

Mobile Madness

March 21, 2018

Remy, Lauren, Jin, and Karen

Goals

1

Understand concepts:

- Churn
- A/B testing

2

Identify behaviors
strongly correlated
with churn.

3

Identify A/B tests to
develop moving
forward

Understanding Churn

What is Churn?

- Customers stop doing business with a company/service.

Impact

- More cost-effective to retain existing customers than acquire new ones

By managing effectively:

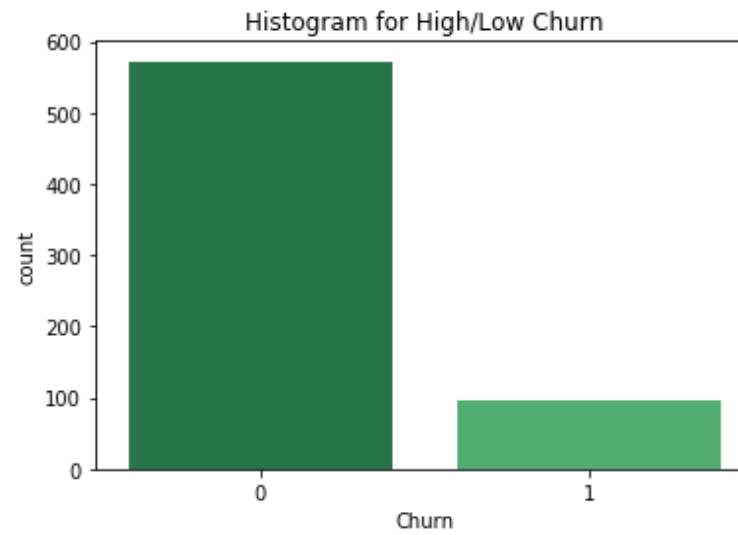
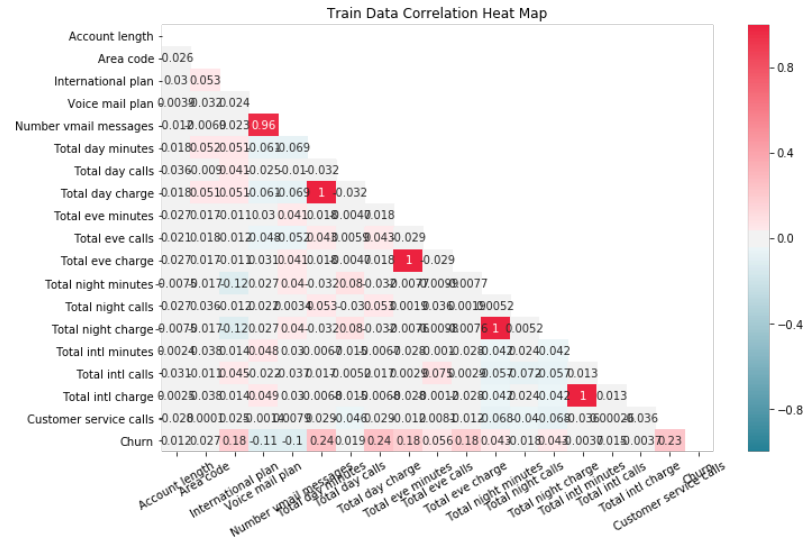
- Anticipate customer needs
- Meet/exceed expectations
- Retain business
- Speed growth

Dataset

State	Account length	Area code	International plan	Voice mail plan	Number vmail messages	Total day minutes	Total day calls	Total day charge	Total eve minutes	Total eve calls	Total eve charge	Total night minutes	Total night calls	Total night charge
LA	117	408	No	No	0	184.5	97	31.37	351.6	80	29.89	215.8	90	9.71
IN	65	415	No	No	0	129.1	137	21.95	228.5	83	19.42	208.8	111	9.40
NY	161	415	No	No	0	332.9	67	56.59	317.8	97	27.01	160.6	128	7.23
SC	111	415	No	No	0	110.4	103	18.77	137.3	102	11.67	189.6	105	8.53
HI	49	510	No	No	0	119.3	117	20.28	215.1	109	18.28	178.7	90	8.04

