

Jessie Searles, Natalie Gurnik, Boyang Yu, Nick Omer, Xiao Ma, Matthew A. Lanham

Purdue University, Krannert School of Management

jsearles@purdue.edu; ngurnik@purdue.edu; yu713@purdue.edu; omer@purdue.edu; ma453@purdue.edu; lanhamm@purdue.edu

## ABSTRACT

This study provides a statistical deep dive into the components of program and student demographics of graduate data analytics and data science programs in the United States. We examine the effects of program location, type (full/part-time, online), degree (analytics vs data science), costs, and student features (age, % women, % international, professional experience, GPA, GMAT/GRE) on student outcomes (e.g., placement rate, starting salary). The purpose of this study is to objectively quantify empirically if any of these variables have a general association with outcomes and if so to what degree. Academic program administrators can use our findings to make strategic recruiting and admission decisions to improve outcomes.

## INTRODUCTION

Purdue (Krannert) currently offers 13 different Masters Programs, and one of those programs is the Masters in Business Analytics & Information Management (BAIM). Some rankings put BAIM as a top 10 program, and some consider demographics as a component in their ranking methodologies and program success. Our study investigates the impact demographics have on measurable program outcomes like placement and starting salary.

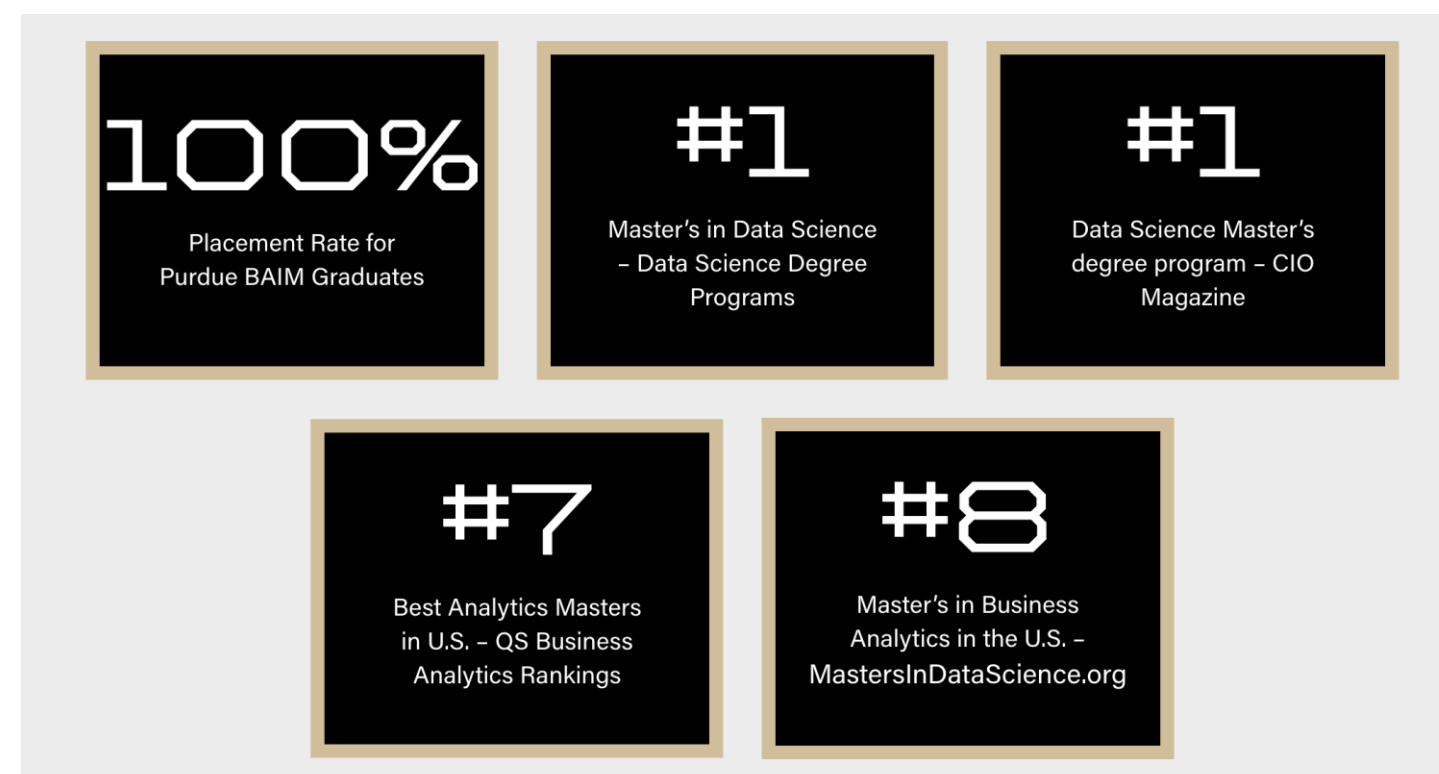


Fig 1. Purdue Rankings for BAIM Source:

