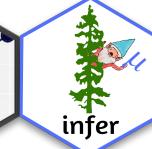




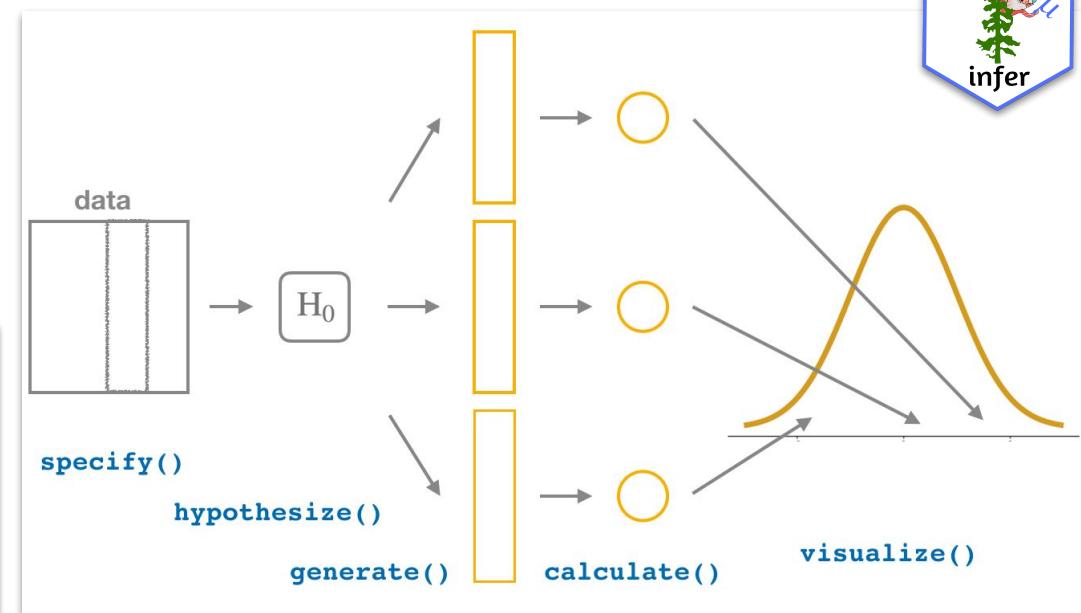
Optimize



A/B Testing

& Statistical Inference

Special Topic: Website Optimization



Matt Dancho & David Curry
Business Science Learning Lab





Learning Lab Structure

- **Presentation**
(20 min)
- **Demo's**
(30 min)
- **Pro-Tips**
(15 mins)



Matt Dancho

Founder of Business Science, Matt designs and executes educational courses and workshops that deliver immediate value to organizations. His passion is up-leveling future data scientists coming from untraditional backgrounds.

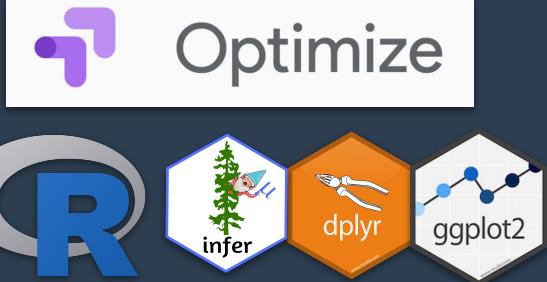


David Curry

Founder of Sure Optimize, David works with businesses to help improve website performance and SEO using data science. His passion is ethical Machine Learning initiatives.

Agenda

The image shows two side-by-side promotional banners. The left banner has a yellow background and says "Join over 15,000 Data Scientists and DOWNLOAD OUR ULTIMATE DATA SCIENCE CHEAT SHEETS FOR FREE". It includes a list of interests: Interested in R, Python, Segmentation, and Machine Learning, with a "Click Here to Download!" button. The right banner has a black background and says "Join over 20,000 Data Scientists and DOWNLOAD MY ULTIMATE DATA SCIENCE CHEAT SHEETS FOR FREE". It includes a similar list of interests and a "Get My Cheat Sheets" button.



- **Business Case Study**
 - Website Optimization
 - A/B Testing
- **30-Min Demo**
 - A/B Test Analysis in R w/ infer
- **A/B Testing**
 - What is it?
 - What can you do?
- **Pro-Tips & Learning Guide**
 - Recap + Pro-Tips
 - Learning Plan
- **Google Optimize**
 - Free tool for A/B Testing
- **Statistical Inferences**
 - 80/20 Concepts
 - R package infer



Learning Labs PRO

Every 2-Weeks

1-Hour Course

Recordings + Code + Slack

\$19/month

university.business-science.io

Lab 23
BigQuery & Conversion Funnel

Lab 22
SQL for Time Series

Lab 21
SQL for Data Science

Lab 20
Explainable Machine Learning

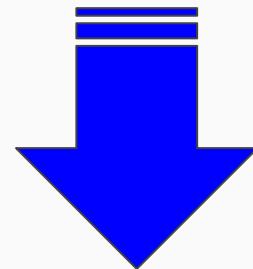
Lab 19
Using Customer Credit Card History for Networks Analysis

Lab 18
Time Series Anomaly Detection with anomalize

Lab 17
Anomaly Detection with H2O Machine Learning



Continuous Learning
Jet Fuel for your Brain



Learning Labs Pro

Community-Driven Data Science Courses

 Matt Dancho

\$19/m

Website Optimization

Business Case



Business Science Website

Email Signup

- Email is critical to **my business**
- Way to connect with **my people**
- Relationship + Value = **Trust**
- Email is the **1st Step** towards building trust

<https://www.business-science.io/>

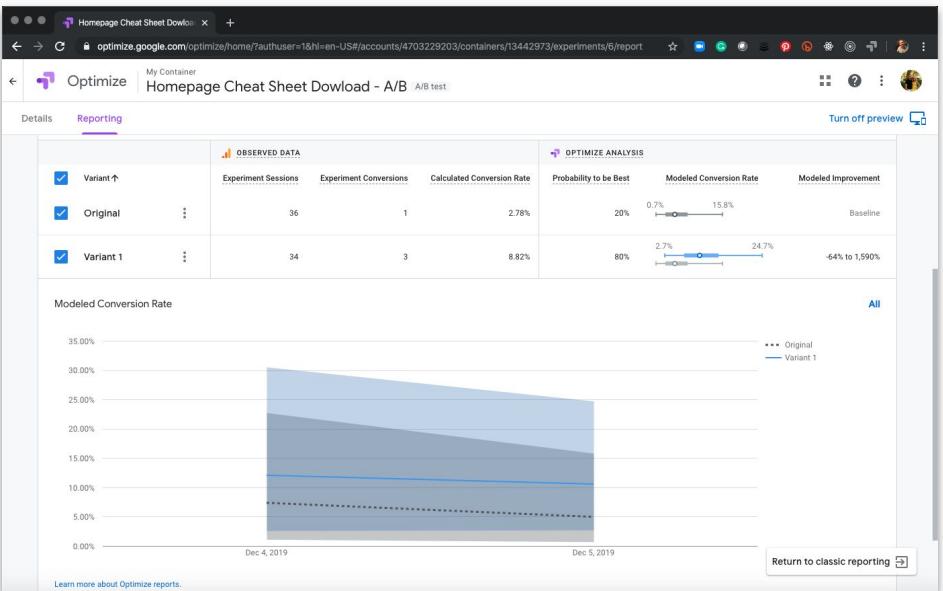
The screenshot shows the homepage of business-science.io. At the top, there's a navigation bar with links for 'START HERE', 'COURSES', 'TESTIMONIALS', 'CONTACT', and 'LOG IN'. A search bar is also present. The main content area features a dark background image of two people working at a computer. Overlaid on this image is the text 'Business Science' and 'Data Science Courses for Business'. To the left, there's a social sharing sidebar with icons for Facebook, Twitter, Google+, and Pinterest, showing 4 shares. To the right, a yellow call-to-action box contains the text 'Join over 15,000 Data Scientists and DOWNLOAD OUR ULTIMATE DATA SCIENCE CHEAT SHEETS FOR FREE'. Below this, there's a section for answering quick questions to receive a cheat sheet. At the bottom right of the page, there's a blue button labeled 'Click Here to Download!'

