Predicting Graduation Rates at Dutch Secondary Schools

Matthijs Toorenburg
Research Paper Business Analytics

2015-07-03

Supervisor: B.T. Knapik
Vrije Universiteit Amsterdam
Faculteit der Exacte Wetenschappen
Area of Business Analytics
De Boelelaan 1081a
1081 HV Amsterdam

Abstract

Almost all students in the Netherlands finish their Secondary Education by doing their Central Exams. Schools use the graduation rates of these exams as educational quality measure, and thereby try to attract new students to their schools. But are graduation rates a reliable quality measure? What if a graduation rate could be predicted according to some features. Then a school could positively influence the graduation rate by increasing the positive correlated features. But these positive correlated features are not always a measurement of educational quality, in fact, it could be opposite. This research has modelled the graduation rate as a logistic model based upon a set of features. Most important related features are the average School Exam grade, demographic background of students and succeeding rates. Schools are able to influence the average School Exam grade and the different succeeding rates, thereby making it possible to positively influence graduation rates. These related features should be evaluated when ranking schools.

Preface

This research paper has been written as part of my master Business Analytics at the Vrije Universiteit of Amsterdam. Part of this study was to perform a scientific research on a chosen topic. I chose this topic because it annoyed me that many people were talking a lot about graduation rates and almost never about other characteristics of the school. Last, I wish to thank Mr. Bartek Knapik for his supervision during this research.

Contents

Chapter 1

Introduction

Every citizen in the Netherlands is obligated to finish secondary school. Besides being obligated, most of us pursue to finish secondary school at our maximum level. To achieve this, students want to go to high quality schools which match their preferences. But what is a high quality school, and how is this measured? Many schools try to attract new students by advertising with graduation rates. But is this trustworthy? Is a high graduation rate really an indicator of educational quality? For example, which school is better? A school with a graduation rate of 100%, but where on average students went to school for 7 years. Or a school with a graduation rate of 90%, where all students went to school for 6 years?
In a published news article from 2013 Omroep Brabant claims that the Prinsentuin College from Andel is the best secondary school of Noord-Brabant, because it has the highest graduation rate. But everyone acknowledges that several factors, like the background of students, influence the graduation rates. If the graduation rate of a secondary school could be predicted, a school would be able to increase the graduation rate by increasing positively related features or decreasing negatively related features. But a school could not only increase its graduation rate by improving its educational quality, but also in other ways, like increasing the amount of students which descend to a lower level. Then this school would obviously not be the best school, but it would have a higher graduation rate. Several organisations publish a ranking of secondary schools every year. These lists include different features of secondary schools, like the financial details and students background. But no research has been done on what features exactly influence the graduation rate, and to what extend. This knowledge could be used by anyone who is interested in rating secondary schools. If the graduation rates could be predicted, it would also be possible to determine whether a school over/underachieved with respect to the graduation rate.
This report will try to discover what influences the graduation rate of secondary schools, and how this could be used to determine whether a school over/under-achieved with respect to the graduation rate. This report will first focus on what features are used by the different rankings to determine which school is best. Next the relationships between these features and the graduation rates will be investigated. Last, a model will be constructed which predicts the graduation rate and this model will be interpreted.

Chapter 2

Background

The Dutch educational system could be largely separated into three parts, first there is the primary education for children between 4 and 12. Next there is the secondary education for children between 13 and 18 and after there is higher education. Secondary education exists of 4 different levels (low level to high level): Praktijkonderwijs, VMBO, HAVO and VWO. Where VMBO can be divided in 4 different levels itself: BBL, KBL, GL and TL. Because Praktijkonderwijs does not involve any exams it was excluded from this research. Praktijkonderwijs and VMBO last 4 years, HAVO lasts 5 years and VWO lasts 6 years. For students to graduate, they need to pass their School Exams (SE) and Central Exams (CE). A student with low Central Exam grades can compensate this partially by high School Exam grades.

Most Dutch Secondary Schools are part of an Educational Institution. Schools inside these Educational Institutions join their forces to be able to work more efficient. Not all schools offer all different education levels. Some schools specialize in high level education, where other schools only offer low level education.

Chapter 3

Literature study

3.1 School Rankings

Several organisations rank secondary schools according to their performance. These organisations all use different criteria and obtain different results. This chapter will analyse which features are used to rank secondary schools. 4 lists will be analysed: *Keuzegids Middelbare Scholen, Elsevier, Excellente scholen, Schoolprestaties.*

Keuzegids Middelbare Scholen is a list which is published since 2014 by Keuzegids. This list grades each school according to 6 different features on a 5-point scale. The first feature is the ratio between the amount of students and the amount of staff. To be able to score the schools, 5 intervals where created and each school gets the points related to its corresponding interval. These 5 intervals differ for certain types of schools. The second score is based upon the financial situation of the school. The financial situation is rated according to 4 characteristics of the financial situation: liquidity, solvability, resistive power and profitability. A school is penalized if these factors are too low (financial instability) and if these factors are too high (schools should not hoard). The 4 other features are education related. The first one is the success of students in their first 2 years. Where the success is determined by the amount of students which proceed to the next year at the same level, or at an even higher level. The background of students is taken into account. The second education factor is the success of students after the first 2 years. Schools get additional points if they have students from"poor neighbourhoods", student which switched from school or students which get additional support (LWOO). The third score is the average grade of the Central Exams and the last score is the difference of the average grade of the School Exam and Central Exam. The 4 education grades each count for 15%, the other grades for 10%.

Elsevier has created a secondary school ranking for 14 years. It also grades school according to a 5 point scale. The rating of schools is based upon a number of scores over the last 3 years. Some of these scores are combined such that 4 scores remain at end. The first score is the success of students in the first 2 years, where the success is determined by the percentage of students which finish the first 2 years without delay, and the level of students after 2 years compared to the level at the start of the first year. The second score is the percentage of students which graduated their last 2,3 or 4 (dependent on level) years without delay. The third score are the exam scores, consisting of: average School Exam grade, average Central Exam grade, difference of School Exam and Central Exam, average grades of Dutch, English and Mathematics and the graduation rate.

Excellente scholen is a list containing schools which performed excellent, according to the school inspection. A jury determines whether a school receives the predicate "Excellent school" by taking into account the following criteria: Education results (with respect to amount of students), clear vision at education, self-learning ability of school, conditions of the school and the way in which the school distinguishes itself from other schools.

Schoolcijferlijst was a ranking created from 2006 until 2013 by Jaap Dronkers. Dronkers graded each score based upon 4 scores, which were all based upon a set of features. The first score was the School Exam grade. The second score is based upon the average grade of each course, where additional points are added for the percentage of students which failed their exams. The third score are bonus-points. This bonus points are based upon the following features: average School Exam, average Central Exam, graduation rate, percentage of students which graduate without delay. This score is then compared with the expectation, which is based upon the start level of students, the percentage of students which switched to this school and the socio-economic background of the students. The last part of the score is based upon the difference between the average School Exam and the average Central Exam. Additional points are added/subtracted if a lot of courses have a difference larger than 0.5 between the average School Exam and average Central Exam.

These 4 rankings all use different measurements and transformations. Besides, top performers in one list are below-average performers in another list. On top of this, schools which perform well in one year, perform worse the next year.

3.2 Statistical Methods

Fitting a statistical model to a dataset is a multistep process which involves different techniques. The dataset can be summarized by graphical and numerical displays. Examples of graphical displays of the data are histograms, boxplots and scatter plots. Examples of numerical displays are mean, mode, variation, quartiles, skewness and kurtosis. This summarizing is done to know what data is available and what are its characteristics.
There are several tests which can be used to determine whether a dataset stems from a certain distribution. First there is the Shapiro-Wilk test. This test can be used to determine whether a dataset stems from a normal distribution. The Kolmogorov-Smirnov test can be used to determine whether a dataset stems from a "simple" distribution, where simple means one specific distribution, like the Exp(5) distribution. The Kolmogorov-Smirnov is not able to deal with location-scale families of distributions. When testing for normality, the Shapiro-Wilk test tends to be stronger than the Kolmogorov-Smirnov test. The Chi-Square Goodness of fit test can be used to determine whether some categorical data stems from a certain distribution. Downside is that at least 5 measurements are needed per category.
Some of the attributes could be correlated to each other. It is important to know which attributes are correlated to each other, because the correlation between attributes could influence the accuracy of the results. Different correlation tests are Rank Correlation test of Spearman, Kendall Correlation and Pearson Correlation. Disadvantage of Pearson's Correlation is that it assumes normality of both datasets, linearity between variables and homoscedasticity of data.
The objective of this paper is not only to discover relationships in the dataset, but also to construct a model which can predict graduation rates. There are different models which can be used during this process. First there is Linear Regression. Assumptions of Linear Regression is that the remaining errors are independent and normally distributed. Besides, it is important to test whether there is a linear relation, and not a different relation, like quadratic. Another model is Analysis of Variance (ANOVA). ANOVA can be used for categorical data to determine whether different categories stem from the same distribution. ANOVA does the following assumptions: independence of measurements, normality and independence of the residuals and homoscedasticity. Analysis of Covariances (ANCOVA) is a model which blends ANOVA and regression. Important assumptions are linearity of regression, homogeneity of error variances, independence and normality of errors and homogeneity of regression slopes. The last model which can be used are the Generalized linear models (GLM). These models allow the errors to be of any distribution of the exponential family, not only the normal distribution. The 4 assumptions of GLM are, independence of each data point, correct distribution of the residuals,

correct specification of the variance structure and a linear relationship between the response and the linear predictor.

