Analytical Evidence To Justify Legislative Changes In State Government Policies

Parijat Rai, Xin (April) Ye, Snigdha Bhargava, Matthew A. Lanham

Purdue University Krannert School of Management

rai21@purdue.edu; ye117@purdue.edu; bharga16@purdue.edu; lanhamm@purdue.edu

Abstract

This study provides a predictive and prescriptive analytics solution to justify the proposal of legislative changes to current state government policies for credentials. The predictive component provides a model that predicts expected future demand for various customer credential types. The prescriptive component provides a workforce capacity planning decision model that identifies how to modify credential expiration dates. If this decision can be made via legislative change, it will inherently lead to a more balanced demand profile year-on-year, as well as facilitate a state government agency's staffing decision. In collaboration with this governing body, our analyses will help provide the empirical evidence they require to propose state-wide changes to credential dates.

Introduction

Every year a state government agency office receives enormous traffic and requires an optimum workforce to manage customers. The motivation for this study comes from the fact that services associated with credentials in a state authority such as issuance, duplication, renewal, amendments, etc. bring significant variation in customer traffic to their offices to date. With limited resources available, the government agency needs to not only balance the yearly customer traffic with workforce capacity, but also explore opportunities to hire additional staff to satisfy demand. In this study, the variation in credential renewals varies widely year-to-year. This occurs because **most people wait until their credential is about to expire to renew it**.

Credential Issue Date
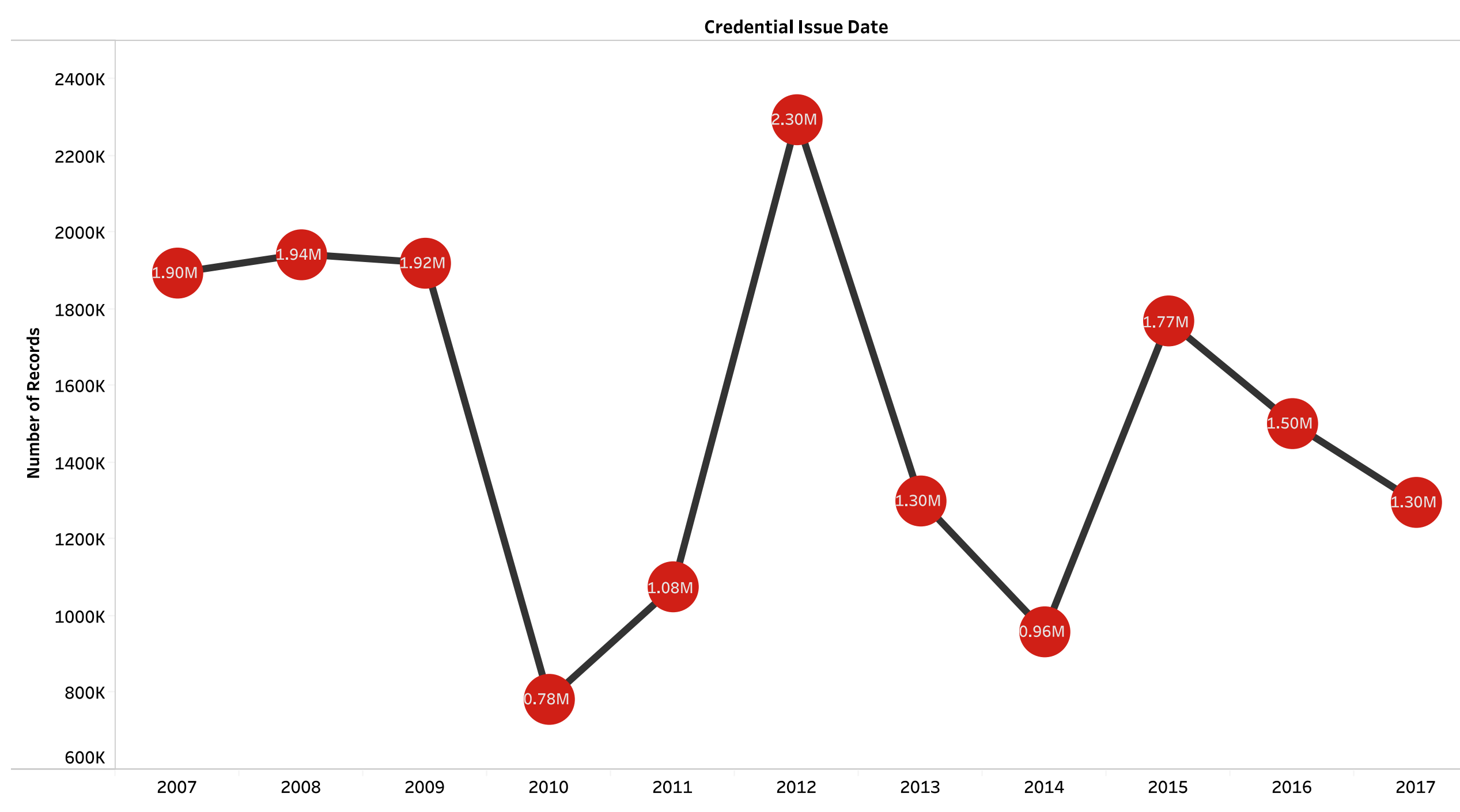
|	Year	Count	Pct_Inc_Dec
0	2007	1895583	0.00
1	2008	1944206	2.57
2	2009	1921912	-1.15
3	2010	784076	-59.20
4	2011	1075532	37.17
5	2012	2296134	113.49
6	2013	1301828	-43.30
7	2014	961223	-26.16
8	2015	1769268	84.06
9	2016	1503316	-15.03
10	2017	1297960	-13.66

