Building with Data: Lessons from Etsy

Strata NY 2014

Nellwyn Thomas
Director of Analytics, Etsy
nellwyn@etsy.com
@nellwyn
Etsy
Shop directly from people around the world.

Recent Favorites
Discover finds from around the marketplace.
Over 25M items from 1M shops in 150 countries.

big community
$1.35B sales in 2013

big community + lots of actions
150GB weblogs daily

big community + lots of actions = lots of data
hadoop cluster “etsydoop” vertica warehouse

big community + lots of actions = lots of data
+ great stack
30 deploys by 200+ people per day

big community + lots of actions = lots of data
+ great stack + continuous delivery
dozens of experiments running simultaneously
big community + lots of actions = lots of data
+ great stack + continuous delivery +
continuous experimentation
today

what data we collect
who uses it
who makes it useful
how it works
and why
the users

product development

member operations

marketing

finance
the users

product development
ranking, relevancy, recommendation
opportunity identification, impact analysis

member operations
queue prioritization, customer experience

marketing
return on investment analysis, customer lifetime value

finance
forecasting, strategy
types of data

clickstream

transactional

survey

interview / usability study / market research

Etsy
types of data

clickstream
feature usage, item preference, search terms, device usage, visit frequency, referral source

transactional
order value, category of sale, purchase frequency, currency paid, shipping preferences

survey
attitudinal, net promoter score, brand awareness, customer service experience, aspirations

interview / usability study / market research
individual experience, industry trends
the stack

desktop & mobile website
/apps
emails

Etsydoop (HDFS)
(event logs
visit logs

bcn.etsy.com

Production Databases

Vertica

A/B Analyzer

SQL Builder

Dashboards

scalding skills

SQL skills

everyone!
the data makers

data-engineering & hadoop team

data science team

analytics
the data makers

data-engineering & hadoop team
ETL, visit serialization, hadoop, vertica, front-end tooling

data science team
personalization algorithms and machine learning models

analytics
opportunity, feature, impact analysis, a/b testing
the data makers

data–engineering & hadoop team
event logging, ETL, serialization, storage, front-end tooling

data science team
personalization algorithms and machine learning models

analytics
opportunity, feature, impact analysis, a/b testing

user research
why three teams?

Etsy

raymondbiesinger

Like this item?
Add it to your favorites to revisit it later.

New York City American Spaces 17x10.75" Art Print by Raymond Biesinger

$27.38 USD

Overview
- Handmade item
- Materials: Epson Ultrachrome K3 Ink, Epson Ultra Premium Matte paper
- Made to order
- Feedback: 100 reviews
- Ships worldwide from Montreal, Canada

This shop accepts Etsy Gift Cards

Add to Cart

Item Details

Shipping & Policies

This was originally made as part of a set of concept illustrations for a 2011 Toyota Avalon ad campaign, and it's meant to show the New York City complete with the Statue of Liberty, Brooklyn Bridge, and a few things in between. What I made would eventually inspire and guide a full package of billboards, magazine ads, web things and television commercials for the US market, and if you search mightily you just might be able to find a few of the ads online.

Printed in an open edition via Epson Stylus Pro 3880 at 17x10.75" on Epson Ultra Premium 192 GSM Enhanced Matte paper. Dimensions include a 1.5" margin. Signed and dated.
why three teams?

Data Science

Analytics

Data-Eng
the users

product development
ranking, relevancy, recommendation
opportunity identification, impact analysis

member operations
queue prioritization, customer experience

marketing
return on investment analysis, customer lifetime value

finance
forecasting, strategy
data analyst organization

product analytics

member operations

marketing analytics

finance insights

Etsy
data analyst organization
sample project team

- reports into eng
- reports into each separate group

- Tech Lead
- Eng
- Eng
- Product Manager
- Designer
- Analyst
- PMM
- Research

- dedicated to a single effort
- works on multiple efforts
the skill set

analytical skill
to understand the problem. and the opportunity.

math/stats skill
to understand the data.

technical skill
to get the data, parse it, and visualize it.

communication skill
to communicate what matters and
and to not communicate what doesn’t.
today

what data we collect
who uses it
who makes it useful
how it works
and some takeaways
data is not a single team
data needs interpreters
technical skill = analytical skill
what we’re really talking about:
better inputs for decisions
more goodness

Steve Mardenfeld on Experiments
http://www.slideshare.net/smardenfeld

Dan McKinley on Experiments
http://mcfunley.com/design-for-continuous-experimentation

Decided by Data: Case Studies from a Data Driven Product Culture
Strata, Thursday at 5:05pm
Learn more
etsy.com/careers
github.com/etsy

Nellwyn Thomas
nellwyn@etsy.com
@nellwyn