

cm005 Worksheet: Exploring Geometric Objects

Preliminary

Begin by loading the required packages. If you don't have these installed (or don't know whether you have them installed), you can install them by executing the following code in your console:

```
install.packages("tidyverse")
install.packages("scales")
install.packages("tsibble")
```

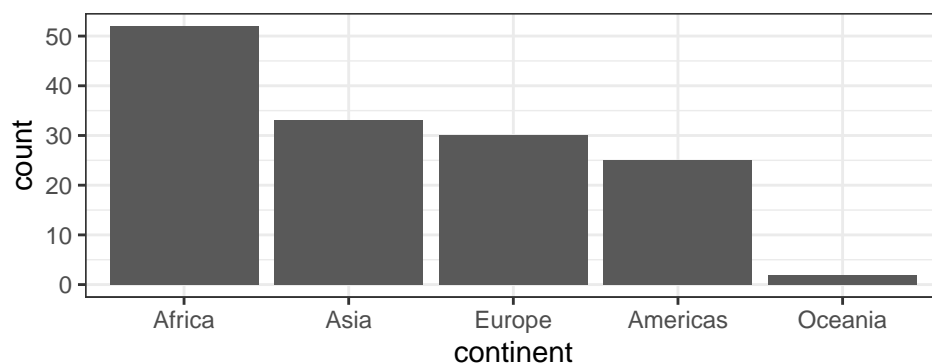
Now run this code chunk to load the packages:

```
suppressPackageStartupMessages(library(tidyverse))
suppressPackageStartupMessages(library(gapminder))
suppressPackageStartupMessages(library(scales))
suppressPackageStartupMessages(library(tsibble))
knitr::opts_chunk$set(fig.align = "center")
```

Exercise 1: Bar Chart Grammar (Together)

Consider the following plot. Don't concern yourself with the code at this point.

```
gapminder %>%
  filter(year == 2007) %>%
  mutate(continent = fct_infreq(continent)) %>%
  ggplot(aes(continent)) +
  geom_bar() +
  theme_bw()
```



Fill in the seven grammar components for this plot.

Grammar Component	Specification
data	gapminder
aesthetic mapping	x:continent, y: count
geometric object	bar

Grammar Component	Specification
scale	linear
statistical transform	count of countries in each continent
coordinate system	rectangular
facetting	none

Exercise 2: ggplot2 Syntax (Your Turn)

The following is a tsibble (a special type of tibble containing time series data, which we'll see more of later), stored in the variable `mauna`, of CO₂ concentrations collected monthly at the Mauna Loa station.

Execute this code to store the data in `mauna`:

```
(mauna <- tsibble::as_tsibble(co2) %>%
  rename(month = index, conc = value))
```

```
## # A tsibble: 468 x 2 [1M]
##   month conc
##   <mth> <dbl>
## 1 1959 Jan  315.
## 2 1959 Feb  316.
## 3 1959 Mar  316.
## 4 1959 Apr  318.
## 5 1959 May  318.
## 6 1959 Jun  318
## 7 1959 Jul  316.
## 8 1959 Aug  315.
## 9 1959 Sep  314.
## 10 1959 Oct 313.
## # ... with 458 more rows
```

2(a)

Produce a line chart showing the concentration over time. Specifically, the plot should have the following grammar components:

Grammar Component	Specification
data	<code>mauna</code>
aesthetic mapping	x: month, y: conc
geometric object	lines
scale	linear
statistical transform	none
coordinate system	rectangular
facetting	none

Fill in the blanks to obtain the plot:

```
p <- ggplot(mauna, aes(month, conc)) +
  geom_line()
```

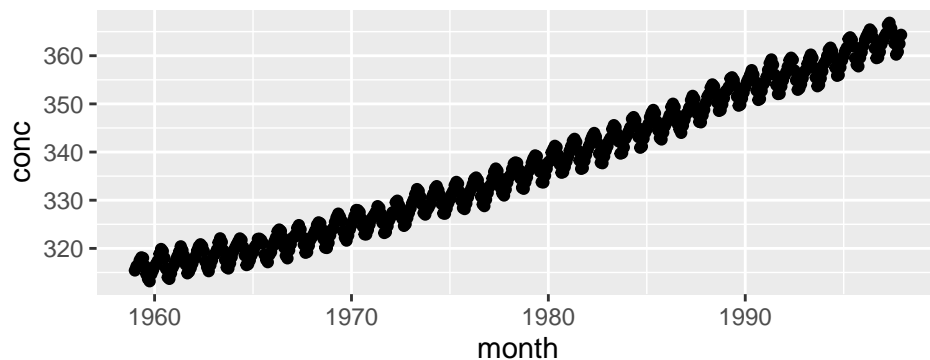
2(b)

It turns out that you're allowed to specify the aesthetic mappings in a `geom` layer instead of, or in addition to, in the `ggplot()` function, with the following rules:

- Aesthetics appearing in a `geom` layer apply only to that layer.
- If there are conflicting aesthetics in both the `ggplot()` function and the `geom` layer, the `geom` layer takes precedence.

