CUNY SCHOOL OF PROFESSIONAL STUDIES MASTER'S IN DATA SCIENCE

Data 621 – Final Project

Prime awards model for the New York Department of Education grants from the federal government

Albina Gallyavova, Chad Smith, Duubar Villalobos Jimenez , Nkasi Nedd, Raghunathan Ramnath

12/20/2018

Table of Contents

Prime awards model for the New York Department of Education grants from the federal government

Abstract

Education investment is crucial to the well-being of nation's future. It has already proven that investing in education help reduce unemployment; fight against crime and over all strengthens communities.

On the other hand, as we also know very well, how schools are currently struggling to keep up with the high demand of quality education, inclusiveness and no child left behind policies, thus, add extra expenses and require extra resources to satisfy the need, securing funds for investing in education is highly desirable.

Child care expenses, characteristics of teachers and student loans are some of the spotlight that affects the condition of education. There is an assumption that the funds are not equally spread or spent towards certain geographies.

In this Project, we were trying to get an insight on the grants provided historically and based on that predict for the next two years.

Key words

New York, education, federal and non-federal grants, funding allocation.

Introduction

As we all know, we can all come to an agreement that education investment is crucial to the well-being of our future. It has already proven that investing in education help reduce unemployment; fight against crime and over all strengthens communities. On the other hand, as we also know very well, are how schools are currently struggling to keep up with the high demand of quality education, inclusiveness and no child left behind policies, thus, add extra expenses and require extra resources to satisfy the need, securing funds for investing in education is highly desirable. An example can be seeing when teachers -in many cases- are obligated to put part of their salary in order to purchase classroom supplies, another example will be to purchase meals in order to keep the young minds alert and ready to learn.

From the above, we can deduct that lack of funds and high education expectations seems to be in opposite sides and currently not necessarily working together. High demand of resources and great diversity experienced -especially here in the great state of New York- claim for a fair need of funds in order to satisfy this quest.

The proposal review will analyze and discuss the economical, geographical and spatial aspects for government education funding for the state of New York.

Literature review

Education impact on economic growth

The importance of education for the society overall cannot be underestimated. Education allows us to grow and develop as individuals, provides skills to become eligible for the workforce, trains self-discipline to be able to achieve the life goals, and much more. More educated people create a platform for a higher quality labor, living standards and solving problems more efficiently. In general, education—as a critical component of a country's human capital—increases the efficiency of each individual worker and helps economies to move up the value chain beyond manual tasks or simple production processes (WEF 2016). Education has been long viewed as a backbone of economic well-being. Throughout his career, Robert Barro produced numerous papers suggesting positive relations between years of schooling and economic growth (Barro, 1991, 1997). 2013 meta-regression analysis conducted by Benos and Zotou (Benos, Zotou, 2013) provided evidence that different measures of education give rise to different coefficients of the size effect of education on growth [BRIDGE].

Hanushek and Woessmann (2010) provide an extensive literature review on the topic, however, they argue the empirical evidence on the impact of education on economic growth has long been mixed mostly due to measurement problems [BRIDGE].

Although most studies use measurements such as years of schooling or overall cognitive skills based on test results, a question arises whether government spending has an impact on quality of education, i.e., students' performance, and consequently on economic growth.

Funding impact on education

The early years of research were not supportive of the view that government spending could have a positive impact on the quality of education. The famous Coleman report (Coleman et al., 1966) suggested that families are the most important determinant of student achievement. Despite the report, the government has implemented a number of educational reforms affecting spending structure.

The resources have not necessarily yielded positive results in general improvement of student achievement (Hanushek, 2003).

More recent studies, however, provide an opposite view. Using data available from The World Bank's World Development Indicators (WDI) database, it is possible to estimate the bivariate relationship between government education expenditure and GDP across a large sample of countries. The estimates show that for every dollar the government spends on education, GDP grows on average by $20. (https://theconversation.com/does-government-spending-on-education-promote-economic-growth-60229).

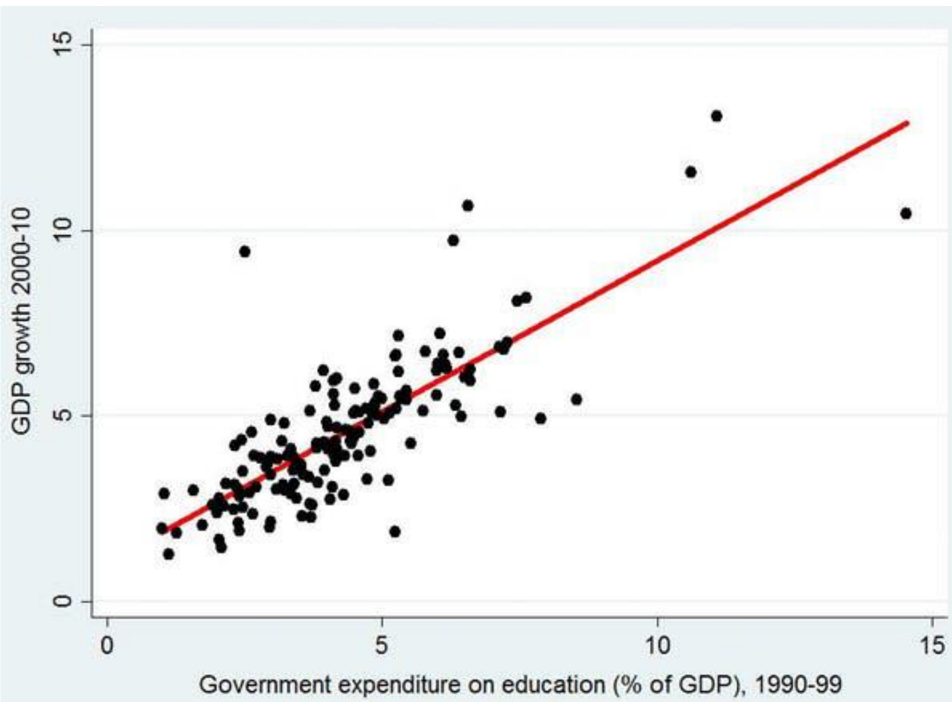

Figure 1: Government expenditure on education (% of GDP), 1990-99.

According to Jackson, Johnson and Persico (2015), event-study and instrumental variable models reveal that a 10 percent increase in per-pupil spending each year for all twelve years of public school leads to 0.27 more completed years of education, 7.25 percent higher wages, and a 3.67 percentage-point reduction in the annual incidence of adult poverty; effects are much more pronounced for children from low-income families. Exogenous spending increases were associated with sizable improvements in measured school quality, including reductions in student-to-teacher ratios, increases in teacher salaries, and longer school years.

In most recent studies of spending impact on student performance, further evidence was provided that following court-ordered finance reforms taking place between 1989 and 2000, the highest poverty quartile in a treated state experienced a 11.5 to 12.1 percent increase in per-pupil spending and a 6.8 to 11.5 percentage point increase in graduation rates (Candelaria and Shores, 2017).

Funding allocation

Education is one of the largest single components of government spending, amassing 6.1% of GDP across federal, state, and local expenditures (OECD, 2015). According to the US Department of Education, funds come predominantly from State, local and private sources, and this is especially true at the elementary and secondary level, where about 92 percent of the funds will come from non-Federal sources. However, due to those resources mostly coming from taxes, many argue that the system creates inequality in funding allocation often causing underprivileged layers of population also be educationally underfunded. In the Funding Gaps 2018 report, The Education Trust highlights that school districts that serve large populations of students of color and students from low-income families receive far less funding than those serving white and more affluent students.

