The University of Massachusetts Cashflows on Tuition Revenue

Alain Duplan, Ariel Lutati *University Of Massachuessetts, Amherst*

The University of Massachusetts is a public institution providing collegiate education to tens of thousands of students spread out between its campuses statewide. Billions of dollars of revenue is accumulated by the institution each and every year. The use of numerous sources of income has resulted in a net gain of profit yearly. One of the most prominent sources is the net total of tuition and fees. Given as the main functionality of the UMass schools is provided higher education to pursuing students, the revenue made off tuition could be stated as the most important source, even if it may not be the biggest. However, one may question if other cash flows would affect this source. A good practice in business is to keep all income independent of one another to avoid a chain reaction of effects. However shifts in other sources of income could lead to the institution making decisions that affect the student body leading to an adjustment in revenue as more students may want to attend or leave or causing a revaluation of tuition costs. **We hypothesize that the revenue coming from tuition and fees for the University of Massachusetts is dependent on the revenue of its other cash flows.** Through the use of random foresting and linear regression we aimed to reject our hypothesis. Upon testing we weren't able to formulate a model that was able to accurately predict revenue of tuition in fees, therefore correlation between the cash flows could not be drawn.

Introduction

The UMass system allows for a quality education for an economical price (for instate students). UMass's tuition and fees pricing is primarily decided by the Department of Education. Nonetheless, we wanted to explore for any internal factors within UMass' cashflows that could influence their revenue produced by tuition and fees. We believe that that if UMass acquires higher revenue in cashflows such as grants and contracts, operating, and non-operating activities, then quality of life on campus would increase consequently leading to higher revenue in tuition and fees. Likewise, if losses in cashflows increases, campus quality of life will decrease and thus, revenue made via tuition and fees will decrease.

Taking into consideration that overall student debt is at an all time high, it is hard not to question whether there is a correlation between other cashflows and tuition and fees. Given that students pay entirely towards tuition and fees. It would be within the students' best interests to understand what factors affect the students' largest contributer to academic finances and loan debt.

Data

The data we used was found on the University of Massachusetts office of the President website. Using the single audit reports of each year, parsed the cash flows for the years 2015 to 2019 into a single balance sheet. We focus on using only the net values of each category of the cash flows, along with the overall total cash flows, and the changes compared the previous years.

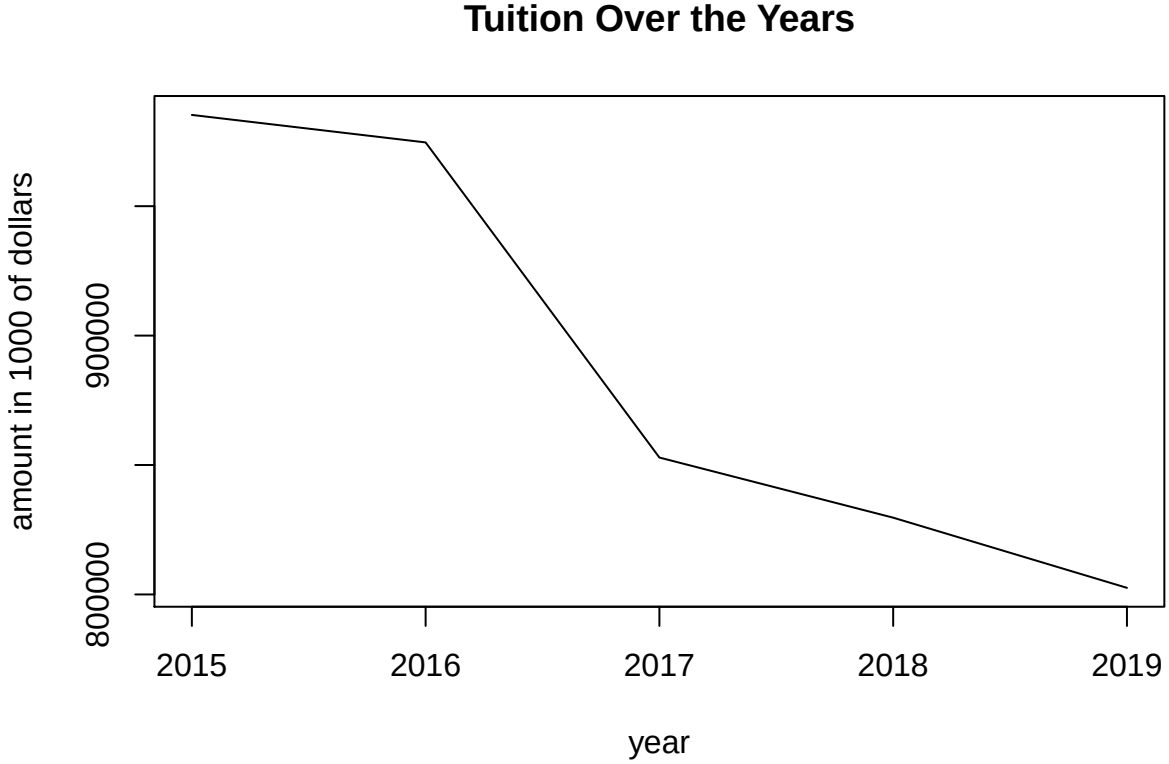
Year	2019	2018	2017	2016	2015
CashFlowsFromOperatingActivities	508876	342853	488027	570853	412417
CashFlowsFromNoncapitalFinancingActivities	922690	880530	834480	804330	754975
CashFlowsFromCapitalAndOtherFincancingActivities	519087	528300	344331	516310	273212
CashFlowFromInvestingActivities	7224	132630	52932	57259	43554
NetDecreaseInCashAndCashEquivalents	112497	123253	50810	225574	25792
CashAndCashEquivalentsBeginingOfYear	430389	553642	604452	830026	804234
CashAndCashEquivalentsEndOfYear	317892	430389	553642	604452	830026

Year	2019	2018	2017	2016	2015
ReconciliationsOfOperatingLossNetUsedByOperatingActivities	508876	342853	488027	570853	412417
OperatingLoss	817436	831578	718364	737257	669513
DeprecationAndAmortizationExpense	276638	261417	245300	240865	221043
SupplementalDisclosureOfNoncashActivities	52279	57532	191178	63109	362114
TuitionAndFees	985247	974631	852889	829654	802554

We will also look deeper at cash flows from non financing activities as it was recently discovered that due to the recent coronavirus outbreak, UMass is using the revenue from grants to refund their students. So we would like to see if there was any correlation between the two.

Year	2019	2018	2017	2016	2015
TuitionAndFees	985247	974631	852889	829654	802554
Grants and contracts	587539	569408	814018	791553	778001
State appropriations	780221	105380	720817	724985	676197
federal appropriations	7004	6688	6602	6827	6619
Grants, contracts, gifts, for other capital purposes	52308	38809	32856	26831	27106
Nonoperating federal grants	84454	81590	74050	75743	76539
Student Orginazation agency transactions	1297	1549	155	527	431

Upon initial inspection we see that the revenue of tution looks to be as if it is steadily decreasing over the years. The loss between 2015 and 2019 was almost 150 million dollars in fact.



In order to make that dataframe more suitable for modeling, we transposed the table to create a column that corresponds to each indiviual row in our data set.

Modeling

To begin we will reserve the last 2 years of the dataframe as our testing set, and everything else as our training set.

We first approached our modeling using linear regression models. When we created a model using all the variables made available to us, the model resulted in statistically insignificant coefficient estimates. So instead we approached this by making single variate regression models on all the variables we used in our data frame. As a result we found that 5 of the variables were significant on their own: sumplimental disclosures of noncash activities, net decrease in cash and noncash equivalents, cash flows from capital and other financing activities, grants and contracts, and student orginzation agency transactions.

