

Build exercises (and provide software)





**When you save into a
computer's memory,
you do the programming.**

**When you save into a
student's memory,
you do the programming.**

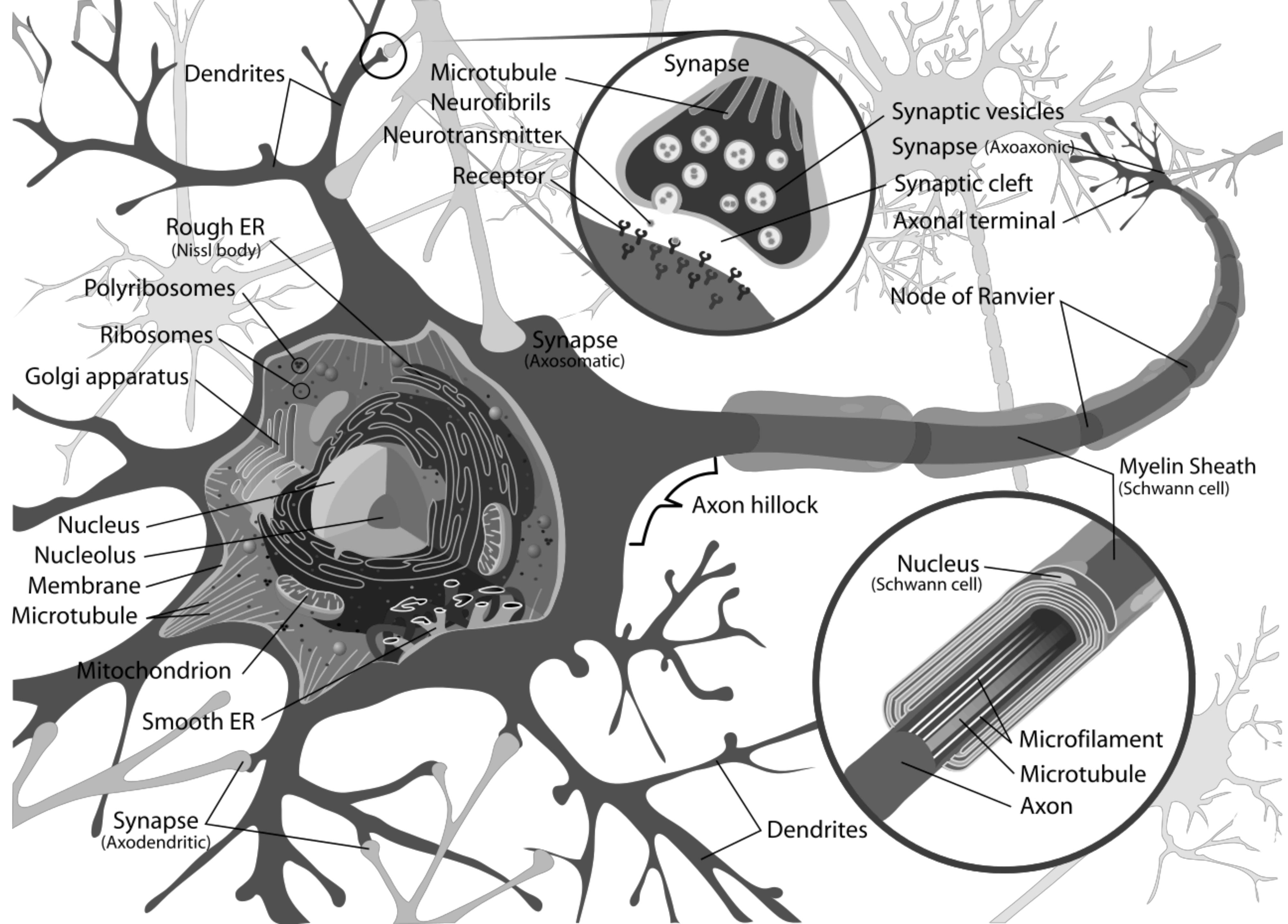


Goal

Rewrite your lesson to include 2-5 topical, illuminating, progressive formative assessments (exercises).

Why exercises? (Why germane load?)





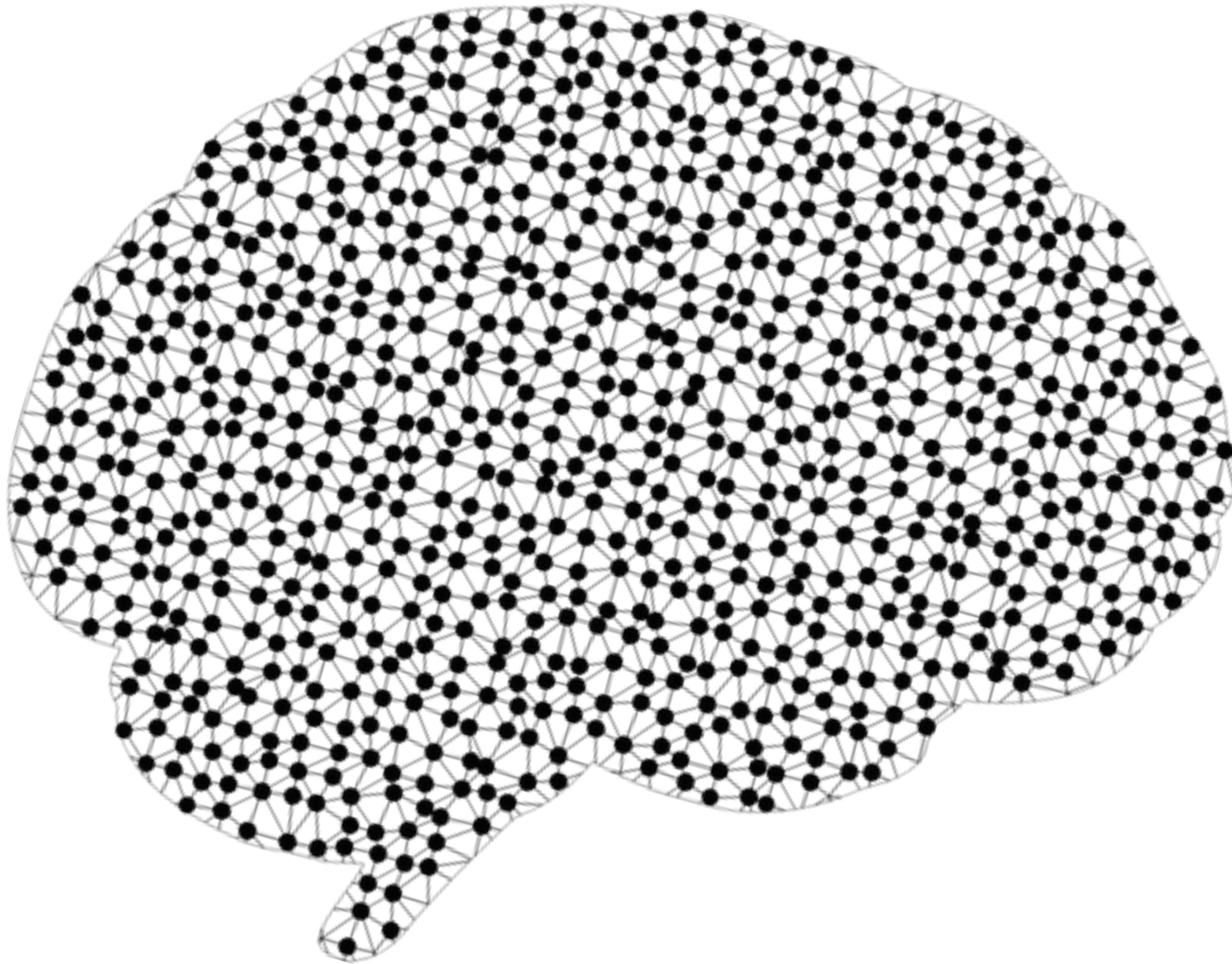
Quiz

How long is a neuron?

Quiz

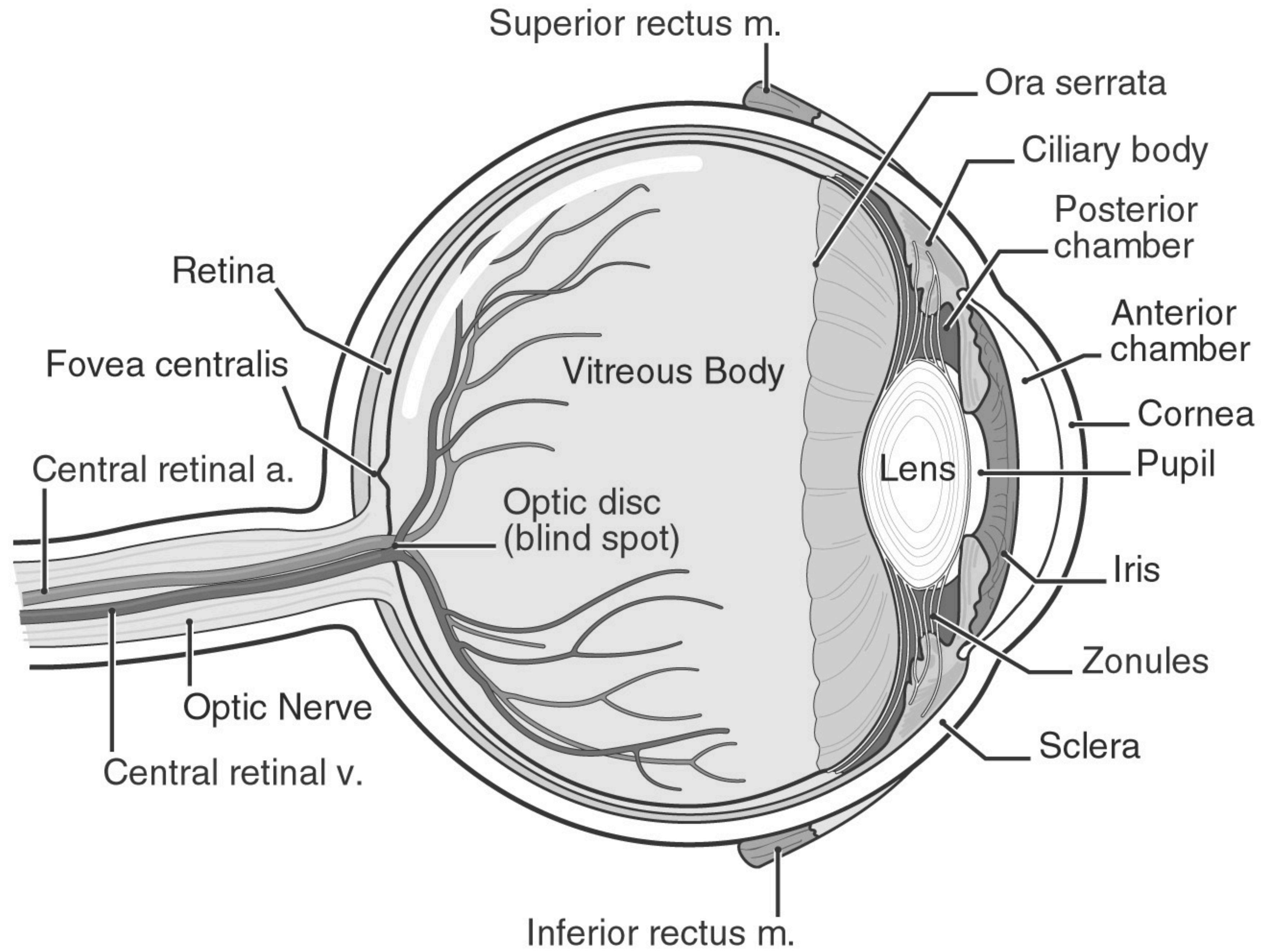
How long is a neuron?

