Summary

Business Problem

• A telecommunication major wanted to understand whether there is a problem on its website
• Current process was manual and the expected state is to derive anomalous behavior across major KPIs in an automated manner

Methodology

Identify key KPIs

Identify business drivers and their impacts on conversion rate

Detect anomaly

Find good and bad anomalies across major KPIs

Forecast behavior

Predict the following 7 days and detect anomalies in the forecast

Provide insights

Derive insights from analysis and give recommendations through Tableau visualizations

Some Insights

• There's a potential customer tagging problem - whenever undetermined increases, customer and prospect decrease, and vice versa.
• Many visitors browse from phones but do not place orders through mobile that often compared to desktop.
• More customer browse on desktop while more undetermined browse on mobile.
• Traffic of undetermined > customer > prospect. Avg conversion rate of customer > prospect > undetermined.
• Traffic of desktop > mobile > tablet. Avg conversion rate of desktop > tablet > mobile.
• Changes of traffic and orders stay consistent across device types but vary from customer types.
• On 2/3, three metrics all go down:
Visitors and orders from Prospect and Customer decrease
Visitors and orders from Undetermined experience a little bump up but not enough to affect the downward trend
• On 2/22, visitors, digital orders and conversion rate all spike up:
Large volumes from Undetermined visit the website and place orders
Fewer visitors and orders from Prospect and Customer
Conversion rate of Prospect on this day increases while that of Customer decreases

Executive Overview

Count of Date	Visitors	Orders	Avg. Conversion Rate	Avg. Ind Conversion Rate
59	38,321,205	368,753	0.96%	2.43%

Visitors

	Device		
Customer	Desktop	Mobile	Tablet
Customer	7.64M	4.79M	0.79M
Prospect	1.00M	1.16M	0.18M
Undetermined	9.41M	12.07M	1.34M

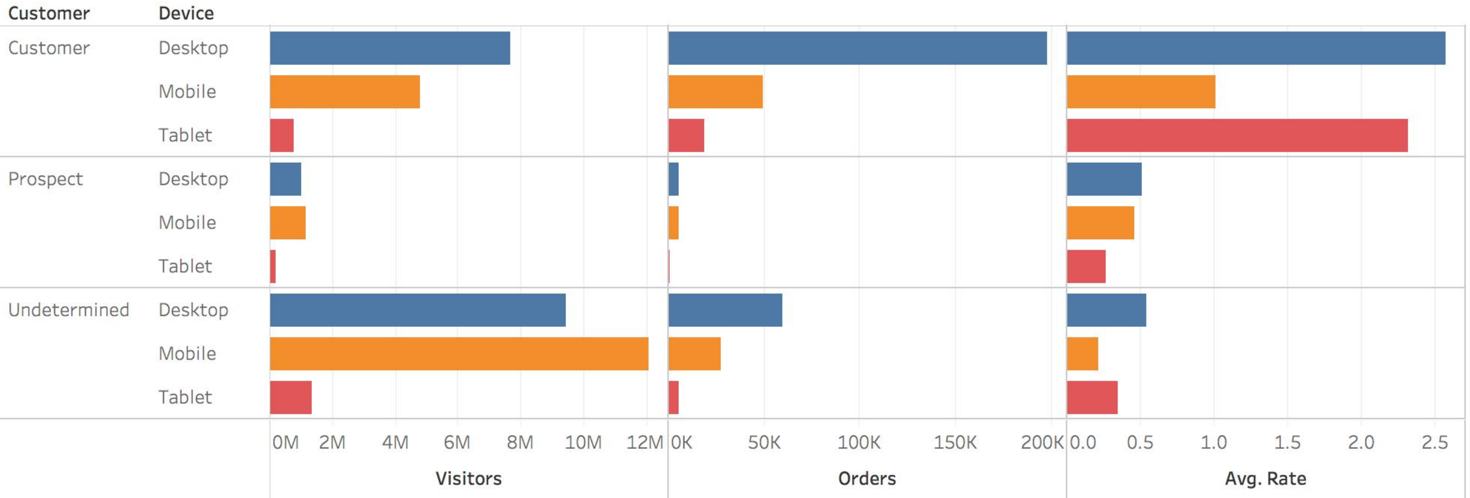
Orders

	Device		
Customer	Desktop	Mobile	Tablet
Customer	197.98K	49.10K	18.77K
Prospect	5.14K	5.29K	0.47K
Undetermined	59.36K	27.40K	5.20K

