

高级软件工程

第13周 (Nov. 22)

主讲：罗铁坚

助教：周文璋、俞永生、姚敏

上课：周一、三 上午 10:30—12:10 (教1-209)

答疑：周一、三 下午 2:30—3:30 (学园2-485)

联系：tjluo@ucas.ac.cn 69671829



提纲

1、A/B 实验检验

2、Split 实验检验的框架

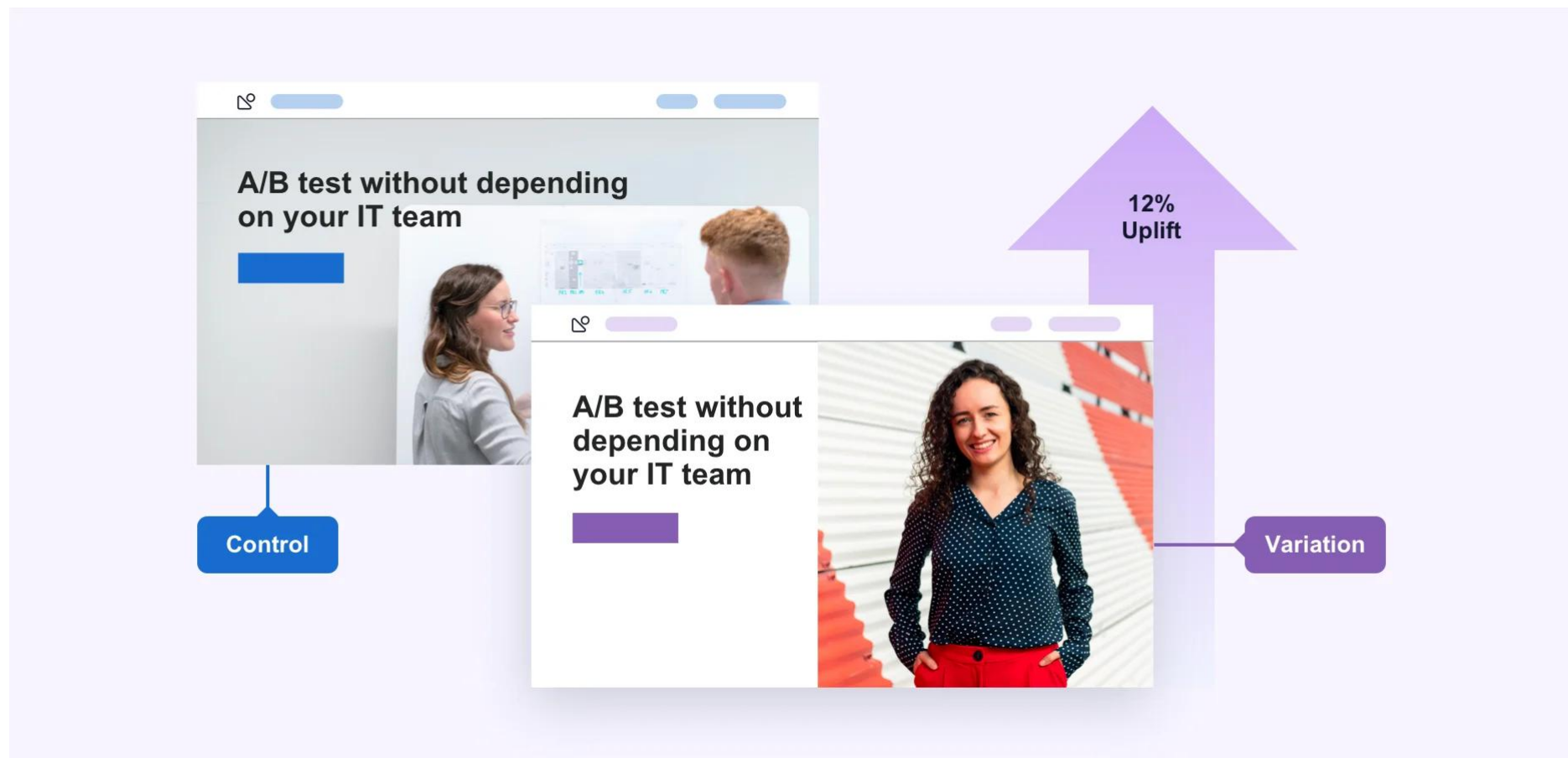
3、Scientists 实验检验的框架

软件质量保障和新功能的实验检验

QA method	Test against	Phase	How to get test data
Unit testing	Test spec	Dev	Manual
Integration testing	Test spec	Dev	Manual
Performance testing	Test spec	Tst	Dump production traffic /simulation
Acceptance testing	User spec	Acc	Manual
Feature flags	User expectations	Prd	Segment of production traffic
A/B-testing	Comparing options	Prd	Segment of production traffic
Blue/green deployments	User expectations	Prd	All production traffic
Canary releases	User expectations	Prd	Early segment of production traffic

什么是A/B 测试

A/B testing, also known as split testing, refers to a randomized experimentation process wherein two or more versions of a variable (web page, page element, etc.) are shown to different segments of website visitors at the same time to determine which version leaves the maximum impact and drive business metrics.