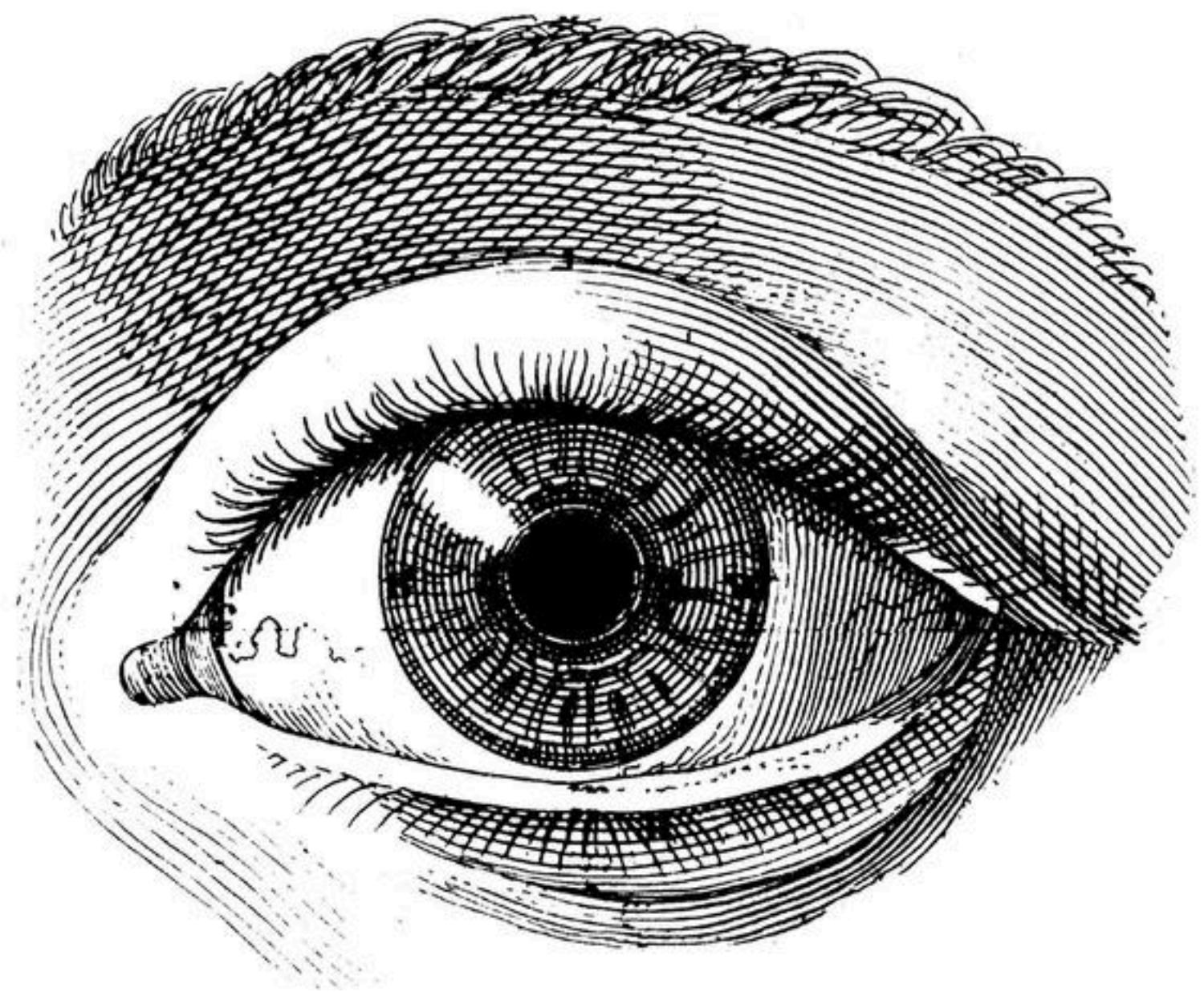
The longest can reach ~5 feet in adults

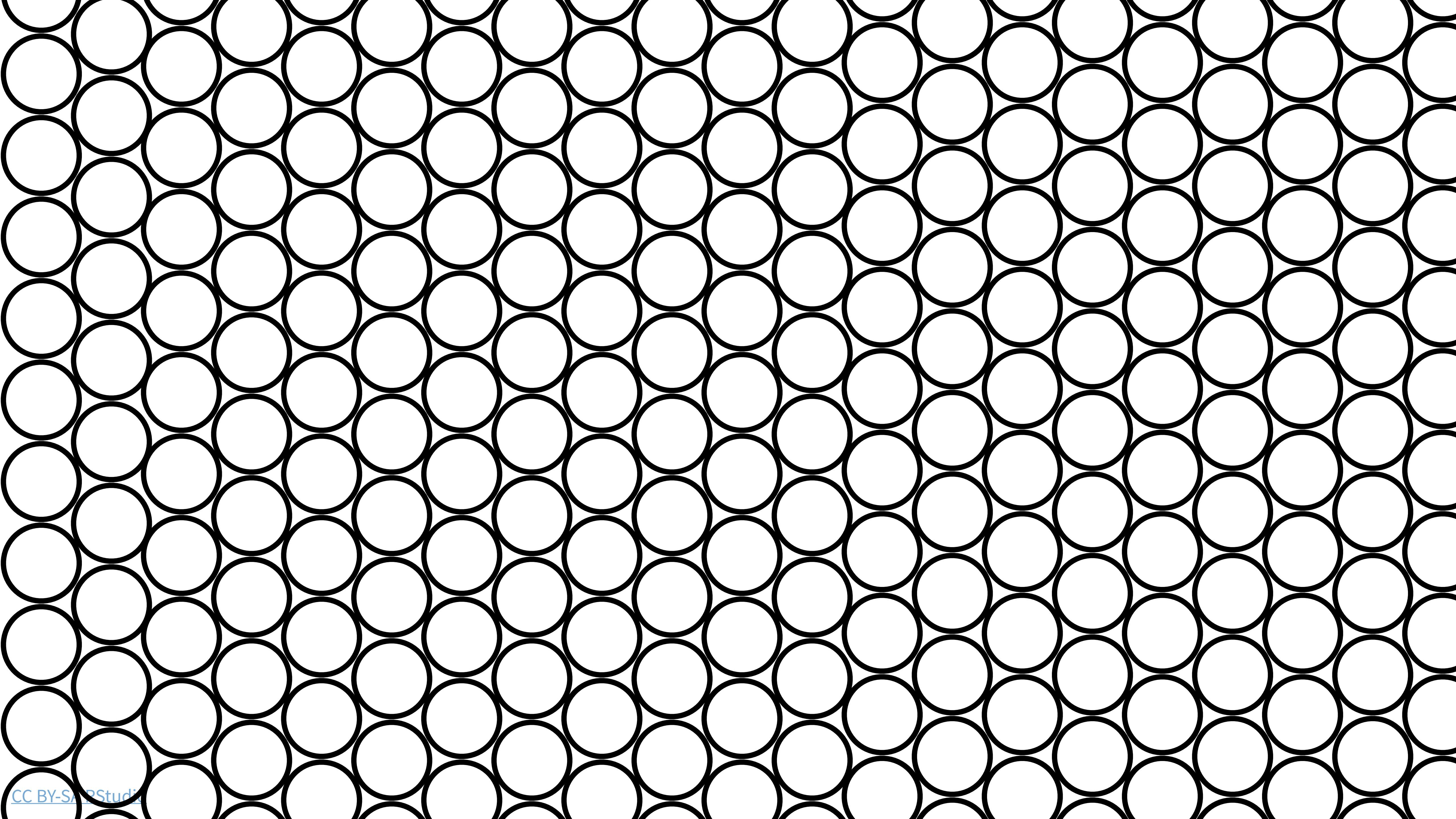


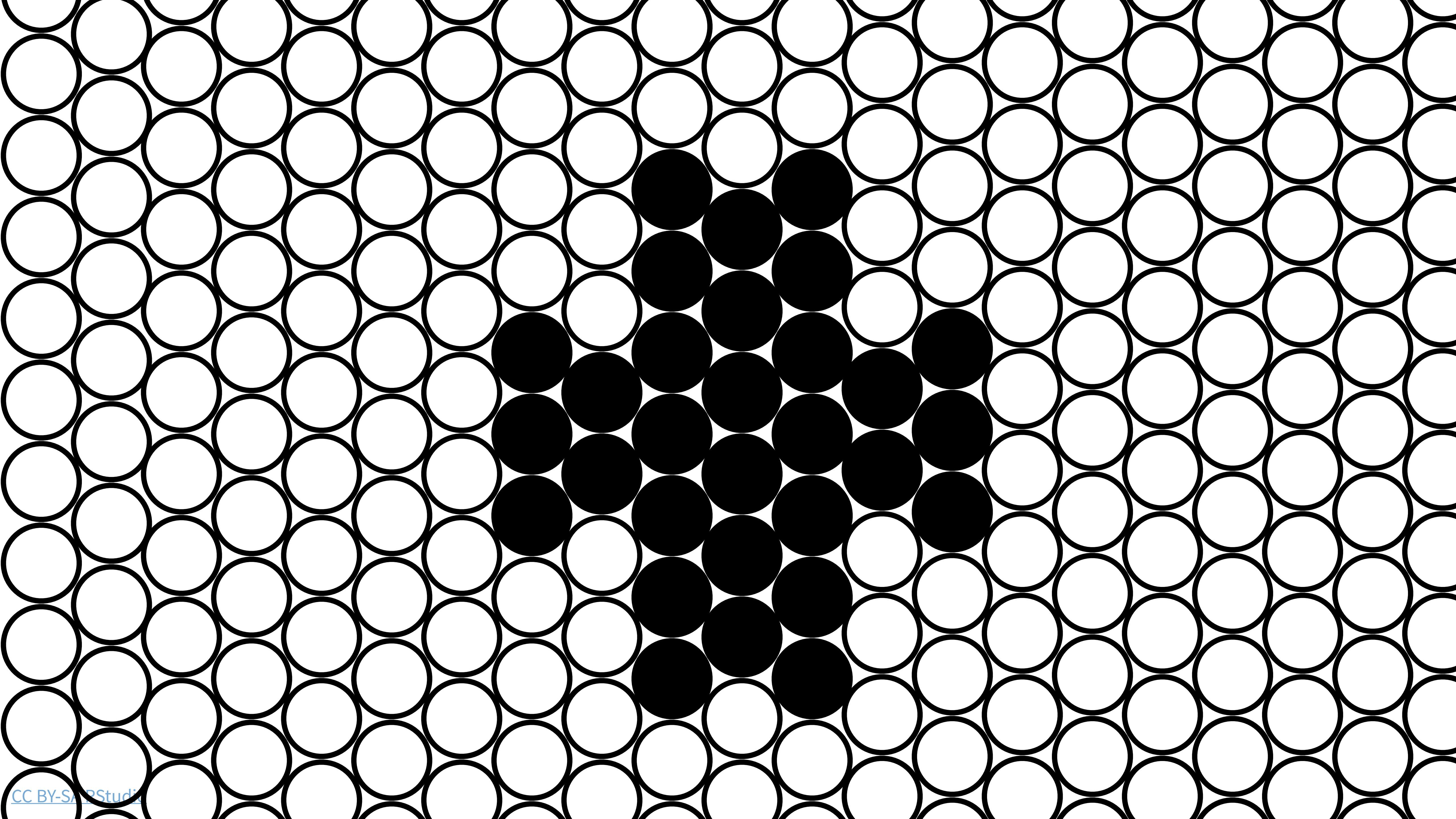
1. A neuron can represent information

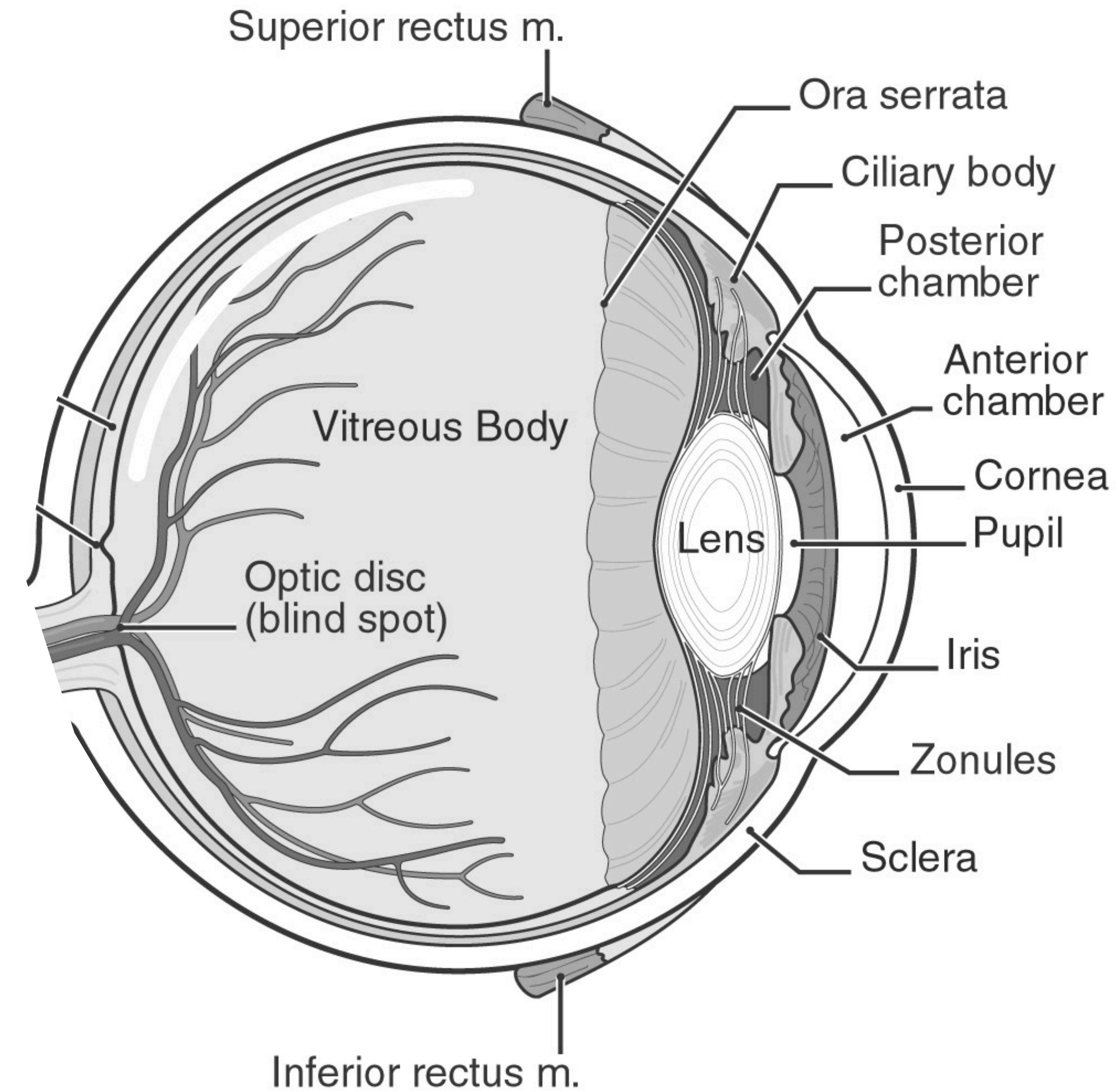
Do you accept that?