Rate

	Device		
Customer	Desktop	Mobile	Tablet
Customer	2.57%	1.01%	2.32%
Prospect	0.51%	0.46%	0.27%
Undetermined	0.55%	0.21%	0.35%

Visitors, Orders and Rate by Customer and Device Type



* Visitors and orders from Gaming Console and E-Reader are low so they are excluded from analysis
* Reference: https://www.growcode.com/blog/ecommerce-conversion-rate/

Overview | Exploratory Analysis | Model Selction | Anomaly Detection | Breakdown | Forecasting | Metric Evaluation | Summary | Story

Summary | Verizon's CR is lower than industry avg | Explore - customer tagging problem | More Undetermined browse through Mobile | While more Customers browse and place orders | When do anoumalous behaviors happen? | Tw du

Exploratory Analysis

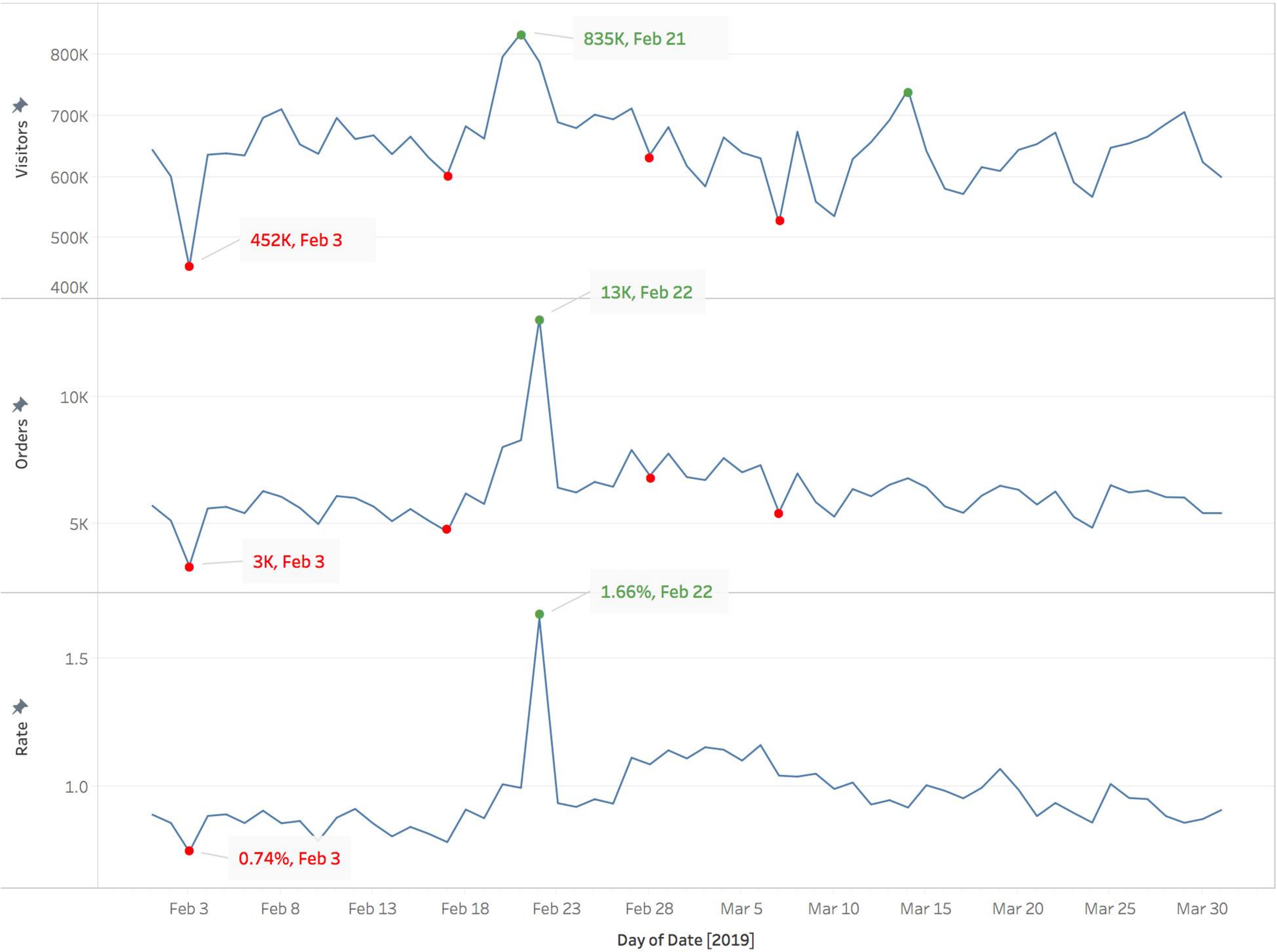
Overview | Exploratory Analysis | Model Selction | Anomaly Detection | Breakdown | Forecasting | Metric Evaluation | Summary | **Story**

