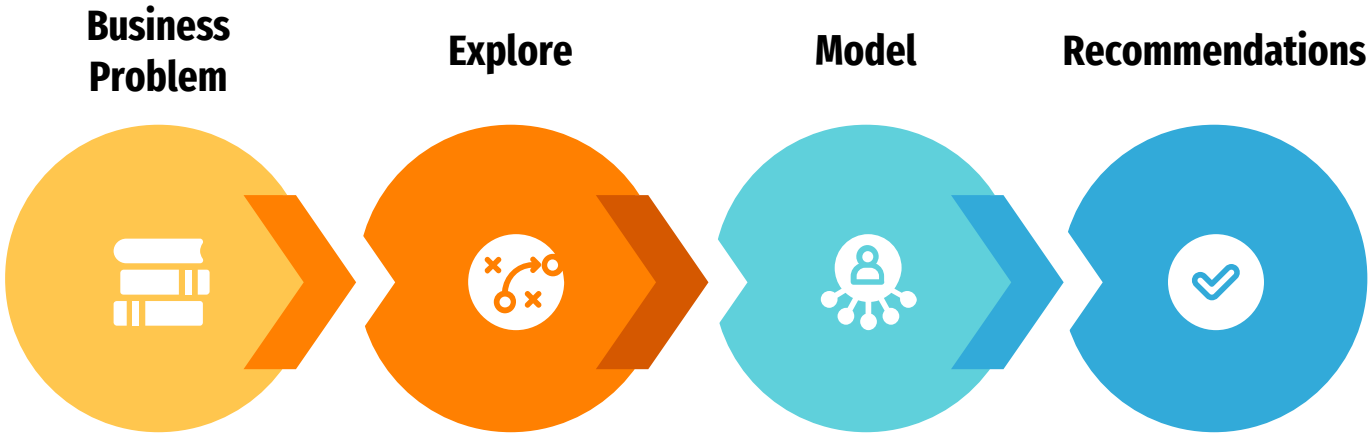
Advanced Business Analytics Final Project

Forecasting Customer Returns

Section 3 Team 9

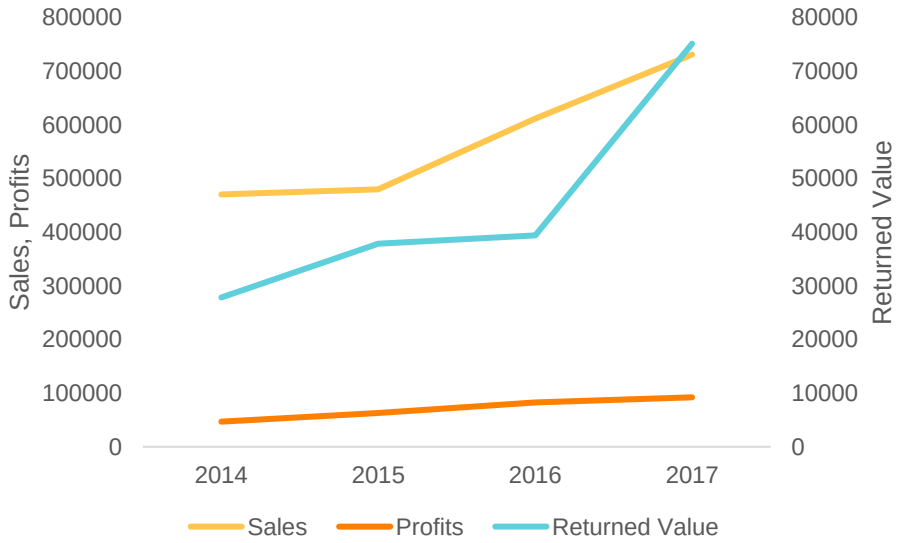
Abhishek Nambiar
Aditya Uppuluri
Lavanya Karthikeyan
Madhulika Chilla
Mandeep Singh Rahi
Tanveer Singh Sapra

