Personalized Customer Experience through Data Science

About Me

Digital Data Science &
Analytics professional
Ex-@Expedia, @Adobe

http://linkedin.com/in/rohitgossain
http://github.com/rohitgossain

Agenda

* Discussion
  * Digital Analytics Journey
  * Stack to build
  * Why Data Science
  * Common Examples
* Case Studies with R Code
  * Predicted Personalized Notification improving Mobile Conversion Rate
  * Understanding the impact of new feature release on business outcome

Stack to build

Data Capturing Layer

Enrichment of data from various sources

Building Visualization Layer for Data Exploration

Machine Learning or Predictive Modeling Layer

Integrating Model Output into Application using API (Integrated Experience)

Why Data Science?

Prescriptive Personalization

OR

Predictive Personalization

Segment Targeting

OR

Individual Targeting

Common Examples

* Recommendation Engine (You may like)
* Pop-ups Triggered by visitor's actions (Good Experience vs Bad Experience)
* Personalized Email offers
* Personalized User Experience (Prescriptive or Predictive)

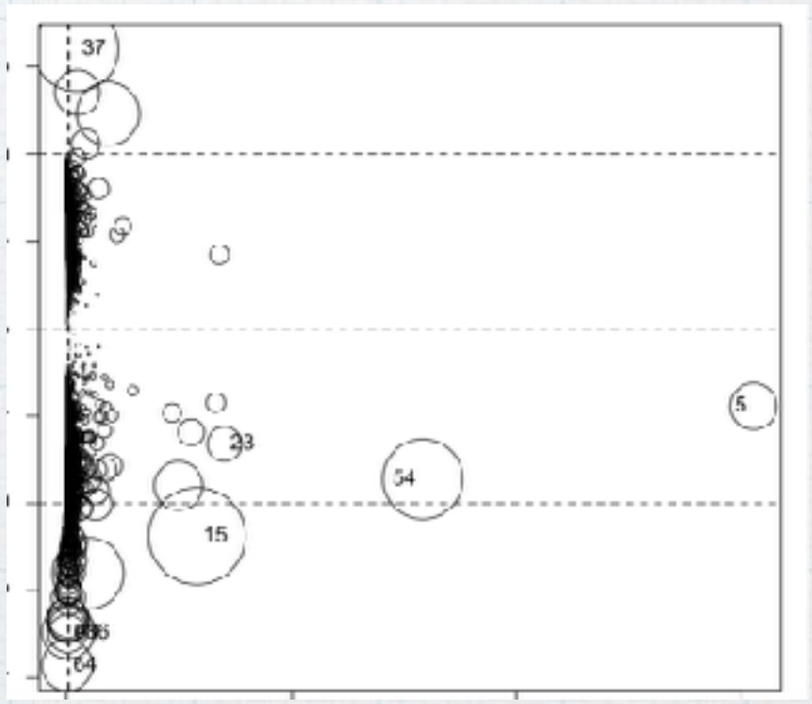
Too much personalization?

Test-Measure-Iterate

Case Study - Improving the odds of purchase through personalized notifications

* Hypothesis - Customers who purchase a product has different browsing pattern than who do not.
* Data Collection - Browsing/purchase pattern of 50k users from last 6 months who registered with us (customer ids)
    * Customer ids and their respective Visits, page-views, product views, number of cart additions, visits to login/account page, shipping page, payment page and units purchased
    * 31,049 did not purchase anything and 18950 purchase at least one product.