Tracking Conversions

For Business Science



1. Google made a new tool **Google Optimize** that helps with A/B Testing for websites
2. **All-in-one tool** which helps streamline the A/B test process
3. **Integrates** with Google Tag Manager & Google Analytics
4. **Other types of Conversions**
 - **Email Marketing** - Test email variants

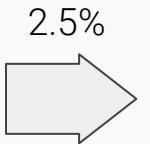
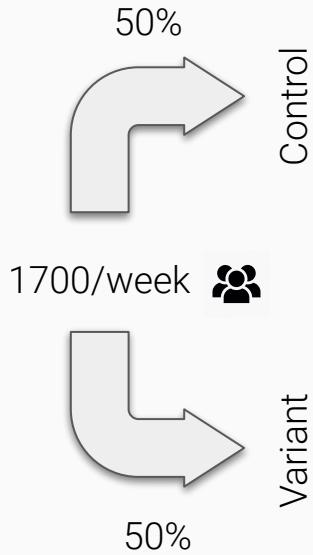


A/B Testing

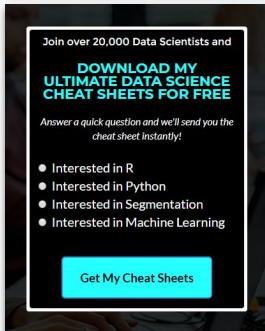
80/20 Concepts



What is A/B Testing?



$1700/\text{wk} \times 2.5\% = 43 \text{ emails/wk}$



3X Improvement
Versus Control

$1700/\text{wk} \times 7.5\% = 128 \text{ emails/wk}$



Small Change, Big Effect

Join over 15,000 Data Scientists and
**DOWNLOAD OUR
ULTIMATE DATA SCIENCE
CHEAT SHEETS FOR FREE**

Answer a quick question and we'll send you the
cheat sheet instantly!

- Interested in R
- Interested in Python
- Interested in Segmentation
- Interested in Machine Learning

[Click Here to Download!](#)

2.5% Conv.
42 emails/week

Join over 20,000 Data Scientists and
**DOWNLOAD MY
ULTIMATE DATA SCIENCE
CHEAT SHEETS FOR FREE**

Answer a quick question and we'll send you the
cheat sheet instantly!

- Interested in R
- Interested in Python
- Interested in Segmentation
- Interested in Machine Learning

[Get My Cheat Sheets](#)

7.5% Conv.
128 emails/week

What Can You A/B Test?



Website

How to optimize?

Improving your website using
**statistically significant
results**

- Solve pain points
- Improve ROI from web traffic
- Reduce bounce rate
- Improve customer experience
- Achieve better results

Blog

What interests your
prospective subscribers?

Finding content & topics that
connect you to your people.

- Improve web traffic & click rates
- Optimize title (traffic)
- Optimize image (traffic)
- Optimize content (click)

Email

What interests your current
subscribers?

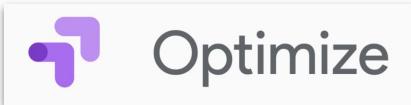
Finding content & topics that
grow your relationship with
your people.

- Improve Open & Click Rates
- Optimize title (traffic)
- Optimize image (click)
- Optimize content (click)



A/B Test Tools

Google Optimize



Pros - Fast & Easy to Set Up

Cons:

- Web Only (No Email)
- No advanced ML

R Ecosystem



Pros:

- Any Data (Email, Web, Other)
- Answer more questions:
 - What correlates to success?
- Apply advanced tools
 - Machine Learning

Cons: Learn R Ecosystem

Google Optimize

Concepts & Short Demo

Google Marketing Platform



Platform Home x + marketingplatform.google.com/info?authuser=1

Google Marketing Platform ? ... Profile picture

Looking for insights and data?

 **Analytics**
Get a deeper understanding of your customers. Google Analytics gives you the free tools you need to analyze data for your business in one place.

[Set up](#) [Learn more](#)

 **Data Studio**
Unlock the power of your data with interactive dashboards and engaging reports that inspire smarter business decisions. It's easy and free.

[Create report](#) [Learn more](#)

 **Optimize**
Make a great impression on each and every visitor. Easily run tests on your website — for free — so it works better for your customers and your business.

[Set up](#) [Learn more](#)

 **Surveys**
Google Surveys gives you a quick, cost-effective way to get valuable insights into the minds of your target audience. Gather the insights you need to make smarter and faster business decisions.

[Set up](#) [Learn more](#)

 **Tag Manager**
Get up and running with measurement faster. Tag Manager delivers simple, reliable, easily integrated tag management — for free.

[Set up](#) [Learn more](#)



Google Optimize

Run A/B Tests

Integrates with:

- Google Analytics
- Google Tag Manager
- And more

The screenshot shows the 'Platform Home' page at marketingplatform.google.com/info?authuser=1. The page features a header with the Google Marketing Platform logo and a search bar. Below the header, there are several cards for different tools:

- Analytics**: Get a deeper understanding of your customers. Google Analytics gives you the free tools you need to analyze data for your business in one place. Includes 'Set up' and 'Learn more' buttons.
- Data Studio**: Unlock the power of your data with interactive dashboards and engaging reports that inspire smarter business decisions. It's easy and free. Includes 'Create report' and 'Learn more' buttons.
- Optimize** (highlighted with a green circle): Make a great impression on each and every visitor. Easily run tests on your website — for free — so it works better for your customers and your business. Includes 'Set up' and 'Learn more' buttons.
- Surveys**: Google Surveys gives you a quick, cost-effective way to get valuable insights into the minds of your target audience. Gather the insights you need to make smarter and faster business decisions. Includes 'Set up' and 'Learn more' buttons.
- Tag Manager**: Get up and running with measurement faster. Tag Manager delivers simple, reliable, easily integrated tag management — for free. Includes 'Set up' and 'Learn more' buttons.

DEMO

Statistical Inference

80/20 Data Concepts

Infer R Package



```
# A tibble: 290,585 x 7
  user_id timestamp      group landing_page converted label      visit_count
  <dbl> <dttm>     <chr>   <chr>        <dbl> <chr>      <dbl>
1 922696 2017-01-02 13:42:05 treatment new_page    0 treatment-new_page 1
2 781507 2017-01-02 13:42:15 control  old_page     0 control-old_page 1
3 737319 2017-01-02 13:42:21 control  old_page     0 control-old_page 1
4 818377 2017-01-02 13:42:26 treatment new_page    0 treatment-new_page 1
5 725857 2017-01-02 13:42:27 treatment new_page    0 treatment-new_page 1
6 762651 2017-01-02 13:42:28 treatment new_page    0 treatment-new_page 1
7 722516 2017-01-02 13:42:34 treatment new_page    0 treatment-new_page 1
8 688937 2017-01-02 13:42:41 treatment new_page    0 treatment-new_page 1
9 781125 2017-01-02 13:42:45 treatment new_page    0 treatment-new_page 1
10 799109 2017-01-02 13:42:45 control  old_page     0 control-old_page 1
# ... with 290,575 more rows
```

Infer R Package



```
# A tibble: 290,585 x 7
  user_id timestamp      group landing_page converted label      visit_count
  <dbl> <dttm>     <chr>   <chr>        <dbl> <chr>        <dbl>
1 922696 2017-01-02 13:42:05 treatment new_page    0 treatment-new_page 1
2 781507 2017-01-02 13:42:15 control  old_page    0 control-old_page 1
3 737319 2017-01-02 13:42:21 control  old_page    0 control-old_page 1
4 818377 2017-01-02 13:42:26 treatment new_page    0 treatment-new_page 1
5 725857 2017-01-02 13:42:27 treatment new_page    0 treatment-new_page 1
6 762651 2017-01-02 13:42:28 treatment new_page    0 treatment-new_page 1
7 722516 2017-01-02 13:42:34 treatment new_page    0 treatment-new_page 1
8 688937 2017-01-02 13:42:41 treatment new_page    0 treatment-new_page 1
9 781125 2017-01-02 13:42:45 treatment new_page    0 treatment-new_page 1
10 799109 2017-01-02 13:42:45 control  old_page    0 control-old_page 1
# ... with 290,575 more rows
```



```
# A tibble: 46 x 5
  label          date      count conversion prop
  <chr> <dttm>    <int>      <dbl> <dbl>
1 control-old_page 2017-01-02 00:00:00  2859    359 0.126
2 control-old_page 2017-01-03 00:00:00  6590    750 0.114
3 control-old_page 2017-01-04 00:00:00  6578    802 0.122
4 control-old_page 2017-01-05 00:00:00  6427    792 0.123
5 control-old_page 2017-01-06 00:00:00  6606    762 0.115
6 control-old_page 2017-01-07 00:00:00  6604    799 0.121
7 control-old_page 2017-01-08 00:00:00  6687    795 0.119
8 control-old_page 2017-01-09 00:00:00  6628    793 0.120
9 control-old_page 2017-01-10 00:00:00  6654    751 0.113
10 control-old_page 2017-01-11 00:00:00  6688    795 0.119
# ... with 36 more rows
```