Model Selection

Pros	Cons
ARIMA	
• Well-studied • Parameters(p, d, q) -> more accurate	• Model assumption – stationary • Parameter tuning
Twitter+IQR	
• Catch seasonality • No loops – easily scaled	• No long term trend • IQR is easily skewed by outliers
★ STL+GESD	
• More sensitive to capture anomalies • Less resistant to outliers	• Iterative – more expensive

Detect Good & Bad Anomalies

Using selected model - STL+GESD



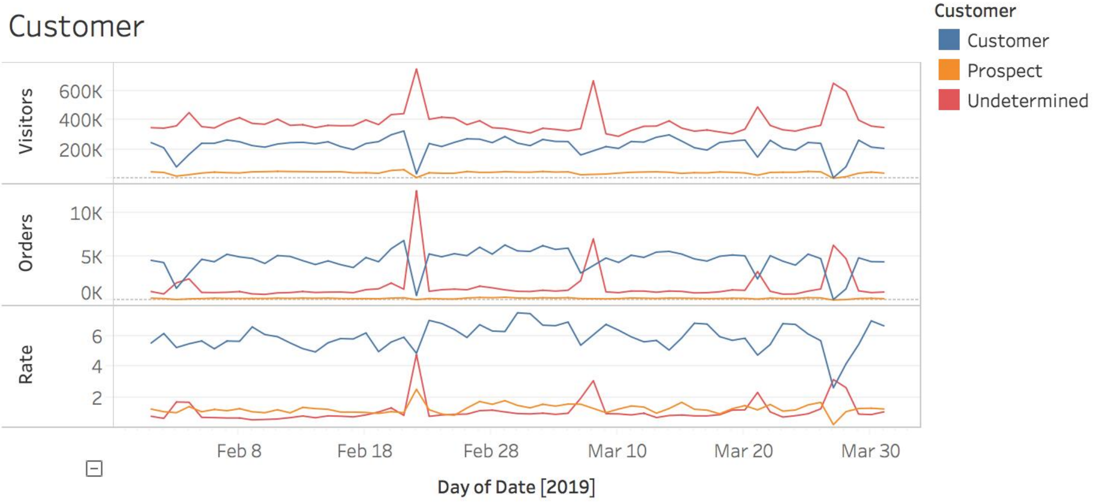
Breakdown

Key Points

• On Feb 3, fewer Customer and Prospect visit the website and place the orders. Those from Undetermined experienced a little bump up but not enough to affect the downward trend.