A/B 测试解决什么问题



**Solve visitor
pain points**



**Get better ROI from
existing traffic**



**Reduce
bounce rate**



**Make low-risk
modifications**

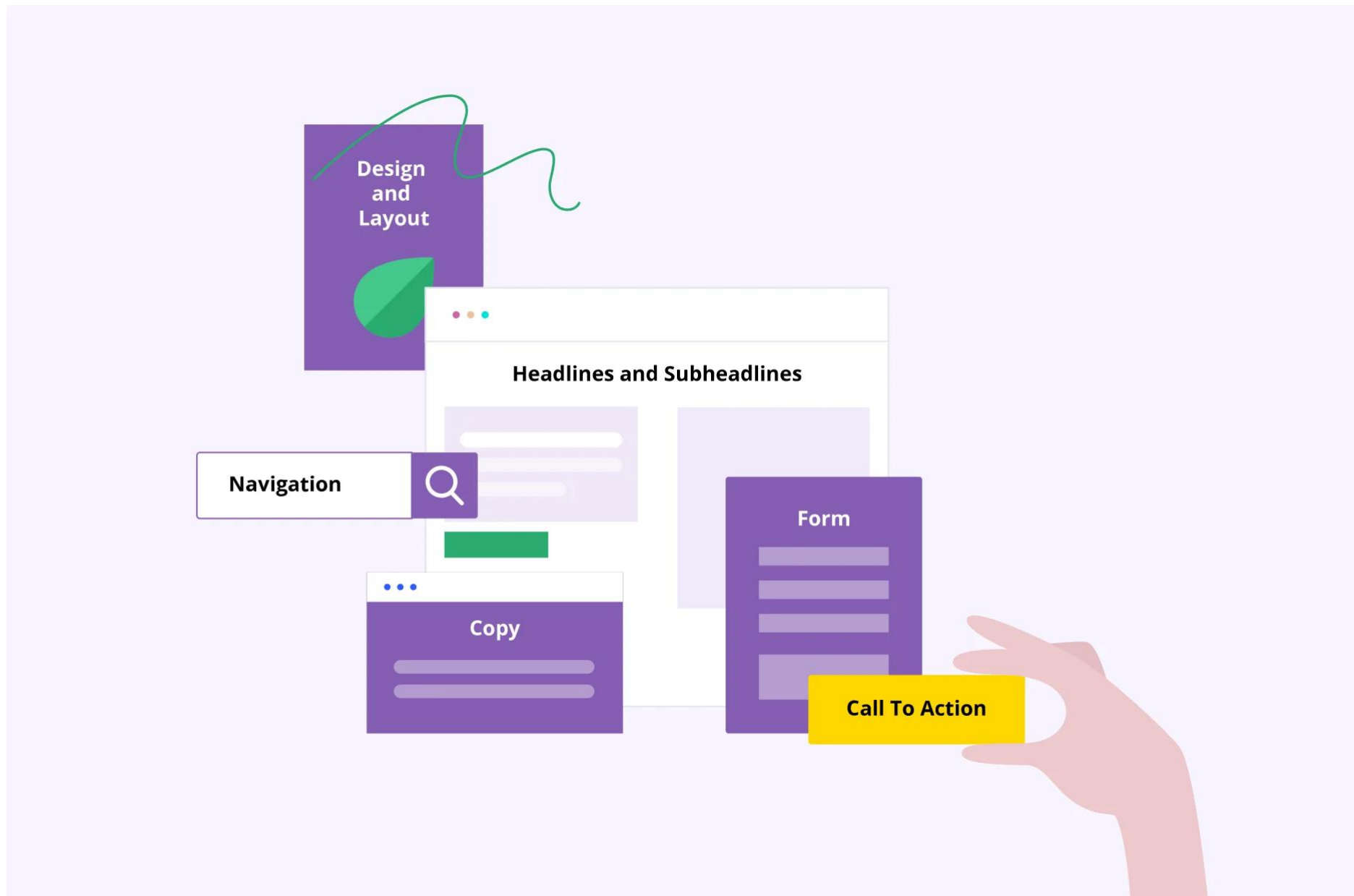


**Achieve statistically
significant
improvements**

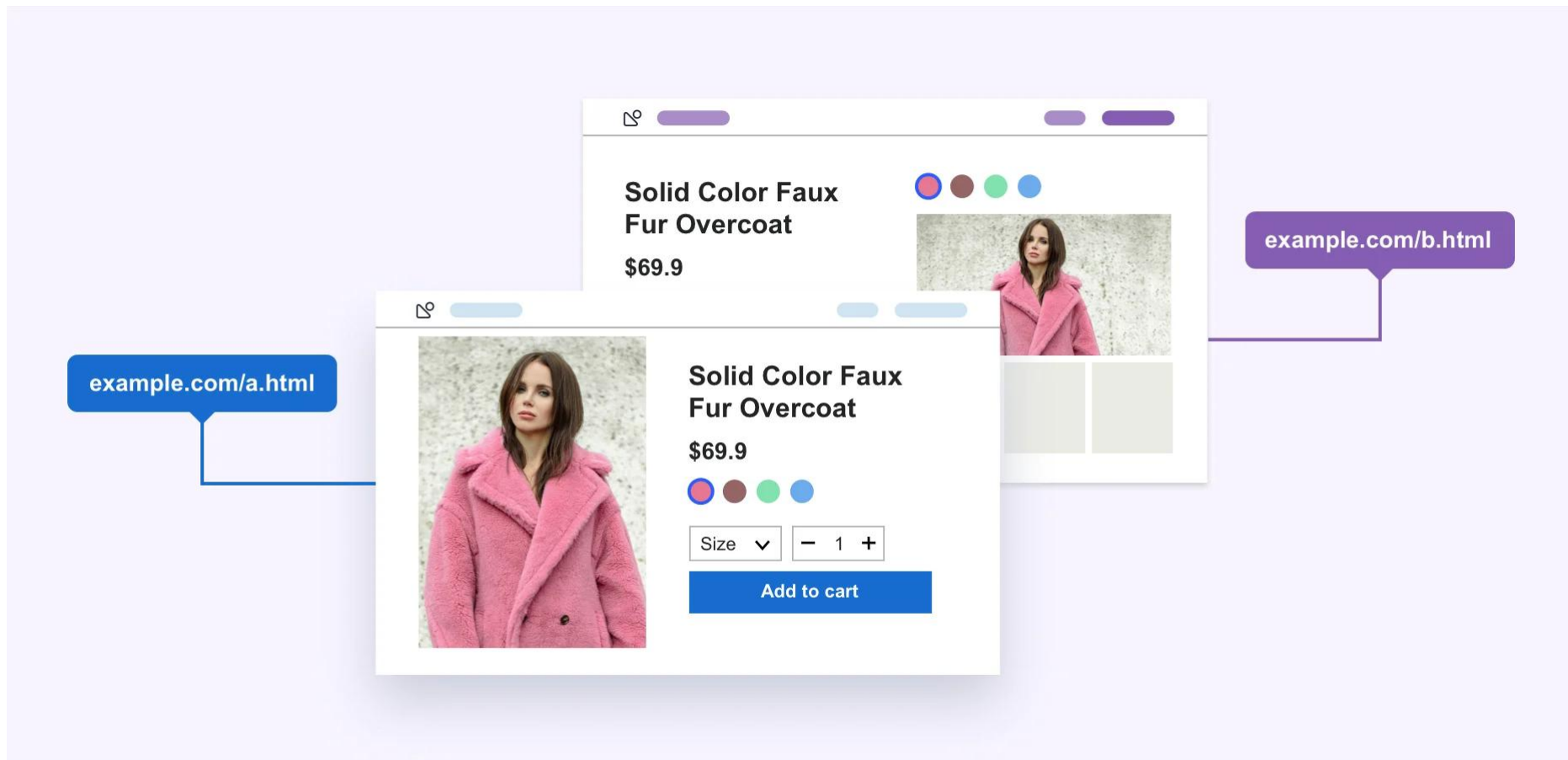


**Redesign website
to increase future
business gains**

