**Student Lending Forecast Case Study**

Background

A forecasting spreadsheet developed in Excel by Capital Markets team and maintained by the Analytics team is currently used by the business to develop projections for Student Loan application and funded volume. The spreadsheet uses historical application and funding data, along with factors such as upcoming Marketing campaigns to develop daily forecasts for applications (number of apps and $ volume) as well as funded loans (number and $ volume).

The forecast is being used by the business for variety of decision making, including:

* Goal setting
* Tracking actual vs. forecast for assessing health of business
* Operations capacity planning
* Inputs into Business Development and Marketing efforts

Over the past few months, especially post-LD1 due to rate changes, the actual funding volume has been tracking significantly higher than the forecast. The immediate impact of this has been felt by the Operations team as the staffing levels are not adequate to process the applications. Processing times have consequently become longer and have impacted customer experience and Net Promoter Scores for the business.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Feb | Mar | Apr | May | Jun | Jul | Aug | Sep | Oct | Nov | Dec |
| Forecast | $125.6 | $82.6 | $88.5 | $132.7 | $156.6 | $148.6 | $160.6 | $159.3 | $163.0 | $167.0 | $157.6 |
| Actual / Projected | $120.9 | $69.3 | $101.9 | $145.8 | $191.3 | $145.8 | $164.2 | $198.4 | $223.4 |  |  |
| Variance | -$4.7 | -$13.3 | $13.5 | $13.1 | $34.7 | -$2.8 | $3.7 | $39.1 | $60.5 |  |  |
| MAPE | 3.7% | 16.1% | 15.2% | 9.8% | 22.1% | 1.9% | 2.3% | 24.6% | 37.1% |  |  |

RMSE% (Feb – Oct): 20.0%

Objective

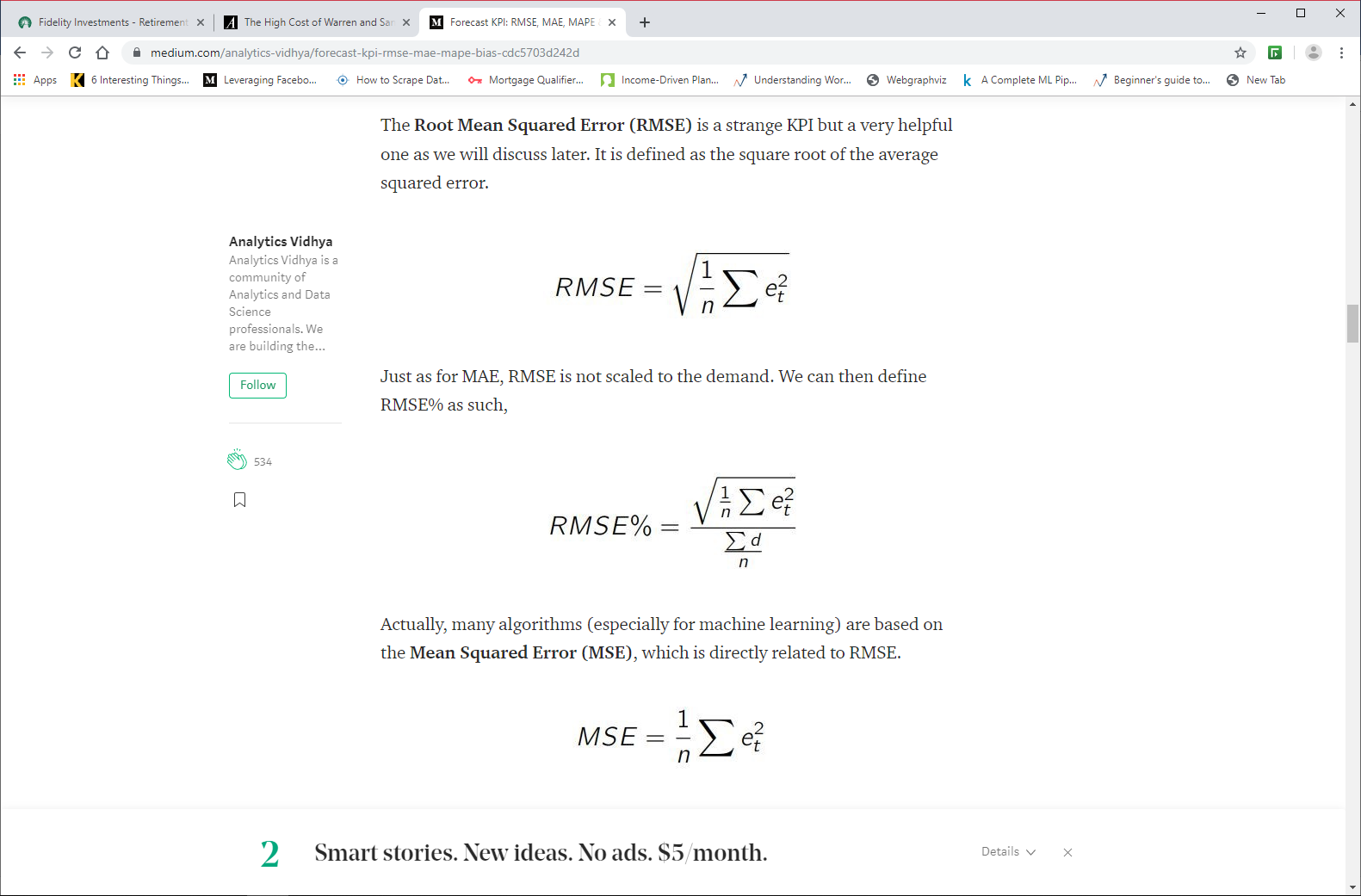
Develop a forecasting tool / model to predict application and funding counts & volume at the daily, weekly or monthly level. Forecasting tool should incorporate the impact of initiatives, such as Marketing campaigns, aggregators etc.