• On Feb 22, large volumes of traffic and orders come from Undetermined, and conversion rate also increases on this day.

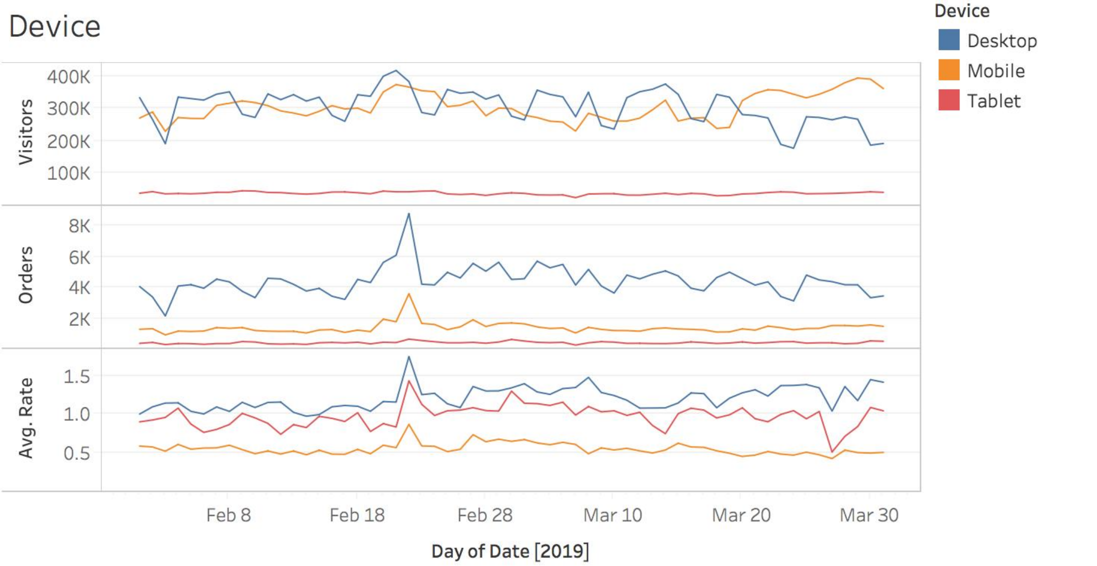
• On Feb 17, Feb 28 and Mar 7, anomalies are found in both visitors and orders but the rates get smoothed out.



• In general, the trends from three devices stay consistent.

• On Feb 3, the decrease in visitors and orders is mainly caused by Desktop. The conversion rates do not experience noticeable changes.

• On Feb 22, visitors, orders and rate increase from all of the three devices.



Forecasting

Pros	Cons

ARIMA

Pros	Cons
• Well-studied • Parameters(p, d, q) -> more accurate	• Model assumption – stationary • Parameter tuning

