2016

OAKLAND AIRPORT PARKING ANALYSIS

PROJECT BY:

MONIKA KATARIYA, TANVI KALEVAR, SESHI HARINATHAN, MANASWANI RAYAPPUREDDI, HONGLUN SU

UCONN SCHOOL OF BUSINESS | ADVANCED BUSINESS ANALYTICS AND PROJECT MANAGEMENT

Table of Contents

Abstract

In the United States, parking represents a $20 billion industry (National Parking Association, 2005), and research shows that a car is parked on average 90 percent of the time. To leverage this situation, parking industry should be built on intelligent data infrastructures and methodologies to explore.

Airport Parking Industry comprises of a major portion of parking industry in United States. In this thesis, we have provided a general approach to provide solutions to parking problems. First, stated is to study the car parking traffic as a function of arrival time. This is to estimate future predictions of the traffic. Second, build time series forecasting model to forecast revenue based on previous revenue, airport passenger traffic and weather forecasts. This model is compared with traditional forecasts only based on past revenue. Third, the competitor pricing and amenities analysis to recommend changes in pricing models.

The Goal

LAZ Parking is more than just the fastest growing and second largest parking company in the country, they believe that the work they do can really **"Make a Difference".** With this focus in mind, our team decided to approach the data in which they could leverage the data they have generated over the years.

Our goal of the project was to make a **"One stop analytical solution for Laz Parking".** It has the following functions:

- Reporting based on capacity, duration and revenue on daily, weekly and yearly basis
- Prediction of number of cars based on hourly traffic from historical data
- Forecasting revenue based on airport passenger traffic and weather forecasts
- Competitor analysis based on 7 competitors near airport

The Dataset

LAZ Parking shared three years (2013- 2015) of revenue data of **Oakland Airport Parking Lots**. The data comprised of four parking facilities namely:

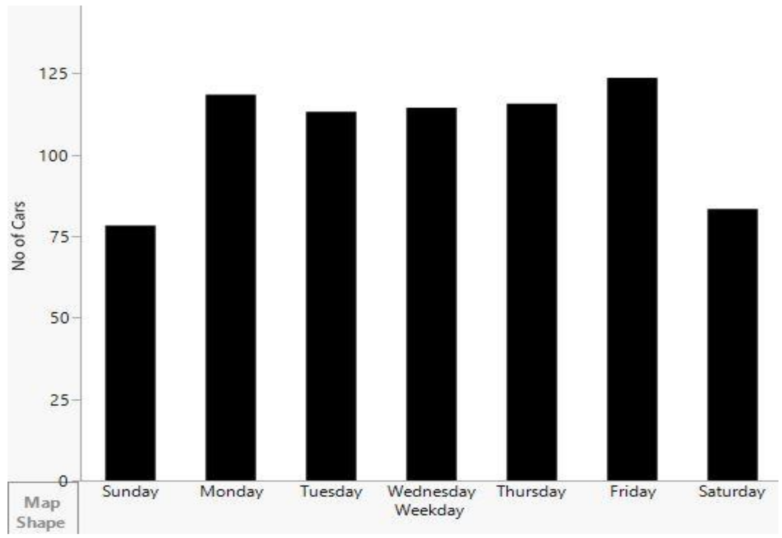
- **Hourly (4D),**
- **Daily (4A),**
- **Economy(4B) and**
- **Premium (4C)**

It had transactional information of each customer including details like entry, exit time, coupons used and license plate information. This data was further supported airport passenger traffic data at Oakland Airport from 1990-2008.

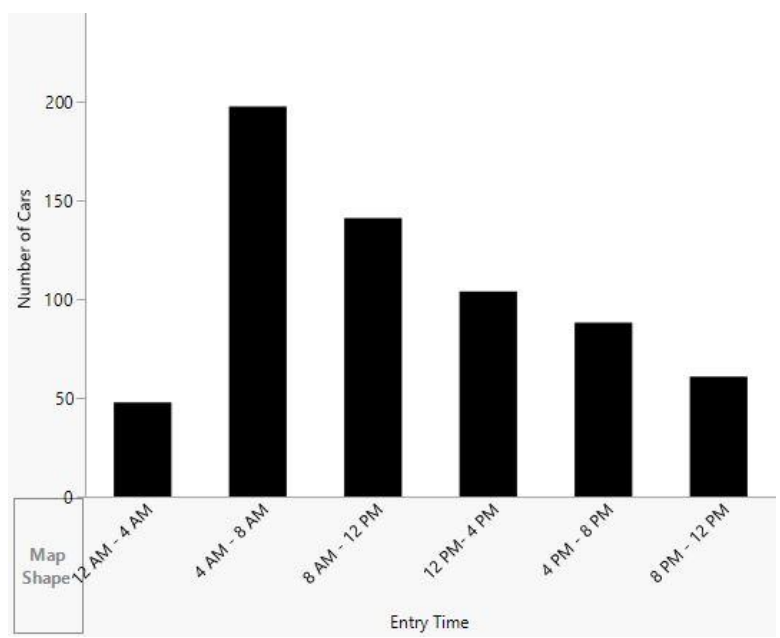
Hourly Prediction of Traffic using Decision Tree Algorithm

Our insights on visualizing the data is as follows:

- **Number of cars entering the parking facility is 78% more on weekday than on weekends.**
- **18% of total revenue comes on a Monday.**
- **A revenue of $1810/ day is generated more on a weekday as compared to weekend**



- On having divide the 24 hours of traffic into 6 bins of 4 hours each we realized, **31% of total cars entering the lots came between 4AM – 8 PM**.
- 4A had almost 50% of its revenue generated on weekdays than on weekends.
- **4B generated more than 60% of revenue on weekdays.**
- 4C has equivalent car traffic throughout.
- Friday generated most revenue for 4B lots specially in the evenings.
- 4A lots had almost 50% of traffic entering on Monday mornings whereas 4C lots had most revenue in the evenings.
- Weekend most of the traffic comes during 8 AM – 12 PM.



- **Friday gets most revenue even though more number of cars enter on Mondays.**

- Specific dates :

2012 : 20 November and 21 December (18% capacity full)

2013 : 10 June - 21 August (25% - 28% capacity full)

2014 : 25 May – Around 99% empty

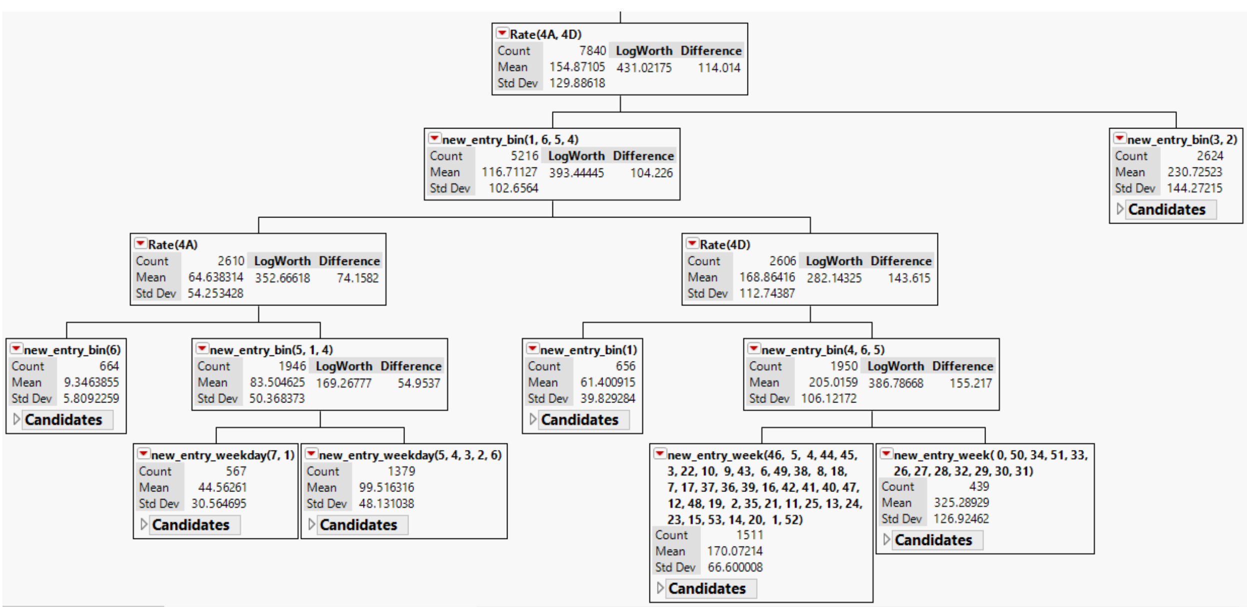
- July and August are busiest months whereas February and September are slow in terms of revenue.

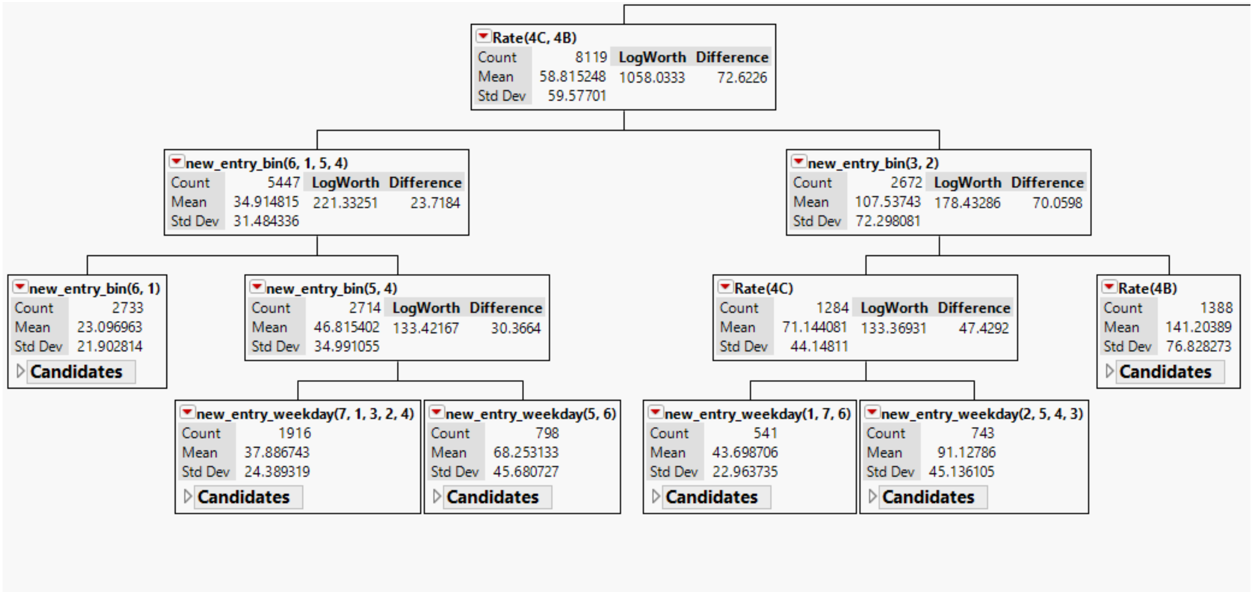
- **Third and twenty first week of the year are the busiest.**

Decision Trees Algorithm:

Based on the above insights, it can be concluded that the entry time bin, weekday or weekend and week of the year play a major role in traffic. Hence, based on these attributes we built a decision tree model with 12 nodes.

