

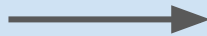
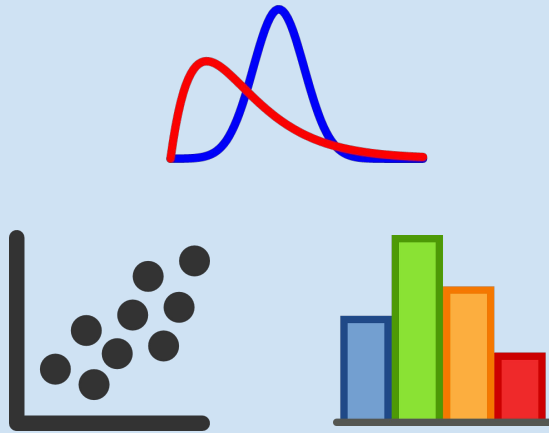


Outbound

Straightforward Insights for Growing Apps

Ben Regner

Problem: Present data-driven insights in a format digestible by non-specialists.



Users are **5x** more likely to perform a “**Checkout**” when...

As an illustrative example, we'll look at anonymized data from “Larry’s Lullabies”

Examples of the events specified by “Larry’s Lullabies” include:

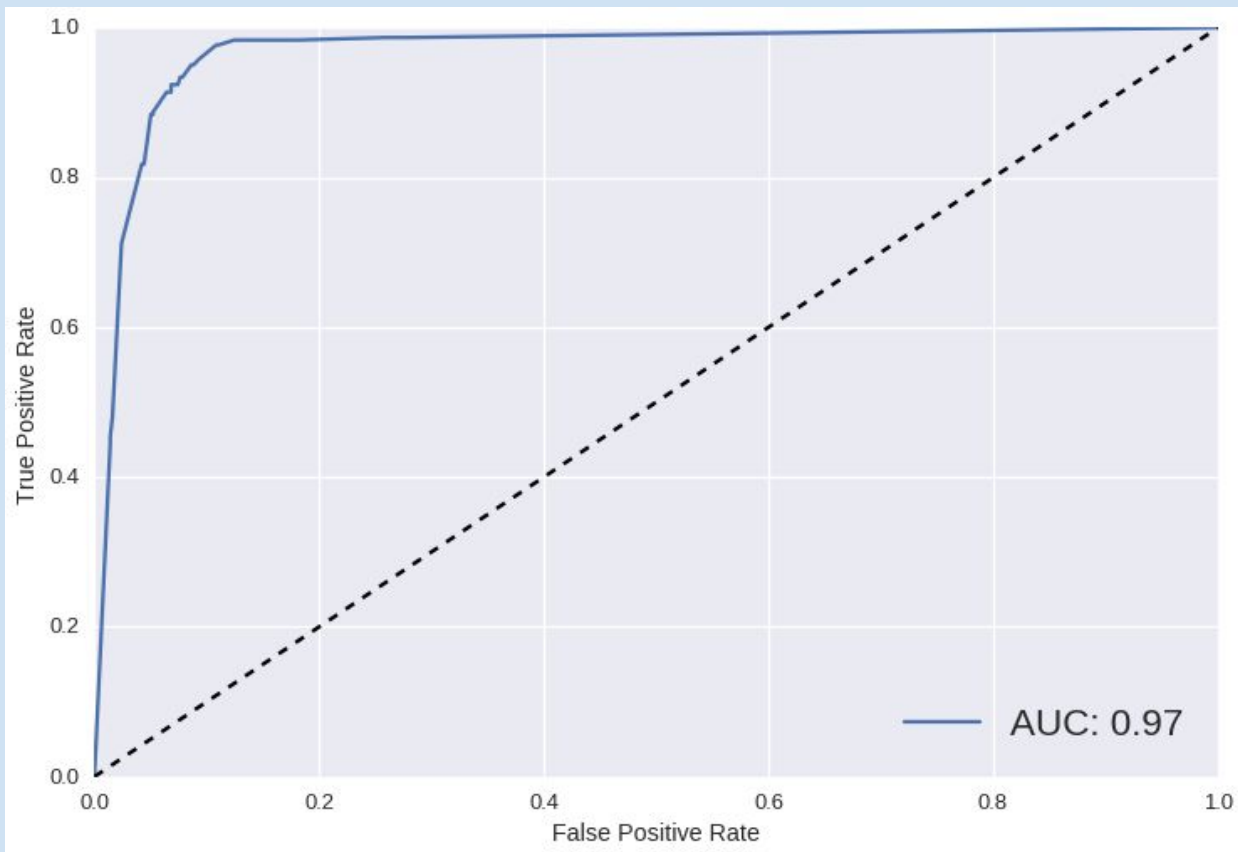
**Checkout, Cart Add, Cart Remove,
Page Viewed, and Signed Up**