Chapter 4

Data Exploration

The graduation rate at secondary schools could be depending on several variables. This chapter contains information about these variables and their interrelations. Data was gathered from 1414 schools for year 2013. The investigated dataset contained a lot of different attributes: demographic information about the school, educational information about the school, financial information about the school, demographic information about the students. Only 2 features were categorical, all other were numerical. Many of these numerical attributes were included for all educational levels, for instance, the amount of students which participated in the Central Exams was included for all 6 educational levels. Appendix A contains a detailed description of all attributes in the final dataset.

There are 57 numerical attributes which had a significant correlation ($\alpha = 99\%$) with at least one of the graduation rates. A lot of these relationships were expected, such as average School Exam grades and graduation rates. For VMBO BBL, KBL and TL the respective graduation rates were negatively correlated with the amount of exam participants. More interesting was that the graduation rates itself were correlated with the levels above and below. For instance, VMBO TL's graduation rate had a positive correlation of 0.1731 with VMBO KBL's graduation rate and 0.2197 with HAVO's graduation rate. Most features from the liveability survey were correlated to the graduation rates. *bev, veilig* and *won* were significantly correlated with all graduation rates. Besides, these liveability scores were mutually correlated.

The dataset contained two categorical features: provinces and Educational Area Codes. The graduation rates differed significantly for some of the provinces: The graduation rates of state Noord-Brabant were significantly higher than the graduation rates of state Noord-Holland for levels VMBO TL and VWO.($\alpha = 95\%$) When looking at the Educational Area Codes, similar results were found for levels VMBO KBL, VMBO TL and VWO. This significant differences could also exist between other provinces and Educational Area Codes, but because there were too little data points for some provinces and Educational Are Codes, the confidence intervals were quite large and it was invalid to draw these conclusions.

At this moment, the dataset consisted of 139 attributes. Because both Spearman's and Kendall's correlation coefficients are generally accepted, I chose to calculate them both and average them to obtain one coefficient and p-value for all combinations of the numerical attributes. 2723 combinations of attributes, about 27% of the total combinations, had a significant correlation($\alpha = 99\%$). For several reasons, not all of these attributes could be used in the eventual model. For example, the dependent variables were included in the dataset.

All graduation rates are values between 0 and 100%. The distribution of the graduation rates should be a finite distribution. One of the possibilities would be that the graduation rates are Beta-distributed. The Beta-distribution depends on 2 different parameters α and β. These parameters could be estimated by Maximum Likelihood Estimation (MLE) and by the method of moments. Below are the formulas for the method of moments:

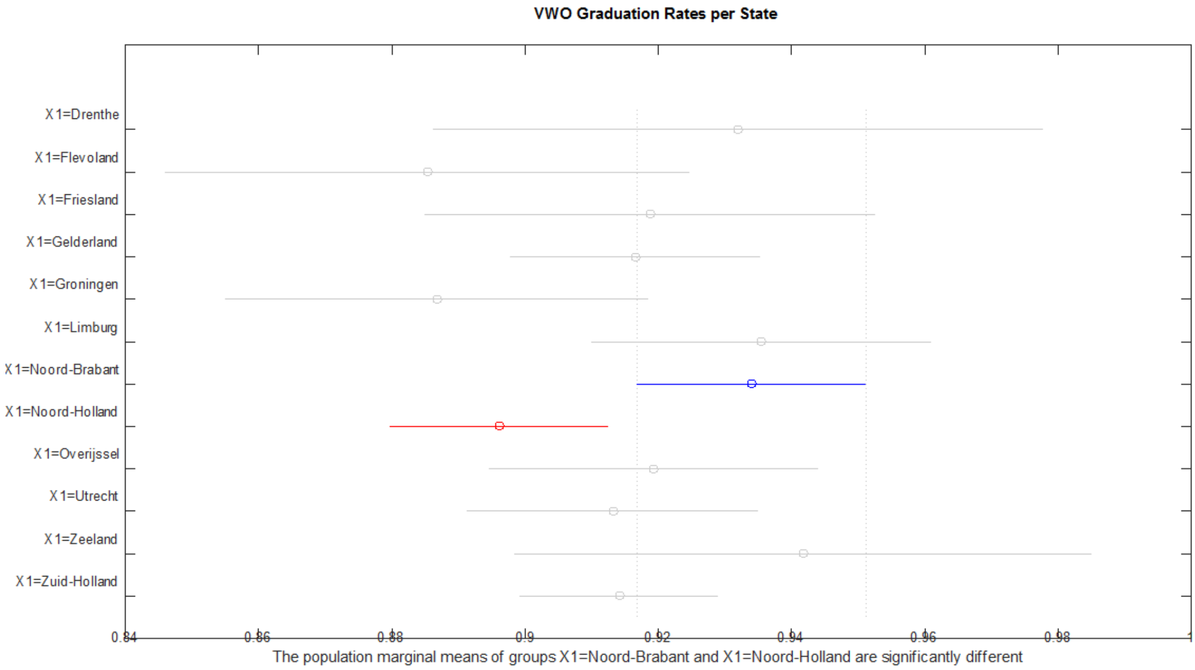

Figure 4.1: VWO Graduation Rates per State

$$\hat{\alpha} = \bar{x}\left(\frac{\bar{x}(1-\bar{x})}{\bar{v}} - 1\right),\ if\ \bar{v} < \bar{x}(1-\bar{x})$$
$$\hat{\beta} = (1-\bar{x})\left(\frac{\bar{x}(1-\bar{x})}{\bar{v}} - 1\right),\ if\ \bar{v} < \bar{x}(1-\bar{x})$$

Where $\bar{x}$ is the sample mean and $\bar{v}$ is the sample variance. For all 6 graduation rates, a sample was drawn 1000 times from the set of graduation rates, and α's and β's where estimated for each sample by MLE and method of moments. The following hypothesis was then tested with the Kolmogorov-Smirnov test:

$$H_0 : \mathbb{P} = B(\hat{\alpha}, \hat{\beta}), \tag{4.1}$$
$$H_1 : \mathbb{P} \neq B(\hat{\alpha}, \hat{\beta}) \tag{4.2}$$

H_0 was rejected 1000 times for all levels when using MLE as estimator. When using method of moments, all VMBO levels were rejected 1000 times($\alpha = 95\%$), so there was no evidence of any of the VMBO graduation rates being Beta-distributed with parameters estimated by the method of moments. HAVO was rejected 666 times and VWO 660 times, so these graduation rates could be Beta distributed, but there is no strong evidence towards it.

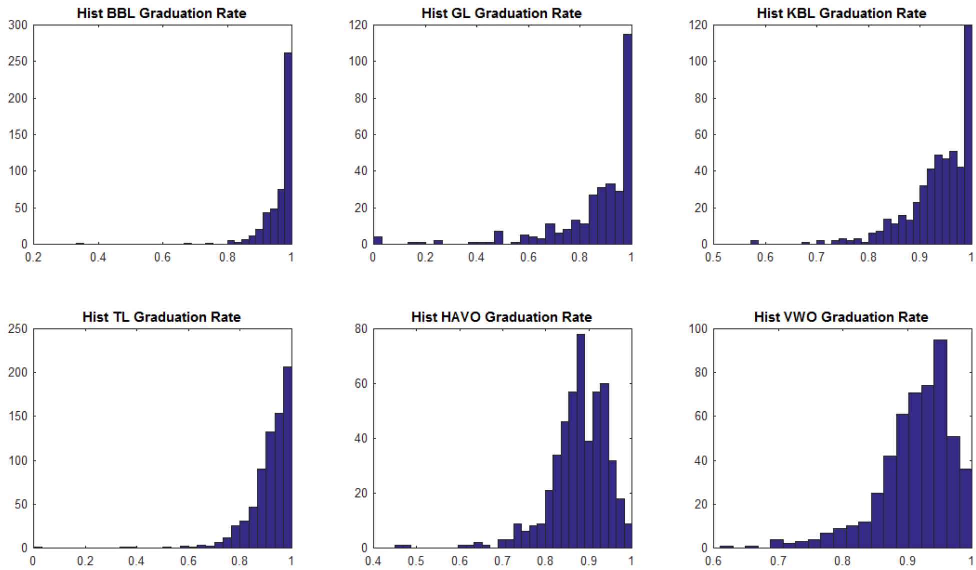

Figure 4.2: Graduation Rates per Level

Chapter 5

Modelling

The distributions of the graduation rates were unknown, so only methods which were not depending on a certain class of distributions could be used. Generalized Linear Models (GLM) allow models to have non-normal dependent variables. GLM maps a linear predictor onto a certain interval by using a certain link function. One extension of GLM is logistic regression. Logistic regression maps a linear predictor onto $[0, 1]$. This can be done by using a logit or a probit link function. There are different techniques to determine whether to include a variable in the model or not. At first I tried to start with the 40 most correlated attributes and fit a model onto these attributes. Next I stepwise removed all attributes which were not significantly related. This method resulted in an empty model for all levels. Therefore I chose to evaluate all combinations of the 15 most correlated attributes with respect to the specific graduation rate and then save the 10 models with the lowest AIC. I chose 15 attributes because of computation time. Besides, I chose to only evaluate attributes which corresponded to their own level. So for predicting the graduation rate of VMBO TL, no attributes from VMBO BBL were considered. This way it was prevented that the eventual model was only applicable on a subset of schools. For cross-validation, the dataset was divided into k-folds. For each fold, the model was trained on all folds except the current fold. Next predictions were made by this model onto the current fold. These predictions were then scored by Mean Square Error (MSE). The score of each model was calculated as the average MSE over all folds. The model with lowest MSE was further investigated.

