Demand Forecasting to Increase Profits on Perishable Items

Ankur Pandey

Arun Chaubey

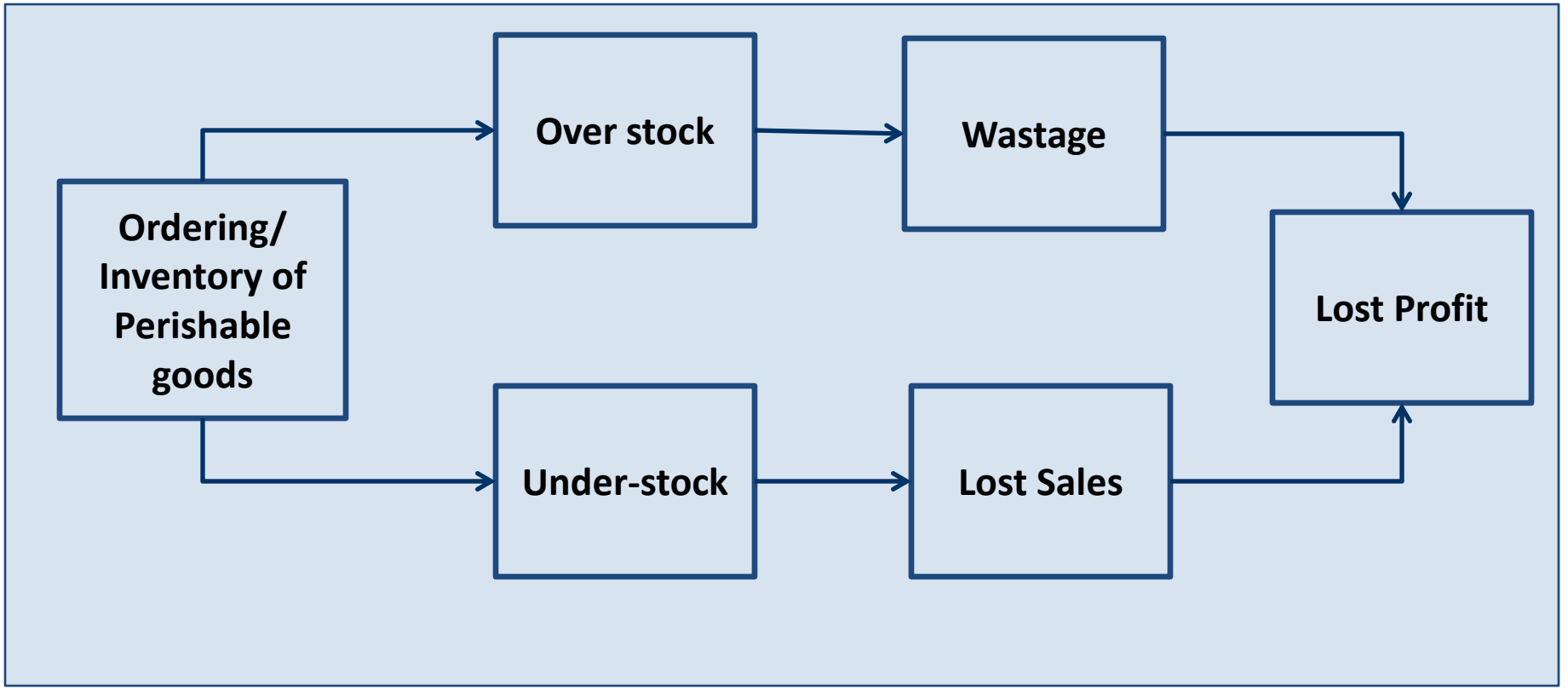
Sanchit Garg

Shahid Siddiqui

Sharath Srinivas

Forecasting Analytics, ISB

Goal | Data | Explore & Visualize | Pre-process | Forecasting | Evaluate



GOAL: Maximize profits when selling perishable items such as fruits and vegetables

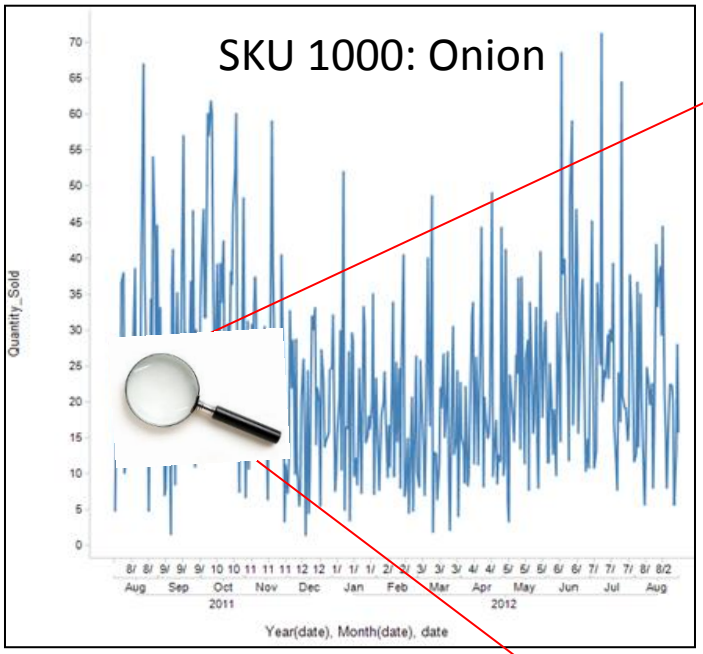
Data:

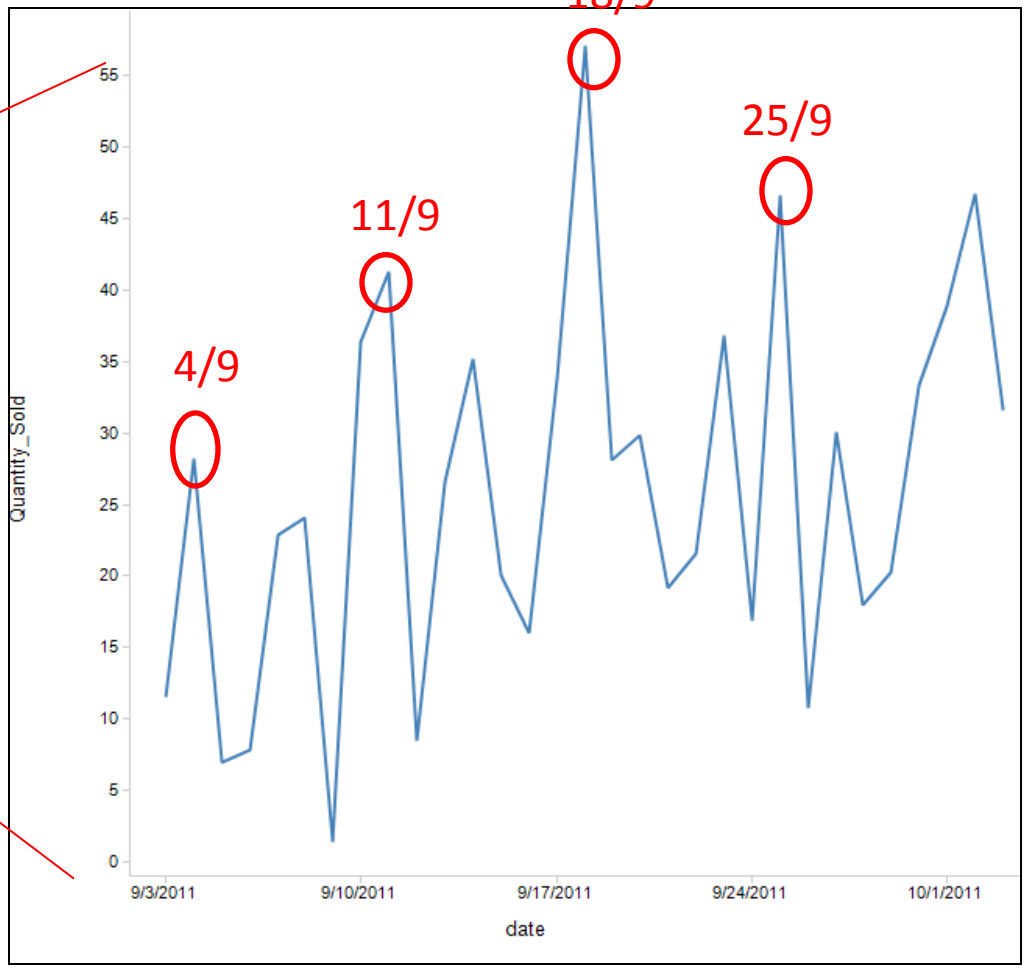
- Hypermart customer Transaction data from 8/1/2011 to 8/31/2012
- Each transaction includes the customer ID, SKU and purchase quantity
- 5 SKUs were explored: Banana, Apple, Onion, Tomato, Papaya

Caveats:

- Transaction data includes only loyalty card purchases
- Data does not include promotions
- Data includes only customer demand and not indicate inventory levels, procurement etc.
- Infrequent visits by customers to the store