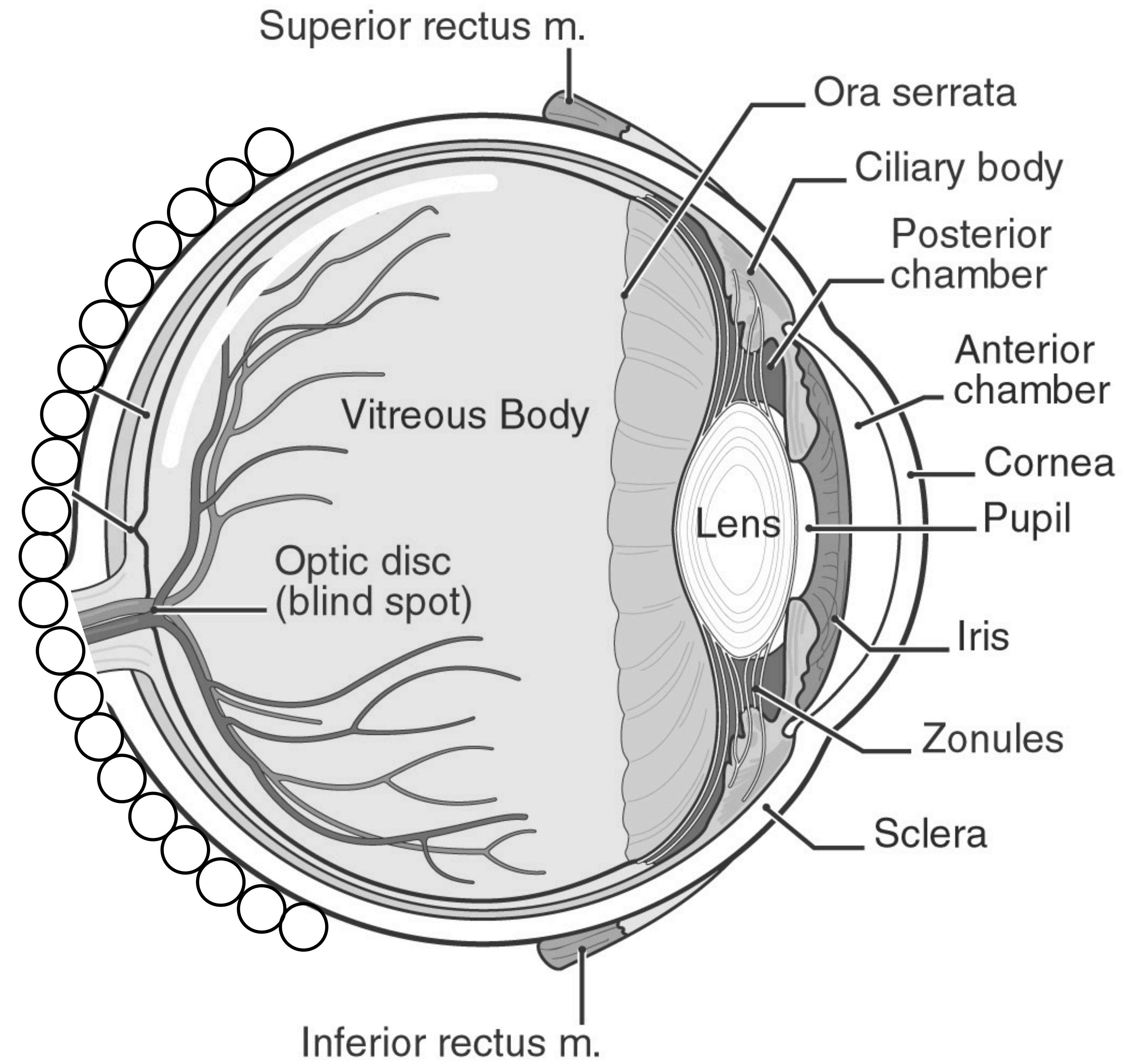


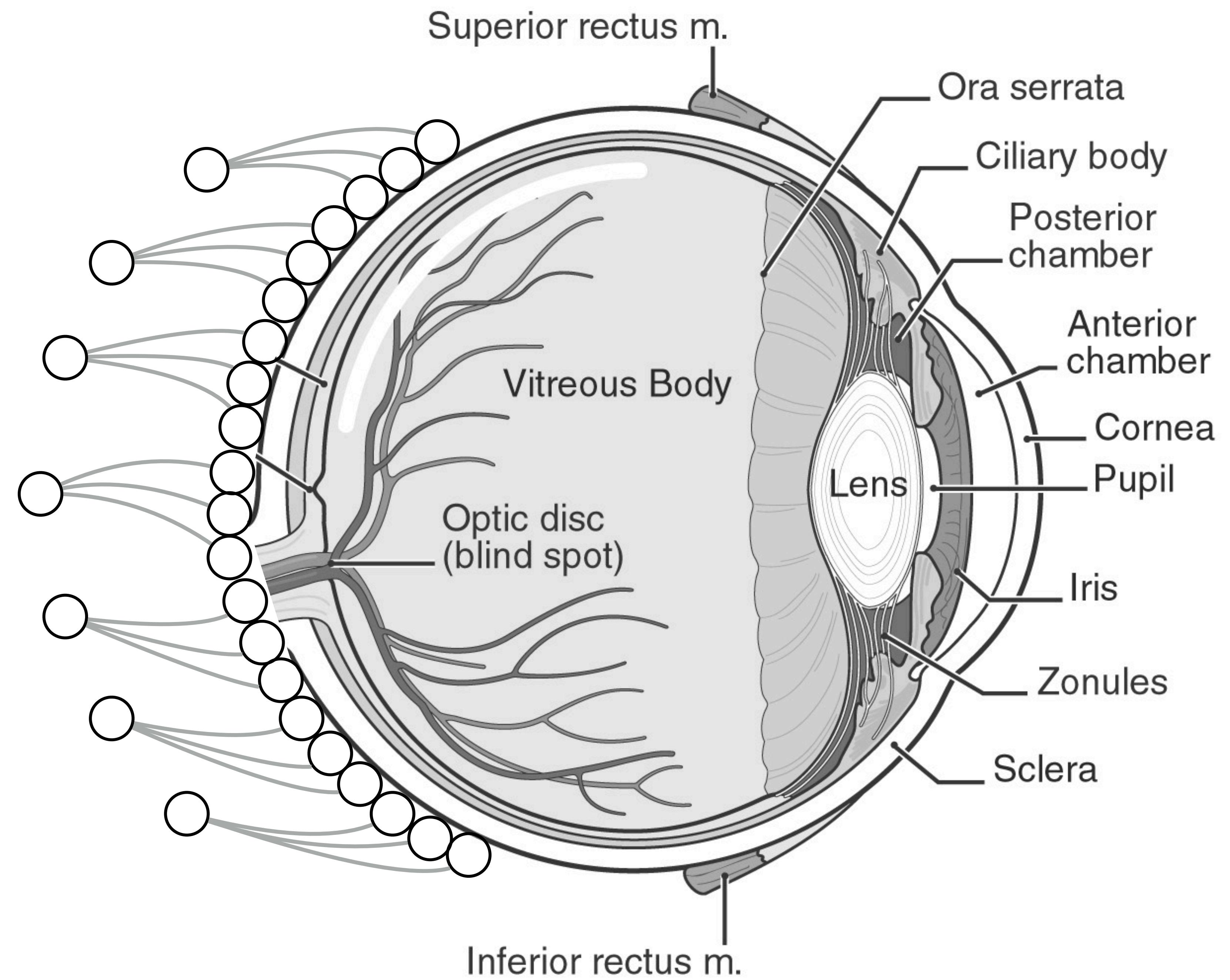


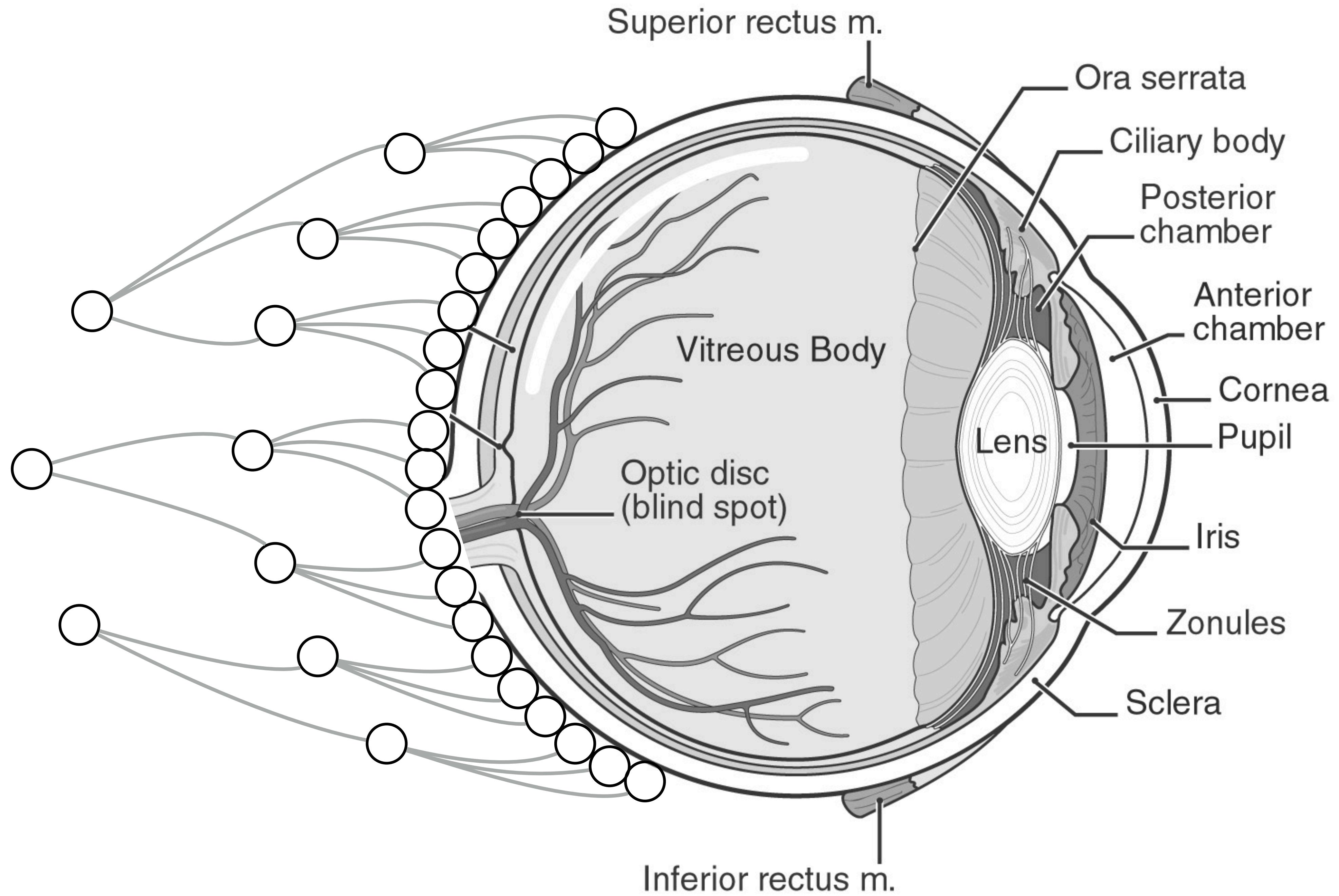


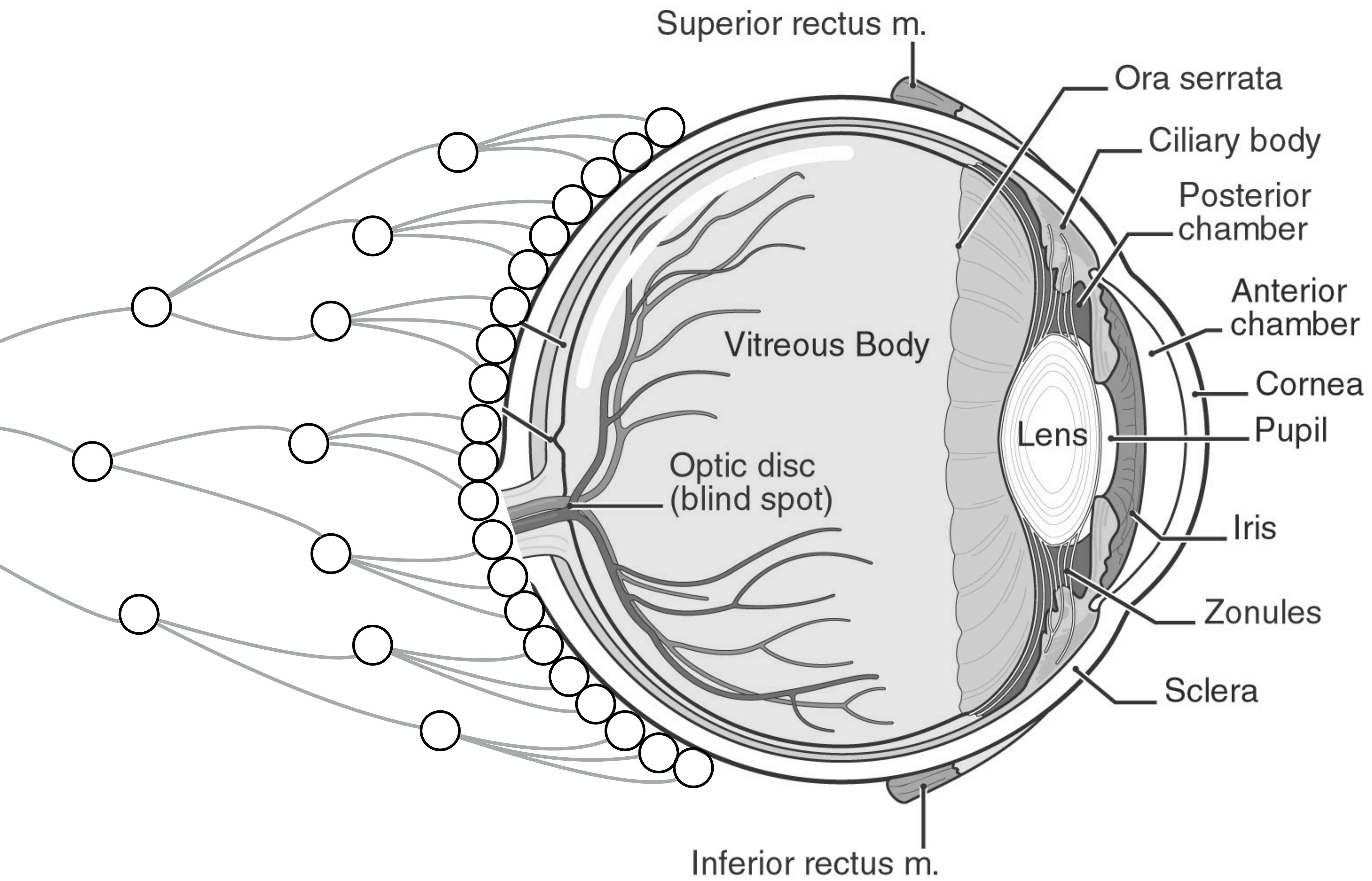


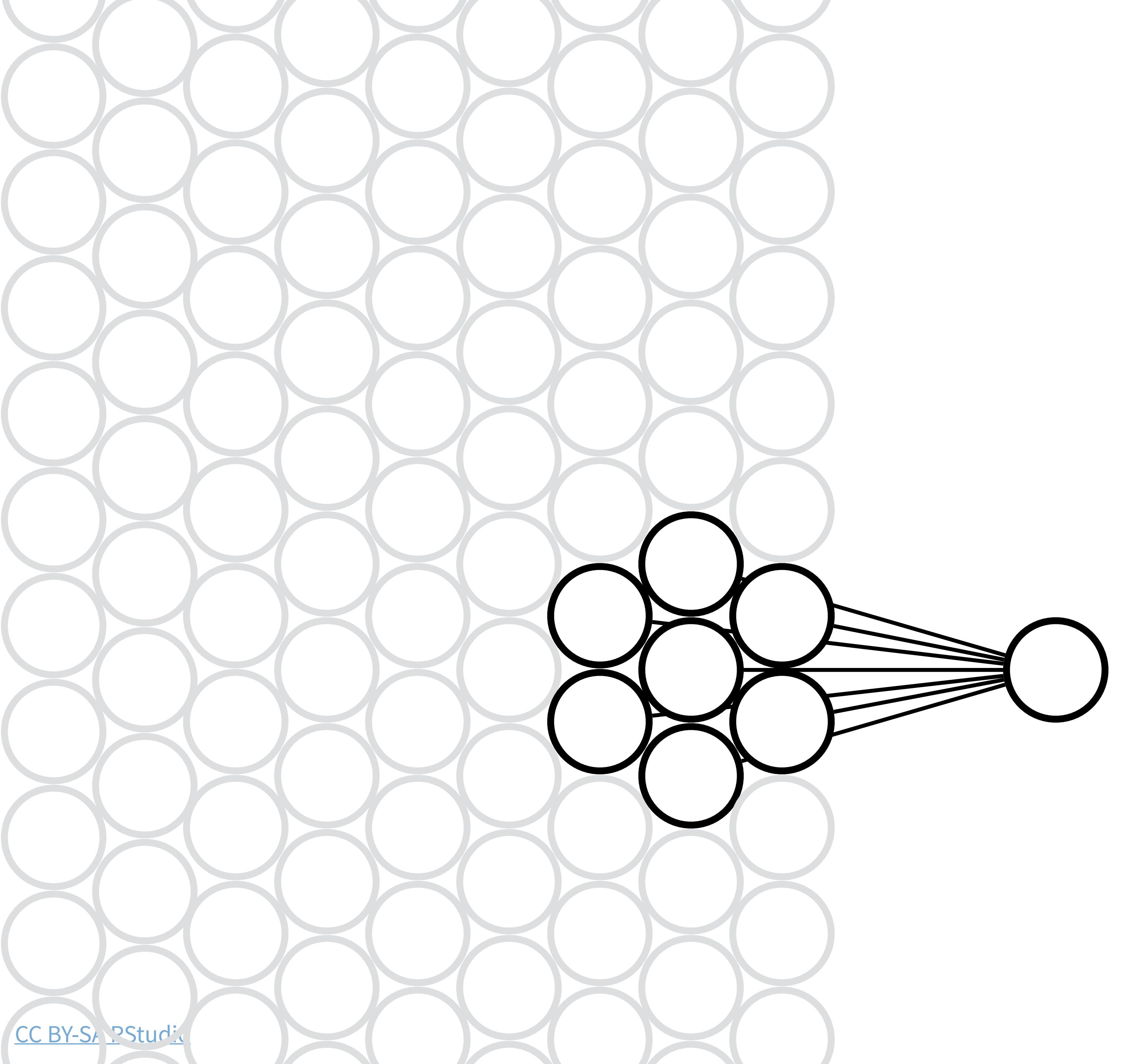


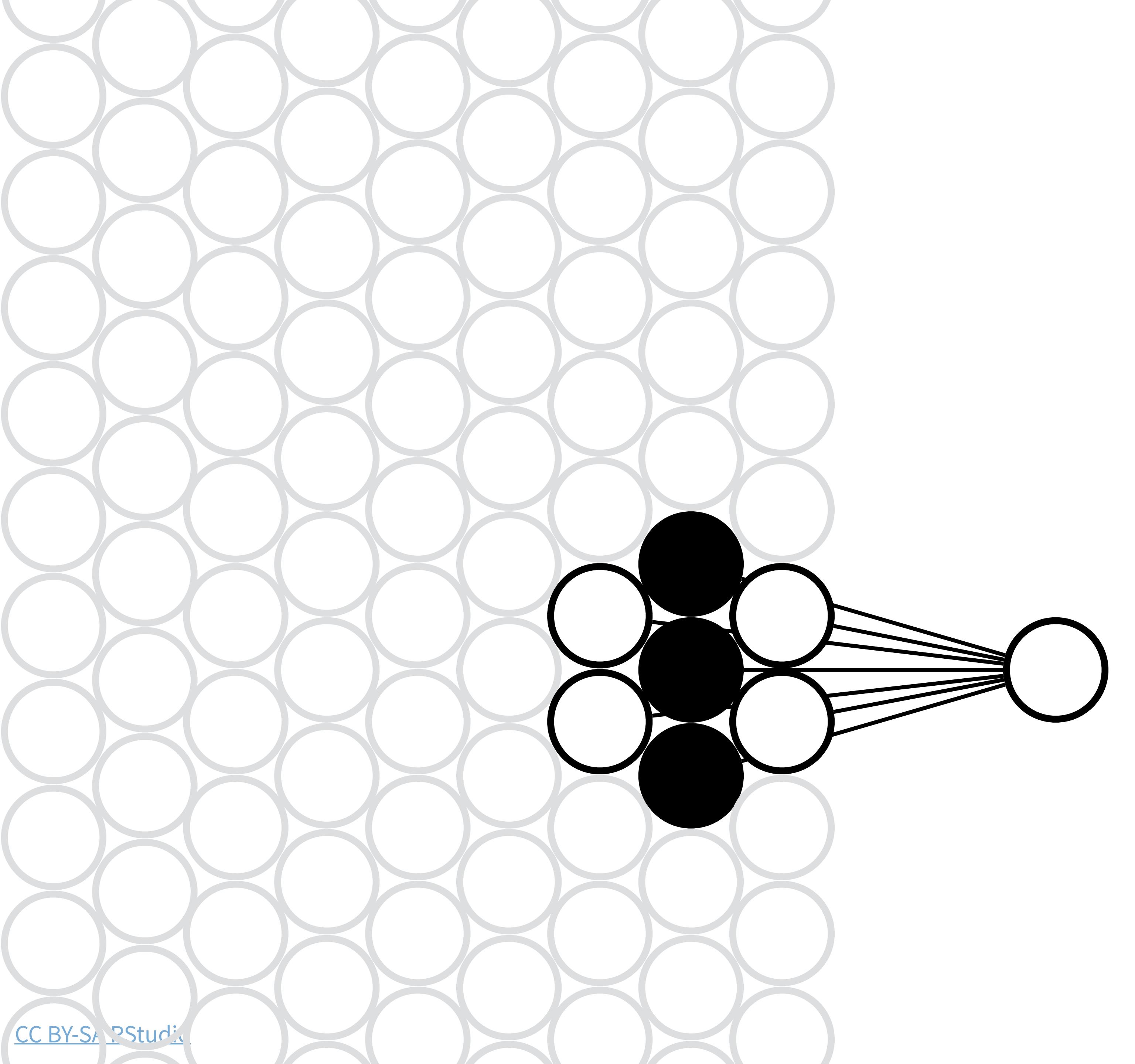


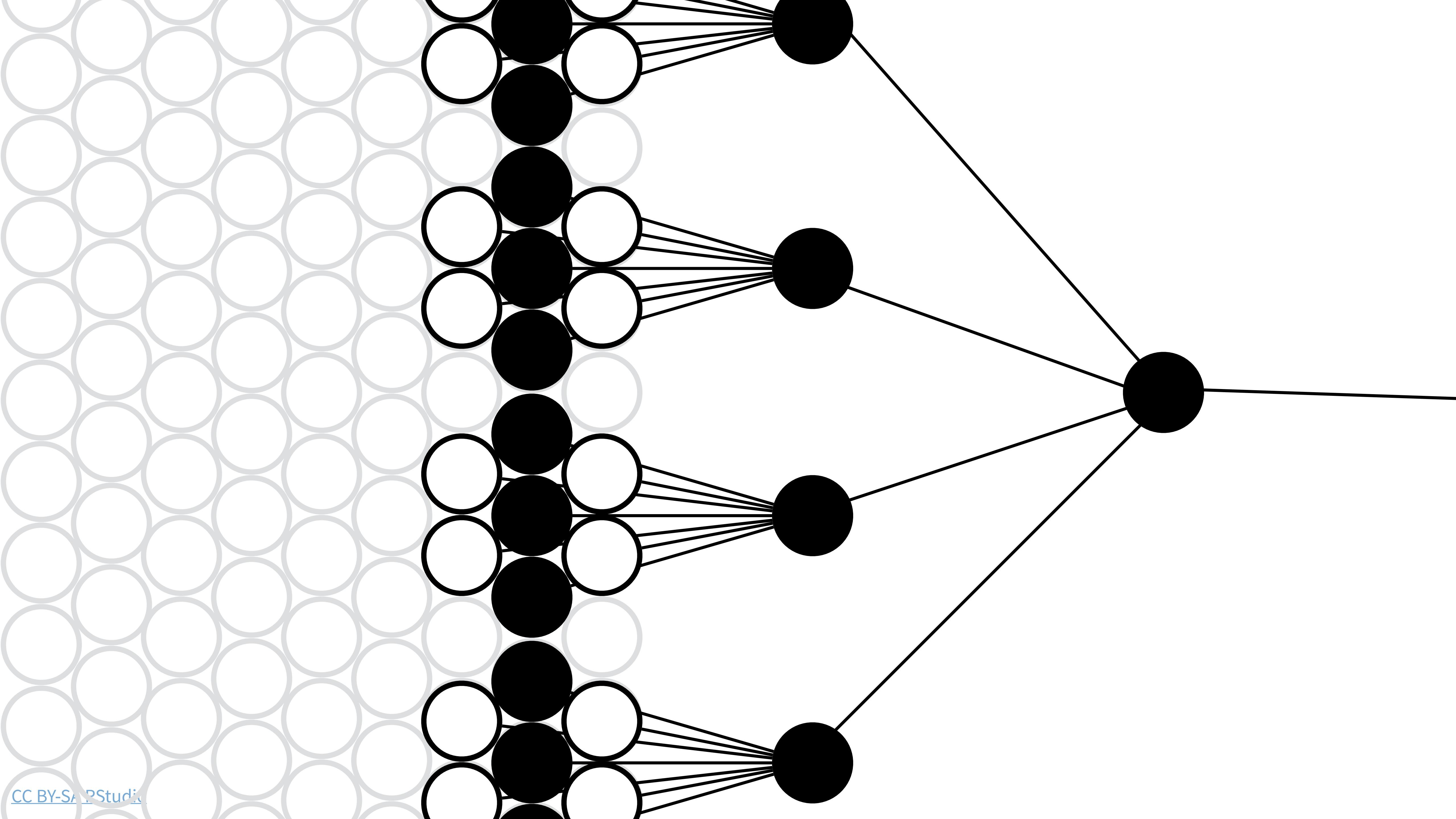