Methodology

Initially, we looked at the data and tried to determine if multiple linear regression models were appropriate to model the grants data-set. First, we tried getting insights from the data-set; some of them are as follows:

- The data-set is currently composed of 18160 records and 73 variables.
- Many variables had high percentages of missing values.
- The data-set did not include a dictionary and additional research had to be done in order to find meaningful insights from the given variables.
- Many records were duplicated and a “cleanup” process had to be implemented.

Based on the number of variables and data, first we tried to build a multiple linear regression model from the given variables in order to predict the grant values. Since this data-set is not complete, we had to abandon the idea due to lack of data and improper results. Based on the above, we decided that in order to obtain the maximum information possible, we had to discard the use of many variables and put our focus into the following variables:

- **action_date**: Original request date in which a grant application was initiated. The original date format was in the standard “mm/dd/yyyy” format; from there, we focused and extracted the year only in the following format “yyyy”.
- **award_id:** Individual grant identifier. It was quickly identified that many records presented modifications and the “duplicate” records were discarded.

- **modification_number**: Incremental value that indicates the number of modifications a grant has had over time. It is assumed that the higher the number the most recent the modification was. Based on that, only the records that have the unique **award_id** with the highest **modification_number** were kept for the analysis.
- **recipient_zip_code**: Variable that denotes the zip code of the school requesting the grant.
- **recipient_county_code**: Variable that denotes the County code of the school requesting the grant.
- **recipient_congressional_district**: Variable that denotes the Congressional District code of the school requesting the grant.
- **total_funding_amount**: Considered to be the variable that holds the grant amount. Later on, due to the magnitude of the values given, we had to divide by one million (1 000 000) since much of the data was reported in the millions.

It is noted that the variable **total_funding_amount** presented some negative values which were surprising to us; the reason why, is because we did not expect to find negative grants.

Experimentation and Results

Once we thought-out of the negative values, we decided to treat these negative grants as a separate entity called "Adjustments" or "Subtractions" in this way, we separate the data in two ways as follows:

1) **Positive Grants**: Holding the positive grants for which the schools applied.

2) **Negative Grants**: Holding the negative values listed in the data-set, we treated these to be as "Adjustments" from the yearly grants.

3) **Net Grants**: That is, we subtracted from the positive grants, the negative grants.

From the above, we decided to group our remaining data in groups, that is:

1) **Group by Year and Zip Code**: We summarize the values by adding the **total_funding_amount**.

2) **Group by Year and County Code**: We summarize the values by adding the **total_funding_amount**.

3) **Group by Year and Congressional District Code**: We summarize the values by adding the **total_funding_amount**.

From these 3 groups, we ended up with nine different matrices from which we were able to predict values for the years 2019* and 2020*; that is:

Matrix Group 1:

- **Group by Year and Zip Code <- Add Grants**
- **Group by Year and Zip Code <- Subtract Grants**
- **Group by Year and Zip Code <- Net Grants**

Matrix Group 2:

- **Group by Year and County Code <- Add Grants**
- **Group by Year and County Code <- Subtract Grants**
- **Group by Year and County Code <- Net Grants**

Matrix Group 3:

- **Group by Year and Congressional District Code <- Add Grants**
- **Group by Year and Congressional District Code <- Subtract Grants**
- **Group by Year and Congressional District Code <- Net Grants**

The process to predict the values from the above matrices, were very simple but tedious. The way we implemented, was by predicting unique linear models for each Zip Code, County Code and Congressional District Code available in the data-set. We did this in a one by one basis; that is, every single combination returned a different set of coefficients for every given variable. Once we were able to obtain the unique linear model coefficients, we were able to find the predicted values for the years 2019* and 2020*.

The models employed were very simple linear regression models as a combination. That is, in some cases we used a standard linear regression model when the generalized linear model could not be applied to the given values returned in the data-set due to missing values or linearity in the data, which returns errors.

Example of simple linear models in **R** applied to our data set.

lm(y ~ x) & gls(y ~ x)

y: represented the values returned in the summarized values for each Grouped matrix for each individual column identifier.

$y = \{y_{1,j}, y_{2,j}, \ldots y_{13,j}\}$ for each j column present on each matrix group listed above.

x: represented the number of years given in the data-set. That is,

$x= \{0, 1, 2, \ldots , 12\}$ representing the years from 2006 through 2018 in a consecutive form.

From the above, we ended up calculating as follows:

- 449 different linear models were employed for the Zip code predictions.
- 62 different linear models were employed for the County predictions.
- 29 different linear models were employed for the Congressional District predictions.

All the models transformed our Grant values divided into one million into a log scale for easier linearity transformations.

With the combination of those predictions, we were able to create a predictive visualization of the results as seeing in the appendix tables and maps as seeing below.

Map visualization of predictions for 2019 and 2020.