Goal | Data | **Explore & Visualize** | Pre-process | Forecasting | Evaluate

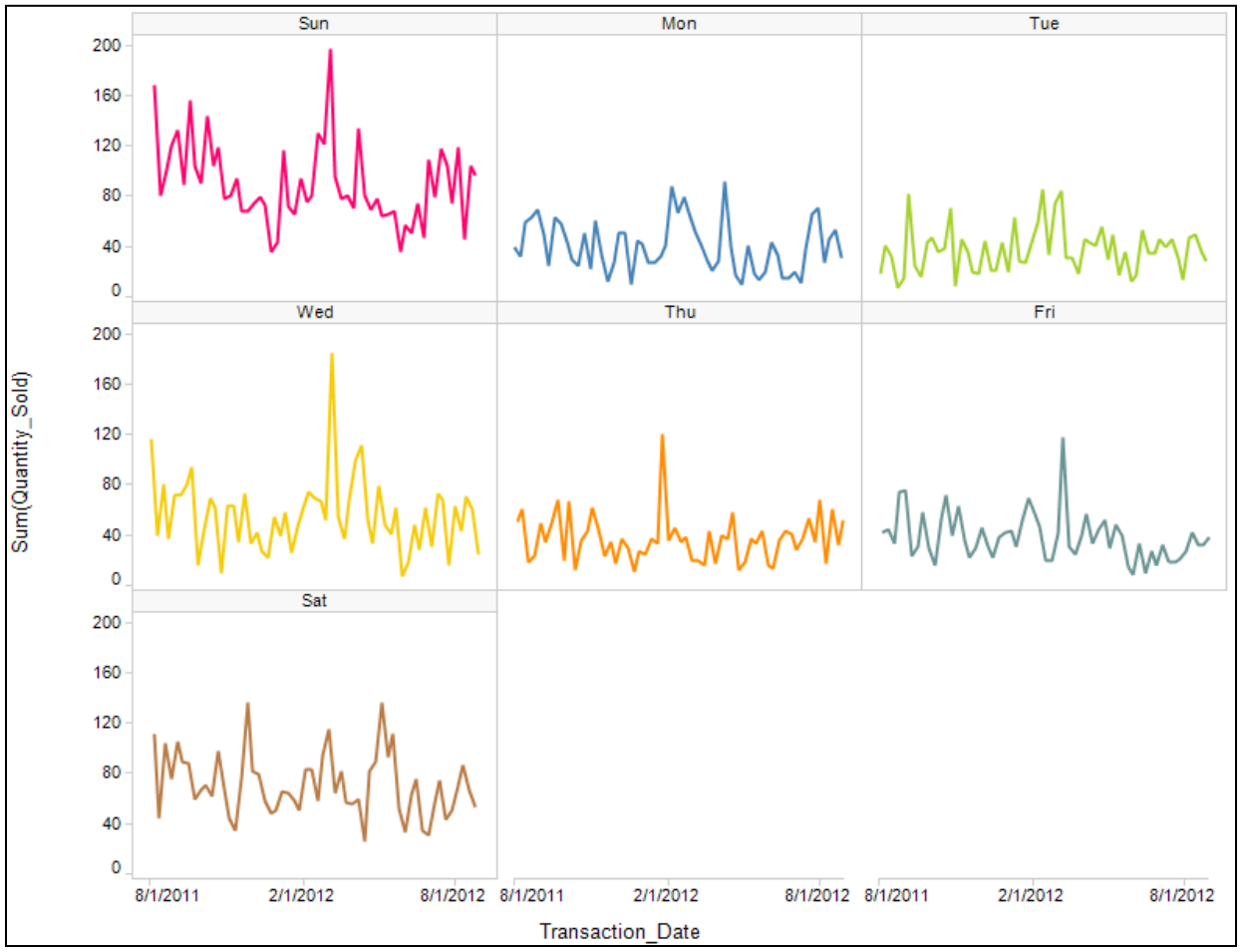


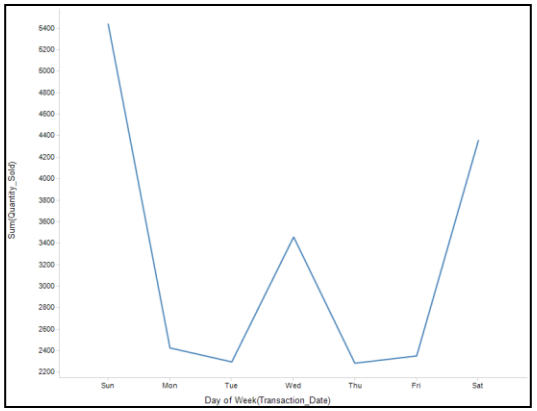


September

S	M	T	W	T	F	S
				1	2	3
4	5	6	7	8	9	10
11	12	13	14	15	16	17
18	19	20	21	22	23	24
25	26	27	28	29	30	

Hypothesis: There is weekly seasonality



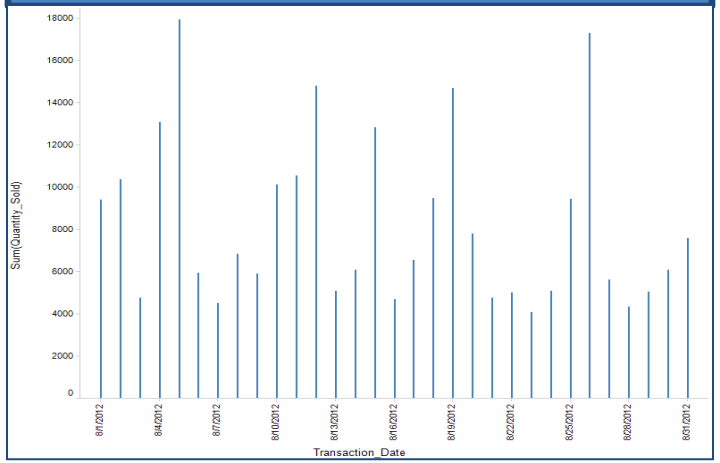


Higher demand on:

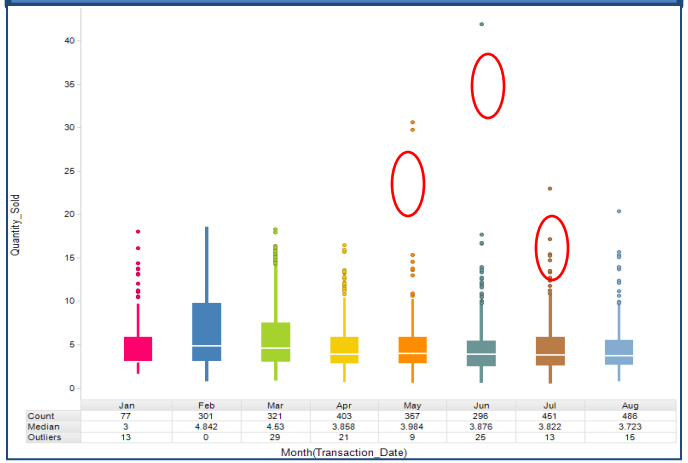
- Sunday
- Saturday
- Wednesday

Goal | Data | Explore & Visualize | **Pre-process** | Forecasting | Evaluate

Aggregate daily customer transaction volumes



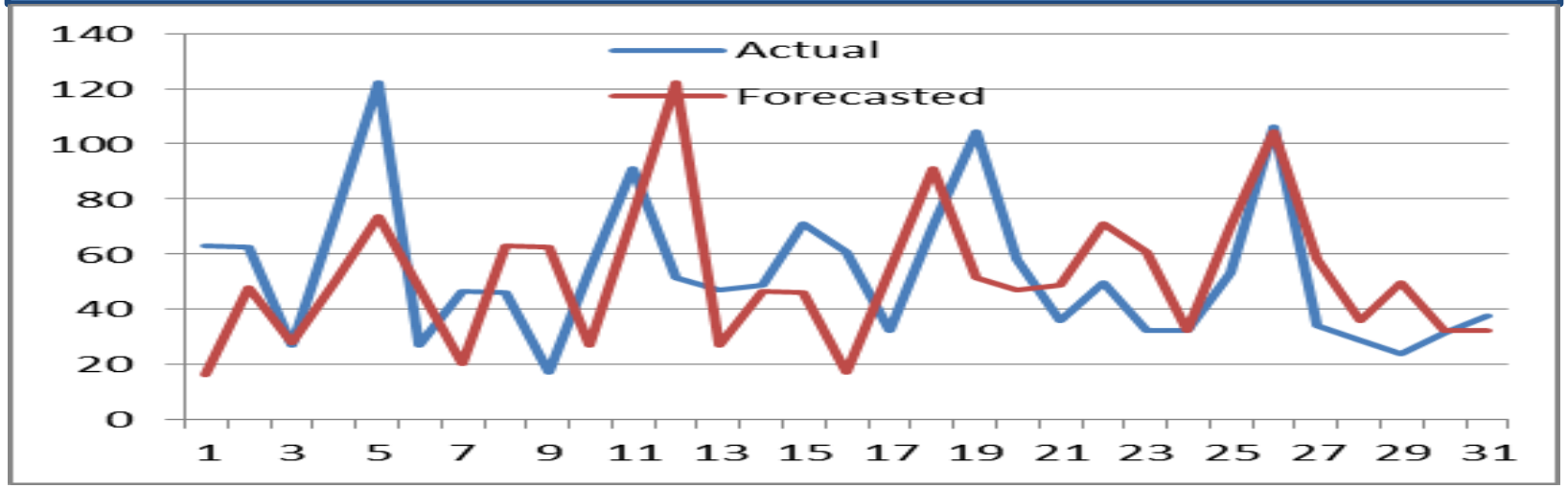
Remove Outliers



	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug
Count	77	301	321	403	357	296	451	486
Median	3	4.842	4.53	3.858	3.984	3.876	3.822	3.723
Outliers	13	0	29	21	9	25	13	15