A/B 测试中，你要做什么



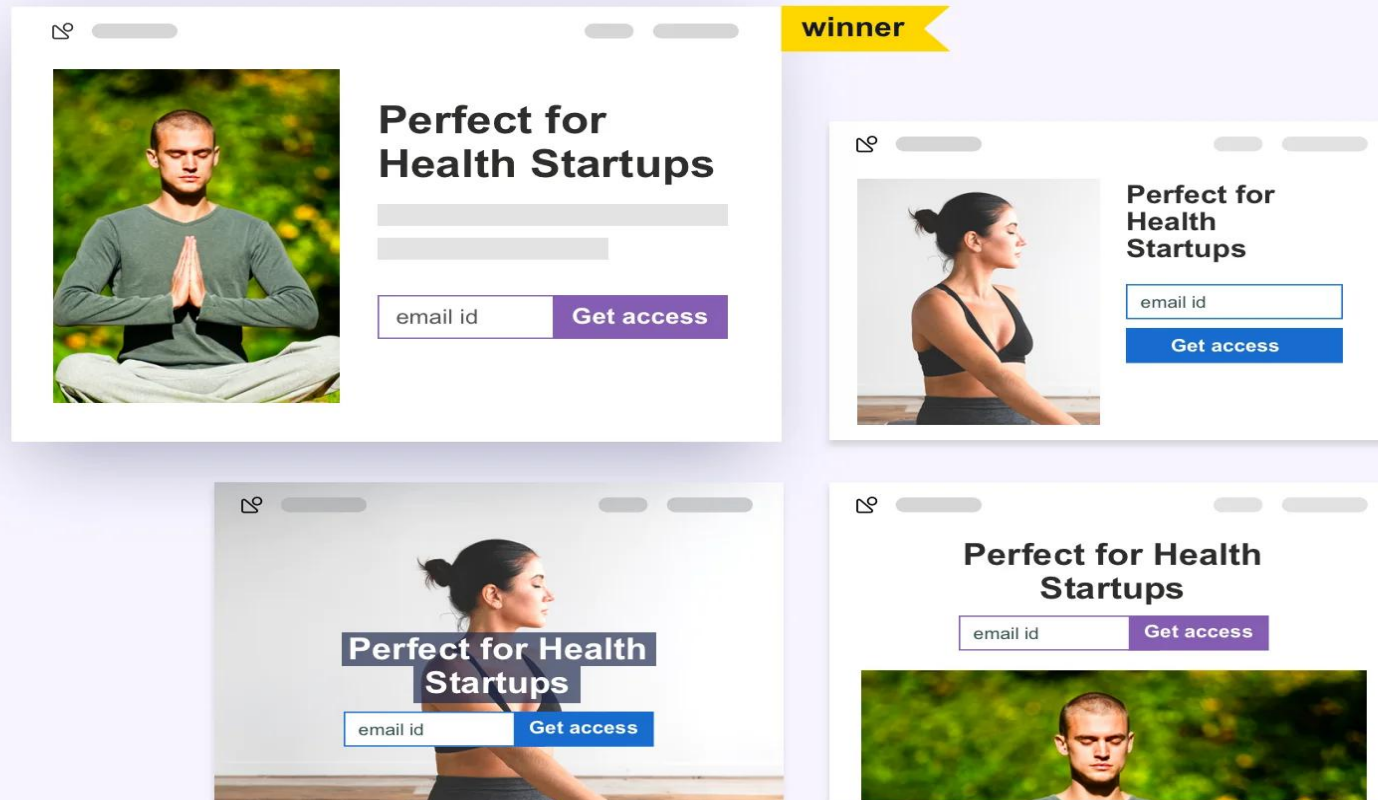
不同类型的A/B 测试-Split URL testing



Ideal for trying out radical new designs while using the existing page design for comparative analysis. Recommended for running tests with non-UI changes, such as switching to a different database, optimizing your page's load time, etc.