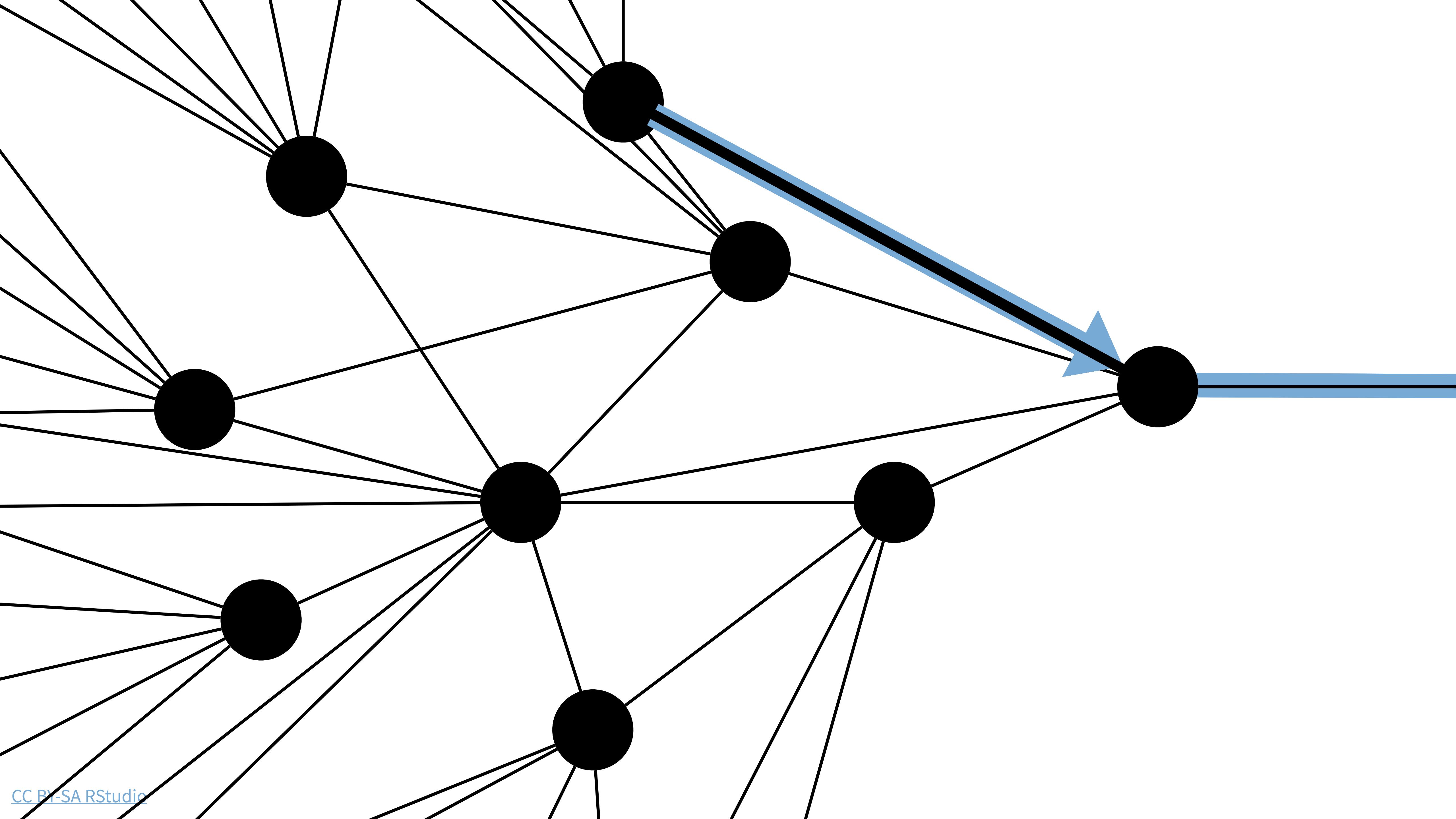


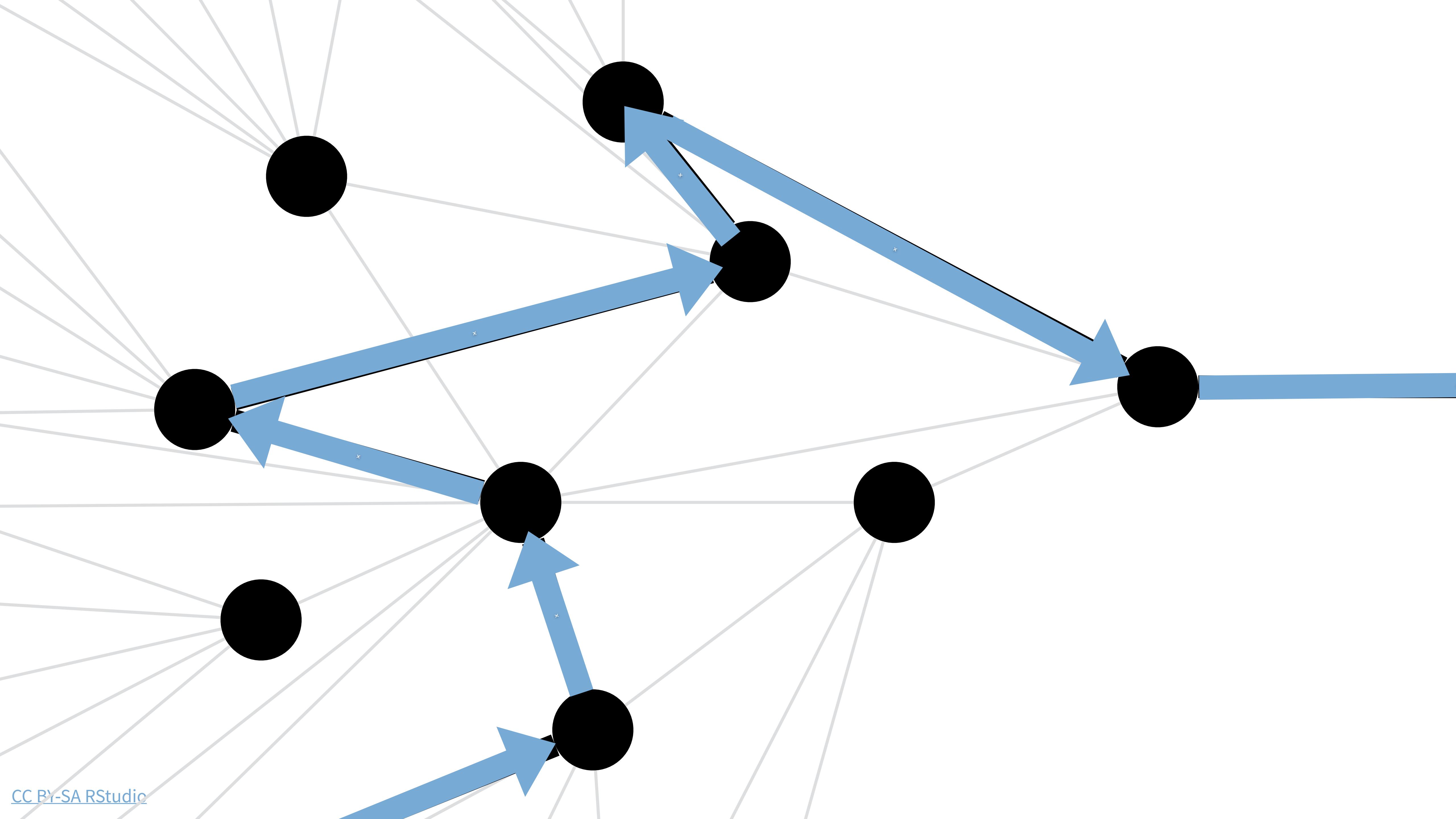




1. A neuron can represent information

1. A neuron can represent information
2. Neurons that fire together wire together





Quiz

What is the key ingredient for strengthening neural connections?

