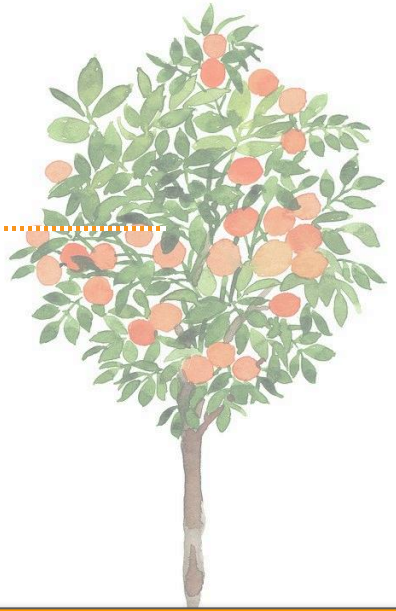
# Leveraging Slate to Calculate Demonstrated Interest and Likelihood to Apply and Enroll



*This session focuses on how Pitzer is leveraging Slate to identify behaviors and model interest and future actions*

# Pitzer College

- Small Liberal Arts College
- Claremont Colleges
- Located in Southern California
- ~1,000 undergraduate students
- On Slate since 2016

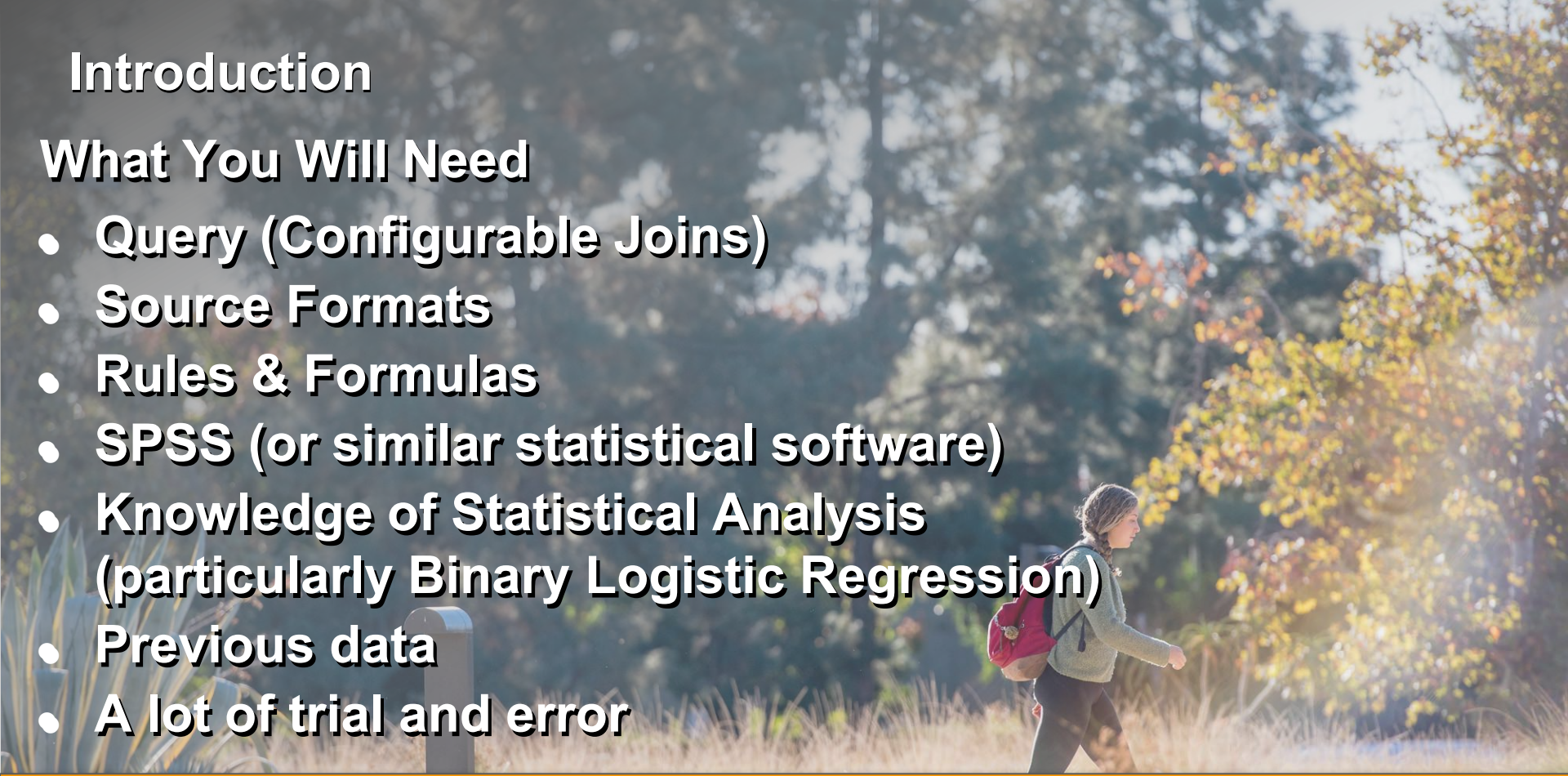


PITZER COLLEGE
OFFICE OF ADMISSION

# Introduction

## What You Will Need

- Query (Configurable Joins)
- Source Formats
- Rules & Formulas
- SPSS (or similar statistical software)
- Knowledge of Statistical Analysis (particularly Binary Logistic Regression)
- Previous data
- A lot of trial and error

