

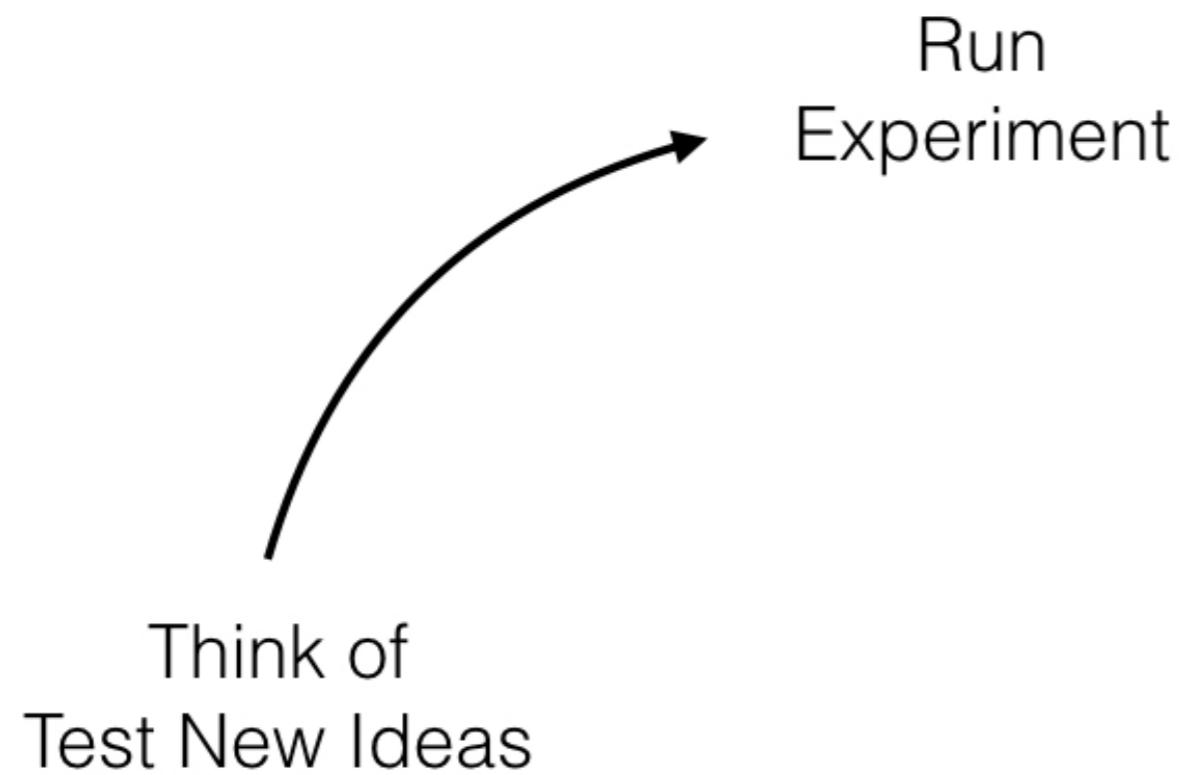


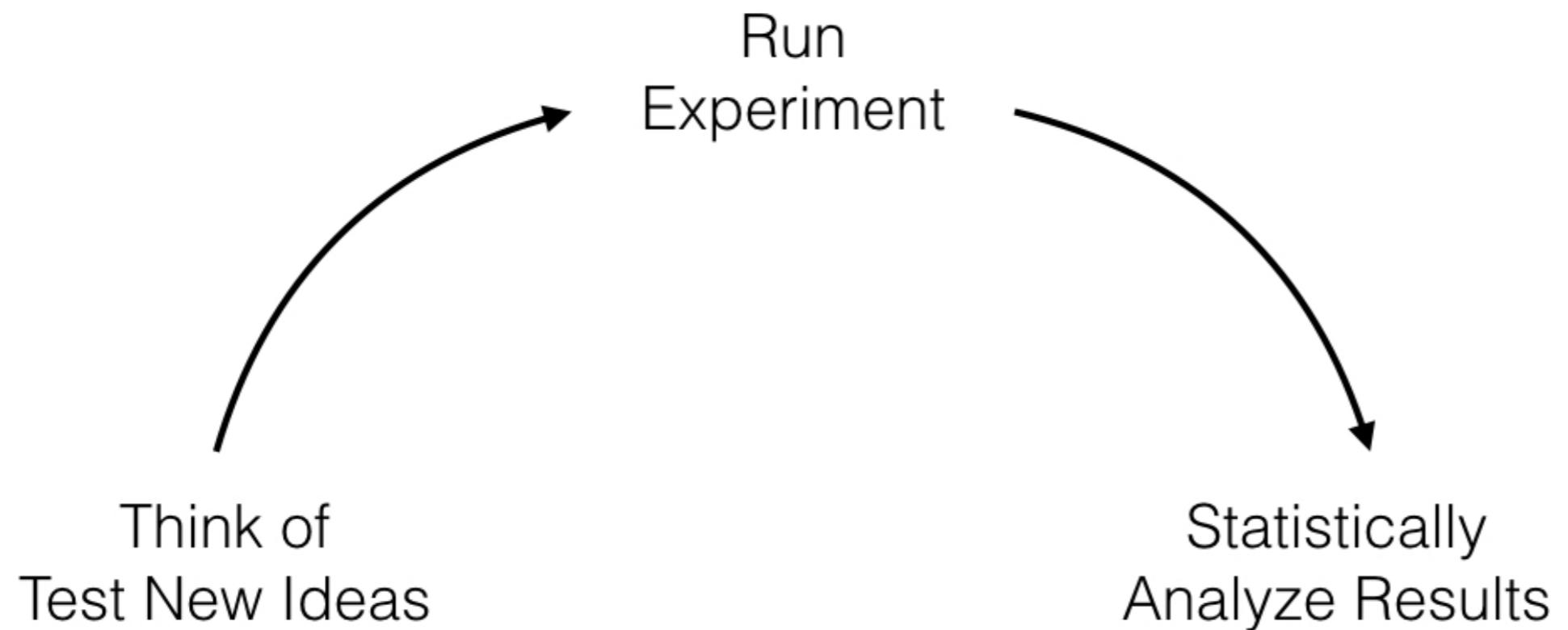
A/B TESTING IN R

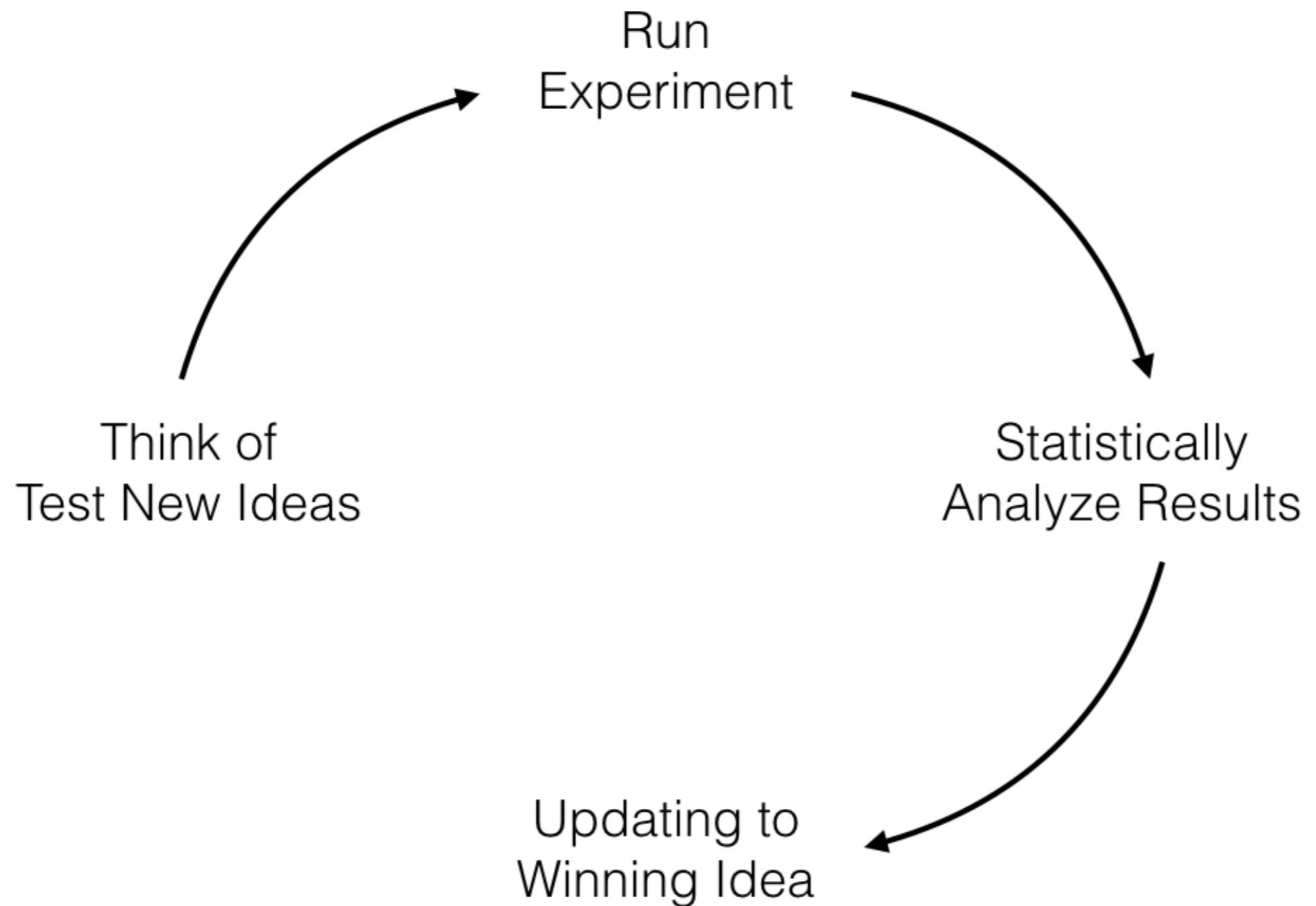
Introduction

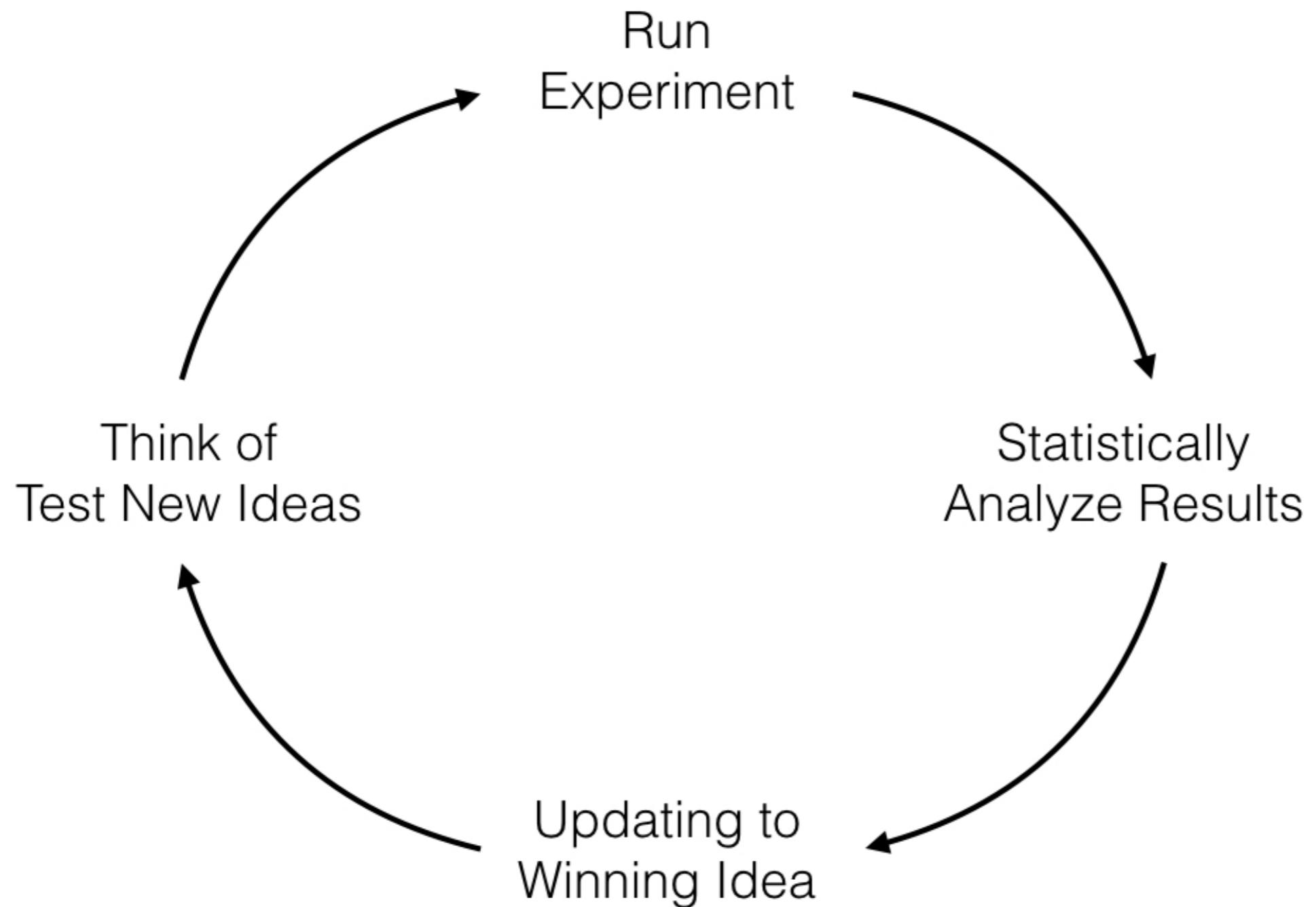
Page Piccinini
Instructor

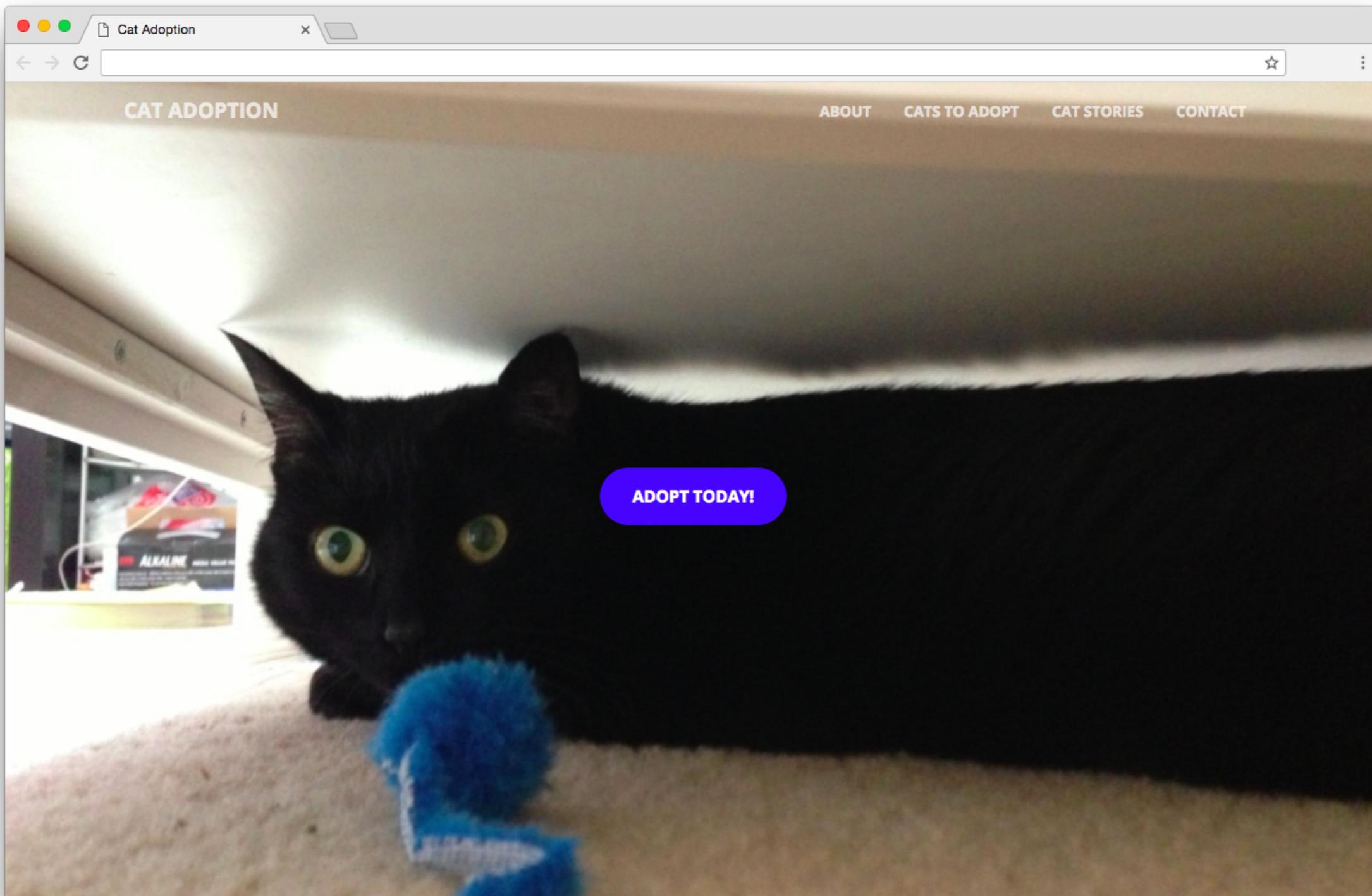
Think of
Test New Ideas

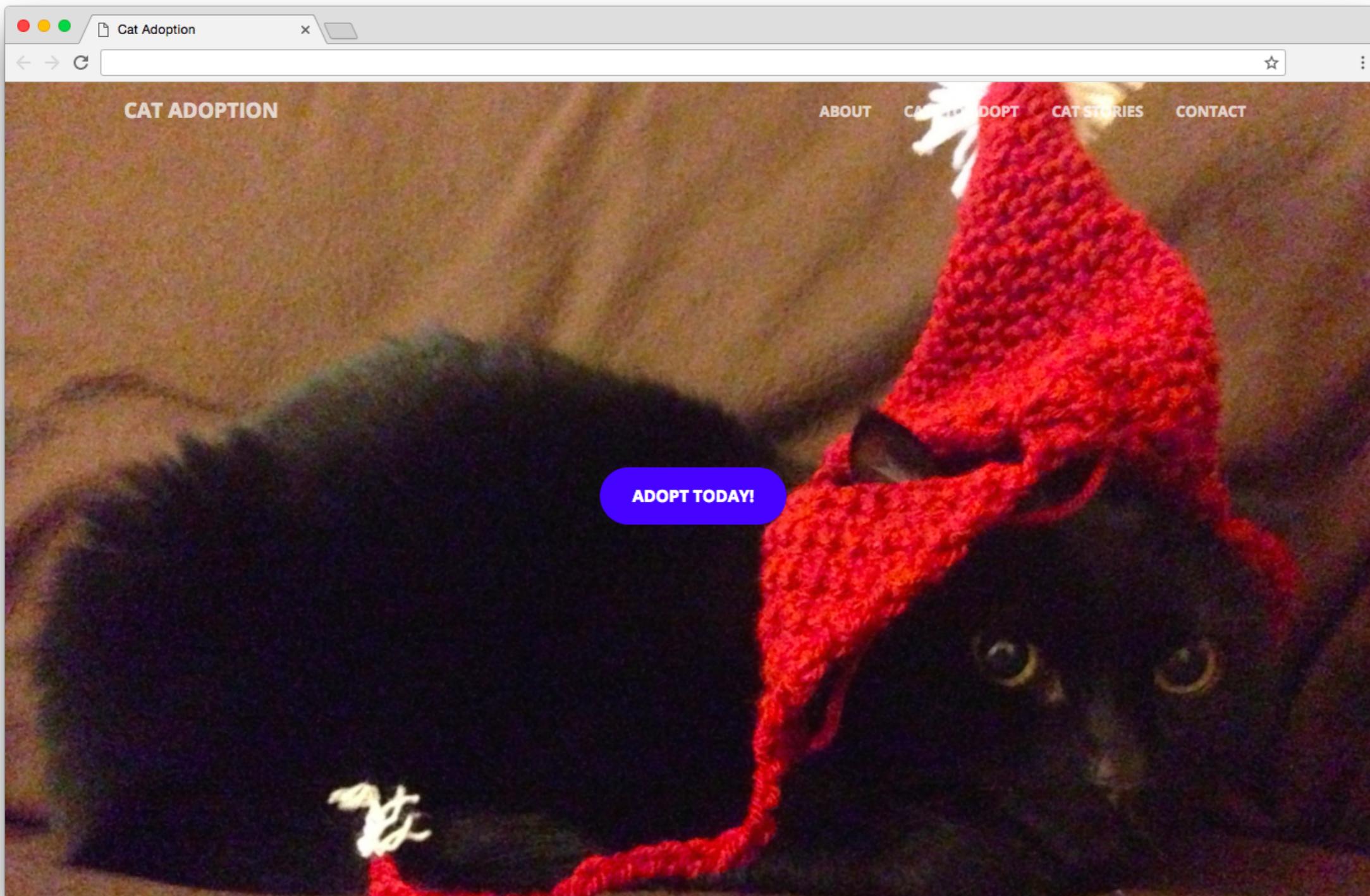












Variables

- **Question:** Will changing the homepage photo result in more "ADOPT TODAY" clicks?
- **Hypothesis:** Using a photo of a cat wearing a hat will result in more "ADOPT TODAY!" clicks.
- **Dependent variable:** Clicked "ADOPT TODAY!" button or not.
- **Independent variable:** Homepage photo.

Preliminary dataset

```
library(tidyverse)  
  
click_data <- read_csv("click_data.csv")  
click_data
```

```
# A tibble: 3,650 x 2  
  visit_date clicked_adopt_today  
  <date>           <int>  