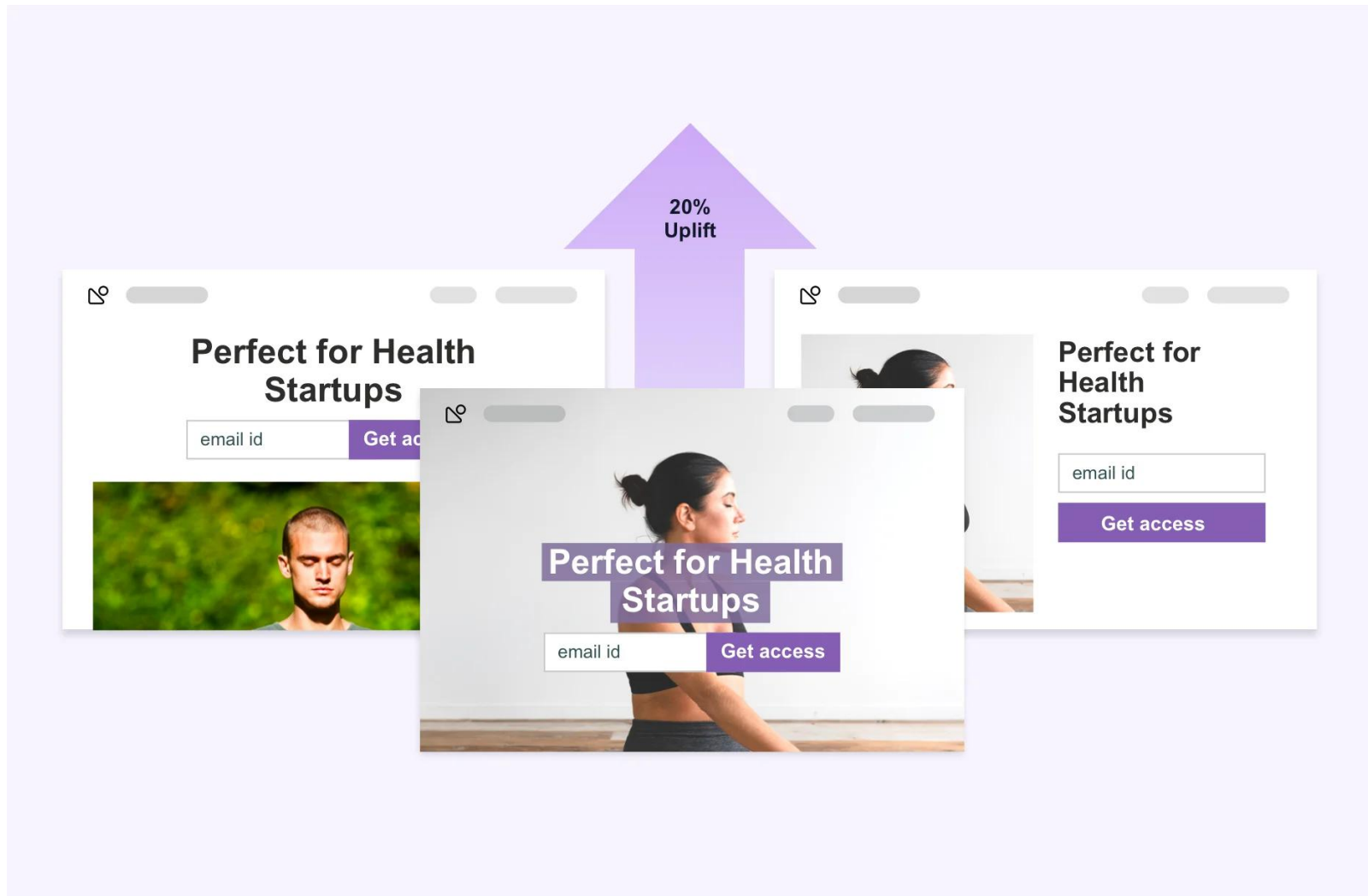
Change up web page workflows. Workflows dramatically affect business conversions, helping test new paths before implementing changes and determine if any of the sticking points were missed. A better and much-recommended testing method for dynamic content.

不同类型的A/B 测试-Multivariate testing



Helps avoid the need to conduct several sequential A/B tests with the same goal and saves time since you can simultaneously track the performance of various tested page elements. Easily analyze and determine the contribution of each page element to the measured gains, Map all the interaction between all independent element variations (page headlines, banner image, etc.).

不同类型的A/B 测试-Multipage testing



It enables you to create consistent experiences for your target audience.