As an illustrative example, we'll look at anonymized data from “Larry’s Lullabies”

Examples of the events specified by “Larry’s Lullabies” include:

**Checkout, Cart Add, Cart Remove,
Page Viewed, and Signed Up**

Predicting Whether a User Will “Checkout”



Predicting Whether a User Will “Checkout”

A statistical tool called a random forest classifier takes user events to predict users who will **Checkout** with 92.6% accuracy. The events that are the most important to distinguish users are:

**Cart Add, Page View, Cart Remove,
Product Details and Filter.**

Predicting Whether a User Will “Checkout”

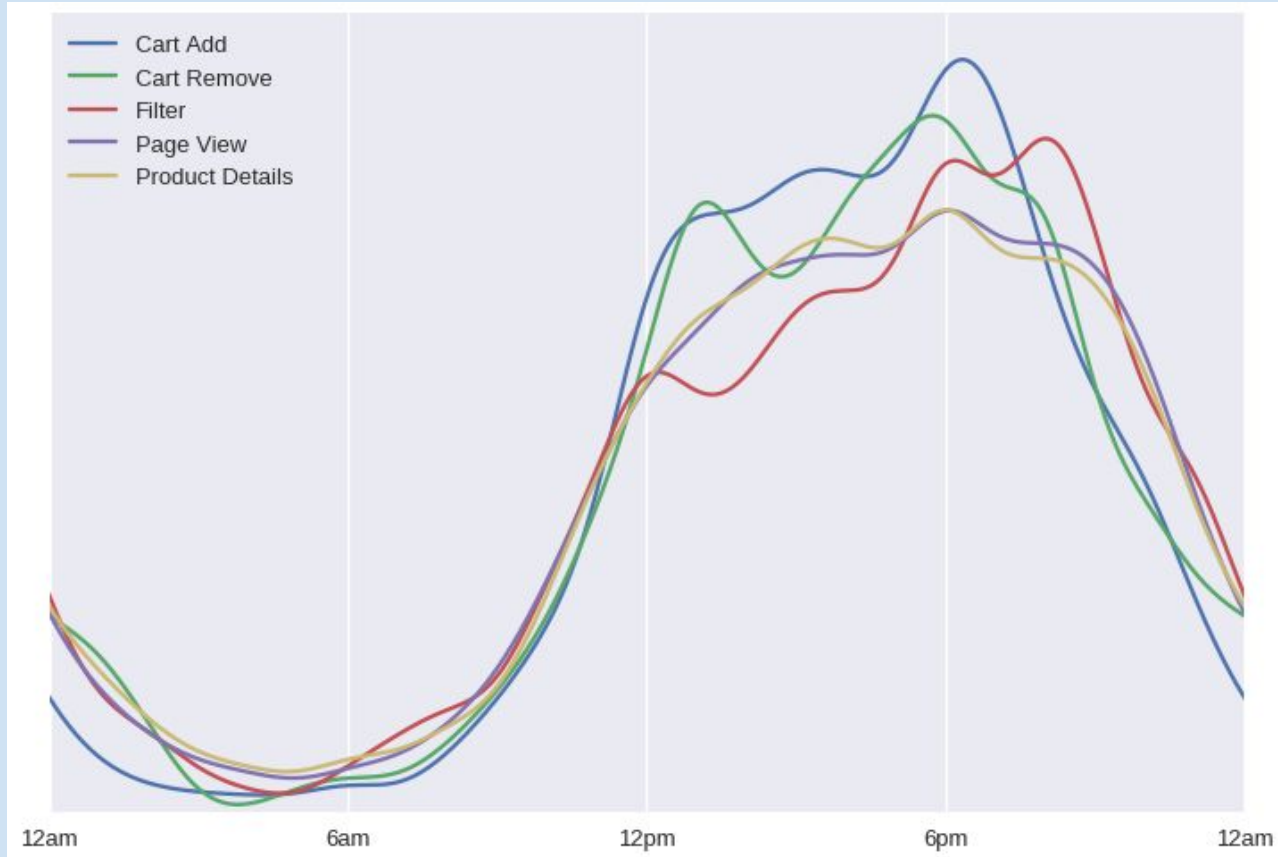
Users who **Cart Add** are 23.9x *more* likely to "Checkout".