Following shows the description of the tree:





The leaf report is as follows:

Leaf Report

Leaf Label	Mean	Count
Rate(4C, 4B)^&new_entry_bin(6, 1)	23.096963	2733
Rate(4C, 4B)^&new_entry_bin(5, 4)&new_entry_weekday(7, 1, 3, 2, 4)	37.8867432	1916
Rate(4C, 4B)^&new_entry_bin(5, 4)&new_entry_weekday(5, 6)	68.2531328	798
^&new_entry_bin(3, 2)&Rate(4C)&new_entry_weekday(1, 7, 6)	43.6987061	541
^&new_entry_bin(3, 2)&Rate(4C)&new_entry_weekday(2, 5, 4, 3)	91.12786	743
^&new_entry_bin(3, 2)&Rate(4B)	141.20389	1388
^^&Rate(4A)&new_entry_bin(6)	9.34638554	664
^^&Rate(4A)&new_entry_bin(5, 1, 4)&new_entry_weekday(7, 1)	44.5626102	567
^^&Rate(4A)&new_entry_bin(5, 1, 4)&new_entry_weekday(5, 4, 3, 2, 6)	99.5163162	1379
^^&Rate(4D)&new_entry_bin(1)	61.4009146	656
^^&Rate(4D)&new_entry_bin(4, 6, 5)&new_entry_week(46, 5, 4, 44, 45, 3, 22, 10, 9, 43, 6, 49, 38, 8, 18, 7, 17, 37, 36, 39, 16, 42, 41, 40, 47, 12, 48, 19, 2, 35, 21, 11, 25, 13, 24, 23, 15, 53, 14, 20, 1, 52)	170.072138	1511
^^&Rate(4D)&new_entry_bin(4, 6, 5)&new_entry_week(0, 50, 34, 51, 33, 26, 27, 28, 32, 29, 30, 31)	325.289294	439
Rate(4A, 4D)&new_entry_bin(3, 2)	230.725229	2624

<u>Attributes included in the decision trees:</u>

Parking Lot

Month

Week of the Year

Weekday

Is Weekend ?

Time Bin:

1 – (Midnight – 4 AM),

2 – (4 AM – 8 AM),

3 – (8 AM – 12pm),

4 – (12 PM – 4 PM),

5 – (4 PM– 8 PM),

6 – (8 PM– 12 PM)

Following represent the accuracy of the model on training/validation data :

	RSquare	RMSE	N	Number of Splits	AICc
Training	0.559	73.931077	15959	12	182665
Validation	0.568	74.365172	10639		

Benefits to LAZ:

- Based on this system, LAZ can define strategies to face best/worst situation.
- This could be integrated with mobile applications to enhance the pre-booking services for LAZ.

Revenue Forecasts Using Time Series Forecasting

Forecasting revenue is of at most importance to understand the returns of Laz parking at Oakland airport.

Firstly, in order to forecast data, we prepared data and aggregated it to month to calculate net revenue of every month. Once we have the necessary data ready to be forecast, we employed the famous ARIMA models to predict revenue for months from May 2015 to 2017 May by adding appoint interactions at July 2013 and August 2013. We found out that the forecasts were decent however the confidence interval is large. We aim to perform more accurate predictions of revenue.

This reason lead to some secondary research. We gathered information about the traffic of Oakland airport with respect to number of passengers. Now that we have more information that contributes to predict net revenue more accurately. We built another ARIMA model with the same old point interactions but regressing with number of passenger. The results were much better with low confidence intervals.

Below is a comparison chart of net revenue with passengers and without passengers.

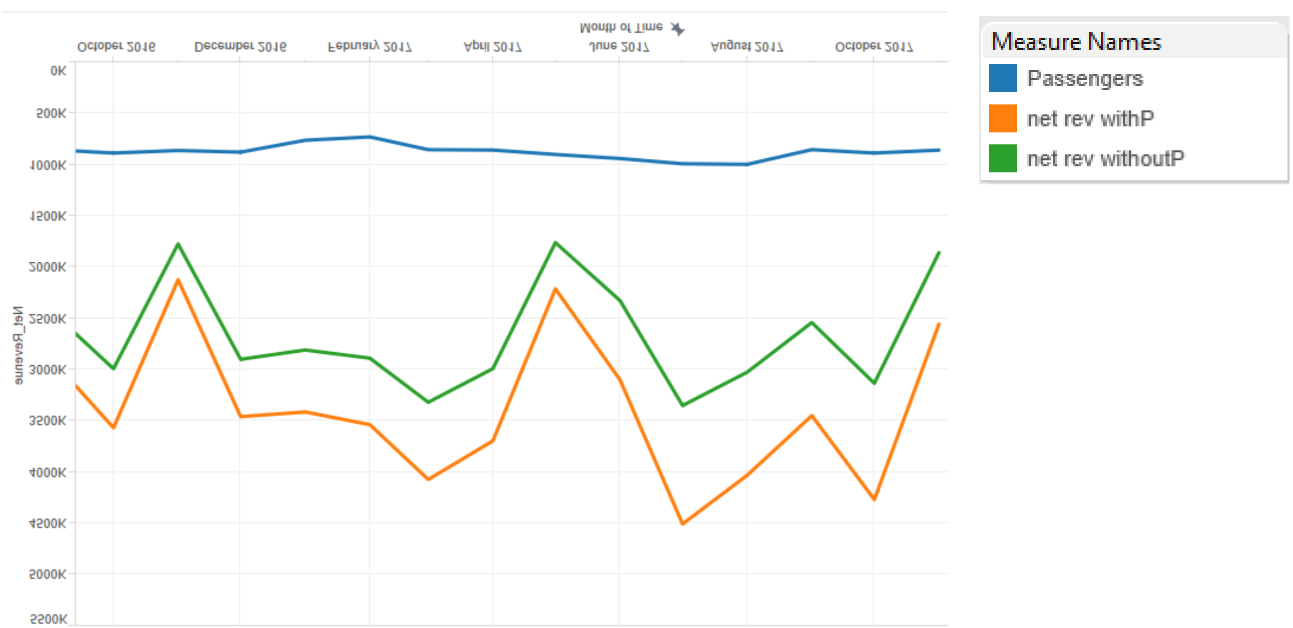


Fig: Net revenue with passengers and without passengers.

To improve our model, we also extracted weather data of all 2013 to 2015 and performed the same procedure to forecast net revenue. An interesting element in forecasting would be to combine both passenger information and weather information as regressors to forecast.

Below is a comparison chart of all the various models we built.

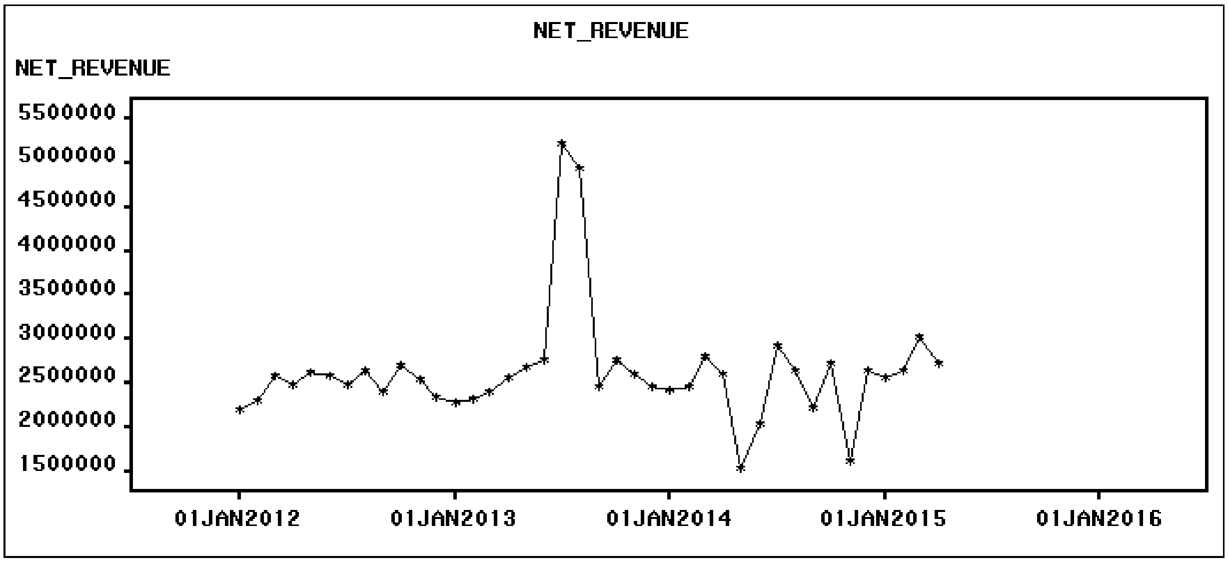
Fig: Comparison of all the models.

Results:

- $800K More Revenue Predicted for 2017 Based on Airport Passenger
- $137K More Revenue Predicted 2017 Based on Weather Forecasts
- $715K More Revenue Predicted 2017 Based on Airport Passenger and Weather Forecasts

MODEL

Trend chart of Net revenue over 2012 to 2013



Below is the autocorrelation plot where we find significant correlation at lag1

Analyzing the Autocorrelation Plots:

ACF: The autocorrelation function at lag k, ACF(k),represents the correlation of a time series with itself

lagged by k time units. The autocorrelation function is one of the primary tools for diagnosing trend, seasonality, and candidate forecast models.

PACF: The partial autocorrelation function at lag k, PACF(k), represents the correlation of a time series with itself at lag k adjusted for lags 1 through k-1. PACF(k) is the coefficient of the k-th order autoregressive term in an autoregressive order-k model. PACF(k) is calculated using a fast recursion equation

IACF: The inverse autocorrelation function at lag k, IACF(k), is the autocorrelation function of the inverse model. The IACF has the same interpretation as the PACF for diagnosing time series features. The IACF does not have an exact formula, so it can be different for different software or different settings.

There is a high autocorrelation in the lead production at lag1 and 2, and also the trend of auto correlation function is sinusoidal in nature and is not random.