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	6.106895e+05	3.89530e+04	15.677600	0.0405520
V3	7.046675e-01	8.26039e-02	8.530676	0.0742881

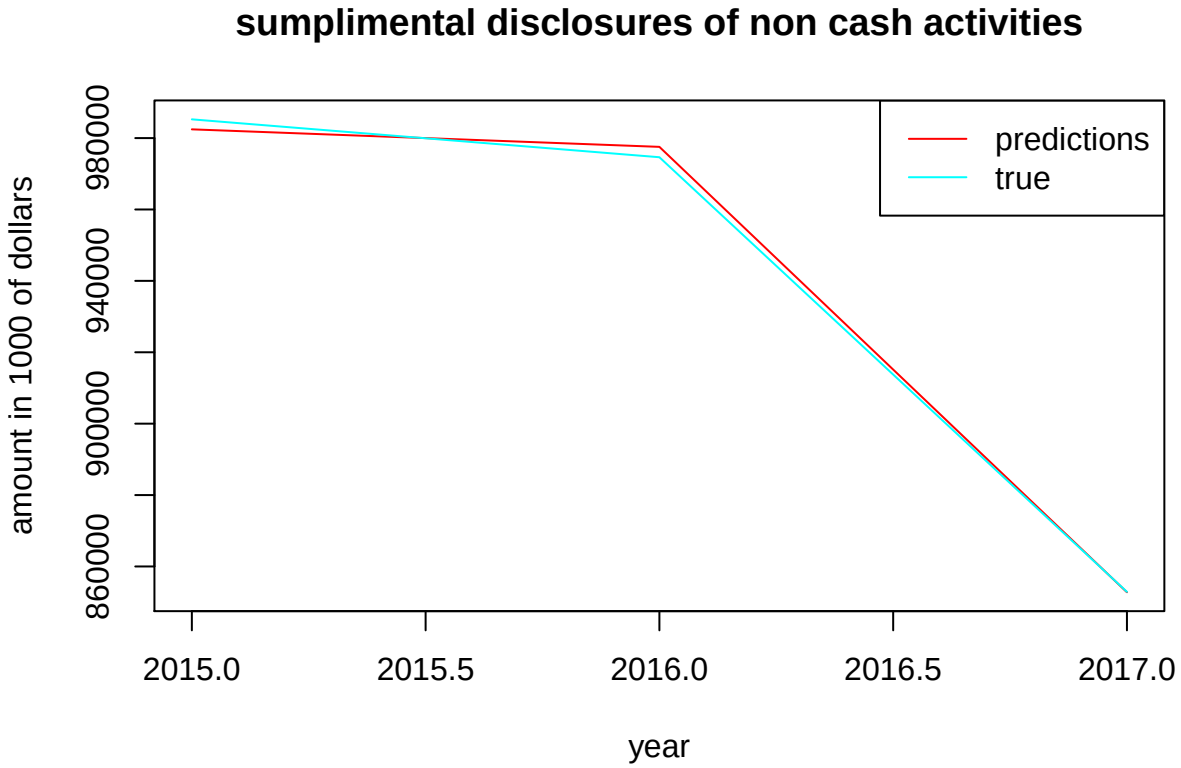
	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	7.618422e+05	3.954328e+04	19.266036	0.0330140
V5	1.839895e+00	3.926384e-01	4.685979	0.1338485

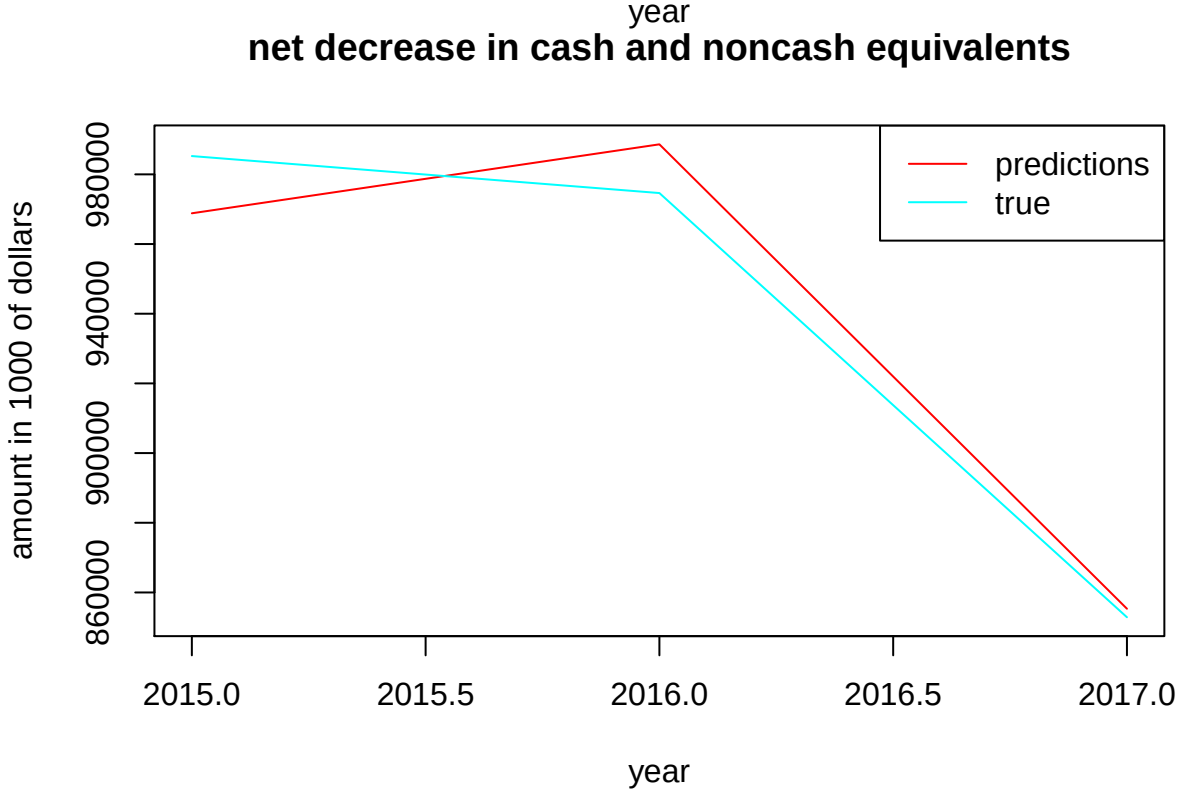
	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	1.03125e+06	4325.3870368	238.41803	0.0026702
V11	-9.33535e-01	0.0363012	-25.71635	0.0247430

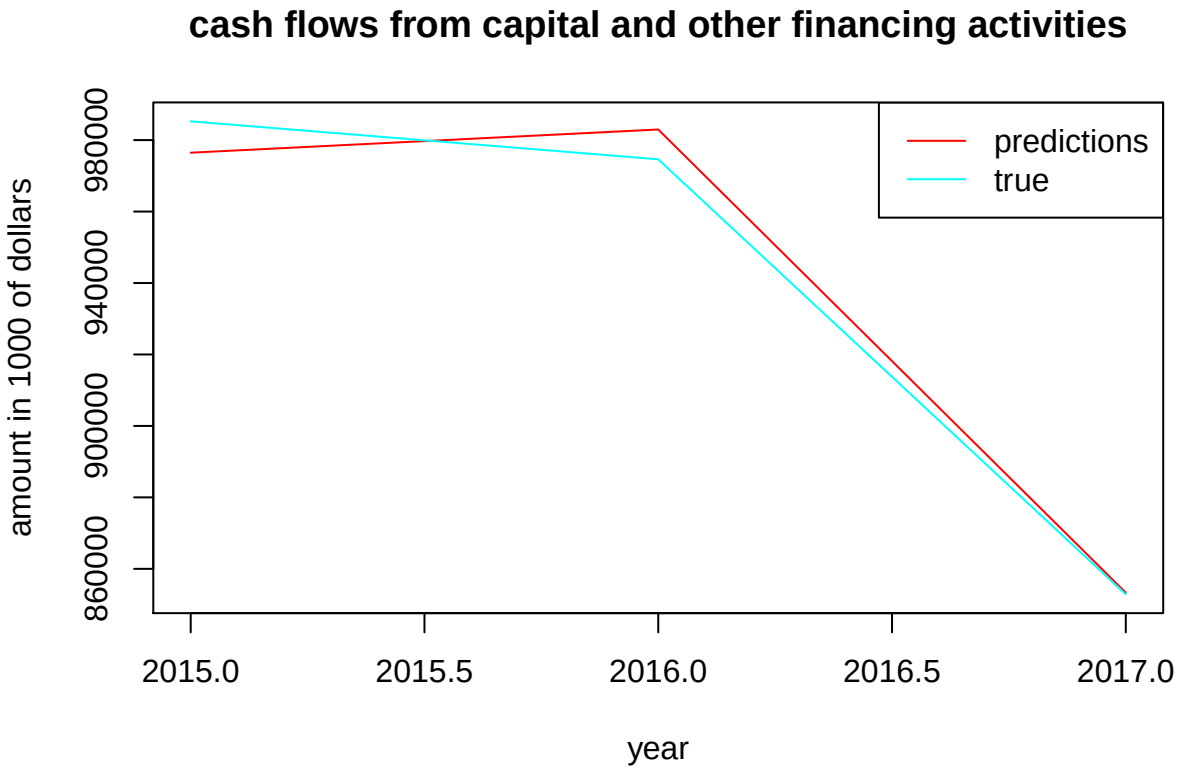
	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	1.288691e+06	4.974732e+04	25.904736	0.0245632
V2	-5.344117e-01	7.465680e-02	-7.158242	0.0883633