Quiz

What is the key ingredient for strengthening neural connections?

repetition
(repeated firing)

Quiz

What is the key ingredient for strengthening neural connections?

ignition + repetition
(repeated firing)

When you save into a student's memory,
the student does the programming (**or doesn't**).



How exercises? (Providing Software)

R

Quiz

Where do you get R?

Quiz

Where do you get R?

cran.r-project.org

Quiz

Where do you get R Packages?

Quiz

Where do you get R Packages?

cran.r-project.org

github.com

(from R)

Quiz

Where do you get RStudio?

Quiz

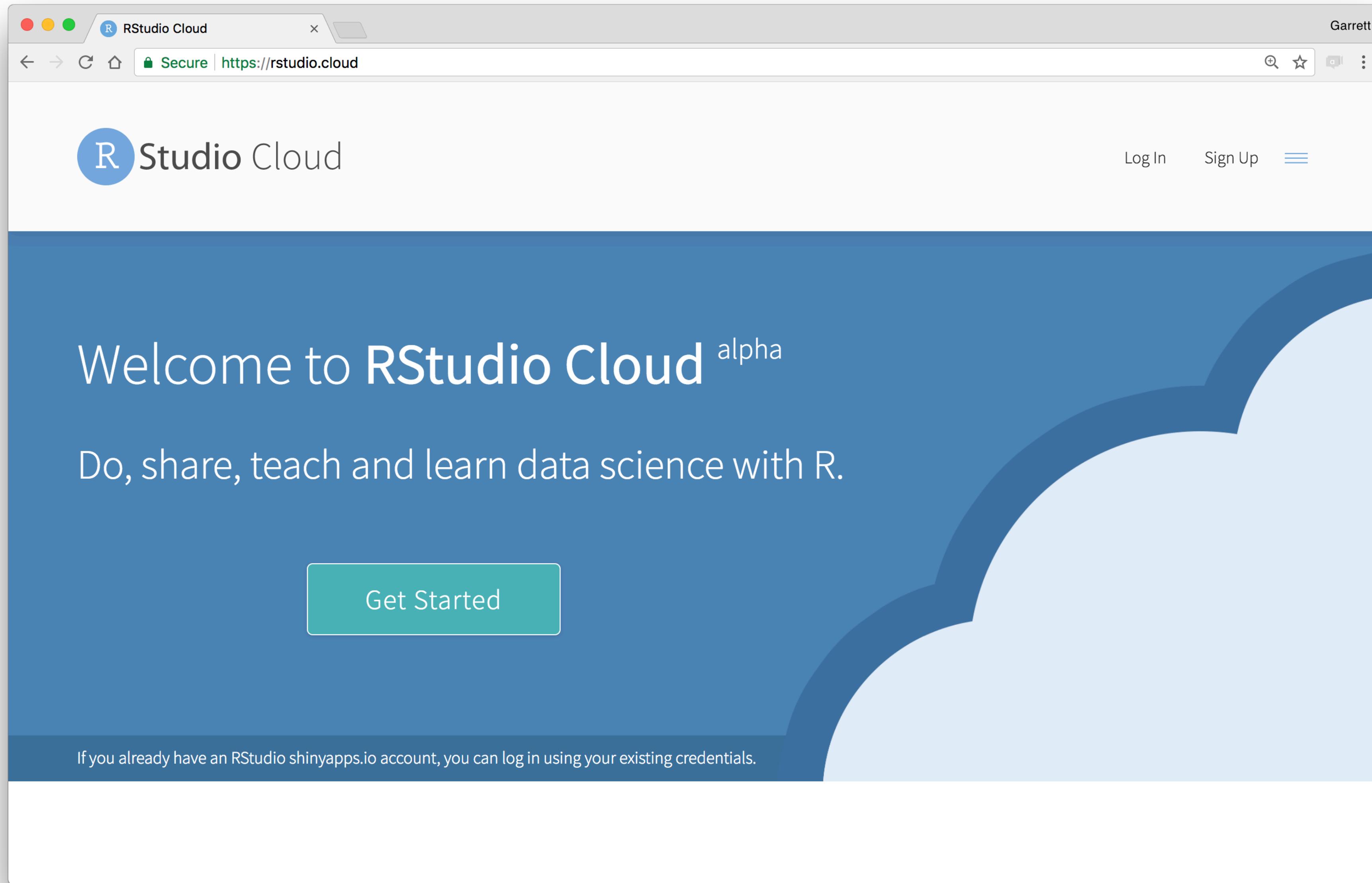
Where do you get RStudio?

www.rstudio.com/download

Go here now:

<https://rstudio.cloud/project/168691>

<https://rstudio.cloud>



Your Turn

Save a permanent copy of
<https://rstudio.cloud/project/168691>.

Read the RStudio Cloud handout. Use it to access the online guide and complete the handout challenges.



Quiz

Should you have students install R?

Your Turn

Should you have students install R?

Write down of a couple of good reasons to have students install R and a couple of good reasons to **not** have them install R.

Think

01 : 00

Share

03 : 00

Your Turn

Should you have students install R?

Debate in your group and vote on one or another.

Think

01 : 00

Pair

03 : 00

Installation

Thank you for enrolling in Master the Tidyverse.

Please bring a laptop to class that has the following installed:

1. A recent version of R (~3.4.3), which is available for free at <cran.r-project.org>
2. A recent version of RStudio IDE (~1.1.4), available for free at <www.rstudio.com/download>
3. The R packages we will use, which you can install by connecting to the internet, opening R, and running:

```
install.packages(c("babynames", "formatR", "gapminder", "hexbin", "mgcv", "maps",
"mapproj", "nycflights13", "tidyverse", "viridis"))
```

4. The class materials, which can be downloaded at
<<https://github.com/rstudio/master-the-tidyverse/archive/master.zip>>

And don't forget your power cord!

pre-workshop-instructions.md

R Markdown

A large, semi-transparent watermark of the R logo is positioned in the bottom right corner. The logo consists of a dark green circle containing a white letter 'R' with a diagonal line through it.

R

R Markdown

An authoring format for Data Science.

The screenshot shows the RStudio interface with an R Markdown file open. The file contains the following content:

```
1 ---  
2 title: "R Notebook"  
3 output: html_notebook  
4 ---  
5  
6 Text written in **markdown**  
7  
8 ```{r}  
9 # code written in R  
10 (x <- rnorm(7))  
11 ````  
12  
13 Text written in _markdown_  
14  
15 ```{r}  
16 # code written in R  
17 hist(x)  
18 ````  
19  
20 (Top Level) ◊
```

The R console window below shows the output of the R code:

```
[1] -1.2  1.0 -0.5  0.9 -0.6 -1.1 -1.5
```

Three callout boxes highlight features of the interface:

- A grey callout box points to the green triangle icon in the toolbar, labeled "Click to run all code chunks above".
- A grey callout box points to the green triangle icon in the chunk header bar, labeled "Click to run code in chunk".
- A dark grey callout box points to the R console output, labeled "Code result".

Your Turn

(Return to your copy of
<https://rstudio.cloud/project/168691>).

Open **exercises > 00-Introduction.Rmd.**

Read through the file and do everything it tells you to do.



Your Turn

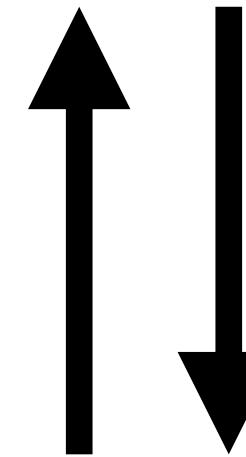
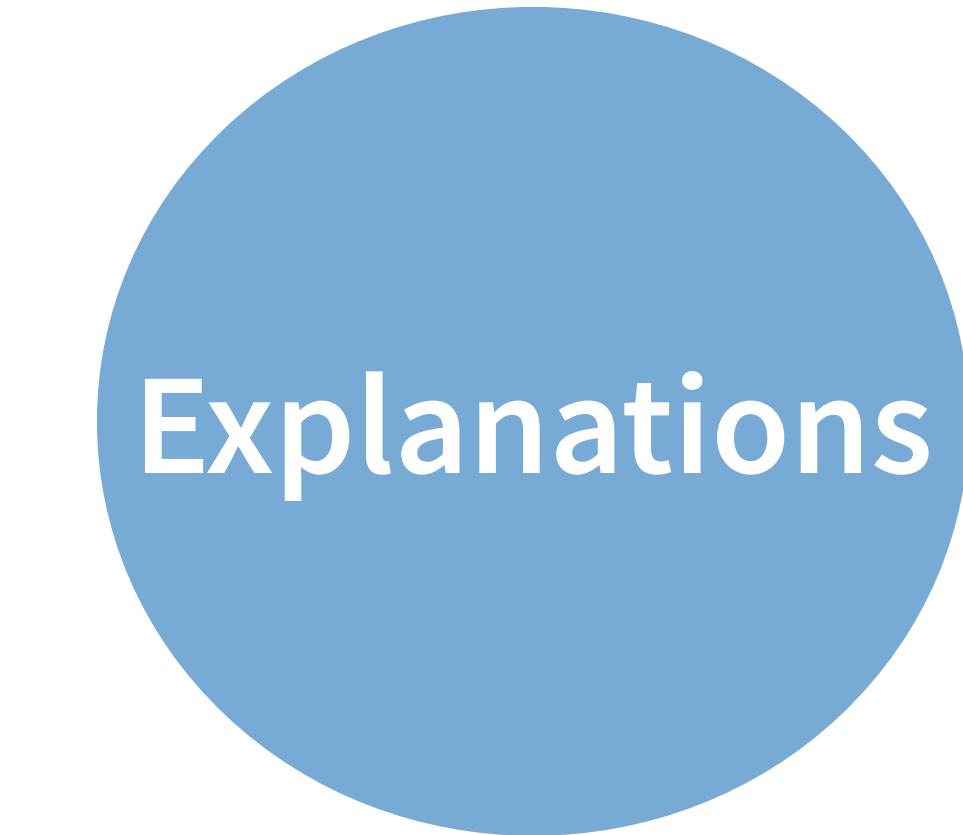
Open a new RStudio Cloud project.

Then open a new R Markdown file to record the exercises that you will write. Save the file as **exercises.rmd**.



Writing Exercises

R
R



**START
EARLY !!!**

Useful Data Sets

R

“Data! Data! Data!” he cried impatiently.
"I can't make bricks without clay."

- Sherlock Holmes, The Adventure of the Copper Beeches

In-Package Data sets

dplyr

band_instruments
band_instruments2
band_members
starwars
storms
nasa (array)

forcats

gss_cat

ggplot2

diamonds
economics
economics_long
faithful
midwest
mpg
msleep
presidential
seals
txhousing
luv_colours

lubridate

lakers

modelr

heights
sim1, sim2,
sim3, sim4

stringr

fruit (vector)
sentences (vector)
words (vector)

tidyr

population
smiths
table1
table2
table3
table4a
table4b
table5
who

Data Set Packages

babynames

applicants
babynames
births
lifetables

gapminder

continent_colors
country_codes
country_colors
gapminder
gapminder_unfiltered

nasaweather

atmos
borders
elev
glaciers
storms

nycflights13

airlines
airports
flights
planes
weather

repurrsive

gap_simple
gap_nested
gap_split (list)
gh_repos (list)
gh_repos_json
gh_repos_xml
gh_users (list)
gh_users_json
gh_users_xml
got_chars (list)
got_chars_json
got_chars_xml

sw_films (list)
sw_people (list)
sw_planets (list)
sw_species (list)
sw_starships (list)
sw_vehicles (list)
wesanderson (list)
wesanderson_json
wesanderson_xml

Unpackaged Data

Tidy Tuesday

<https://github.com/rfordatascience/tidytuesday>

Your Turn

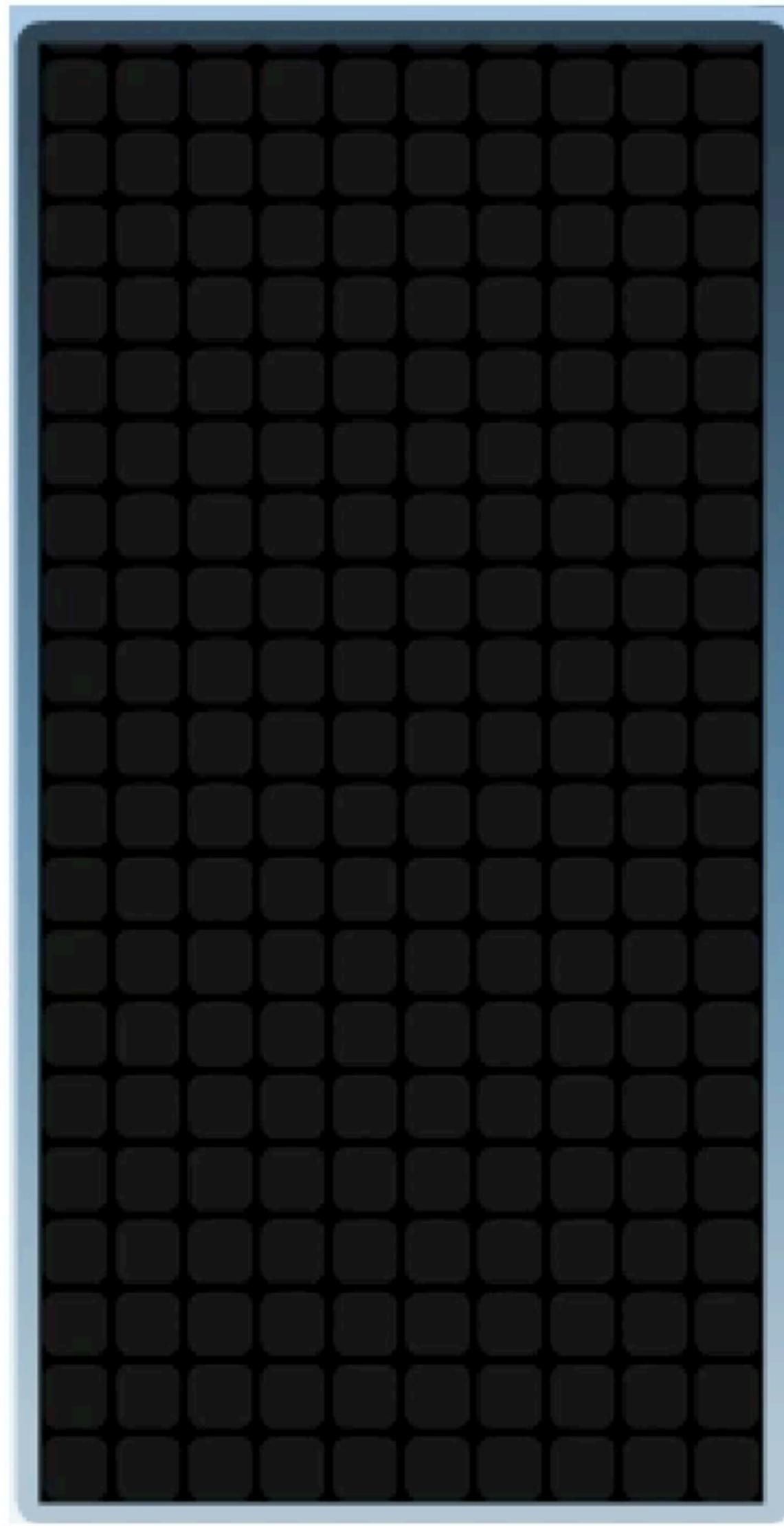
Use one of the data sets to design a realistic exercise
that applies the knowledge you wish to cover.

