



A/B TESTING IN R

Analyzing Results

Page Piccinini
Instructor



Experiment results

```
library(tidyverse)
```

```
experiment_data <- read_csv("experiment_data.csv")  
experiment_data
```

```
# A tibble: 588 x 3  
  visit_date condition clicked_adopt_today  
    <date>      <chr>          <int>  
1 2018-01-01 control            0  
2 2018-01-01 control            1  
3 2018-01-01 control            0  
4 2018-01-01 control            0  
5 2018-01-01 test              0  
6 2018-01-01 test              0  
7 2018-01-01 test              1  
8 2018-01-01 test              0  
9 2018-01-01 test              0  
10 2018-01-01 test              1  
# ... with 578 more rows
```



Experiment results

```
library(tidyverse)

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

experiment_data %>%
  group_by(condition) %>%
  summarize(conversion_rate = mean(clicked_adopt_today))
```

```
# A tibble: 2 x 2
  condition conversion_rate
  <chr>         <dbl>
1 control      0.1666667
2 test        0.3843537
```



Plotting results

```
library(tidyverse)

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

experiment_data %>%
  group_by(condition) %>%
  summarize(conversion_rate = mean(clicked_adopt_today))
```



Plotting results

```
library(tidyverse)

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

experiment_data %>%
  group_by(visit_date, condition) %>%
  summarize(conversion_rate = mean(clicked_adopt_today))
```



Plotting results

```
library(tidyverse)

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

experiment_data_sum <- experiment_data %>%
  group_by(visit_date, condition) %>%
  summarize(conversion_rate = mean(clicked_adopt_today))

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



Plotting results

```
library(tidyverse)

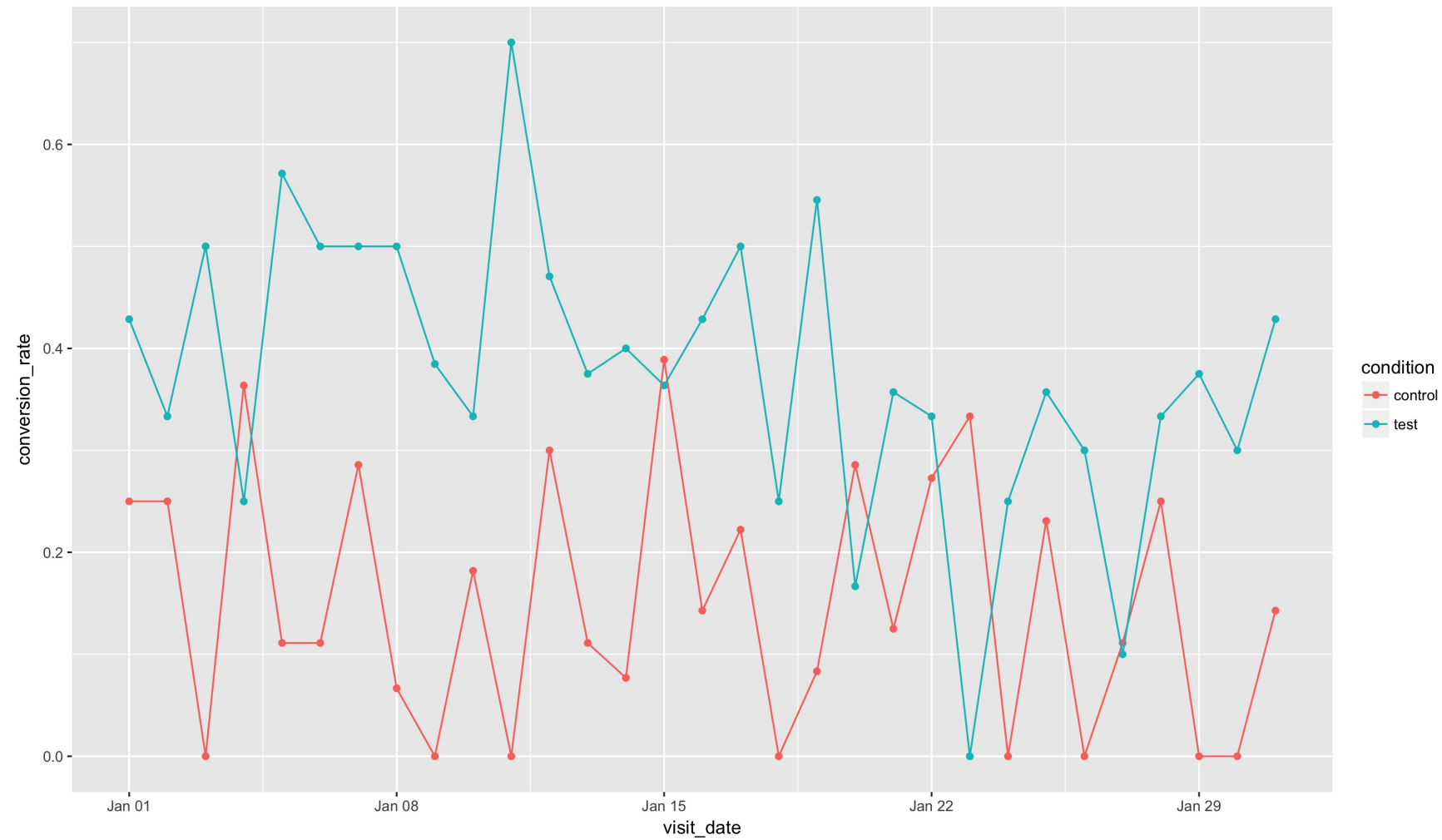
experiment_data <- read_csv("experiment_data.csv")
experiment_data