ETS

Pros	Cons
• No requirement for stationarity	• Parameter selection

Holt-Winter

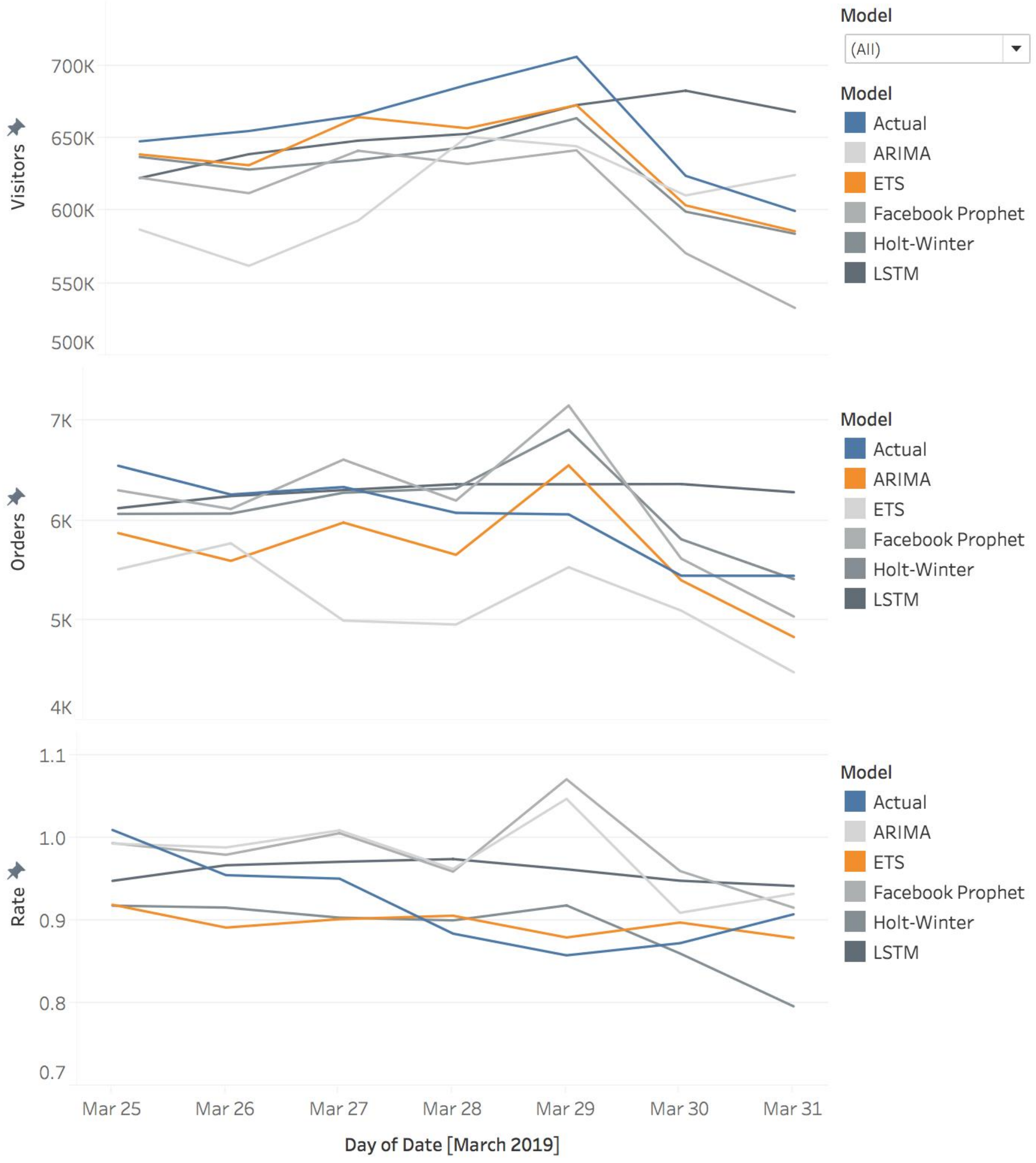
Pros	Cons
• Triple exponential smoothing • Catch seasonal components	• Model complexity - overfitting

LSTM

Pros	Cons
• Time dependency • Perform well with large data	• Hyper parameter tuning • Require large data

Facebook Prophet

Pros	Cons
• Strong seasonal effects • Robust to missing data and outliers	• Overfitting • Require large data



Overview | Exploratory Analysis | Model Selction | Anomaly Detection | Breakdown | **Forecasting** | Metric Evaluation | Summary | Story

Metric Evaluation

Visitors

	Metric		
Model	MAPE	MASE	RMSE
ARIMA	7.35%	0.51	132365
ETS	2.94%	0.21	57044.3
FB	7.91%	0.54	132491
Holt-Winter	4.36%	0.45	78978.5
LSTM	6.22%	0.45	109967

