

Statistical Modelling for Rate of Traffic Volume

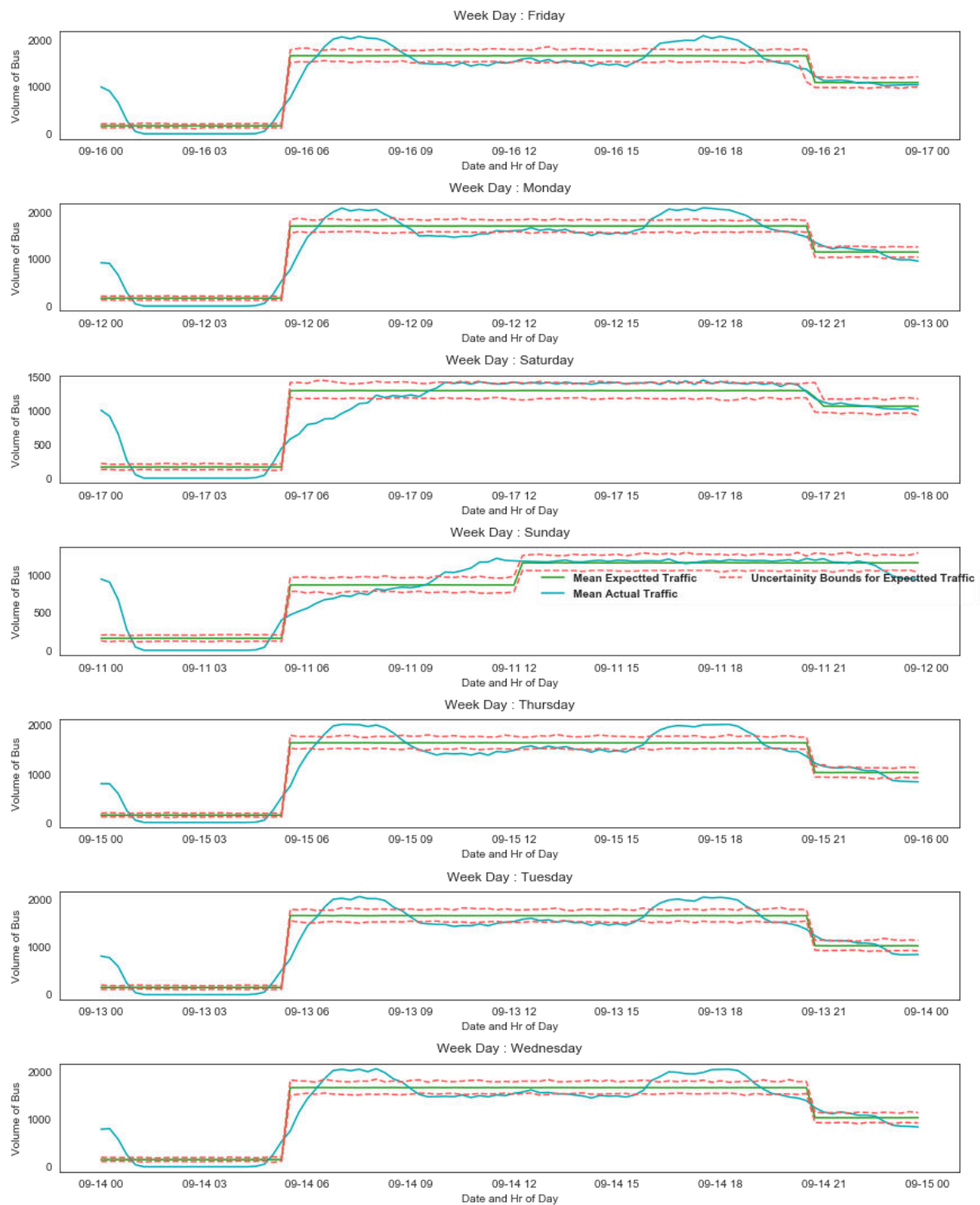
The first method we explore to model the traffic in Zurich is a statistical inference method which roots in a field of modelling techniques called Bayesian Inference. The basic Idea originates from understanding the Baye's theorem that is widely used in probability and statistics. The figure below shows the formula.

$$p(\Theta | \mathbf{x}) = \frac{p_{\mathbf{x}}(\mathbf{x} | \Theta)p(\Theta)}{\int p_{\mathbf{x}}(\mathbf{x} | \Theta')p(\Theta')d\Theta'}$$

The main goal of this method is to generate probability distributions for the model predictions which are also known as posteriors and is left-hand term in the formula. The numerator in the formula also has two probability distributions. The first, known as the Likelihood, is what the model will infer and update from the data set. The second term, known as the Prior, is a distribution that the modeler defines according to any information they know about the data.