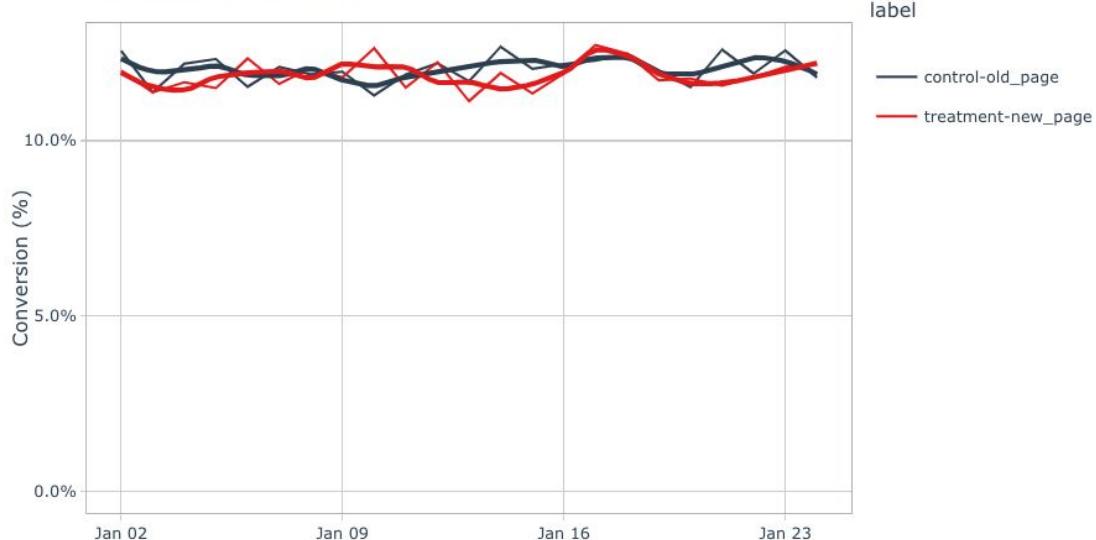
Infer R Package

```
# A tibble: 290,585 x 7
```

```
  user_id timestamp      group landing_page converted label
  <dbl> <dttm>        <chr>   <chr>       <dbl> <chr>
1 922696 2017-01-02 13:42:05 treatment new_page
2 781507 2017-01-02 13:42:15 control  old_page
3 737319 2017-01-02 13:42:21 control  old_page
4 818377 2017-01-02 13:42:26 treatment new_page
5 725857 2017-01-02 13:42:27 treatment new_page
6 762651 2017-01-02 13:42:28 treatment new_page
7 722516 2017-01-02 13:42:34 treatment new_pa
8 688937 2017-01-02 13:42:41 treatment new_pa
9 781125 2017-01-02 13:42:45 treatment new_pa
10 799109 2017-01-02 13:42:45 control  old_pa
# ... with 290,575 more rows
```

```
  visit_count
  <dbl>
0 treatment-new_page     1
0 control-old_page       1
0 control-old_page       1
0 treatment-new_page     1
0 treatment-new_page     1
0 treatment-new_page     1
```

Conversion Over Time



```
# A tibble: 46 x 2
  label      date
  <chr>    <dttm>
1 control-old_page 2017-01-02 00:00:
2 control-old_page 2017-01-03 00:00:
3 control-old_page 2017-01-04 00:00:
4 control-old_page 2017-01-05 00:00:
5 control-old_page 2017-01-06 00:00:
6 control-old_page 2017-01-07 00:00:
7 control-old_page 2017-01-08 00:00:
8 control-old_page 2017-01-09 00:00:
9 control-old_page 2017-01-10 00:00:
10 control-old_page 2017-01-11 00:00:
# ... with 36 more rows
```

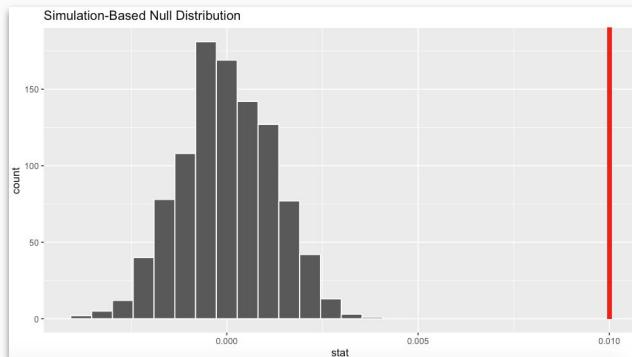




Types of Statistical Inference Tests

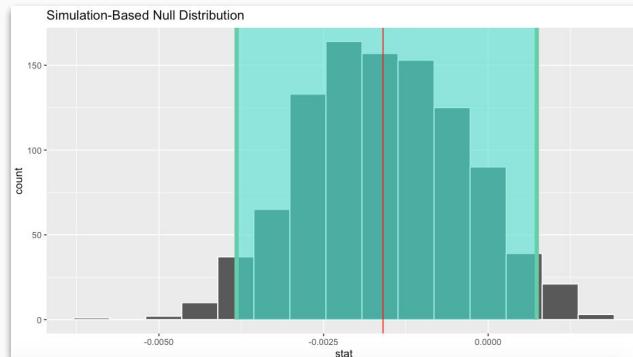
Hypothesis (Permutation)

Probability of a conversion more extreme than 1.0%

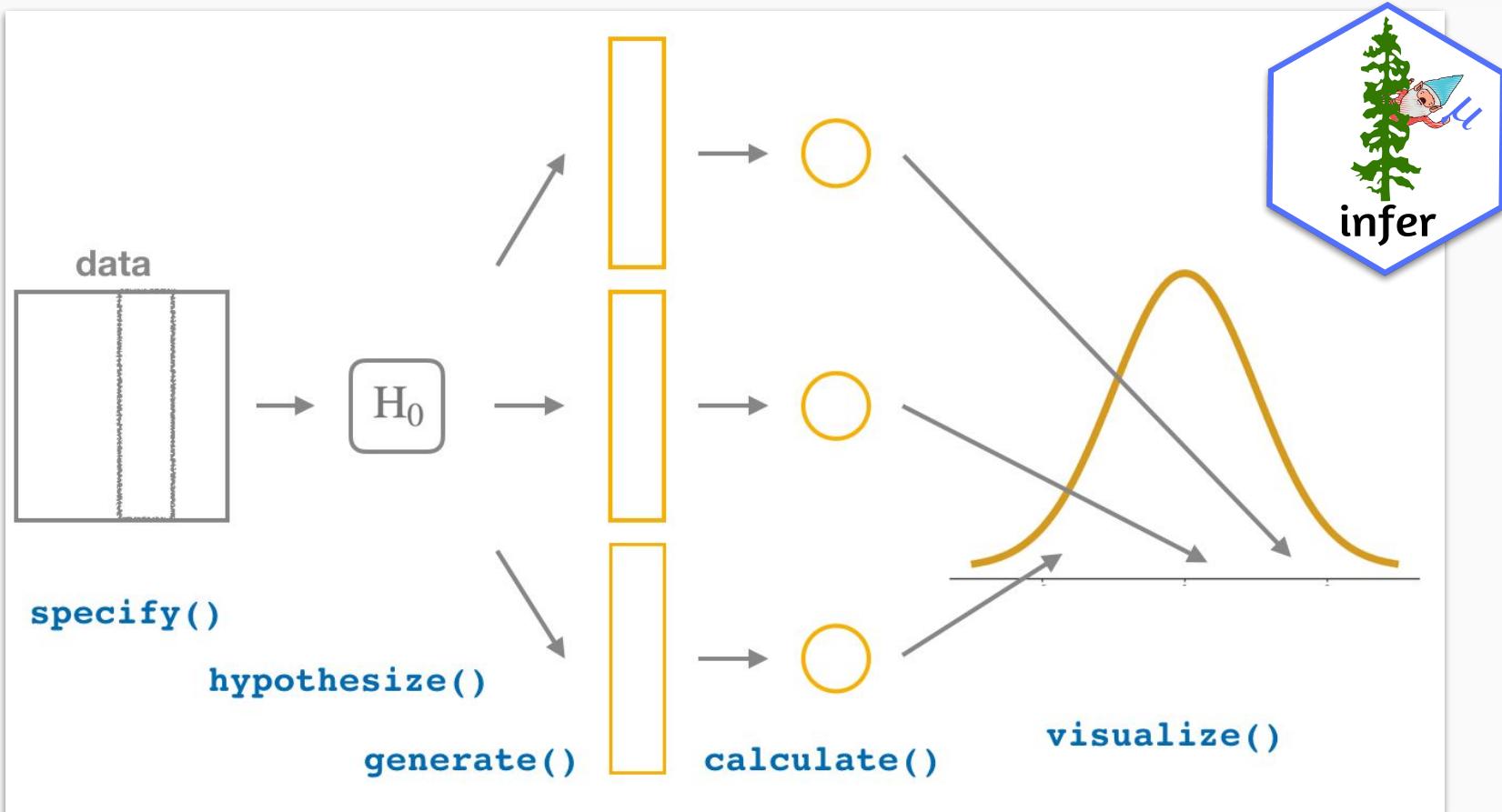


Confidence Interval (Bootstrap)