MS BAIM - Purdue Krannert. (2020, July 29). Retrieved November 11, 2020

### Research Questions

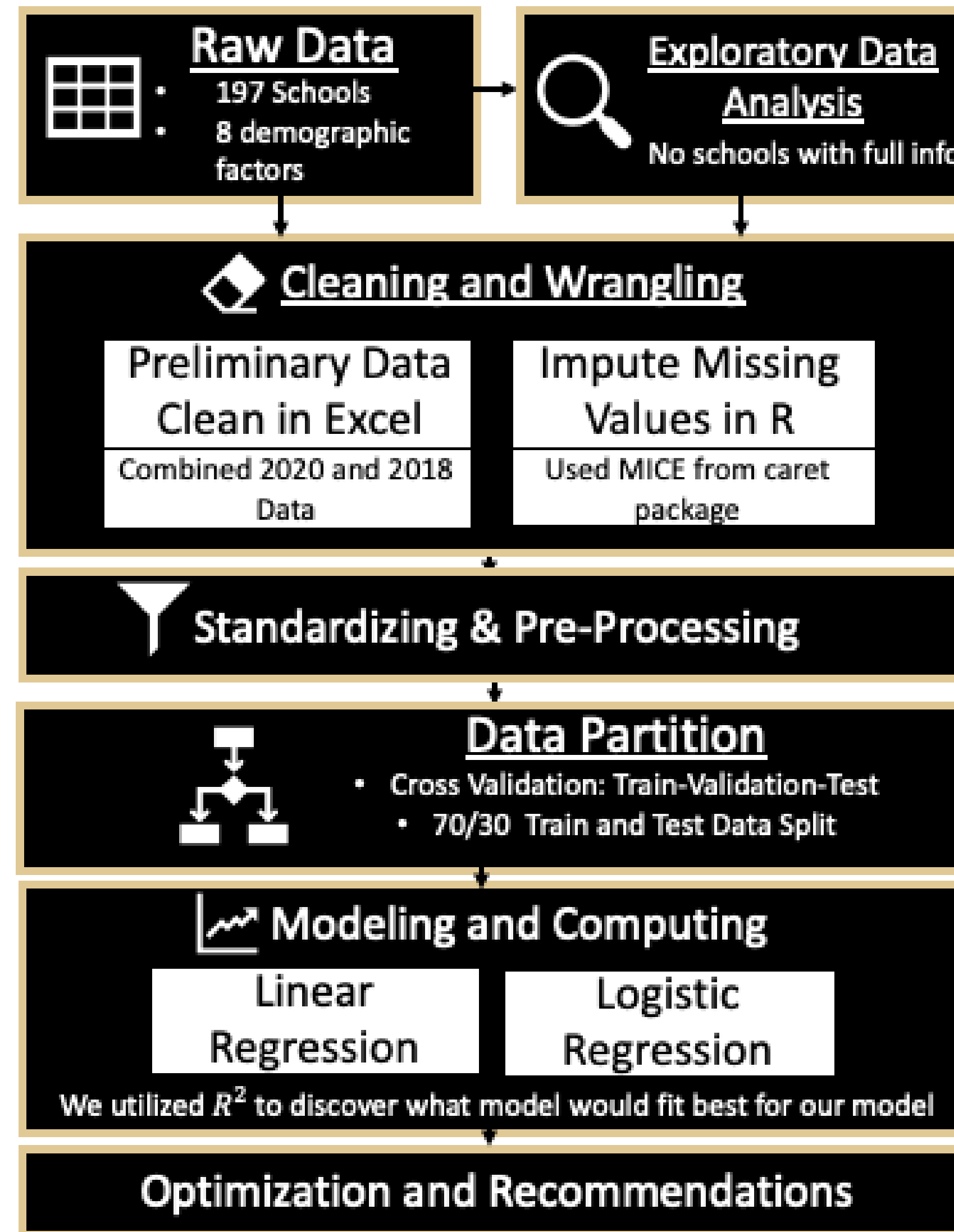
- Is there an association with student demographics and analytics program outcomes?
- Does a program's greater diversity affect placement rate and starting salary?