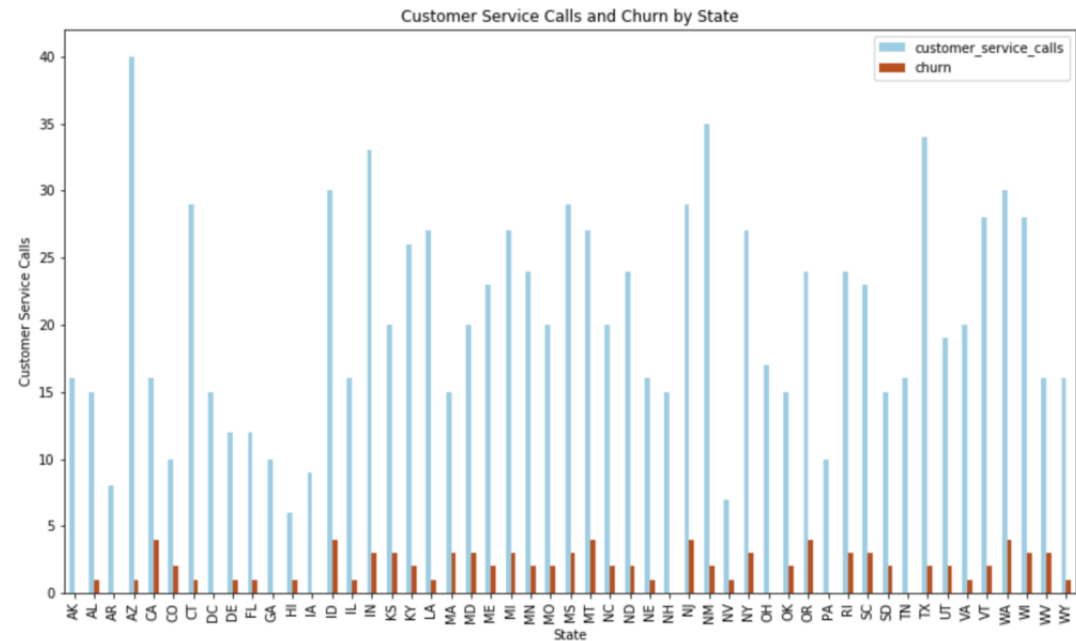
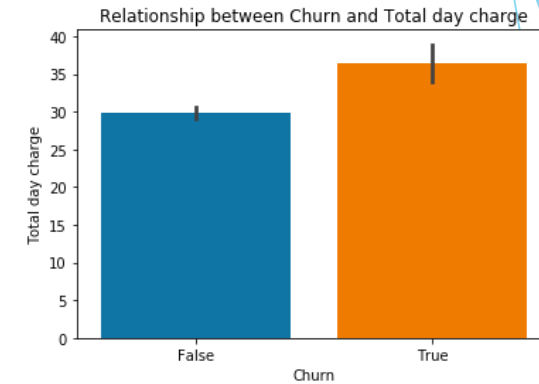
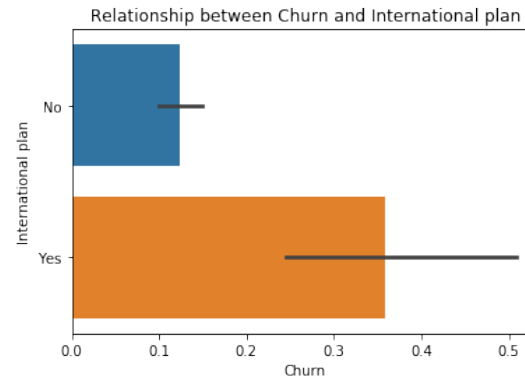
Exploratory Data Analysis

- ▶ To build: classification model to predict churn
- ▶ To identify: features that are strongly correlated with churn