Partition – Training and Validation

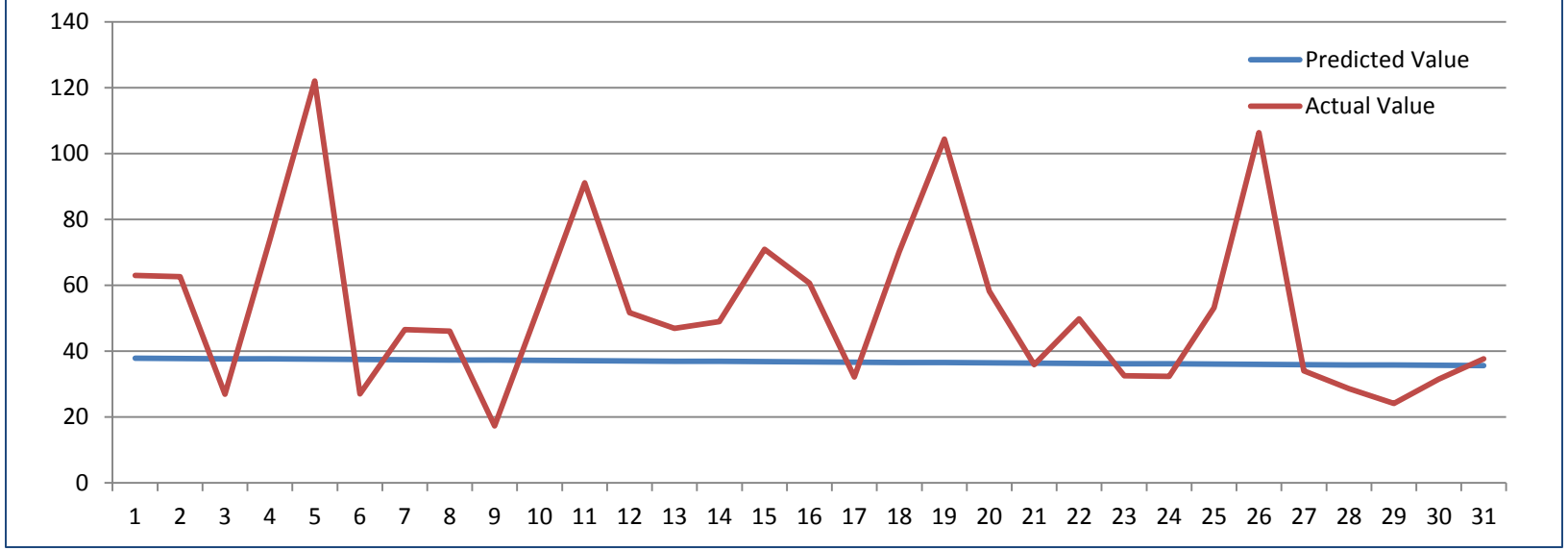
Naïve Model: Forecast for Aug 2012

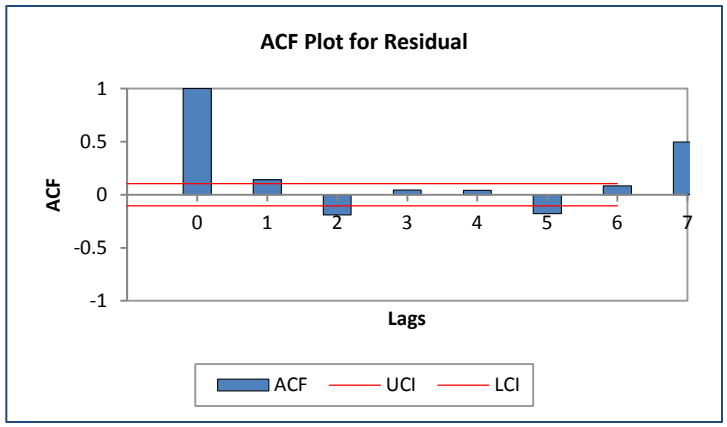


MAE	22.632
Average Error	1.004903
MAPE	50.15%
RMSE	791.4196

- Naïve model can only be used as a benchmark
- Accuracy of the model is very low