1 2017-01-01             1  
2 2017-01-02             1  
3 2017-01-03             0  
4 2017-01-04             1  
5 2017-01-05             1  
6 2017-01-06             0  
7 2017-01-07             0  
8 2017-01-08             0  
9 2017-01-09             0  
10 2017-01-10            0  
# ... with 3,640 more rows
```



A/B TESTING IN R

Let's practice!



A/B TESTING IN R

Baseline Conversion Rates

Page Piccinini
Instructor

Hypothesis: Using a photo of a cat wearing a hat will result in more "ADOPT TODAY!" clicks.

Hypothesis: Using a photo of a cat wearing a hat will result in **more** "ADOPT TODAY!" clicks.

Hypothesis: Using a photo of a cat wearing a hat will result in **more** "ADOPT TODAY!" clicks.

conversion rate

last year?

Hypothesis: Using a photo of a cat wearing a hat will result in **more** "ADOPT TODAY!" clicks.

conversion rate

last year?

conversion rate

today?

Hypothesis: Using a photo of a cat wearing a hat will result in **more** "ADOPT TODAY!" clicks.

*conversion rate
last year?*

*conversion rate
today?*

*conversion rate
next week?*

Hypothesis: Using a photo of a cat wearing a hat will result in **more** "ADOPT TODAY!" clicks.

*conversion rate
last year?*

*conversion rate
today?*

*conversion rate
next week?*

*conversion rate
when run experiment?*

Current conversion rate

```
library(tidyverse)

click_data <- read_csv("click_data.csv")
click_data

click_data %>%
  summarize(conversion_rate = mean(clicked_adopt_today))
```

```
# A tibble: 1 x 1
  conversion_rate
  <dbl>
1 0.2772603
```

Current conversion rate seasonality

```
library(tidyverse)

click_data <- read_csv("click_data.csv")
click_data

click_data %>%
    summarize(conversion_rate = mean(clicked_adopt_today))
```

Current conversion rate seasonality

```
library(tidyverse)
library(lubridate)

click_data <- read_csv("click_data.csv")
click_data

click_data %>%
  group_by(month(visit_date)) %>%
  summarize(conversion_rate = mean(clicked_adopt_today))
```

```
# A tibble: 12 x 2
`month(visit_date)` conversion_rate
                <dbl>            <dbl>
1                  1      0.1967742
2                  2      0.1892857
3                  3      0.1451613
4                  4      0.1500000
5                  5      0.2580645
6                  6      0.3333333
7                  7      0.3483871
8                  8      0.5419355
9                  9      0.2933333
10                 10     0.1612903
11                 11     0.2333333
12                 12     0.4645161
```