Steps followed

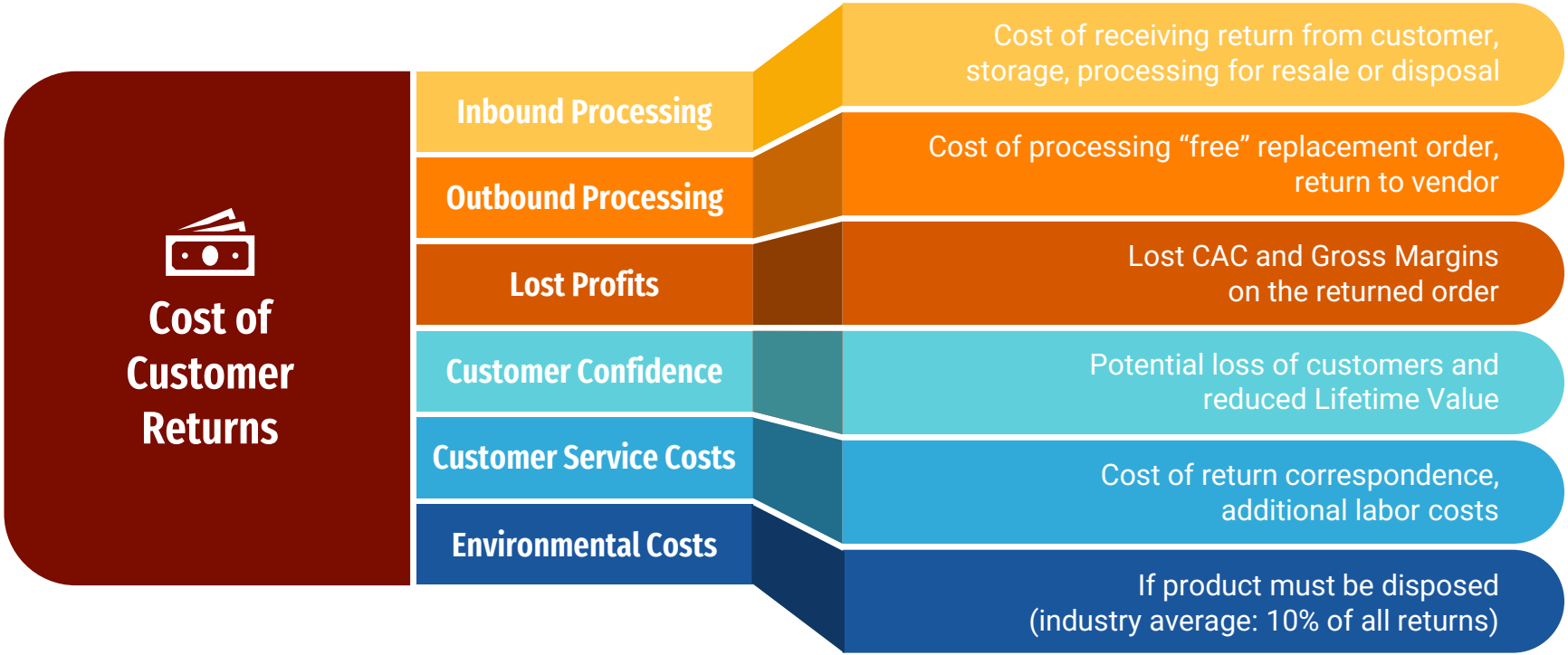


Defining the business problem

- Steady increase in sales and profits for the business
- But, returned value crossing more than **10%** of sales value in 2017
- Puts future sales planning at risk
- Expected to increase in the future with eCommerce



Other than losing almost 10% of revenue, additional costs are incurred on customer returns