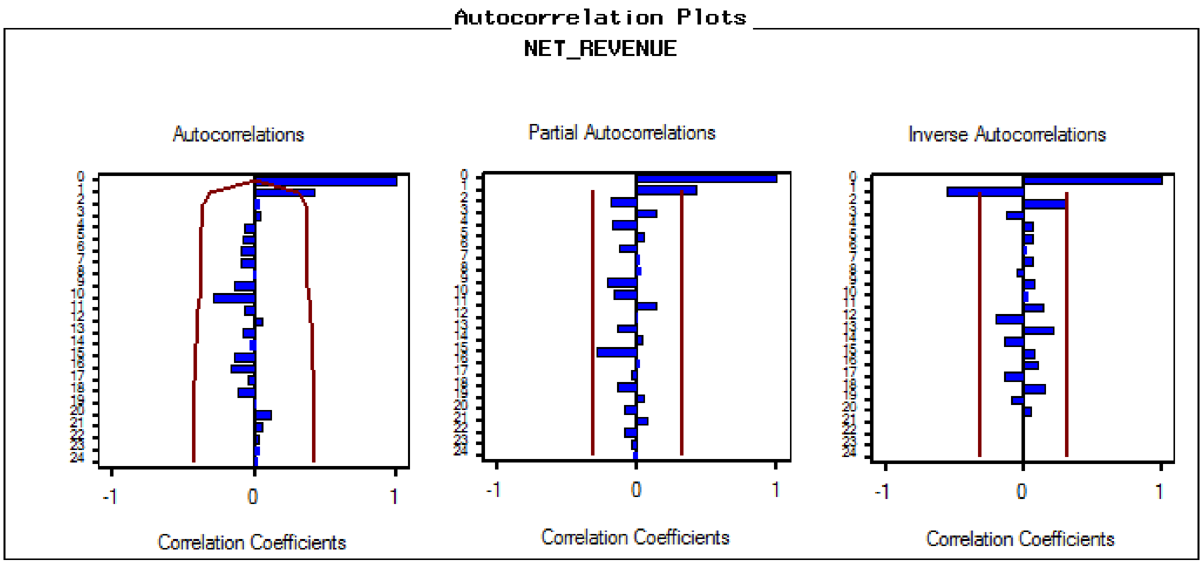


Fig: Auto correlation plots.

Analyzing the White Noise and Stationary tests:

White Noise test: Pass

Interpretation of White noise test:

Null hypothesis: Prediction errors represent white noise.

The white noise test bar chart shows significance probabilities of the Ljung-Box chi square statistic. Each bar shows the probability computed on autocorrelations up to the given lag. Longer bars favor rejection of the null hypothesis that the prediction errors represent white noise. In the below screenshot all the computed probabilities are insignificant leading to the fail to rejection of null hypothesis.

Interpretation of Stationarity tests:

Null Hypothesis: Series is non stationary.

The Unit root tests show the significance probabilities of the augmented Dickey-Fuller test for unit roots. As we can see that the computed probabilities for lag 0 and 1 are close to 0.01 so we reject the null hypothesis that the series is non stationary.

Unit Root test: Pass

Seasonal Root test: pass

Seasonal root tests also test for the stationarity in the residuals except that it represents the seasonal lags.

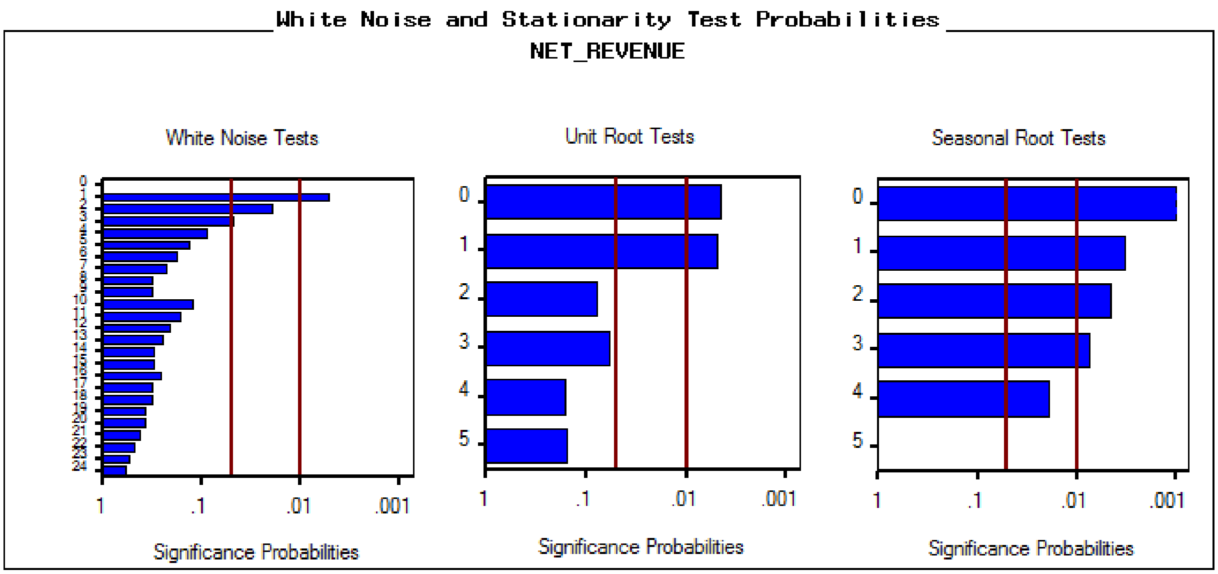
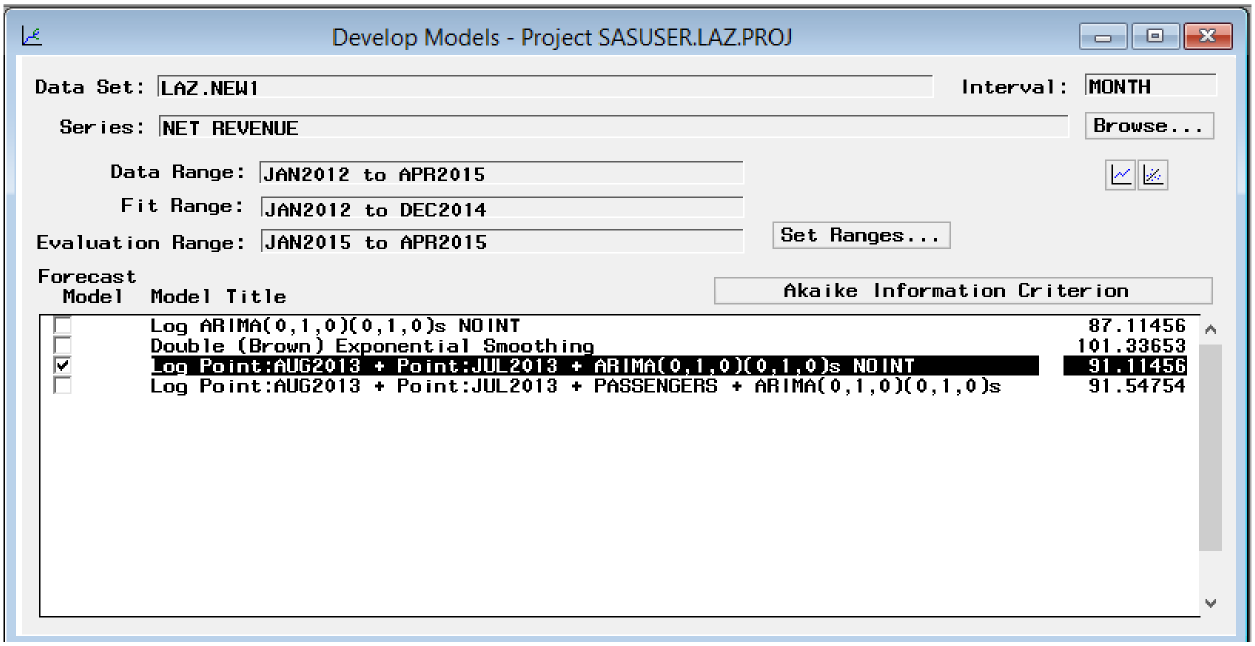
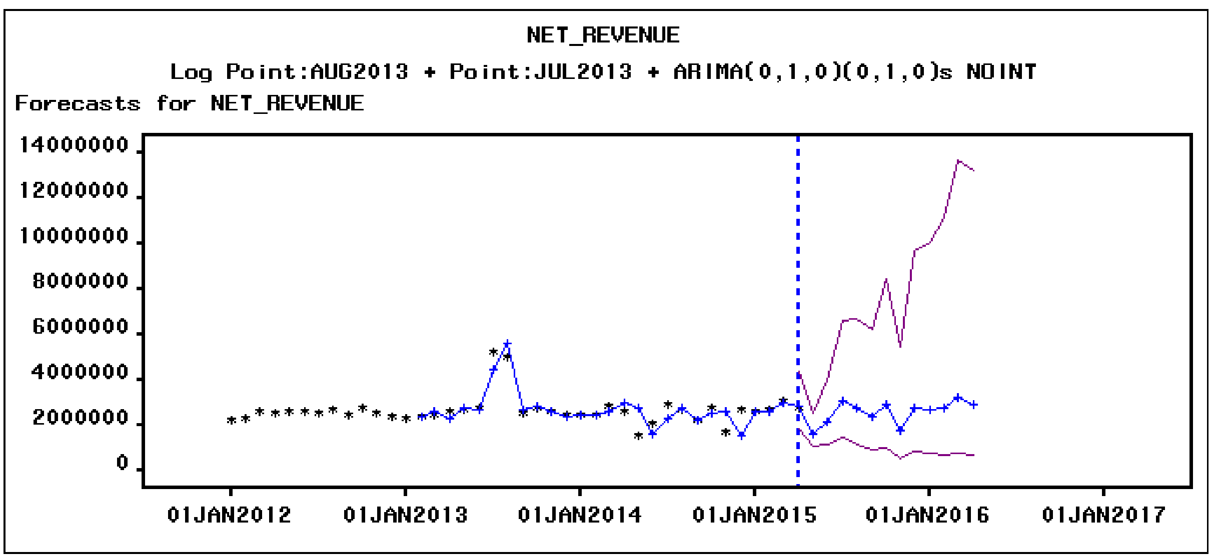
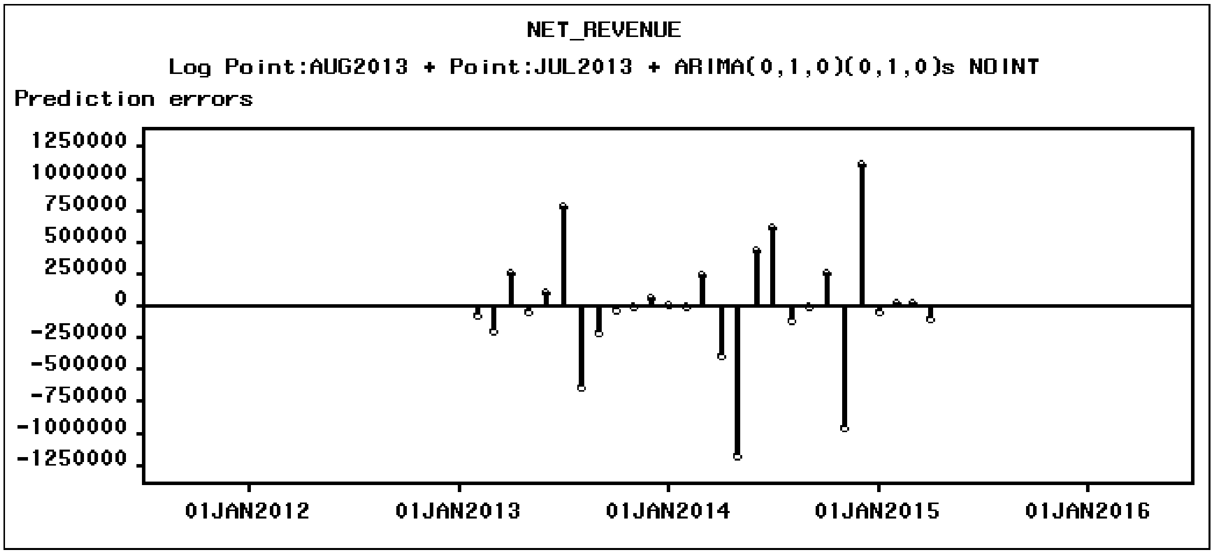


Fig: testing stationarity.