The remaining model was tested upon how well it fitted the data. First, the Hosmer-Lemeshow goodness of fit test was performed. This test indicates whether a model is correctly specified. The Hosmer-Lemeshow test statistic is given by:

$$H = \sum_{g=1}^{G} \frac{(O_g - E_g)^2}{E_g} \tag{5.1}$$

O_g are the measured y-values, E_g are the predicted y-values and G are the number of groups. Then H is approximately distributed chi-quared with $(G - 2)$ degrees of freedom. Second, the linear predictor and the individual variables were plotted against the Pearson residuals. If one of these plots revealed curvature, this would be an indication that the linear predictor and/or one or more of the variables were not linearly related to the logit/probit of the dependent variable. Besides, these plots could reveal non-constant variance. Third, the model was checked upon overdispersion, and if present, the results were corrected for this overdispersion. Last, leverage points were detected using Cook's distances. If there were any errors in the data, these points were corrected or removed.

Chapter 6

Results

All final models and coefficients can be found in Appendix B.

6.1 VMBO TL

This chapter includes the results and procedures for level VMBO TL. First, all combinations of the 15 most significantly related attributes were evaluated. The 10 models with the lowest AIC were saved. All attributes which were not present in one of the 10 remaining models were checked upon a non-linear relationship between these attributes and the logit of the graduation rate. *VMBOTLDeelnemers* seem to have an increasing relationship with respect to the graduation rate. Several transformations were evaluated, but none had a positive effect, so this and all other attributes were removed. The remaining 10 attributes were plotted against the graduation rate to detect any non-linear relationships. No evidence were found for a non-linear relationship.

The remaining10 models had an AIC between 2496.09 and 2497.76. The dataset was now divided into 10 random folds. For each combination of fold and model, the model was trained upon all data except the current fold. This model was then used to make a prediction of the graduation rates of the current fold, and the MSE was calculated. Eventually, the average MSE per model was determined. Because the average MSE was depending on the separation of the folds, this procedure was repeated multiple times. Model 1 had the lowest average MSE and therefore I chose to continue investigating the first model. For model 1, the Pearson's residuals were plotted against the fitted values and against the variables. These plots could reveal inhomogeneity of errors and a non-linear relationship between the fitted value/variable and the linear predictor. For variables *onderbouwpercinstroom, TLpercafstroom, TLpercopstroom* and *VMBOTLSE*, the errors seemed to be heterogeneous. Data transformations could reduce this heterogeneity of errors. For all 4 attributes, the following 4 transformations were tried: square root of x, square of x, exp(x) and log(1+x). Each of these transformations was tested on significance. This was done by adding the transformed variable to the model, and comparing this model to the model without the transformed variable. This comparison was done by performing ANOVA with Chi-Square test. None of the transformed attributes were significant, so none were included. ($\alpha = 95\%$) The Goodness of fit of the remaining model was tested using the Hosmer-Lemeshow goodness of fit test with number of bins 5...15. For all tested number of bins, this test returned a p-value larger than 0.99, so the Hosmer-Lemeshow goodness of fit test does not indicate a lack of fit. The overdispersion of the model was equal to 2.11. Because this number is larger than 1, this influences the test statistics. These test statistics were adjusted for overdispersion. Figure 4 shows Cook's distances. The points with the highest Cook's distance were investigated and adjusted/removed if there were inconsistencies. After adjusting/removing these points, the model was again fitted to the data. Eventually, the following attributes were present in at least one of the 10 models: *VMBOTLSE, vrz, won, gt1e3e, onderbouwpercinstroom, TLpercover, TLpercafstroom, TLpercopstroom, pinstgt* and *publ*.

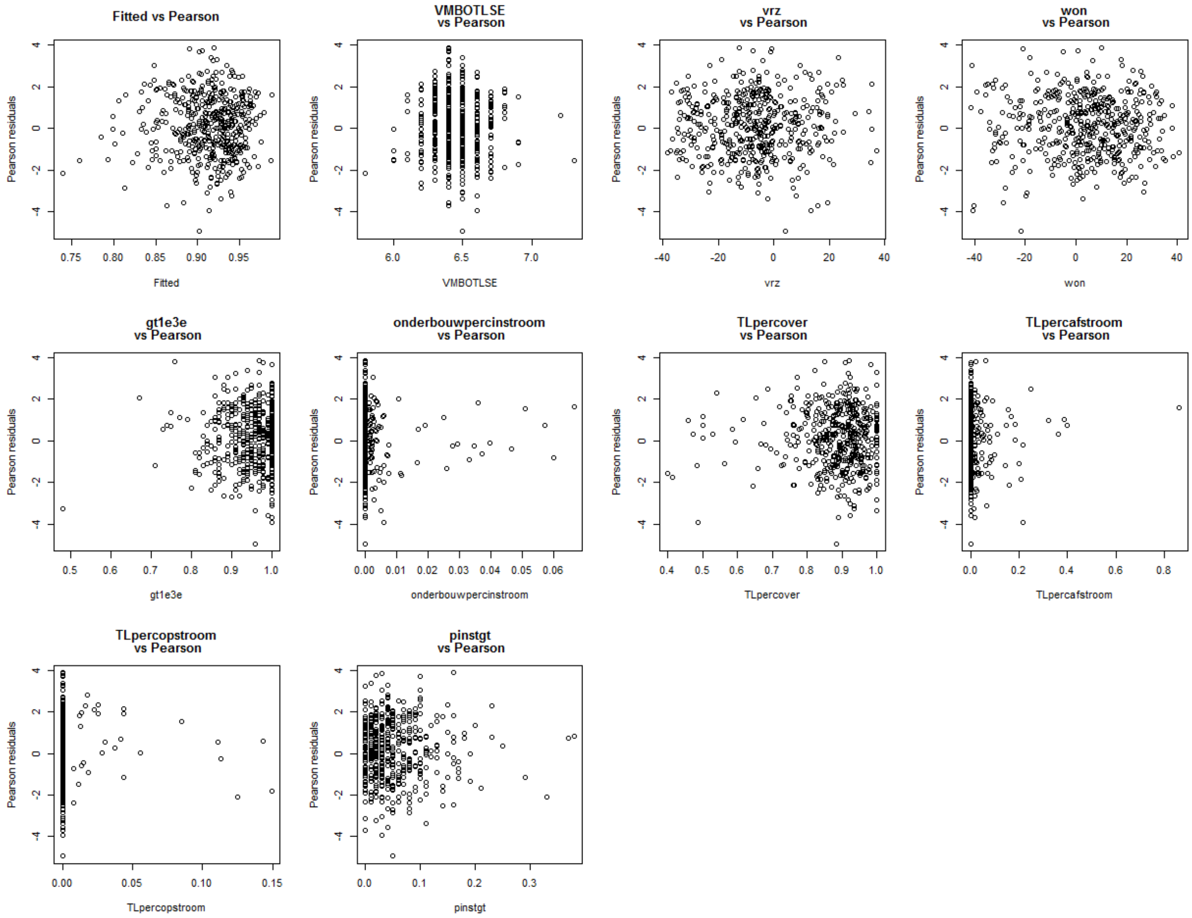

Figure 6.1: Pearson Plots

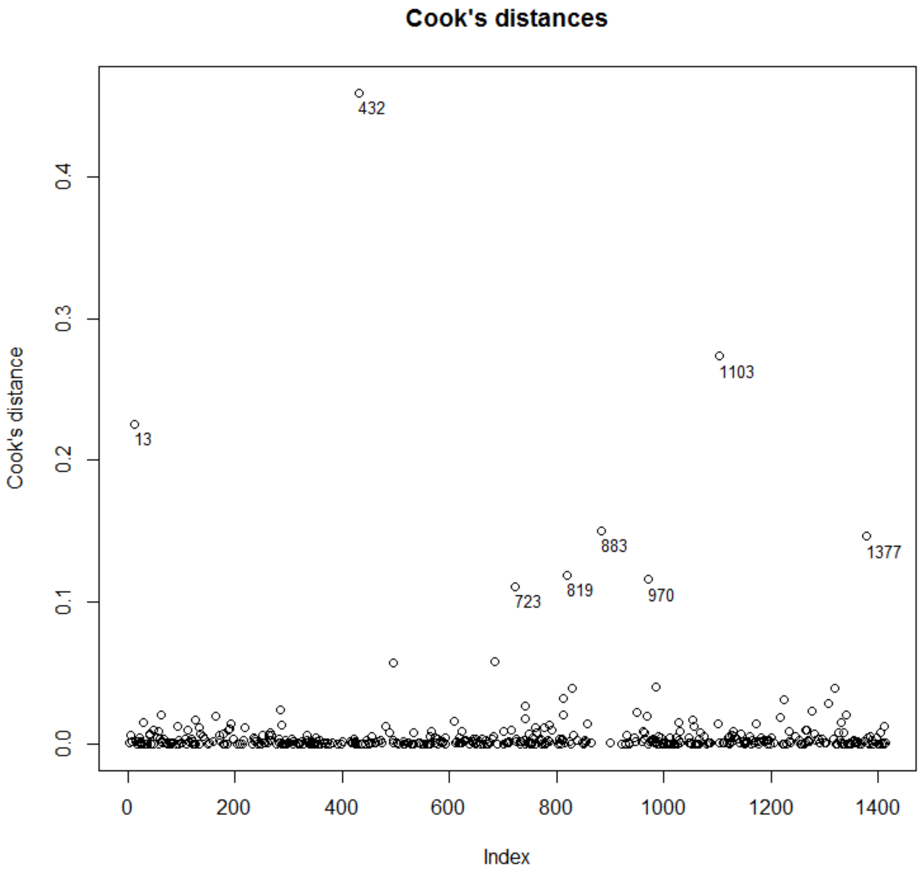

Figure 6.2: Cook's distances

Eventually there were 10 models left. Table 7.1 shows the statistics for the model, which had the lowest MSE after cross-validation. With the following coefficients and p-values (adjusted for overdispersion):