experiment_data_sum <- experiment_data %>%
  group_by(visit_date, condition) %>%
  summarize(conversion_rate = mean(clicked_adopt_today))

ggplot(experiment_data_sum,
       aes(x = visit_date,
           y = conversion_rate,
           color = condition,
           group = condition)) +
  geom_point() +
  geom_line()
```



Plotting results





Analyzing results

```
library(tidyverse)
library(broom)

experiment_data <- read_csv("experiment_data.csv")

glm(
    ~
)
```



Analyzing results

```
library(tidyverse)
library(broom)

experiment_data <- read_csv("experiment_data.csv")

glm(clicked_adopt_today ~
    )
```



Analyzing results

```
library(tidyverse)
library(broom)

experiment_data <- read_csv("experiment_data.csv")

glm(clicked_adopt_today ~ condition,
      )
```



Analyzing results

```
library(tidyverse)
library(broom)

experiment_data <- read_csv("experiment_data.csv")

glm(clicked_adopt_today ~ condition,
     family = "binomial",
     )
```



Analyzing results

```
library(tidyverse)
library(broom)

experiment_data <- read_csv("experiment_data.csv")

glm(clicked_adopt_today ~ condition,
     family = "binomial",
     data = experiment_data)
```



Analyzing results

```
library(tidyverse)
library(broom)

experiment_data <- read_csv("experiment_data.csv")