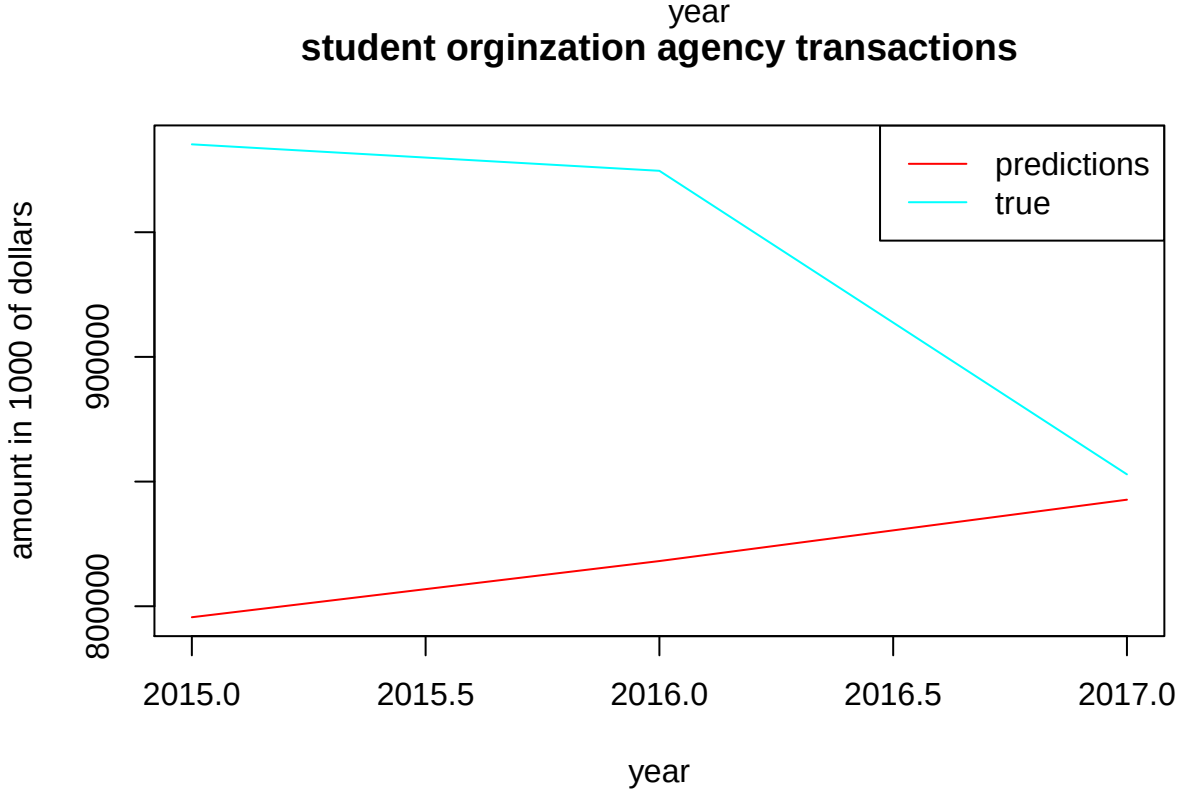
	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	841454.52616	27831.84249	30.233519	0.0210491
V7	96.10244	23.79098	4.039448	0.1544946

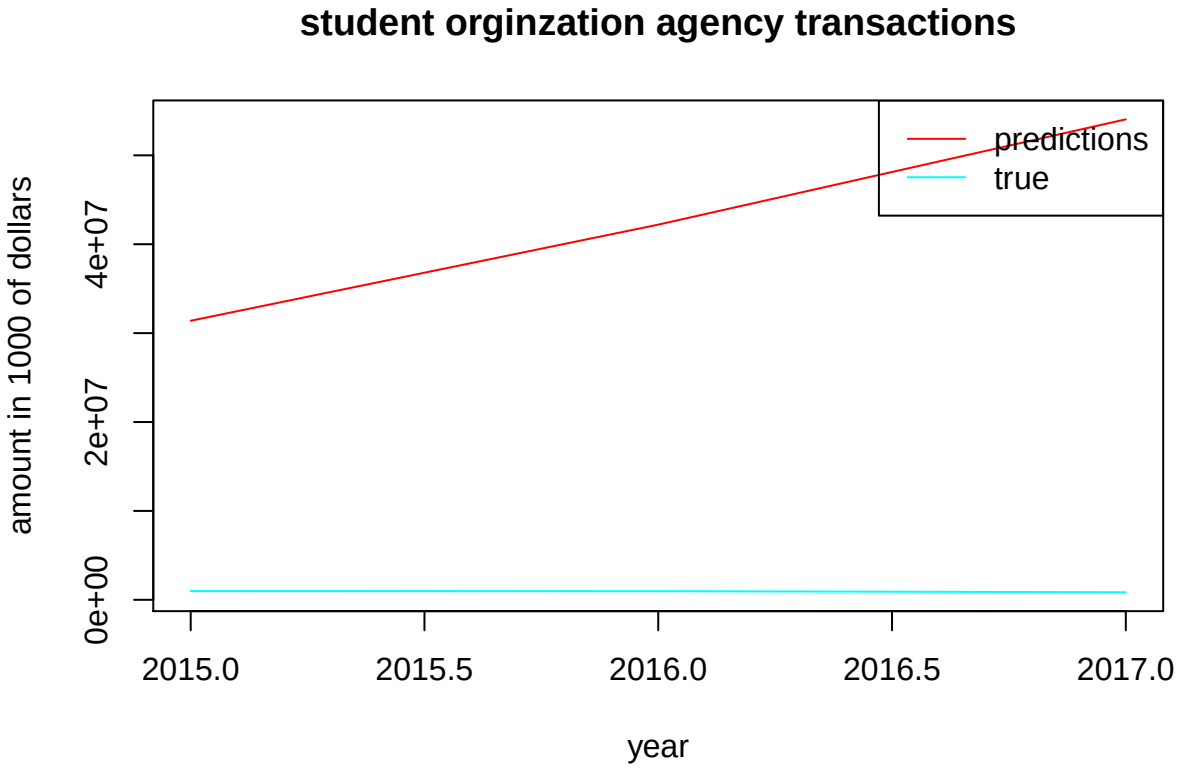
We then try to fit the model over our training set an we found that only sumplimental disclosures of non cash activities, net decrease in cash and noncash equivalents and cash flows from capital and other financing activities were accurate predictors with sumplimental disclosures of non cash activities being the closest to the true values.


sumplimental disclosures of non cash activities
amount in 1000 of dollars
860000
900000
940000
980000
2015.0
2015.5
2016.0
2016.5
2017.0
year
predictions
true


net decrease in cash and noncash equivalents
amount in 1000 of dollars
860000
900000
940000
980000
2015.0
2015.5
2016.0
2016.5
2017.0
year
predictions
true

