

Figuring Out Which Patients to Outreach Back To



Stephen Soboslay

Choose **Ohio** First

INTRODUCTION

I'm Stephen Soboslay and I am working on the Business Intelligence team for Cleveland Clinic working as a Data Scientist Intern. This was a remote internship that has lasted from last May and is still ongoing. I got the internship from an email sent out to math majors at Cleveland State University.

OBJECTIVES

- The goal of this project was to figure out which patients had the better chance of converting a follow-up appointment if outreached. They may be outreached via a phone call or their MyChart.
- Another objective of this project was to learn Python as I had no prior experience with it.

METHODS

- The first thing we tried running was a logistic regression. But when we ran the logistic regression, we found only one of the variables from our dataset (If the patient used Medicaid MCP as their insurance). Other variables included in the dataset were race, patients residence of county, insurance company and what department the order came from.
- So we decided to switch to a subjective scoring method using three variables: the age of the order, the patients appointment completion rate, and the department in which the order came from.

PAT_ID	order_age_percentile	completion_percentile	department_percentile	harmonic
44010613213	41.79497339522246	26.721698027250852	0.5	1.4553575661332945
44010043574	34.31167213857127	26.730494426978808	0.5	1.4516913008479608
14021397431	51.55383222008378	88.8076740544605	46.45898828236639	57.49125579729712
1404808453	19.585644741310986	11.396197457307903	32.386396113175195	17.67997135619391
14021803752	23.964111853277487	29.768505287224922	7.936472417335727	14.901487783254883
1406819914	11.448545228121816	55.62931754404185	6.796188808448536	11.8828462151466
44010597177	38.06181365334541	48.91390397256359	0.500000000002733	1.4657619005783515
44011970871	31.518170496999886	64.01413207911432	29.608459559874248	36.980693204363234
1404632949	7.166308162572179	72.0234352672192	69.20262647841766	17.87027406231172
14021358868	11.196648930148314	47.33895199432762	0.5000009080554275	1.4215091964995379

Figure 1. Python output of a sample output of all the scores.