In the model, the posterior distribution is a Poisson Distribution that models the count of traffic in Zurich Central Station in a particular time. The parameter that defines a Poisson distribution is λ which is defines the rate of traffic in a particular time and is the Parameter our model needs to learn. The prior distribution that we define in for the model helps it to learn λ and in turn the likelihood distribution. From our initial exploration of the model in section III, gives us an initial belief that the rates change Three times a day. The rate as we believe is further affected by the delays, which make the number of traffic in the Zurich Station. Therefore, for the prior we define λ to originate from a mixture exponential distribution. We use the 2016 dataset to generate seven different posteriors for each day of the week. To understand how good the generated Posterior is we sample 1000 prediction of traffic volume in a day. Using these 1000 samples we can plot the mean prediction value and put uncertainty bounds on them using the maximum and minimum sampled values of predictions. Further the actual traffic volume from the 2017 dataset, is plotted against the posterior mean prediction and the uncertainty bounds to see how well the learnt distribution is able to simulate future data. The plot can be seen below.

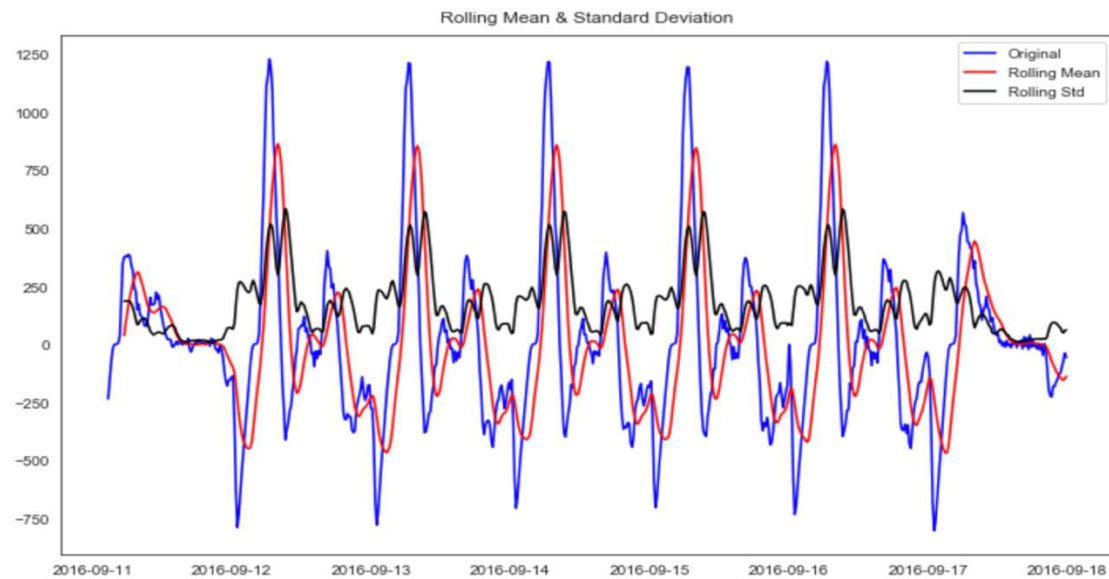
Model Prediction and Uncertainty: Bus volume in Zurich Bahnhof (01-15-2017 - 01-21-2017)



i. *Arima Model for traffic volume*

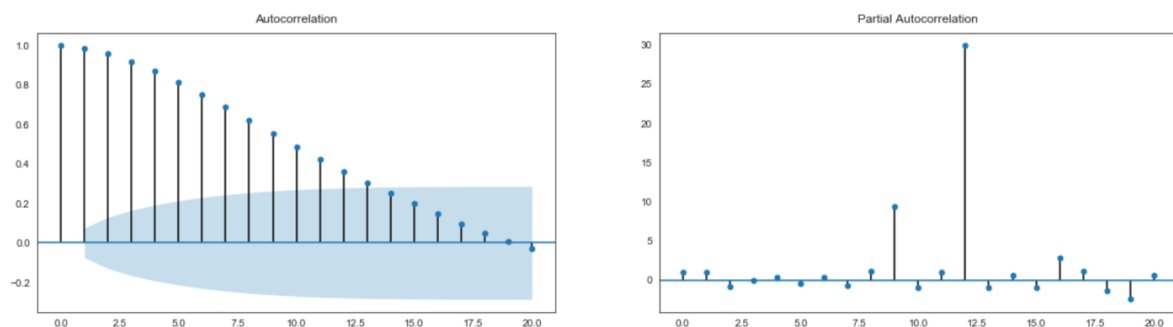
The second modelling technique that we implemented to predict the traffic volume was an Arima Model. We use the 2016 dataset to run the Arima model. Since the dataset comes from public transport that run-in schedules, we would expect the data to have a trend. Before applying the Arima model on the time series, we want to ensure it's stationary. For this we run an Augmented Dickey-Fuller Test where: the time series is considered stationary if the p-value is low (according to the null hypothesis) and the critical values at 1%, 5%, 10%