Your Turn

Trade exercises with a partner. Work through their exercise, and try to spot areas of excess load (tell them).



Intrinsic Load



Goal

1 self-contained idea = 1 chance for feedback

Find the most popular names by sex
from 1880 to 2015.

```
babynames %>%  
  group_by(name, sex) %>%  
  summarise(total = sum(n)) %>%  
  arrange(desc(total))  
  
# # name sex total  
# 1 James M 5120990  
# 2 John M 5095674  
# 3 Robert M 4803068  
# 4 Michael M 4323928  
# 5 Mary F 4118058  
# 6 William M 4071645  
# 7 David M 3589754  
# 8 Joseph M 2581785  
# 9 Richard M 2558165  
# 10 Charles M 2371621  
# ... with 105,376 more rows
```

Calculate the total number of children
with each name/sex combination

```
babynames %>%  
  group_by(name, sex) %>%  
  summarise(total = sum(n)) %>%  
  arrange(desc(total))
```

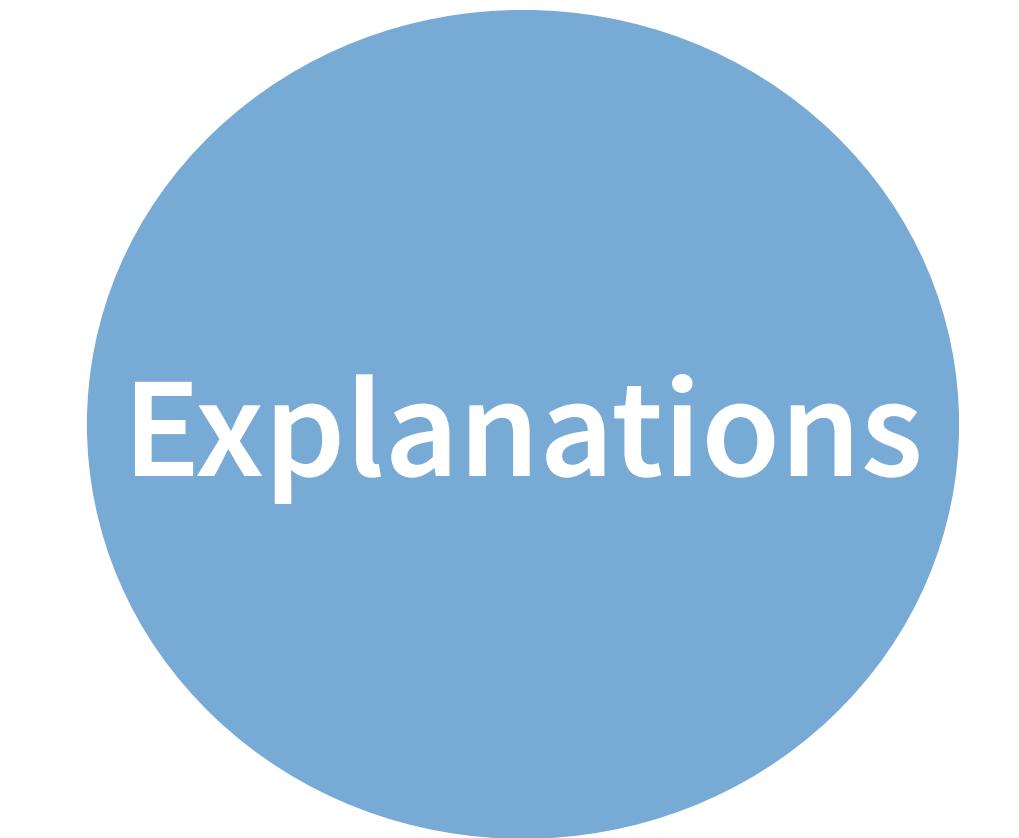
```
#          name sex total  
# 1      James M 5120990  
# 2      John  M 5095674  
# 3    Robert  M 4803068  
# 4 Michael  M 4323928  
# 5     Mary F 4118058  
# 6  William  M 4071645  
# 7     David  M 3589754  
# 8   Joseph  M 2581785  
# 9 Richard  M 2558165  
# 10  Charles M 2371621  
# ... with 105,376 more rows
```

Extend the pipe to arrange the results by total n.

```
babynames %>%  
  group_by(name, sex) %>%  
  summarise(total = sum(n))
```

```
babynames %>%  
  group_by(name, sex) %>%  
  summarise(total = sum(n)) %>%  
  arrange(desc(total))
```

```
#          name sex total  
# 1      James M 5120990  
# 2      John  M 5095674  
# 3    Robert  M 4803068  
# 4 Michael  M 4323928  
# 5     Mary F 4118058  
# 6  William  M 4071645  
# 7    David  M 3589754  
# 8   Joseph  M 2581785  
# 9 Richard  M 2558165  
# 10 Charles M 2371621  
# ... with 105,376 more rows
```



Your Turn

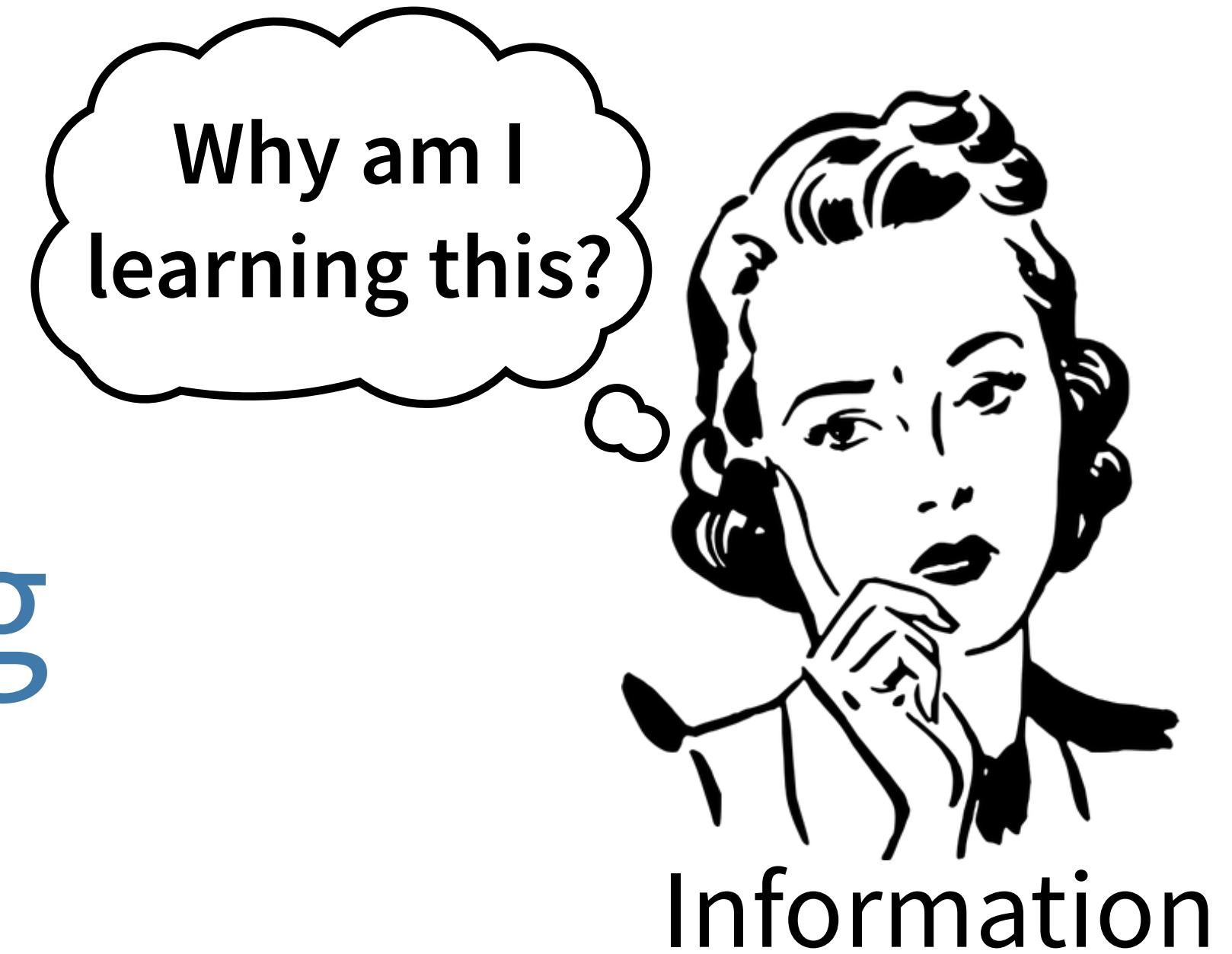
Divide your exercise into smaller steps and then add the steps to your explanation outline. Rewrite the outline as necessary.

Notice if your exercise requires any knowledge you didn't plan to cover.

Quiz

What if your students do need to know something pre-requisite and unrelated to do the exercise?

Ahead of time learning



Information

Just in time learning



Task



Task



Information

(Difficulty)

R

A

Subtract `sched_dep_time` from `dep_time` to calculate the delay for each flight.

B

Which airline's performance suffered the most whenever it rained?



A

Subtract `sched_dep_time` from `dep_time` to calculate the delay for each flight.

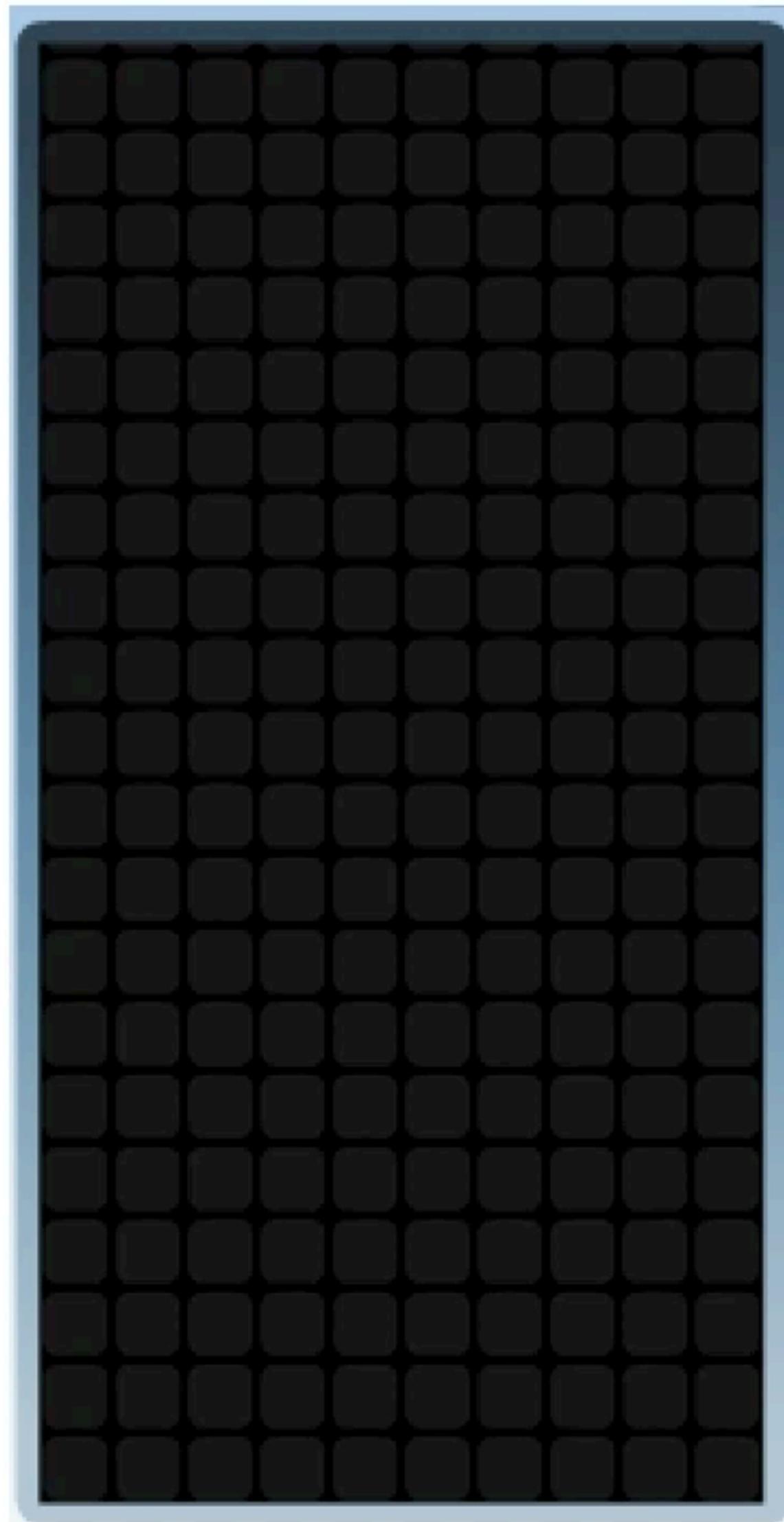


n

Which airline performed the most when untrained?

A green circular icon with a thick dark green border. Inside the circle is a stylized face with two dark green eyebrows pointing downwards, a single dark green eye, and a small dark green mouth.

Intrinsic Load



A

Subtract `sched_dep_time` from `dep_time` to calculate the delay for each flight.

B

Which airline's performance suffered the most whenever it rained?