Variable	**Estimate**	**Std.Error**	**z value**	**p-value**
(Intercept)	-11.745525	1.342532	-8.749	<2e-16
VMBOTLSE	2.177046	0.1977	11.012	<2e-16
vrz	0.007254	0.002966	2.446	0.0144
won	0.015095	0.002559	5.9	3.64E-09
gt1e3e	-0.434109	0.489799	-0.886	0.3755
onderbouwpercinstroom	-4.109379	4.103958	-1.001	0.3167
TLpercover	0.582989	0.387533	1.504	0.1325
TLpercafstroom	1.815154	0.879693	2.063	0.0391

Table 6.1: Model VMBO TL

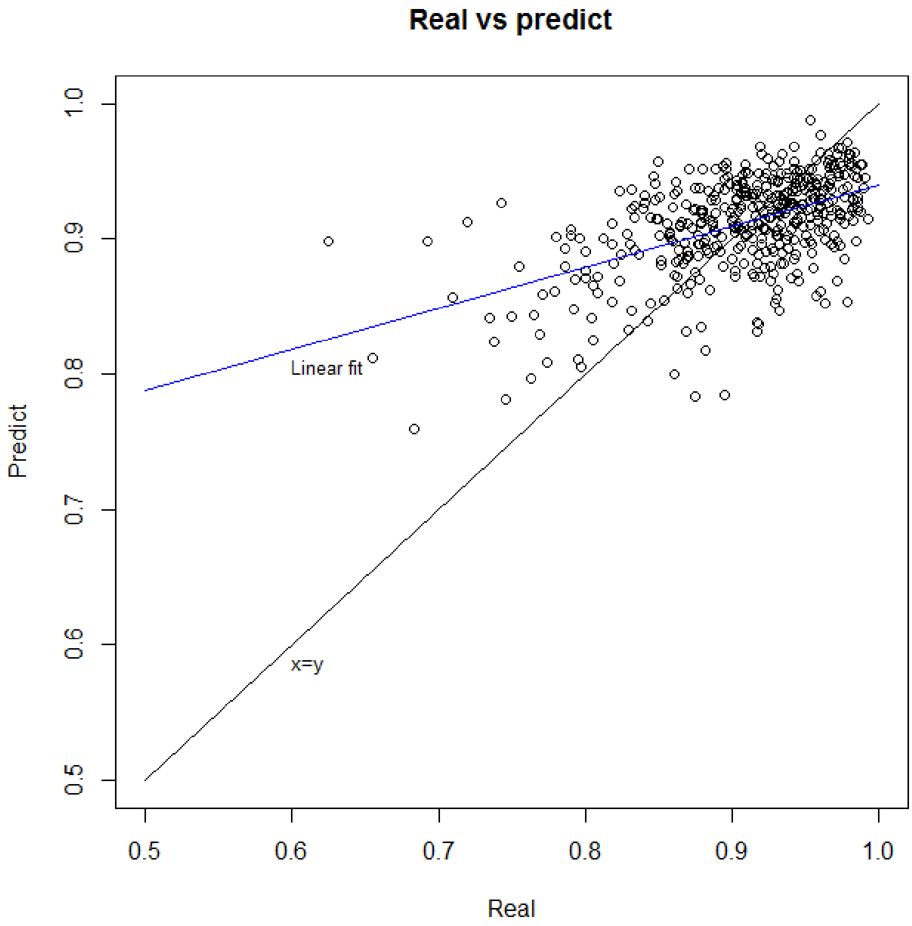

Figure 6.3: Prediction (Points with 100% graduation rate are removed from this figure)

6.2 Other levels

This section contains the main results for all other levels.

6.2.1 VMBO BBL

Analysed model 1, model had lowest MSE in cross-validation. No transformations, AIC's between 902.3195 and 904.263. Hosmer-Lemeshow goodness of fit test with number of bins 5...15 gave a p-value larger than 0.98 for all combinations. Overdispersion equal to 1.46. The following attributes were present in one of the 10 remaining models: *VMBOBLSE, onderbouwpercininstroom, BBLpercover, bev, onderbouwpercover, papcgba, VMBOBLDeelnemers, onderbouwpercuitstroom* and *veilig*.

	Estimate	**Std. Error**	**z value**	**p-Value**
(Intercept)	-12.7338	2.2761	-5.595	2.21E-08
VMBOBLSE	2.8086	0.3523	7.971	1.57E-15
onderbouwpercinstroom	-4.5246	2.4602	-1.839	0.0659
BBLpercover	-1.9821	0.8457	-2.344	0.0191
papcgba	-0.4337	0.191	-2.271	0.0232

Table 6.2: Model VMBO BBL

6.2.2 VMBO KBL

Analysed model 1, model had lowest MSE in cross-validation. No transformations, AIC's between 1549.661 and 1551.793. Hosmer-Lemeshow goodness of fit test with number of bins 5...15 gave a p-value larger than 0.99 for all combinations. Overdispersion equal to 2.30. The following attributes were present in one of the 10 remaining models: *VMBOKLSE, KaOnv, won, veilig, vrz, VMBOKLDeelnemers, kl1e3e, onderbouwpercuitstroom, KBLpercover, KBLpercuistroom, bev* and *onderbouwpercinstroom.*

	Estimate	**Std. Error**	**z value**	**p-value**
(Intercept)	-8.78513	1.886694	-4.656	3.22E-06
VMBOKLSE	2.058064	0.31822	6.467	9.97E-11
won	0.013317	0.005843	2.279	0.0227
veilig	-0.00967	0.004824	-2.005	0.0449
vrz	-0.00692	0.005018	-1.378	0.1681
VMBOKLDeelnemers	-0.0023	0.00121	-1.903	0.057
kl1e3e	-0.12486	0.755756	-0.165	0.8688
onderbouwpercuitstroom	-1.79316	0.839668	-2.136	0.0327
KBLpercover	-1.47173	0.611467	-2.407	0.0161
KBLpercuitstroom	-3.4059	2.093218	-1.627	0.1037

Table 6.3: Model VMBO KBL

6.2.3 VMBO GL

Analysed model 9, model had lowest MSE in cross-validation. No transformations, AIC's between 749.8611 and 751.2658. Hosmer-Lemeshow goodness of fit test with number of bins 5...15 gave a p-value larger than 0.53 for all combinations and larger than 0.72 for model 5. Overdispersion equal to 1.79. The following attributes were present in one of the 10 remaining models: *GLpercuitstroom, Papcggt, VMBOGLSE, onderbouwpercover, GLpercover, GLpercopstroom, onderbouwpercuitstroom, onderbouwpercinstroom, won* and *GLpercafstroom.*

	Estimate	**Std. Error**	**z value**	**p-value**
(Intercept)	-12.5554	1.788502	-7.02	2.22E-12
GLpercuitstroom	-7.66536	2.76997	-2.767	5.65E-03
won	0.016739	0.004029	4.155	3.26E-05
VMBOGLSE	2.271991	0.278914	8.146	3.77E-16
onderbouwpercuitstroom	1.263865	1.378998	0.917	0.3594

Table 6.4: Model VMBO GL

6.2.4 HAVO

Analysed model 2, model had lowest MSE in cross-validation. No transformations, AIC's between 2005.494 and 2006.788. Hosmer-Lemeshow goodness of fit test with number of bins 5...15 gave a p-value larger than 0.99 for all combinations. Overdispersion equal to 1.92. The following attributes were present in one of the 10 remaining models: *HAVOSE, veilig, papcgha , won, ha1e3e, onderbouwpercover, HAVOpercuitstroom, onderbouwpercinstroom, HAVOpercover, HAVOpercafstroom* and *HAVOpercinstroom.*

	Estimate	Std. Error	z value	p-value)
(Intercept)	-16.073	1.574624	-10.208	<2e-16
HAVOSE	2.928707	0.233114	12.563	<2e-16
papcgha	-0.55907	0.19422	-2.879	4.00E-03
won	-0.00247	0.002166	-1.14	2.54E-01
onderbouwpercuitstroom	-1.08201	0.558294	-1.938	0.05262
ha1e3e	-0.75561	0.459554	-1.644	0.10013
HAVOpercover	0.756873	0.577221	1.311	0.18978
HAVOpercopstroom	-0.99195	0.444289	-2.233	0.02557
HAVOpercafstroom	1.214046	0.44918	2.703	0.00688
HAVOpercuitstroom	2.216969	1.17978	1.879	0.06023