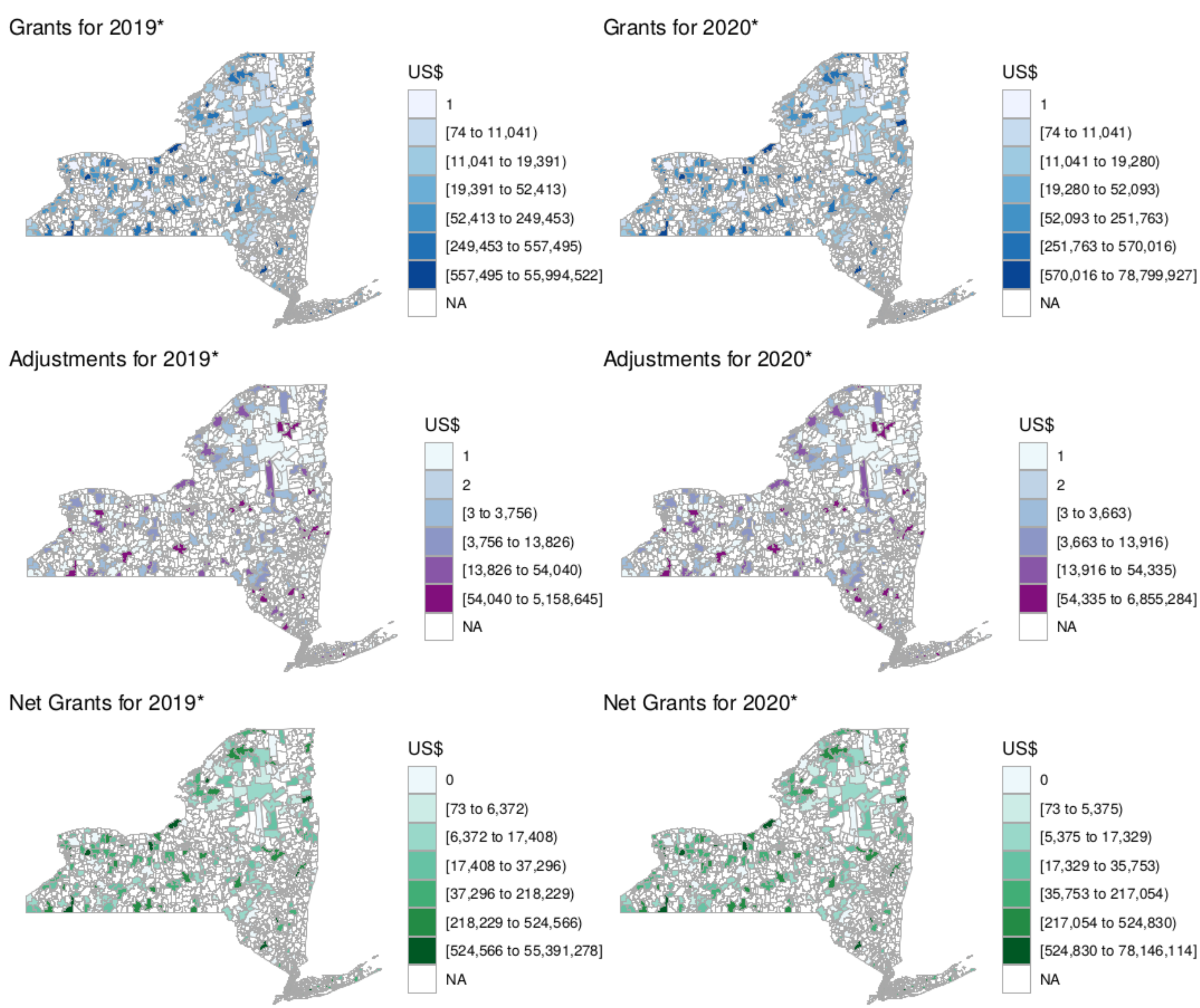

Figure 2: Zip code Grants and predicted values in a map.

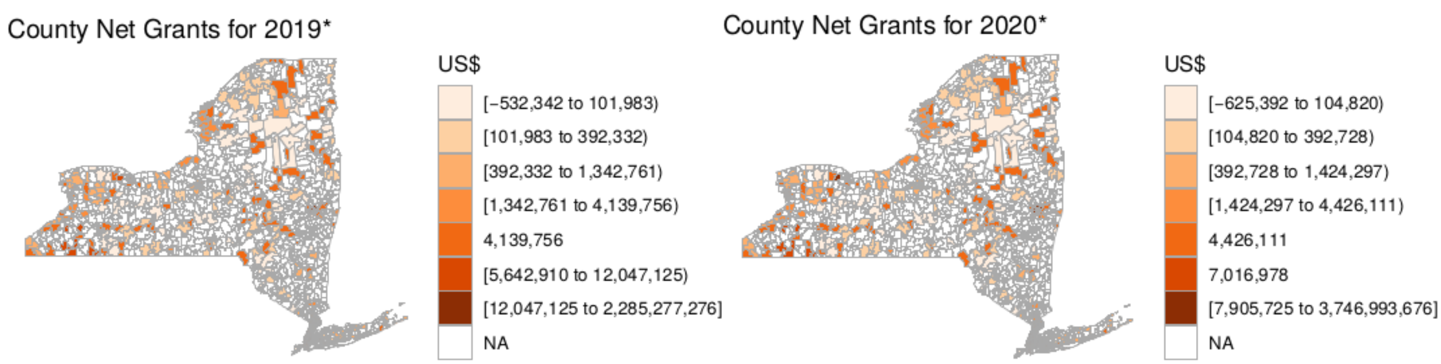

Figure 3: County Grants and predicted values in a map.

Bar visualization of past data and predictions.

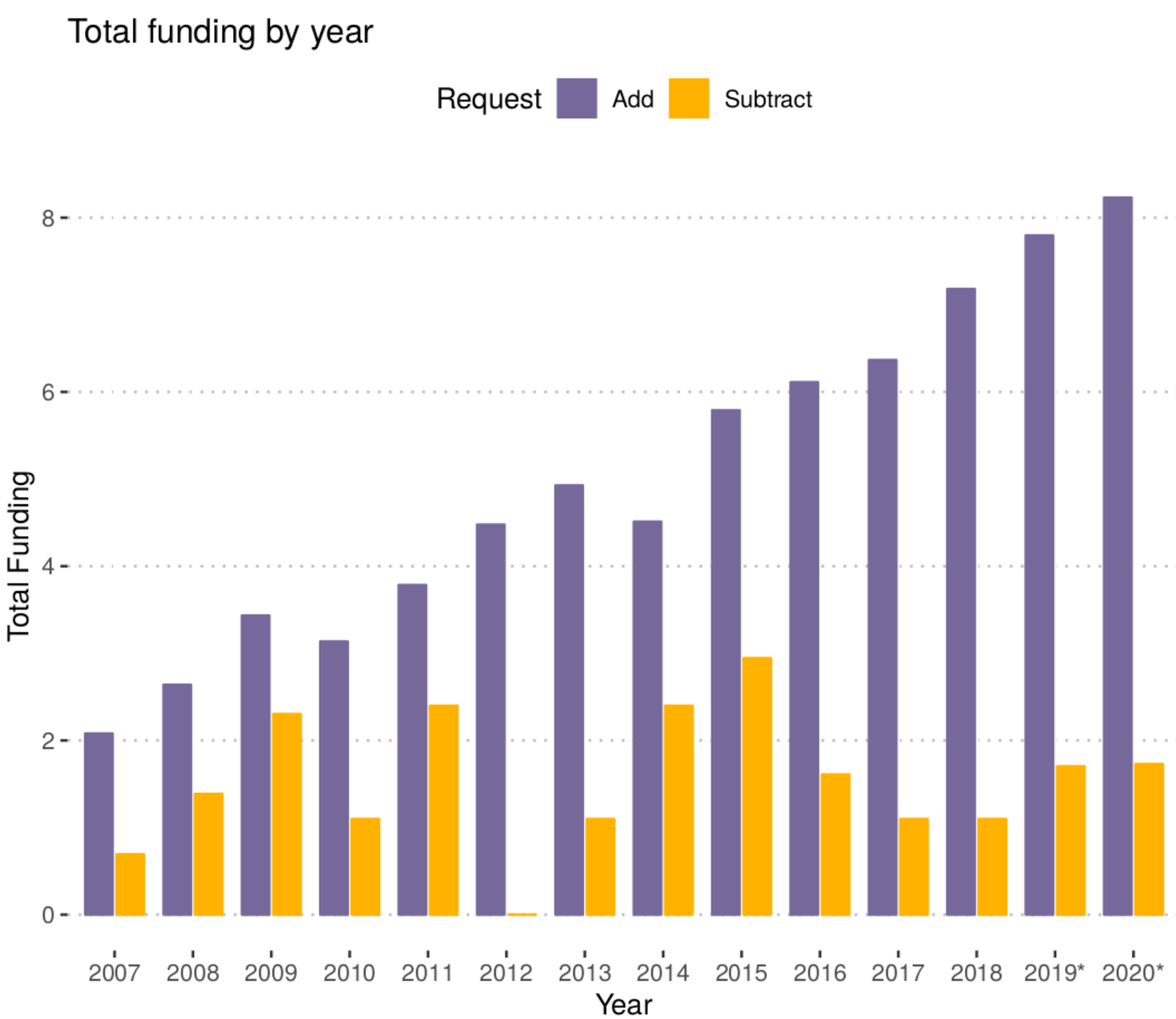

Figure 4: Total funding in millions of US dollars by request type. The scale is set to logarithmic. The graph includes predictions for 2019 and 2020. Prediction years show a star (*) next to the year.

Net funding visualization of past data and predictive values.