Historically expiration dates have been fixed (e.g. 8 years), which leads to similar high peak years that are challenging to staff. Having the ability to modify expirations to certain ranges (e.g. 4-6 years) rather than fixed values would help minimize the variation in future customer arrival times and increase the overall experience with higher-quality service for customers (e.g. faster service, reduced anxiety). In collaboration with a state agency, we ran a pilot project for comparing yearly demands (before and after optimizing credential expiration dates).

Our empirical results provide motivation to justify proposing legislative modifications to credential expiration dates. Our solution provides a predictive model using time-series analysis to forecast growth in customer traffic due to services provided by the agency. Furthermore, an optimization model serves to minimize future demand variation and staff offices so that demand is satisfied.

Methodology

Data Source:
The data given by the state government agency consists of the type of transaction made by a customer at the agency office. The transaction type includes whether a customer came to renew/amend/duplicate/issue a credential. There are various types and classes of credentials issued to customers. The data also mentions when the customer came to office and when the credential expiration date was, spanning across 11 years from 2007 to 2017. There are certain columns which also help us understand whether the customer is a foreign national, Real ID compliant, or whether a customer's transaction was cancelled by the office. In totality, there are over 23 million rows and 11 columns
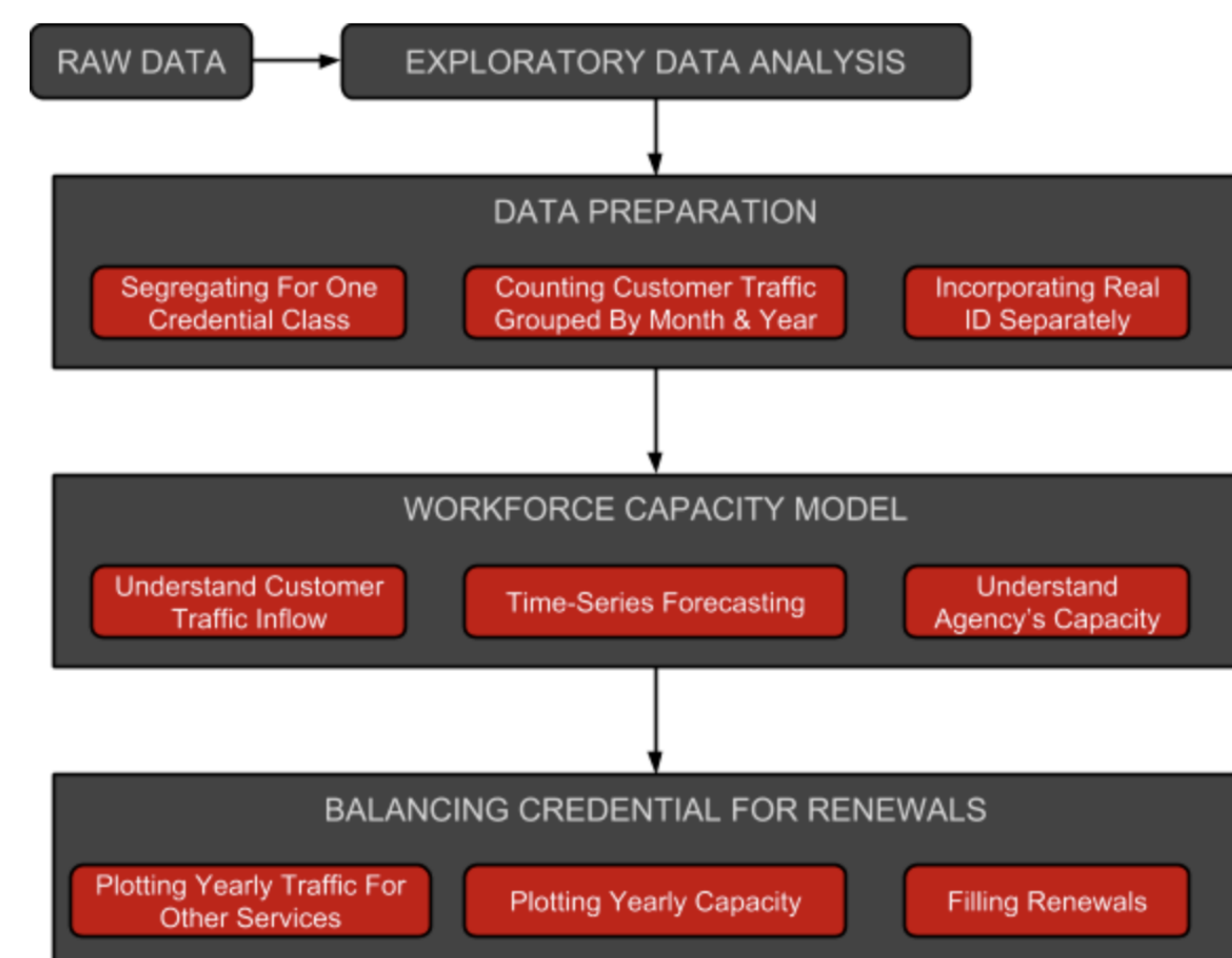
Exploratory Data Analysis:
The analysis began by looking at the distribution of customers when they came to office. The distribution can be gauged at the year level as well as month level. The year level distribution suggests the % increase/decrease across years, while the month level distribution suggests seasonality effect. Besides, the quantity of Real Id compliant customers and foreign nationals were also noticed.
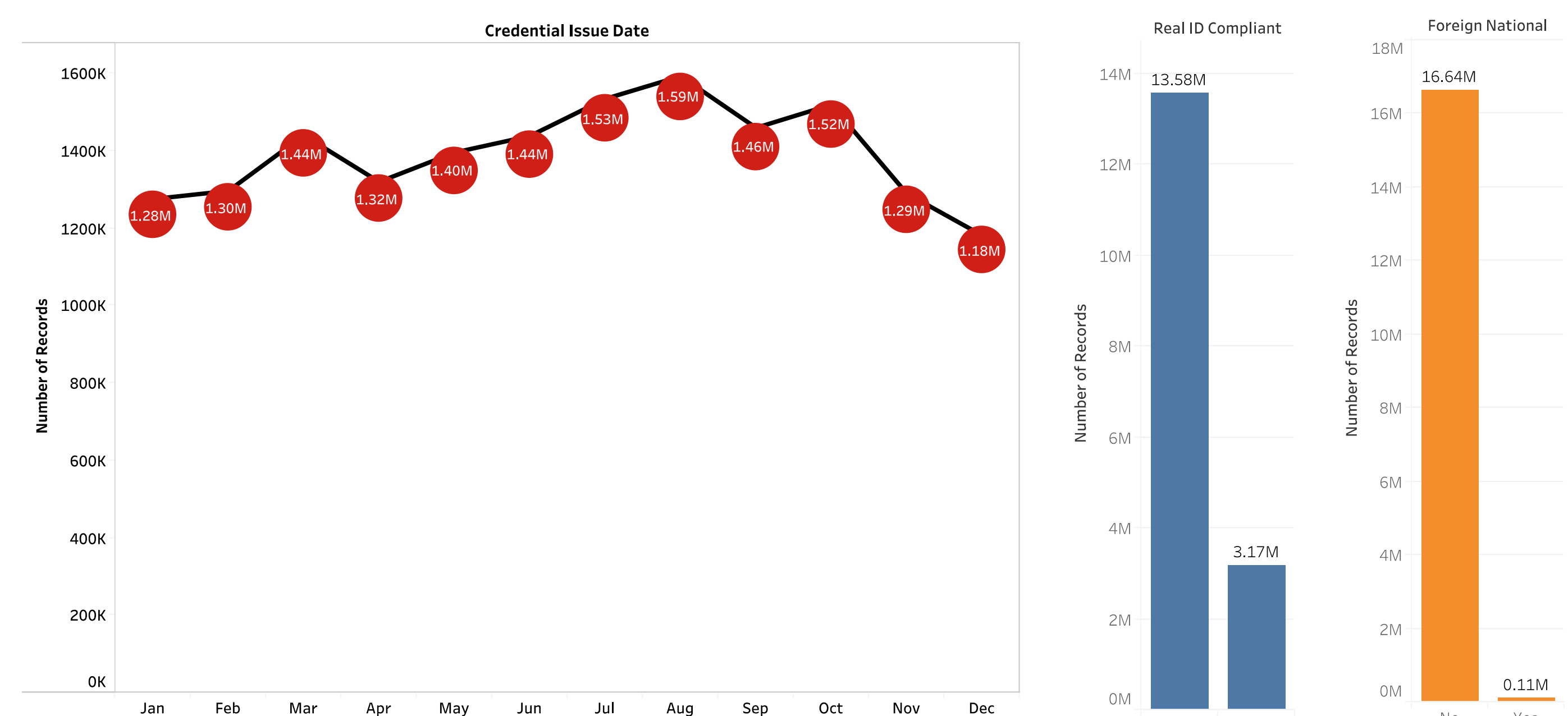
**Data Preparation & Partition:**
The next step involved preparing the data for building workforce data was partitioned to validate time-series forecastcapacity model. Although there are multiple credential classes, the agency wants to notice the results of only one type of credential class which generates 70% of the incoming customer traffic. The data was grouped by month and year and consequent traffic was calculated. It was then fed in R to conduct time-series analysis. The ing models. The train dataset consisted of traffic inflow count from years 2007 to 2016, while the test dataset included traffic inflow of 2017. Post validation, traffic was forecasted for year 2018.