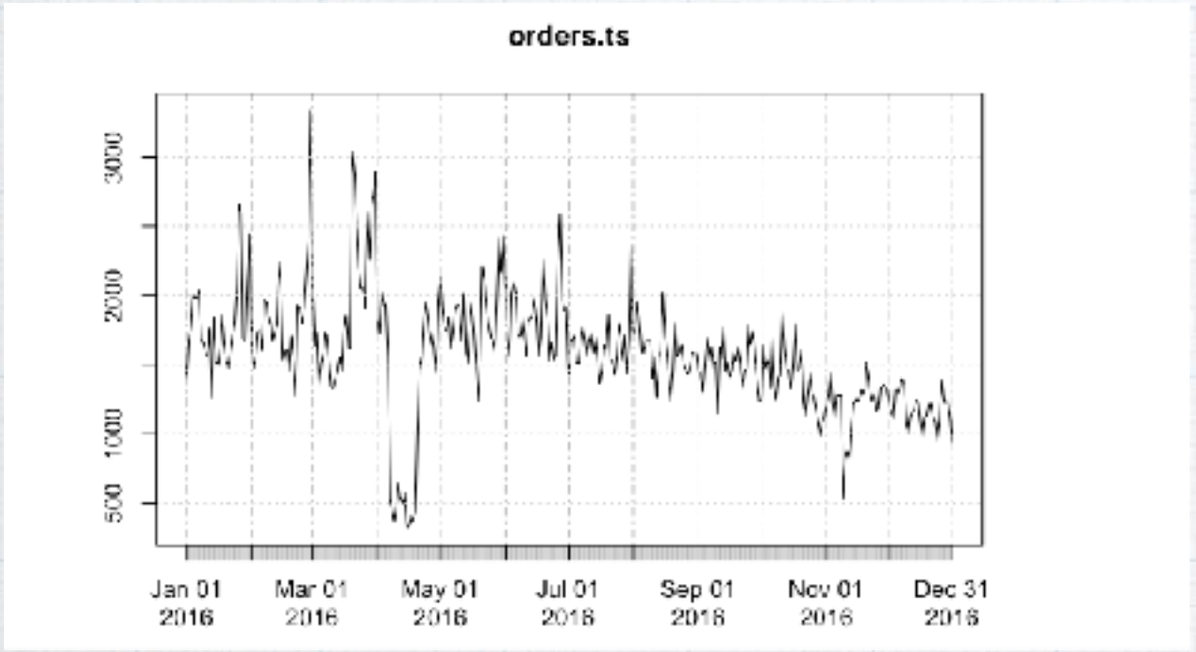
	custid	visits	pageviews	productviews	cartaddition	checkout1	checkout2	checkout3	units
2	12212183	2112	11572	1643	768	2	3667	704	459
3	12054313	1885	8402	189	98	0	3470	2912	223
4	11884269	747	2527	3	2	1	1521	2	108
5	113152	668	12668	2361	820	80	2164	1437	329
6	3292681	404	2833	1454	13	0	4	0	0
7	12283859	325	2701	395	164	17	1046	479	63