Table 6.5: Model HAVO

6.2.5 VWO

Analysed model 4, model had lowest MSE in cross-validation. No transformations, AIC's between 1909.896 and 1911.8. Hosmer-Lemeshow goodness of fit test with number of bins 5...15 gave a p-value larger than 0.98 for all combinations. Overdispersion equal to 1.67. The following attributes were present in one of the 10 remaining models: *VWOSE, bev, onderbouwpercover, VWOpercover, VWOpercafstroom, veilig, won, VWOpercinstroom, VWOpercuitstroom* and *vrz.*

	Estimate	Std. Error	z value	p-value
(Intercept)	-14.979	1.457053	-10.28	<2e-16
VWOSE	2.787254	0.226283	12.318	<2e-16
bev	0.011441	0.002666	4.292	1.77E-05
onderbouwpercuitstroom	-0.33716	0.788656	-0.428	6.69E-01
VWOpercover	-1.4481	0.774196	-1.87	0.0614
VWOpercafstroom	1.712407	0.882034	1.941	0.0522

Table 6.6: VWO

Chapter 7

Discussion

For each school level one model was constructed and evaluated which can predict the graduation rate of a certain school. Besides, there are 9 other models which are closely related to the final model. Because a logistic model is used, it is not possible to linearly relate the linear predictor to the graduation rate. Instead, the magnitude of several variables will be compared mutually. All 6 final models include the average School Exam grade for the specific model. These average SE grades all have a positive effect on the graduation rate, which suggests that schools should grade their students as high as possible. Currently, schools are penalized if the difference between CE and SE grades is too large. Besides, because of new CE regulations, it is not longer possible to fully compensate low CE grades with high SE grades.

All 6 models include one or more features which are related to the background of the students. *vrz, won, veilig, papcgba, papcgha* and *bev* are present in one or more models. *papcgba* and *papcgha* respectively have a negative effect on the graduation rate of VMBO BBL and HAVO. These features indicate the percentage of students from "poor" areas. *vrz, won, veilig* and *bev* are mostly positively related. When more than 1 of these attributes were present in a model, some of the coefficients became negative because of collinearity. So, when judging schools, the background of the students should certainly be taken into account. The highest coefficient for *vrz, won, veilig* and *bev* over all models is 0.0167. The range of these variables is [-50, 50], so the maximum difference of the linear predictor could be 1.67, which at maximum could result in an increase of predicted graduation rate of 40%. For the two lowest levels, VMBO BBL and VMBO KBL, the total amount of exam participants is negatively related to the graduation rate. But, these corresponding coefficients are very small and do not influence the graduation rate much.

Most inflow and outflow rates are negatively related to the graduation rates. High inflow would result in a lot of students which have to adapt to a new school, which would probably result in a lower graduation rate. Outflow of students may be do to several reasons: it could be that students move to another city. More interesting are the students which are leaving school because they are not satisfied with the school and the students which are expelled from school. A lot of students which leave school because they are not satisfied is an indicator that the quality of the school is low, so this would negatively influence the graduation rate. Students which are expelled from school are most times students which have low motivation. These students would have had a high chance of not passing their Central Exams because of motivation issues, therefore them being expelled would probably have a positive effect on the graduation rate. *HAVOpercuitstroom* has a positive effect on the graduation rate, where all other outflow rates are negatively related to the graduation rate. This could indicate that a lot of studens at level HAVO are being expelled from school. These in- and outflow rates of schools should be considered when judging schools, but more important, the underlying reasons of these outflows.

Another important factor is the amount of students which ascend to a higher level or descend to a lower level. If a lot of students ascend to a higher level, the graduation rate for the current

level decreases because students who would have probably graduated leave. For descending to a lower level it is the other way around. It is important to evaluate the percentage of students which ascend or descend to another level and to evaluate the amount of students which flow in from other levels. A school can increase its graduation rates by forcing a lot of students to descend to a lower level, which would not be an indicator of high educational quality.
Last, there are the percentages of students which pass their year. For low levels VMBO BBL, VMBO KBL and VMBO GL, the proceeding rate in the first 2 years is slightly negatively related or non-related to their respective graduation rates. For high levels HAVO and VWO, the succeeding rate in the first 2 years is positively related to the graduation rate. For low levels this could be because the students would have an additional year to graduate. For high levels it could mean that if a student cannot pass his first 2 years without delay, he/she probably won't graduate easily at all. For VMBO BBL, VMBO KBL, VMBO GL and VWO, the percentage of students which proceed to the next year is negatively related to the graduation rate. This is not surprising, if a student can go to school one year longer, the probability of graduating would increase. For VMBO TL, this is the other way around, which is quite surprising. Overall, schools have a lot of possibilities to positively influence their graduation rates. When judging a school, not only graduation rates should be evaluated, but much more other characteristics, and especially the reasons behind the numbers. Besides, it would be interesting to have specific data on the underlying reasons of in- and outflows.

The resulting models could be compared to the available rankings. *Keuzegids Middelbare Scholen* includes financial variables and the amount of FTE per student. None of these factors is significantly related to any of the graduation rates. Therefore there is no evidence that any of these factors is related to educational quality. *Elsevier* and *Dronkers Lijst* do not take into account the percentage of students which ascend or descend a level.

Bibliography

[1] Agresti, A. (2015).*Foundations of linear and generalized linear models.* Hoboken, New Jersey: Wiley.

[2] Breslow, N.E. *Generalized Linear Models: Checking Assumptions and Strenghening Conclusions.*(1996, July 31).

[3] Dronkers, J. (2014, June 6). *Schoolcijferlijst.* Retrieved June 25, 2015, from http://www.schoolcijferlijst.nl/

[4] *Elsevier - Beste Scholen.* (2015). Retrieved June 25, 2015, from http://bestescholen.elsevier.nl/

[5] *Excellente Scholen.* (2015). Retrieved June 25, 2015, from http://www.excellentescholen.nl/

[6] Hosmer, D., Lemeshow, S., & Sturdivant, R. (2013). *Applied logistic regression (Third ed.).* Joboken, New Jersey: Wiley.

[7] *Keuzegids Middelbare Scholen.* (2015). Retrieved June 25, 2015, from http://www.keuzegids.org/middelbarescholen

[8] Ummels, J. (2013, April 5). *Slagingspercentages Brabantse scholen: Meeste leerlingen slagen bij Prinsentuin College in Andel.* Omroep Brabant. Retrieved June 25, 2015, from http://www.omroepbrabant.nl/?news/1913971213/Slagingspercentages Brabantse scholen meeste leerlingen slagen bij Prinsentuin College in Andel.aspx

Appendices

Appendix A

Data description

Column D/I indicates whether the variable is dependent (D) or independent (I). Column source contains the source, *DUO* is *Dienst Uitvoerend Onderwijs*, *OI* is *Onderwijsinspectie* and *LB* is *Leefbaarheidsonderzoek*.