Workforce Capacity Model:
To build a workforce capacity model, it was imperative to first understand the customer traffic inflow. A temporal graph was developed to notice a trend and seasonality effect of traffic into the agency offices. Predictive modelling technique such as Time-Series forecasting (Holt-Winters, ARIMA, etc.) was used to estimate traffic inflow and the results were validated through data partitioning. Besides, historical capacity evolutional trends of the agency's offices were stochastically modelled over time to generate a similar temporal graph as traffic. Difference between traffic and capacity were assessed to present strategic and optimum workforce planning in-terms of internal transitions or hires/attritions.
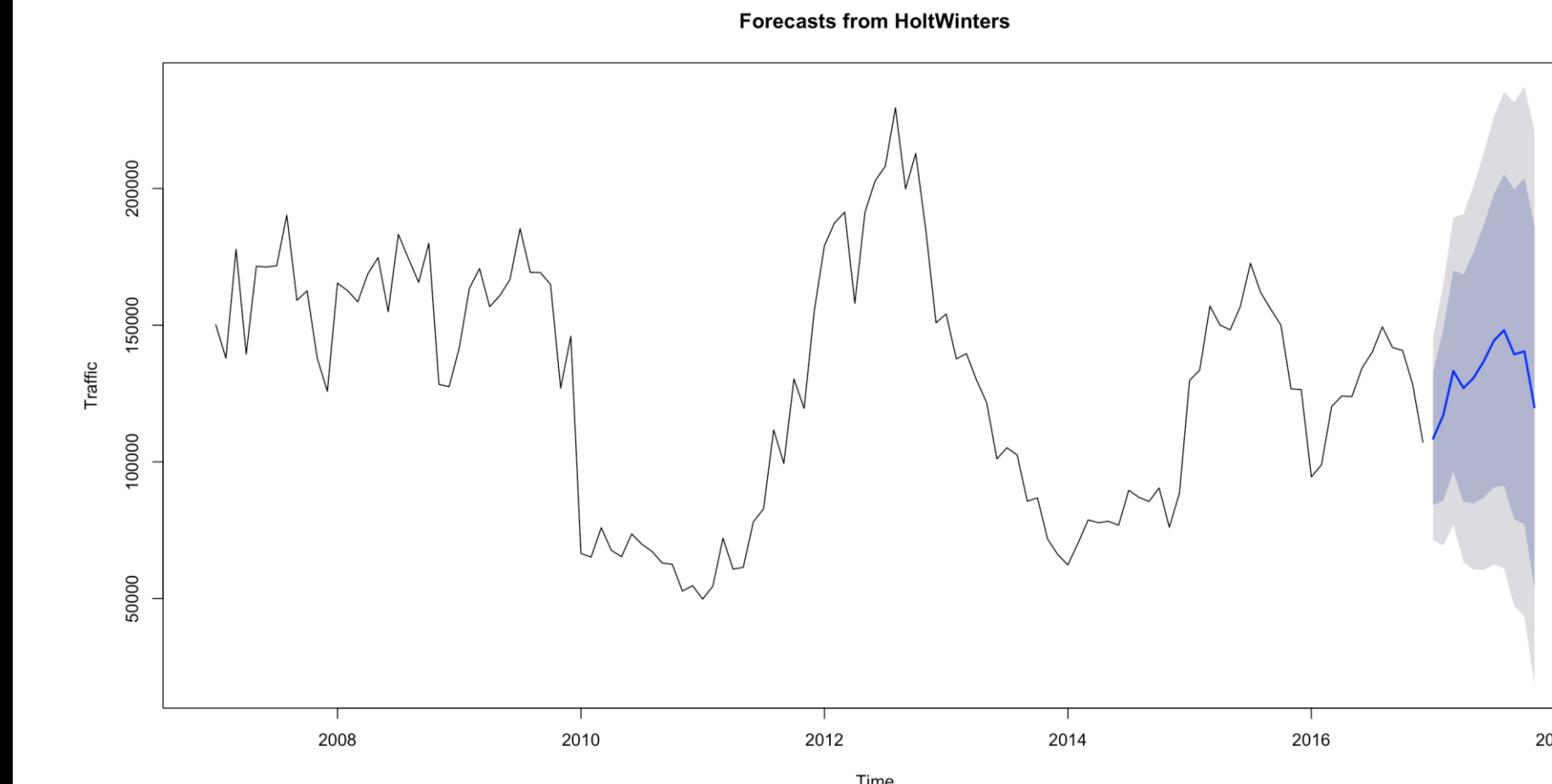
Results