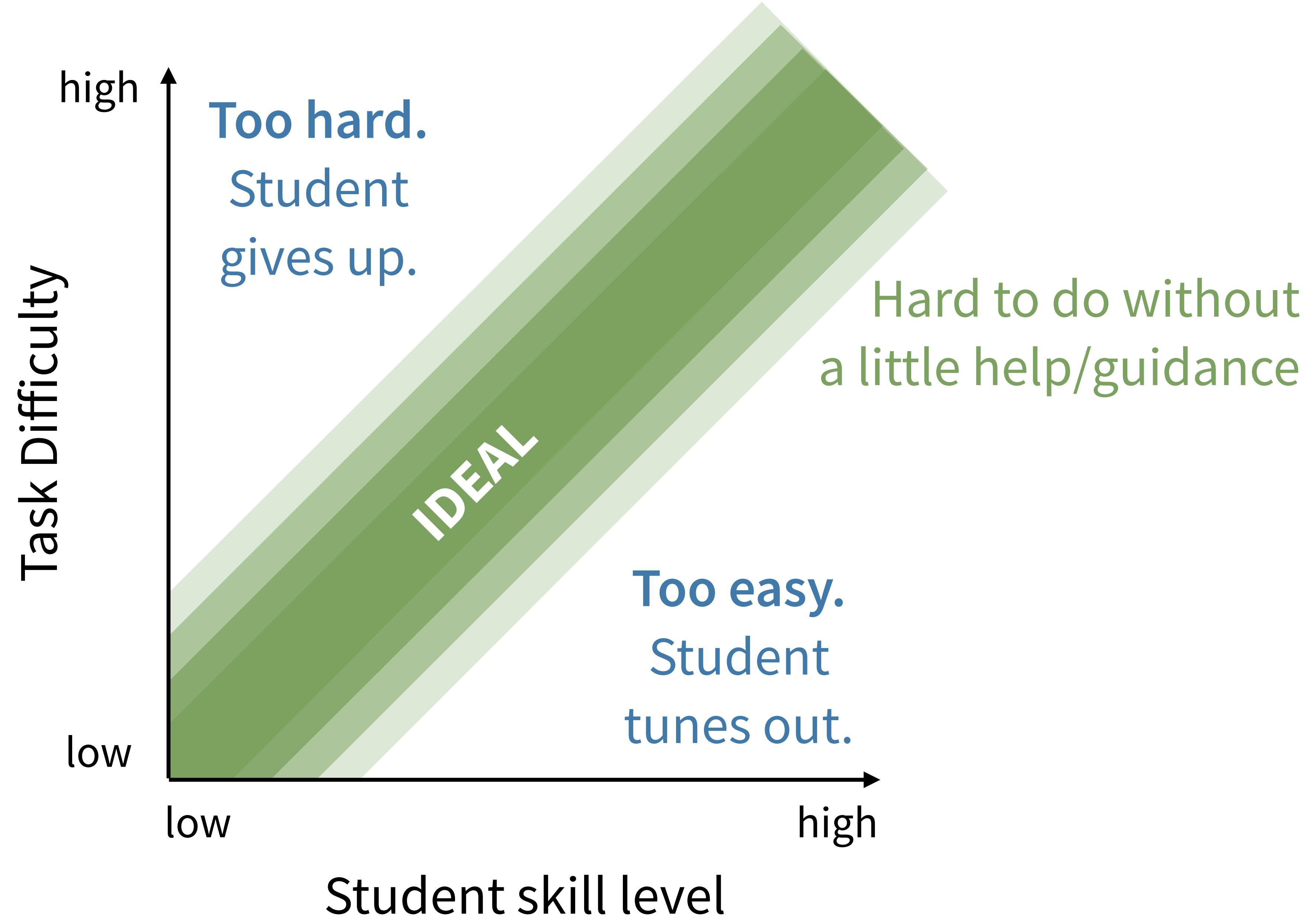


B

Subtract `dep_time`
from `arr_time` to calculate
the delay in each flight.

Which airline's
performance suffered the
most whenever it rained?





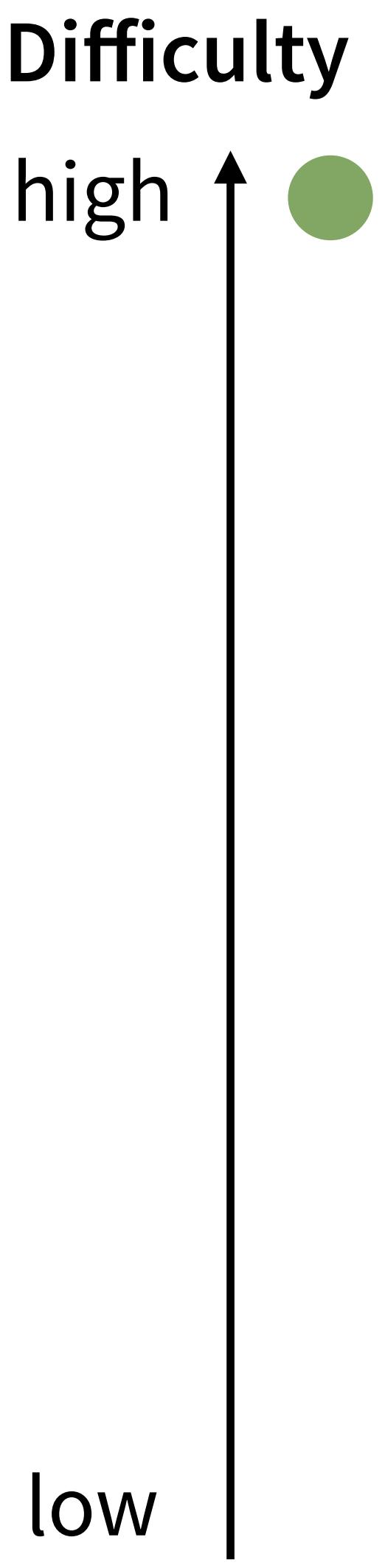
Scaffolding Exercises

Find the most popular names by sex
from 1880 to 2015.

```
babynames %>%  
  group_by(name, sex) %>%  
  summarise(total = sum(n)) %>%  
  arrange(desc(total))  
  
# # name sex total  
# 1 James M 5120990  
# 2 John M 5095674  
# 3 Robert M 4803068  
# 4 Michael M 4323928  
# 5 Mary F 4118058  
# 6 William M 4071645  
# 7 David M 3589754  
# 8 Joseph M 2581785  
# 9 Richard M 2558165  
# 10 Charles M 2371621  
# ... with 105,376 more rows
```

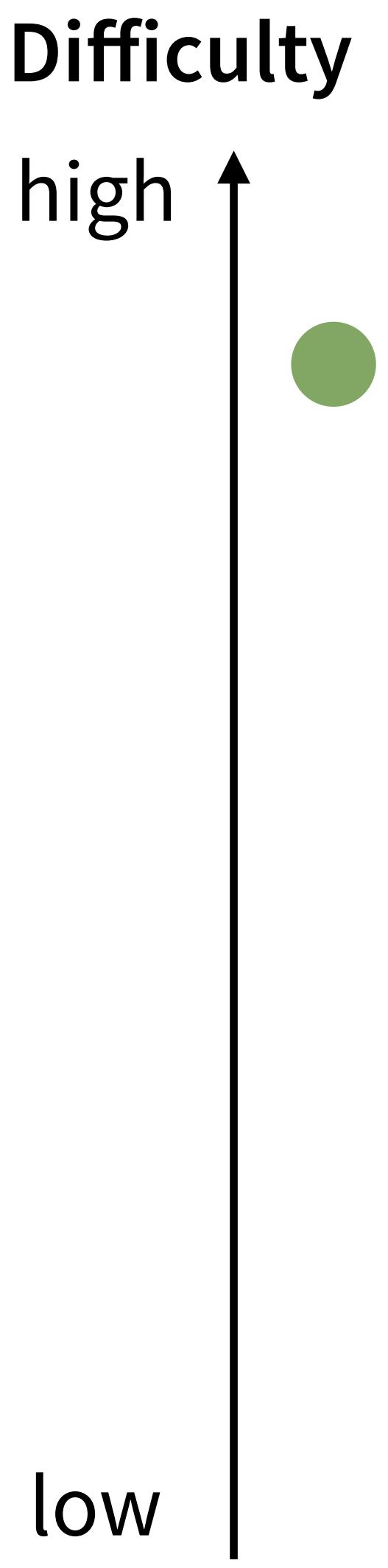
Scaffolding Exercises

Find the most popular names by sex
from 1880 to 2015.



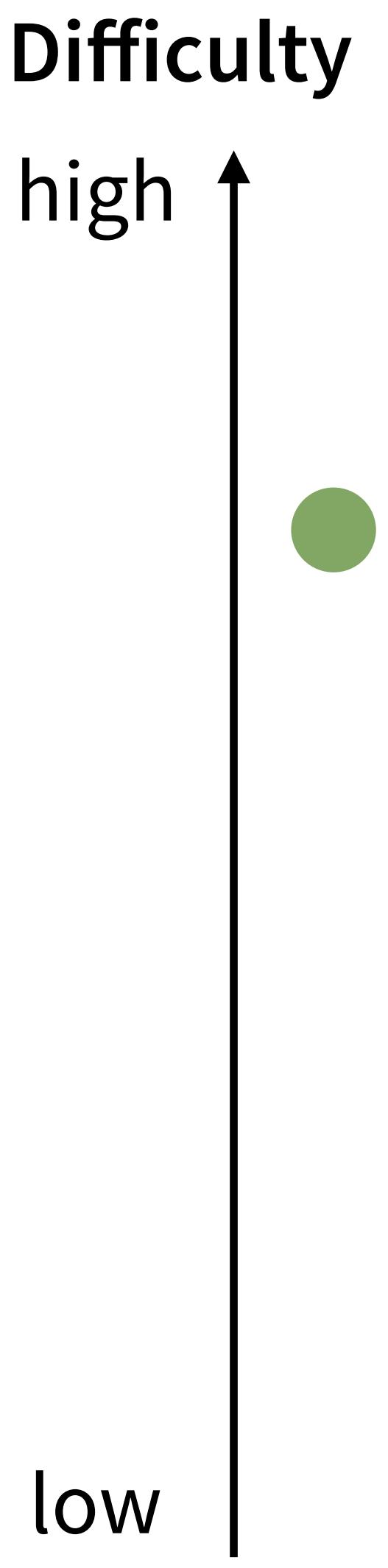
Scaffolding Exercises

Use **dplyr** commands and the **babynames** data set to find the most popular names by sex from 1880 to 2015.



Scaffolding Exercises

Use **group_by()** and **summarise()** to find the most popular names by sex from 1880 to 2015.



Scaffolding Exercises

Fill in the blanks to....

----- %>%

----- %>%

----- %>%

----- %>%

Difficulty

high

low



Scaffolding Exercises

Fill in the blanks to....

```
babynames %>%
```

```
----- %>%
```

```
arrange(desc(total))
```

Difficulty

high ↑



low

Scaffolding Exercises

Fill in the blanks to....

```
babynames %>%  
  group_by(____, ____) %>%  
  summarise(____) %>%  
  arrange(desc(total))
```

Difficulty

high

low



Scaffolding Exercises

What goes wrong? Can you fix it?

```
babynames %>%  
  group_by(name) %>%  
  summarise(total = sum(n)) %>%  
  arrange(desc(total))
```

Difficulty

high ↑

low



Scaffolding Exercises

Predict what the code will do

```
babynames %>%  
  group_by(name, sex) %>%  
  summarise(total = sum(n)) %>%  
  arrange(desc(total))
```

Difficulty

high ↑

low



Scaffolding Exercises

Difficulty

high ↑

Which chunk will find the most popular names?

A

```
babynames %>%  
  group_by(name, sex) %>%  
  summarise(total = sum(n)) %>%  
  arrange(desc(total))
```

B

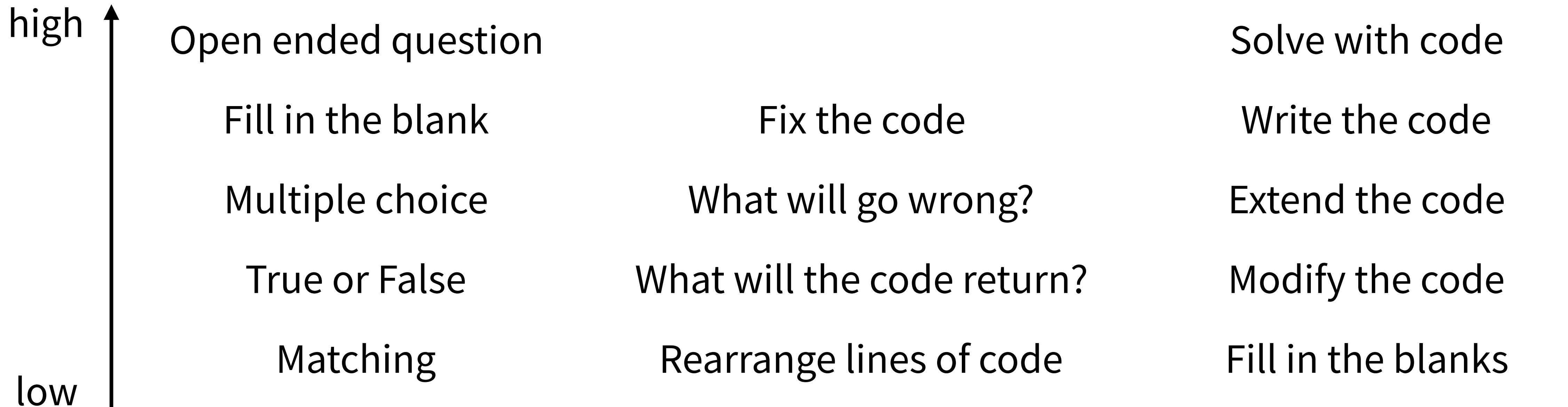
```
babynames %>%  
  filter(year == 2015,  
         sex == "F") %>%  
  select(name, n)
```