#	Name	D/I	Source	Description
1	PROVINCIE	I	DUO	Province
2	BRINNUMMER	I	DUO	EI ID
3	VESTIGINGSNUMMER	I	DUO	School ID
4	ONDERWIJSGEBIEDNAAM	I	DUO	Education Area Code
5	Leerlingen2013	I	DUO	Total Students School in 2013
6	LWOO2013	I	DUO	Total Students with LWOO School in 2013
7	LWOO20134eklas	I	DUO	Total Students with LWOO School in 4th year 2013
8	VMBOBLDeelnemers	I	DUO	VMBO BBL Exam Participants
9	VMBOBLGeslaagden	D	DUO	VMBO BBL Exam Graduated
10	VMBOBLSE	I	DUO	VMBO BBL Average SE
11	VMBOBLCE	D	DUO	VMBO BBL Average CE
12	VMBOGLDeelnemers	I	DUO	VMBO GL Exam Participants
13	VMBOGLGeslaagden	D	DUO	VMBO GL Exam Graduated
14	VMBOGLSE	I	DUO	VMBO GL Average SE
15	VMBOGLCE	D	DUO	VMBO GL Average CE
16	VMBOKLDeelnemers	I	DUO	VMBO KBL Exam Participants
17	VMBOKLGeslaagden	D	DUO	VMBO KBL Exam Graduated
18	VMBOKLSE	I	DUO	VMBO KBL Average SE
19	VMBOKLCE	D	DUO	VMBO KBL Average CE
20	VMBOTLDeelnemers	I	DUO	VMBO TL Exam Participants
21	VMBOTLGeslaagden	D	DUO	VMBO TL Exam Graduated
22	VMBOTLSE	I	DUO	VMBO TL Average SE
23	VMBOTLCE	D	DUO	VMBO TL Average CE
24	HAVODeelnemers	I	DUO	HAVO Exam Participants
25	HAVOGeslaagd	D	DUO	HAVO Exam Graduated
26	HAVOSE	I	DUO	HAVO Average SE
27	HAVOCE	D	DUO	HAVO Average CE
28	VWODeelnemers	I	DUO	VWO Exam Participants
29	VWOGeslaagd	D	DUO	VWO Exam Graduated
30	VWOSE	I	DUO	VWO Average SE
31	VWOCE	D	DUO	VWO Average CE
32	SlagingspercentageVMBOBL	D	DUO	Graduation Rate VMBO BBL
33	SlagingspercentageVMBOGL	D	DUO	Graduation Rate VMBO GL
34	SlagingspercentageVMBOKL	D	DUO	Graduation Rate VMBO KBL
35	SlagingspercentageVMBOTL	D	DUO	Graduation Rate VMBO TL
36	SlagingspercentageHAVO	D	DUO	Graduation Rate HAVO
37	SlagingspercentageVWO	D	DUO	Graduation Rate VWO
38	FTE2013Instelling	I	DUO	Amount of FTE EI
39	FTE2013School	I	DUO	(Amount of Students School) / (Amount of Students EI) * (Amount of FTE EI)
40	Leeftijd2013	I	DUO	Average age employees EI
41	LIQUIDITEITCURRENTRATIO	I	DUO	Financial Characteristics EI
42	RENTABILITEIT	I	DUO	See DUO
43	SOLVABILITEIT1	I	DUO	
44	SOLVABILITEIT2	I	DUO	
45	ALGEMENERESERVETOTALEBATEN	I	DUO	
46	BELEGGINGENTOVEV	I	DUO	
47	CONTRACTACTIVITEITENRIJKSBIJDRAGE	I	DUO	
48	CONTRACTACTIVITEITENTOTALEBATEN	I	DUO	
49	EIGENVERMOGENTOTALEBATEN	I	DUO	
50	INVESTERINGHUISVESTINGTOTALEBATEN	I	DUO	
51	INVESTERINGENINVENTAPPTOTALEBATEN	I	DUO	
52	KAPITALISATIEFACTOR	I	DUO	
53	LIQUIDITEITQUICKRATIO	I	DUO	
54	OVOVERHEIDSBIJDRAGENTOTBATEN	I	DUO	
55	PERSONEELRIJKSBIJDRAGEN	I	DUO	
56	PERSONELELASTENTOTALELASTEN	I	DUO	
57	RIJKSBIJDRAGENTOTALEBATEN	I	DUO	
58	VOORZIENINGENTOTALEBATEN	I	DUO	
59	WEERSTANDSVERMOGENMVA	I	DUO	
60	WEERSTANDSVERMOGENVOTOTALEBN	I	DUO	
61	WERKKAPITAALTOTALEBATEN	I	DUO	
62	HUISVESTINGSLASTENTOTALELASTEN	I	DUO	
63	WERKKAPITAAL	I	DUO	

Continued on next page

Table A.1 – continued from previous page

#	Name	D/I	Source	Description
64	(KORTL. SCHULDEN / TOTALE BATEN) * 365	I	DUO	
65	(VORDERINGEN / TOTALE BATEN) * 365	I	DUO	
66	bev	I	LB	Lifeability score population composition. Calculated as Weighted Average per zip code.
67	lftsam	I	LB	Lifeability score social coherence. Calculated as Weighted Average per zip code.
68	publ	I	LB	Lifeability score public space. Calculated as Weighted Average per zip code.
69	veilig	I	LB	Lifeability score safety. Calculated as Weighted Average per zip code.
70	vrz	I	LB	Lifeability score population facilities. Calculated as Weighted Average per zip code.
71	won	I	LB	Lifeability score housing. Calculated as Weighted Average per zip code.
72	bl1e3e	I	OI	Percentage of VMBO BBL students which succeed from year 1 till 3 without delay.
73	kl1e3e	I	OI	Percentage of VMBO KBL students which succeed from year 1 till 3 without delay.
74	gt1e3e	I	OI	Percentage of VMBO GL/TL students which succeed from year 1 till 3 without delay.
75	ha1e3e	I	OI	Percentage of HAVO students which succeed from year 1 till 3 without delay.
76	vw1e3e	I	OI	Percentage of VWO students which succeed from year 1 till 3 without delay.
77	BaOnv	D	OI	Percentage of VMBO BBL students which graduate without delay.
78	KaOnv	D	OI	Percentage of VMBO KBL students which graduate without delay.
79	GtOnv	D	OI	Percentage of VMBO GL/TL students which graduate without delay.
80	HaOnv	D	OI	Percentage of HAVO students which graduate without delay.
81	VwOnv	D	OI	Percentage of VWO students which graduate without delay.
82	plwooba	I	OI	Percentage of VMBO BBL students with LWOO.
83	plwooka	I	OI	Percentage of VMBO KBL students with LWOO.
84	plwoogt	I	OI	Percentage of VMBO GL/TL students with LWOO.
85	papcgba	I	OI	Percentage of VMBO BBL students from poor areas.
86	papcgka	I	OI	Percentage of VMBO KBL Students from poor areas.
87	Papcggt	I	OI	Percentage of VMBO GL/TL students from poor areas.
88	papcgha	I	OI	Percentage of HAVO students from poor areas.
89	papcgvw	I	OI	Percentage of VWO students from poor areas.
90	pinstba	I	OI	Percentage inflow VMBO BBL.
91	pinstka	I	OI	Percentage inflow VMBO KBL
92	pinstgt	I	OI	Percentage inflow VMBO GL/TL
93	pinstha	I	OI	Percentage inflow HAVO
94	pinstvw	I	OI	Percentage inflow VWO
95	fteperleerling	I	DUO	Amount of FTE School / Amount of Students
96	VMBOBLDIFFSECE	I	DUO	VMBOBLSE - VMBOBLCE
97	VMBOGLDIFFSECE	I	DUO	VMBOGLSE - VMBOGLCE
98	VMBOKLDIFFSECE	I	DUO	VMBOKLSE - VMBOKLCE
99	VMBOTLDIFFSECE	I	DUO	VMBOTLSE - VMBOTLCE
100	HAVODIFFSECE	I	DUO	HAVOSE - HAVOCE
101	VWODIFFSECE	I	DUO	VWOSE - VWOCE
102	onderbouwpercover	I	OI	Percentage of students in year 1 and 2 which succeed till next year.
103	onderbouwpercblijvenzitten	I	OI	1 - onderbouwpercover
104	onderbouwpercinstroom	I	OI	Percentage inflow in year 1 and 2
105	onderbouwpercuitstroom	I	OI	Percentage outflow in year 1 and 2
106	BBLpercover	I	OI	Percentage of students in year 3 and 4 VMBO BBL which succeed till next year.
107	BBLpercblijvenzitten	I	OI	1 - BBLpercover
108	BBLpercopstroom	I	OI	Percentage of students in year 3 and 4 VMBO BBL which succeed to higher level.
109	BBLpercinstroom	I	OI	Percentage inflow in year 3 and 4 VMBO BBL
110	BBLpercuitstroom	I	OI	Percentage outflow in year 3 and 4 VMBO BBL
111	GLpercover	I	OI	Percentage of students in year 3 and 4 VMBO GL which succeed till next year.
112	GLpercblijvenzitten	I	OI	1 - GLpercover
113	GLpercopstroom	I	OI	Percentage of students in year 3 and 4 VMBO GL which succeed to higher level.
114	GLpercafstroom	I	OI	Percentage of students in year 3 and 4 VMBO GL which succeed to lower level.
115	GLpercinstroom	I	OI	Percentage inflow in year 3 and 4 VMBO GL
116	GLpercuitstroom	I	OI	Percentage outflow in year 3 and 4 VMBO GL
117	KBLpercover	I	OI	Percentage of students in year 3 and 4 VMBO KBL which succeed till next year.
118	KBLpercblijvenzitten	I	OI	1 - KBLpercover
119	KBLpercopstroom	I	OI	Percentage of students in year 3 and 4 VMBO KBL which succeed to higher level.
120	KBLpercafstroom	I	OI	Percentage of students in year 3 and 4 VMBO KBL which succeed to lower level.
121	KBLpercinstroom	I	OI	Percentage inflow in year 3 and 4 VMBO KBL
122	KBLpercuitstroom	I	OI	Percentage outflow in year 3 and 4 VMBO KBL
123	TLpercover	I	OI	Percentage of students in year 3 and 4 VMBO TL which succeed till next year.
124	TLpercblijvenzitten	I	OI	1 - TLpercover
125	TLpercopstroom	I	OI	Percentage of students in year 3 and 4 VMBO TL which succeed to higher level.
126	TLpercafstroom	I	OI	Percentage of students in year 3 and 4 VMBO TL which succeed to lower level.
127	TLpercinstroom	I	OI	Percentage inflow in year 3 and 4 VMBO TL
128	TLpercuitstroom	I	OI	Percentage outflow in year 3 and 4 VMBO TL
129	HAVOpercover	I	OI	Percentage of students in year 3, 4 and 5 HAVO which succeed till next year.
130	HAVOpercblijvenzitten	I	OI	1 - HAVOpercover
131	HAVOpercopstroom	I	OI	Percentage of students in year 3, 4 and 5 HAVO which succeed to higher level.
132	HAVOpercafstroom	I	OI	Percentage of students in year 3, 4 and 5 HAVO which succeed to lower level.
133	HAVOpercinstroom	I	OI	Percentage inflow in year 3, 4 and 5 HAVO
134	HAVOpercuitstroom	I	OI	Percentage outflow in year 3, 4 and 5 HAVO