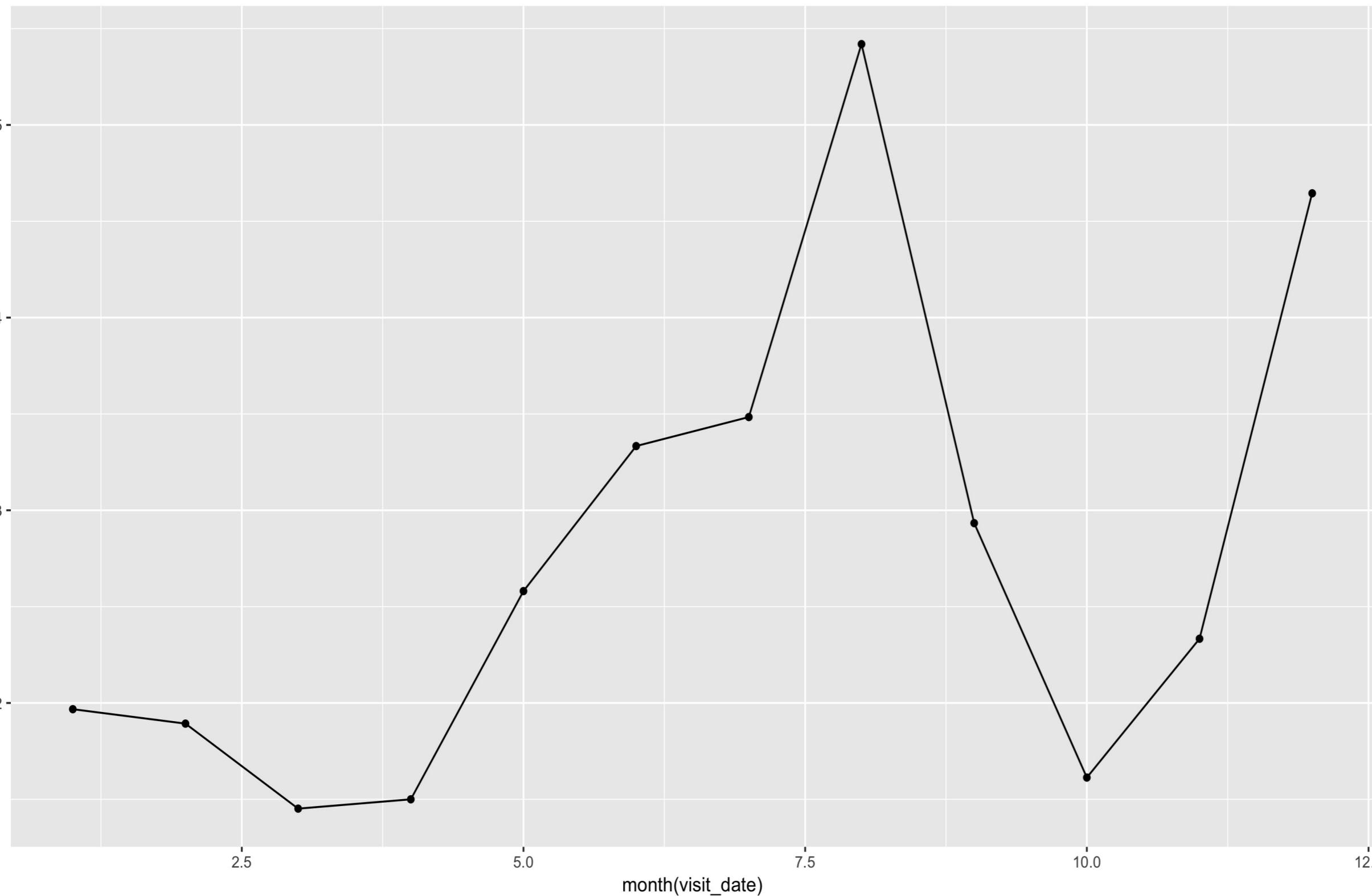
Plotting current conversion rate seasonality

```
library(tidyverse)
library(lubridate)

click_data <- read_csv("click_data.csv")
click_data

click_data_sum <- click_data %>%
  group_by(month(visit_date)) %>%
  summarize(conversion_rate = mean(clicked_adopt_today))

ggplot(click_data_sum, aes(x = `month(visit_date)`, y = conversion_rate)) +
  geom_point() +
  geom_line()
```





A/B TESTING IN R

Let's practice!

