

# Final Project Proposal

CUNY School of Professional Studies - DATA 621 – Fall 2018

*Group 5*

*October 22, 2018*

## Contents

|   |          |
|---|----------|
| <b>Team members</b>   | <b>1</b> |
| <b>Final Project Proposal</b>   | <b>2</b> |
| Title: Prime awards model for the New York Department of Education grants from the federal government. . . . .  | 2        |
| Project requester . . . . .   | 2        |
| <b>Introduction</b>   | <b>2</b> |
| <b>Statement of the Problem or Need</b>   | <b>2</b> |
| <b>Objectives</b>   | <b>2</b> |
| <b>Plan of action</b>   | <b>3</b> |
| What is award data? . . . . .   | 3        |
| Why would we be interested in this data? . . . . .  | 3        |
| How do I use this form? . . . . .   | 3        |
| OUR SELECTED OPTIONS ARE... . . . .   | 3        |
| <b>Management Plan</b>  | <b>4</b> |
| Week 10/28/18 – 11/04/18: Gather data. . . . .  | 4        |
| Week 11/05/18 – 11/11/18: Introductory Analysis. . . . .  | 4        |
| Week 11/12/18 – 11/18/18: Building Insights. . . . .  | 4        |
| Week 11/19/18 – 11/25/18: Building Models and starting to prepare documentation. . . . .                        | 4        |
| Week 11/26/18 – 11/02/18: Introduction, Literature Review. . . . .  | 4        |
| Week 11/03/18 – 11/09/18: Methodology, Experimentation and Results, Discussion and Conclusions. . . . .         | 4        |
| Week 11/10/18 - 11/ 18/18: Abstract, Key Words, Prepare Presentation slides and submission, References. . . . . | 5        |
| <b>Conclusion</b>   | <b>5</b> |
| <b>Resources</b>  | <b>5</b> |

## Team members

Group 5:

- Albina Gallyavova albina.gallyavova68@spsmail.cuny.edu
- Chad Smith chad.smith62@spsmail.cuny.edu
- Nkasi Nedd nkasi.nedd@spsmail.cuny.edu
- Raghunathan Ramnath raghunathan.ramnath@spsmail.cuny.edu
- Duubar Villalobos Jimenez duubar.villalobosjimenez@spsmail.cuny.edu

# Final Project Proposal

**Title: Prime awards model for the New York Department of Education grants from the federal government.**

*The justification and validity of the project needs to be confirmed before the project proceeds. This document is used to clarify the project purpose and justification and to gain approval to proceed towards developing the Project Charter.*

## Project requester

- Nasrin Khansari, Professor of DATA 621 - Business Analytics and Data Mining.

## Introduction

This is our final group project proposal. The purpose of this project is to explore, analyze and model a real-world data set of our own interest using the regression modeling techniques learned in the course.

The idea is to develop a problem statement and research question based on the data set that we have obtained. We must survey the state-of-the-art literature and research developments dealing with empirical studies, algorithms, or methodologies related to our problem.

## Statement of the Problem or Need

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 strengths communities. On the other hand, 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. 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.

## Objectives

We propose to review the data set provided by USAspending.gov[1] related to grants provided to schools in the state of New York in order to:

- 1) Create an awards model able to predict government funds by school.
- 2) Create an awards model able to predict government funds by zip code.
- 3) Create an awards model able to predict government funds by city.

- 4) Create an awards model able to predict government funds by county.
- 5) Create an awards model able to predict government funds by congressional district.
- \*6) Map current awards over time (Not sure if we will have enough time).

According to the National Association of Elementary Schools Principals, they advocate for increased federal funding for cornerstone programs that have the most impact in our nation's schools. Title I grants, the main source of funding for No Child Left Behind, helps low-income and disadvantaged students improve their academic performance. This, along with IDEA grants, which fund services for special education, cannot meet the demand for these services and have been underfunded by the federal government for decades[2].

## Plan of action

This section presents our plan for obtaining the objectives discussed in the previous section.

There's an award data site named: USASPENDING.gov[1] from which we are able to extract the data.

## What is award data?

Award data contains all the details of the prime award and sub-award records from the federal government.

## Why would we be interested in this data?

Downloading this data gives us access to every attribute of any particular award, including data that may not be surfaced on this site.

## How do I use this form?

This form allows us to select specific awards by type, agency and sub-agency, and date range. Select an option in each section and click the "Download" button at the bottom. Heads up: all fields are required. We'll only be able to start the download when all sections are properly filled.

In particular, in order to download our data set, we will select as follows: from the following link:

[https://www.usaspending.gov/#/download\\_center/custom\\_award\\_data](https://www.usaspending.gov/#/download_center/custom_award_data)

## OUR SELECTED OPTIONS ARE...

- **Award Levels:** Prime Awards.
- **Agency:** Department of Education.
- **Recipient Location:** New York, United States.
- **Dates:** 10/01/2009 to present.
- **File Format:** CSV

The above request will take some time since we will look at the most collected data.

# Management Plan

This section presents our schedule for completing our proposed research. This research culminates in a formal report and project presentation slides, which will be completed by December 20th, 2018, 11:59pm EST. The final project report will be similar to technical papers we read in the literature. The report will not exceed 12 single-spaced pages. To reach this goal, we will follow collaboration techniques by contributing and dividing tasks in between team members.

## **Week 10/28/18 – 11/04/18: Gather data.**

- Download data and make it accessible for our team members for future analysis.
- Create a GitHub repository and download scripts.

## **Week 11/05/18 – 11/11/18: Introductory Analysis.**

- First impressions from our data.
- Visualizations.
- Summaries and cleaning strategies.

## **Week 11/12/18 – 11/18/18: Building Insights.**

- Best approaches.
- Experimentation.

## **Week 11/19/18 – 11/25/18: Building Models and starting to prepare documentation.**

- Discuss how other researchers have addressed similar problems such as:
  - what their achievements are.
  - what the advantage and drawbacks of each reviewed approach are.

## **Week 11/26/18 – 11/02/18: Introduction, Literature Review.**

- Discuss and explain how our investigation is similar or different to the state-of-the art.
- Cite relevant papers where appropriate.

## **Week 11/03/18 – 11/09/18: Methodology, Experimentation and Results, Discussion and Conclusions.**

- Discuss key aspects of our problem, data set and regression model(s).
- Explain at a high-level our exploratory data analysis.
- Explain how we prepared the data for regression modeling.
- Explain our process for building regression models, and our model selection.
- Explain the specifics of what we did for:
  - data exploration.
  - data preparation.
  - model building.
  - model selection.

- model evaluation.
  - model comparison.
- Explain the specifics of what we did for:
  - statistical analyses.
  - interpretation.
  - discussion of the results.
- Limitations
- Suggested areas for future work.

## **Week 11/10/18 - 11/ 18/18: Abstract, Key Words, Prepare Presentation slides and submission, References.**

- Select up to 5 key words.
- Abstract.
- Review and final edits.
- References: Cite relevant papers where appropriate in APA format.

## **Conclusion**

Less resources and more needs are linked to our New York stated education system. This proposed research will achieve the following goals: (1) To explain how much funding the federal government is providing our education system, and (2) determine how much federal funding will be provided in the near future. The research will include a formal presentation slides and a formal report on December 20.

## **Resources**

[1] <https://www.usaspending.gov>

[2] <https://www.naesp.org/advocacy/education-funding>