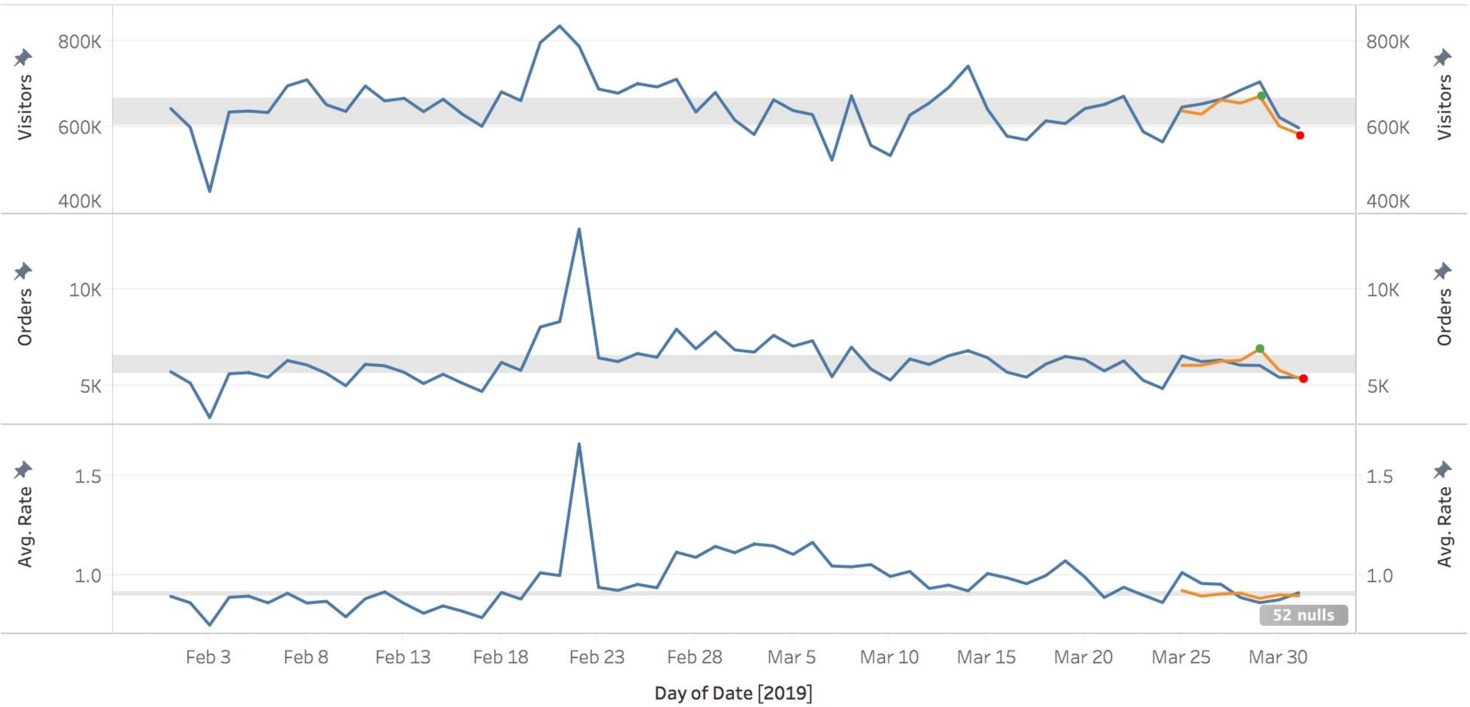
Orders

	Metric		
Model	MAPE	MASE	RMSE
ARIMA	16.30%	1.59	2376.38
ETS	8.22%	0.89	1343.48
FB	5.22%	0.67	1242.65
Holt-Winter	4.99%	0.6	1083.92
LSTM	6.36%	0.77	1373.22

Rate

	Metric		
Model	MAPE	MASE	RMSE
ARIMA	4.76%	0.48	1.3E-03
ETS	6.26%	0.69	2.10E-03
FB	7.66%	0.76	2.5E-03
Holt-Winter	6.19%	0.69	1.69E-03
LSTM	5.94%	0.63	1.7E-03

Anomalies in Forecast



* Anomalies: points lie out of ± 1 standard deviation, which includes 65% data.