Linear Regression Model (With no Dummy Variables): Forecast for Aug 2012

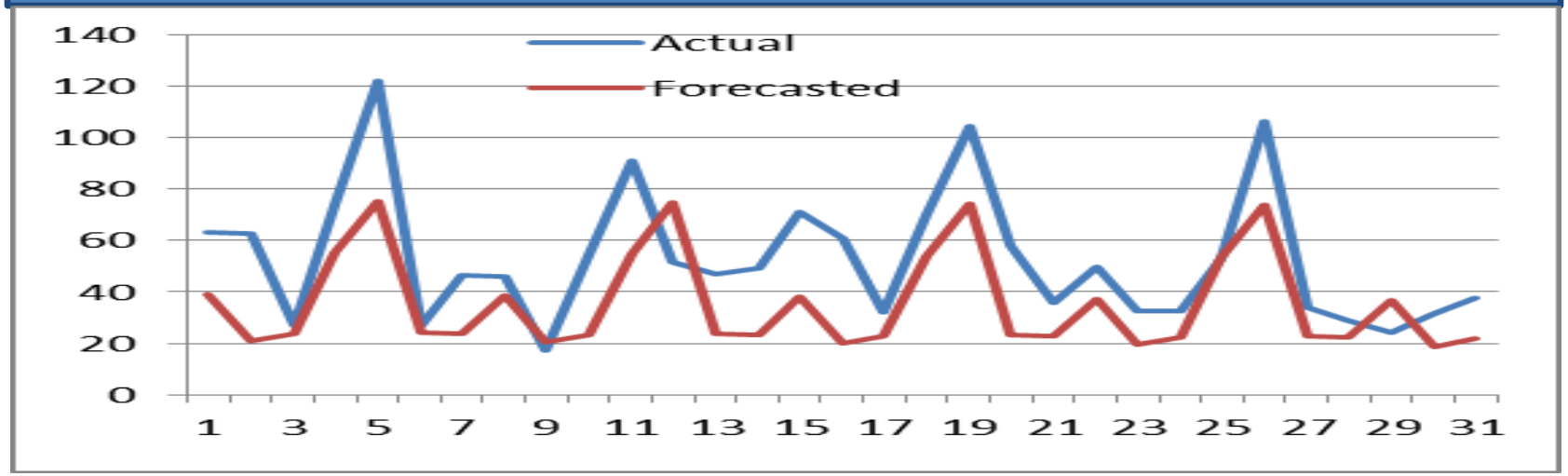


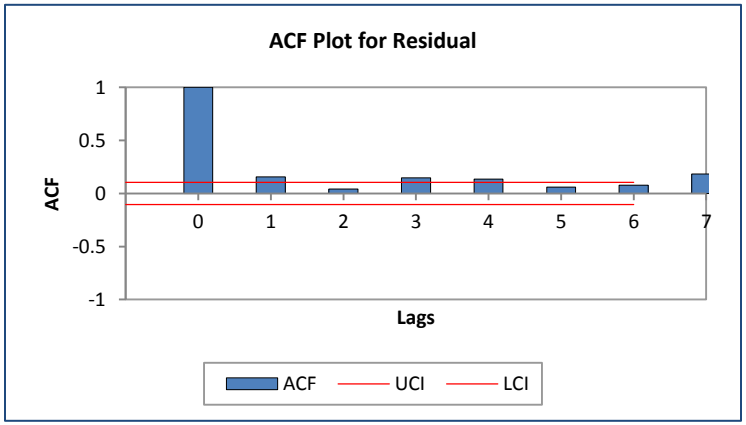


MAE	21.26407
Average Error	16.19199
MAPE	35.33%
RMSE	896.3684

- **ACF Plot shows that there is seasonality left in the residuals**

Linear Regression Model (With Dummy Variables): Forecast for Aug 2012

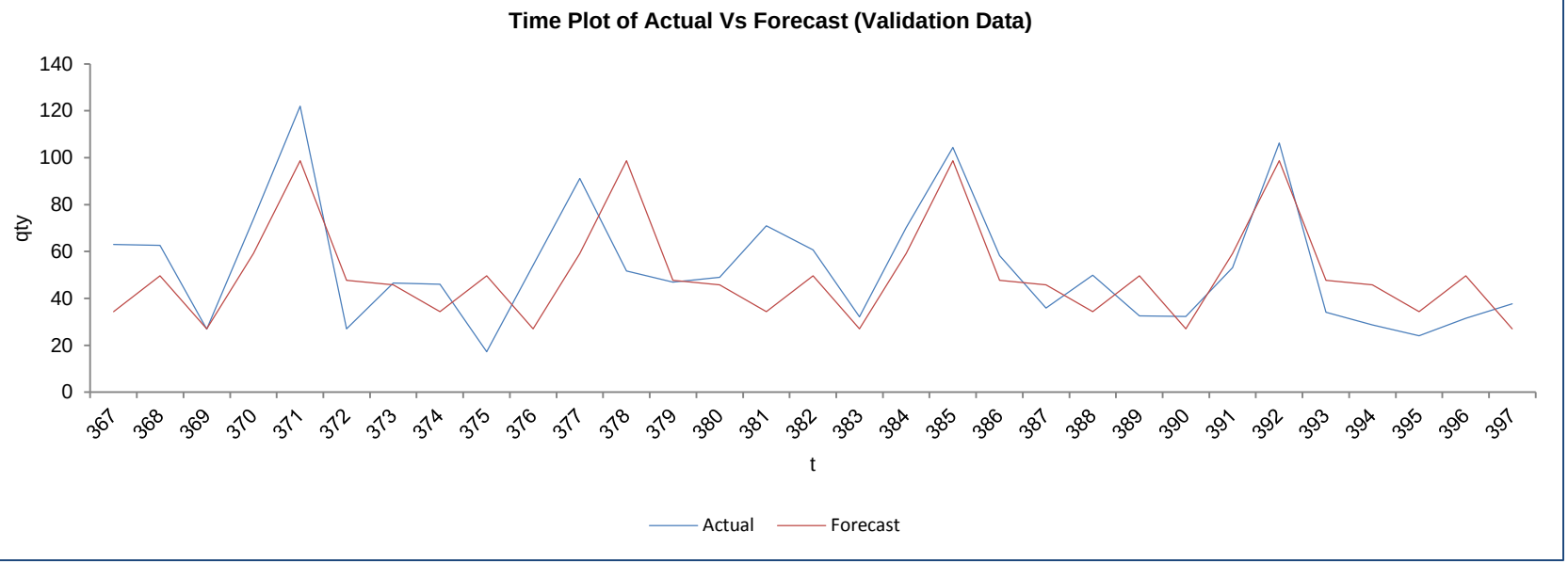


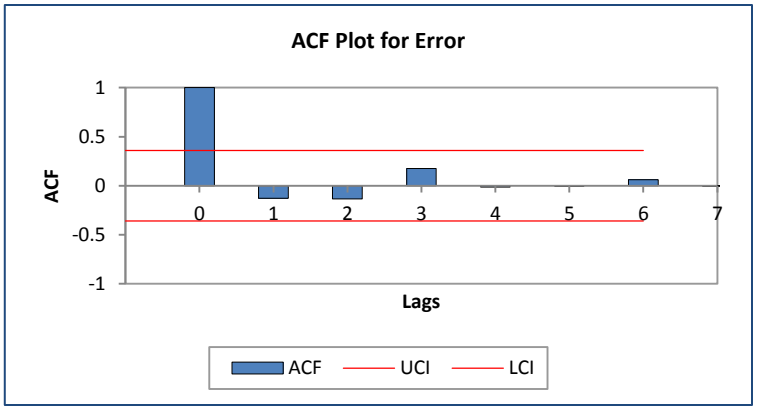


MAE	19.8724
Average Error	17.33157
MAPE	36.46%
RMSE	552.8134

- **ACF Plot shows that there is seasonality left in the residuals**

Holt Winters: Forecast for Aug 2012





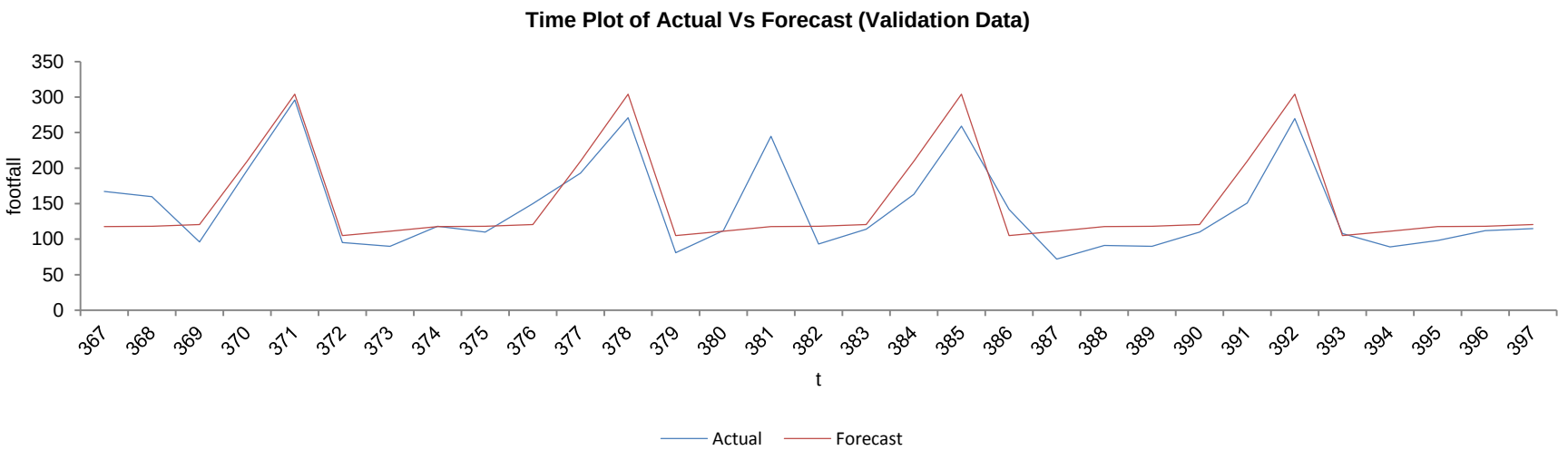
MAE	15.04821
Average Error	2.589459
MAPE	34.84%
RMSE	350.3683