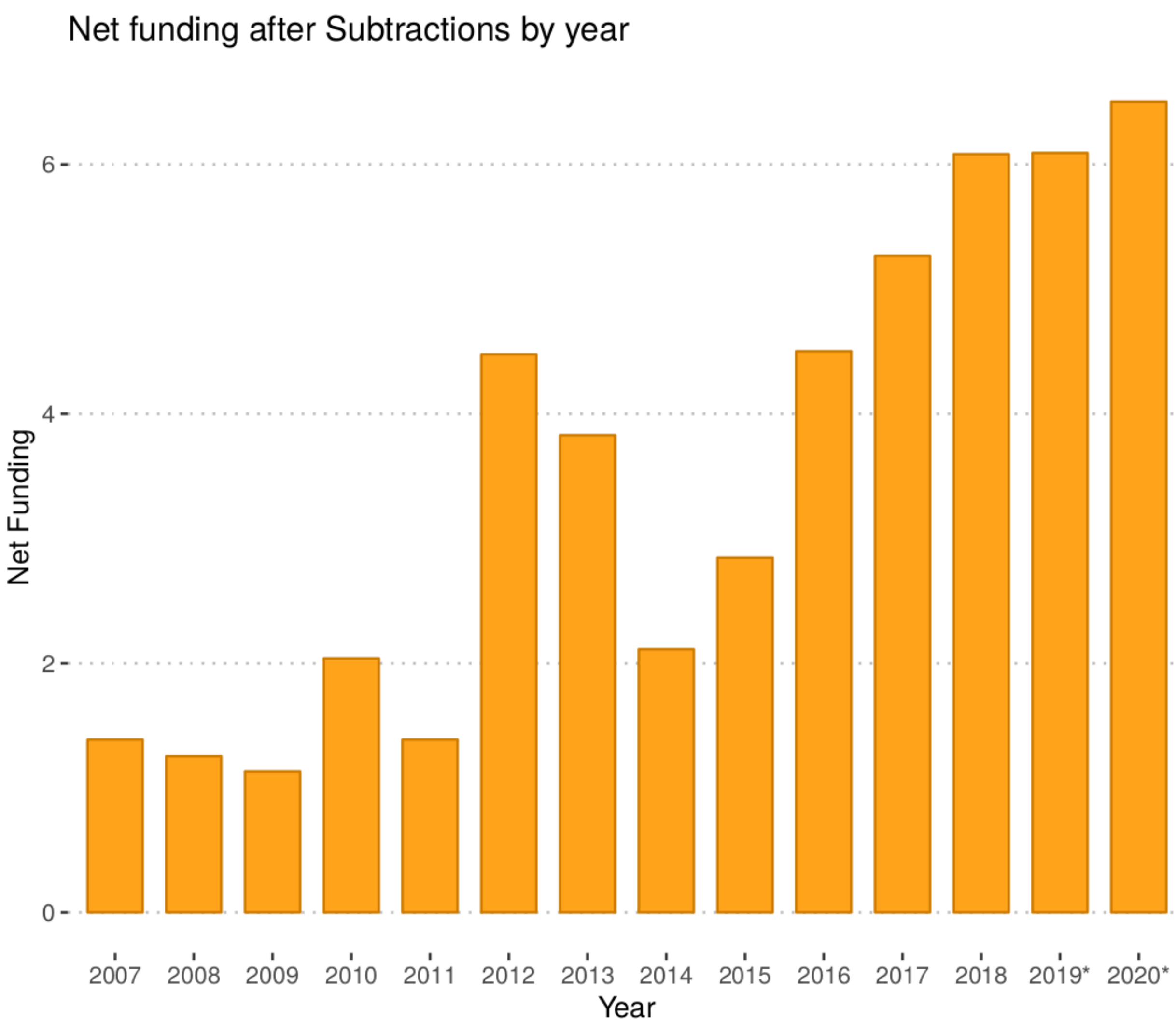

Figure 5: Net funding in millions of US dollars. The scale is set to logarithmic. The graph includes predictions for 2019 and 2020. Prediction years show a star (*) next to the year.

Conclusions

Based on the grant analysis, the only thing we are sure of is that grant funding has gradually increased for the last ten years and that the location of where grants are rewarded have mostly been rural areas. Since many high-population areas are funded by non-federal sources, this makes sense to spread out grant funding to low-populated areas. What is difficult to define is whether the funding has made a significant difference in the level of education in these areas.

Based on the initial exploratory analysis, we were encouraged to see so many different variables that described the dataset. However, shortly after we discovered most were unusable. It's highly probable that anyone you ask would say that education is a high priority, thought the lack of effort in organizing the data and explaining it is poor. With millions of dollars being distributed every year, it should be rare to find a record with as much missing data as we've seen. And with our analysis consisting of only New York grants, standardization across other states may or may not be needed if other states collect grant information with different variables.

With the complicated process of applying for grants, many schools or districts may be discouraged to apply for them at all. We judge this by an increasing trend of asking money from the community through websites like DonorsChoose.org. and GoFundMe.com

This dataset should be very valuable when analyzed with other data about test scores and graduation rates across New York. Future analysis of this dataset combined with other data could be used to see how grant funding impacts the lives of teachers, researchers and students. The issue is whether the data set can be relied upon as more care in the backend of data collection is needed.

R Code:

https://raw.githubusercontent.com/raghu74us/DATA-621/master/FinalProject/Findings_Final_Project_DATA621.Rmd

References

https://www.grants.gov/web/grants/learn-grants/grant-eligibility.html

https://www.grants.gov/applicants/organization-registration/step-1-obtain-duns-number.html

https://beta.sam.gov/

http://reports.weforum.org/global-competitiveness-report-2015-2016/education/

https://opendocs.ids.ac.uk/opendocs/ds2/stream/?#/documents/3546245/page/5

http://hanushek.stanford.edu/sites/default/files/publications/Hanushek%2BWoessmann%202010%20IntEncEduc%202.pdf

https://mpra.ub.uni-muenchen.de/46143/1/MPRA_paper_46143.pdf

https://www.nber.org/papers/w20847.pdf - !!!

http://www.oecd.org/innovation/research/1825455.pdf - barro 2001

http://hanushek.stanford.edu/sites/default/files/publications/Hanushek%202012%20FourPercentSolution_0.pdf - chapter in book hanushek

http://theconversation.com/does-government-spending-on-education-promote-economic-growth-60229

http://edpolicy.education.jhu.edu/wp-content/uploads/2017/05/Does-Money-Matter-Commentary.pdf

http://www.ascd.org/publications/educational-leadership/may02/vol59/num08/Unequal-School-Funding-in-the-United-States.aspx

https://edtrust.org/resource/funding-gaps-2018/

http://hanushek.stanford.edu/sites/default/files/publications/Hanushek%202003%20EJ%20113%28485%29.pdf - failure of input-based schooling policies

does money matter http://edpolicy.education.jhu.edu/wp-content/uploads/2017/05/Does-Money-Matter-Commentary.pdf

https://cepa.stanford.edu/sites/default/files/cofr-efp.pdf

unequal funding in the us

https://www.usnews.com/news/best-states/articles/2018-02-27/in-most-states-poorest-school-districts-get-less-funding

https://www.nap.edu/read/10023/chapter/9

funding methodology

https://www.usnews.com/news/best-states/articles/2017-12-08/state-education-finance-systems-take-center-stage

Appendices

Supplemental tables and/or figures

Table 1: Net funding by year with respective request type in millions of US dollars. Prediction years show a star (*) next to the year.

year	Add	Subtract	Net
2007	8	2	6
2008	14	4	10
2009	31	10	21
2010	23	3	20
2011	44	11	33
2012	88	1	87
2013	138	3	135
2014	91	11	80
2015	327	19	308
2016	451	5	446
2017	582	3	579
2018	1315	3	1312
2019*	2431.91	5.49	2426.42
2020*	3755.97	5.63	3750.34

Table 2: Data table with predictions for 2019 and 2020 for each county.