Time-Series Analysis:
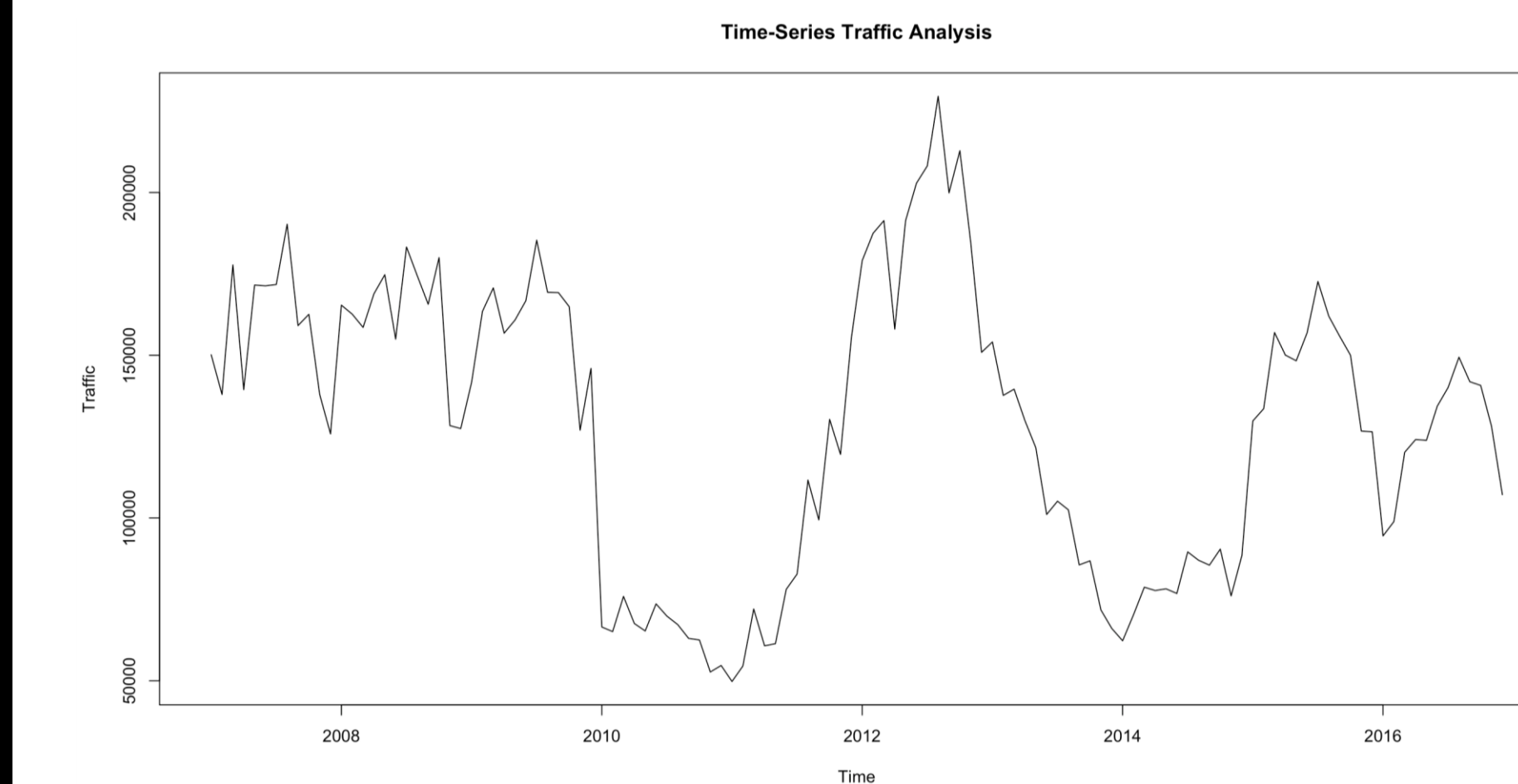
The time-series analysis was conducted on the train dataset which was later validated on test dataset. The graph conforms to the fact that traffic inflow varies significantly over each month. Hence, balancing the traffic becomes imperative.