Users who **Checkout** have always done "Page View".

Users who **Cart Remove** are 2x *less* likely to "Checkout".

Users who **Filter** are 1.25x *less* likely to **Checkout**.

Timing of Daily Events



Timing of Daily Events

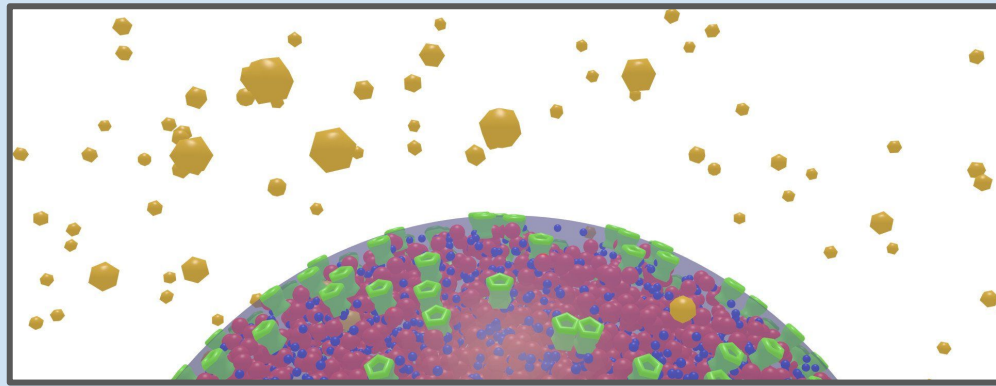
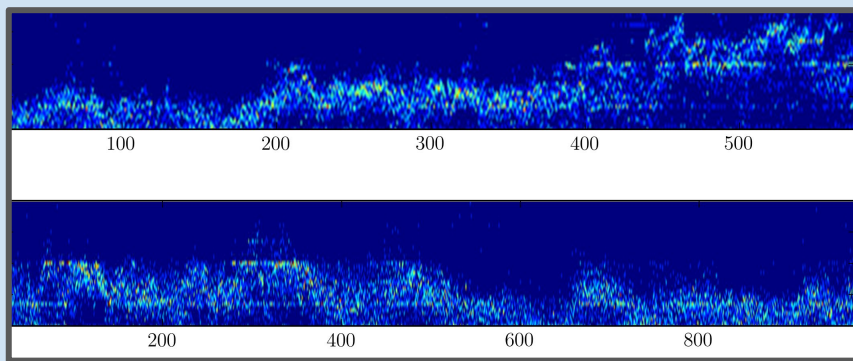
Users are mostly likely to perform...