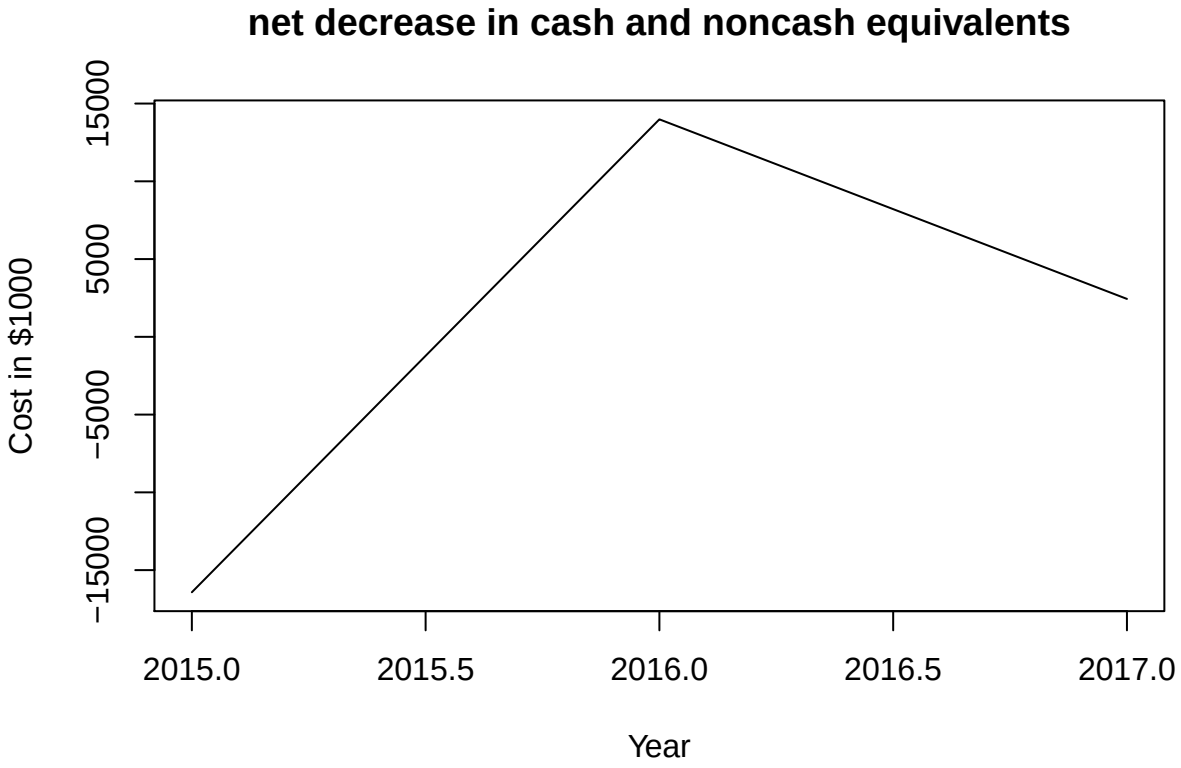

cash flows from capital and other financing activities
amount in 1000 of dollars
860000
900000
940000
980000
2015.0
2015.5
2016.0
2016.5
2017.0
year
predictions
true

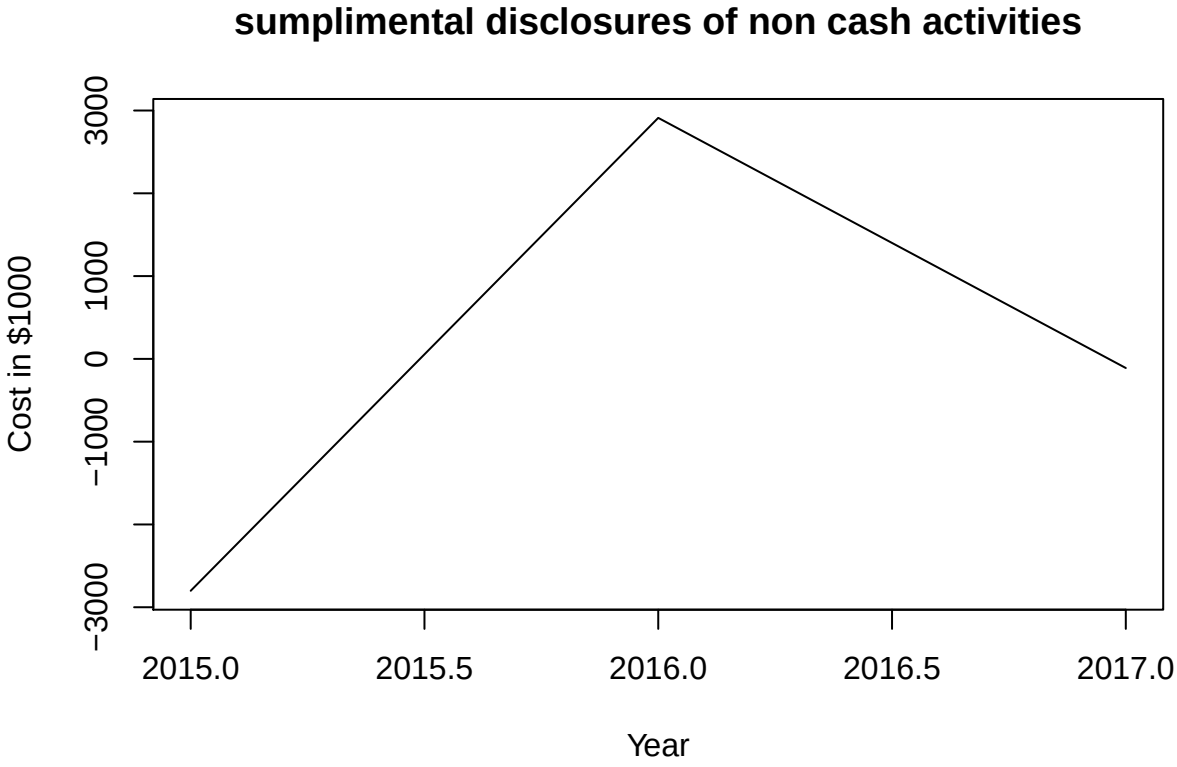

student orginzation agency transactions
amount in 1000 of dollars
800000
900000
2015.0
2015.5
2016.0
2016.5
2017.0
year
predictions
true