Dataset description and aggregation

Order Data

Collection of sales orders with relevant features

Returns Data

Join collection of returned sales orders

Calculate

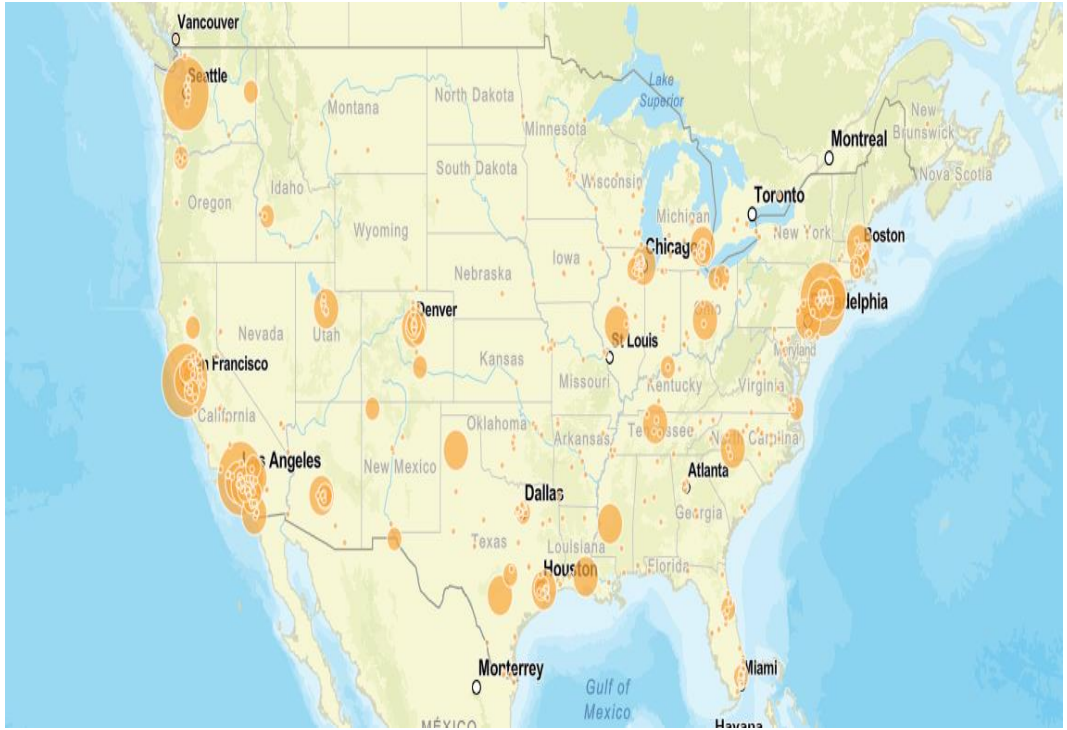
% return rate and returned value

Aggregate

At monthly level granularity

Exploratory data analysis

Business overview



55%+ discount

Average sales were highest when discount offered was greater than 55%

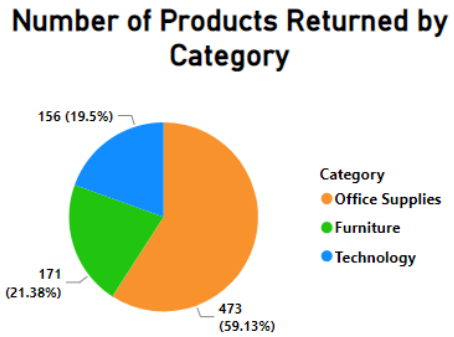
531 cities

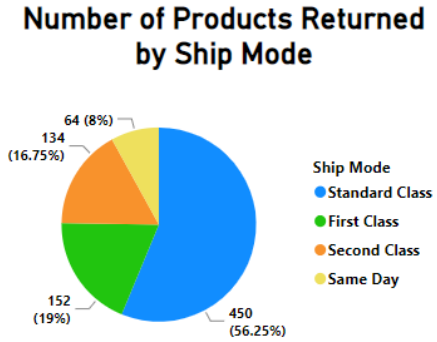
These cities were the main hubs of customers ordering products

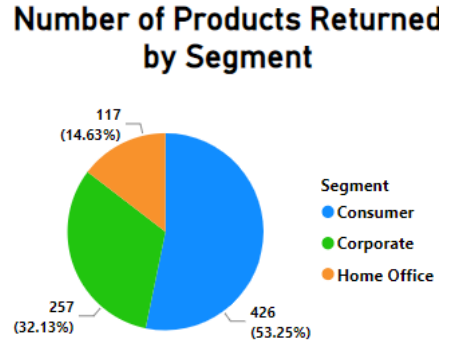
28% profit

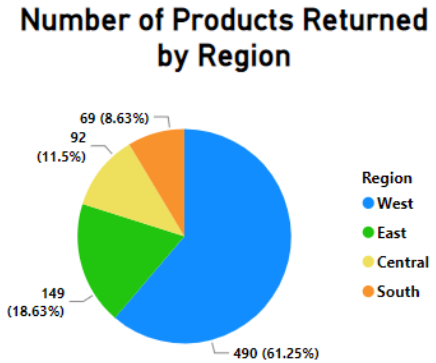
About 700 customers contribute to over 28% profit in sales through purchases