Training dataset with daily application counts, app volume, funded counts and funded volumes are included with this case study. The dataset comprises actual data from Jan ‘16 – Oct ‘ 19. The dataset should be split into training data (Jan ’16 – Jun ’19) and testing data (Jul ’19 – Oct ’19). The forecast tool / model should be developed on the training dataset and validated against the test dataset. It is recommended that Refinances and In-School loans be handled separately.

Once the forecasting methodology / model has been developed using the training dataset, participants will use it to predict the application and funded counts and volume for Jul ‘ 19 – Oct ’19 time period. Only attributes that are known as of Jun 30, 2019 should be used in prediction i.e. actual LIBOR rates from Jul – Oct ’19 should not be used to make the prediction. The winning model will be selected based upon the evaluation criteria described below.

Evaluation Criteria

Root Mean Square Error (RMSE%) for funded volume ($) forecast in the test period (Jul ’19 – Nov ’19) will be the primary metric to evaluate the best forecast. RMSE is calculated as follows:

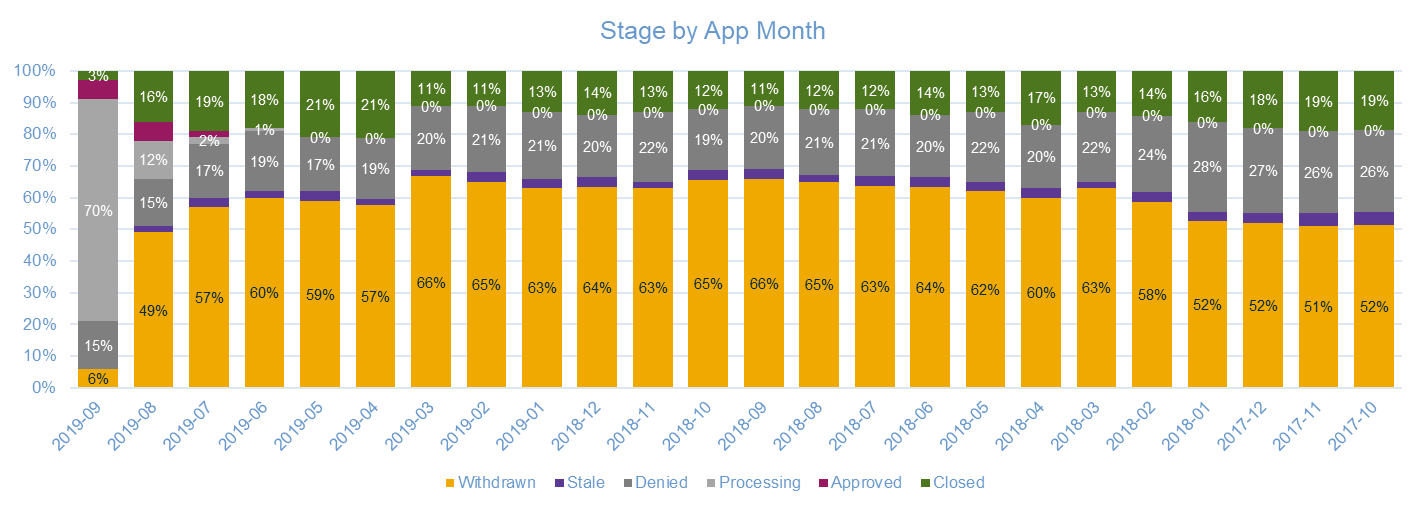


In addition to RMSE for funded volume, RMSE for application count, application volume and funded count may be used to additional evaluation metrics to break ties.