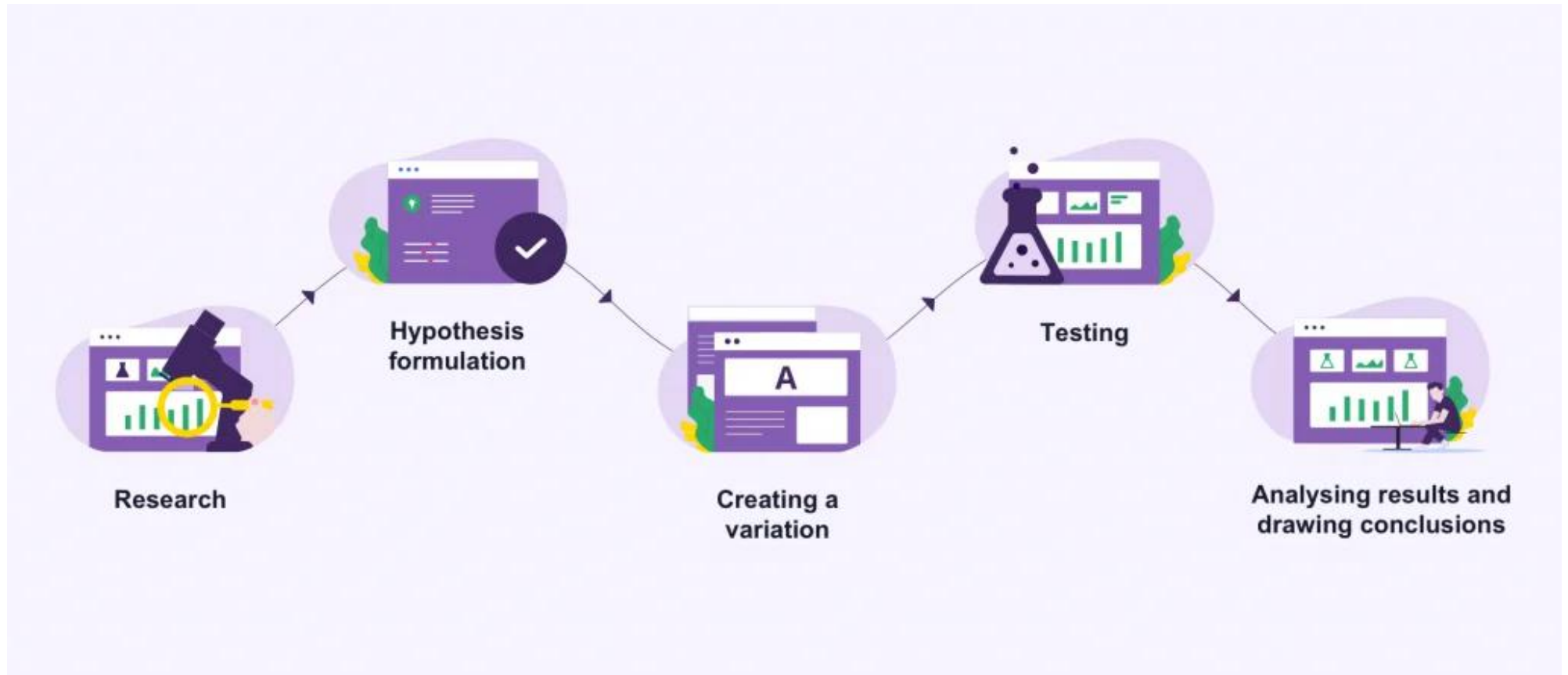
It helps your target audience see a consistent set of pages, no matter if it's the control or one of its variations.

It enables you to implement the same change on several pages to ensure that your website visitors don't get distracted and bounce off between different variations and designs when navigating through your website.