## LITERATURE REVIEW

We reviewed several articles that discussed BA/DS curriculum suggestions and employer's desired skills. However, we found no published research that models relationships between student demographics and a program's success (as we define it). Our study is novel since we utilize predictive analytics to see what effect, if any, demographics have on tangible, measurable program outcomes.

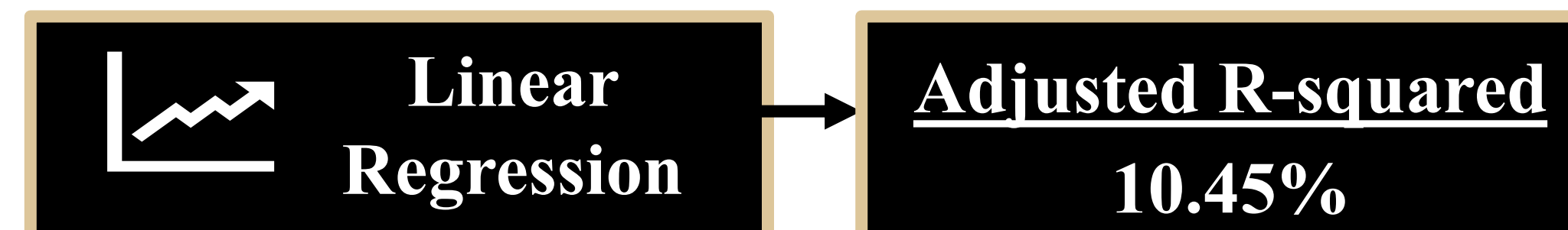
Study	Skills	Age	%Women	GPA	Work Experience	Duration
(2016) Turel & Kapoor	✓					
(2015) Gupta, Goul & Dinter	✓					
(2015) Schoenherr & Speier-Pero	✓					
(2014) Wixom, Douglas & Goul	✓					
Our Study		✓	✓	✓	✓	✓

## METHODOLOGY



## STATISTICAL RESULTS

We experimented with various models and decided that the linear regression on the Average Starting Salary (target) provided the interpretability needed to answer our research questions. We found only 10.45% of the variation in starting salary was explained by our predictors, suggesting other factors were not measured, potentially non-demographic factors, that were important.



Unexpectedly, our model suggested that younger and less experienced candidates tended to get higher salaries.

### Regression Equation

$$\text{Avg Starting Salary} = 71,569 + 0.0906 (\text{Avg Cost}) - 495(\text{Student Age}) + 9,943(\text{Student GPA}) - 122.1(\text{Student Work Exp in Months}) - 12,930(\% \text{ women}) - 3,061(\% \text{ International})$$

## PREDICTIVE TO PRESCRIPTIVE INSIGHTS

Using our parametric predictive model, we formulated it into a simple optimization model to identify the optimal demographics that would yield the greatest average starting salary. We then identified the top five programs that matched this demographic profile.

### Target Demographic Profile

Age	%Women	GPA	% International	Work Experience
21	0.25	4	0.03	0

The optimization results indicate that the top 5 programs with this demographic information to receive a high starting salary are:

Programs with these parameters

- Duke University
- Massachusetts Institute of Technology
- Harvard University
- Carnegie Mellon University
- Boston University

While our prescribed profile seems unintuitive at first glance, the schools that match our optimization prove otherwise. These schools are some of the most expensive in the world, and they report their placement statistics, where many programs do not. Our model suggests that the more money you spend on a program, the greater chance you will land a higher paying position. Since many schools do not report their placement statistics, these schools create a bias in our model that the higher the program's price, the more likely you are to get a higher starting salary.

### Average Cost of Programs

Boston	\$102,527
Carnegie Mellon	\$115,207
Harvard	\$73,800
MIT	\$82,000
Duke	\$69,800

## CONCLUSIONS

Our study found that we could predict starting salary; however, it has a weak association with the demographic information. Our model suggests that GPA and work experience have the most robust linear relationship. There are other critical casual factors, which are not demographics that are not measured in this study contributing to starting salary. Our data used to estimate demographic effects is biased toward schools that report their performance. Those tend to be more expensive programs, which tend to admit younger, less experienced domestic candidates.

## ACKNOWLEDGEMENTS

We would like to thank Professor Matthew Lanham, Xinyu Wang, Theo Ginting, and our graduate student mentors student Hsin Yu Pan (Candice) and Oladimeji Adekoya (Dayo) for their guidance and support on this project.