Returns overview

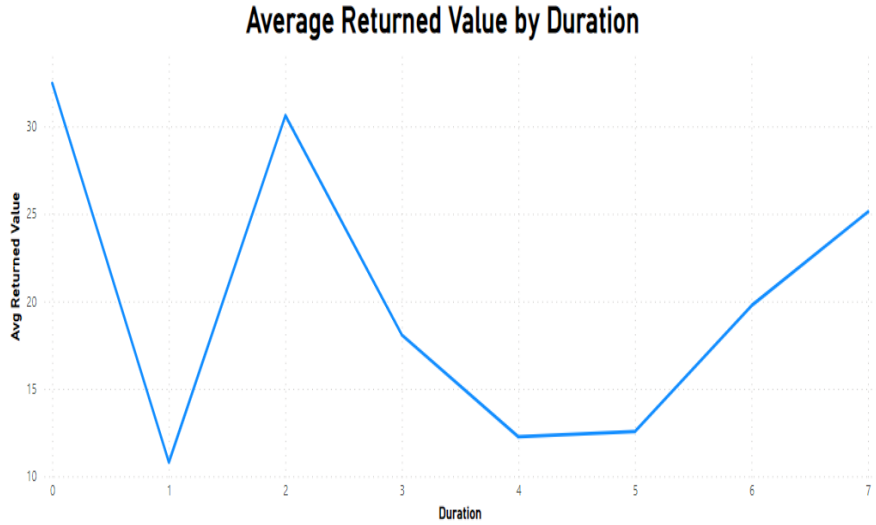






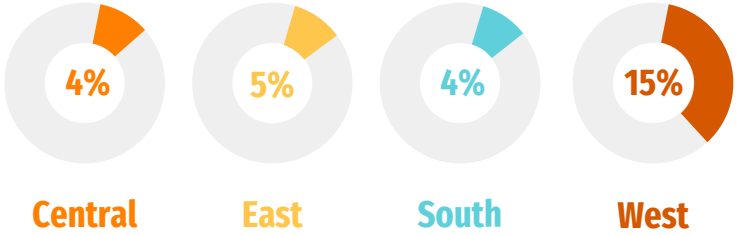


Delivery duration



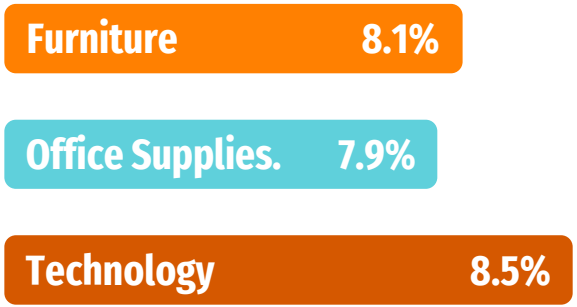
Returns rate varying by region and category

Region



West has more than 3x the average returns rate of other regions.

