



Catholic Family Center

Team 11:

Luke Bamburoske	Tianbi Guo
Pramath Kashyap	Chenghao Li
Wenjie Li	Jonathan Teoh



Overview



Data Cleaning & Preprocessing

Goal Achievement Analysis

Model Analysis

► Data Cleaning & Preprocessing

- **Joined admission & discharge database:**
Include encounters with both admission and discharge time
- **Redundant features from admission & discharge:**
Step 1: Drop 100% duplicate columns;
Step 2: Consolidate similar columns but with contradictory info;
Step 3: Drop features with bare variations.
- **Feature conversion for modeling:**
Feature 1 -> Feature 1 (alter)
Feature 2 -> Feature 2 (alter), Feature 3
Feature 5-> Feature 5 (alter)
...

		Dependent Variable	Unique Identifier	Independent Variables				
		Come Back?	Cust	Encounter	Program	Demo Info	Other Info	Previous Programs
1	C00X	CXXX_XXX	CFC00X	None	
0	C00X	CXXX_XXX	CFC00X	CFC00X	
0	C00X	CXXX_XXX	CFC00X	None	
1	C00X	CXXX_XXX	CFC00X	None	
1	C00X	CXXX_XXX	CFC00X	CFC00X	
...	

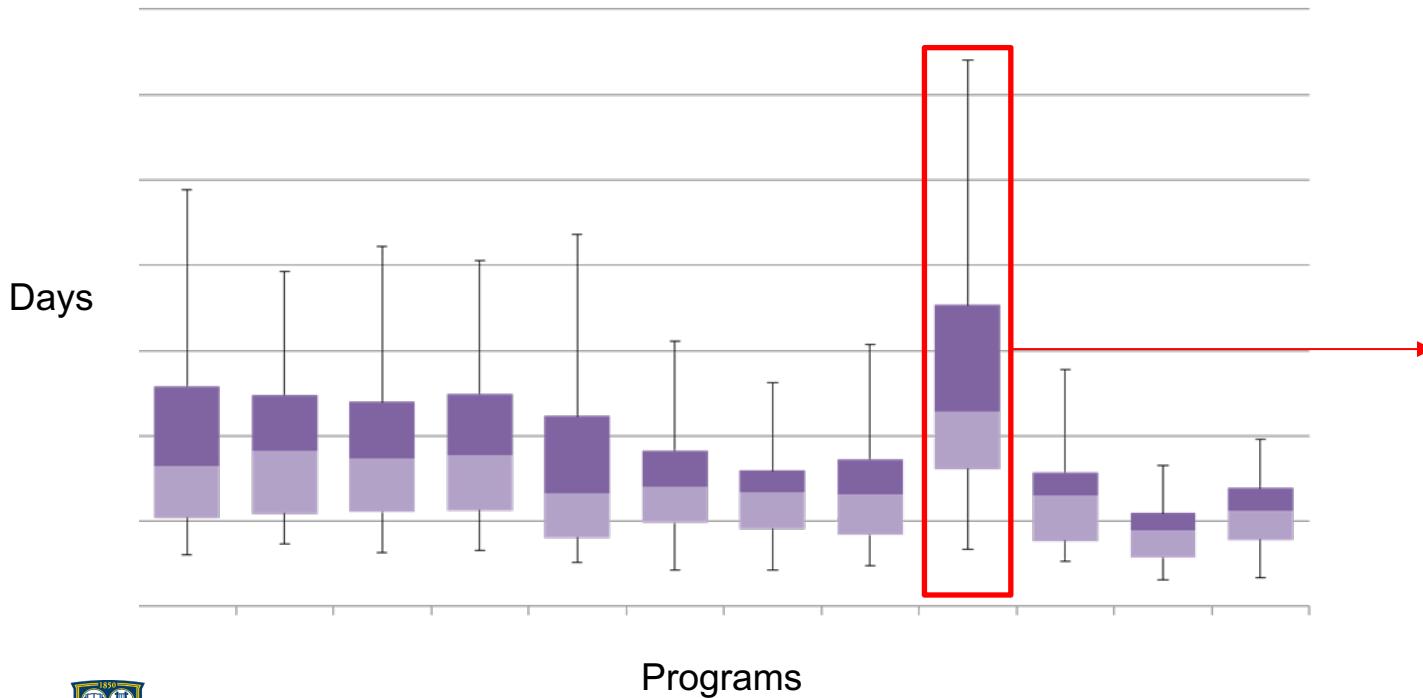


PART I: Goal Achievement Analysis





Program Stay Length

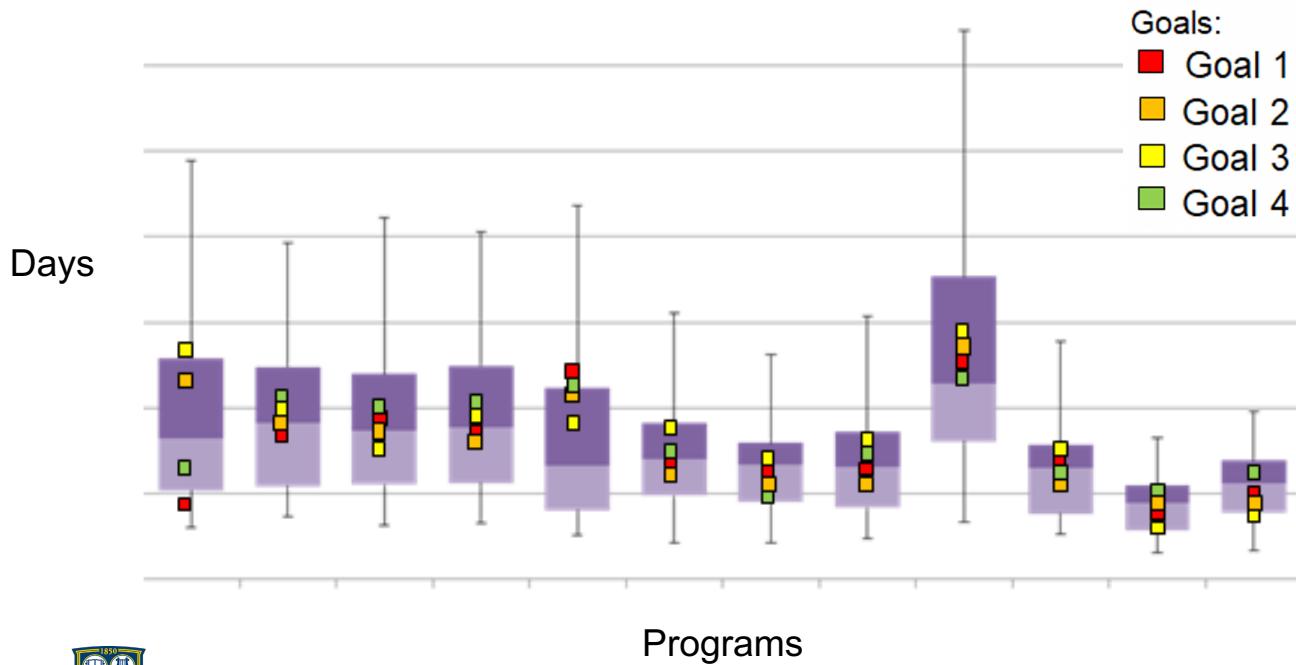


Patients from this program stay much longer and have obviously more outliers.





Program Stay Length & Goal Achievement



Takeaways:

- ❑ Patients achieving Goal 3 stay longer.
- ❑ Different programs focus on different goals.