glm(clicked_adopt_today ~ condition,
     family = "binomial",
     data = experiment_data) %>%
  tidy()
```

	term	estimate	std.error	statistic	p.value
1	(Intercept)	-1.609438	0.1564922	-10.284464	8.280185e-25
2	conditiontest	1.138329	0.1971401	5.774212	7.731397e-09



A/B TESTING IN R

Let's practice!



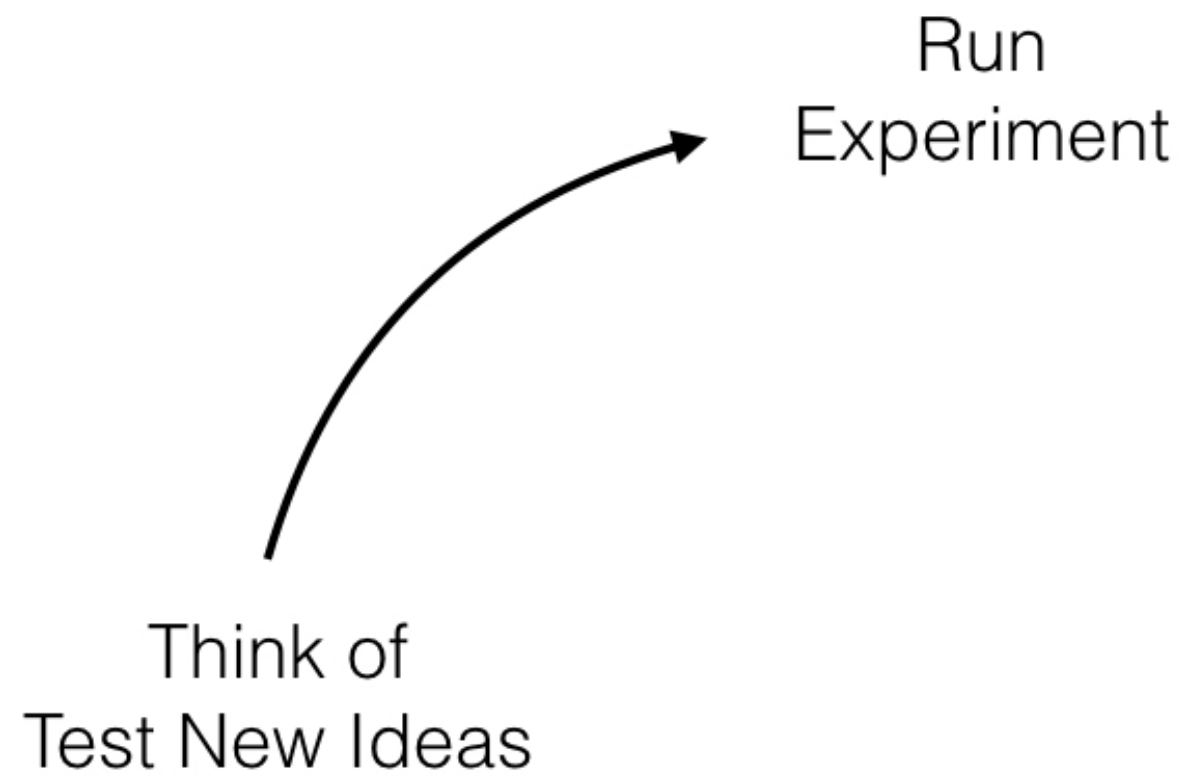
A/B TESTING IN R

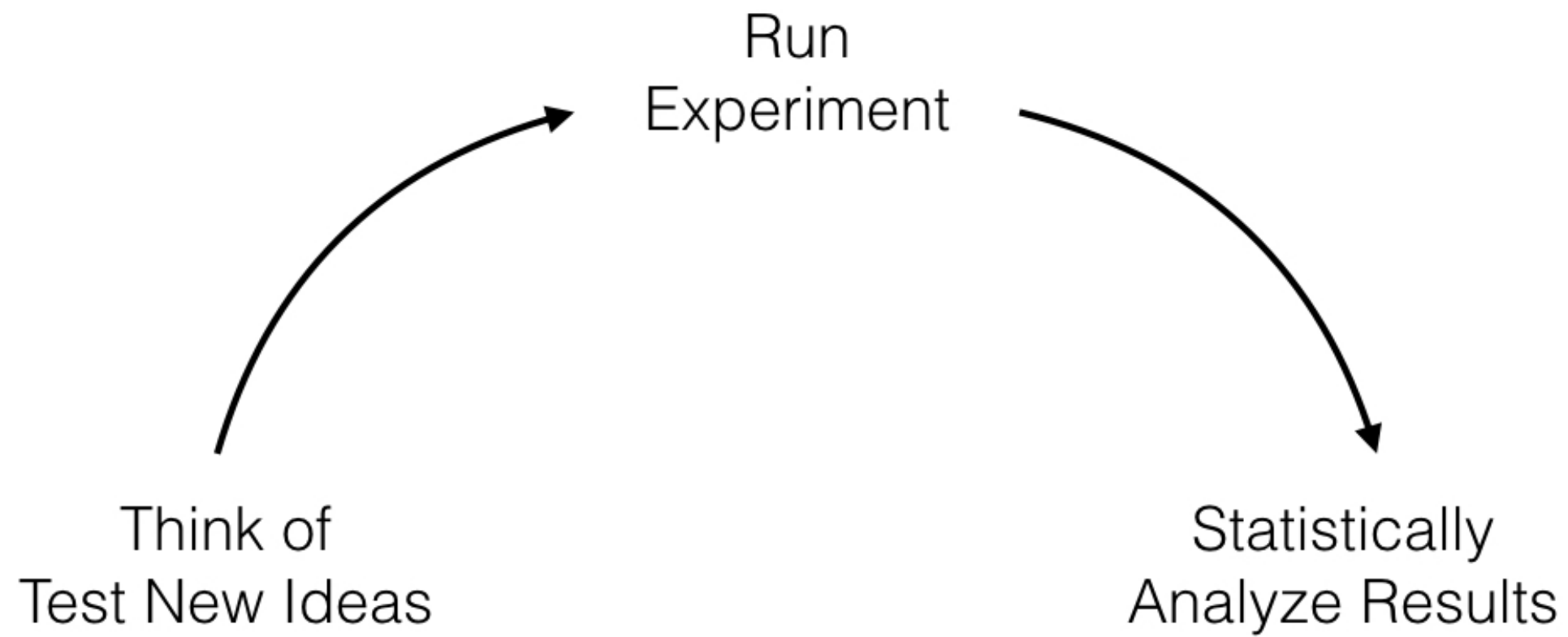
Designing Follow-up Experiments

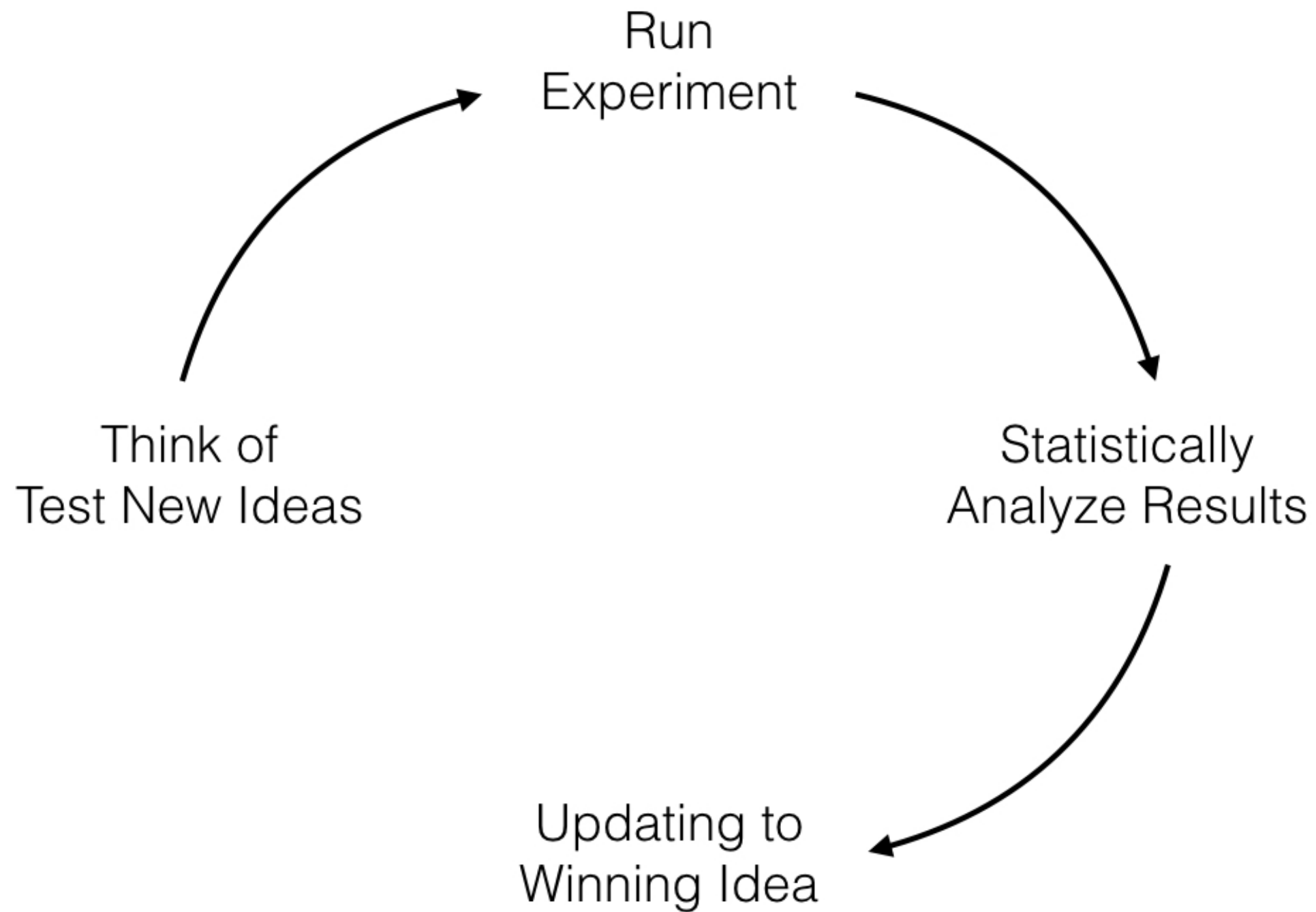
Page Piccinini
Instructor

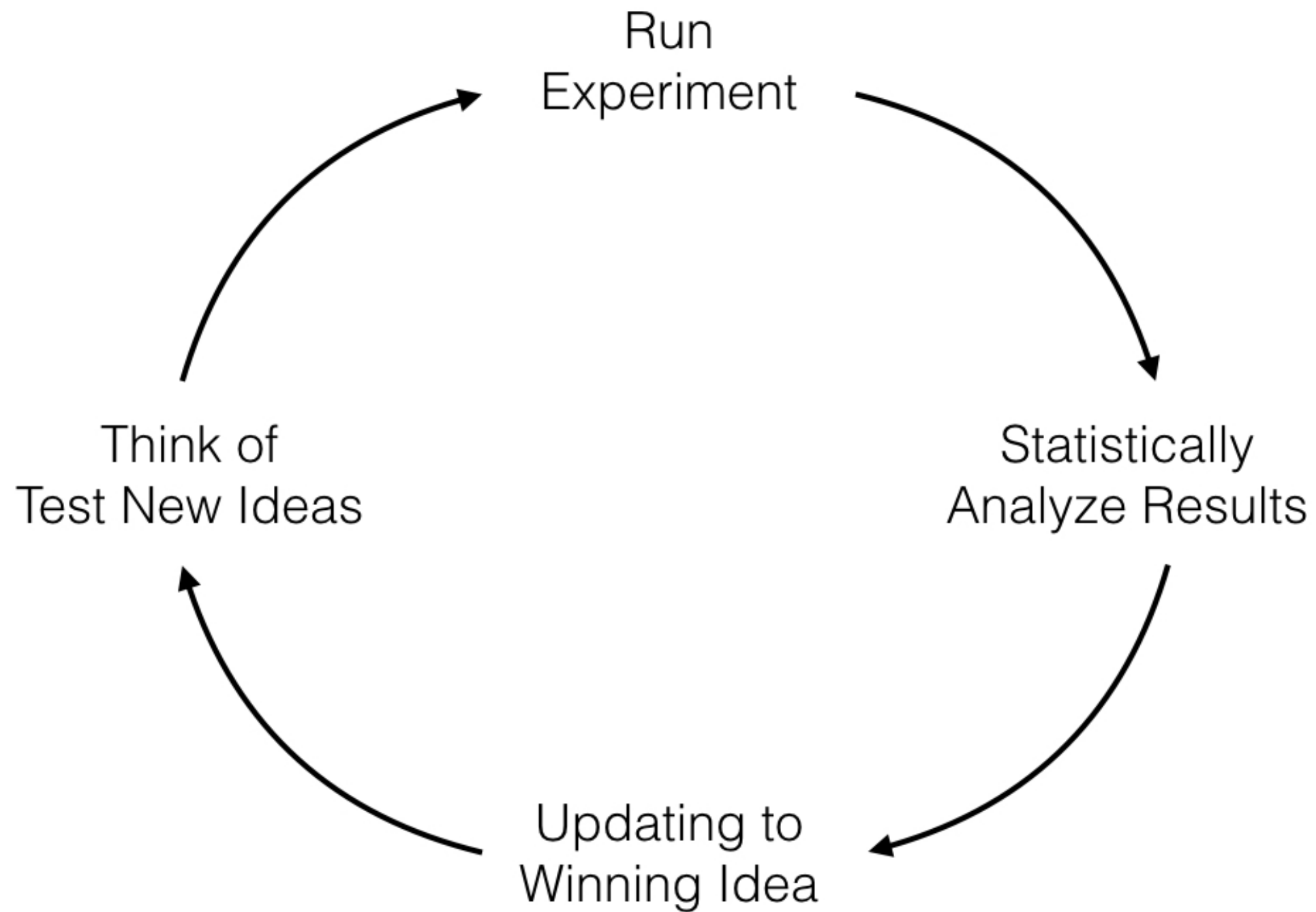


Think of
Test New Ideas











Tips for designing a new experiment

