000 ~ "billions",
              pop >= 100000000 ~ "hundred millions",
              pop < 100000000 ~ "fewer than hundred million")
  mutate(pop_category = fct_relevel(pop_category,
                                    level = c("fewer than hundred million",
                                                "hundred millions",
                                                "billions"))) %>%

ggplot() +
  aes(y = pop_category) +
  aes(x = gdpPercap) +
  geom_point()

```




```

my_countries <- c("Brazil", "France", "Russia",
                 "United Kingdom", "United States",
                 "China", "India", "Japan")

gapminder %>%
  filter(year == 2002) %>%
  filter(country %in% my_countries) %>%
  mutate(pop_category =
    case_when(pop >= 1000000000 ~ "billions",
              pop >= 100000000 ~ "hundred millions",
              pop < 100000000 ~ "fewer than hundred million")
  mutate(pop_category = fct_relevel(pop_category,
                                   level = c("fewer than hundred million",
                                             "hundred millions",
                                             "billions"))) %>%

  ggplot() +
  aes(y = pop_category) +
  aes(x = gdpPercap) +
  geom_point() +
  ggrepel::geom_text_repel(aes(label = country))

```