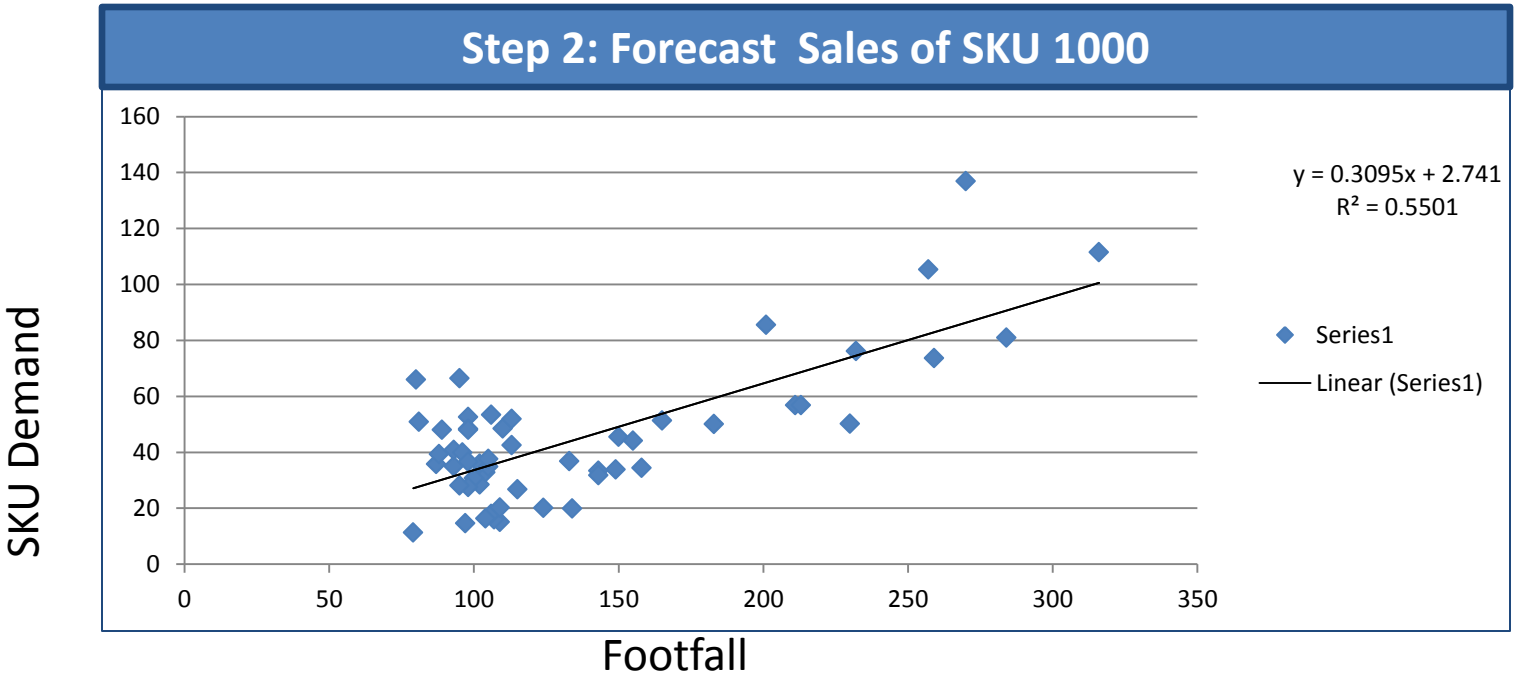
Seasonality handled by Holt Winters method