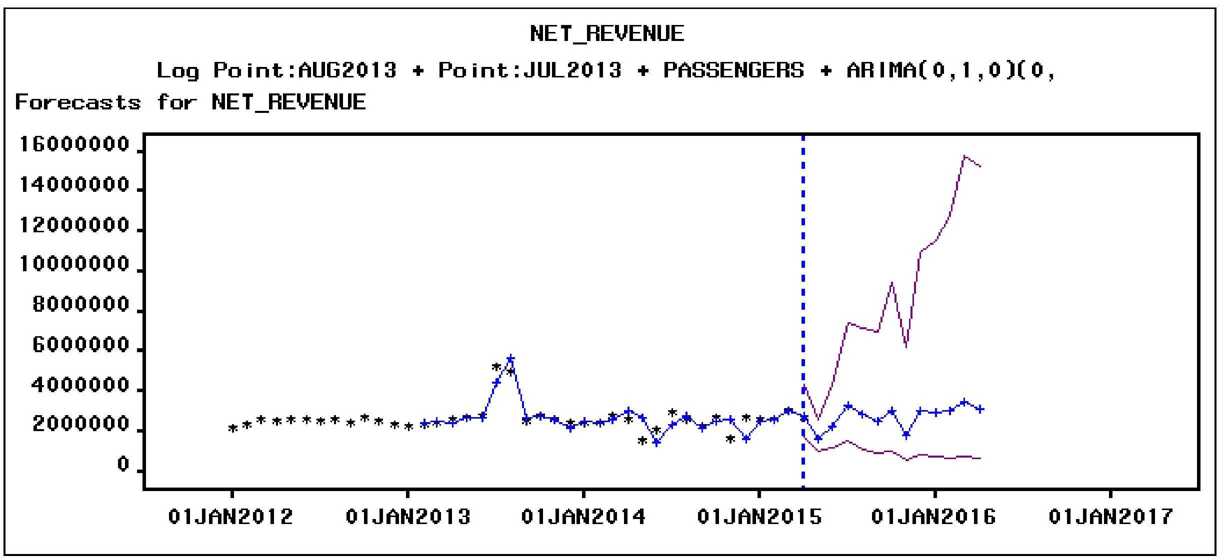


Without passengers

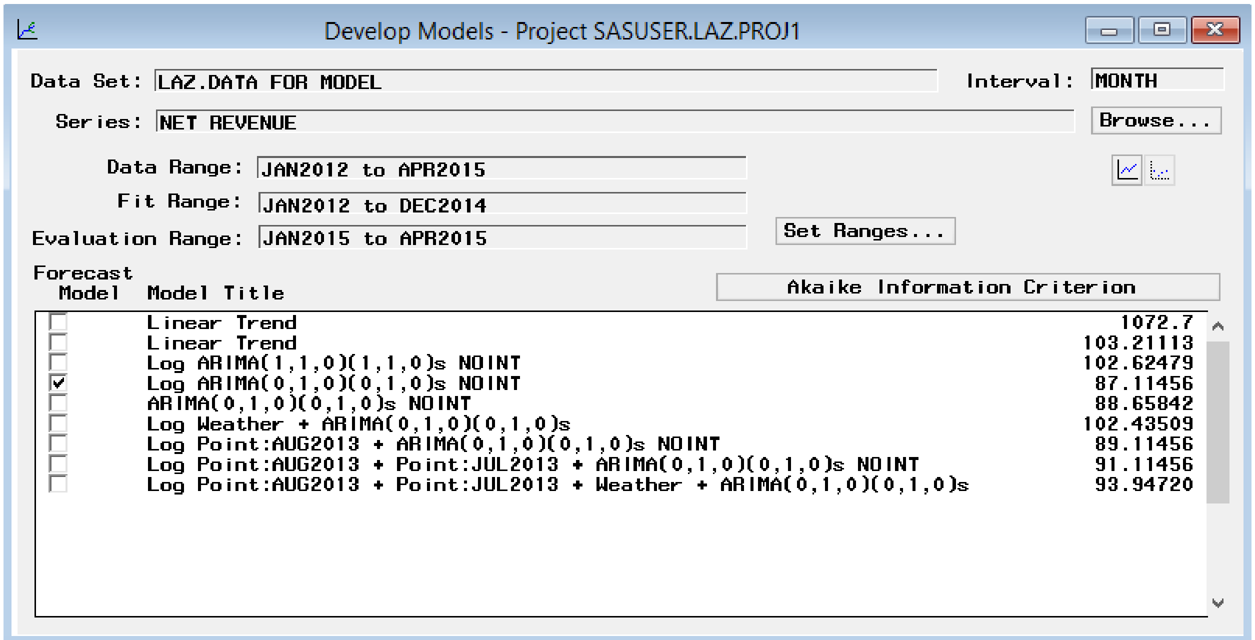




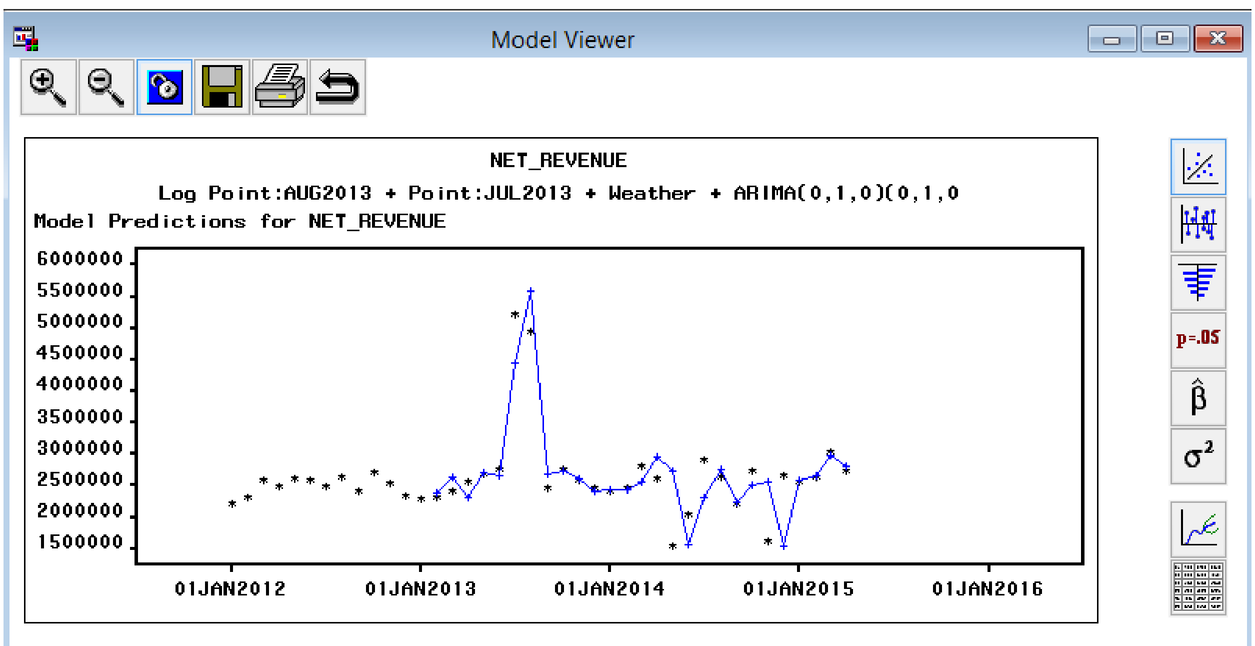
With Passengers:

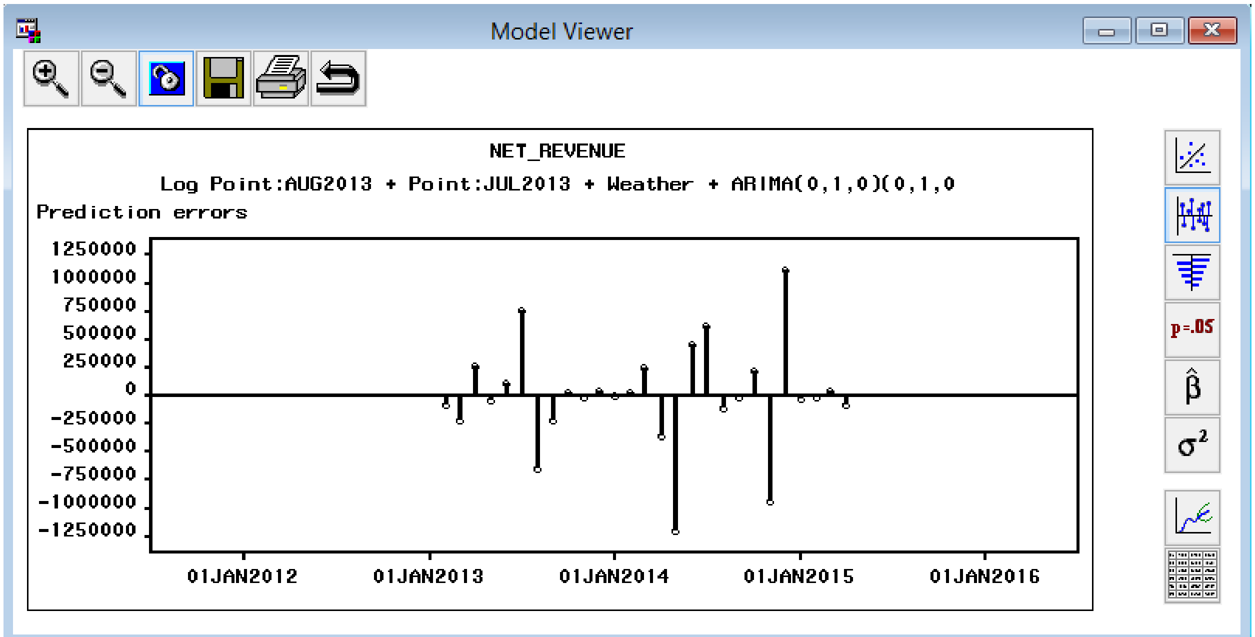


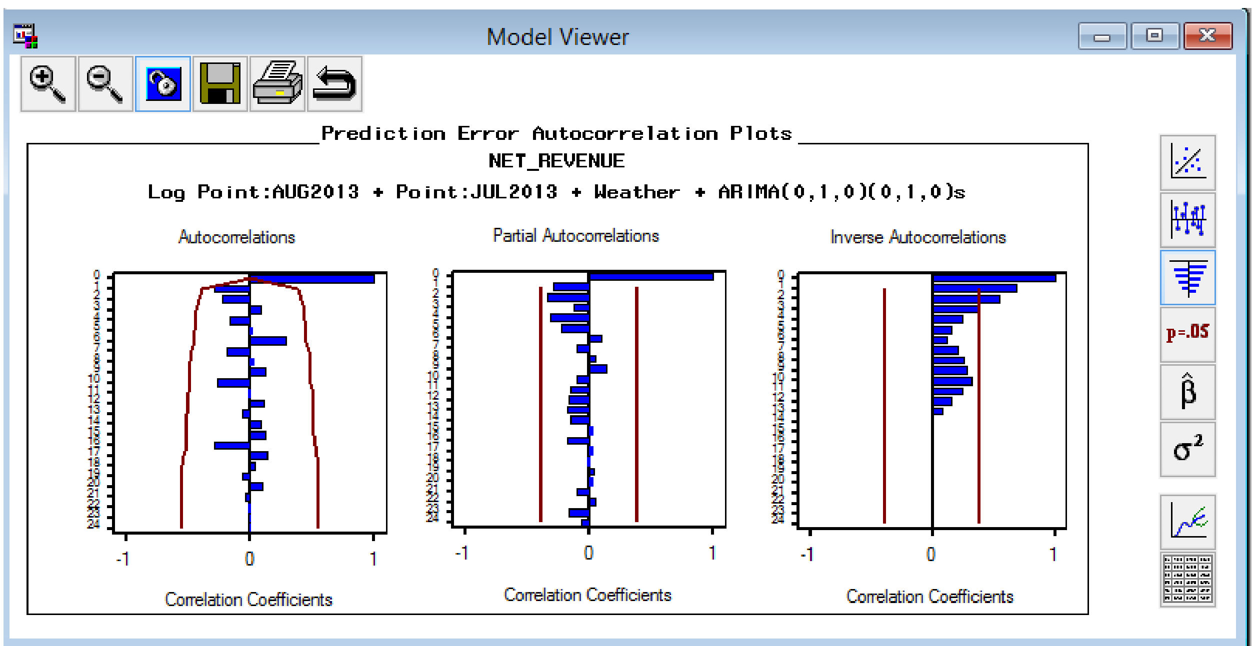
WEATHER

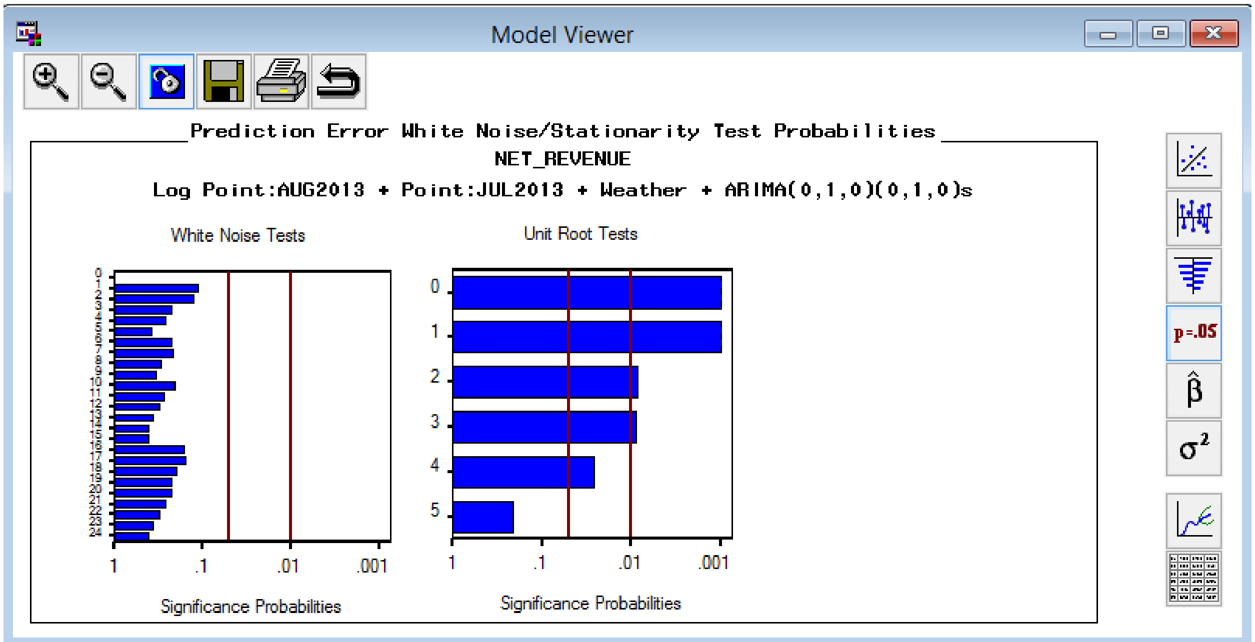


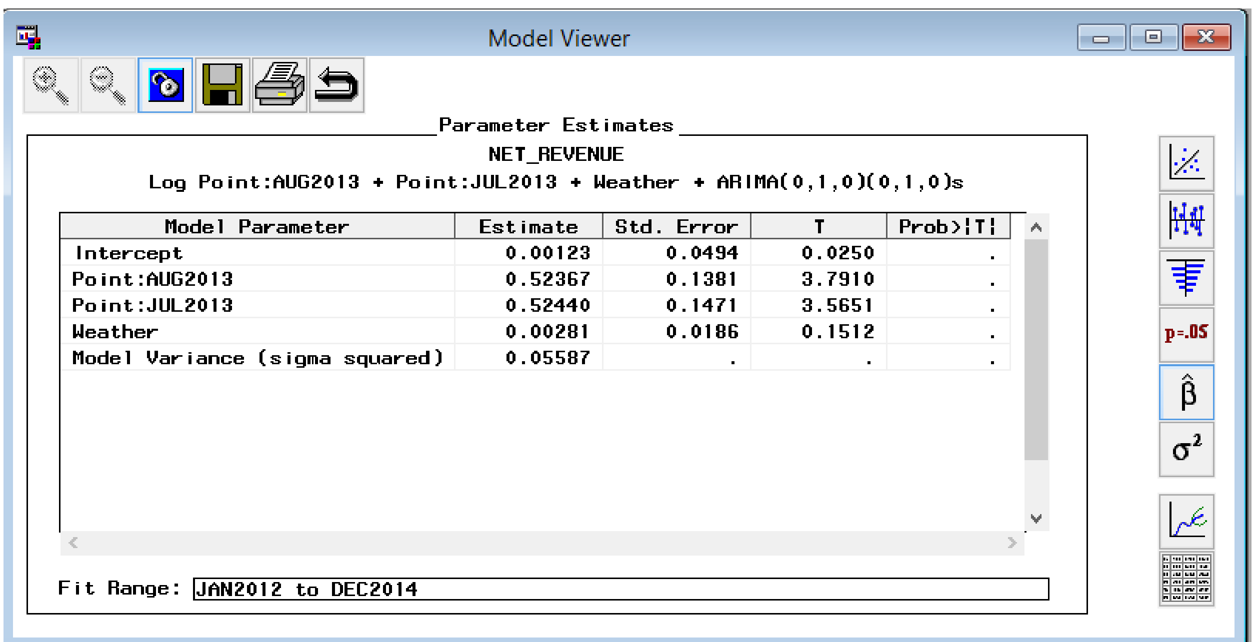
With weather model




Model Viewer
NET_REVENUE
Log Point:AUG2013 + Point:JUL2013 + Weather + ARIMA(0,1,0)(0,1,0
Prediction errors
1250000
1000000
750000
500000
250000
0
-250000
-500000
-750000
-1000000
-1250000
01JAN2012
01JAN2013
01JAN2014
01JAN2015
01JAN2016
p=.05

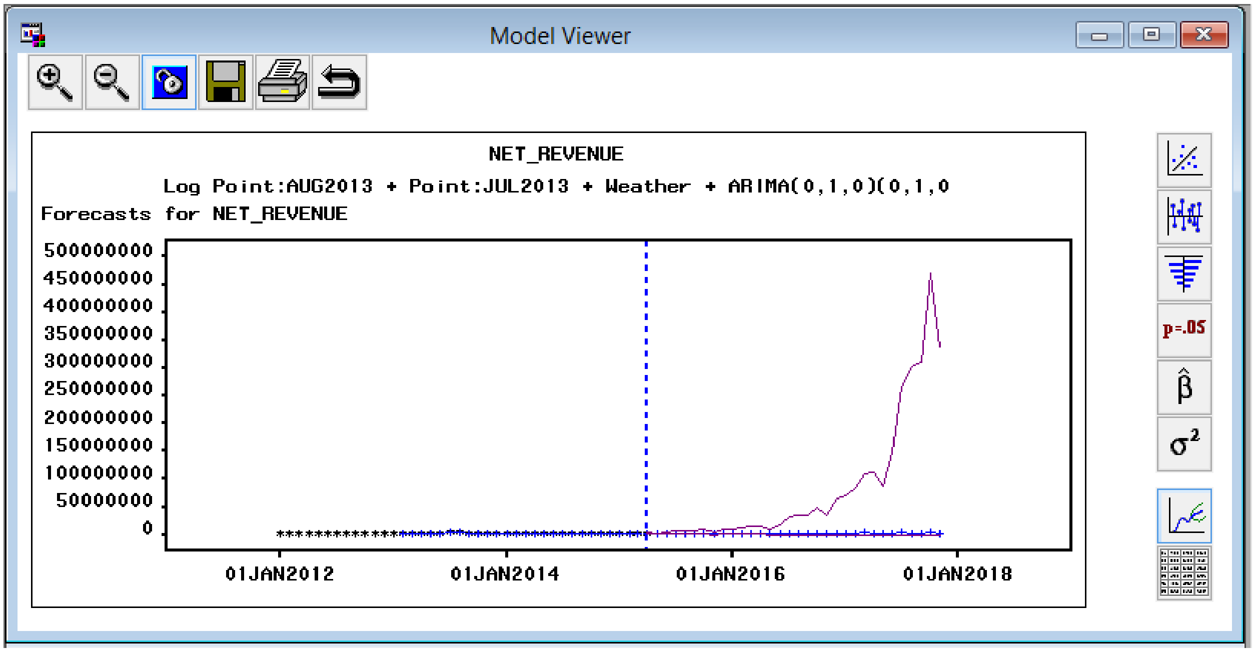

Model Viewer
Prediction Error Autocorrelation Plots
NET_REVENUE
Log Point:AUG2013 + Point:JUL2013 + Weather + ARIMA(0,1,0)(0,1,0)s
Autocorrelations
Partial Autocorrelations
Inverse Autocorrelations
-1
0
1
Correlation Coefficients
p=.05