- Build several small follow-up experiments
- Avoid **confounding variables**
- Test small changes



Follow-up experiment #1

1. Use a picture of a **kitten** in a hat instead of an adult cat.
2. Use a picture of **two cats or kittens** in hats instead of one.



A/B TESTING IN R

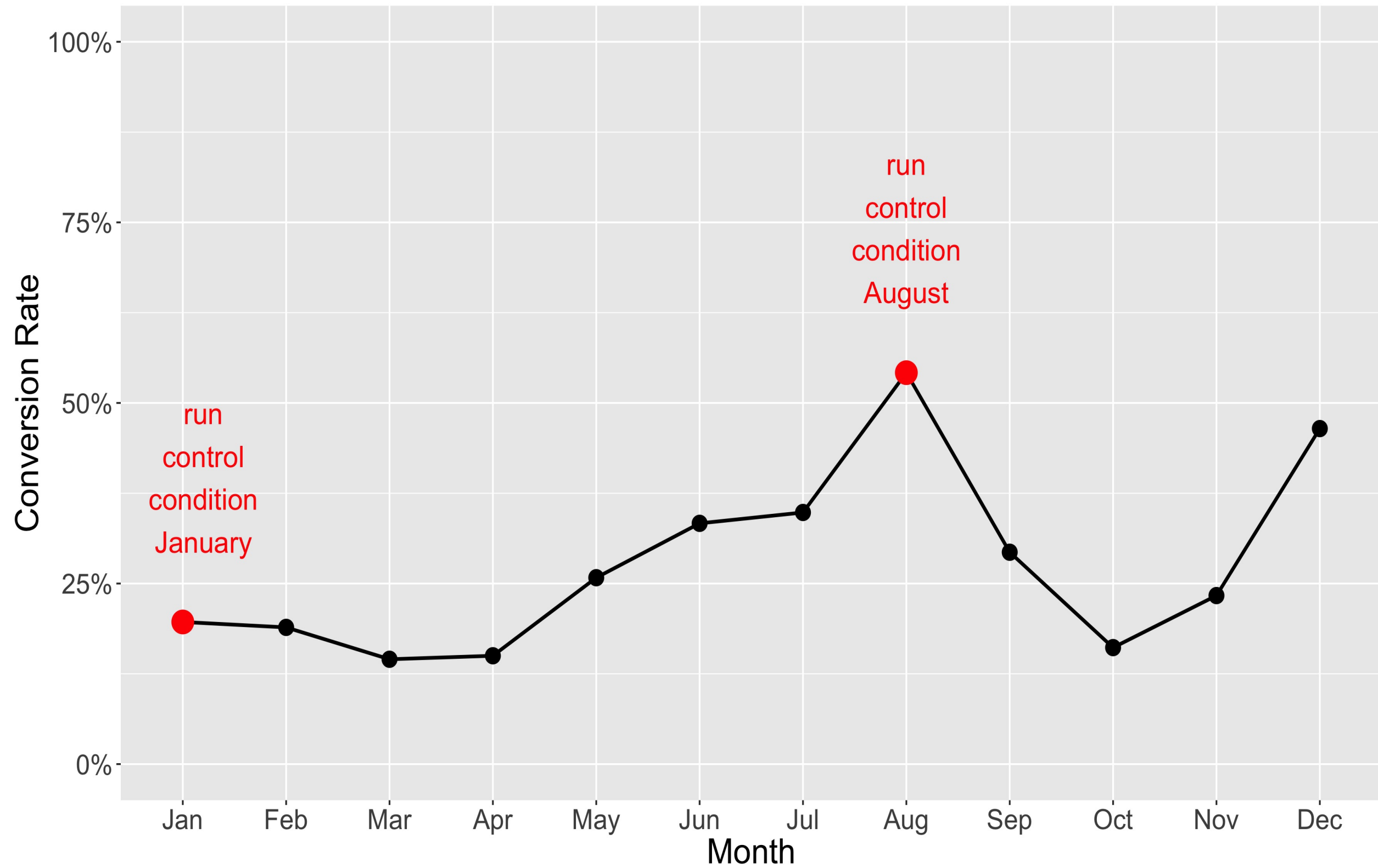