Two Stage Model

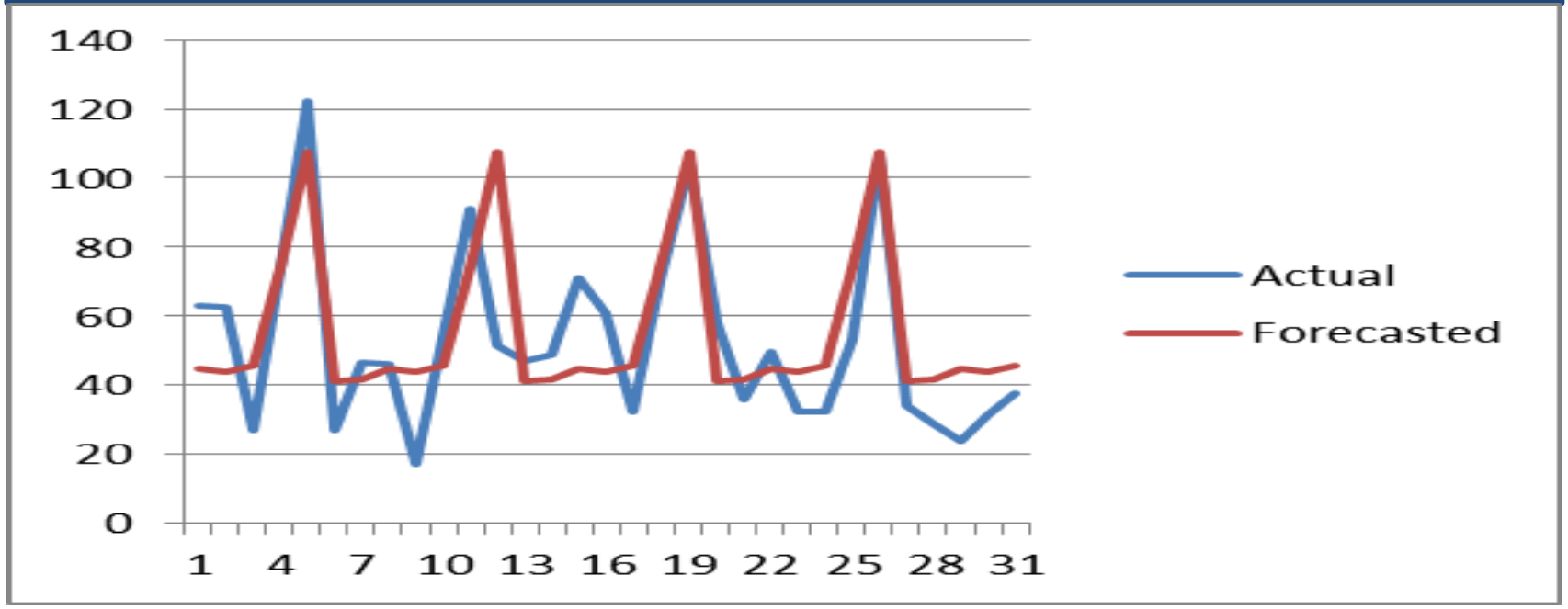
- **Sales of individual SKUs categorized on Day of the week is noisy**
- **Instead we forecast amount of Footfall in the store**
- **Use footfall as a proxy to forecast the SKU quantity demand**

Step 1: Forecast Footfall




Step 2: Forecast Sales of SKU 1000
SKU Demand
Footfall
0
20
40
60
80
100
120
140
160
0
50
100
150
200
250
300
350
y = 0.3095x + 2.741
R² = 0.5501
Series1
Linear (Series1)

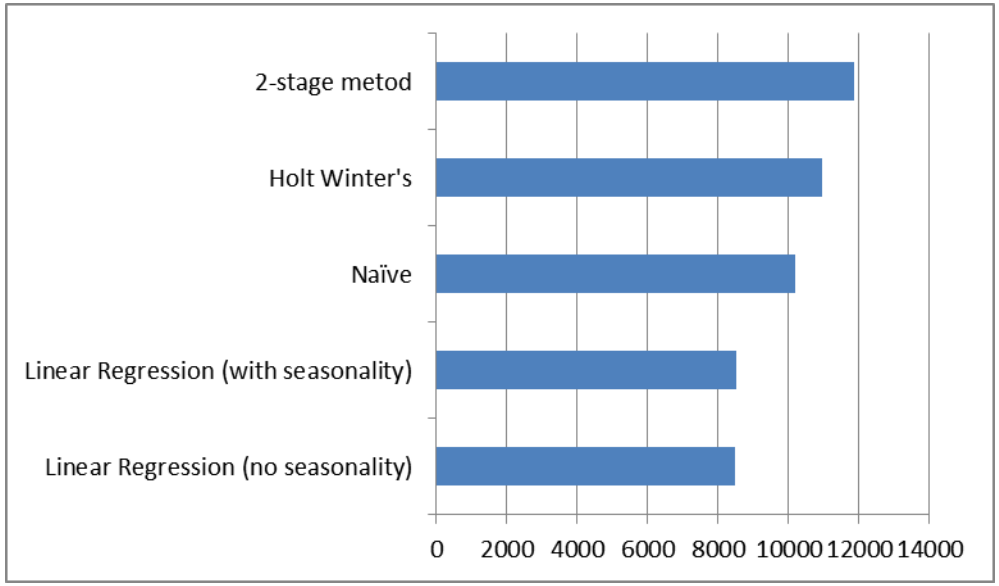
2 Staged Model: Forecast for Aug 2012



MAE	13.27165
Average Error	-2.95764
MAPE	33.33%
RMSE	286.4652

Two step process:

1. Determine a cost-metric (e.g. profit: See Appendix for profit calculations)
2. Evaluate the effect of different forecasting methods on the metric



Achieved **17% improvement** in profitability by **leveraging** advanced forecasting techniques at SKU level (compared to baseline-naïve forecasting

Recommendations

- By forecasting demand at the SKU level, the store can increase profitability by:
  - Reducing wastage
  - Reducing lost sales

- Two-staged model offers best performance in terms of profitability improvement

- **Key Learning:** Handling noisy data, "*Torture the data, and it will confess to anything!*"

Appendix

A Model of System Mechanics

- Suppose $Q = 100$
  - suppose demand turns out to be $D = 110$
  - profit = \$70,000(100) - \$55,000(100) = \$1,500,000
- Now suppose $Q = 100$ and demand turns out to be $D = 90$
  - profit = \$70,000(90) + \$20,000(10) - \$55,000(100) = \$1,000,000
- The General Formula:

	Sales	Salvage	COGS
if $D \geq Q$	$P*Q$	0	$C * Q$
if $D < Q$	$P*D$	$Sx (Q - D)$	$C * Q$

- As a single formula…
  - $profit(Q, D) = IF(D<Q, (P \times D)+S \times (Q-D)-(C \times Q), (P \times Q)-(C \times Q))$