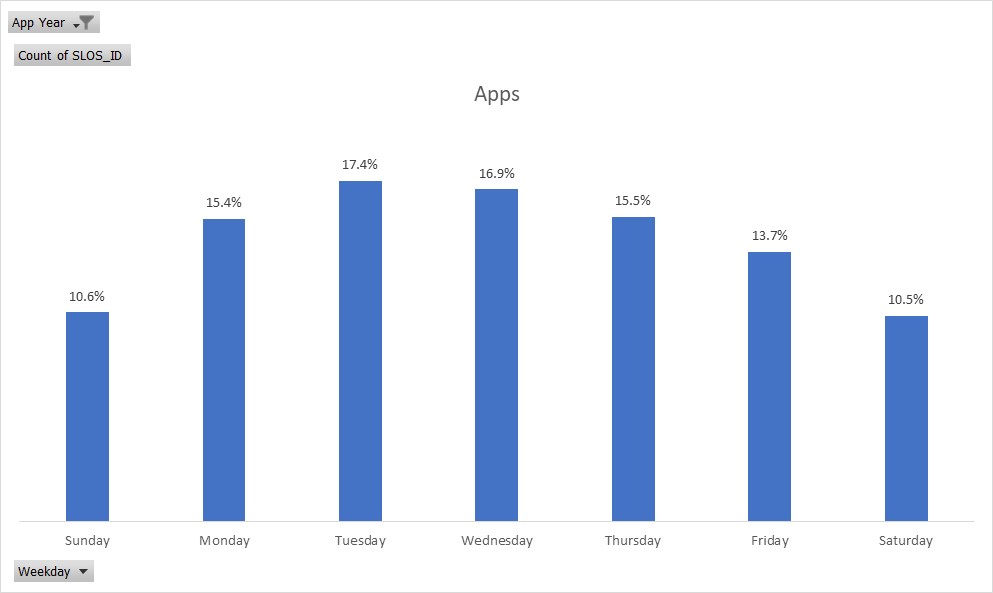
Factors to Consider

Participants should consider the following factors that may influence the forecast accuracy.

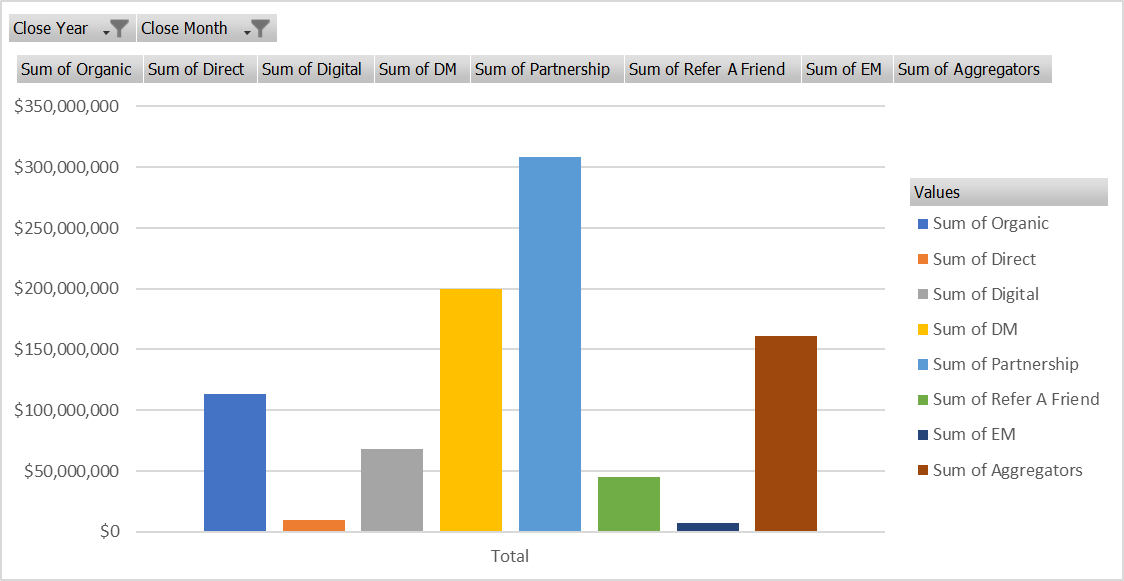
1. Macro-Economic Environment: overall macro factors, such as unemployment rate, interest rate, inflation rate etc. might play a role in borrower’s decision to refinance. Data.Gov, TreasuryDirect.Gov and CBO.Gov are great sources for macro variables to include in the forecast. A few macro attributes are included as part of the case study dataset
2. Competitive Rate Environment: Laurel Road Marketing team generates a weekly report with comparison of Laurel Road rates vs. select competitors. Borrowers decision to use Laurel Road may be influenced by competitor’s rates.
3. Laurel Road Rates: Drop in rates post-LD1 had significant impact on the close rate as illustrated in the chart below