Let's practice!



A/B TESTING IN R

Pre-follow-up Experiment Assumptions

Page Piccinini
Instructor





A/B TESTING IN R

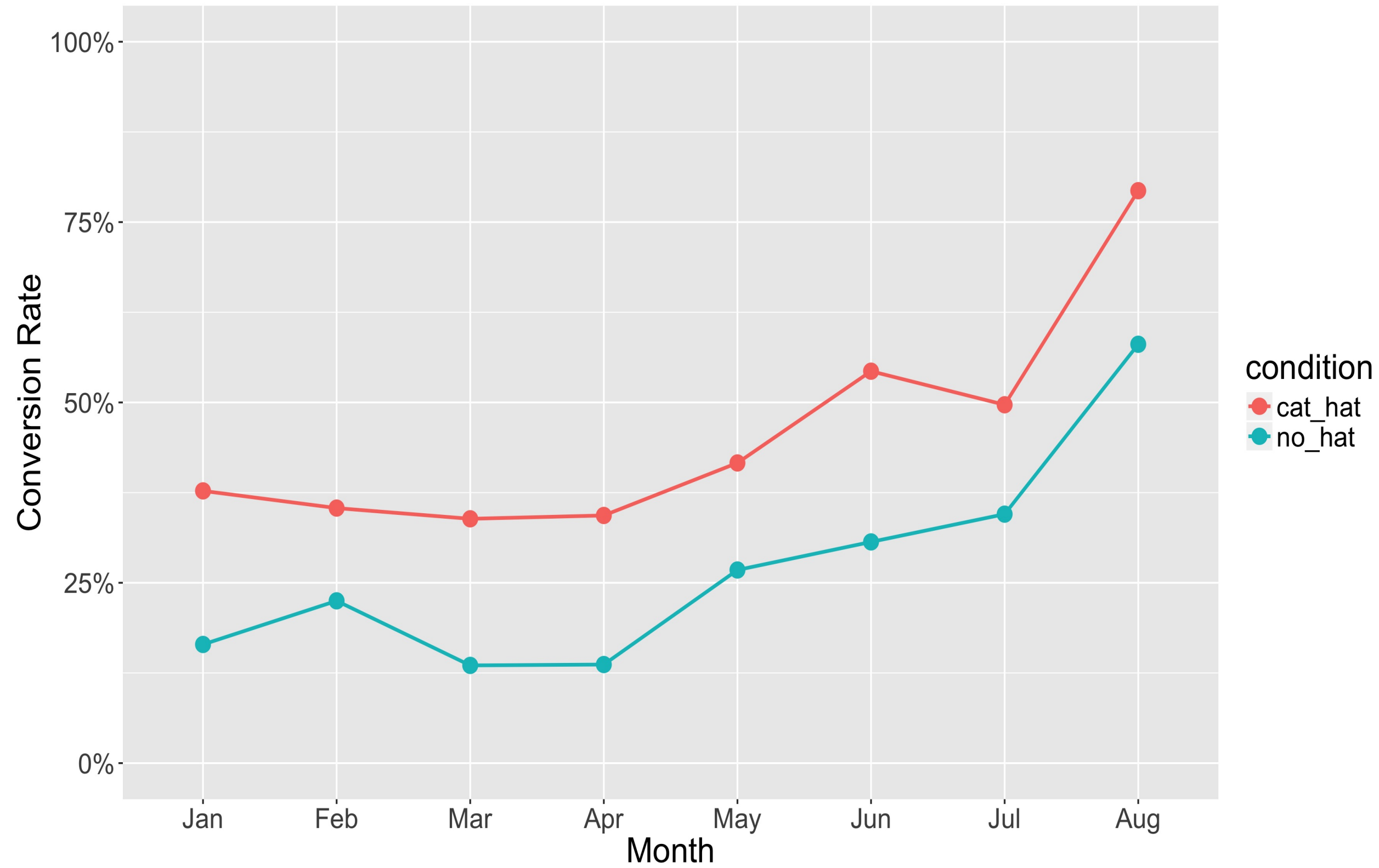
Let's practice!



A/B TESTING IN R

Follow-up Experiment Assumptions

Page Piccinini
Instructor





Computing conversion rate difference