low



Scaffolding Ideas

Difficulty



General
Knolwedge

Thinking
about code

Writing
code

Scaffolding Ideas

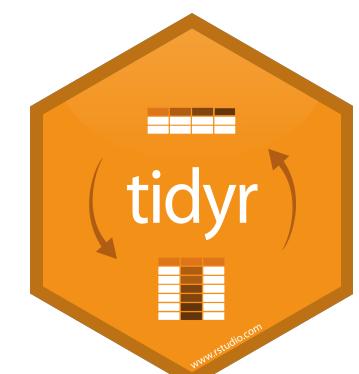
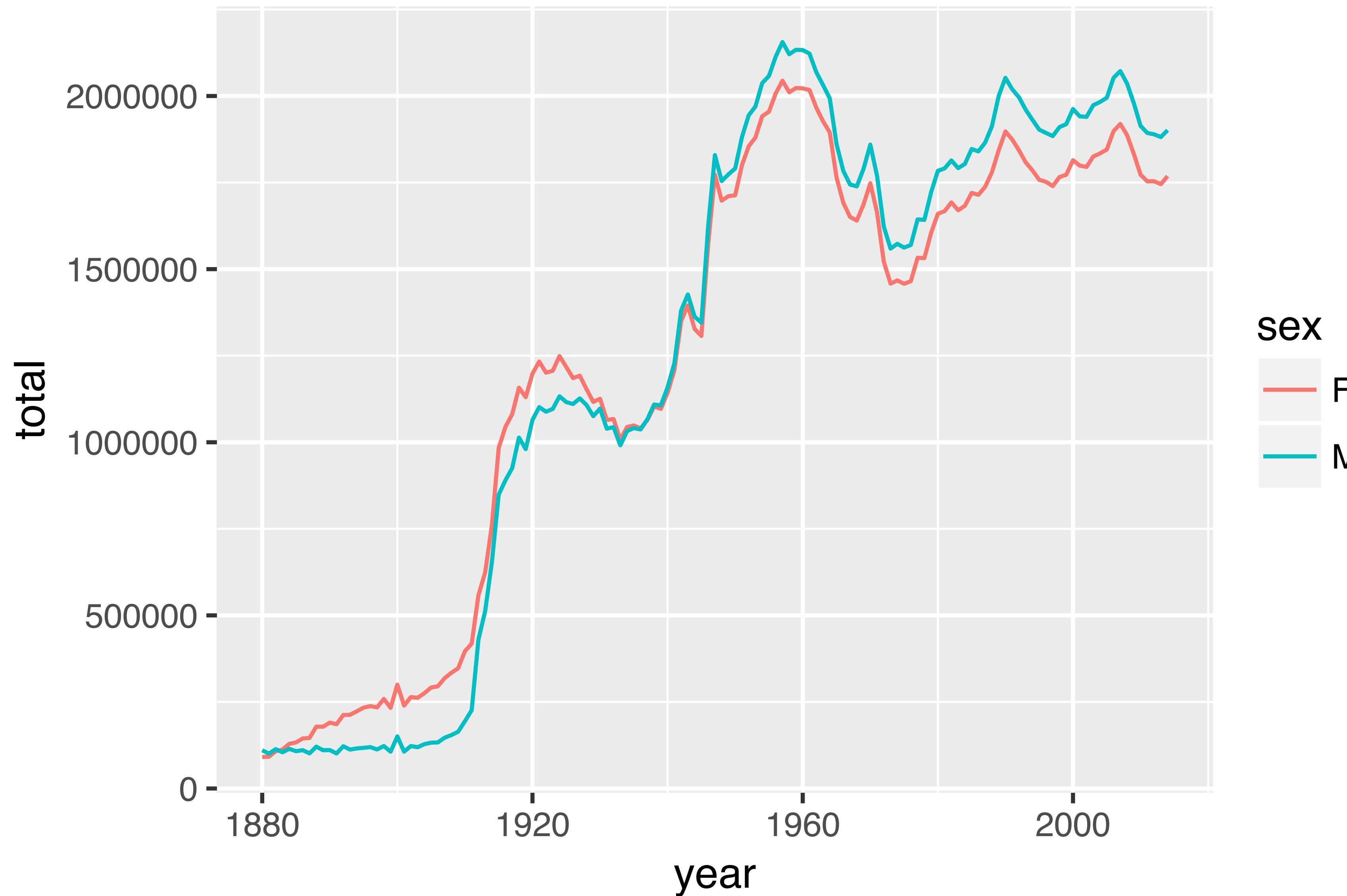
Extend the pipe to arrange the results by total n.

```
babynames %>%  
  group_by(name, sex) %>%  
  summarise(total = sum(n))
```

Your Turn 7

First, the context...

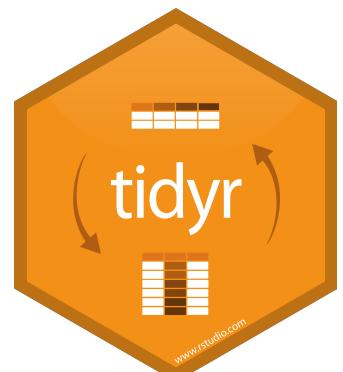
Number of children by year and gender



Can we calculate the ratio of boys to girls?

```
babynames %>%  
  group_by(year, sex) %>%  
  summarise(n = sum(n))
```

	year	sex	n
	<dbl>	<chr>	<int>
1	1880	F	90993
2	1880	M	110491
3	1881	F	91954
4	1881	M	100745
5	1882	F	107850
6	1882	M	113688

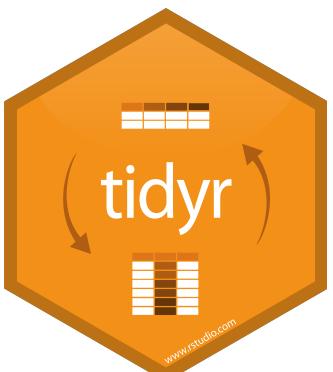


Can we calculate the ratio of boys to girls?

```
babynames %>%  
  group_by(year, sex) %>%  
  summarise(n = sum(n))
```

	year	sex	n
	<dbl>	<chr>	<int>
1	1880	F	90993
2	1880	M	110491
3	1881	F	91954
4	1881	M	100745
5	1882	F	107850
6	1882	M	113688

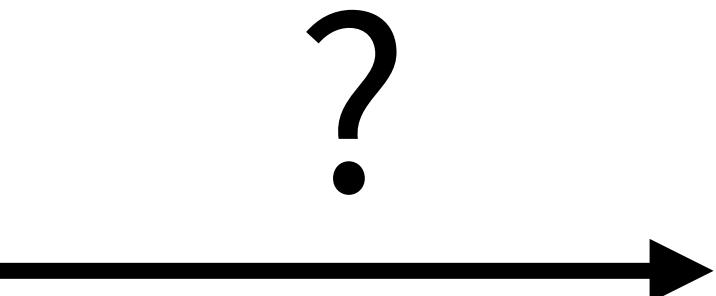
Now
what?



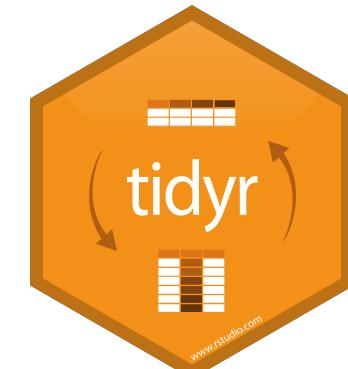
Can we calculate the ratio of boys to girls?

```
better_layout %>%  
  mutate(percent_male = M / (M + F) * 100)
```

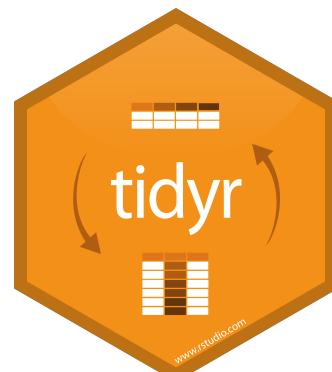
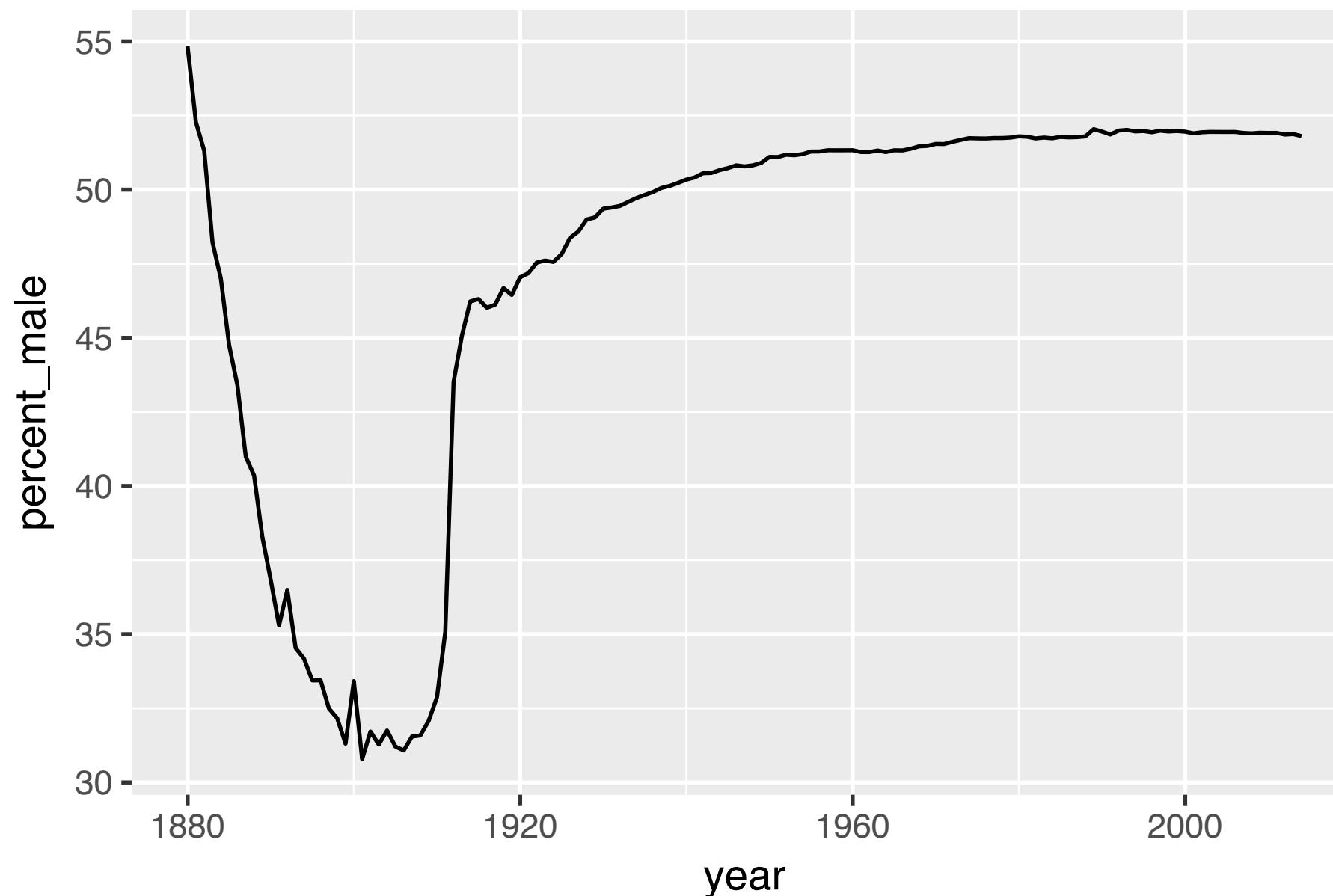
	year	sex	n
	<dbl>	<chr>	<int>
1	1880	F	90993
2	1880	M	110491
3	1881	F	91954
4	1881	M	100745
5	1882	F	107850
6	1882	M	113688



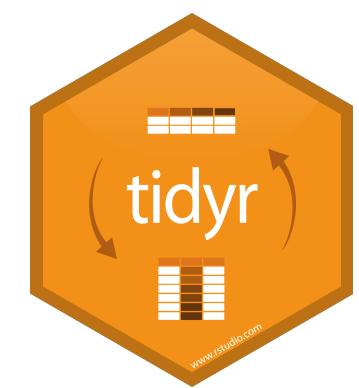
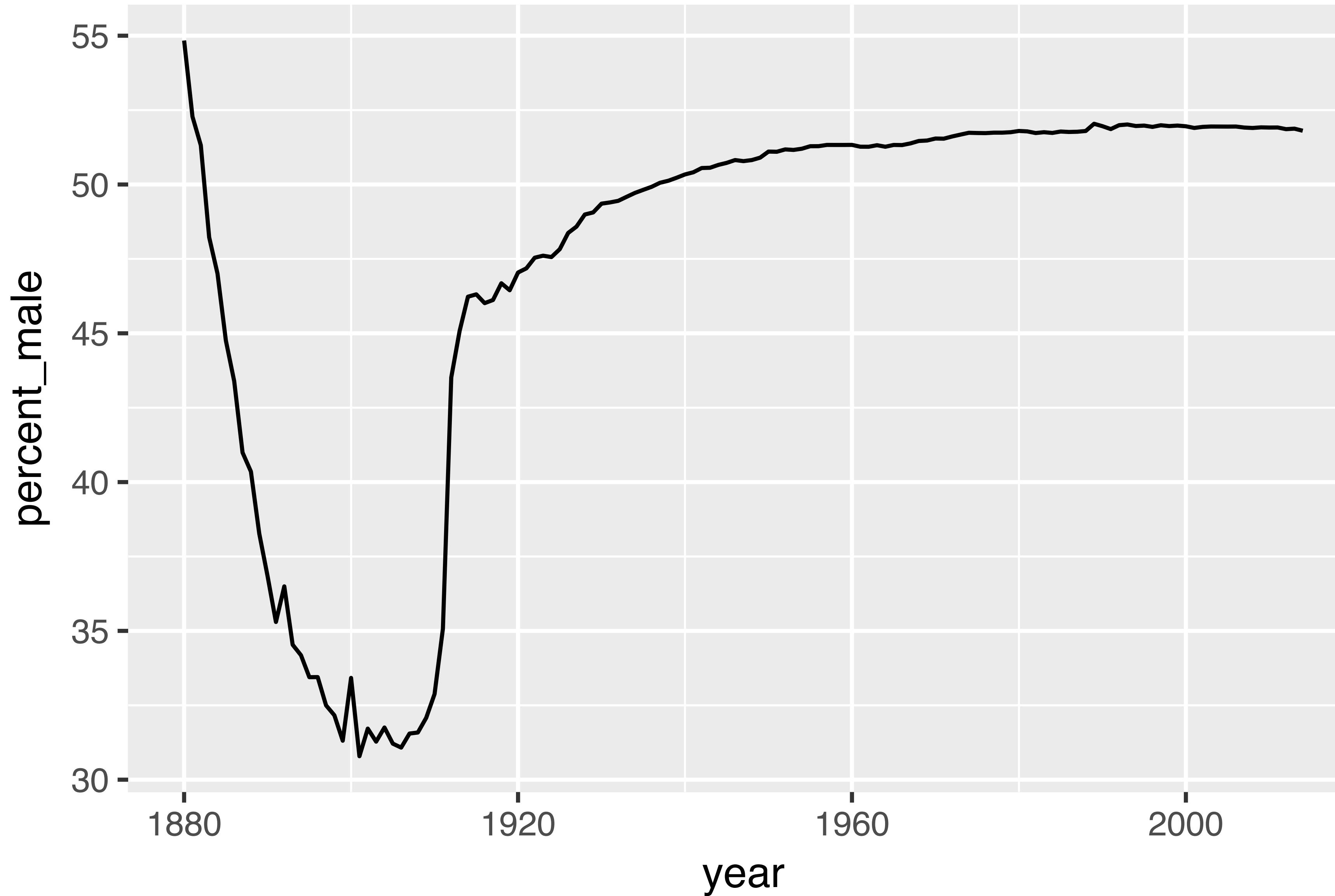
*	year	F	M
*	<dbl>	<int>	<int>
1	1880	90993	110491
2	1881	91954	100745
3	1882	107850	113688
4	1883	112321	104629
5	1884	129022	114445
6	1885	133055	107800



```
babynames %>%  
  group_by(year, sex) %>%  
  summarise(n = sum(n)) %>%  
  spread(sex, n) %>%  
  mutate(percent_male = M / (M + F) * 100) %>%  
  ggplot(aes(year, percent_male)) + geom_line()
```



Percent of children that are male by year



Your Turn

Reshape the layout of this data. Calculate the percent of male (or female) children by year. Then plot the percent over time.

```
babynames %>%  
  group_by(year, sex) %>%  
  summarise(n = sum(n))
```

*Make a harder version.
Make an easier version.*