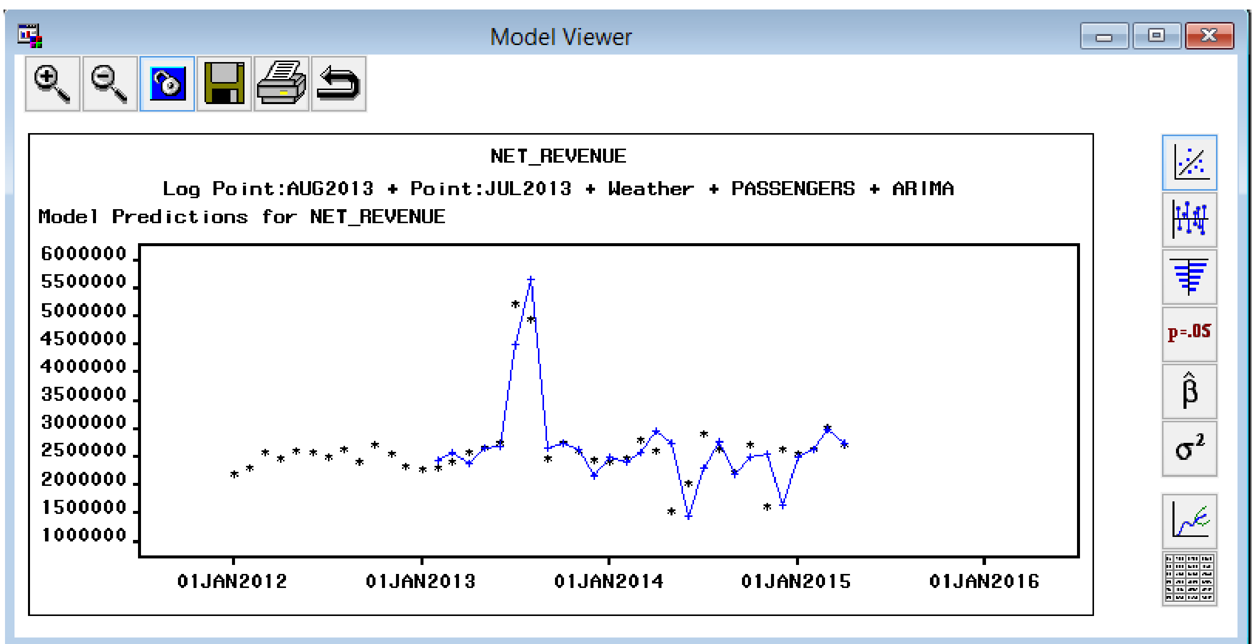


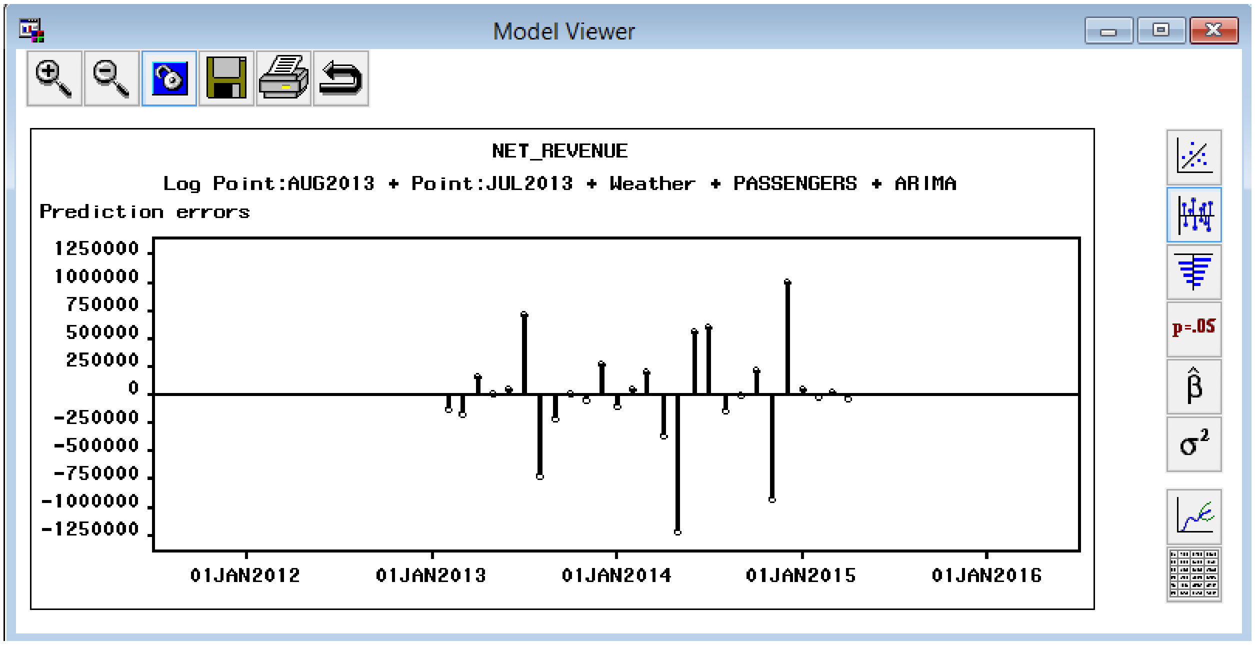
Model Parameter	Estimate	Std. Error	T	Prob>|T|
Intercept	0.00123	0.0494	0.0250	.
Point:AUG2013	0.52367	0.1381	3.7910	.
Point:JUL2013	0.52440	0.1471	3.5651	.
Weather	0.00281	0.0186	0.1512	.
Model Variance (sigma squared)	0.05587	.	.	.

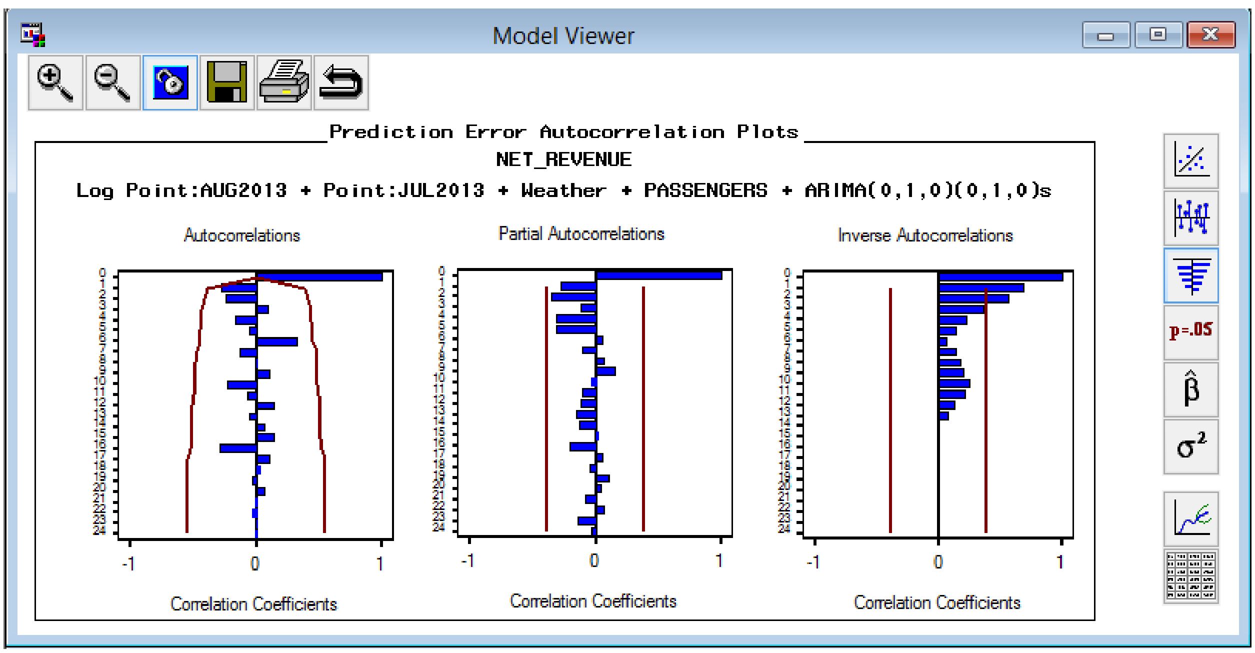
Fit Range: JAN2012 to DEC2014

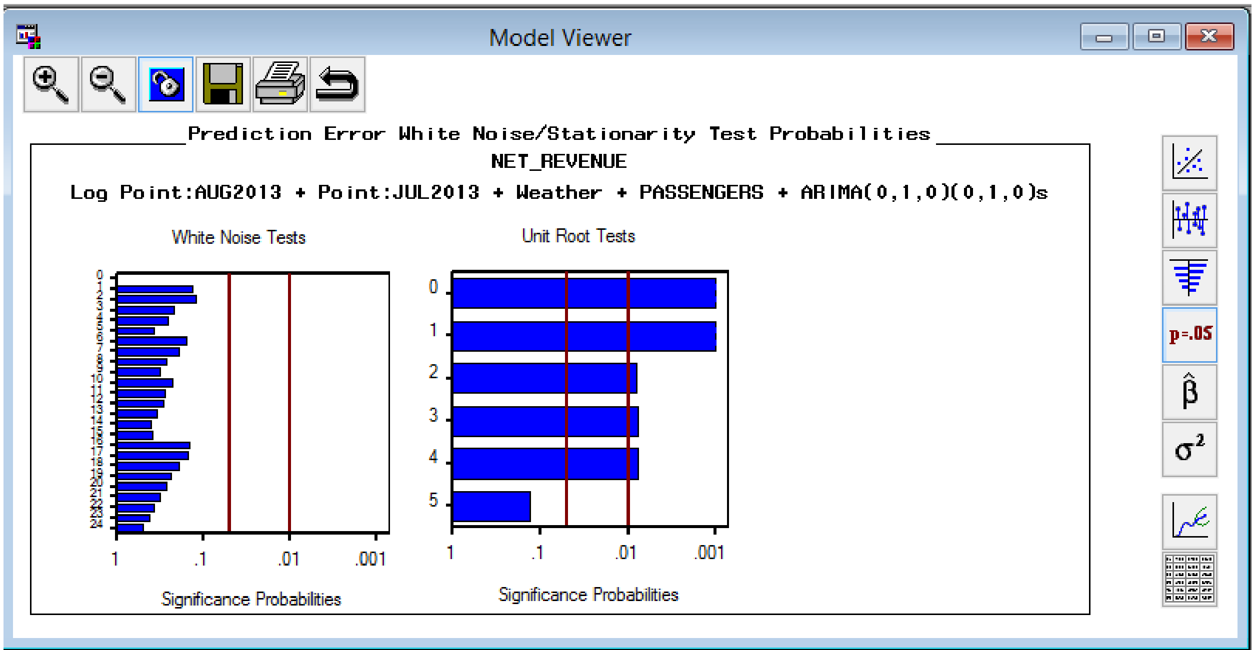


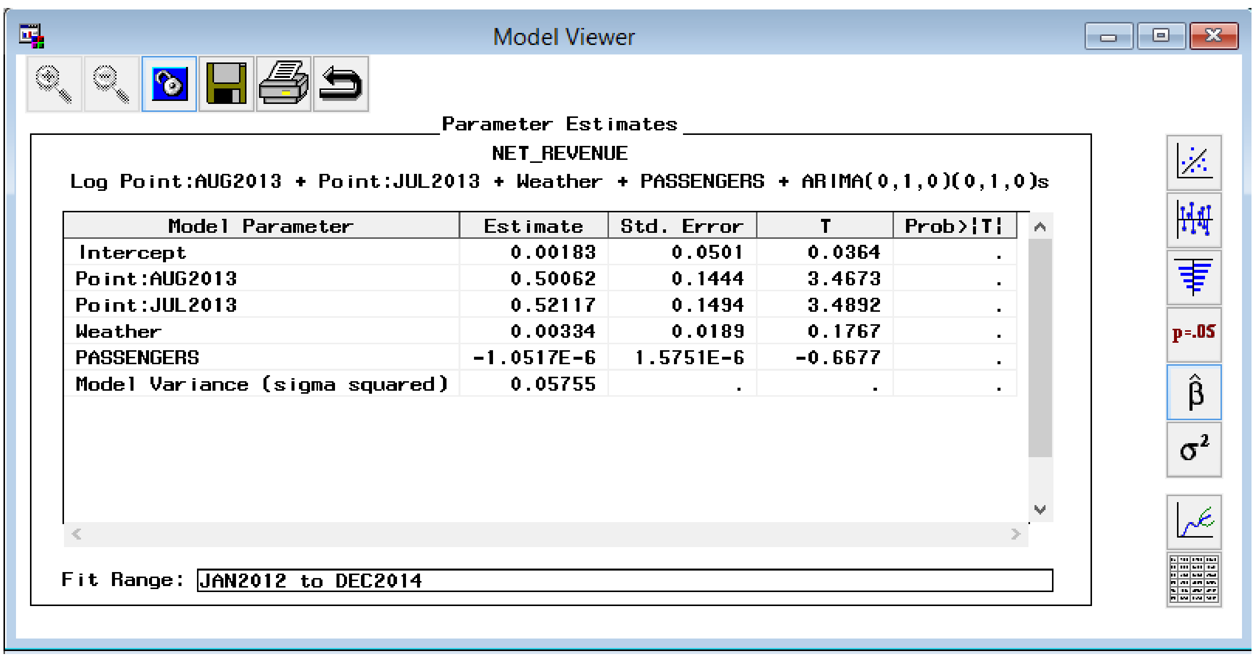
Both Passengers and weather:




Model Viewer
NET_REVENUE
Log Point:AUG2013 + Point:JUL2013 + Weather + PASSENGERS + ARIMA
Prediction errors
1250000
1000000
750000
500000
250000
0
-250000
-500000
-750000
-1000000
-1250000
01JAN2012
01JAN2013
01JAN2014
01JAN2015
01JAN2016
p=.05


Model Viewer
Prediction Error Autocorrelation Plots
NET_REVENUE
Log Point:AUG2013 + Point:JUL2013 + Weather + PASSENGERS + ARIMA(0,1,0)(0,1,0)s
Autocorrelations
Partial Autocorrelations
Inverse Autocorrelations
Correlation Coefficients
Correlation Coefficients
Correlation Coefficients
p=.05





Model Viewer

Parameter Estimates

NET_REVENUE

Log Point:AUG2013 + Point:JUL2013 + Weather + PASSENGERS + ARIMA(0,1,0)(0,1,0)s

Model Parameter	Estimate	Std. Error	T	Prob>|T|
Intercept	0.00183	0.0501	0.0364	.
Point:AUG2013	0.50062	0.1444	3.4673	.
Point:JUL2013	0.52117	0.1494	3.4892	.
Weather	0.00334	0.0189	0.1767	.
PASSENGERS	-1.0517E-6	1.5751E-6	-0.6677	.
Model Variance (sigma squared)	0.05755	.	.	.

Fit Range: JAN2012 to DEC2014

data laz.for_with_weather_and_passenger;

set work.forecast;

run;

```
PROC SQL;

create table x as

select a.time,a.net_rev_withoutP,a.net_rev_withP,a.Passengers ,

b.net_revenue as net_rev_withWeather,b.weather,c.net_revenue as net_rev_withWP

from laz.new4 a,laz.for_with_weather b,laz.for_with_weather_and_passenger c

where a.time= b.time and b.time=c.time;

quit;

data laz.Final_Model_Tab;

set x;

run;
```

Competitor Analysis:

1) Comparison based on the different transportation systems:

We have compared the average prices of the different transportation systems such as BART, Uber, Lyft with the rates of LAZ parking. The below table shows the details of this analysis. We have considered three LAZ scenarios for this comparison namely four-hour parking, daily parking and the three day – weekend parking. We have calculated the fares of the different transport systems to nearby places to the airport such as Richmond, Millbrae, Pittsburg, Dublin and Freemont. From this analysis, we can see that LAZ parking is cheaper for hourly and daily parking whereas it is a bit expensive for weekend parking. The LAZ parking rates are considered for the economy lots and the two-way charges are considered for the transportation systems. We have assumed that only one person is traveling and the average UberX and standard Lyft rates are considered.

Price Comparision (Average)	BART	Uber	LYFT	LAZ Scenario 1 (4hours) : $20	LAZ Scenario 2 (1 Day) : $16	LAZ Scenario 3 (3 Days) : $48
Richmond	$ 18.80	$ 58.00	$ 54.00	$ (23.60)	$ (27.60)	$ 4.40
Millbrae	$ 22.90	$ 86.00	$ 72.00	$ (40.30)	$ (44.30)	$ (12.30)
Pittsburg	$ 22.20	$ 100.00	$ 96.00	$ (52.73)	$ (56.73)	$ (24.73)
Dublin	$ 19.90	$ 50.00	$ 46.00	$ (18.63)	$ (22.63)	$ 9.37
Freemont	$ 19.90	$ 58.00	$ 52.00	$ (23.30)	$ (27.30)	$ 4.70

2) Comparison based on the different parking companies:

We have compared the various parking companies which are within two-mile radius from the Oakland airport. The below bar graph shows the daily parking rates of the economy lot for these parking companies. This graph shows that there is only

competitor namely Park N Fly which is more expensive than LAZ. Whereas the others provide cheaper parking rates than LAZ parking. These competitors provide 15% to 28% cheaper parking rates than LAZ. From this analysis, we recommend LAZ to build pricing models based on the competitor prices.

3) Comparison based on the amenities provided by different parking companies:

We have compared the amenities offered by the parking companies and the results are displayed below. The table shows that the competitors provide services such as car wash and oil change services, online parking reservation and complimentary luggage assistance to their customers. These services are not being offered by LAZ parking and we recommend that LAZ needs to focus on these services in order to attract more new customers. The services such as Loyalty programs which include the Frequent parker program, Corporate discount program and other coupons and discounts based on airline partnership deals are being offered by all the parking companies.

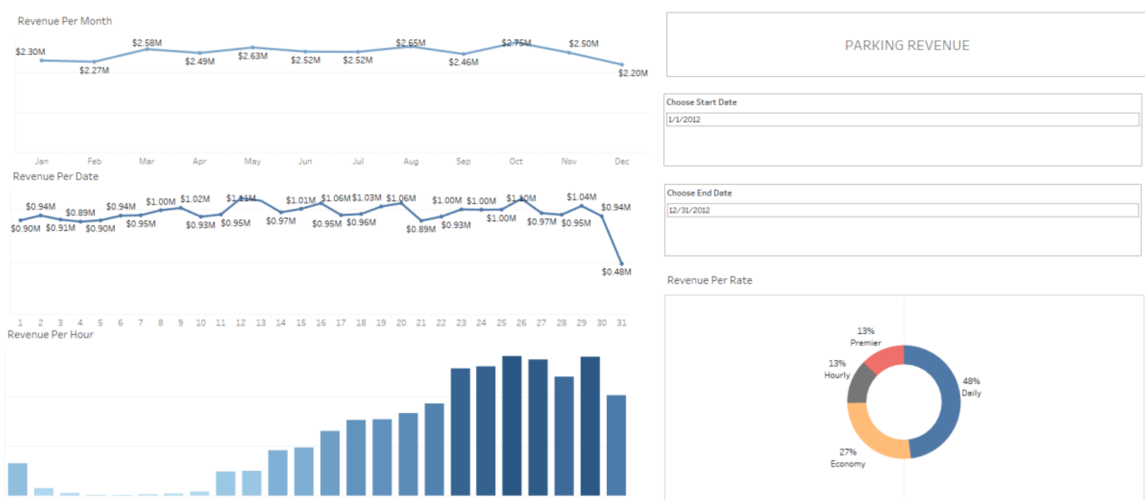
	AIRPARK airport parking	Expresso Airport Parking	FASTRACK AIRPORT PARKING	PARK 'N FLY More than just parking.™	LAZ PARKING
Car wash & oil change	★	★		★	
Online reservation	★	★	★	★	
Luggage assistance			★	★	
Loyalty programs	★	★	★	★	★
Coupons & discounts	★	★	★	★	★

Tableau Dashboard Implementation:

After the data was clean, we pulled the data into Tableau to look for patterns and observations. For all the dashboards, we added time filter and other parameters to help our sponsor simply change the time range or dimensions depending on what they need and do the analysis.

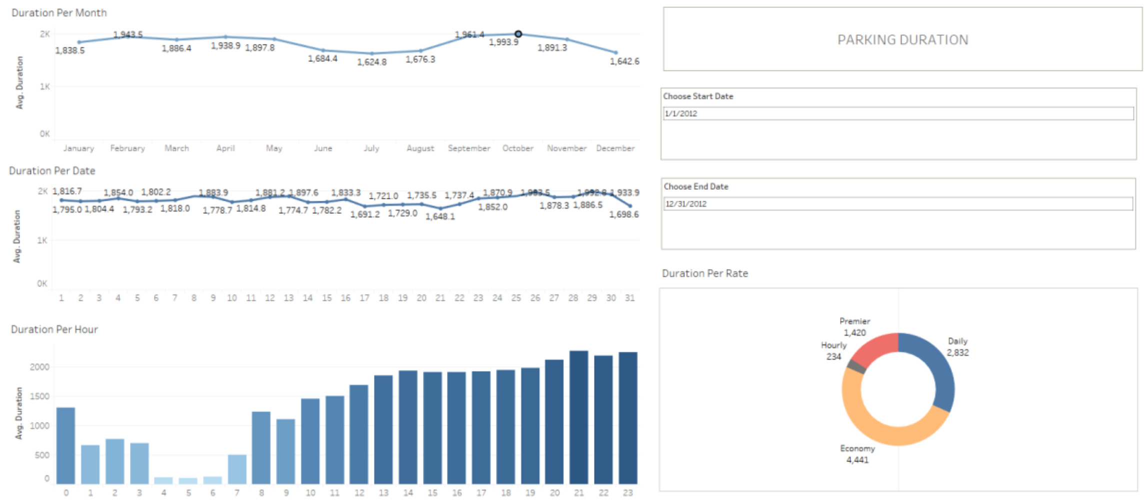
Parking Revenue Dashboard

Since the revenue is our main objective, we built our first dashboard with parking lots revenue sheets base on monthly, daily and hourly. Also, we added a donut chart showing the differences between four kinds of parking lots. It is obvious to see that the most revenue is coming between 4 pm to midnight of a day, and winter season has the lowest revenue of each year.