confidence intervals are as close as possible to the Augmented Dickey-Fuller Statistics. The snapshot below shows the Augmented Dickey-Fuller Statistics for our model.

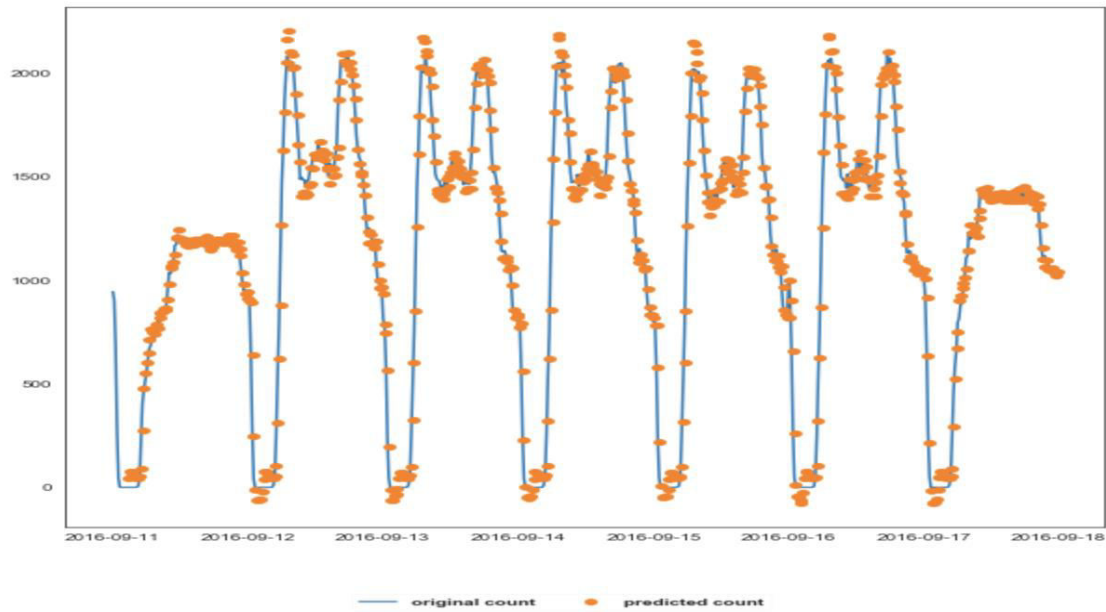


ADF Statistic: -5.532437908861344
p-value: 1.7750756230902862e-06
Critical Values:
1%: -3.440560883168159
5%: -2.8660454146233434
10%: -2.569169329058723

The Augmented Dickey-Fuller Statistics is far from the critical values and the p-value is greater than the threshold (0.05). Thus, we can conclude that the time series is not stationary. Next we plot the autocorrelation and partial correlation plots to understand the lags and seasonality.



Since the AR model implemented in the stats model's library allows to include only consecutive seasonality terms, we implement the model manually. In this way we can flexibly use nonconsecutive seasonality terms like weeks 1, 9,11 as observed from the Autocorrelation and Partial Autocorrelation analysis. The plot below shows the result of running the ARIMA model implemented with the respective AR order, differencing order and MA order of 1,9,11.