The seasonality plot suggests that there is a dip in traffic in the months of April and November across all years, while there is surge in in-flow of traffic in months of August and March. This gives us unique insights while balancing the credentials
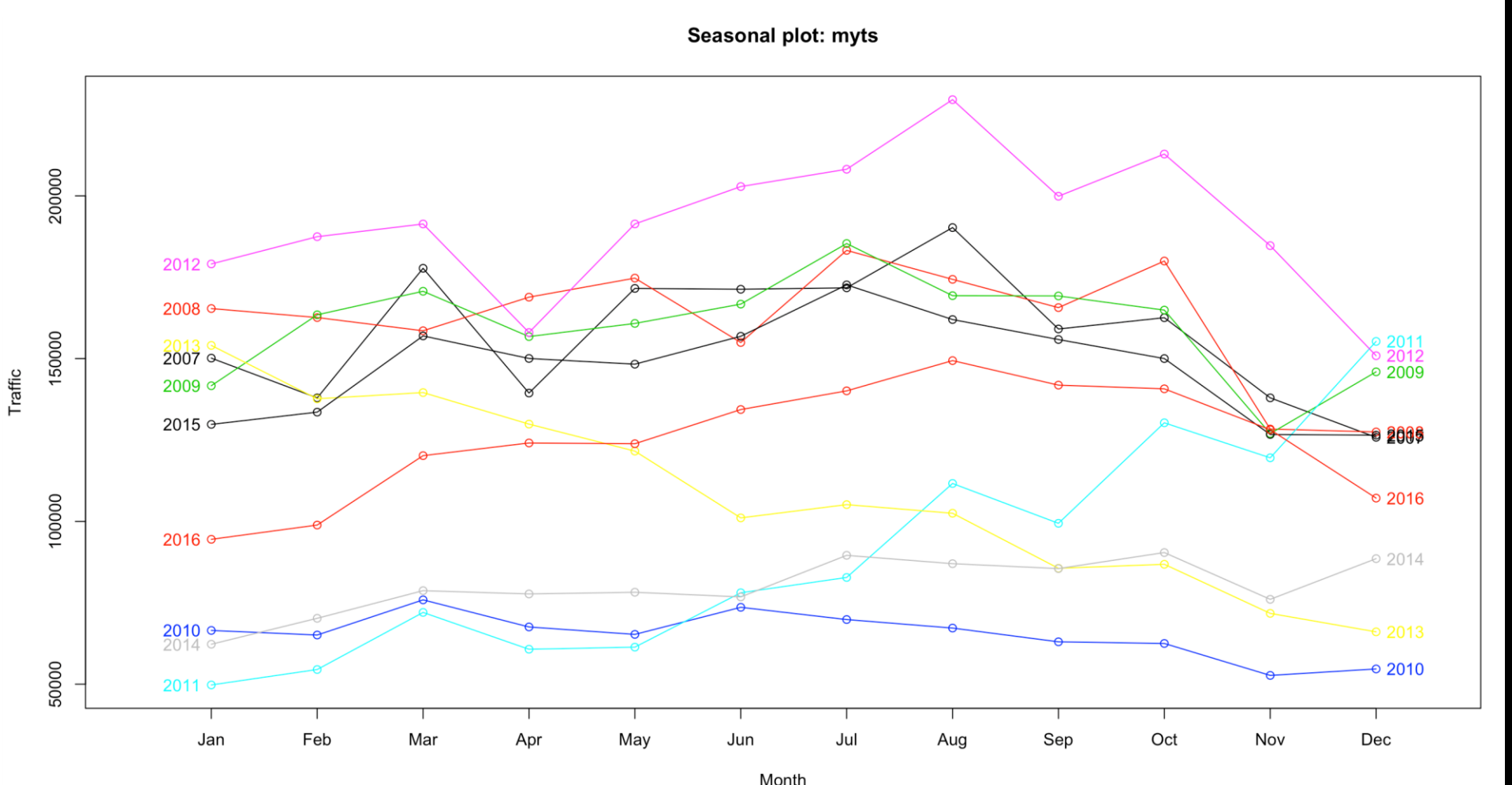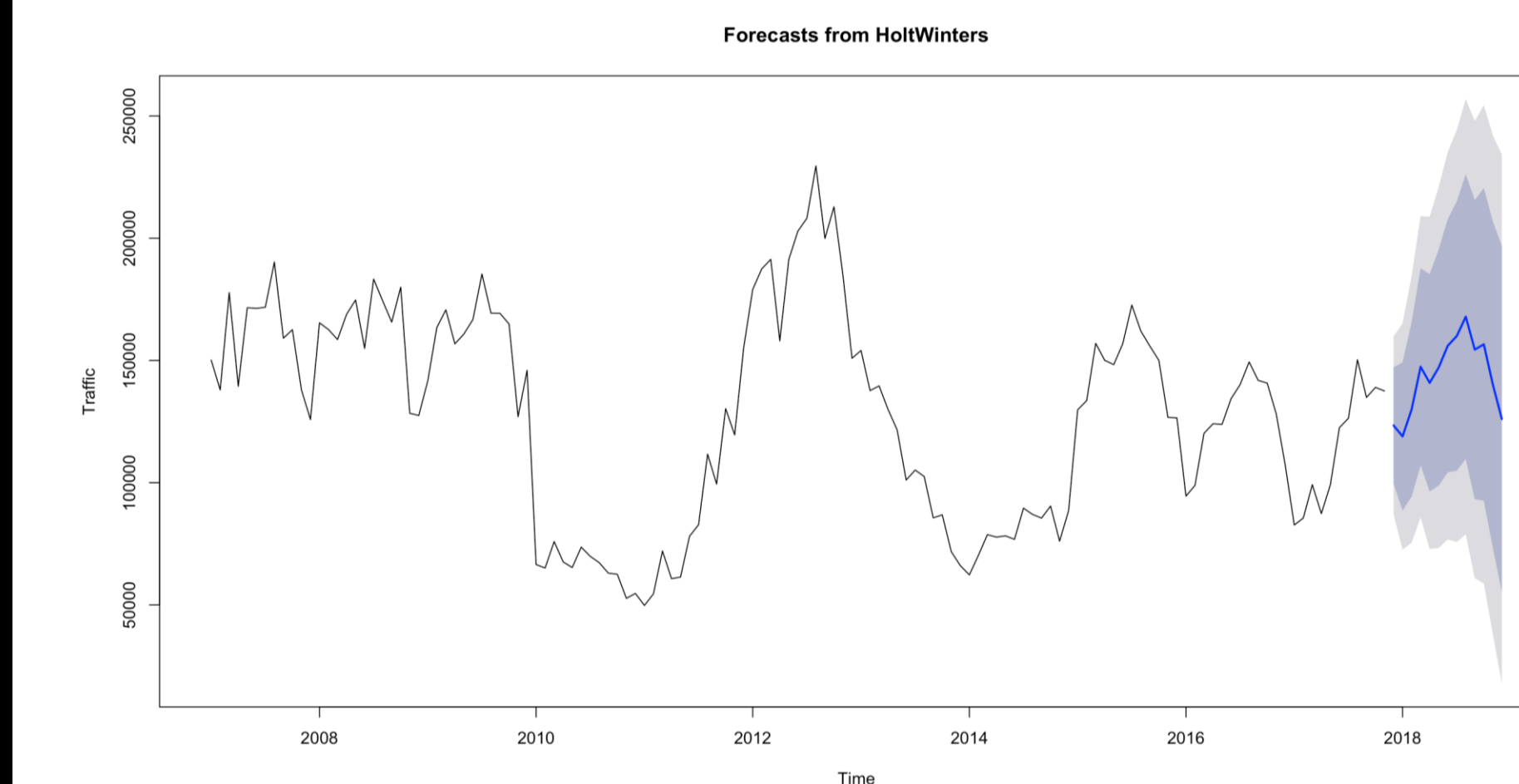