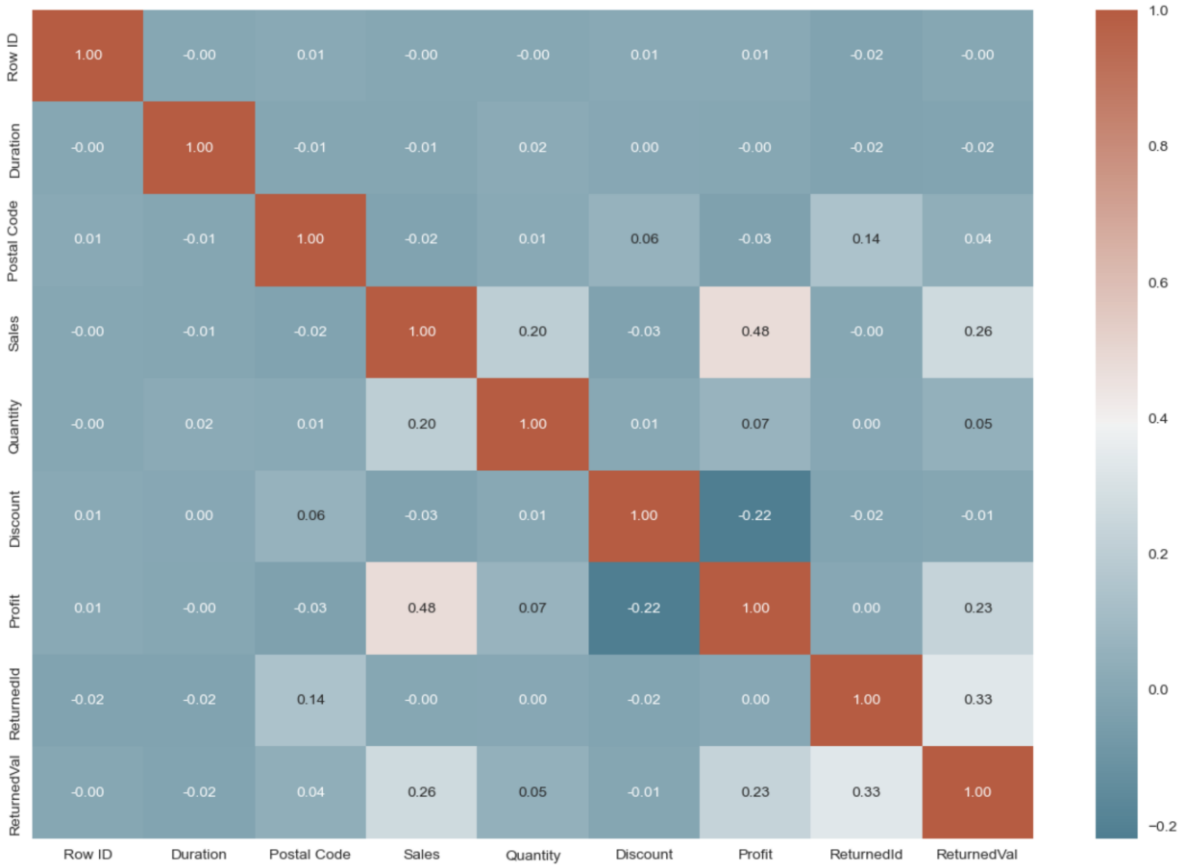
Category



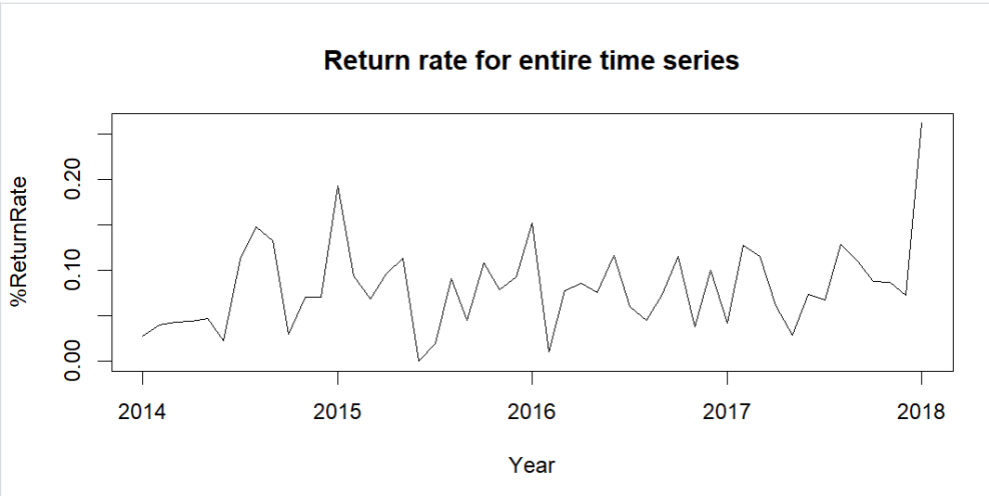
Technology, responsible for 51% of total profits, has the highest returns rate among categories.