County	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019*	2020*
0	1.51	0.59	4.09	1.95	7.04	5.09	1.15	0	0	0	0	0	4.19	4.48
1	0.14	0	0	0.06	0	42.97	125	27.35	298.52	327.92	445.13	1048.43	2286.73	3748.49
101	0	0	0	0	0.05	0.06	0.21	0.06	0.06	0.19	0.44	0.73	0.23	0.24
103	0.15	0	0	0	0.09	0.57	0.48	0.23	0.31	0.52	1.53	2.46	1.19	1.34
105	0	0	0	0	0.01	0.03	0.02	0.02	0.05	0.04	0.05	0.05	0.04	0.04
109	0	0	0	0.05	0	0.03	0.45	0.01	0.08	0.39	1.62	2.61	0.39	0.41
11	0	0	0	0.05	0	0	0	0	0	0.55	0	0.19	0.31	0.32
111	0.03	0	0	0	0	0	0	0.29	0	0	0	1.05	0.99	1.13
113	0	0	0	0	0.25	0.07	0.01	0.02	0.04	0.31	0.05	0.12	0.05	0.05
115	0	0.2	0	0	0.06	0.18	0.15	0.13	0.13	0.13	0.14	0.1	0.11	0.11
117	0	0	0	0	0	0	0	0.37	0	0.19	0	0.73	0.41	0.41
119	0.05	0.1	0.17	0.02	1.22	0.27	0	4.01	0.61	1.04	0.33	5.02	2.34	3.21
121	0	0	0	0	0	0.03	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
123	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13	0.05	0.02	0	0	0.04	0.19	0.13	0.08	0.57	0.8	0.41	0.34	0.82	1.02
15	0	0	0	0	0	0	0	0	0	0	0	1.6	1.6	1.6
17	0	0	0	0	0	0.02	0.02	0.03	0.03	0.03	0.02	0	0.02	0.02
19	0	0	0	0	0.24	0.04	0	0.04	0.03	0.94	0.14	0.68	0.17	0.16
21	0	0	0	0	0.02	0.05	0.04	0.09	0.06	0.07	0.03	0.07	0.06	0.06
23	0.05	0	0	0	0.03	0.03	0.3	0.33	0.04	0.04	0.04	0.03	0.05	0.04

25	0	0	0	0	0.06	0.1	0.1	0.1	0.11	0.11	0.16	0.14	0.13	0.13
27	0	0	0.04	0.25	0.64	0.24	1.51	0	0.09	0.68	0.36	1.34	0.68	0.73
29	1.52	2.36	0.37	0.45	0.5	1.2	0.22	0.51	4	5.04	1.75	7.03	3.62	4.19
3	0	1.79	0	0	0.02	0.05	0.24	0.07	0.07	0.08	0.12	0.1	0.04	0.03
31	0	0	0	0	0.05	0.12	0.08	0.09	0.09	0.09	0.11	3.41	0.16	0.15
33	0	0	0	0	0	0.03	0.04	0.72	0.5	0.51	0.51	1.27	0.81	0.96
35	0	0	0	0	0	0.04	0.01	0.02	0.02	0.03	0.03	0.53	0.04	0.04
37	0	0.02	0.06	0.06	0.04	0.29	0.43	0.32	0.03	2.13	0.04	1.06	0.46	0.53
39	0	0	0	0	0	0.02	0.01	0.02	0.02	0.02	0.03	0.03	0.02	0.02
41	0	0	0	0	0.09	0.07	0.06	0.07	0.07	0.07	0.08	0.06	0.07	0.07
43	0	0	0	0	0	0.03	0.05	0.05	0.05	0.05	0.04	0.05	0.05	0.05
45	0	0	0	0.15	0.08	0.32	0.43	4.04	3.32	8.79	0.23	7.75	2.23	2.42
47	0	0	0	0	0	0	0	0	0	0	35.96	75.91	54.42	54.3
49	0	0	0	0	0.05	0.06	0.04	0.05	0.04	0.06	0.05	0.07	0.05	0.05
5	0	0	0	0	0	0	0.3	0	0.3	1.01	1.42	10.7	1.4	1.48
51	0.21	0	0	0	0	0	0	0	0.05	0.04	0.05	0.72	0.1	0.09
53	0	0	0	0	0.03	0.09	0.09	0.06	0.08	0.08	0.07	0.44	0.12	0.12
55	0.18	0.03	0	0	0.16	0	0.32	0.9	0.84	3.04	7.31	4.59	5.71	7.96
57	0	0	0	0	0	0.04	0.02	0	0.48	0	0	0	0.16	0.16
59	0.32	0	0	0	0	0	0	0.26	3.11	0.38	0.3	2.84	0.27	0.28
61	2.62	6.4	21.75	17.54	24.69	8.47	4.24	32.64	6.34	28.5	46.23	93.18	62.37	75.34
63	0.01	0.15	0.2	0.05	0.02	0.03	0.02	0.51	0	0.47	0.11	0.31	0.42	0.51
65	0	0.86	0	0	1.04	0.03	0.01	0.01	0.01	0.27	0.02	1.99	0.04	0.03
67	0.19	0	0	0.07	0.38	0.08	0	2.58	0.28	0.99	0.43	2.75	0.75	0.85
69	0	0	0	0	0	0	0	0	0	0.16	0.01	0.11	0.06	0.06
7	0.04	0	0	0	0.24	0.44	0	0	0	0.53	0.16	1.38	0.18	0.19
71	0	0	0	0	0.41	20	0.01	0	0.04	0.64	4.69	0.27	0.2	0.16
73	0	0	0	0	0	0	0	0.02	0.02	0.15	0	0	0.05	0.05
75	0	0	0	0	0	0	0	2.89	0	0	0	0	2.89	2.89
77	0	0	0	0	0.02	0.11	0.11	0.15	0.58	0.19	0.19	0.6	0.28	0.3
79	0	0	0	0	0.03	0.03	0.02	0.02	0.02	0.02	0.03	0.02	0.02	0.02
81	0.06	0.87	0.1	0.89	3.65	0.73	0.09	5.19	1.29	3	0	14.83	12.1	16.87
83	0	0	0	0	0	0	0.91	1.59	1.71	10.68	29.34	11.81	4.01	4.18
85	0	0	0	0	0	0	0	0	0	0	0	0	0	0
87	0.04	0	0	0	0	0	0	0	0	0.25	0.28	0.32	0.47	0.56
89	0	0	0.17	0	0.29	0.11	0.19	0.08	0.08	0.67	0.36	1.61	0.39	0.4
9	1.12	0.96	3.97	1.27	2.55	5.72	0.9	4.47	2.61	49.27	1.06	3.1	6.36	7.14
91	0	0	0	0	0	0	0	0	0	0	0.01	0.01	0.01	0.01
93	0.01	0	0	0	0	0	0	0.28	0.38	0.01	0	0.46	0.03	0.03
95	0	0	0	0	0	0.06	0.03	0.03	0.04	0.04	0.04	0.03	0.03	0.03
97	0	0	0	0	0	0	0	0	0.25	0	0.02	0.02	0.04	0.03
99	0	0	0	0	0	0.02	0.01	0.01	0.01	0.01	0.02	0.02	0.01	0.01