95% of time the difference in conversion is within this **Confidence Interval** window



Infer R Package

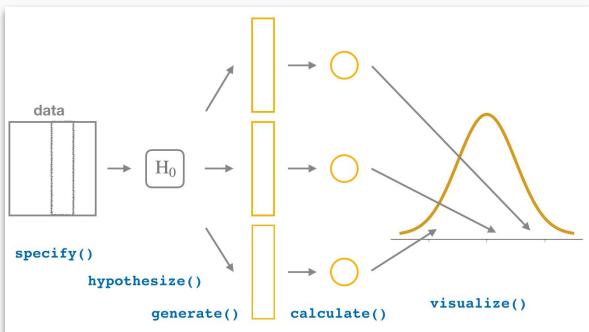


Infer Workflow



specify()

Specify the variables in the A/B Test



```
> comparison_1_tbl %>%
+   ...
+   # Infer
+   specify(prop ~ label)
Response: prop (numeric)
Explanatory: label (factor)
# A tibble: 46 x 2
  prop      label
  <dbl> <fct>
1 0.126 control-old_page
2 0.114 control-old_page
3 0.122 control-old_page
4 0.123 control-old_page
5 0.115 control-old_page
6 0.121 control-old_page
7 0.119 control-old_page
8 0.120 control-old_page
9 0.113 control-old_page
10 0.119 control-old_page
# ... with 36 more rows
```



Infer Workflow



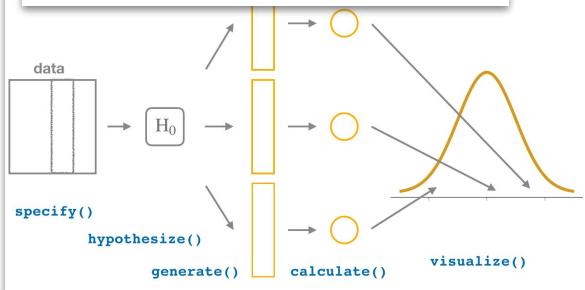
hypothesize()

For Hypothesis Testing
Only

(Not for Bootstrap CI
Estimation)

$$H_0 : p_m - p_f = 0$$

vs $H_A : p_m - p_f > 0$



```
> comparison_1_tbl %>%
+   # Infer
+   specify(prop ~ label) %>%
+   hypothesize(null = "independence")
Response: prop (numeric)
Explanatory: label (factor)
Null Hypothesis: independence
# A tibble: 46 x 2
  prop    label
  <dbl> <fct>
1 0.126 control-old_page
2 0.114 control-old_page
3 0.122 control-old_page
4 0.123 control-old_page
5 0.115 control-old_page
6 0.121 control-old_page
7 0.119 control-old_page
8 0.120 control-old_page
9 0.113 control-old_page
10 0.119 control-old_page
# ... with 36 more rows
```

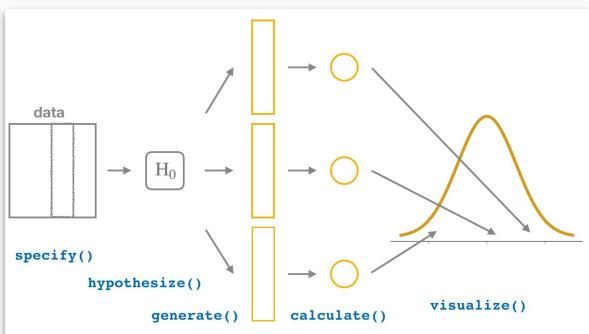


Infer Workflow



generate()

Make permutations or
bootstrap replicates



```
> comparison_1_tbl %>%
+   # Infer
+   specify(prop ~ label) %>%
+   hypothesize(null = "independence") %>%
+   generate(reps = 1000, type = "permute")
Response: prop (numeric)
Explanatory: label (factor)
Null Hypothesis: independence
# A tibble: 46,000 x 3
# Groups:   replicate [1,000]
  prop label      replicate
  <dbl> <fct>        <int>
1 0.126 control-old_page     1
2 0.125 control-old_page     1
3 0.119 control-old_page     1
4 0.113 control-old_page     1
5 0.113 control-old_page     1
6 0.119 control-old_page     1
7 0.122 control-old_page     1
8 0.119 control-old_page     1
9 0.126 control-old_page     1
10 0.117 control-old_page    1
# ... with 45,990 more rows
```



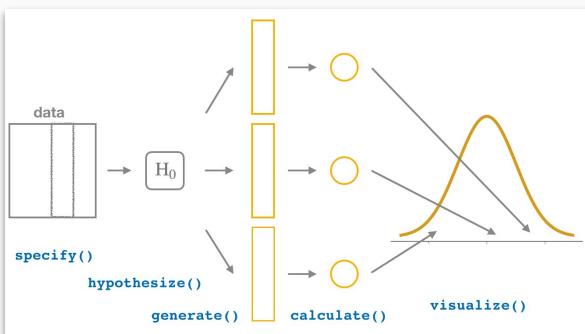
Infer Workflow



calculate()

Apply a summary statistic:

- “diff in means” for difference in means



```
> comparison_1_tbl %>%
+   # Infer
+   specify(prop ~ label) %>%
+   hypothesize(null = "independence") %>%
+   generate(reps = 1000, type = "permute") %>%
+   calculate(stat = "diff in means",
+             order = c("treatment-new_page", "control-old_page"))
# A tibble: 1,000 x 2
  replicate      stat
     <int>    <dbl>
1       1  0.000822
2       2 -0.00131
3       3 -0.000462
4       4  0.00197
5       5  0.0000738
6       6  0.00156
7       7  0.000606
8       8  0.00172
9       9  0.00173
10      10 -0.00000660
# ... with 990 more rows
```

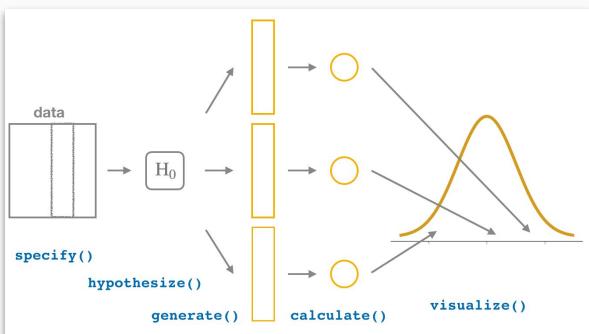


Infer Workflow

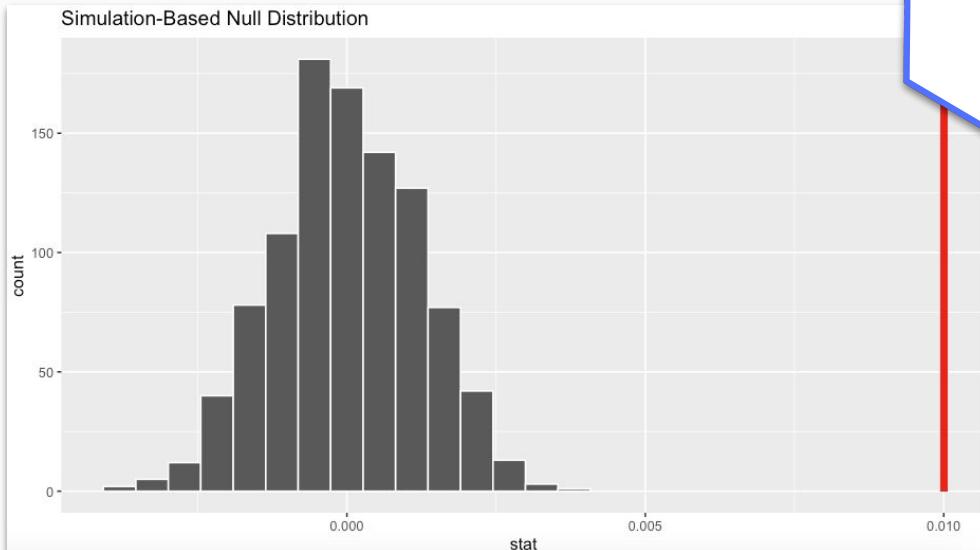


visualize()

- View the probability of a value more extreme



```
193 comp_1_diff_means_permuted_tbl %>%
194   visualize() +
195   shade_p_value(obs_stat = 0.01, direction = "right")
```



```
> comp_1_diff_means_permuted_tbl %>%
+   get_p_value(obs_stat = 0.01, direction = "right")
# A tibble: 1 x 1
  p_value
  <dbl>
1     0
```

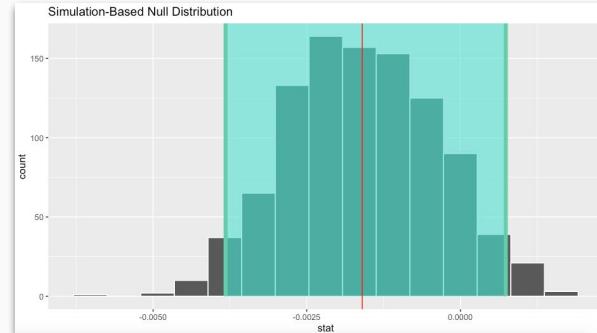
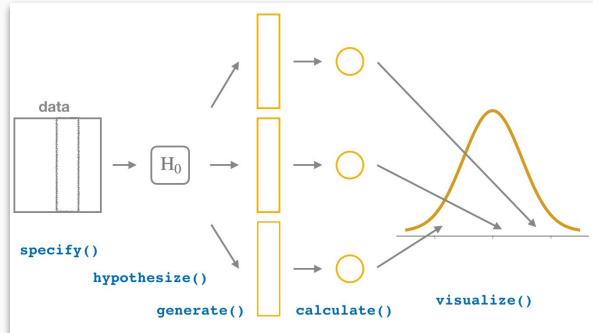




What about Confidence Intervals?