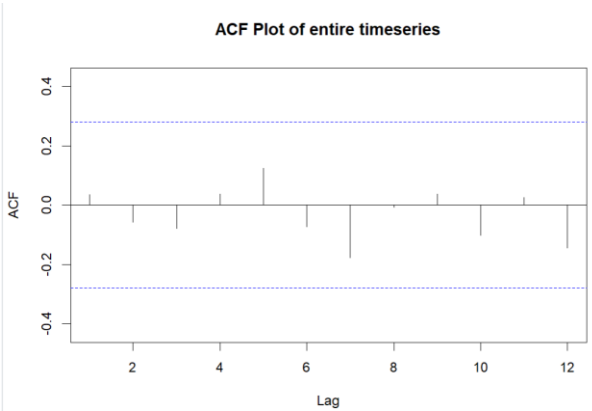
No significant correlation found

(>= 0.5 or <= -0.5)



Non-stationary time series observed





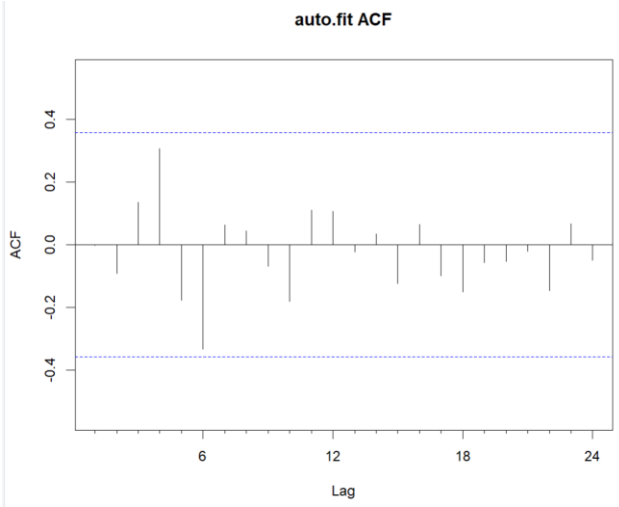
No significant ACF observed

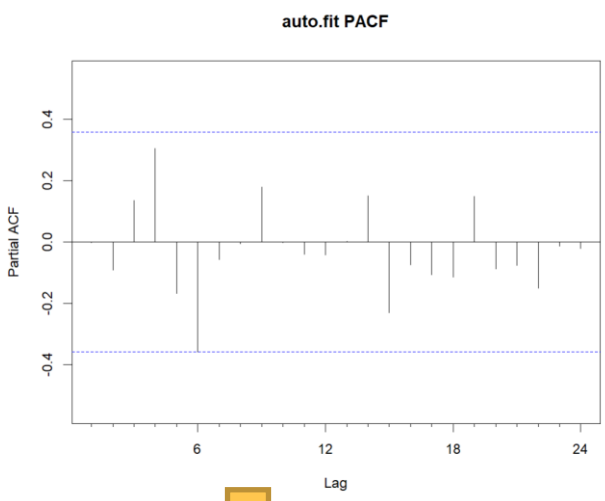
```
        Augmented Dickey-Fuller Test

data:  return.ts
Dickey-Fuller = -3.2748, Lag order = 3, p-value =
0.08608
alternative hypothesis: stationary
```