Table 3: Data table with predictions for 2019 and 2020 for each Congressional District. (Part 1)

District	2007	2008	2009	2010	2011	2012	2013
1	0	0	0	0	0.14	0.57	0.55
10	0.05	0	0.21	0.13	0.49	0.57	0
11	1.12	0.96	0.02	0.54	3.66	4.68	0.51
12	0	0	3.74	0.67	0	1.23	0
13	0	0	0	0	0	0	0
14	1.26	1.98	11.31	0.04	12.57	1.19	1.41
15	0	0	0.45	0	0.84	0.7	1.13
16	0	0	0	0	0	0	0.3
17	0.04	0.1	1.43	0.69	0	0	0
18	0.07	0	0.17	0.02	1.22	0.27	0
19	0	0	0	0	1.08	20.27	0.03
2	0.15	0	0	0	0	0.02	0
20	0	0.2	0.04	0.25	0.73	0.55	1.93
21	0.32	0	2.32	1.23	0.05	45.57	125.96
22	0.07	0	0	0.05	0.26	0.46	0.46
23	0	0	0.17	0.15	1.61	1.31	1.36
24	0.19	0.86	0	0.05	1.26	0.29	0.46
25	0.19	0.51	0	0.07	0.38	0.08	0
26	0.43	0.05	0.06	0.06	0.5	1.42	0.54
27	1.4	2.38	0.24	0.28	0.21	0.39	0.27
28	0.13	0.17	0.2	0.25	1.13	0.11	0
29	0.06	1.79	0	0	1.04	0.66	1.48
3	0	0	0	0	0	0	0
4	0.04	0.05	0	0	0	0	0
5	0.32	0.73	0	0.45	3	0	0
6	0	0	0	0.22	0.08	0	0.09
7	0	0	0	0.22	0	0	0
8	2.47	4.56	10.54	17.5	13.85	7.81	1.98
9	0	0	0	0	0	0	0

Table 4: Data table with predictions for 2019 and 2020 for each Congressional District. (Part 2)

District	2014	2015	2016	2017	2018	2019*	2020*
1	0.21	0.29	0.34	1.53	2.19	0.73	0.75
10	0.5	2.57	8.34	25.48	38.18	43.59	67.82
11	3.56	2.21	0	0	0	4.17	4.96
12	0	0.44	13.82	16.66	48.65	3.09	3.13
13	0	0.17	6.46	4.09	13.33	3.39	3.25
14	3.29	1.28	0.26	0	5.51	2.1	2.17
15	3.33	0.51	0.39	1.36	2.12	1.19	1.22
16	0	0.3	0.31	0.06	3.85	0.34	0.31
17	0	0	0.6	0.6	2.33	0.07	0.07
18	4.01	0.61	1.76	2.46	1.28	4.33	5.65
19	0.02	0.2	0.66	3.12	2.22	0.29	0.24
2	0.01	2.94	0.18	0	0.27	0.12	0.12
20	0.45	0.48	335.95	474.47	1061.28	229.96	396.78
21	29.25	301.13	14.13	1.73	14.95	51.25	66
22	0.33	0.11	0.91	0.33	3.84	1.11	1.29
23	8.03	4.19	2.34	3.69	8.77	7.39	8.88
24	0.52	0.64	1.86	0.43	3.71	1.72	1.99
25	2.95	0.44	3.04	7.31	4.59	3.34	4.34
26	0.77	1.72	6.5	1.5	6.54	10.85	15.92
27	0.51	2.77	2.86	0.47	2.62	1.35	1.47
28	0.52	0.37	0.02	0	0	0.11	0.11
29	1.09	1.09	0.18	0	0	1.04	1.09
3	0	0	0	0	0.13	0	0
4	0.25	0.19	0.38	0.29	2.71	1.32	1.59
5	0	0	1.07	0	2.48	2.47	2.74
6	1.03	0	1.07	0	4.89	1.55	1.71
7	0.65	1.04	46.01	23.16	73.21	15.18	15.91
8	28.92	1.46	1.48	9.14	3.8	3.85	3.66
9	0.6	0.25	0.28	3.66	1.64	0.82	0.79

Table 5: Data table with predictions for 2019 and 2020 for each Zip Code. Due to the long length, only years 2017 – 2020 are provided*

ZipCode	2017	2018	2019*	2020*	ZipCode	2017	2018	2019*	2020*
06390	0	0	0.02	0.02	**12865**	0.02	0.02	0.02	0.02
10001	0	1.59	0.57	0.51	**12870**	0	0	0	0
10002	0	0.6	0.6	0.6	**12883**	0	3.32	3.32	3.32
10003	2.12	0.23	1.02	0.92	**12901**	0.11	0.65	0.39	0.39
10004	18.16	12	12.97	14.86	**12920**	0.03	0.03	0.03	0.03
10005	3.23	6	4.33	4.2	**12921**	0.04	0.03	0.03	0.03
10006	0	1.14	1.14	1.14	**12928**	0	0.01	0.01	0.01
10007	0.25	3.01	1.03	1.02	**12932**	0.02	0.01	0.01	0.01
10010	11.06	0.14	0.5	0.42	**12937**	0.46	0.5	0.47	0.47
10011	0.82	5.12	2.8	2.96	**12943**	0.02	0	0.02	0.02
10012	1.19	5.48	0.27	0.26	**12953**	0	0	0	0
10013	0	0.67	0.97	0.95	**12970**	0	0.26	0.3	0.33
10014	0	0	1.09	0.92	**12980**	0.01	0.01	0.01	0.01
10016	1.94	4.11	0.63	0.6	**12983**	0	0	0	0
10017	0.04	1.29	0.09	0.09	**12986**	0	0	0	0
10018	0.5	28.62	3.89	3.98	**12993**	0.02	0.02	0.01	0.01
10019	0.35	1.43	0.66	0.65	**12996**	0.02	0.02	0.02	0.02
10021	0	0	0.08	0.08	**13021**	0	0.19	0.15	0.16
10022	0	0	0.07	0.07	**13031**	0	0	0	0
10023	0	0	0.51	0.51	**13035**	0	0.32	0.32	0.32
10024	0	1.51	1.68	1.77	**13040**	0.01	0.01	0.01	0.01
10025	0	0	0.3	0.3	**13045**	0	0	0.24	0.26
10027	0.73	8.5	2.18	2.32	**13052**	0.01	0.01	0.01	0.01
10029	0.54	0.55	0.67	0.66	**13068**	0	0	0.02	0.02
10030	2.62	0	2.13	2.13	**13090**	0	0	0.03	0.03
10031	0.3	2.8	0.74	0.71	**13101**	0.03	0.03	0.02	0.02
10032	0	0	0	0	**13114**	0	0	0	0
10035	0	0	0	0	**13126**	0	0	2.89	2.89
10036	0	5.81	2.82	3.4	**13139**	0	0	0.4	0.4
10038	2.21	0.63	0.96	0.95	**13155**	0	0	0.02	0.02
10040	0	0	0.08	0.08	**13160**	0	0	0.16	0.16
10065	0.19	0.89	0.41	0.42	**13203**	0	0	0	0
10115	0	0	3	3	**13207**	0.01	0.02	0.01	0.01
10121	0	0	0.01	0.01	**13210**	0.14	0.16	0.2	0.18
10122	0	0	0.19	0.19	**13211**	0	0	0.07	0.07
10128	0	0	0.42	0.42	**13212**	0.03	0.03	0.03	0.03
10153	0	0	0	0	**13214**	0	0.04	0.17	0.16
10166	0	0	0	0	**13215**	0.09	0.73	0.27	0.27
10168	0	0	0.94	0.95	**13244**	0.16	1.77	0.5	0.51
10281	0	1.07	1.07	1.07	**13305**	0	0	0	0
10314	0	0	0	0	**13314**	0	0.01	0.01	0.01
10452	0.74	0	0.74	0.74	**13320**	0	0	0.02	0.02
10453	0	0.78	0.78	0.78	**13323**	0	0	0	0
10457	0	0	0	0	**13335**	0.02	0.02	0.02	0.02

