Building it Beautiful

Analyzing the Effectiveness of Products at Scale

Josh Laurito
Strata NYC 2018

(Text to go with slides at blog.joshlaurito.com)
Keeping it Simple
Squarespace

Company mission:

Squarespace empowers people with creative ideas to succeed
From websites and online stores, to marketing tools and analytics, Squarespace is the all-in-one platform to build a beautiful online presence.

START A FREE TRIAL

No credit card required
Control the web.

Squarespace is a powerful new way to create and manage your web site (blog included).

Try Squarespace for Free
One step signup. Name not required.

What is Squarespace?

Squarespace is the universal answer to the question: How do I put my ___ online?
They’re all covered. We let you snap together a web site in a flash that will allow you to manage all of your unique content right in one place. It can be as professional or personal as you want.

Who’s Squarespace for?

Bloggers. Independent professionals.
Small businesses. Anyone who needs to maintain a web presence, but wants exacoting control over their site, and powerful publishing features that cover everything from blogs to files. No technical skill is required. You provide the ideas, we turn those into your site. It’s that easy. You can get started right away, or learn more.

2005

SMARter Publishing

Squarespace is the best way to put your website and blog on the web.

What is Squarespace?

Squarespace is an exceptionally tuned publishing system for managing websites and blogs. Need to create a blog? Build a website? Our software helps discerning individuals build stylish, easy to manage websites complete with state-of-the-art blogging tools. Squarespace hosts thousands of high-quality websites — both large and small.

Who's Squarespace for?

Bloggers, Authors, Teachers, Lawyers, Doctors, Musicians, Small Businesses. Your web presence is a reflection of who you are — and Squarespace makes sure you look good doing what you love. We're the publishing choice for anyone who is sick of bargain bin services and is ready for an elite solution to their publishing needs. No technical skill is required.

Try Squarespace for Free

One step signup. Name not required.

“This is the kind of software the internet has been crying out for.”
— The Wall Street Journal Online

“This is the best content management system I have ever seen within its class.”
— Gadgetopia

2007

The secret behind exceptional websites.

What is it?
A fully hosted, completely managed environment for creating and maintaining a beautiful website, blog, or portfolio.

Who is it for?

Why use it?
We give you full control over your content and site customization. You can build a website or create a blog 10X faster on our comprehensive platform.
SET YOUR WEBSITE APART

Create a beautiful website. Start your free trial today. No credit card required.

GET STARTED
User Journeys: A/B Testing

Version A

100 Landings

50 Trials

25 Signups
User Journeys: A/B Testing

Version A
- 100 Landings
- 50 Trials
- 25 Signups

Version B
- 100 Landings
- 70 Trials
- 35 Signups
User Journeys: A/B Testing

Version A
- 100 Landings
- 50 Trials
- 25 Signups

Version B
- 100 Landings
- 70 Trials
- 35 Signups
User Journeys: A/B Testing

**Version A**
- 100 Landings
- 50 Trials
- 25 Signups
- 20 Retained

**Version B**
- 100 Landings
- 70 Trials
- 35 Signups
- 25 Retained

*(not immediately observable)*
A New Product
Make it beautiful.
Make your own website.

START A FREE TRIAL
No credit card required.
Adding Domains - Some Considerations

- Complement to websites (in the economic sense)
- Rivalrous (websites are not)
- Payment processor implications
- New partnerships
- New homepage flow potentially cannibalizes website trials & signups
How Much More Complex?

- Potential States: 4 \(\rightarrow\) 8
  (theoretically scales at \(2^n\) if all products are totally independent)

- Potential Transitions: 3 \(\rightarrow\) 11
  (theoretically scales at \(n \times 2^n\) if all single-step transitions are legal)
How Much More Complex?

- This **vastly understates** the actual technical & state complexity
  - Transferring domains/websites in from other vendors
  - Transferring domains/websites out to other vendors
  - Differences in marketing messages and landing users
  - Customers who buy multiple sites
  - Transitions between plan types
  - Transitions between payment cadence (annual/monthly)
  - etc.

- Fortunately, A/B testing gives us some confidence these considerations will be wrapped into the total numbers
The Test

- Do we make more money with domain search on our homepage?
Major Challenges

- Tracking users through the different services
- Modeling difficult to observe values
Tracking
Tracking Users

- The homepage is managed by one engineering group
- The domain experience is managed by another
- Lifecycle events (eg. trialing, signups) are managed by a third
- The mobile versions of these processes are managed by yet another
- We need all of them to implement events in roughly the same way
We’re Not the First

The Unified Logging Infrastructure for Data Analytics at Twitter

George Lee, Jimmy Lin, Chuang Liu, Andrew Lorek, and Dmitriy Ryaboy
Twitter, Inc.
@GeorgeJLee @lintool @chuangl4 @mrtall @squarecog

Scalable Schema Management for Hadoop and Spark Applications

Kelvin Chu, Evan Richards
Signs You Might Need An Event Schema

- You have your own homegrown eventing system
- Your company is somewhere on the SOA - Microservices spectrum
- You have multiple event logs, or an event log that is sparse & more than 100 columns wide
- No one on your team knows the name of the person implementing the events!
High Level Ideas

- Single, unified logging architecture
- Split between ‘common’ & ‘custom’ schemas
- Events validated against versioned schemas on write
Structure

Data Producers

Analytics Service → Kafka → HDFS → Airflow → PostgreSQL

Hive → Presto
Unified Logging

- Maybe obvious but deserves to said: have services record data using the same endpoint and going to the same place