Model Validation:

The model was validated to using test dataset and metrics such as RMSE, MAPE, MPE, etc. were calculated and best performing model was used for forecasting.

Predicted			Actual	
month	Point Forecast		month	final_count
			yearmon	<int>
Jan 2017	108501.3		Jan 2017	82687
Feb 2017	116984.3		Feb 2017	85524
Mar 2017	133233.7		Mar 2017	99167
Apr 2017	126979.4		Apr 2017	87347
May 2017	130750.0		May 2017	99153
Jun 2017	136808.9		Jun 2017	122509
Jul 2017	144354.4		Jul 2017	126332
Aug 2017	148179.7		Aug 2017	150286
Sep 2017	139351.1		Sep 2017	134874
Oct 2017	140407.0		Oct 2017	138966
Nov 2017	119950.5		Nov 2017	137509

Time-Series Forecasting For Traffic:

Forecast For Year 2018	
	Point Forecast
Dec 2017	123419.4
Jan 2018	118937.5
Feb 2018	129934.9
Mar 2018	147400.9
Apr 2018	140836.3
May 2018	147087.7
Jun 2018	155976.5
Jul 2018	159952.5
Aug 2018	167881.9
Sep 2018	154430.5
Oct 2018	156581.0
Nov 2018	140125.8
Dec 2018	126036.1

Conclusions

In collaboration with a state government agency, we ran a pilot project for comparing yearly demands (before and after optimizing credential expiration dates). Our empirical results provide motivation to justify proposing legislative modifications to license expiration dates. Our solution provides a predictive model using time-series analysis to forecast growth in customer traffic due to services provided by the agency. Furthermore, an optimization model serves to minimize future demand variation and staff offices so that demand is satisfied. Currently, only traffic modelling is conducted. We are working on building a stochastic capacity model to give staffing recommendations to the agency and on balancing the credential renewals to even-out the traffic inflow in the office.

Acknowledgements