10458	0.62	8.69	0.82	0.77
10459	0	0	0	0
10460	0	0	0	0
10465	0	0	0	0
10468	0	0	0.62	0.62
10471	0.06	0.83	0.26	0.27
10474	0	0.4	0.4	0.4
10522	0.33	1.08	1.69	2.12
10524	0.03	0.02	0.02	0.02
10553	0	0	0.26	0.26
10562	0	0	0	0
10566	0	0	0.16	0.16
10567	0	0	0	0
10570	0	0	0	0
10573	0	0	0	0
10577	0	0	0	0
10591	0	0	0.01	0.01
10595	0	0.93	0.34	0.34
10598	0	0	0	0
10601	0	0	0.02	0.02
10603	0	0	0	0
10701	0	0	0	0
10708	0	0	0.4	0.4
10710	0	0.66	0.52	0.52
10801	0	0	0.17	0.17
10805	0	2.35	2.95	2.92
10901	0.28	0.32	0.21	0.21
10913	0	0	0	0
10922	4.64	0	0.35	0.33
10923	0	0	0	0
10924	0	0	0	0
10940	0	0.24	0.94	0.76
10987	0.04	0.02	0.04	0.04
10994	0	0	0.11	0.12
11020	0	0	0	0
11101	0	1.96	3.62	4.66
11106	0	0	0	0
11201	32.3	73.06	55.99	78.8
11203	0	0	0	0
11205	0	0	0.04	0.04
11206	0	0.25	0.18	0.18
11207	0	0.36	0.24	0.24
11208	0	0.26	0.26	0.26
11210	0	0.56	0.38	0.39
11212	3.66	0	3.66	3.66
11216	0	0	0	0

13346	0	0.04	0.03	0.03
13350	0	0	0	0
13360	0.02	0.02	0.02	0.02
13367	0	0	0	0
13402	0.03	0.02	0.02	0.02
13409	0.03	0.03	0.02	0.02
13413	0	0	0.91	0.91
13420	0.02	0.02	0.02	0.02
13431	0.02	0.02	0.02	0.02
13438	0	0	0.01	0.01
13439	0.01	0.01	0	0
13452	0	0	0.03	0.03
13459	0	0	0.02	0.02
13475	0.01	0.01	0.01	0.01
13501	0	1.99	0.01	0.01
13502	0	0	0	0
13601	0.03	0.41	0.11	0.1
13605	0	0.01	0.01	0.01
13606	0	0	0	0
13607	0	0	0.03	0.03
13611	0.01	0	0	0
13617	0	0.54	0.38	0.38
13619	0.02	4.16	0.48	0.46
13622	0.03	0.03	0.03	0.03
13624	0	0.01	0	0
13625	0.02	0.02	0.02	0.02
13626	0.03	0.04	0.03	0.03
13630	0.01	0.01	0.01	0.01
13634	0	0.01	0.01	0.01
13642	0	0	0	0
13646	0.02	0.02	0.02	0.02
13648	0.02	0.02	0.02	0.02
13652	0.01	0.01	0.01	0.01
13655	0	0.47	0.47	0.47
13656	0.04	0.04	0.04	0.04
13658	0.02	0.02	0.02	0.02
13660	0.02	0	0.02	0.02
13662	0.11	0.11	0.11	0.11
13664	0.01	0.01	0.01	0.01
13672	0.02	0.02	0.02	0.02
13673	0.05	3.03	0.56	0.52
13676	0.11	0.84	0.38	0.38
13684	0	0	0.01	0.01
13685	0.03	0.04	0.03	0.03
13690	0.01	0.01	0	0
13699	0	0	0.22	0.22

11217	0	0	0.06	0.06
11218	0	0.64	0.64	0.64
11220	0	0	0.34	0.34
11224	0	0	0	0
11225	0	0.44	0.44	0.44
11226	0	0	0.5	0.5
11235	0	0.33	0.33	0.33
11238	0	0	0	0
11361	0	0.1	0.1	0.1
11364	0	0	0.25	0.25
11367	0	2.22	0.87	0.87
11368	0	3.08	2.72	2.91
11369	0	2.42	0.86	0.86
11372	0	0	0.22	0.22
11373	0	0.52	0.52	0.51
11378	0	0	0.22	0.22
11434	0	0	0	0
11439	0	2.05	0.77	0.83
11451	0	2.48	1.73	1.75
11507	0	0	0.32	0.32
11520	0	0	0.19	0.19
11530	0	1.57	0.82	0.92
11542	0	0	0	0
11548	0	0.13	0.13	0.13
11549	0.29	0.89	0.68	0.74
11550	0	0	0	0
11553	0	0	0.38	0.38
11568	0	0	0	0
11570	0	0	0.23	0.23
11571	0	0.26	0.26	0.26
11575	0	0	0	0
11703	0	0	0	0
11704	0	0	0	0
11706	0	0	0	0
11717	0	0	0.16	0.16
11735	0	0.24	0.9	0.85
11741	0	0	0.32	0.32
11746	0	0	0	0
11752	0	0	0.01	0.01
11769	0	0	0.01	0.01
11770	0	0.03	0.01	0.01
11772	0.33	0	0.32	0.32
11778	0	0	0.14	0.14
11784	0.57	0.5	0.53	0.53
11787	0	0	0	0
11794	0.49	1.06	0.19	0.19

13730	0.02	0	0.02	0.02
13731	0.01	0.01	0.01	0.01
13750	0.02	0.02	0.01	0.01
13755	0.02	0.02	0.02	0.02
13775	0.02	0.02	0.02	0.02
13776	0.02	0.02	0.02	0.02
13783	0.02	0	0.01	0.01
13790	0	0	0.04	0.04
13796	0.02	0.02	0.02	0.02
13807	0.03	0.03	0.03	0.03
13808	0.03	0.03	0.03	0.03
13820	0	0.42	0.42	0.42
13842	0.02	0.02	0.02	0.02
13850	0	0	0.35	0.35
13901	0	0	0	0
13902	0.16	1.38	0.52	0.51
13905	0	0	0.07	0.07
14001	0.13	0.18	0.12	0.12
14004	0	0	0	0
14005	0	0	0.06	0.06
14006	0.21	0.31	0.2	0.21
14011	0	0	0	0
14020	0.01	0.96	0.13	0.11
14036	0	0	1.97	2.25
14043	0	0	1.1	1.1
14048	0	0	0.01	0.01
14058	0.03	0.03	0.03	0.03
14062	0.02	0.02	0.02	0.02
14063	0	0.04	0.18	0.17
14068	0	0	0.01	0.01
14070	0.25	0.37	0.22	0.23
14075	0	0	0	0
14080	0	0.01	0.01	0.01
14081	0	0	0.05	0.05
14094	0	0	0	0
14098	0	0	0.02	0.02
14105	0	0	0.01	0.01
14108	0	0	0.02	0.02
14109	0	0	0	0
14111	0.03	0.03	0.03	0.03
14120	0	0	0.16	0.16
14125	0	0	0.06	0.06
14127	0	0.08	0.09	0.09
14132	0	0.1	0.1	0.1
14136	0.13	0.22	0.13	0.13
14171	0.02	0.02	0.02	0.02