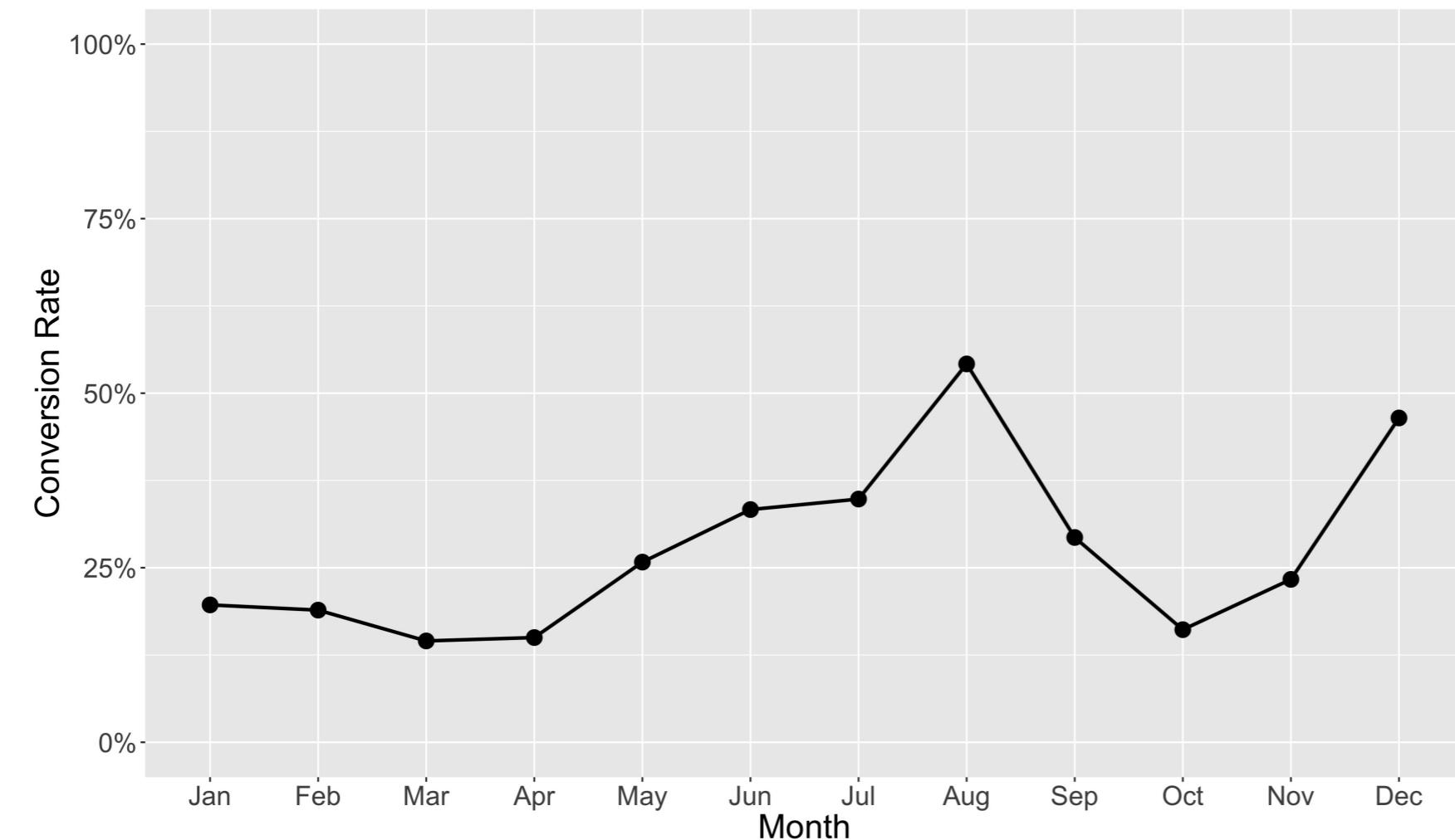


A/B TESTING IN R

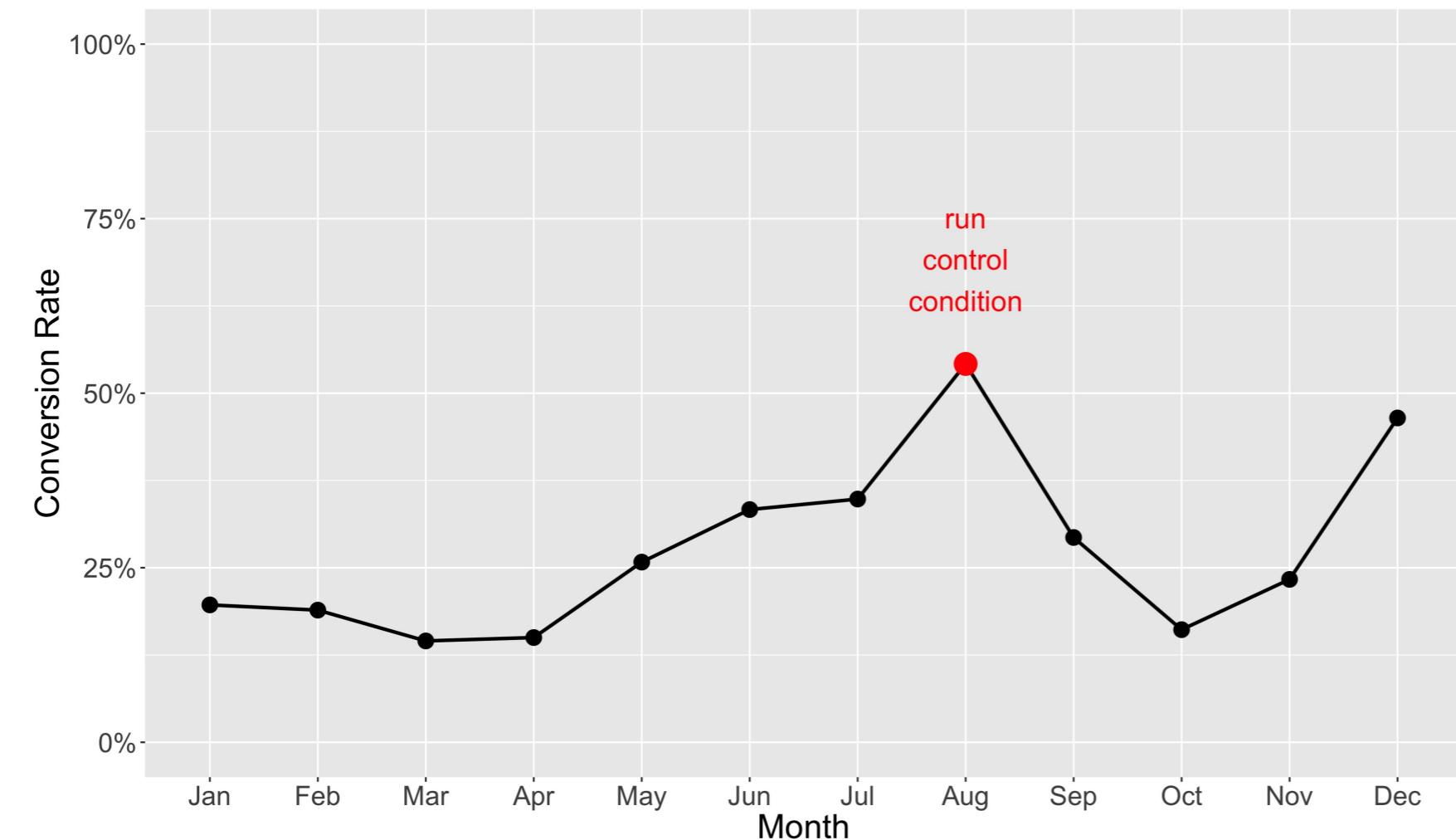
Designing an Experiment - Power Analysis