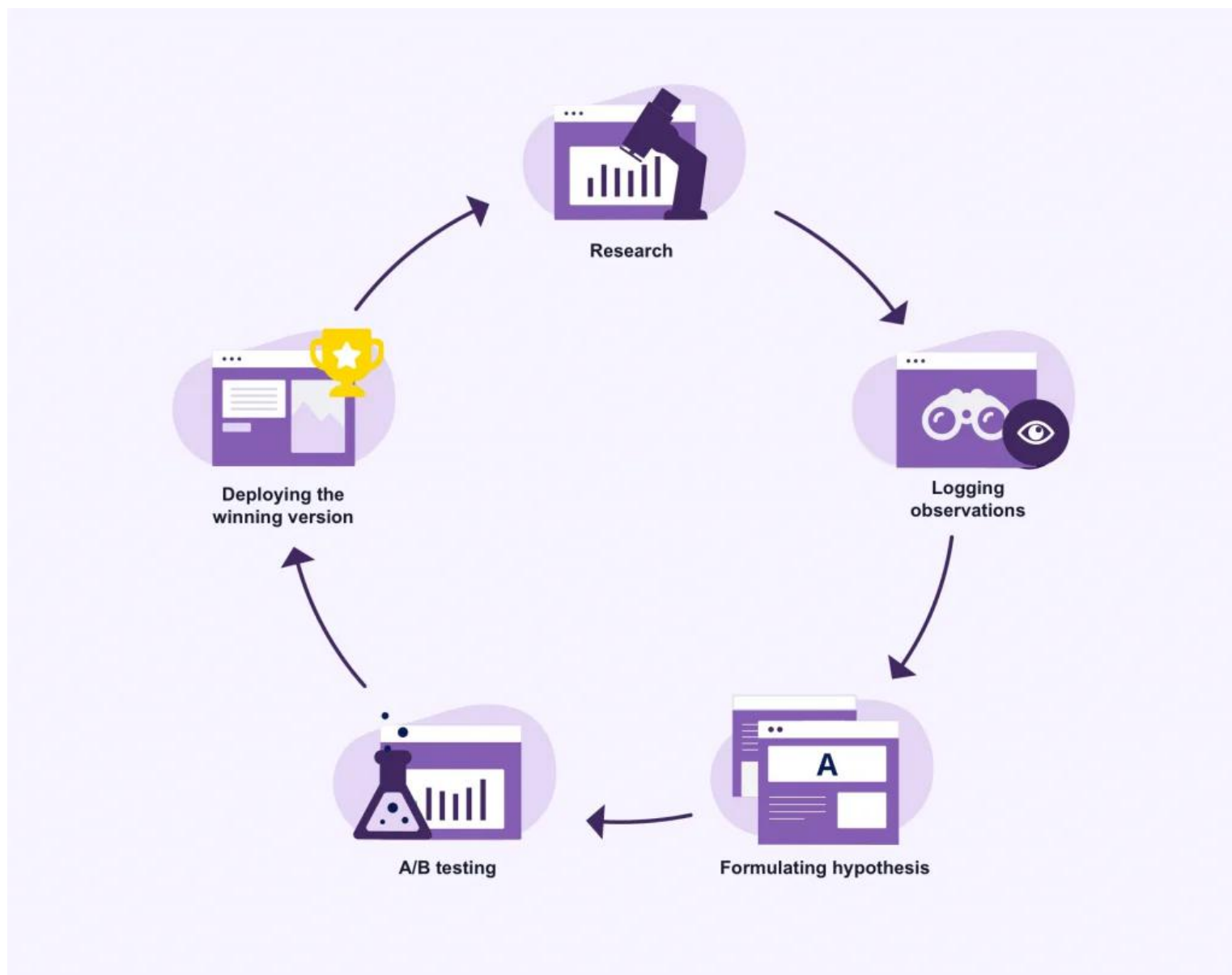
用什么理论方法

Frequentist Approach	Bayesian Approach
Frequentist Statistics follow the 'Probability as Long-Term Frequency' definition of probability.	Bayesian Statistics follow the notions of 'Probability as Degree of Belief' and 'Logical Probability.'
In this approach, you only use data from your current experiment. The frequentist solution is to conduct tests and draw conclusions.	In this approach, you use your prior knowledge from the previous experiments and try to incorporate that information into your current data. The Bayesian solution is to use existing data to draw conclusions.
Give an estimated mean (and standard deviation) of samples where A beats B but completely ignores the cases when B beats A.	It takes into account the possibility of A beating B and also calculates the range of the improvement you can expect.
Requires the test to run for a set period to get correct data from it but can't figure out how close or far A and B actually are. It fails to tell you the probability of A beating B.	Gives you more control over testing. You can now plan better, have a more accurate reason to end tests, and get into the nitty-gritty of how close or far apart A and B are.

A/B 测试步骤



A/B 测试步骤



More likely
to see



Less likely

10,000s of
possible
rows



Variable number of
possible videos per
row (up to thousands)



1 personalized page



10-40
rows

per device



V1 Variant

C Control



Join the 1.5M other property owners listing with Booking.com

List anything on Booking.com

Find guests with ease and earn more

Why Booking.com?

A larger global audience

We connect your property to a huge audience of potential guests, your way.

Free registration

There are absolutely no sign-up fees or commission costs when you advertise with us.

Booking.com and Airbnb all in one calendar

With our calendar, you can manage both your Booking.com and Airbnb listings in one place. It's the easiest way to manage your property.

Stories from newly listed properties



It's not just about the view

We make the most of the space of your property. Booking.com and Airbnb are the best way to get the most of your property. We help you manage all your listings in one place. We help you manage all your listings in one place.



It's not just about the view

We make the most of the space of your property. Booking.com and Airbnb are the best way to get the most of your property. We help you manage all your listings in one place. We help you manage all your listings in one place.


How it works


 Open and close your property when you want

We give you the freedom to open or close your property when you want. We give you the freedom to open or close your property when you want.

 We don't charge guests

We don't charge guests. We don't charge guests. We don't charge guests.






Rent out your place and pay for your passions with Booking.com

People all across the United Kingdom are earning easy money every week

Fancy making money from your place while you're away? Whether it's a primary residence or a secondary property, increasing numbers of property owners are renting out their places for short periods and earning easy money every week with Booking.com.



Register your place

Number of properties at different addresses

1

First and last name

Email

Register now

Already started a registration? [Continue here](#)

By continuing, you agree to our Booking.com terms and conditions regarding your property registration.

POSist

FEATURES

PRODUCTS

ENTERPRISE

BENEFITS

INTEGRATIONS

CUSTOMERS

CONTACT US



FOCUS ON RUNNING YOUR FOOD OUTLET

Leave The Technology Management To Us!

Learn how POSist can help you run your restaurant better

Name

Restaurant Name

Email Id

Phone Number

City

Schedule Demo



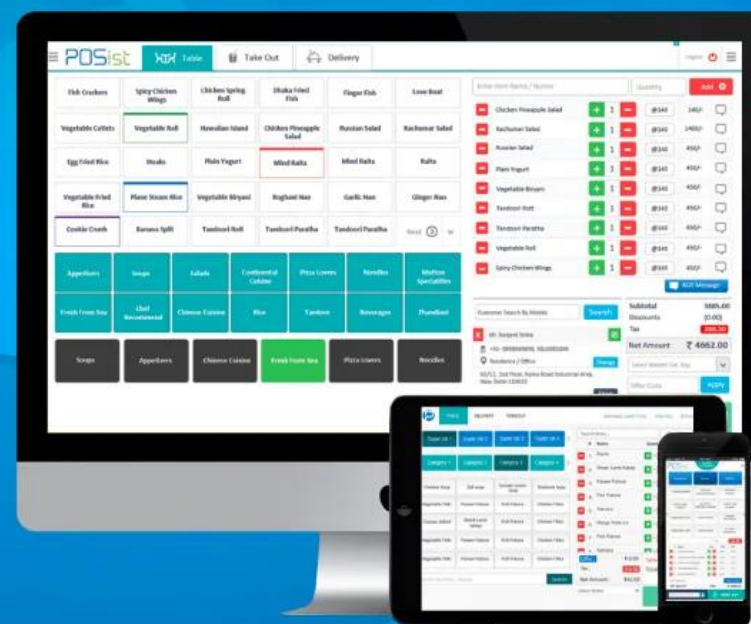
POSist

FEATURES ▾ PRODUCTS ▾ ENTERPRISE ▾ BENEFITS ▾ INTEGRATIONS CUSTOMERS CONTACT US

POS Software For Ever Growing Restaurants!

Trusted by over 5000+ Restaurants, POSist Simplifies Restaurant Operations & Boosts your Restaurant Business!