Bootstrap Workflow

Just skip `hypothesize()` and set `generate(type = "bootstrap")`



Infer Workflow

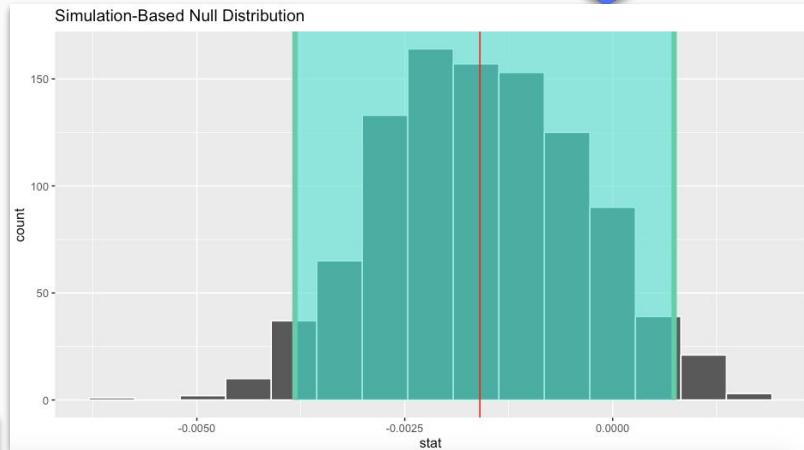


Bootstrap

- Skip `hypothesize()`
- Set `generate(type = "bootstrap")`

```
205 comp_1_diff_bootstrap_tbl <- comparison_1_tbl %>%
206   # Infer
207   specify(prop ~ label) %>%
208
209   # Remove Hypothesis:
210   # hypothesize(null = "independence") %>%
211
212   # Change generate(): type = "bootstrap"
213   generate(reps = 1000, type = "bootstrap") %>%
214     calculate(stat = "diff in means",
215       order = c("treatment-new_page", "control-old_page"))
```

```
> bootstrap_ci_tbl
# A tibble: 1 x 2
`2.5%` `97.5%
<dbl>    <dbl>
1 -0.00380 0.000720
```

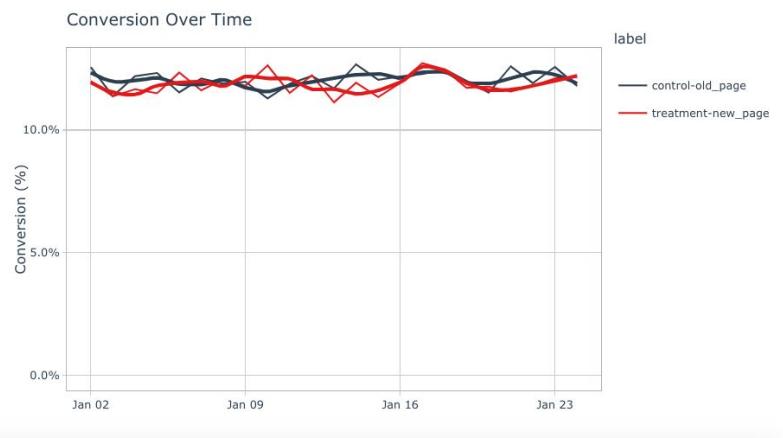


30-Min Demo

A/B Testing

PRO-TIPS

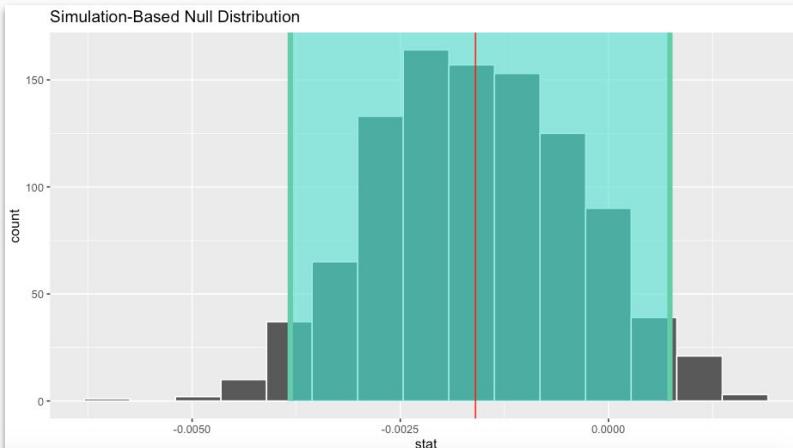
Yeahhhhhh!



#1. Visualize Over Time

If conversion aggregation by day, **run for at least 2 weeks**

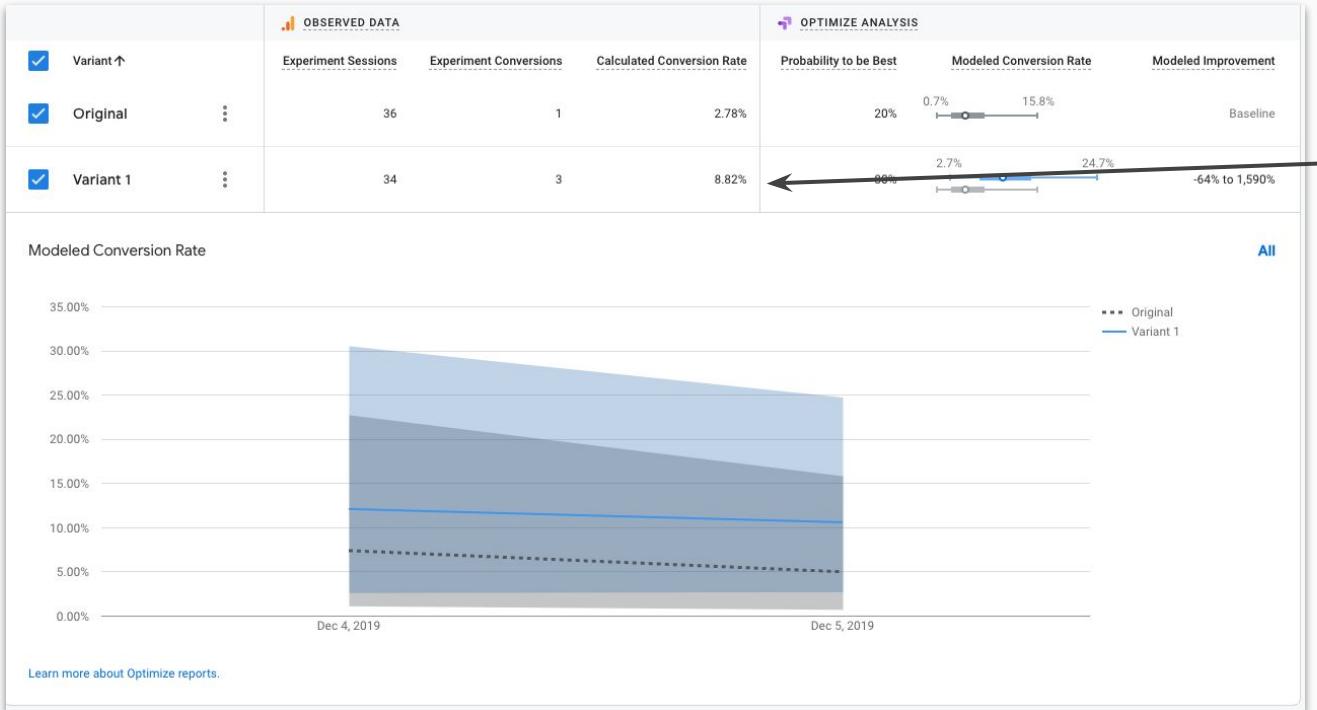
Don't just go off of mean: Outliers, changes in rates, can affect .