```
eight_month_checkin_data_sum <- eight_month_checkin_data %>%  
  mutate(month_text = month(visit_date, label = TRUE)) %>%  
  group_by(month_text, condition) %>%  
  summarize(conversion_rate = mean(clicked_adopt_today))
```

```
# A tibble: 16 x 3  
  month_text condition conversion_rate  
  <ord>      <chr>      <dbl>  
1 Jan      cat_hat      0.3774194  
2 Jan      no_hat       0.1645161  
3 Feb      cat_hat      0.3535714  
4 Feb      no_hat       0.2250000  
5 Mar      cat_hat      0.3387097  
6 Mar      no_hat       0.1354839  
7 Apr      cat_hat      0.3433333  
8 Apr      no_hat       0.1366667  
9 May      cat_hat      0.4161290  
10 May     no_hat       0.2677419  
11 Jun     cat_hat      0.5433333  
12 Jun     no_hat       0.3066667  
13 Jul     cat_hat      0.4967742  
14 Jul     no_hat       0.3451613  
15 Aug     cat_hat      0.7935484  
16 Aug     no_hat       0.5806452
```



Computing conversion rate difference

```
eight_month_checkin_data_sum <- eight_month_checkin_data %>%  
  mutate(month_text = month(visit_date, label = TRUE)) %>%  
  group_by(month_text, condition) %>%  
  summarize(conversion_rate = mean(clicked_adopt_today))  
  