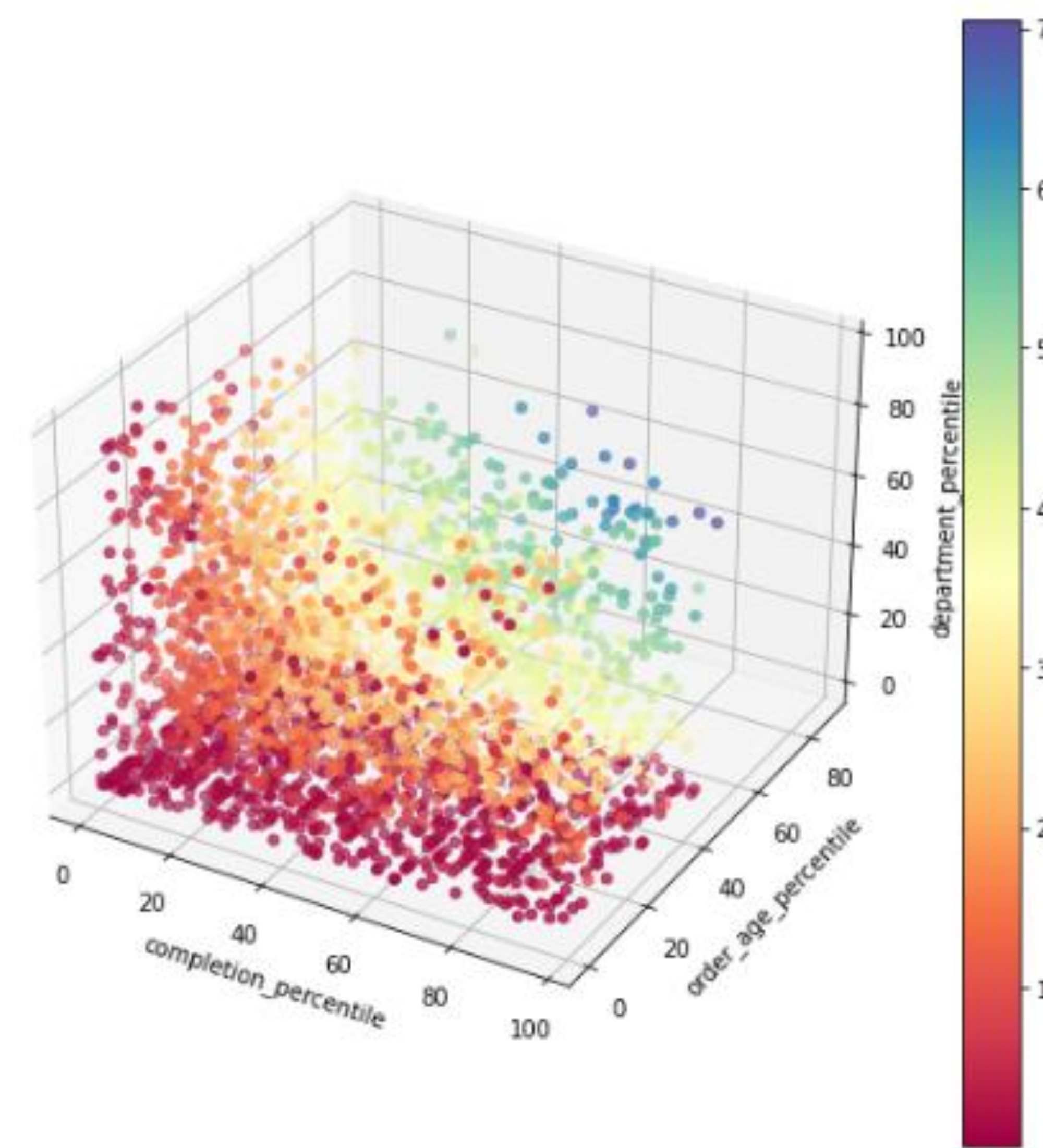


Figure 2. Python graph of the harmonic mean scores

RESULTS

- Age of Order:** If the patient scheduled a follow-up appointment: follow-up date minus order date. If the patient did not schedule a follow-up: Date the data was collected minus order date. From there we converted the dates into percentiles where if the age of the order was small, you were in a higher percentile, and if the order age was high, you were in a lower percentile
- Appointment Completion Rate:** Converted the percentages into percentile, higher percentage completed equals higher percentile. For patients who hadn't had appointments before or had low completion percentages, we used median blending to help. The median blending helped bring your score up.
- Order Department:** We took the appointment completion variable and converted it into a binary variable (1=yes, 0=no). From there we found the mean appointment completion rate for each department and converted it into a percentile where higher percentile=higher completion rate.
- We now have three different scores, to give the patient one final score, we take the harmonic mean of the three scores.

RESULTS (continued)

- Harmonic Mean divides the number of scores by the reciprocal of each number of scores.
- Ex: We have the numbers 4,7 and 9, then $3/((1/4)+(1/7)+(1/9))=5.953$

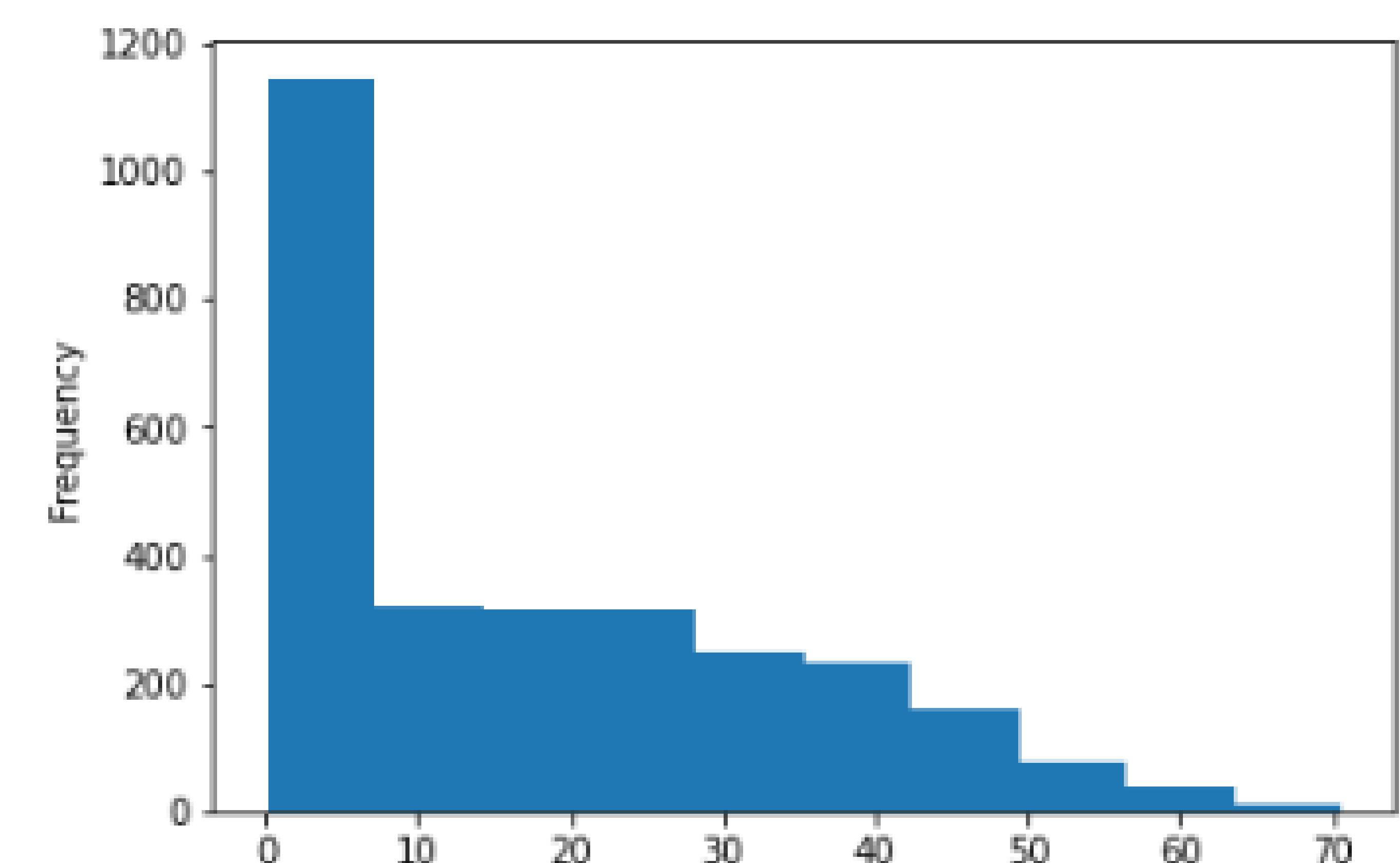


Figure 3. Histogram of Harmonic Mean Scores

FUTURE WORK and CONCLUSIONS

- If the patient had a higher score, that patient was more likely to be outreached.
- What we are currently doing now is trying to create a dashboard so it will be visible to multiple institutes throughout Cleveland Clinic so they can outreach to patients to schedule a follow up appointment.

Acknowledgments

I would like to thank everyone on the Business Intelligence team but especially Matt Steele, my mentor for this project and fellow intern Caitlin Gibson.