#2. Use Bootstrap for Confidence Intervals

Explain: 95% of time the range is between -0.38% and +0.01% with the mean of -0.16%.

Use Significance: Say with significance that we should or should not make the change.



Can I trust this number?

Absolutely not.

I haven't run for 2 weeks.

I don't know if this trend will persist.

What We Just Did

And how WE did it!



A/B Testing Workflow

Step-By-Step



Google Optimize

Perform A/B Testing

dplyr & ggplot2

A/B Test Data Wrangling and
Visualization

infer

Hypothesis Testing &
Confidence Interval Estimation

Google Optimize



Optimize

		OBSERVED DATA			OPTIMIZE ANALYSIS		
<input checked="" type="checkbox"/>	Variant ↑	Experiment Sessions	Experiment Conversions	Calculated Conversion Rate	Probability to be Best	Modeled Conversion Rate	Modeled Improvement
<input checked="" type="checkbox"/>	Original	36	1	2.78%	20%	0.7% - 15.8%	Baseline
<input checked="" type="checkbox"/>	Variant 1	34	3	8.82%	80%	2.7% - 24.7%	-64% to 1,590%

Modeled Conversion Rate All

The chart displays the modeled conversion rate for two variants from December 4, 2019, to December 5, 2019. The Y-axis represents the conversion rate in percentages, ranging from 0.00% to 35.00%. The X-axis shows the dates Dec 4, 2019, and Dec 5, 2019. The 'Original' variant (dashed line) starts at approximately 12.0% and remains relatively flat. The 'Variant 1' variant (solid blue line) starts at approximately 12.0% and shows a slight downward trend to about 10.5%. Shaded areas around the lines represent confidence intervals.

Learn more about Optimize reports.

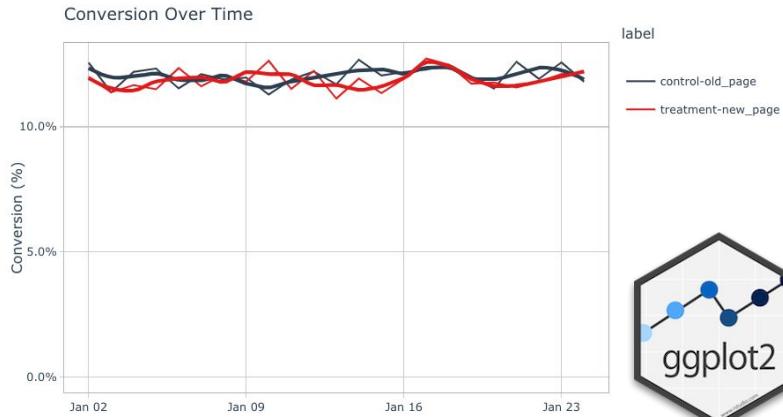
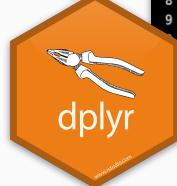


dplyr & ggplot2

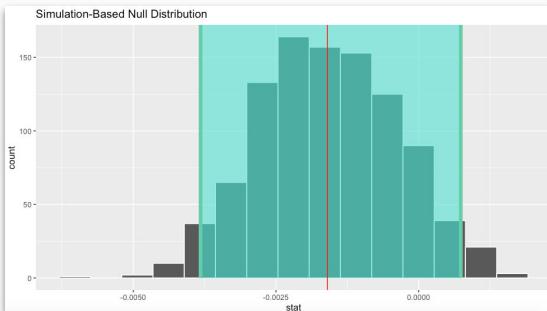
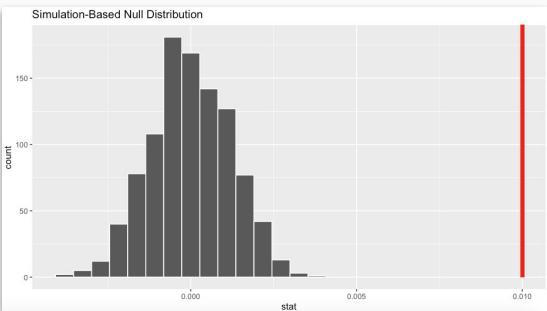
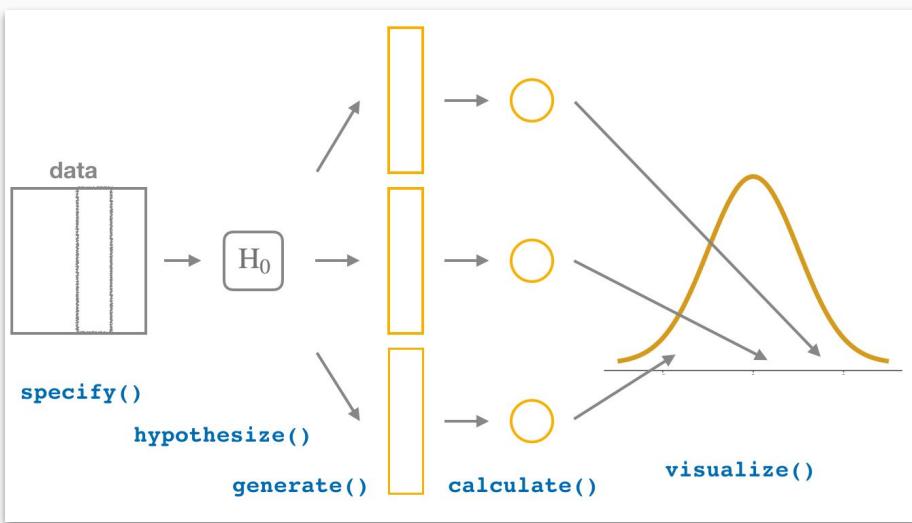
```
# A tibble: 290,585 x 7
  user_id timestamp      group landing_page converted label    visit_count
  <dbl>   <dttm>     <chr>   <chr>        <dbl> <chr>        <dbl>
1 922696 2017-01-02 13:42:05 treatment new_page 0 treatment-new_page 1
2 781507 2017-01-02 13:42:15 control  old_page 0 control-old_page 1
3 737319 2017-01-02 13:42:21 control  old_page 0 control-old_page 1
4 818377 2017-01-02 13:42:26 treatment new_page 0 treatment-new_page 1
5 725857 2017-01-02 13:42:27 treatment new_page 0 treatment-new_page 1
6 762651 2017-01-02 13:42:28 treatment new_page 0 treatment-new_page 1
7 722516 2017-01-02 13:42:34 treatment new_page 0 treatment-new_page 1
8 688937 2017-01-02 13:42:41 treatment new_page 0 treatment-new_page 1
9 781125 2017-01-02 13:42:45 treatment new_page 0 treatment-new_page 1
10 799109 2017-01-02 13:42:45 control  old_page 0 control-old_page 1
# ... with 290,575 more rows
```

With dplyr, you can easily filter, transform, and summarize data.

```
# A tibble: 46 x 5
  label      date       count conversion prop
  <chr>     <dttm>     <int>    <dbl> <dbl>
1 control-old_page 2017-01-02 00:00:00 2859  0.126
2 control-old_page 2017-01-03 00:00:00 6590  0.114
3 control-old_page 2017-01-04 00:00:00 6578  0.122
4 control-old_page 2017-01-05 00:00:00 6427  0.123
5 control-old_page 2017-01-06 00:00:00 6606  0.115
6 control-old_page 2017-01-07 00:00:00 6604  0.121
7 control-old_page 2017-01-08 00:00:00 6687  0.119
8 control-old_page 2017-01-09 00:00:00 6628  0.120
9 control-old_page 2017-01-10 00:00:00 6654  0.113
10 control-old_page 2017-01-11 00:00:00 6688  0.119
# ... with 36 more rows
```