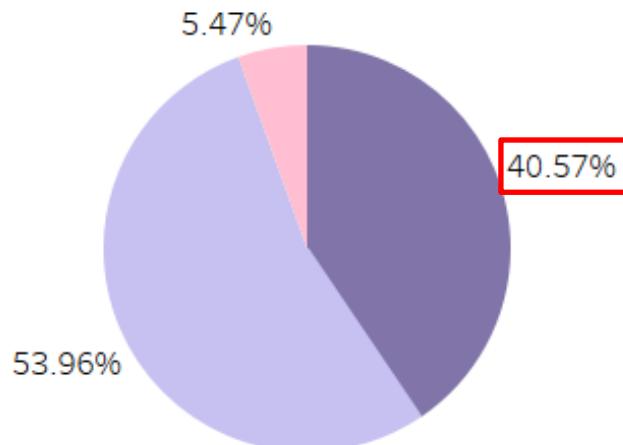
SIMON
BUSINESS SCHOOL

UNIVERSITY of ROCHESTER

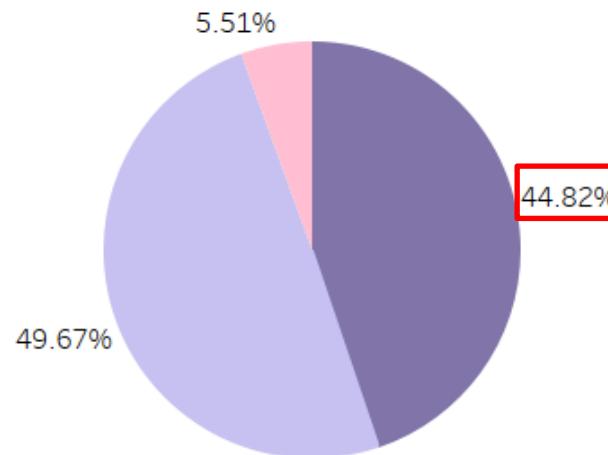


Goal Achievement: Substance X & Y

Substance X



Substance Y



Achieved
Partial Achievement
Not Achieved

Substance Y is more difficult to get rid of than Substance X.



Substance Z Goal Achievement & Previous Encounter

Returning clients who partially achieved Substance Z goals on their previous encounter had a lower Substance Z goal achievement rate again.

Previous Program

Achieved

Partially Achieved

No Previous Programs

Current Program
Goal Achievement
Rate



70%

24%*

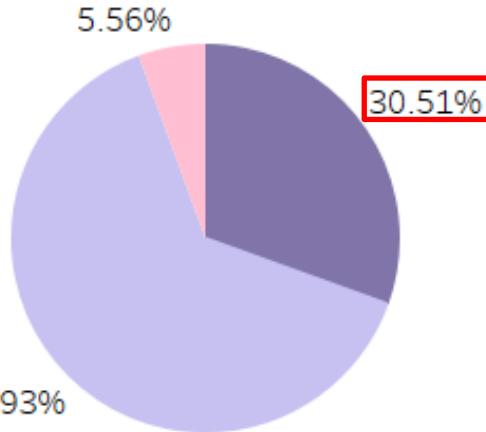
33%

Substance P Goal Achievement & Previous Program Type

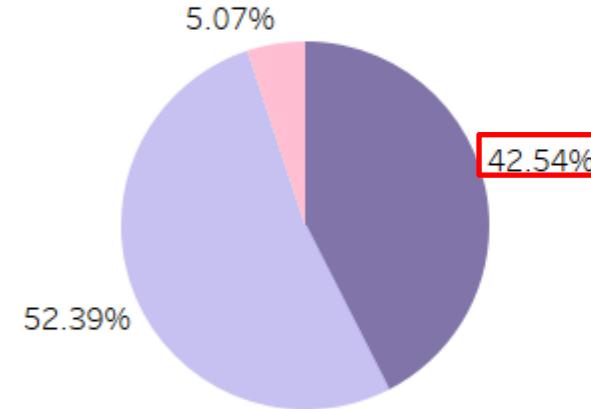
Clients who completed a Type 1 program had a higher Substance P goal achievement rate in an Type 2 program.

Type 2 Program Goal Completion

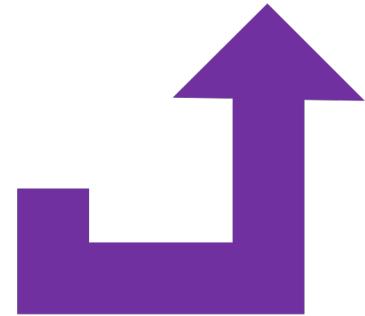
Did Not Complete
A Type 1 Program



Complete
A Type 1 Program



PART II: Modeling Analysis



Probability of Return: Methodology of Modeling

Model: Logistic Regression

- **Feature Selection**

Attempt 1: BIC (Bayes Information Criterion)

Attempt 2: MARS Model

Attempt 3: Business Intuition

Decision: Run selected features in regression;

Evaluate by coefficient > 0.02 (feature impact) and
p-value < 0.07 (coefficient reliability)

- **Model Performance**

Accuracy: 82.7%

- **Slice & Dice on Selected Features**

Representative feature visualizations in next slides.