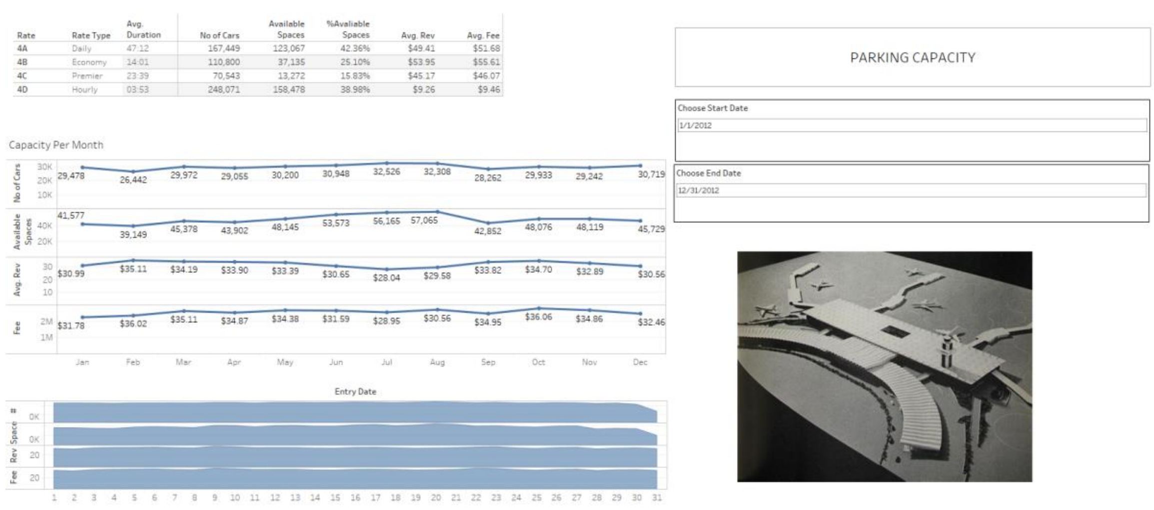
<u>Parking Duration Dashboard:</u>

Then, we built the duration dashboard which used the same way with revenue. We want to seek the average duration base on time and different parking lots. It was just as we expected that spring and autumn has the most average duration of the year. On the contrary, winter as the lowest average duration time. In addition, both of these two dashboards indicated that economy and daily parking lots has much more revenue and average duration time than hourly and premier parking lots.



Parking Capacity Dashboard:

After that, we built our parking capacity dashboard which showed the capacity monthly and daily to help audiences to understand how is the trending of the entire year. Different with the previous dashboards, we collected the exactly numbers of average duration time, numbers of cars, spaces, revenues and fees to complete the capacity sheet.



Forecasting Dashboard:

After completing these approaches, we met again to discuss our next steps. With the results of our revenue forecast from January 2015 to December 2017 with different weather and passengers options, we decided to show it on our tableau file and to see what we can explore more. What we did is build a revenue forecast based on month and the number of passenger trends of the entire year.

Competitor Analysis:

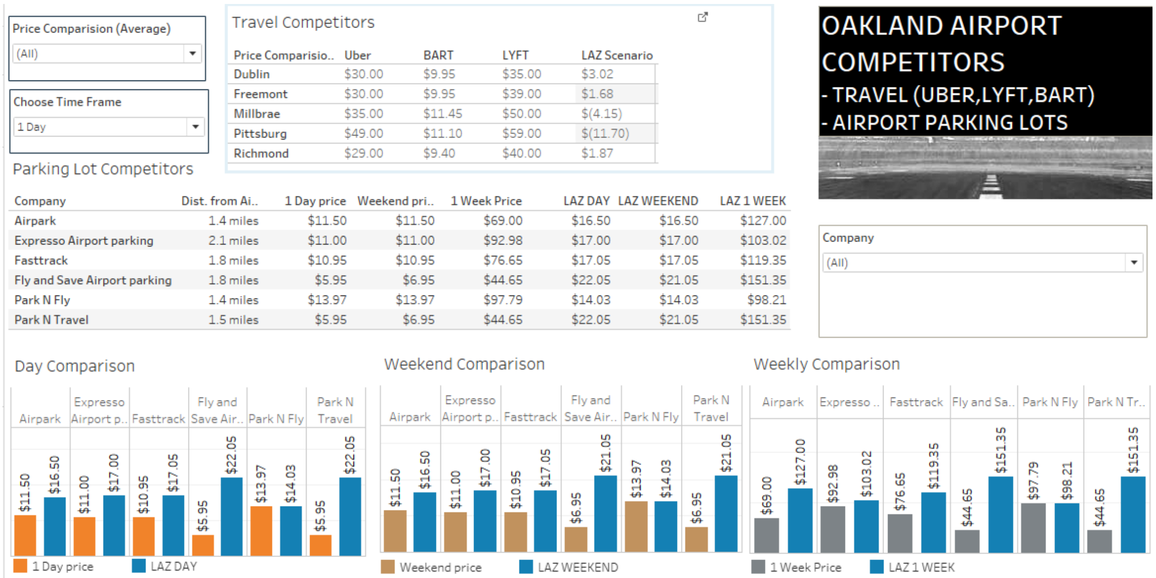
After we implemented our forecasting model into our Tableau dashboards, we decided that implementing our decision trees and Competitor analysis would be the perfect package for LAZ to understand the reason behind each decision they made and prepare for the same in the most strategic manner, thus saving costs and increasing revenue.

We then created the Competitor Analysis Dashboard. This dashboard gave us all the features and data we needed to approach a decision regarding our competitors. To start with we had divided our analysis on competitors on :

1. Travel Competitors: Uber, BART and LYFT
2. Parking Lot Competitors: Airpark, Expresso Parking, Fast Track, Fly and Save, Park N Fly and Park N Travel.

It was interesting to see how each of these competitors stood against LAZ in their pricing models and amenities. Our Primary Filter for this dashboard were – Price Comparison between Laz and a travel competitor. This filter changed the graphs below to show how LAZ's pricing model differed from the travel at the chosen time range. However, the Company Filter showed the difference between LAZ and the chosen

Parking Facility for the Airport. This helped us understand on which occasions LAZ was beneficial to a customer and when they were, whether the weekend, weekday or a longer duration mattered in the total pricing for a customer.



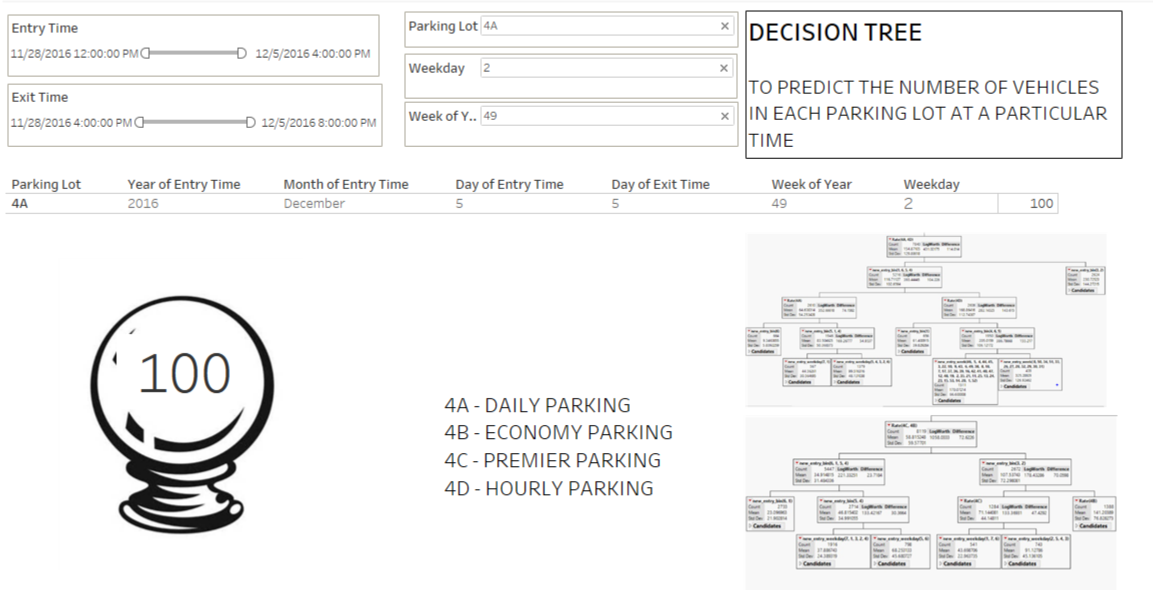
Price Comparisio..	Uber	BART	LYFT	LAZ Scenario
Dublin	$30.00	$9.95	$35.00	$3.02
Freemont	$30.00	$9.95	$39.00	$1.68
Millbrae	$35.00	$11.45	$50.00	$(4.15)
Pittsburg	$49.00	$11.10	$59.00	$(11.70)
Richmond	$29.00	$9.40	$40.00	$1.87

Company	Dist. from Ai..	1 Day price	Weekend pri..	1 Week Price	LAZ DAY	LAZ WEEKEND	LAZ 1 WEEK
Airpark	1.4 miles	$11.50	$11.50	$69.00	$16.50	$16.50	$127.00
Expresso Airport parking	2.1 miles	$11.00	$11.00	$92.98	$17.00	$17.00	$103.02
Fasttrack	1.8 miles	$10.95	$10.95	$76.65	$17.05	$17.05	$119.35
Fly and Save Airport parking	1.8 miles	$5.95	$6.95	$44.65	$22.05	$21.05	$151.35
Park N Fly	1.4 miles	$13.97	$13.97	$97.79	$14.03	$14.03	$98.21
Park N Travel	1.5 miles	$5.95	$6.95	$44.65	$22.05	$21.05	$151.35

The purpose of this dashboard was to help Laz understand why they must have coupons, variations in prices at the different durations of time so that they stand as a strong candidate with their competitors.

Decision Trees:

Next, we went on to implement our decision tree. This was indeed the most time-consuming dashboard created. First, we had to aggregate data for two years for different bins of days for each day and collect the predicted number of cars depending on the bin chosen and the parking lot chosen in the dashboard.



This dashboard showed the user the multiple slits that our model had been created to show the predicted cars depending on what the user wanted. The filters are primary here to give the user the options of:

1. Entry Time and Day
2. Exit Time and Day
3. Parking Lot
4. Weekday Number
5. Week of the year Number

Once the user had inputted the values the splits were made and connected to the data we had aggregated for the two years at hand and a table was created to tell the user the details of the Parking Lot as well as the Fortune ball predicted the number of cars.

This is a very important dashboard as it helps LAZ make predictions for each lot depending on the time of day and date. It will help them make staff management changes, advertise for coupons and most important help them pre-book in their mobile application. They would always know if there was a dearth or an excess of cars and this is a very important aspect in any parking business.

Conclusion

Finally, we recommend LAZ for the following:

- Use our product for easy handling of data and reporting purpose
- Integrate decision tree prediction algorithm with mobile application for pre-booking services
- Leverage Oakland airport passenger traffic and weather forecasts for accurate predictions
- Build pricing models and provide service with competitors in mind

Also some recommendation which if implemented could reap short term benefits are :

- 4A sees 18.4% of traffic on Monday mornings between 8 am to 12pm
- 4B sees 21.6 % of the total revenue coming on weekends
- 4C can have luxury valet services
- 4D sees 22% of the lots cars come between 8pm to 12 am
- Also Friday night sees 31% of Friday's traffic which generally comes in after 5pm
- Jan and March have the highest capacity and sept and October the lowest.
- We can have staff management and pricing solutions based on our models and recommendations.

Thus, we believe that the parking industry requires to implement an accurate methodology of data collection which analytics could leverage and make decision-making easy. Our one stop analytics solution would help LAZ to see real time patterns in traffic and also, future analysis of revenue and hourly traffic.

References

- *Revenue management techniques applied to the parking industry by Daniel Rojas*
- *www.airpark.com*
- *www.expressoparking.com*
- *www.fasttrackoak.com*
- *www.pnf.com*