Infer Workflow



A lot more to learn... Advanced ML

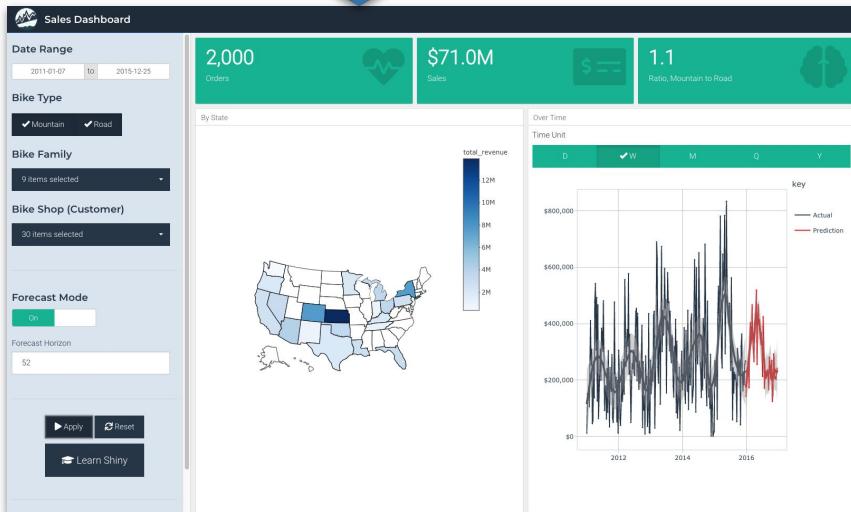


```
> h2o.predict(h2o_model, newdata = as.h2o(credit_card_group_tbl)) %>%
+   as_tibble()
# A tibble: 1,125 x 7
  predict      p1      p2      p3      p4      p5 Other
  <fct>    <dbl>    <dbl>    <dbl>    <dbl>    <dbl>    <dbl>
1 Other     0.0000704 0.0228     0       0     0.977
2 Other    0.0232    0.0000717 0.0000553 0       0.00376 0.973
3 Other     0       0.0000737 0.0238     0       0.000107 0.976
4 Other    0.00643   0.0000724 0.0000558 0.00343  0.000105 0.990
5 Other     0       0.0000720 0.0000555 0       0.000104 1.000
6 3        0       0.0000704 0.909      0       0.000102 0.0909
7 3        0       0.0000761 0.995      0       0.000110 0.00491
8 1        0.984    0.0000735 0.0000567 0.00349  0.000106 0.0127
9 Other    0.195    0.0000602 0.0000464 0.00285  0.0000870 0.802
10 Other     0       0.0000737 0.0000568 0       0       1.000
# ... with 1,115 more rows
```

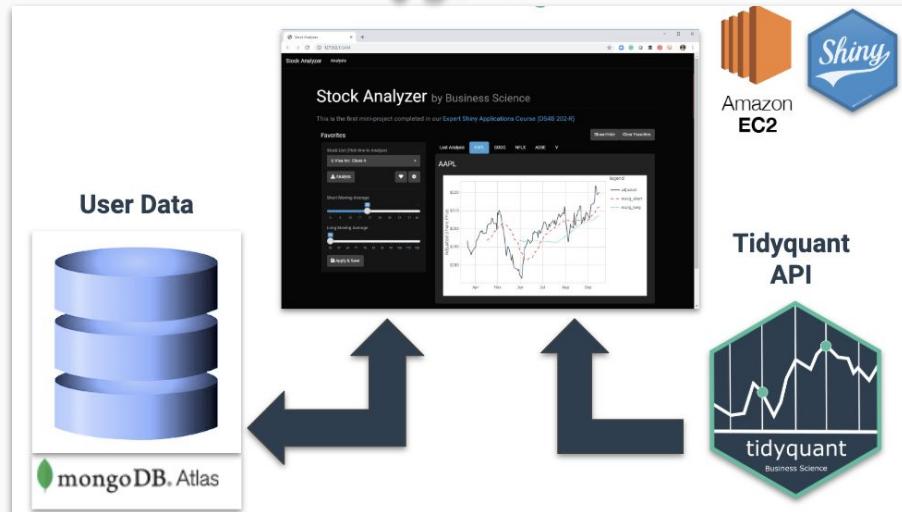
The image shows a tablet displaying the H2O AutoML interface. The screen displays four performance plots: ROC, Precision vs Recall, Gain, and Lift. Below these plots, a message box contains the text "H2O AutoML". The background of the tablet screen is white.

```
1 #!/usr/bin/env python
2
3 import sys
4 from h2o.grid.grid_search import H2OGridSearch
5 from h2o.estimators import H2OAutoML
6 from h2o import H2OContext
7
8 H2OContext.init(log_level="INFO")
9
10 automl = H2OAutoML(max_runtime_secs=3600)
11 automl.train(x=[1, 2, 3, 4, 5], y="y", training_frame=h2o_train, seed=123456789)
12
13 print(automl.gridsearch_leaderboard())
14
15 print(automl.leaderboard().sort("auc").model_ids[-1])
16
17 print(automl.leaderboard().sort("auc").model_ids[-2])
```

A lot more to learn... Web Apps



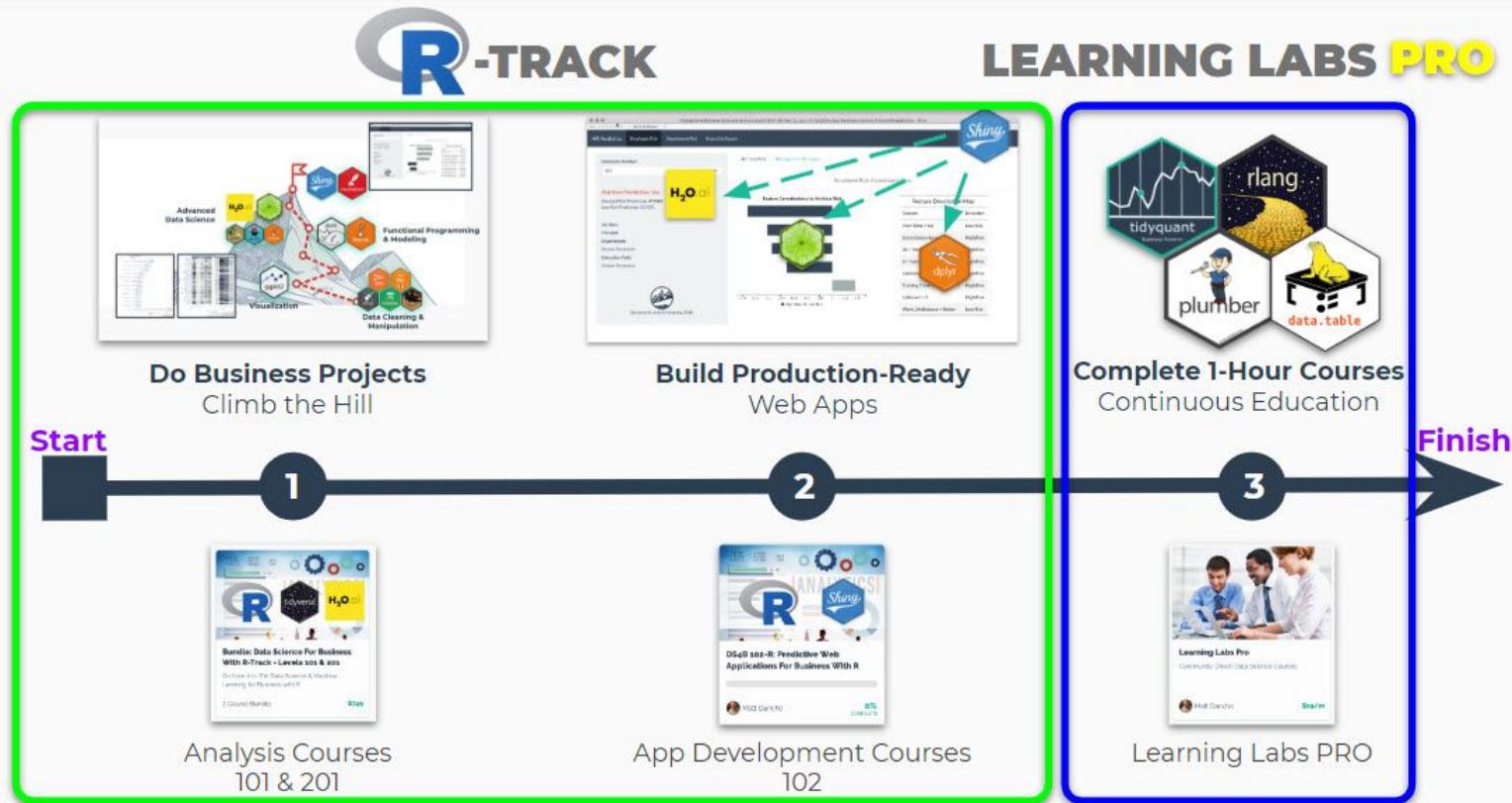
102 & 202A



Business Science University

Learn Data Science for Business in 6-Months

The program that will deliver YOUR Transformation



4-Course R-Track System



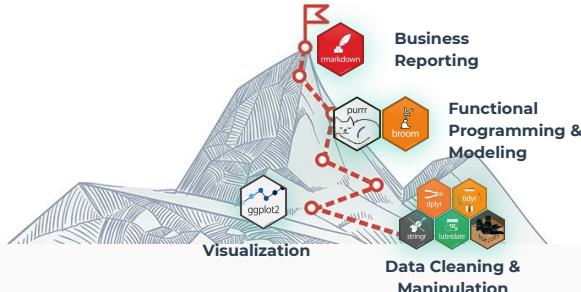
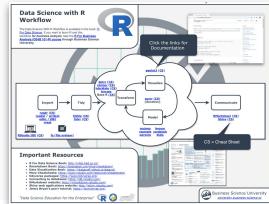
Business Analysis with R (DS4B 101-R)