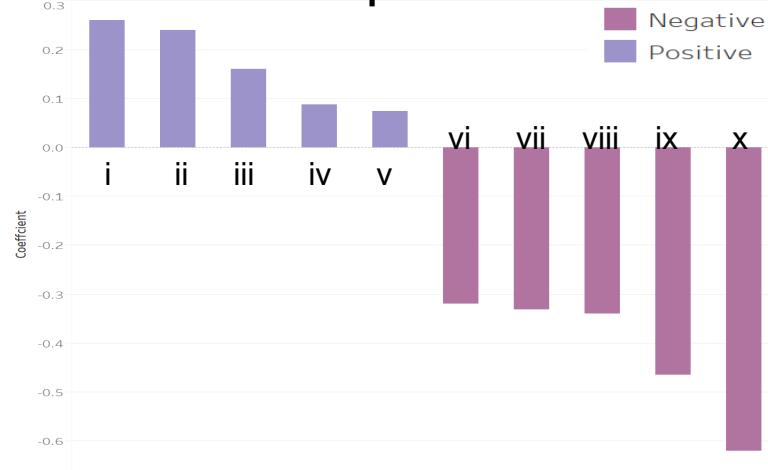


Model Demonstration

Come Back Or Not = **Coefficient 1 * Feature 1 + Coefficient 2 * Feature 2 + ... Coefficient N * Feature N**



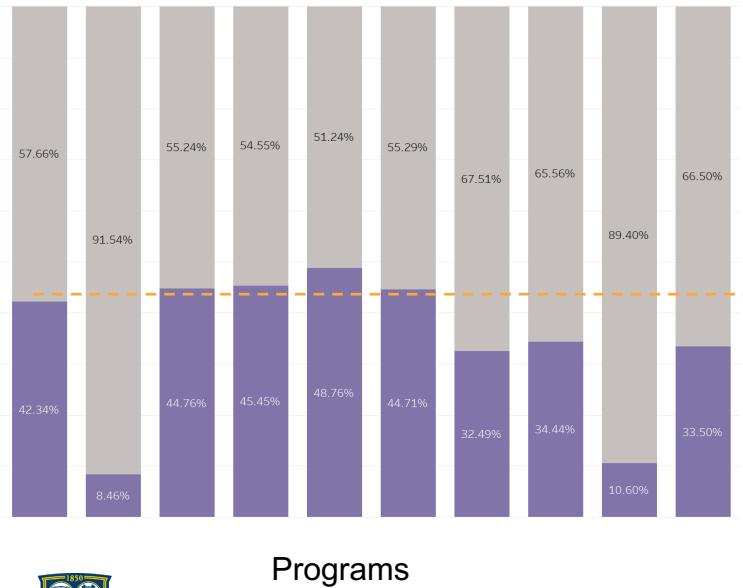
Top Features





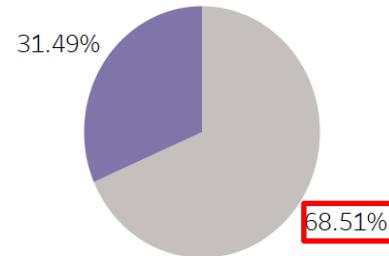
Probability of Return: Key Features of Program Attendance

Prob. of Return by Program Attendance

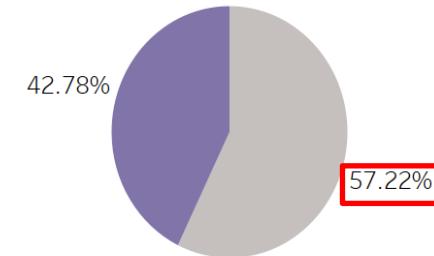


Previous Program Attendance
– Most Immediate Attendance

No Previous Programs



With Previous Programs



Summary: Review of Findings



Goal Achievement Analysis

- 1) Length of the Programs
- 2) Program Class

Probability of Return Analysis

- 1) Program Attendance
- 2) Case #24
- 3) Status
- 4) Omitted Variables





THANK YOU

Team 11:

Luke Bamburoske	Tianbi Guo
Pramath Kashyap	Chenghao Li
Wenjie Li	Jonathan Teoh