- Tough questions:
  - Where do networking/SRE monitoring events go?
  - How do you maintain continuity with older system?
Unified Log Table Design

- Wide & Sparse

- Nested

- Multiples

<table>
<thead>
<tr>
<th>User</th>
<th>Service</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Domain</td>
</tr>
<tr>
<td>2</td>
<td>Domain</td>
</tr>
<tr>
<td>3</td>
<td>Domain</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>User</th>
<th>Service</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>Mobile</td>
</tr>
<tr>
<td>4</td>
<td>Mobile</td>
</tr>
<tr>
<td>5</td>
<td>Mobile</td>
</tr>
</tbody>
</table>
Split Schemas

- **Common schema**: universal data, usually generated from global variables or enrichment process
  - Changes rarely

- **Custom schema**: data important for a specific producer/service
  - Changes more frequently

```plaintext
<table>
<thead>
<tr>
<th>Key</th>
<th>User</th>
<th>ts</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A</td>
<td>123</td>
</tr>
<tr>
<td>2</td>
<td>B</td>
<td>123</td>
</tr>
<tr>
<td>3</td>
<td>C</td>
<td>123</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Key</th>
<th>Homepage</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>Grey</td>
</tr>
<tr>
<td>4</td>
<td>Olive</td>
</tr>
<tr>
<td>5</td>
<td>Black</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Key</th>
<th>Domain</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>Click</td>
</tr>
<tr>
<td>6</td>
<td>Search</td>
</tr>
<tr>
<td>7</td>
<td>Cancel</td>
</tr>
</tbody>
</table>
```
Event Validation

- Schema service set up (Avro, including schema registry)
- Server responds with 202 for valid requests
- 400 for schema violations
- In development environment, more detailed error messages
Structure

- Data Producers
- Kafka
- HDFS
- Airflow
- PostgreSQL
- Hive
- Presto
- Analytics Service
- Schema Service
(Hopefully) Future Structure

Data Producers

Analytics Service

Kafka

HDFS

Airflow

PostgreSQL

Hive

Presto
/** Enum of types of direct/indirect objects. */
enum ObjectType {
  banner,
  button,
  carousel,
  currency,
  domain,
  form,
  language,
  link,
Modeling
Modeling Values

100 Landings

40 Website Trials

20 Website Signups

10 Domain Signups

10 Combined Signups

13 Websites Retained

8 Combined Retained

4 Domains Retained

not immediately observable

not immediately observable
Modeling Values

Domain Purchase to Website Sub Attach Curves

Days from Domain Purchase
Modeling Values

- 100 Landings
- 40 Website Trials
- 20 Website Signups
- 10 Domain Signups
- 10 Combined Signups
- 13 Websites Retained
- 8 Combined Retained
- 4 Domains Retained

not immediately observable
not immediately observable
Convert Into Point Estimates of User LTV

- 100 Landings
  - $3,000
  - $30 / Landing

- 40 Website Trials
  - $2,500

- Website Signups
  - $2,100

- Domain Signups
  - $900

- Combined Signups
  - $2,500

- Websites Retained
  - $1,600

- Combined Retained
  - $1,500

- Domains Retained
  - $400

- not immediately observable
- not immediately observable
Modeling Values

- 100 Landings
- 40 Website Trials
- 20 Website Signups
- 10 Domain Signups
- 10 Combined Signups
- 13 Websites Retained
- 8 Combined Retained
- 4 Domains Retained

Not immediately observable

- Combined Retained
- Combined Signups
- Website Signups
- Website Trials
- Landings

Not immediately observable
Numbers inside existing models

- Attribution Model
  - Expected likelihood to convert of different traffic
  - Output format: as credit to attribution channel, **not conversion** likelihood of a visitor

- Lifetime Value Model
  - Expected likelihood of upsell
  - Expected likelihood of retention
  - Output format: value distribution, **not conversions across individual states**
  - Not sure that will stay constant in our new setup
Solution: Decomposing Models

- Break down large models into component parts
- Easier to switch concepts/calculations & keep models consistent with one another
- Explicit propagation of error in more complex inferences
- Problems: proliferation of models means each model gets less attention and may be more difficult to discover
Decomposing Models

**Attribution Model**
How much it costs to acquire a customer via a channel

**Lifetime Value Model**
How much money we expect to earn from a customer
Decomposing Models

- If (Net) Lifetime Value > Cost Acquisition, spend more on that channel!
Decomposing Models

**Attribution Model**
How much it costs to acquire a customer via a channel

**Lifetime Value Model**
How much money we expect to earn from a customer

**Traffic Generation**
How much traffic a channel generates

**Conversion Rate**
How likely that traffic is to convert

**Retention Rate**
How likely those users are to stay with us

** Upsell Rate**
How likely those users are to add additional products
Decomposing Models

**Attribution Model**
How much it costs to acquire a customer via a channel

**Lifetime Value Model**
How much money we expect to earn from a customer

**Traffic Generation**
How much traffic a channel generates

**Conversion Rate**
How likely that traffic is to convert

**Retention Rate**
How likely those users are to stay with us

**Upsell Rate**
How likely those users are to add additional products
The Test

- Do we make more money with domain search on our homepage?

No.
How Can We Make This Simpler?
Things We’re Working On

- More investment in infrastructure
- Modular design of models & building knowledge base
- Better understanding of how people build sites, models of early success on the platform
- We’re hiring! squarespace.com/careers
The End

jlaurito@squarespace.com
@joshlaurito