Non-Stationary

Auto Arima

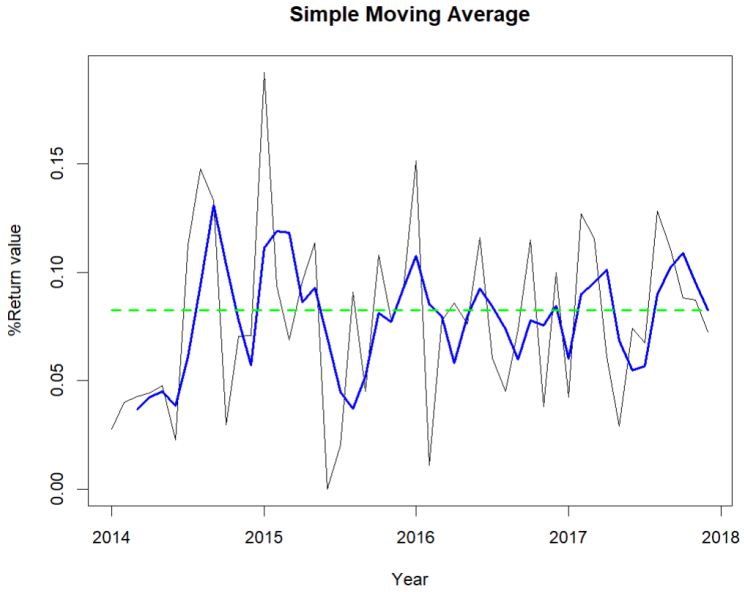


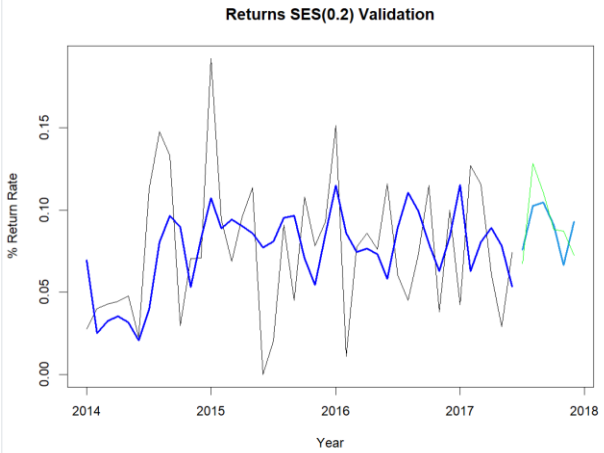


No significant autocorrelation factor observed

```
Training set error measures:
                     ME         RMSE          MAE  MPE
Training set 3.601599e-17 0.04156151 0.03359206 -Inf
                MAPE      MASE        ACF1
Training set     Inf 0.6409524 0.04702396
```

Moving on to smoothing models



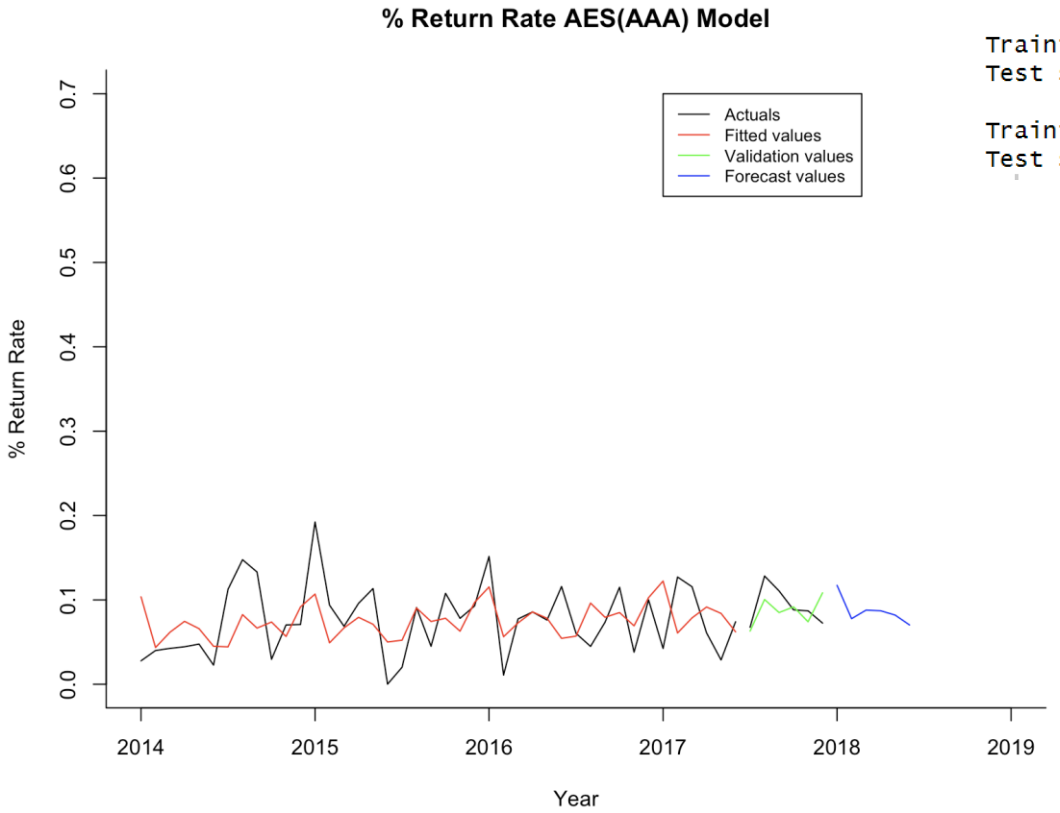


Models not giving best results

	ME	RMSE	MAE	MPE
Training set	0.002109871	0.04420767	0.03304653	-Inf
Test set	0.019561499	0.02880290	0.02140969	17.17331