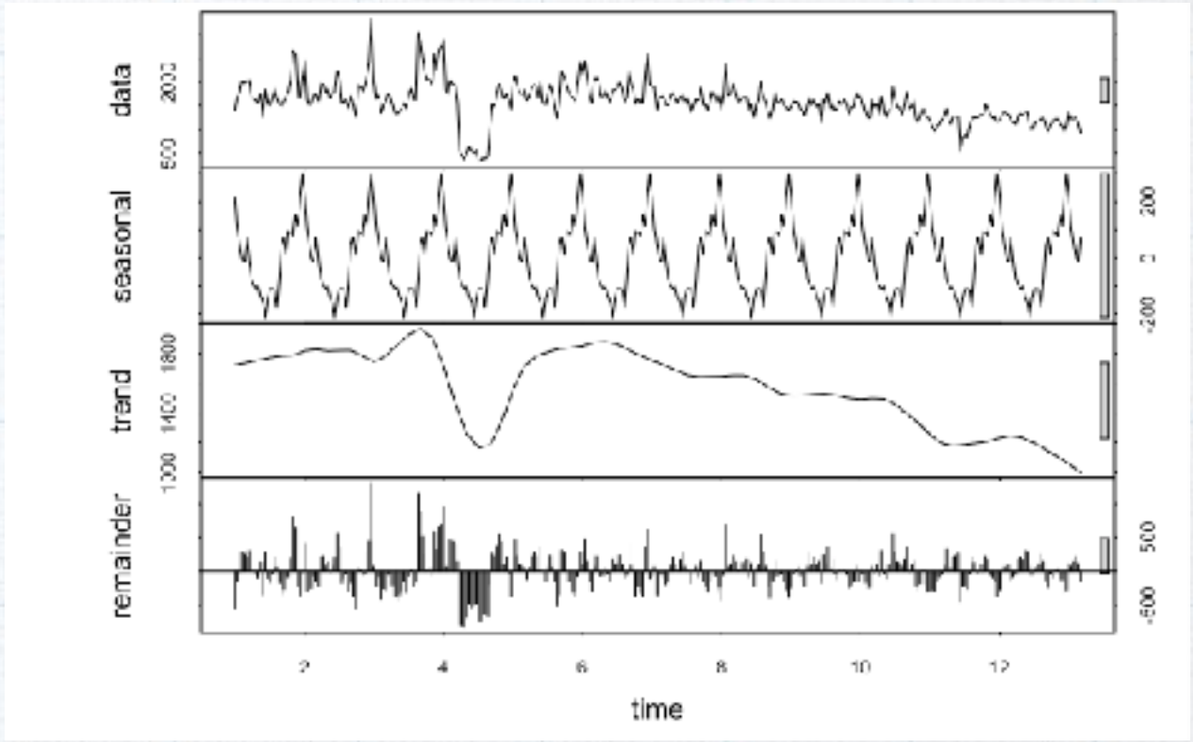


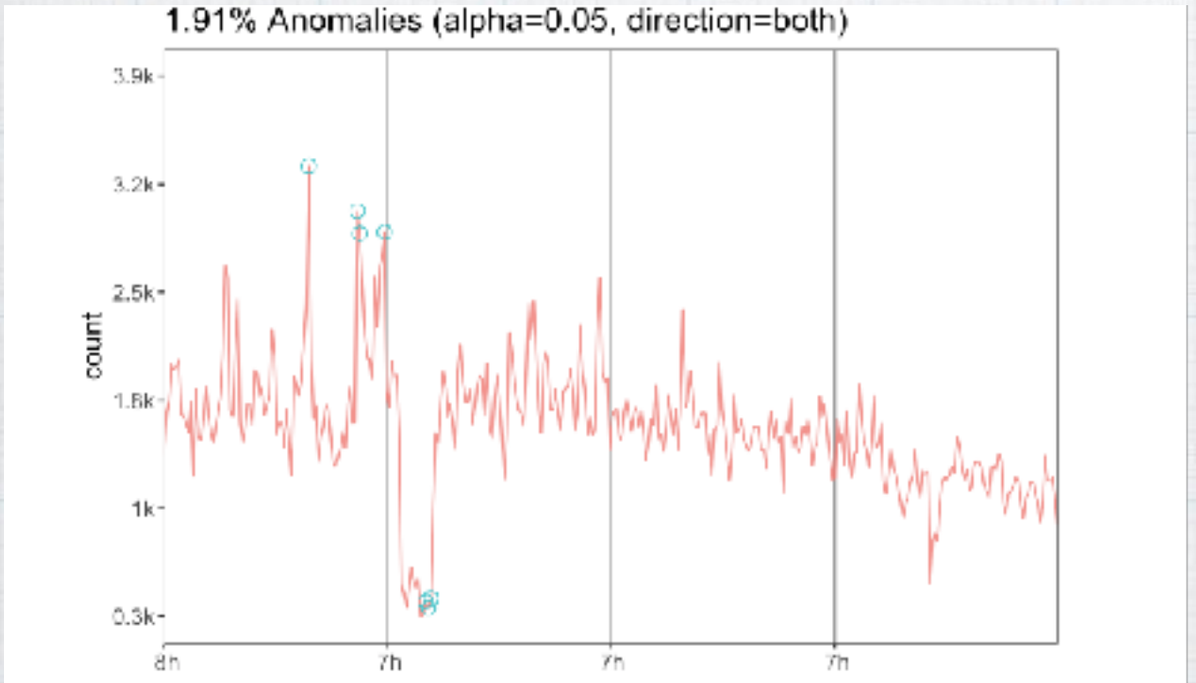
* Findings-
    * Model - Unsupervised k-means Clustering
        * Researchers & One time Buyers (46229 Customers - <1 Units, <3 Visits, >42 page views)
        * Repeat Purchasers (3696 Customers - 2+ Units, average 15 visits, 82 product views, 12 cart add)
        * Software Testers (6 Customers - 210+ Units, 900+ visits, 8730+page views)
        * Resellers (68 Customers - 21+ Units, 140+ visits, 1700+ page views)
    * Model - Logistic Regression on Segment Researchers and One Time Buyers
        * 72% Accuracy, 9.9% Variability
        * One unit less product view, improve the odds of buying the product by factor 0.02
        * One unit more visit to shipping page, improve the odds of buying the product by factor 0.080
        * Making customers to add a product in cart, improve the odds of buying the product by factor 0.78
    * Model - Logistic Regression on Repeat Purchaser
        * 61.83% accuracy, 4.6% variability
        * One unit less product view, improve the odds of buying the product by factor 0.008
        * One unit more visit to shipping page, improve the odds of buying the product by factor 0.012
        * Making customers to add a product in cart, improve the odds of buying the product by factor 1.97