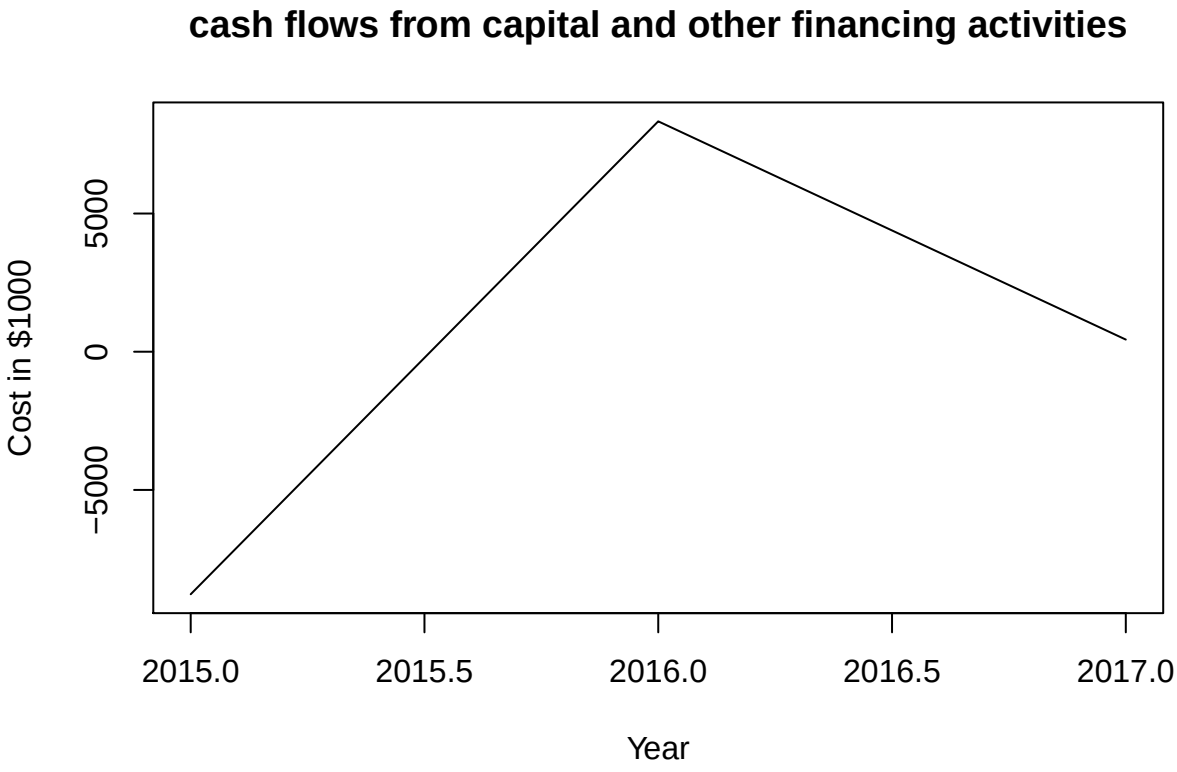


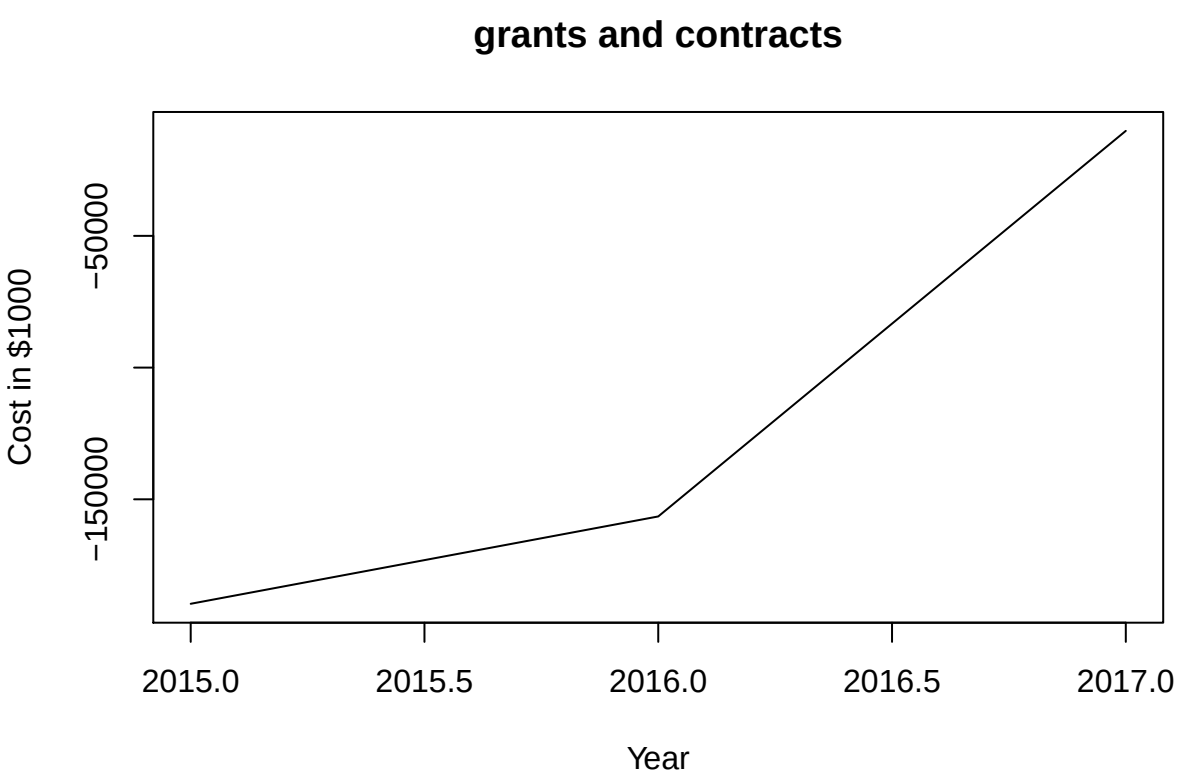
Upon inspection of residuals, we saw that the sum residuals of sumplimental disclosures of non cash activities, net decrease in cash and noncash equivalents and cash flows from capital and other financing activities all are essentially zero, while grants and contracts and student orginzation agency transactions are -356,311 and 12,429,573, respectively, showed poor under and over estimates. The root mean square error of sumplimental disclosures of non cash activities, net decrease in cash and noncash equivalents, cash flows from capital and other financing activities, grants and contracts, and student orginzation agency transactions were 12,532.33, 6,991.26, 2,333.271, 14,2085.6 and 42638079 in the same order. This once again showed grants and student orginzation revenue were extremely poor linear regression predictors. Keep in mind that these numbers were in 1000s of dollars so that showed upwards of millions of dollars in error.

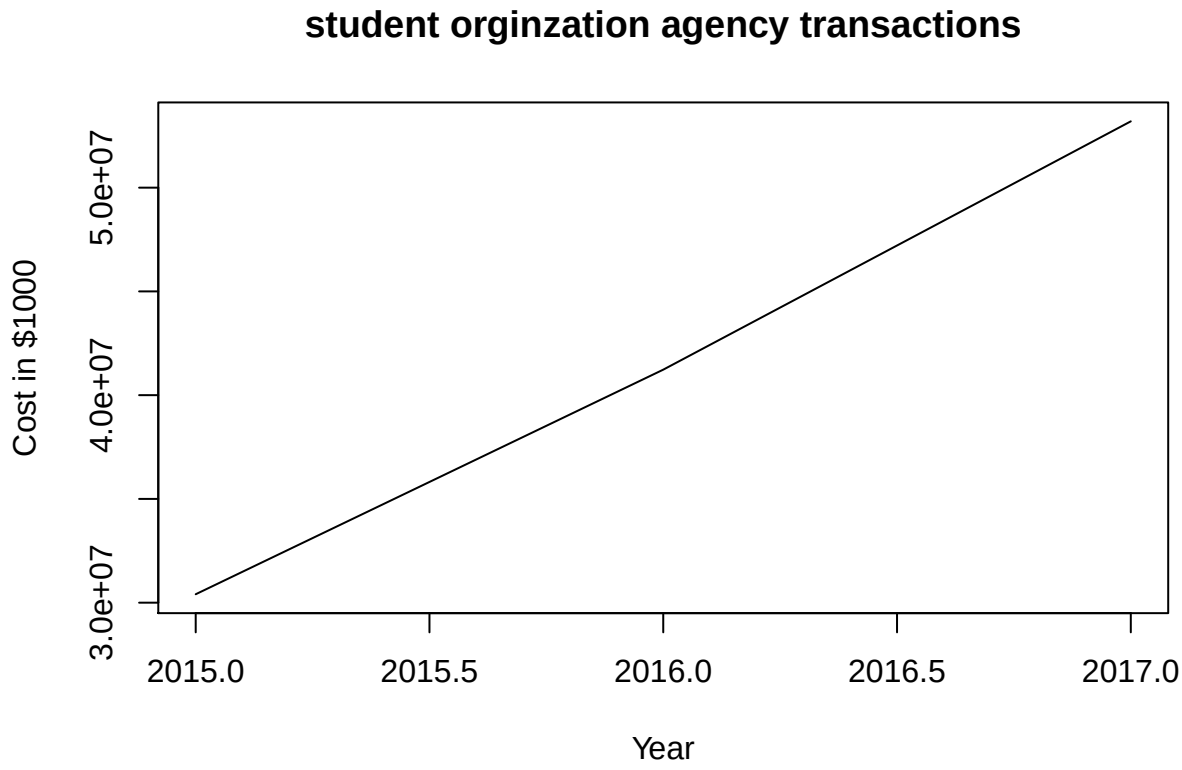
Lastly a final look at our residual plots enforced our previous observation of grants and contracts and student organization agency transactions as extremely poor predictors. In response we didn't continue and tested those two individual variables on our testing set.


net decrease in cash and noncash equivalents
Cost in $1000
15000
5000
−5000
−15000
2015.0
2015.5
2016.0
2016.5
2017.0
Year