1. Seasonality: As illustrated by the chart below, the application and closed volumes are seasonal in nature with the Nov-Dec traditionally being slow months
2. Day of Week: As the chart below shows, the app volume is higher Mon – Thu and lower during the weekends. Participants may want to factor in the day of week if developing the forecast at the daily level.



1. Borrower Segment: Borrower characteristics, such as degree type, in-school, FICO, DTI, income, occupation etc. may influence the forecasts. Participants may want to develop predictions for each borrower segment and combine to develop aggregate forecasts.
2. Marketing DM & EM Campaigns: Marketing campaigns have significant impact on application and funded loan volumes. The dataset included as part of the competition tags the organic apps and those generated due to a Marketing campaign using logic developed by Rachel Young.
   1. Prospect Marketing: Laurel Road Marketing team executes 7-8 DM campaigns annually targeting approx. 1MM prospects in each campaign. The campaigns response rates are ~0.35% resulting in 3,500 apps and ~500 closed loans. 2019 DM calendar is included for reference
   2. Affinity Marketing: Laurel Road partners, such as AMA, ADA, EMRA etc., share their member list that is leveraged by the Marketing team for several DM campaigns. List of campaigns included in attached PDF document
   3. Withdrawn EM: Marketing team executes monthly EM campaign targeting customers who withdrew their applications. Campaign targets ~4,000 withdrawn applicants each month with ~28% open rate and ~1.5% click rate
   4. Other DM/EM: Refer a Friend and other DM/EM campaigns executed by Marketing also have an impact on application volume
3. Digital Marketing Awareness Campaigns: LR Marketing team executed comprehensive digital marketing campaigns, including Search Engine, Banner Ads, Social Media, that drive potential customers to our site. Marketing team generates daily report to show impact of digital campaigns on app volume (see chart of Apr – Sep 2019 funded $ below)



Competition Rules:

1. Create forecast for Student Loan application count, application $, funded loan and funded $ for Jul ’19 – Dec ’19 using Jan ’16 – Jun ’19 data as training sample
2. Generate separate forecasts for Refis and In-School loans
3. Teams can use any of the data provided as part of this competition and incorporate additional attributes from internal or external sources
4. **Forecast for Jul ’19 – Dec ’19 should be based upon information that is known as of Jun 30, 2019. For example, knowledge of upcoming Marketing campaigns or digital spend can be incorporated but the actual LIBOR rates from Jul – Dec ’19 should not be used**
5. Total funded volume (Refi + In-School) by month from Jul – Nov ’19 will be the primary measure to determine effectiveness of model. RMSE% described above will be used to determine the winning team
6. Winning team earns prize money & bragging rights!

Datasets:

1. Application and funded loan volume at the daily level is included in the folder below along with several internal and macro attributes:

Path: \\52.45.41.128\SASShare\ANarayan\Forecast Case Study

1. Files in the folder include:
   1. SL - refi & inschool: Excel workbook with application and funded loan data for Refinances and in-school loans from Jan 1, 2016 onward
   2. Funding Forecast Variables – Marketing: Excel workbook created by Marketing team with weekly data on LR rates, SoFi rates, Digital Media spend, DM, EM campaigns etc.
   3. LIBOR: Excel file with LIBOR-3M daily rates
   4. Unemployment Rate: Monthly unemployment rate data
   5. CPILFESL: Monthly Consumer Price Index for All Urban Consumers. All Items Less Food and Energy, Index 1982-1984=100, Monthly, Seasonally Adjusted
   6. Treasury: Daily US Treasury Yield Curve data

Recommended Reading:

1. <https://www.analyticsvidhya.com/blog/2018/08/auto-arima-time-series-modeling-python-r/>
2. <https://www.analyticsvidhya.com/blog/2016/02/time-series-forecasting-codes-python/>
3. <https://otexts.com/fpp2/intro.html>