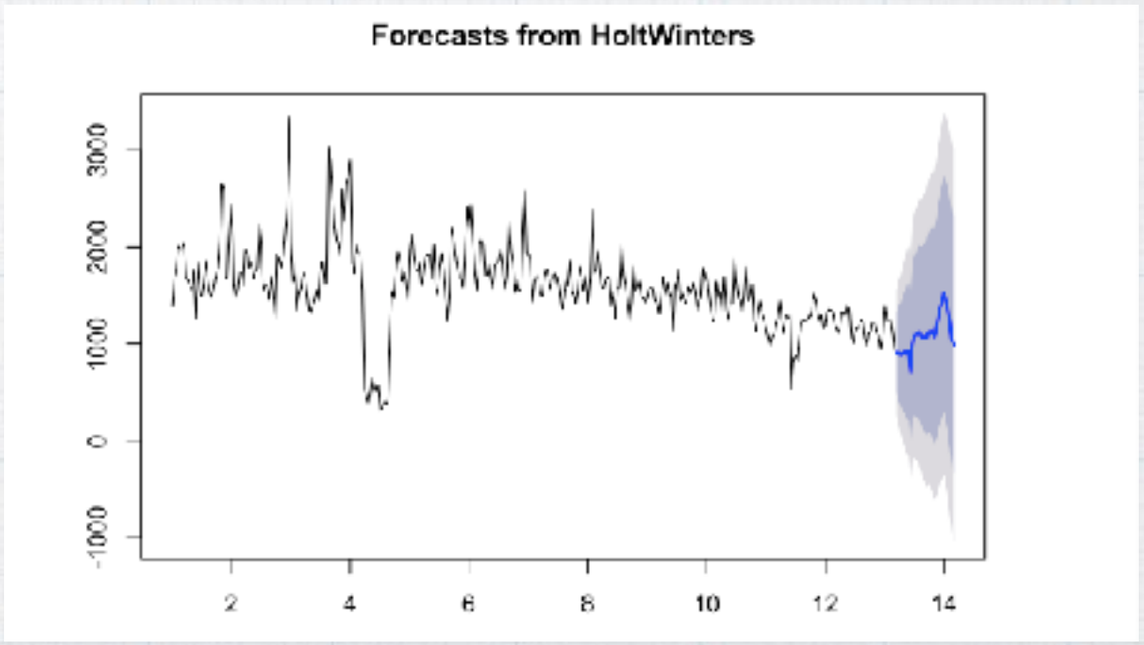
Recommendation
Notification Service for Mobile App at Product View page, Cart Page and Checkout page
Launched A/B test with 10-90 split

Case Study - Understanding Orders Trend and their components


orders.ts
3000
2000
1000
500
Jan 01 2016
Mar 01 2016
May 01 2016
Jul 01 2016
Sep 01 2016
Nov 01 2016
Dec 31 2016


data
seasonal
trend
remainder
time


1.91% Anomalies (alpha=0.05, direction=both)
count
3.9k
3.2k
2.5k
1.8k
1k
0.3k
8h
7h


Forecasts from HoltWinters
3000
2000
1000
0
-1000
2
4
6
8
10
12
14

Questions?