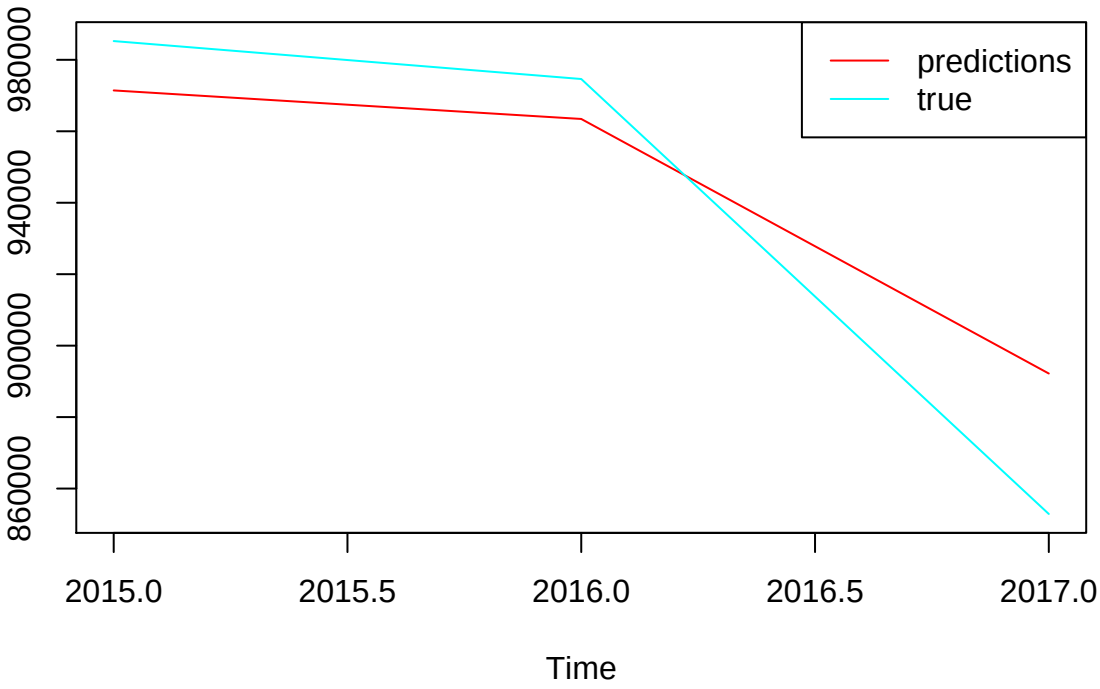

sumplimental disclosures of non cash activities
Cost in $1000
3000
1000
0
−1000
−3000
2015.0
2015.5
2016.0
2016.5
2017.0
Year


cash flows from capital and other financing activities
Cost in $1000
5000
0
−5000
2015.0
2015.5
2016.0
2016.5
2017.0
Year

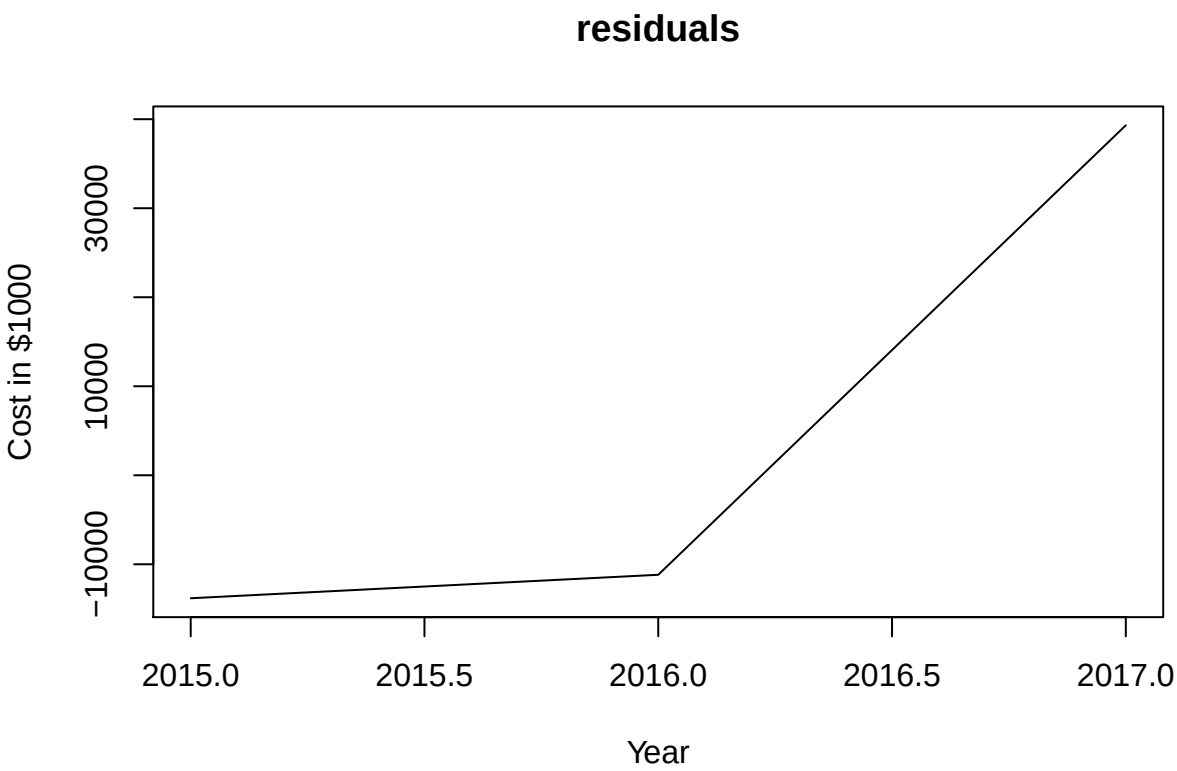

grants and contracts
Cost in $1000
−50000
−150000
2015.0
2015.5
2016.0
2016.5
2017.0
Year



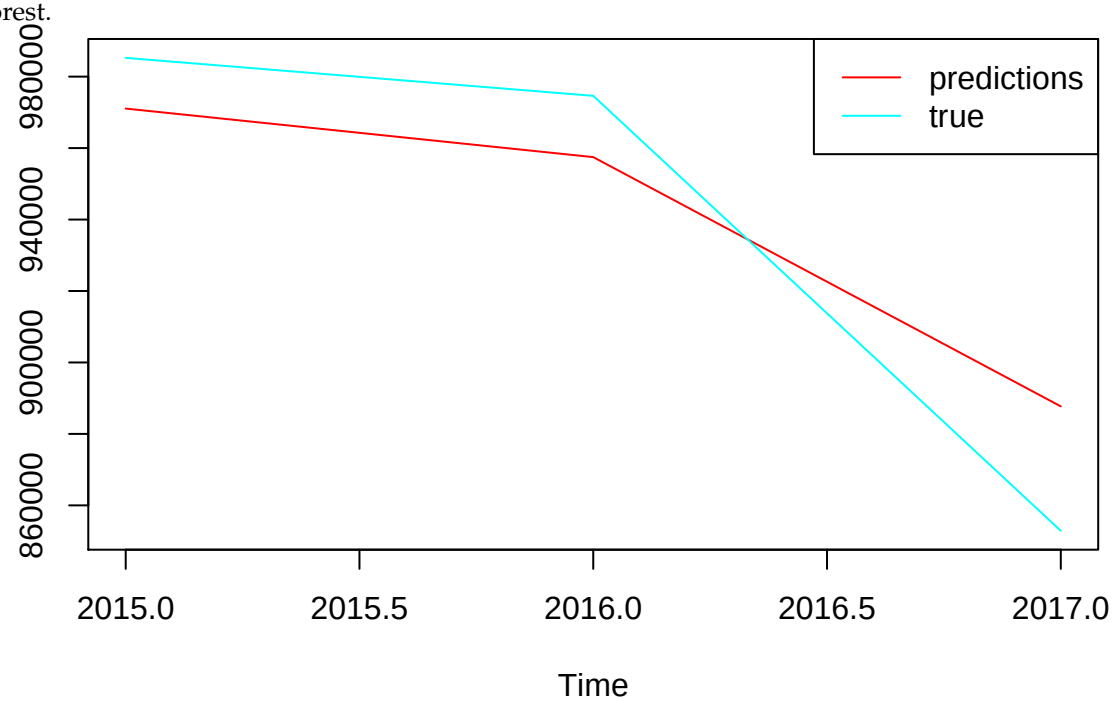
Lets take a look at another approach using random foresting. We will attempt to forest our net cash flows using 4 partitions and 500 trees. Due our small dataset, we are unable to provide the importance of the variables but the model was successfully created. We do see that the model is a fairly accurate predictor for the true values but not as strong as the linear regression models.