Data Science For Business with R (DS4B 201-R)

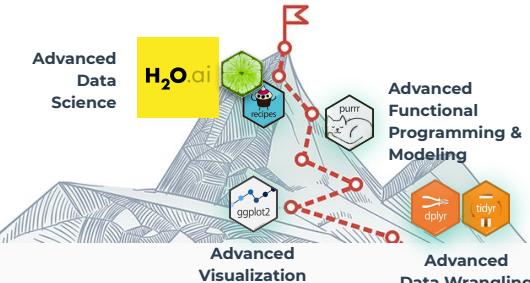
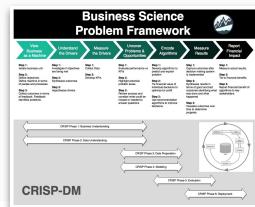
Web Apps & Shiny Developer (DS4B 102-R + DS4B 202A-R)

Project-Based Courses with Business Application

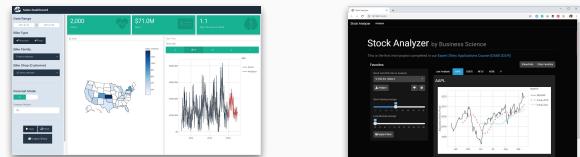
Data Science Foundations
7 Weeks



Machine Learning & Business Consulting
10 Weeks



Web Application Development
12 Weeks

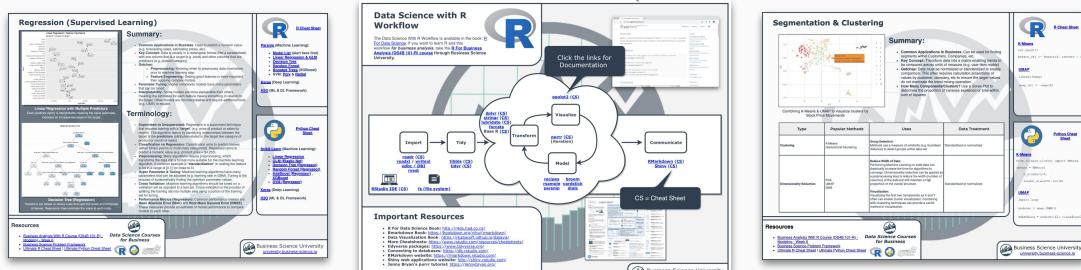


Key Benefits

- Fundamentals - Weeks 1-5 (25 hours of Video Lessons)
 - Data Manipulation (dplyr)
 - Time series (lubridate)
 - Text (stringr)
 - Categorical (forcats)
 - Visualization (ggplot2)
 - Programming & Iteration (purrr)
 - 3 Challenges
- **Machine Learning - Week 6 (8 hours of Video Lessons)**
 - Clustering (3 hours)
 - Regression (5 hours)
 - 2 Challenges
- Learn Business Reporting - Week 7
 - RMarkdown & plotly
 - 2 Project Reports:
 1. Product Pricing Algo
 2. Customer Segmentation

Business Analysis with R (DS4B 101-R)

Data Science Foundations
7 Weeks



Key Benefits

End-to-End Churn Project

Understanding the Problem & Preparing Data - Weeks 1-4

- Project Setup & Framework
- Business Understanding / Sizing Problem
- Tidy Evaluation - rlang
- EDA - Exploring Data -GGally, skimr
- Data Preparation - recipes
- Correlation Analysis
- 3 Challenges

Machine Learning - Weeks 5, 6, 7

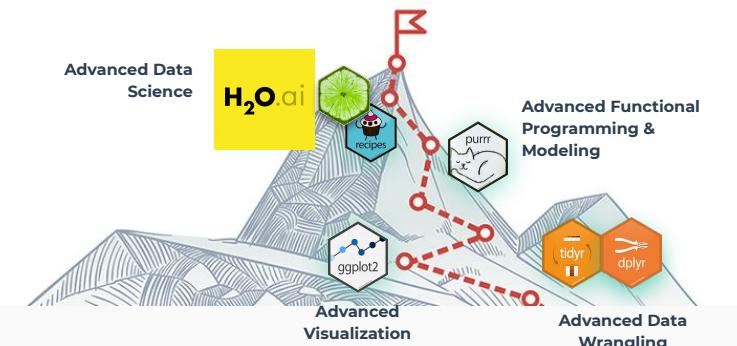
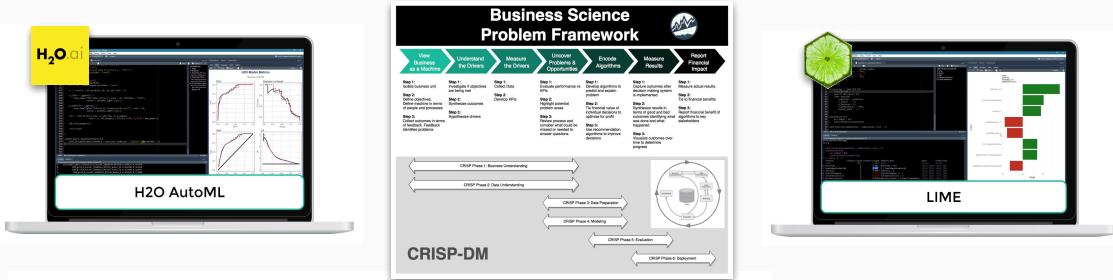
- H2O AutoML - Modeling Churn
- ML Performance
- LIME Feature Explanation

Return-On-Investment - Weeks 7, 8, 9

- Expected Value Framework
- Threshold Optimization
- Sensitivity Analysis
- Recommendation Algorithm

Data Science For Business (DS4B 201-R)

Machine Learning & Business Consulting
10 Weeks



Key Benefits

Learn Shiny & Flexdashboard

- Build Applications
- Learn Reactive Programming
- Integrate Machine Learning

App #1: Predictive Pricing App

- Model Product Portfolio
- XGBoost Pricing Prediction
- Generate new products instantly

App #2: Sales Dashboard with Demand Forecasting

- Model Demand History
- Segment Forecasts by Product & Customer
- XGBoost Time Series Forecast
- Generate new forecasts instantly

Shiny Apps for Business (DS4B 102-R)



Web Application Development
4 Weeks

The image contains three separate screenshots from the DS4B 102-R course. The first screenshot shows a 'Sales Dashboard' with a map of the US, a bar chart for 'Blue Type' (2,000), a line chart for 'Revenue' (\$71.0M), and a scatter plot for 'Profit' (1.1). The second screenshot is a flowchart titled 'Data Science with R' showing the process from 'Start' to 'Publish' through 'Data Cleaning', 'Model Selection', 'Advanced Forecasting', 'Testing', and 'Deployment'. The third screenshot shows a 'Shiny App' interface with sections for 'Flexdashboard Apps' and 'Shiny Apps', and links to 'Themes, Dashboards, & Examples'.



ANALYTICS! DS4B 102-R: Shiny Web Applications For Business (Level 1)
Build a predictive web application using Shiny, Flexdashboard, and XGBoost.
Matt Dancho

Key Benefits

Frontend + Backend + Production Deployment

Frontend for Shiny

- Bootstrap

Backend for Shiny

- MongoDB
- Dynamic UI
- User Authentication
- Store & Write User Data

Production Deployment

- AWS
- EC2 Server
- VPC Connection
- URL Routing

Shiny Apps for Business (DS4B 202A-R)



Web Application Development
6 Weeks





15% OFF PROMO Code: learninglabs

R-TRACK BUNDLE

4-Course Bundle - Machine Learning + Expert Web Applications (R-Track)

Go from Beginner to Expert Data Scientist & Shiny Developer in Under 6-Months

4 Course Bundle ~~\$1,500~~

DS4B 101-R: Business Analysis With R

Your Data Science Journey Starts Now! Learn the fundamentals of data science for business with the tidyverse.

Matt Dancho

DS4B 102-R: Shiny Web Applications For Business (Level 1)

Build a predictive web application using Shiny, Flexdashboard, and XGBoost.

Matt Dancho

DS4B 201-R: Data Science For Business With R

Solve a real-world churn problem with H2O AutoML (automated machine learning) & LIME black-box model explanations using R.

Matt Dancho

DS4B 202A-R: Expert Shiny Developer with AWS

Learn how to build Scalable Data Science Applications using R, Shiny, and AWS Cloud Technology.

Matt Dancho

<input type="radio"/>	Paid Course 15% COUPON DISCOUNT	\$1,596 \$2,356.60
<input checked="" type="radio"/>	12 Low Monthly Payments 15% COUPON DISCOUNT	12 payments of \$149/m 12 payments of \$126.65/m

Begin Learning Today

university.business-science.io