Page Piccinini
Instructor

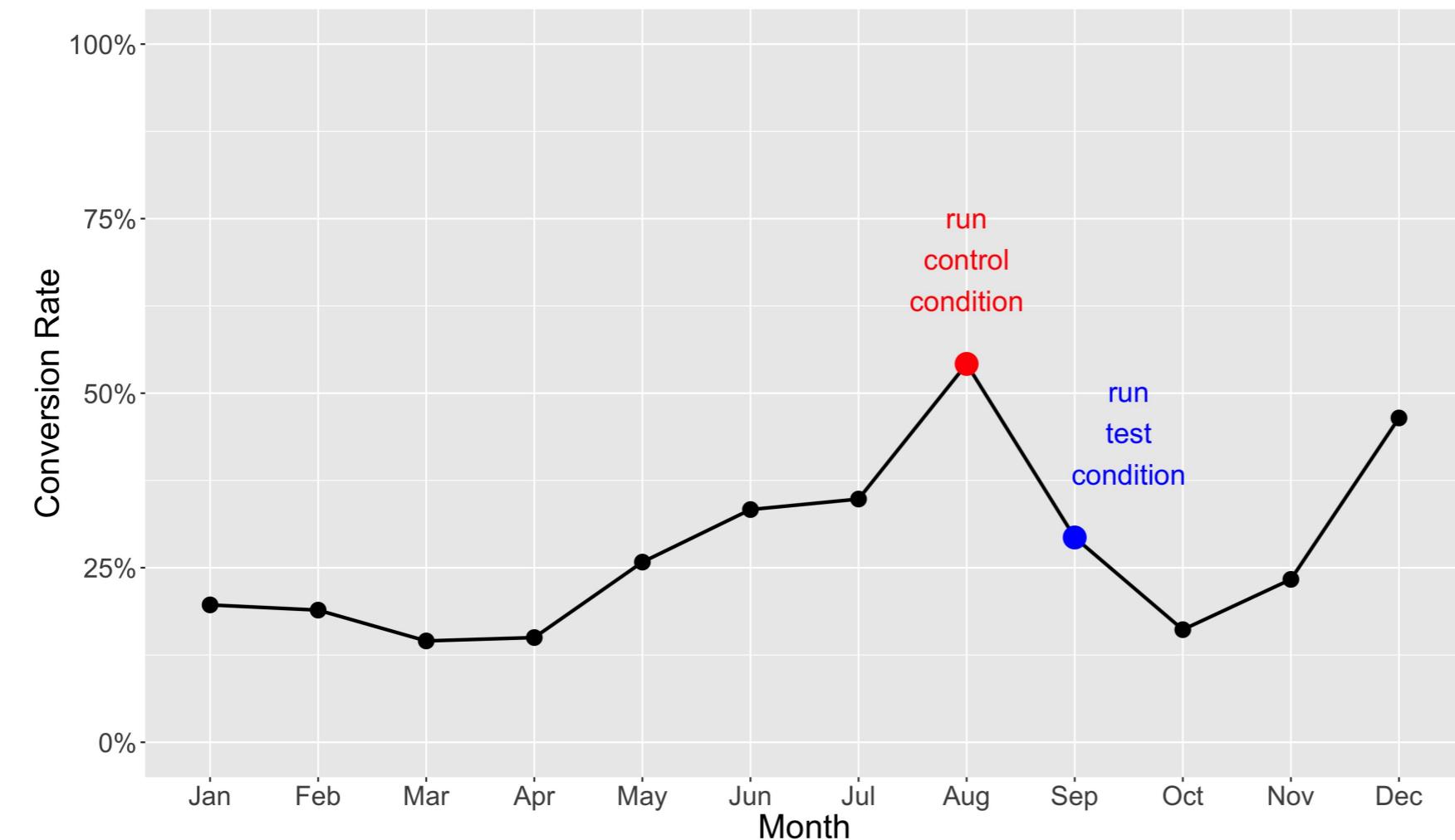
Conversion rate seasonality



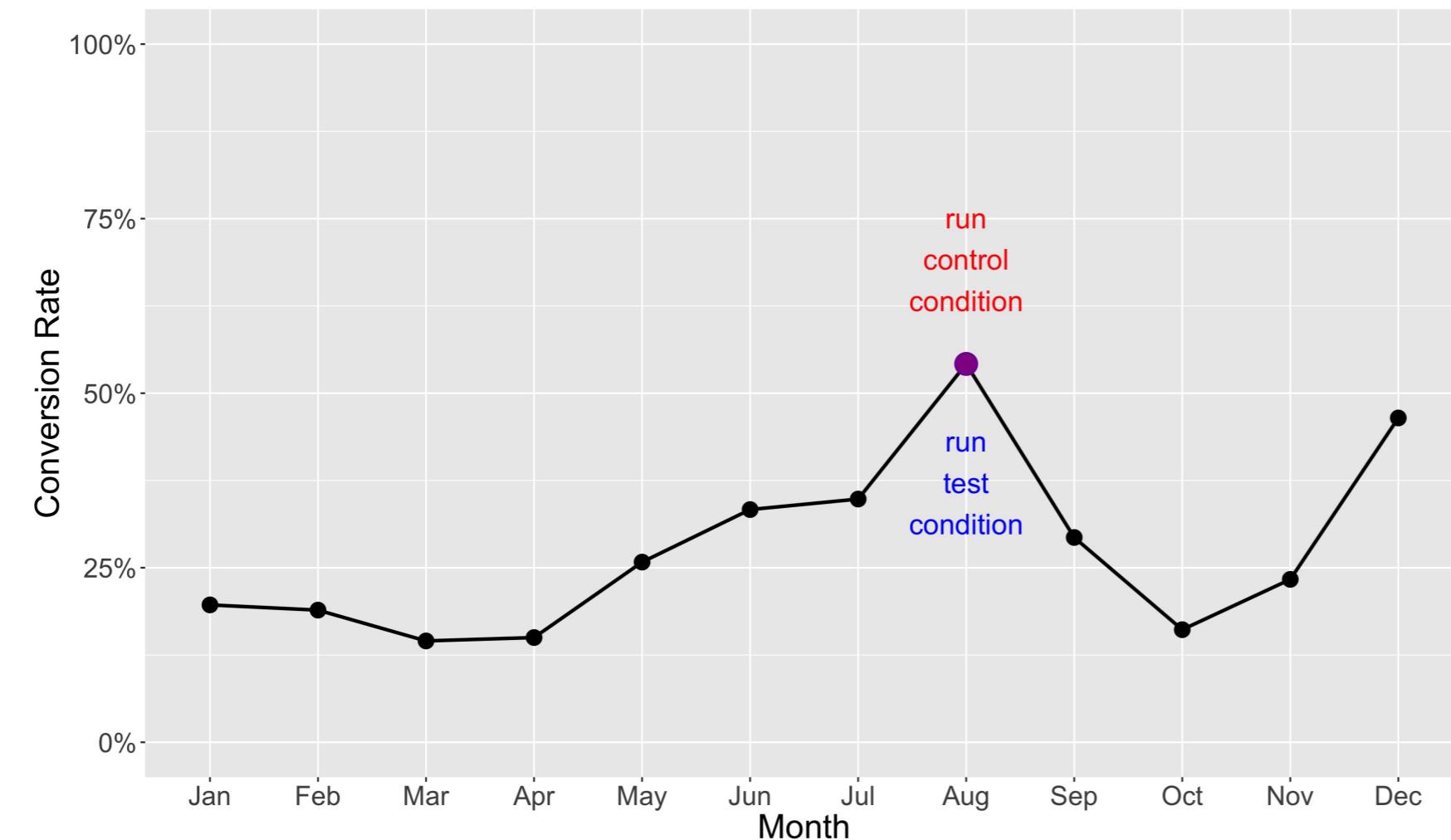
Conversion rate seasonality



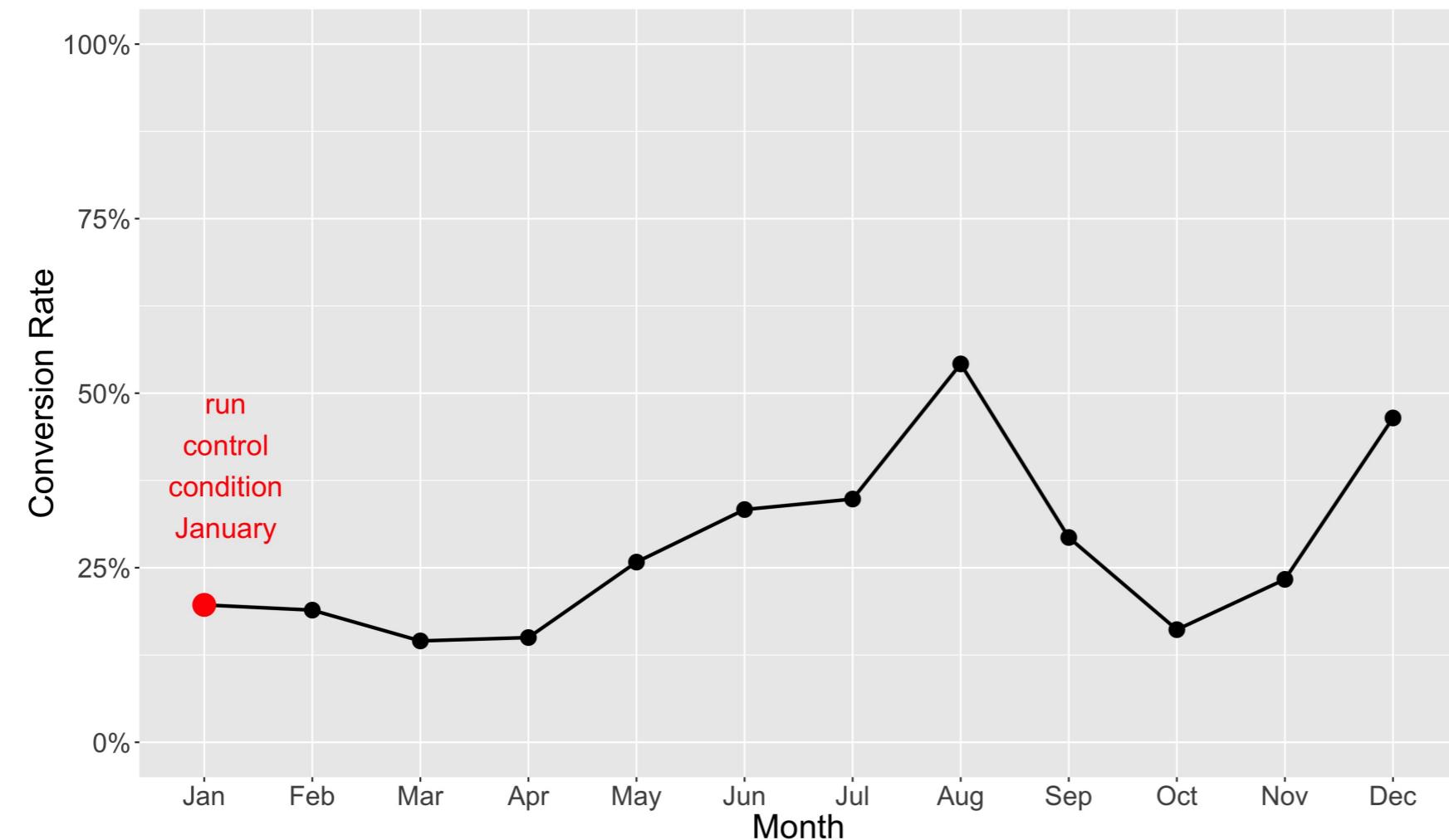
Conversion rate seasonality



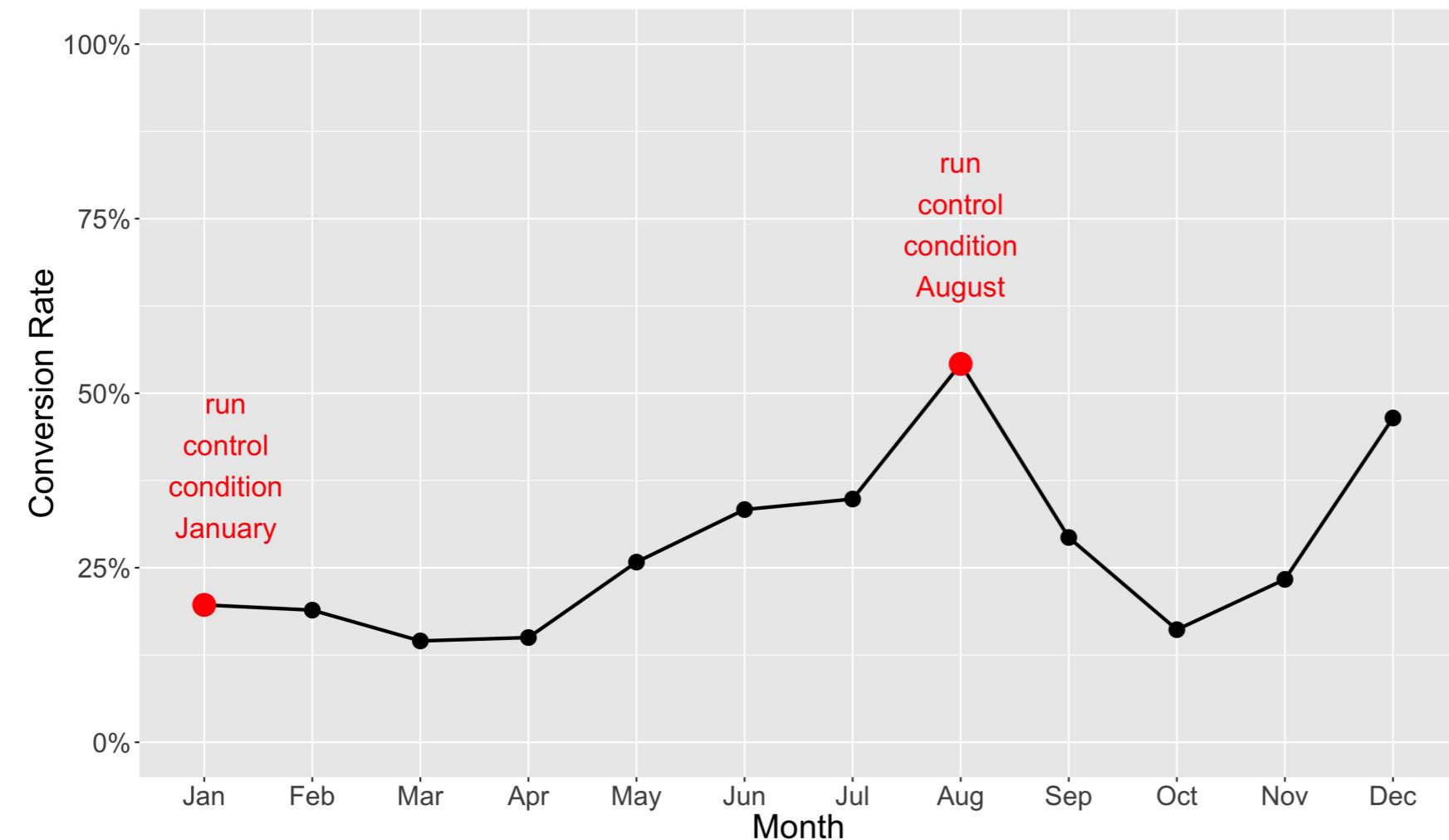
Conversion rate seasonality



Conversion rate seasonality



Conversion rate seasonality



Power analysis

- **statistical test** - statistical test you plan to run
- **baseline value** - value for the current control condition
- **desired value** - expected value for the test condition
- **proportion of the data** from the test condition (ideally 0.5)
- **significance threshold / alpha** - level where effect significant (generally 0.05)
- **power / 1 - beta** - probability correctly rejecting null hypothesis (generally 0.8)

Power analysis in R

```
library(powerMediation)  
total_sample_size <- SSizeLogisticBin(  
    )
```

Power analysis in R

```
library(powerMediation)

total_sample_size <- SSizeLogisticBin(
    B = 0.5,
    alpha = 0.05,
    power = 0.8)
```

Power analysis in R

```
library(powerMediation)

total_sample_size <- SSizeLogisticBin(p1 = 0.2,
                                         p2 = 0.3,
                                         B = 0.5,
                                         alpha = 0.05,
                                         power = 0.8)
```

```
> total_sample_size
[1] 587
> total_sample_size / 2
[1] 293.5
```



A/B TESTING IN R

Let's practice!