Continued on next page

Table A.1 – continued from previous page

#	Name	D/I	Source	Description
135	VWOpercover	I	OI	Percentage of students in year 3, 4, 5 and 6 VWO which succeed till next year.
136	VWOpercblijvenzitten	I	OI	1 - VWOpercover
137	VWOpercafstroom	I	OI	Percentage of students in year 3, 4, 5 and 6 VWO which succeed to lower level.
138	VWOpercinstroom	I	OI	Percentage inflow in year 3, 4, 5 and 6 VWO
139	VWOpercuitstroom	I	OI	Percentage outflow in year 3, 4, 5 and 6 VWO

Appendix B

Models

B.1 VMBOBBL

Model 1	
(Intercept)	-12.7338246
VMBOBLSE	2.808576122
onderbouwpercinstroom	-4.524633691
BBLpercover	-1.982099274
papcgba	-0.433676964

Model 2	
(Intercept)	-11.89073528
VMBOBLSE	2.795944234
onderbouwpercinstroom	-7.534735654
BBLpercover	-1.840141865
onderbouwpercover	-0.916989526
papcgba	-0.442172532

Model 3	
(Intercept)	-12.80155989
VMBOBLSE	2.827703043
onderbouwpercinstroom	-4.566326549
BBLpercover	-1.994544644
papcgba	-0.436741842
VMBOBLDeelnemers	-0.000687993

Model 4	
(Intercept)	-12.82522028
VMBOBLSE	2.810315256
onderbouwpercinstroom	-5.253978309
BBLpercover	-1.923255049
papcgba	-0.44630828
onderbouwpercuitstroom	0.430263475

Model 5	
(Intercept)	-12.72572096
VMBOBLSE	2.809155846
onderbouwpercinstroom	-4.574570051
BBLpercover	-1.96975
bev	-0.001308038
papcgba	-0.502196136

Model 6	
(Intercept)	-12.73619493
VMBOBLSE	2.808934041
onderbouwpercinstroom	-4.525781814
BBLpercover	-1.981323775
papcgba	-0.435647914
veilig	-3.336E-05

Model 7	
(Intercept)	-13.61836777
VMBOBLSE	2.827421007
BBLpercover	-2.118688615
onderbouwpercover	0.909600985
papcgba	-0.448774139

Model 8	
(Intercept)	-13.1813677
VMBOBLSE	2.851866673
onderbouwpercinstroom	-4.6186196
BBLpercover	-1.943895906
bev	0.006025781

Model 9	
(Intercept)	-13.00606046
VMBOBLSE	2.826303496
onderbouwpercinstroom	-4.830123273
BBLpercover	-1.946606267
veilig	0.004533065

Model 10	
(Intercept)	-11.95961273
VMBOBLSE	2.814916437
onderbouwpercinstroom	-7.569414608
BBLpercover	-1.852440633
onderbouwpercover	-0.915394136
papcgba	-0.445381863
VMBOBLDeelnemers	-0.000680081

B.2 VMBOKBL

Model 1	
(Intercept)	-8.78512639
VMBOKLSE	2.058064414
won	0.01331683
veilig	-0.009673556
vrz	-0.006917373
VMBOKLDeelnemers	-0.002303597
kl1e3e	-0.124863391
onderbouwpercuitstroom	-1.793162022
KBLpercover	-1.471727976
KBLpercuitstroom	-3.405895929

Model 2	
(Intercept)	-8.755546641
VMBOKLSE	2.053956936
won	0.013386873
veilig	-0.01046427
vrz	-0.006715936
VMBOKLDeelnemers	-0.002337537
kl1e3e	-0.124523207
onderbouwpercuitstroom	-1.791076679
KBLpercover	-1.477014974
KBLpercuitstroom	-3.402678383
bev	0.001404862

Model 3	
(Intercept)	-8.815847883
VMBOKLSE	2.058598368
won	0.013331205
veilig	-0.009698299
vrz	-0.006943263
VMBOKLDeelnemers	-0.00227497
kl1e3e	-0.108547227
onderbouwpercuitstroom	-1.678998247
KBLpercover	-1.465449145
KBLpercuitstroom	-3.426196121
onderbouwpercinstroom	-0.756262322

Model 4	
(Intercept)	-8.560363236
VMBOKLSE	2.027106338
won	0.015415767
veilig	-0.008471395
VMBOKLDeelnemers	-0.002327448
kl1e3e	-0.127726911
onderbouwpercuitstroom	-1.781079527
KBLpercover	-1.471428645
KBLpercuitstroom	-3.343990921

Model 5	
(Intercept)	-8.786354663
VMBOKLSE	2.054674734
won	0.013398363
veilig	-0.010457162
vrz	-0.006748228
VMBOKLDeelnemers	-0.002308833
kl1e3e	-0.108998045
onderbouwpercuitstroom	-1.681421226
KBLpercover	-1.470786427
KBLpercuitstroom	-3.422357185
bev	0.001349397
onderbouwpercinstroom	-0.724680383

Model 6	
(Intercept)	-8.518157008
VMBOKLSE	2.021508842
won	0.015449413
veilig	-0.010081442
VMBOKLDeelnemers	-0.002392966
kl1e3e	-0.127136284
onderbouwpercuitstroom	-1.777971033
KBLpercover	-1.482887431
KBLpercuitstroom	-3.342544214
bev	0.002723118

Model 7	
(Intercept)	-9.064749179
VMBOKLSE	2.035465296
won	0.013382048
veilig	-0.009235838
vrz	-0.006738547
VMBOKLDeelnemers	-0.002046339
kl1e3e	-0.062485527
onderbouwpercuitstroom	-2.10194119
KBLpercover	-1.181391255

Model 8	
(Intercept)	-8.583467134
VMBOKLSE	2.027236141
won	0.01543249
veilig	-0.008485488
VMBOKLDeelnemers	-0.002303625
kl1e3e	-0.114344321
onderbouwpercuitstroom	-1.682044287
KBLpercover	-1.466559932
KBLpercuitstroom	-3.362892061
onderbouwpercinstroom	-0.652428439

Model 9	
(Intercept)	-9.033521187
VMBOKLSE	2.031263217
won	0.013459048
veilig	-0.010057626
vrz	-0.006531711
VMBOKLDeelnemers	-0.002082942
kl1e3e	-0.063138699
onderbouwpercuitstroom	-2.098975019
KBLpercover	-1.186635634
bev	0.001457314

Model 10	
(Intercept)	-8.540449238
VMBOKLSE	2.021793546
won	0.01546501
veilig	-0.010072868
VMBOKLDeelnemers	-0.002370246
kl1e3e	-0.114979252
onderbouwpercuitstroom	-1.686542607
KBLpercover	-1.478226314
KBLpercuitstroom	-3.360024571
bev	0.002685097
onderbouwpercinstroom	-0.599139421

B.3 VMBOGL

Model 1	
(Intercept)	-11.36376381
GLpercuitstroom	-8.73779459
VMBOGLSE	2.276684625
GLpercover	-1.280411523
GLpercopstroom	0.669976451
onderbouwpercuitstroom	0.728363123
won	0.015748651

Model 2	
(Intercept)	-11.32369252
GLpercuitstroom	-8.475285927
VMBOGLSE	2.251932185
onderbouwpercover	0.326533465
GLpercover	-1.447831471
GLpercopstroom	0.658720843
won	0.014701579

Model 3	
(Intercept)	-11.09093349
GLpercuitstroom	-8.485004771
VMBOGLSE	2.259784991
GLpercover	-1.413269236
GLpercopstroom	0.662313212
onderbouwpercinstroom	-0.325305347
won	0.014955839

Model 4	
(Intercept)	-11.06752574
GLpercuitstroom	-8.541921992
Papcggt	-0.354608016
VMBOGLSE	2.252625466
GLpercover	-1.377846651
GLpercopstroom	0.682233797
onderbouwpercuitstroom	0.785872095
won	0.012406509

Model 5	
(Intercept)	-12.28987297
GLpercuitstroom	-8.921046745
VMBOGLSE	2.258323907
onderbouwpercover	1.045869409
GLpercover	-1.275235409
GLpercopstroom	0.65970494
onderbouwpercuitstroom	1.367783796
won	0.015420592

Model 6	
(Intercept)	-12.32165395
GLpercuitstroom	-7.705400705
VMBOGLSE	2.229914642
GLpercopstroom	0.707035136
onderbouwpercuitstroom	1.092218226
won	0.016436424
GLpercafstroom	1.266370284

Model 7	
(Intercept)	-10.9138886
GLpercuitstroom	-8.271778861
Papcggt	-0.335531989
VMBOGLSE	2.232900422
onderbouwpercover	0.160225595
GLpercover	-1.53040973
GLpercopstroom	0.672538563
won	0.011657056

Model 8	
(Intercept)	-12.55542435
GLpercuitstroom	-7.665359458
VMBOGLSE	2.2719909
GLpercopstroom	0.647830907
onderbouwpercuitstroom	1.263865296
won	0.016738793

Model 9	
(Intercept)	-10.80906769
GLpercuitstroom	-8.270587353
Papcggt	-0.3479122
VMBOGLSE	2.238675263
GLpercover	-1.517179109
GLpercopstroom	0.677154059
onderbouwpercinstroom	0.300517013
won	0.01173555

Model 10	
(Intercept)	-11.36999167
GLpercuitstroom	-8.835257924
VMBOGLSE	2.269448074
GLpercover	-1.228367522
GLpercopstroom	0.660581107
onderbouwpercuitstroom	1.037703791
onderbouwpercinstroom	-2.339229759
won	0.015679106