The following code mistakenly puts the month variable on the y-axis. Fill in the `FILL_THIS_IN` so that you still obtain the same result as above.

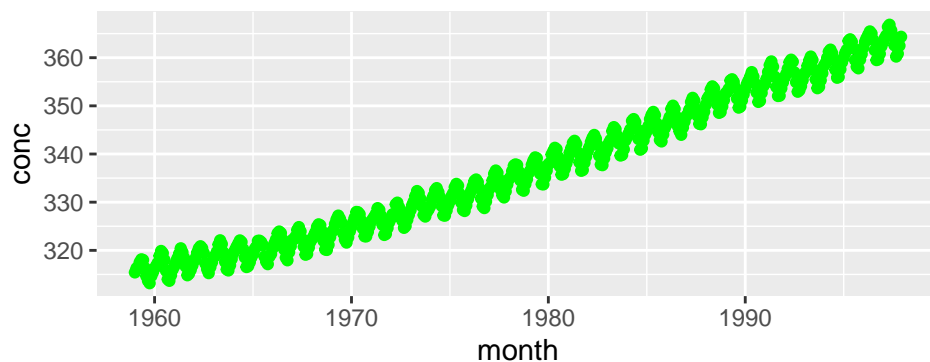
```
ggplot(mauna, aes(y = month)) +  
  geom_point(aes(month, conc)) +  
  labs(y="conc")
```



2(c)

You can store the output of the plot in a variable, too. Store the plot from 2(a) in the variable named `p`, then add a layer to `p` that adds green points to the plot.

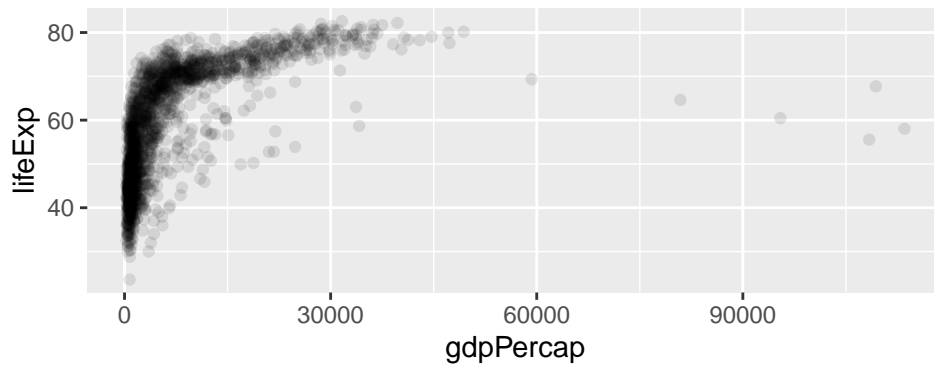
```
p +  
  geom_point(colour = 'green')
```



2(d)

What's wrong with the following code? Fix it.

```
ggplot(gapminder) +  
  geom_point(aes(x = gdpPercap, y = lifeExp), alpha = 0.1)
```



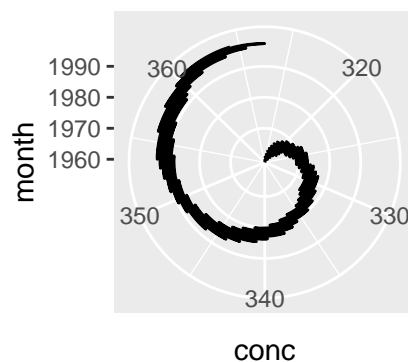
2(e) BONUS

So you're a ggplot2 pro? Then, let's see this plot adapted to polar coordinates. Specifically:

- angle is month (January through December)
- radius is CO₂ concentration

The plot should look like a spiral, or concentric circles.

```
ggplot(mauna, aes(month, conc)) +  
  geom_line() +  
  coord_polar(theta = "y")
```

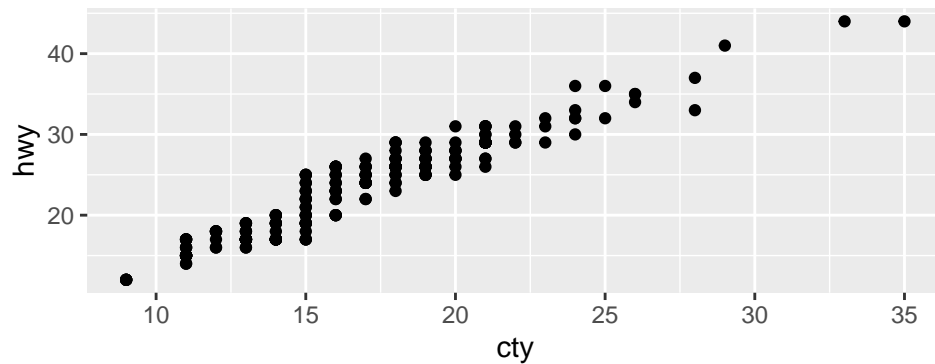


Exercise 3: Fix the plots (Together)

3(a)

Fix the following plot (attribution: “R for data science”).