Since we used two different modelling techniques to model the bus volume, it is difficult to quantify the accuracy of the two models for comparison. Visually comparing the models, we see that the statistical model gives us flexibility in showing the uncertainty in our predictions but is rigid in terms of predicting smaller fluctuations in the count of traffic volume. On the other hand, the ARIMA model is able to predict the smaller fluctuations in the count better than the statistical model but does not give us an idea about uncertainties in our prediction.

i. Osterwalder's Business Model Canvas

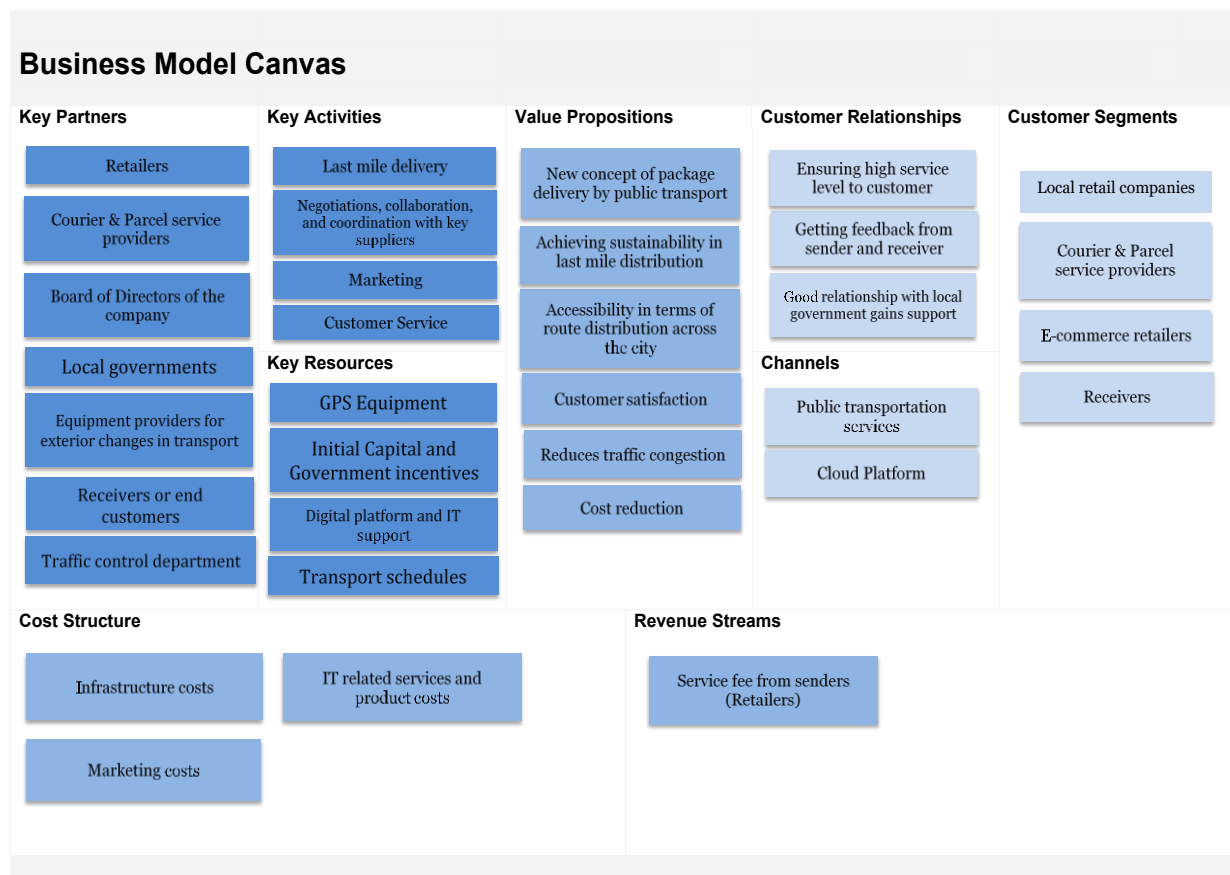


Figure 19: Business Model 3

Strengths: The last mile delivery of the goods, helps in achieving sustainability by utilizing the available transport, i.e. trams and buses, in carrying the packages to the end customer. Zurich, the metropolitan city of Switzerland, is densely populated and eliminating the additional trucks and buses on the road would contribute to lowering congestion and pollution levels. The e-commerce sectors tackle a grave problem of managing the last mile delivery of their products at affordable costs, by using the facility offered by the transport agency. A liaison between the e-commerce retailers as a substitute to the existing operations of providing transport services to the public, generates easy revenue to the company as there are no other huge capital investments or expenses to be borne by the company to provide such services.

Weaknesses: Most of the e-commerce retailers long in the business, possess the necessary transport vehicles to conduct the last mile delivery process. Likewise, the retailers already maintaining long-term relationships with outsourced delivery service partners may be reluctant for a shift. Although, the aim is to deliver a 100% customer service, it is not feasible to provide door delivery of the product. Finally, the modifications made to the existing trams and buses could act as a possible hinderance to the passengers who are travelling by it.