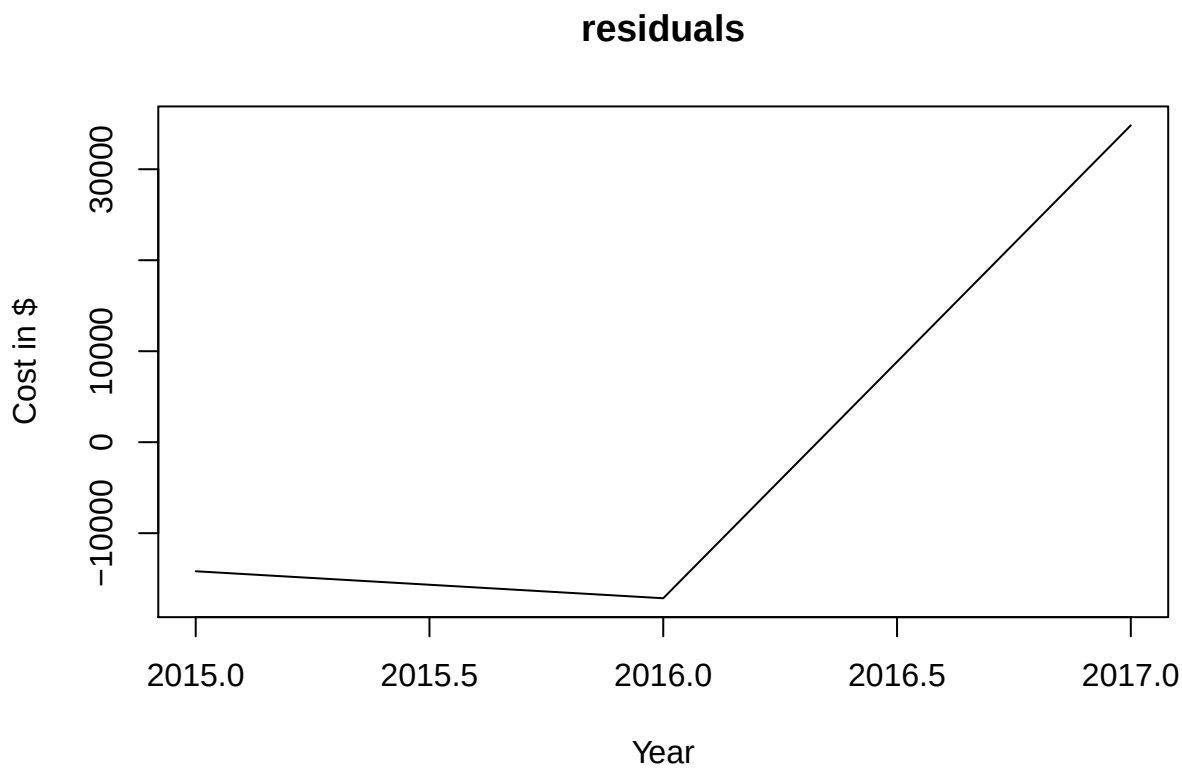
However we see that the root mean square error is an outstanding 20,829.5 thousands of dollars which is huge! The total resiudual is only 1,588.296 thousands dollars however which is signifcantly smaller. We cannot ignore this variance, which is reflected in our residual time seris.



We will also be checking to the cash flows from non financing activities to see if it would a better model, this time using 5 partitions instead of 4. However the same resulting variance seem to be present in this forest.



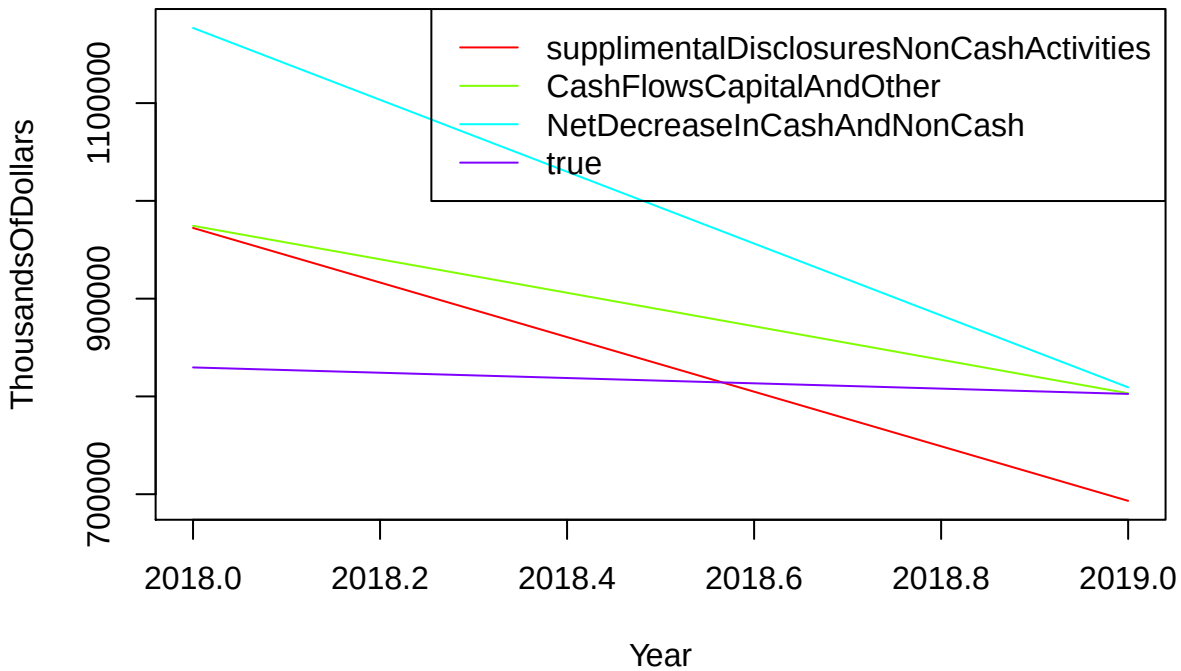
We see that the residuals remain roughly the same distrubtion but the variance does seem slightly higher with a root mean square error of 26,378.29 thousands of dollars and a sum of residuals at 9930.44 thousands of dollars.



Regardless we shall continue to use these models inour testing phase and hope to remain within the same margin of ever.

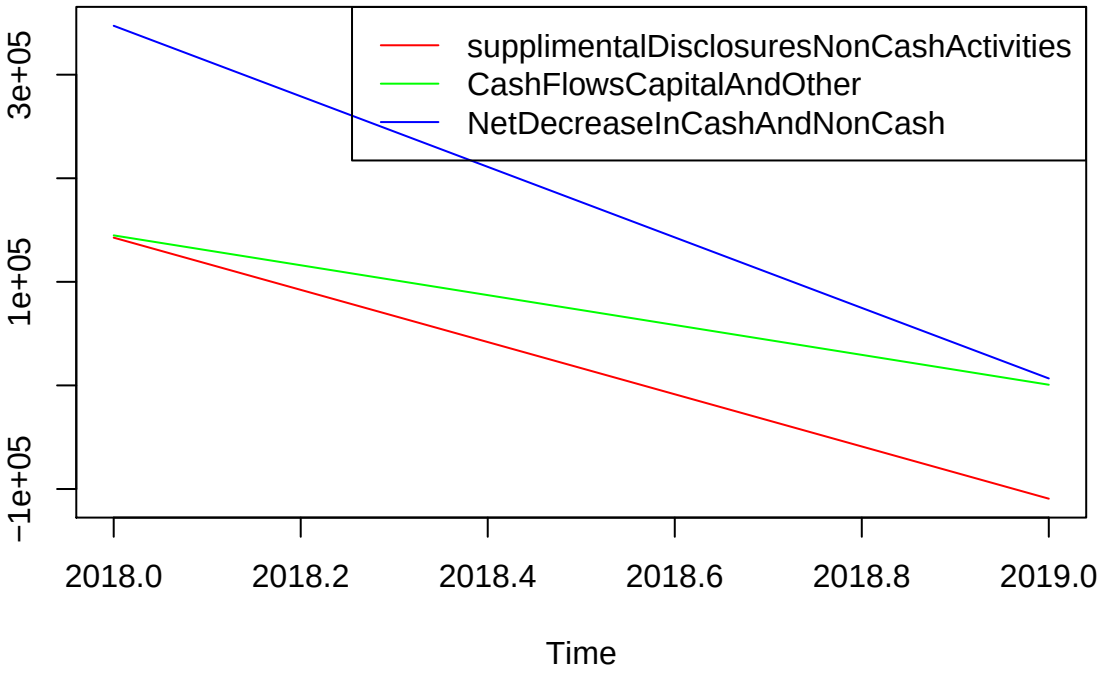
Results

Testing the predictive capabilities of our linear models, we can easily see that none of the models come even close to our true values. These models fit poorly because they produce a lot of error (false values).