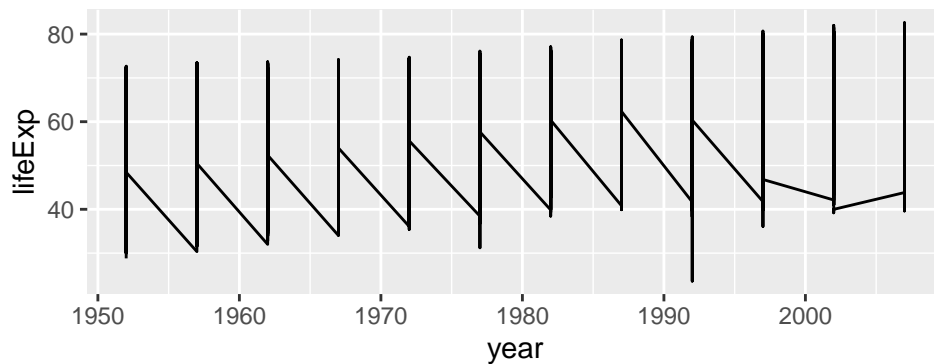
```
ggplot(mpg, aes(cty, hwy)) +  
  geom_point()
```



3(b)

Fix this plot so that it shows life expectancy over time *for each country*.

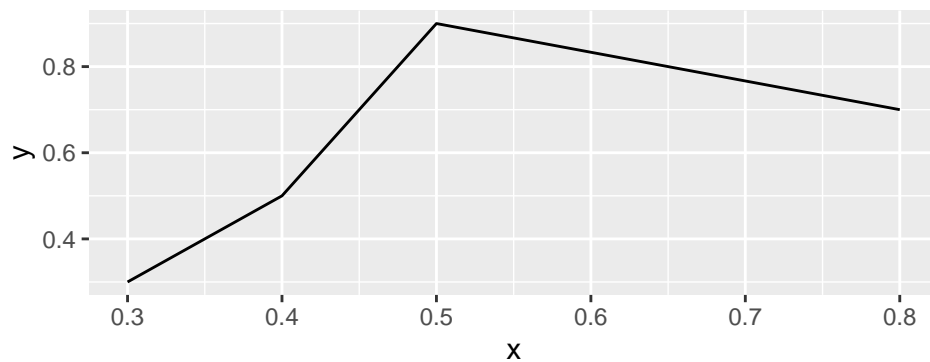
```
ggplot(gapminder, aes(year, lifeExp)) +  
  geom_line()
```



3(c)

The following mock data set marks the (x,y) position of a caribou at four time points. Fix the plot below so that it shows the path of the caribou. Add an arrow with `arrow = arrow()`.

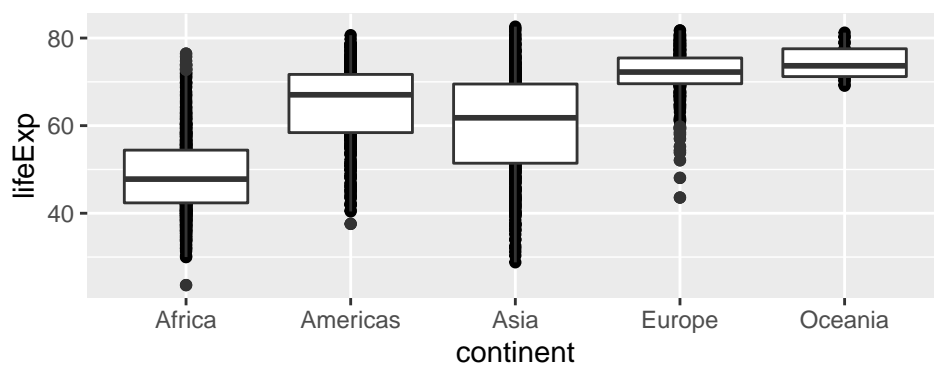
```
motion <- tribble(  
  ~time, ~x, ~y,  
  1, 0.3, 0.3,  
  2, 0.8, 0.7,  
  3, 0.5, 0.9,  
  4, 0.4, 0.5  
)  
ggplot(motion, aes(x, y)) +  
  geom_line()
```



3(d)

The following plot attempts to put both the raw data and boxplots together, but it's hard to see the raw data. Fix the plot.

```
ggplot(gapminder, aes(continent, lifeExp)) +  
  geom_point() +  
  geom_boxplot()
```



3(e)

Change the following plot so that it shows *proportion* on the y-axis, not count.

```
ggplot(mtcars, aes(cyl)) +  
  geom_bar()
```