	MAPE	MASE	ACF1	Theil's U
Training set	Inf	0.6305434	-0.01897547	NA
Test set	19.89829	0.4085071	-0.06877673	0.9487445

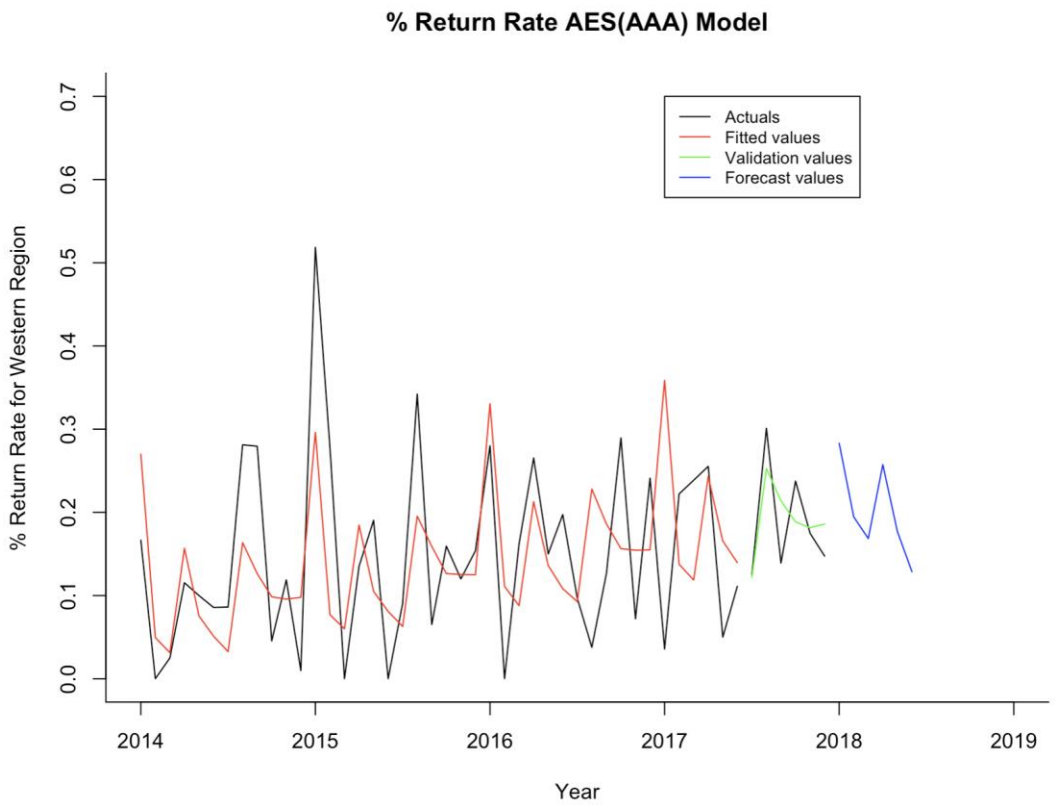
Finalized on Holt-Winters' method




```
                      ME      RMSE       MAE       MPE
Training set 0.003076962 0.04135453 0.03306079      -Inf
Test set     0.012088221 0.02014200 0.01678227 10.48281
                   MAPE      MASE        ACF1 Theil's U
Training set        Inf 0.6308155 -0.04376335        NA
Test set       16.95784 0.3202137 -0.32595043 0.6369183
```

Month-Year	Forecasted values
Jan-2018	0.1172
Feb-2018	0.0777
Mar-2018	0.0879
Apr-2018	0.0871
May-2018	0.0820
June-2018	0.0703

Particularly forecasting for West region



	ME	RMSE	MAE	MPE
Training set	0.005463227	0.10394683	0.08128112	-7220.1262
Test set	-0.003215744	0.04448113	0.03675545	-7.4059

	MAPE	MASE	ACF1	Theil's U
Training set	7258.95704	0.6398539	-0.1042406	NA
Test set	20.48257	0.2893429	-0.6066518	0.3664636

Month-Year	Forecasted values
Jan-2018	0.2930
Feb-2018	0.1980
Mar-2018	0.1727
Apr-2018	0.2567
May-2018	0.1797
June-2018	0.1223

Recommendations

Thank you!