Exploratory Data Analysis

- ▶ Investigating correlation with factor plots



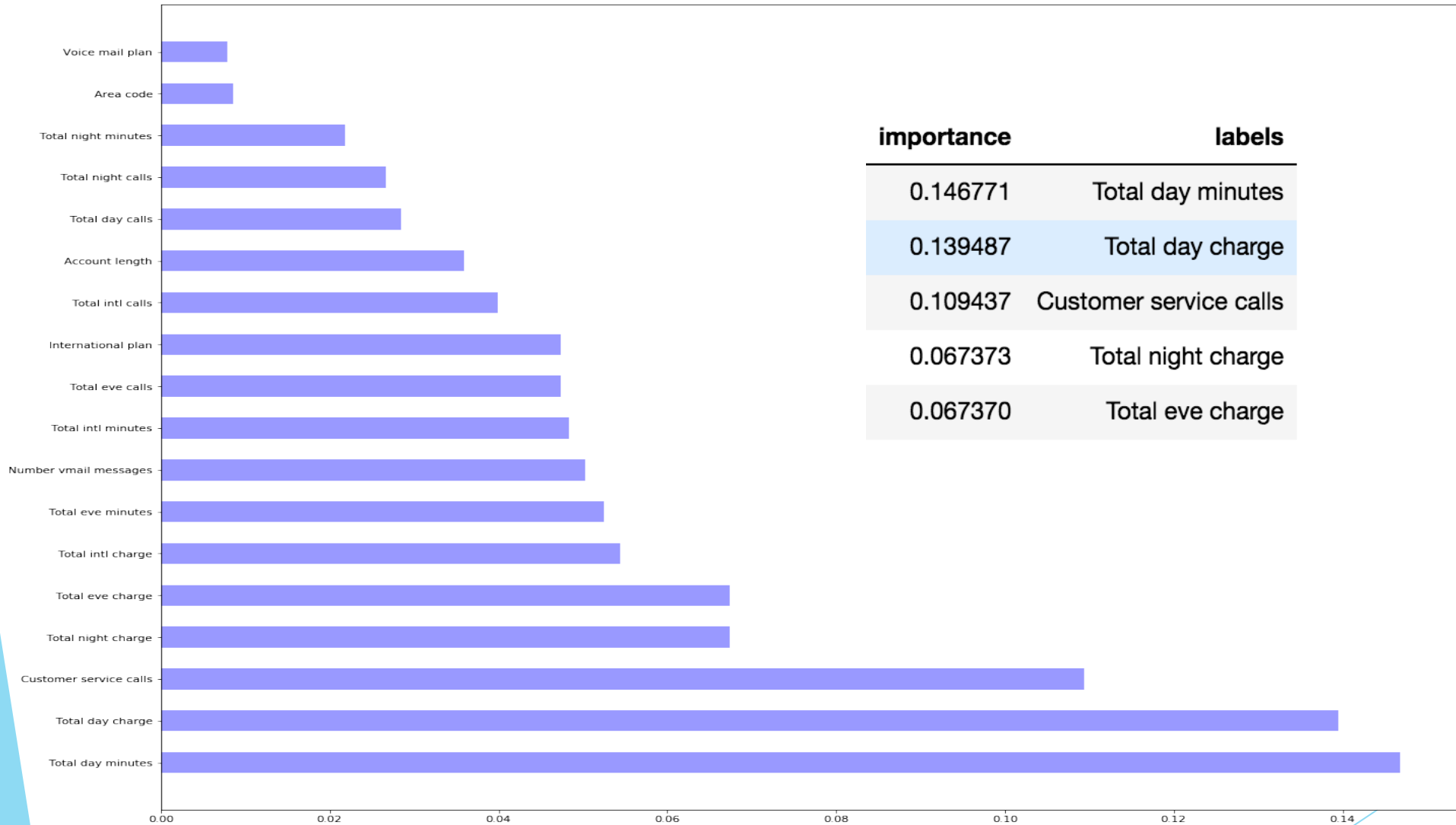
Model Building

Logistic Regression

Random Forest Classifier

GridSearch CV

Feature Importance



Model Scores

Logistic Regression

Logistic Regression Model Score: 0.8982035928143712

Logistic Regression 5 Folds Cross-Validation Score: 0.8807152406417111

Random Forest Classifier

Random Forest Model Score: 0.9461077844311377

Random Forest 5 Folds Cross-Validation Score: 0.8983511586452764

Random Forest Classifier with GridSearch Feature Importances

Random Forest Model Score: 0.9161676646706587

Random Forest 5 Folds Cross-Validation Score: 0.8919786096256684

Understanding A/B Testing

What is A/B Testing?

- Randomized experiment with 2 variants
- Give better level of certainty than observational studies

Questions to Answer

- What factors will be changed?
- Who will be part of the test group?
- How long will the test run?
- Why is this test necessary?

Best Practices:

- Define dependent and independent variables
 - Before test begins
- Randomly assign individuals to groups

Recommendations