SIGN UP FOR FREE DEMO



Learn how POSist can help you run your restaurant better

Name

Restaurant Name

Email Id

Phone Number

City



中国科学院大学

+91-8010133399
info@posist.com

ABOUT US
THE RESTAURANT TIMES
RESELLER
NEWS

POSist

FEATURES
PRODUCTS
ENTERPRISE
BENEFITS
INTEGRATIONS
CUSTOMERS
GET DEMO

Cloud-Based Restaurant POS Software & Management System

SIGN UP FOR FREE DEMO

Trusted by over 5000+ Restaurants, POSist simplifies restaurant operations & boosts your restaurant business!



AB(C) testing with Ruby on Rails

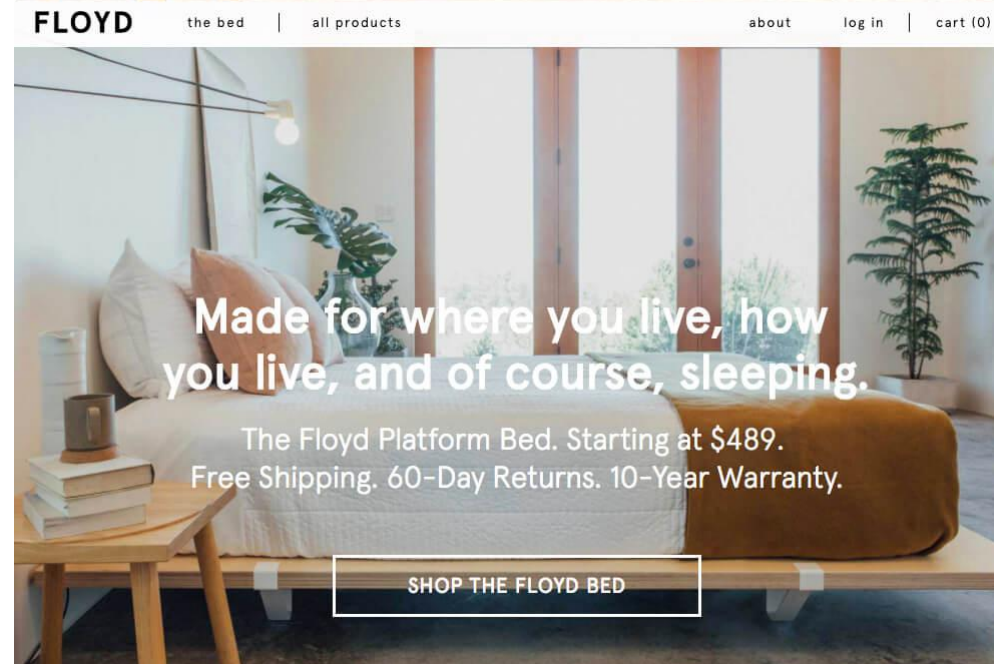
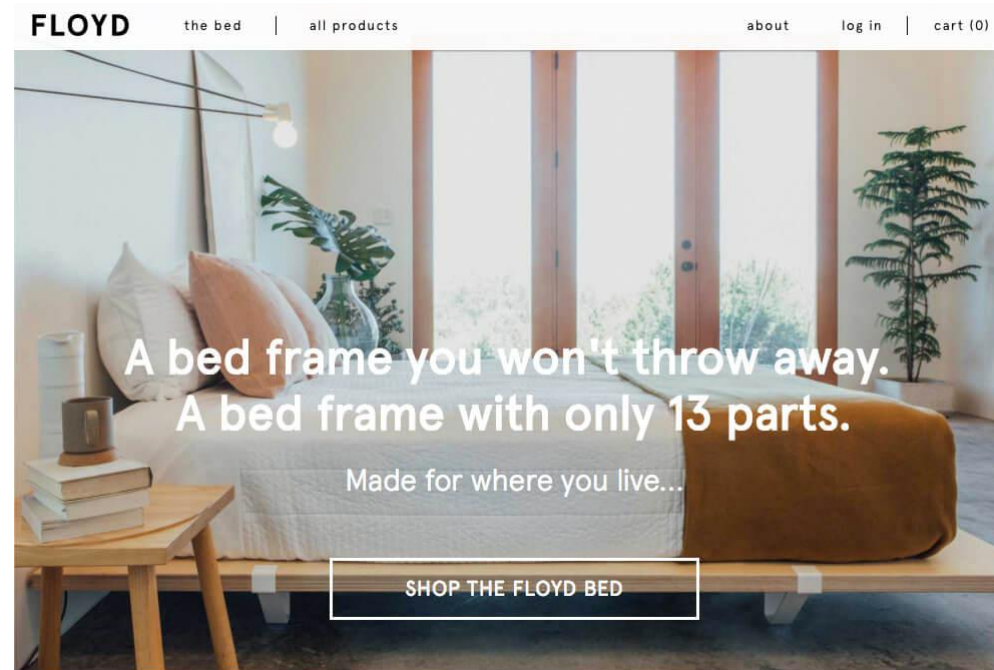
```
gem 'split', '~> 2.2.0', require: 'split/dashboard'
```

```
class YourController < ApplicationController
  def index
    @experiment = ab_test(:my_page, ["simple", "kind", "clickbait"])
  end
end
```

AB(C) testing with Ruby on Rails

```
<% if @experiment == 'simple' %>
  <p>That's a simple tutorial</p>
<% elsif @experiment == 'kind' %>
  <p>Hi, can you please read this tutorial?</p>
<% elsif @experiment == 'clickbait' %>
  <p>This tutorial is so cool, check it out!</p>
<% end %>
```

AB(C) testing with Ruby on Rails



AB(C) testing with Ruby on Rails

```
namespace :admin do
  authenticate :admin_user, ->(u) { u.admin? } do
    mount Split::Dashboard, at: '/split'
  end
end
```


AB(C) testing with Split --实验结果看板

Split Dashboard

Development

The list below contains all the registered experiments along with the number of test participants, completed and conversion rate currently in the system.