11798	0	0	0	0
11930	0	0.04	0.04	0.04
11932	0.02	0.02	0.02	0.02
11933	0	0.08	0.04	0.04
11934	0.03	0.03	0.03	0.03
11953	0	0.37	0.37	0.37
11954	0	0	0.03	0.03
11957	0.02	0.02	0.01	0.01
11959	0	0	0.02	0.02
11963	0	0	0.03	0.03
11964	0.03	0.03	0.03	0.03
11968	0.05	0.05	0.05	0.05
11975	0	0	0.03	0.03
12010	0	0	0.48	0.48
12022	0	0	0	0
12029	0	0	0.02	0.02
12032	0.01	0.01	0.01	0.01
12043	0	0	0	0
12054	0	0	0	0
12061	0	0	0	0
12076	0.02	0.01	0.01	0.01
12093	0.02	0.02	0.02	0.02
12095	0	0.5	0.5	0.5
12110	0	0	0	0
12125	0	0.04	0.02	0.02
12134	0.03	0.02	0.02	0.02
12139	0	0	0	0
12144	29.33	11.51	17.18	23.03
12155	0.03	0.03	0.02	0.02
12164	0.01	0	0.01	0.01
12167	0.02	0.02	0.02	0.02
12180	0	0.3	0.3	0.3
12182	0	0	0	0
12189	0	0	0	0
12190	0.02	0.02	0.02	0.02
12197	0.02	0.02	0.02	0.02
12203	0	0	0.99	1.12
12204	0	0	0	0
12205	0	0.01	0.01	0.01
12206	0	0	0	0
12207	0.66	0.82	1.53	1.64
12208	0	0	0.95	0.95
12210	0	7	2.49	2.63
12222	0.07	0.24	0.12	0.11
12226	0	0	0	0
12234	418.5	1014.48	2072.12	3270.19

14174	0	0	0.18	0.18
14201	0	0.04	0.14	0.12
14202	0.24	0.61	0.76	0.76
14203	0	0	0.31	0.32
14207	0	0	0.5	0.46
14208	0	0.67	1.14	1.23
14209	0	0.5	0.5	0.5
14212	0	0	0	0
14214	0.1	0.88	0.35	0.35
14215	0	0	0	0
14222	0.03	0.88	0.28	0.27
14223	0	0	0	0
14225	0	0.74	0.53	0.53
14226	0	0	0.49	0.57
14228	1.01	2.1	1.52	1.51
14260	0	0	0.93	0.98
14263	0	0	0	0
14304	0.11	0.21	0.15	0.15
14411	0	0	0	0
14420	0.37	0.97	0.51	0.51
14423	0.05	0	0.05	0.05
14424	0	0	0.14	0.14
14433	0	0	0.19	0.19
14450	0	0	0.03	0.03
14454	0	0.24	0.24	0.24
14456	0	0.11	0.11	0.11
14470	0	0	0.13	0.13
14471	0.01	0	0.02	0.02
14481	0	0.47	0.47	0.47
14482	0	0.07	0.07	0.07
14487	0	0	0	0
14489	0	0.73	0.73	0.73
14510	0	0	0.21	0.21
14522	0	0	0	0
14527	0	0	0	0
14541	0.02	0.02	0.01	0.01
14551	0	0	0	0
14559	0	0	0.03	0.03
14569	0	0	0	0
14590	0	0	0.37	0.37
14591	0.01	0.01	0.01	0.01
14607	0	0	0.16	0.16
14608	0	0	0.05	0.05
14611	0	0	0.55	0.7
14614	0	0	0	0
14615	0	0	0.04	0.04

12237	25.89	25.87	25.62	25.7
12248	0	0	0	0
12302	0	0	0.01	0.01
12303	0	0.01	0	0
12305	0	0.46	0.45	0.44
12307	0	0	0	0
12308	0	0	0	0
12455	0.01	0.01	0.01	0.01
12474	0.02	0.02	0.02	0.02
12484	0	1	1	1
12485	0.02	0.02	0.01	0.01
12493	0	0	0	0
12496	0.02	0.01	0.01	0.01
12504	0	0	0.47	0.46
12526	0.03	0.03	0.03	0.03
12538	0.36	0.35	0.05	0.04
12561	0	0.05	0.13	0.14
12571	0	0	0.04	0.04
12601	0	0.99	0.72	0.74
12603	0	0	0	0
12701	0	0	0	0
12732	0.03	0.03	0.03	0.03
12733	0	0	0	0
12748	0	0	0	0
12754	0	0	0	0
12758	0.01	0.01	0.01	0.01
12759	0	0	0	0
12776	0.02	0.01	0.01	0.01
12803	0	0	0	0
12804	0	0.09	0.16	0.16
12809	0.03	0.03	0.03	0.03
12814	0.02	0.02	0.01	0.01
12817	0.02	0	0.02	0.02
12827	0.04	0	0.03	0.03
12828	0	0	0.2	0.2
12838	0.04	0.04	0.04	0.03
12842	0.01	0.01	0.01	0.01
12847	0.01	0.01	0.01	0.01
12852	0.02	0.02	0.02	0.02
12853	0.02	0.02	0.01	0.01
12857	0.02	0.02	0.01	0.01
12861	0.01	0.01	0.01	0.01

14618	0	0.23	0.21	0.22
14620	0	0.47	0.34	0.36
14621	0	0	0	0
14623	5.47	1.07	1.1	1.12
14627	1.33	1.84	1.58	1.58
14642	0.14	0	0.47	0.46
14701	0.19	0	0.3	0.3
14711	0.02	0.01	0	0
14715	0	0	0	0
14716	0	0	0	0
14724	0.02	0.02	0.02	0.02
14731	0.03	0.03	0.03	0.03
14735	0	0	0	0
14739	0.01	0	0.01	0.01
14743	0.02	0.01	0.01	0.01
14744	0	0	0.29	0.29
14760	0	0	0.18	0.18
14767	0.02	0.02	0.02	0.02
14770	0	0	0	0
14772	0	0	0.02	0.02
14775	0	0	0.01	0.01
14779	0.74	2.67	0.67	0.69
14781	0.02	0.02	0.01	0.01
14801	0.37	0	0.37	0.37
14802	0	0	0.19	0.19
14806	0.02	0.02	0.01	0.01
14807	0.03	0.03	0.03	0.03
14809	0.01	0.01	0.01	0.01
14813	0.01	0.01	0.01	0.01
14815	0.02	0.02	0.02	0.02
14822	0.02	0.02	0.02	0.02
14830	0	0.29	0.29	0.29
14840	0.01	0.01	0.01	0.01
14843	0	0.37	0.14	0.14
14850	1.62	2.61	0.41	0.42
14870	0	0	0	0
14873	0.01	0.01	0.01	0.01
14880	0.02	0	0.01	0.01
14891	0	0	0.25	0.25
14895	0	0	0.03	0.03
14897	0.02	0.02	0.02	0.02
14901	0	0.25	0.25	0.25
14905	0	1.34	1.34	1.34