Cart Remove at about 5pm

Cart Add at about 6pm

Product Details at about 6pm

Filter at about 10pm

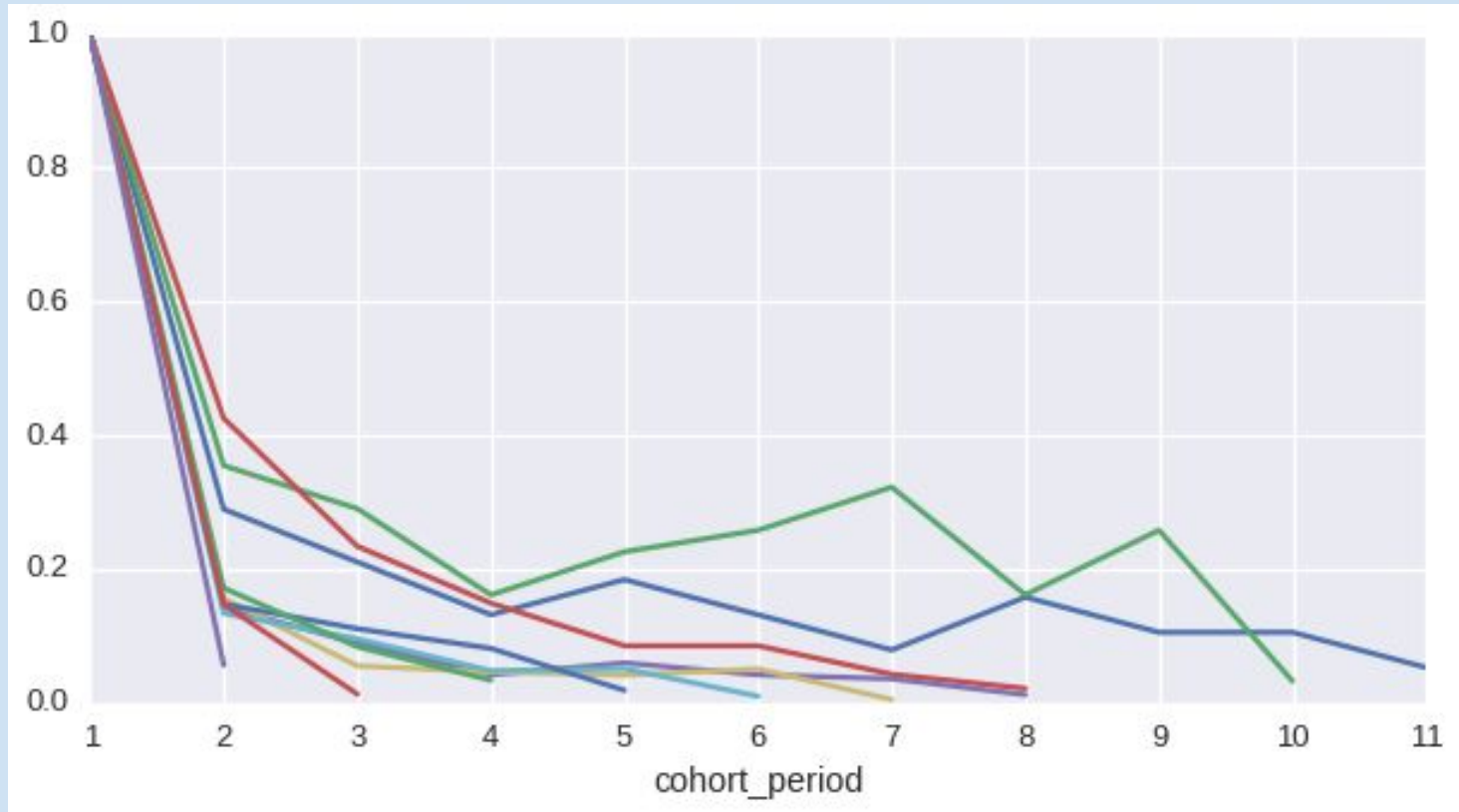


Ben Regner

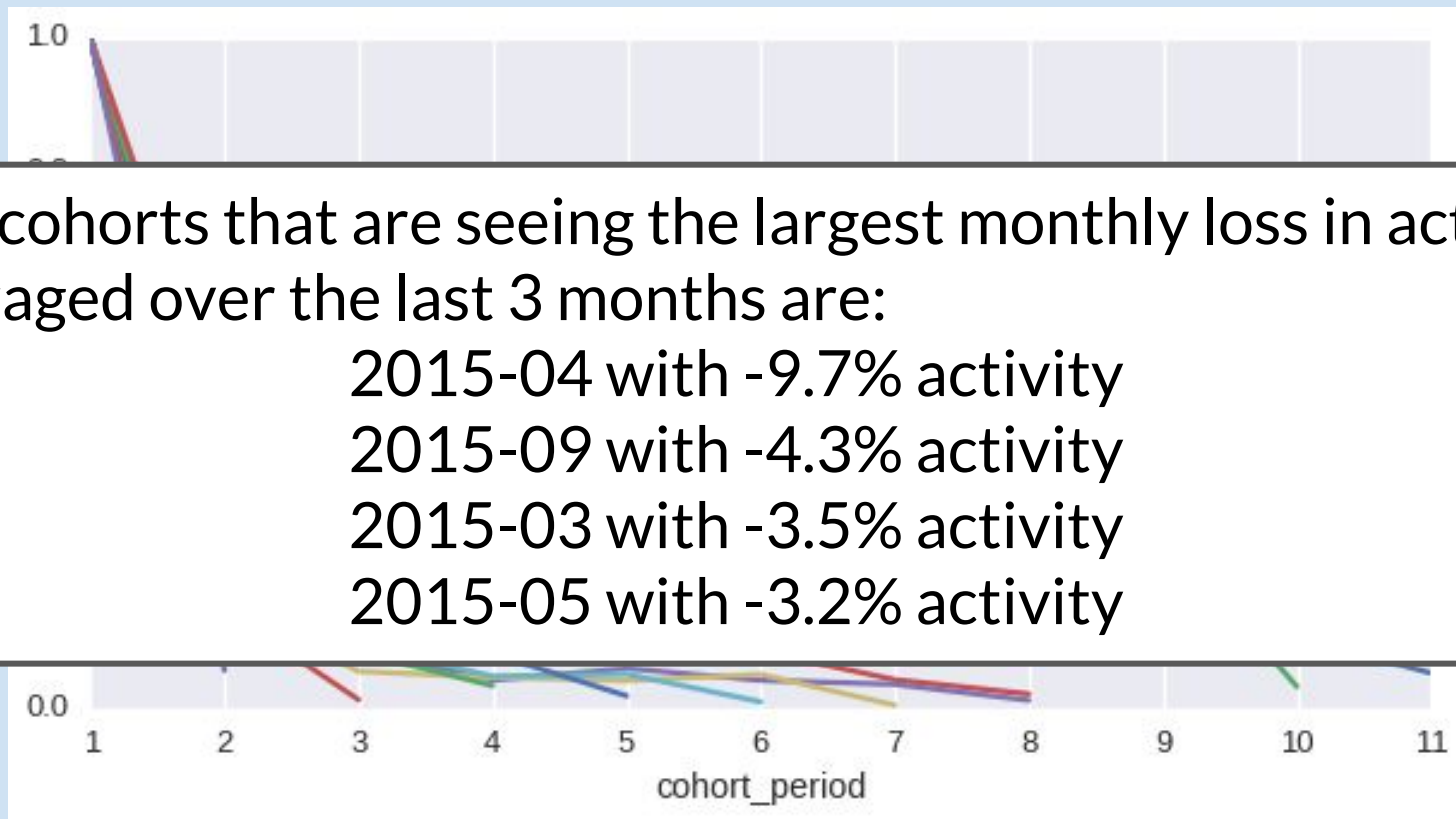
UC San Diego



Cohort Analysis



Cohort Analysis



The cohorts that are seeing the largest monthly loss in activity averaged over the last 3 months are:

2015-04 with -9.7% activity

2015-09 with -4.3% activity

2015-03 with -3.5% activity

2015-05 with -3.2% activity

Algorithms

Random Forest Classifier — scalability, handles irrelevant inputs, interpretable

Kernel Density Estimation — clear visualization of noisy timestamp information

Conditional Probability — suggests important dependencies between events

Cohort Analysis — provides context for previous business decisions