Begin typing to filter

Hide completed

Hide active

Clear filters

Experiment: homepage v2

2017-04-04

Reset Data

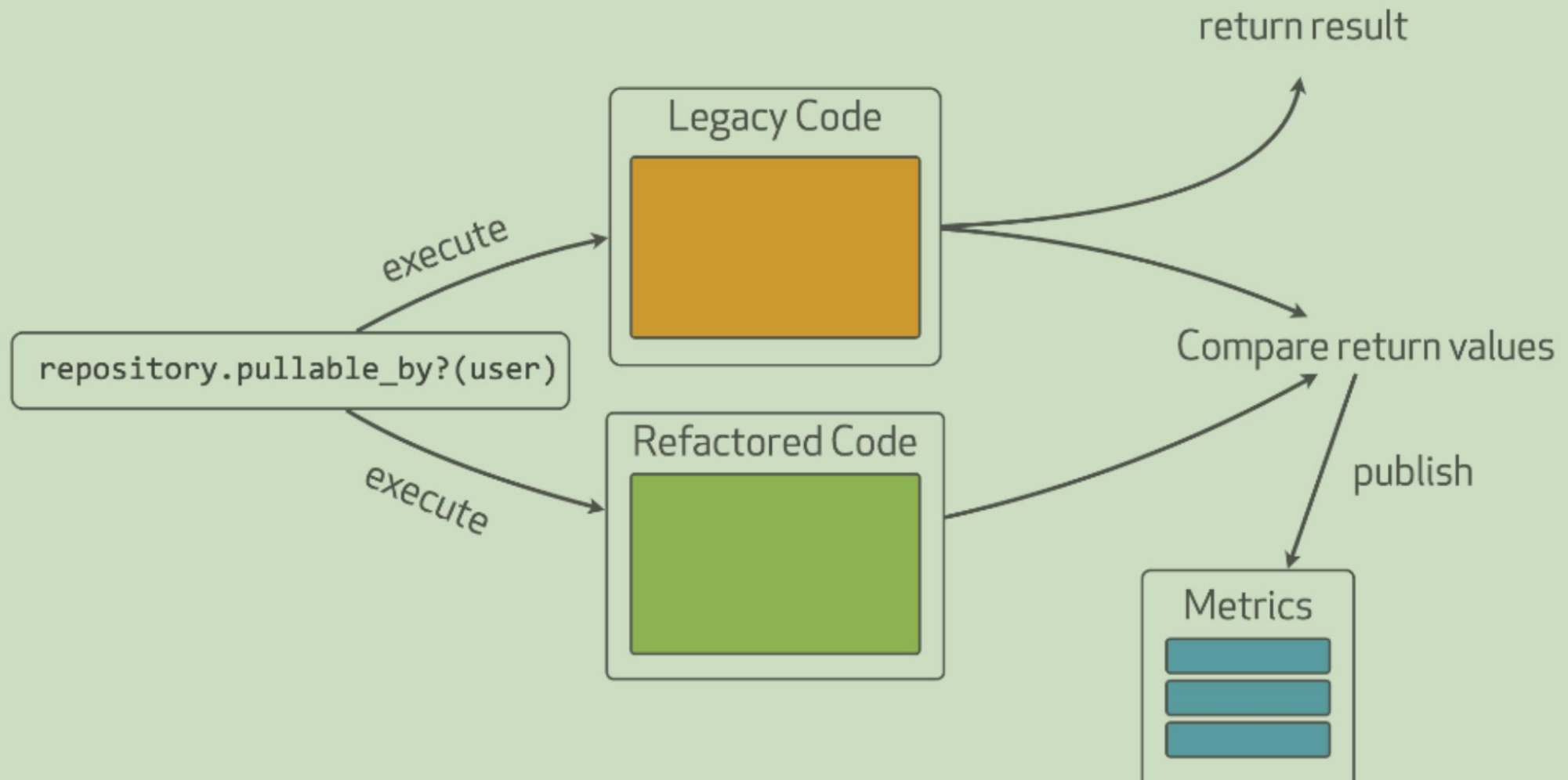
Delete

ALTERNATIVE NAME	PARTICIPANTS	NON-FINISHED	COMPLETED	CONVERSION RATE	Probability of being Winner	FINISH
origin <i>control</i> Force for current user	3	2	1	33.33%	37.5%	Use this
new bedframe copy Force for current user	2	1	1	50.0% +50.0%	62.5%	Use this
Totals	5	3	2	N/A	N/A	N/A

Powered by Split v2.2.0

A/B Test and Scientists

GitHub's Scientist Aims to Help Refactoring Critical Paths



How to ensure Refactoring Critical Path is Good

```
experiment = Scientist::Default.new "my-experiment"  
experiment.use { <call the old code here, the control> }  
experiment.try { <call the new code here, the candidate>  
experiment.run
```

```
#...
```

```
def publish(result)
```

```
#...
```


A/B Test and Scientists

On Scientist

GitHub's original Ruby gem can be found at <https://github.com/github/scientist>

The backstory of GitHub's Scientist can be found at

- "Move Fast and Fix Things",
<https://github.blog/2015-12-15-move-fast/>
- "Scientist: Measure Twice, Cut Over Once",
<https://github.blog/2016-02-03-scientist/>

Implementations / ports of the Scientist gem to other languages are listed at <https://github.com/github/scientist#alternatives>



练习

- 1、 课堂练习：什么是A/B 实验验证，为什么做A/B实验验证？如何做A/B实验验证？
- 2、 课外练习：分别安装使用Scientist 与 Split的实验验证框架，对比分析它们的异同。