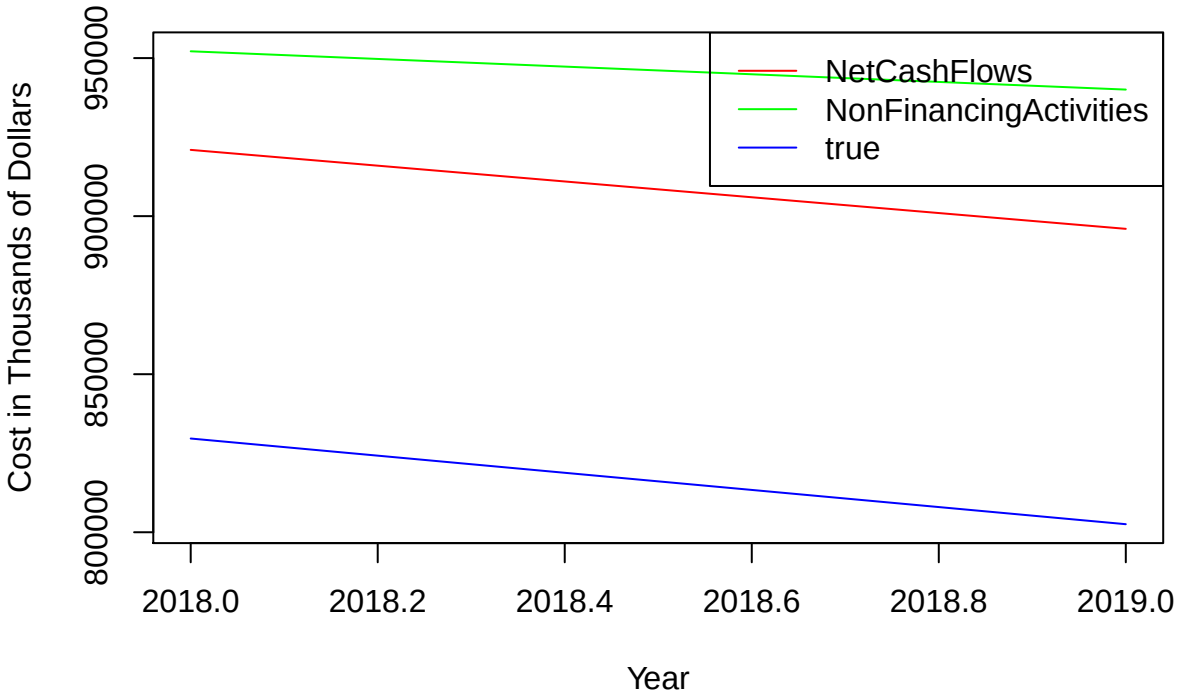


We can even both quantify and plot how abysmal the fits of these linear models are by computing the residual mean squared error. As you can see, the average error of each linear model is astronomically high

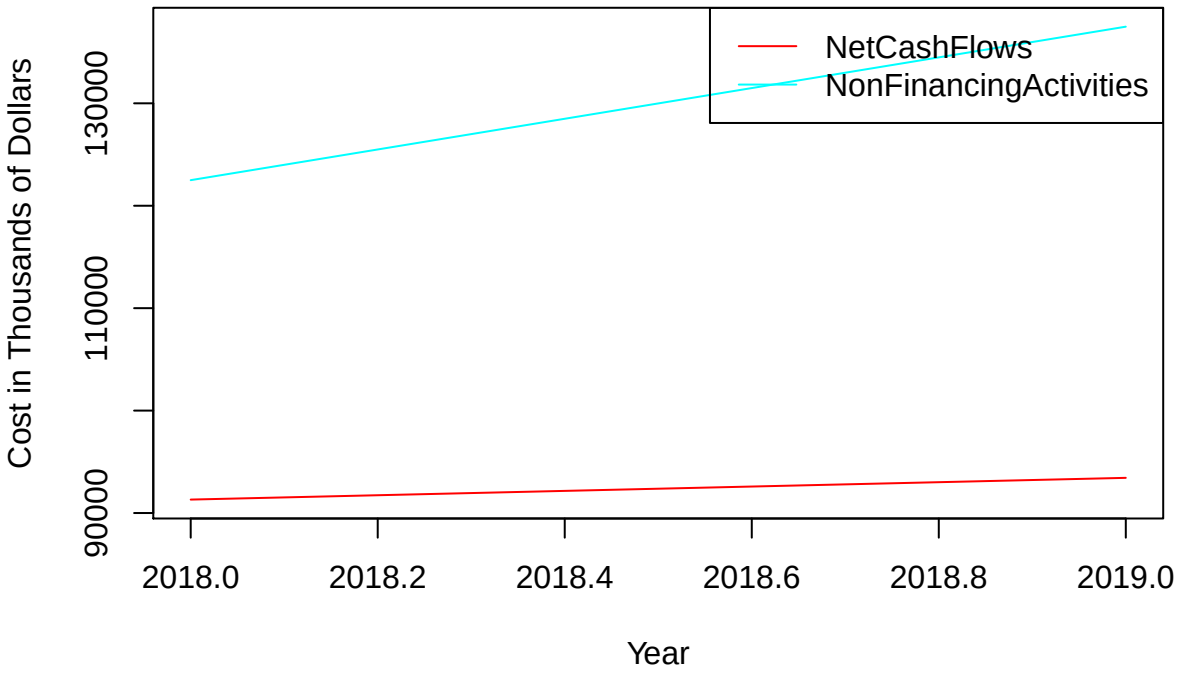
(Billions of dollars in error).



We can run similar tests for out random forest tree model. This model fits less poorly than it's linear regression counterpart since the predictions are off by a lesser degree. This means that the error produced by this models produces lower amounts of false values.



Surely enough, once we plot the residuals, the random forest model produces lower error than the linear models. Respectively, the calculated RMSE is lower than that of the linear models. Meaning, although this model still produces extremely poor predictive capability, it still proves better than our linear model.



Conclusion

We were unable to draw a clear correlation between tuition & fees and other cashflows. none of our models using significant factors produced an RMSE low enough to be accepted. Simply put, at first glance, factors such as Supplemental Disclosure Of Noncash Activities seemed to affect tuition and fees by about a $933 USD decrease in tuition and fees per $1000 increase. However, when the model's fit was eventually tested, it failed horribly by producing astronomically high RMSE values. This seems logical since we must also keep in mind that UMass doesn't directly choose their tuition as that is the decision of the Department of Education.

Taking all of these into consideration, it's undeniable that other cashflows preside no affect over UMass' revenue from tuition and fees. Therefore, we must reject our hypothesis. There is a no significant relationship between other cashflows for the University of Massachusetts system and the cost of tuition and fees.

References

"Reports & Audits." Reports & Audits | University of Massachusetts Office of the President

Thys, Fred. Study: Out-Of-State Students With Lower Test Scores Admitted To UMass Amherst More. WBUR, 29 May 2018

"U.S. Bureau of Labor Statistics." U.S. Bureau of Labor Statistics, U.S. Bureau of Labor Statistics, 28 Apr. 2020