B.4 VMBOTL

Model 1	
(Intercept)	-11.74552486
VMBOTLSE	2.177045779
vrz	0.007254305
won	0.015094832
gt1e3e	-0.434109242
onderbouwpercinstroom	-4.109379307
TLpercover	0.582989001
TLpercafstroom	1.815154489

Model 2	
(Intercept)	-12.25830492
VMBOTLSE	2.170359872
vrz	0.007292047
won	0.014933222
gt1e3e	-0.607344312
onderbouwpercover	0.75987287
TLpercover	0.561895457
TLpercafstroom	1.809321174

Model 3	
(Intercept)	-11.76954355
VMBOTLSE	2.183026209
vrz	0.007332509
won	0.015072595
gt1e3e	-0.432806461
onderbouwpercinstroom	-3.855985107
TLpercover	0.569710919
TLpercopstroom	-1.740210192
TLpercafstroom	1.800954103

Model 4	
(Intercept)	-11.64012524
VMBOTLSE	2.165756765
vrz	0.007332656
won	0.014930239
gt1e3e	-0.435128351
pinstgt	-0.405194516
onderbouwpercinstroom	-3.785584232
TLpercover	0.569196162
TLpercafstroom	1.843840527

Model 5	
(Intercept)	-12.25313968
VMBOTLSE	2.176944356
vrz	0.007371604
won	0.014920349
gt1e3e	-0.59537981
onderbouwpercover	0.714962611
TLpercover	0.549316522
TLpercopstroom	-1.784474947
TLpercafstroom	1.795137704

Model 6	
(Intercept)	-12.12222891
VMBOTLSE	2.158579749
vrz	0.007380186
won	0.014763541
gt1e3e	-0.600258464
pinstgt	-0.433782001
onderbouwpercover	0.726347478
TLpercover	0.547060116
TLpercafstroom	1.83850485

Model 7	
(Intercept)	-11.62272307
VMBOTLSE	2.160000424
bev	0.00164367
vrz	0.007513785
won	0.014172809
gt1e3e	-0.442246366
onderbouwpercinstroom	-3.838777844
TLpercover	0.561052054
TLpercafstroom	1.859098453

Model 8	
(Intercept)	-11.81950135
VMBOTLDeelnemers	-0.000343559
VMBOTLSE	2.188027131
vrz	0.007307998
won	0.015019748
gt1e3e	-0.435069123
onderbouwpercinstroom	-4.345285259
TLpercover	0.624940401
TLpercafstroom	1.795747569

Model 9	
(Intercept)	-11.76114734
VMBOTLSE	2.17469029
publ	-0.001073167
vrz	0.007003278
won	0.014444251
gt1e3e	-0.390950804
onderbouwpercinstroom	-4.216734149
TLpercover	0.580942177
TLpercafstroom	1.814442652

Model 10	
(Intercept)	-12.08990806
VMBOTLSE	2.154339532
bev	0.00160788
vrz	0.007538673
won	0.014049046
gt1e3e	-0.600085398
onderbouwpercover	0.688705666
TLpercover	0.542778244
TLpercafstroom	1.853043211

B.5 HAVO

Model 1	
(Intercept)	-18.15465862
HAVOSE	2.941875195
papcgha	-0.547732481
won	-0.002459788
ha1e3e	-0.805670363
onderbouwpercover	2.251890256
HAVOpercuitstroom	2.040595475
onderbouwpercinstroom	6.1808936
HAVOpercafstroom	0.984125858
HAVOpercinstroom	-17.06212796

Model 2	
(Intercept)	-17.74609756
HAVOSE	2.963533709
papcgha	-0.556183173
won	-0.002247966
ha1e3e	-0.799409239
onderbouwpercover	1.682624678
HAVOpercuitstroom	2.11779931
HAVOpercafstroom	0.983833103
HAVOpercinstroom	-14.96286017

Model 3	
(Intercept)	-17.39728199
HAVOSE	2.933151537
papcgha	-0.418564435
ha1e3e	-0.911589979
onderbouwpercover	1.599095376
HAVOpercuitstroom	1.943055276
HAVOpercafstroom	1.014332189

Model 4	
(Intercept)	-17.63676696
HAVOSE	2.954854424
papcgha	-0.543782653
won	-0.002175142
ha1e3e	-0.849972931
onderbouwpercover	1.665050776
HAVOpercuitstroom	2.042679588
HAVOpercafstroom	1.025841809

Model 5	
(Intercept)	-17.84649186
HAVOSE	2.919741916
papcgha	-0.407788815
ha1e3e	-0.875895921
onderbouwpercover	2.123379081
HAVOpercuitstroom	1.933054232
onderbouwpercinstroom	5.647406936
HAVOpercafstroom	0.972821236
HAVOpercinstroom	-16.33446719

Model 6	
(Intercept)	-17.49746998
HAVOSE	2.941095886
papcgha	-0.426492424
ha1e3e	-0.86462958
onderbouwpercover	1.614596356
HAVOpercuitstroom	2.011051996
HAVOpercafstroom	0.973552645
HAVOpercinstroom	-14.46759062

Model 7	
(Intercept)	-17.97543141
HAVOSE	2.935089118
papcgha	-0.535068492
won	-0.002349388
ha1e3e	-0.861830848
onderbouwpercover	2.153931053
HAVOpercuitstroom	1.967211996
onderbouwpercinstroom	5.370702001
HAVOpercafstroom	1.031157162

Model 8	
(Intercept)	-17.68573267
HAVOSE	2.913843341
papcgha	-0.401643526
ha1e3e	-0.926802991
onderbouwpercover	2.034993027
HAVOpercuitstroom	1.868659742
onderbouwpercinstroom	4.89055246
HAVOpercafstroom	1.018197897

Model 9	
(Intercept)	-18.09065303
HAVOSE	2.937968501
veilig	0.001585376
papcgha	-0.515634264
won	-0.003896862
ha1e3e	-0.837021827
onderbouwpercover	2.225868174
HAVOpercuitstroom	2.075482069
onderbouwpercinstroom	6.294217905
HAVOpercafstroom	0.978893003
HAVOpercinstroom	-17.48839502

Model 10	
(Intercept)	-17.67796253
HAVOSE	2.96012814
veilig	0.001483539
papcgha	-0.526296938
won	-0.003588788
ha1e3e	-0.828661849
onderbouwpercover	1.648075263
HAVOpercuitstroom	2.151505477
HAVOpercafstroom	0.978828261
HAVOpercinstroom	-15.33634021

B.6 VWO

Model 1	
(Intercept)	-14.64257512
VWOSE	2.675855155
bev	0.010101476
onderbouwpercover	0.449904703
VWOpercover	-1.445366701
VWOpercafstroom	2.364367756
VWOpercuitstroom	-4.916582737

Model 2	
(Intercept)	-14.61588565
VWOSE	2.673465476
bev	0.008346419
onderbouwpercover	0.442749277
VWOpercover	-1.428046939
VWOpercafstroom	2.28832907
won	0.001967651
VWOpercuitstroom	-4.8840565

Model 3	
(Intercept)	-14.88669573
VWOSE	2.706915939
bev	0.009857636
onderbouwpercover	0.493991152
VWOpercover	-1.423614353
VWOpercafstroom	2.389333933
veilig	-0.003504763
won	0.004954837
VWOpercuitstroom	-5.178207386

Model 4	
(Intercept)	-14.64420578
VWOSE	2.674371481
bev	0.009621413
onderbouwpercover	0.466792698
VWOpercover	-1.446228168
VWOpercafstroom	2.285730464
VWOpercuitstroom	-4.827559973
vrz	-0.001229483

Model 5	
(Intercept)	-14.61865217
VWOSE	2.674978683
bev	0.010129723
onderbouwpercover	0.43983105
VWOpercover	-1.454096835
VWOpercafstroom	2.354521981
VWOpercinstroom	-2.978966059
VWOpercuitstroom	-4.841167621

Model 6	
(Intercept)	-14.67745684
VWOSE	2.679997786
bev	0.010570989
onderbouwpercover	0.456943505
VWOpercover	-1.44782323
VWOpercafstroom	2.388700366
veilig	-0.000394538
VWOpercuitstroom	-4.95527445

Model 7	
(Intercept)	-14.87612953
VWOSE	2.569690684
bev	0.009977295
onderbouwpercover	0.034623779
VWOpercafstroom	3.419407243
VWOpercuitstroom	-3.974029014

Model 8	
(Intercept)	-14.57635136
VWOSE	2.671999487
bev	0.008228814
onderbouwpercover	0.426064211
VWOpercover	-1.440310209
VWOpercafstroom	2.265693684
won	0.002150573
VWOpercinstroom	-4.644568388
VWOpercuitstroom	-4.765039964

Model 9	
(Intercept)	-14.61765199
VWOSE	2.673371563
bev	0.008367649
onderbouwpercover	0.445735167
VWOpercover	-1.429079055
VWOpercafstroom	2.280393557
won	0.001860811
VWOpercuitstroom	-4.872035745
vrz	-0.000189192

Model 10	
(Intercept)	-14.7544295
VWOSE	2.6864952
bev	0.010806004
onderbouwpercover	0.49914864
VWOpercover	-1.454900174
VWOpercafstroom	2.315308854
veilig	-0.001230416
VWOpercuitstroom	-4.898382956
vrz	-0.0019471