eight_month_checkin_data_diff <- eight_month_checkin_data_sum %>%  
  spread(condition, conversion_rate)
```

```
# A tibble: 8 x 3  
  month_text    cat_hat    no_hat  
  <ord>        <dbl>    <dbl>  
1      Jan 0.3774194 0.1645161  
2      Feb 0.3535714 0.2250000  
3      Mar 0.3387097 0.1354839  
4      Apr 0.3433333 0.1366667  
5      May 0.4161290 0.2677419  
6      Jun 0.5433333 0.3066667  
7      Jul 0.4967742 0.3451613  
8      Aug 0.7935484 0.5806452
```



Computing conversion rate difference

```
eight_month_checkin_data_sum <- eight_month_checkin_data %>%  
  mutate(month_text = month(visit_date, label = TRUE)) %>%  
  group_by(month_text, condition) %>%  
  summarize(conversion_rate = mean(clicked_adopt_today))  
  
eight_month_checkin_data_diff <- eight_month_checkin_data_sum %>%  
  spread(condition, conversion_rate) %>%
```




Computing conversion rate difference

```
eight_month_checkin_data_sum <- eight_month_checkin_data %>%  
  mutate(month_text = month(visit_date, label = TRUE)) %>%  
  group_by(month_text, condition) %>%  
  summarize(conversion_rate = mean(clicked_adopt_today))  
  
eight_month_checkin_data_diff <- eight_month_checkin_data_sum %>%  
  spread(condition, conversion_rate) %>%  
  mutate(condition_diff = cat_hat - no_hat)
```

```
# A tibble: 8 x 4  
  month_text    cat_hat    no_hat condition_diff  
    <ord>      <dbl>      <dbl>      <dbl>  
1      Jan 0.3774194 0.1645161    0.2129032  
2      Feb 0.3535714 0.2250000    0.1285714  
3      Mar 0.3387097 0.1354839    0.2032258  
4      Apr 0.3433333 0.1366667    0.2066667  
5      May 0.4161290 0.2677419    0.1483871  
6      Jun 0.5433333 0.3066667    0.2366667  
7      Jul 0.4967742 0.3451613    0.1516129  
8      Aug 0.7935484 0.5806452    0.2129032
```



Computing conversion rate difference

```
eight_month_checkin_data_sum <- eight_month_checkin_data %>%  
  mutate(month_text = month(visit_date, label = TRUE)) %>%  
  group_by(month_text, condition) %>%  
  summarize(conversion_rate = mean(clicked_adopt_today))  
  
eight_month_checkin_data_diff <- eight_month_checkin_data_sum %>%  
  spread(condition, conversion_rate) %>%  
  mutate(condition_diff = cat_hat - no_hat)
```

```
> mean(eight_month_checkin_data_diff$condition_diff)  
[1] 0.1876171  
> sd(eight_month_checkin_data_diff$condition_diff)  
[1] 0.03893739
```



A/B TESTING IN R

Let's practice!