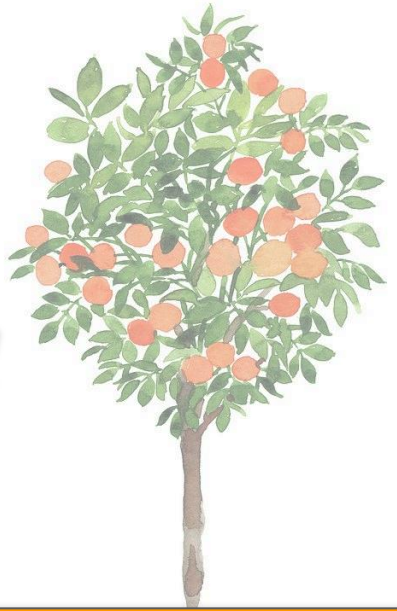


# Focus

1. Exploration process
2. Collaboration
3. Queries, Source Formats, and Rules used
4. Explain some of the statistical principles and processes used

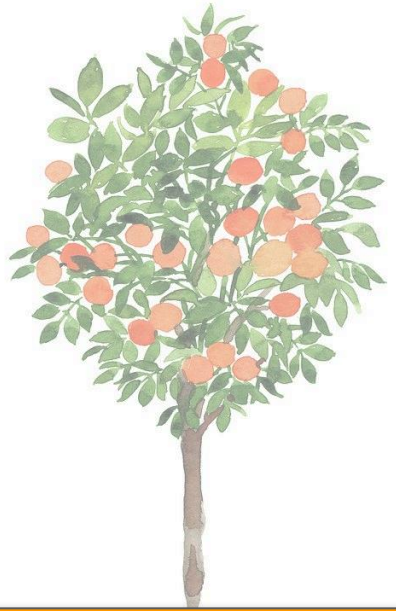
# Demonstrated Interest

*Pitzer-specific scores of behaviors already captured in Slate*
*Calculate automatically and early in the process*



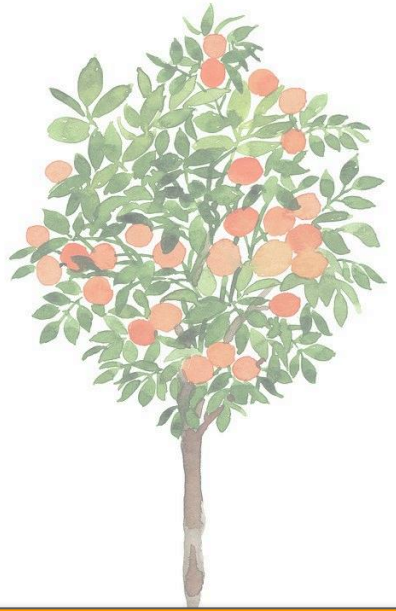
# Demonstrated Interest

- Slate has allowed us to track and catalogue Demonstrated Interest
- I wanted to be able to capture and calculate this early in the process
- To help inform:
  - Strategic Recruitment
  - Application review
  - Reader Training



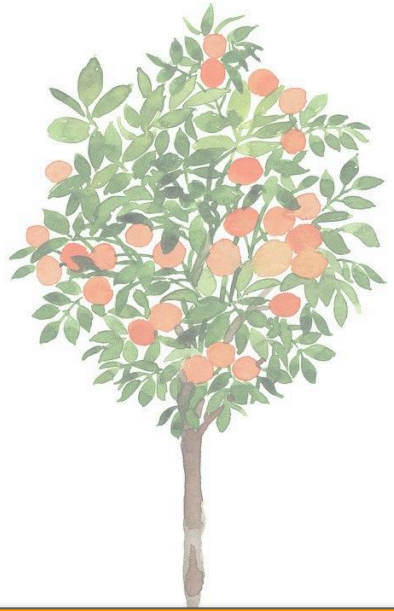
# Demonstrated Interest - Exploration

- I came up with an initial draft of behaviors and point scores
- Start the collaboration with my Admission Officers.
- Sent a Slate Form to collect their thoughts
- They suggested new behaviors and scoring
- I rebuilt my Query and scoring rubric
- I rebalanced and validated the score to the internal score weighting

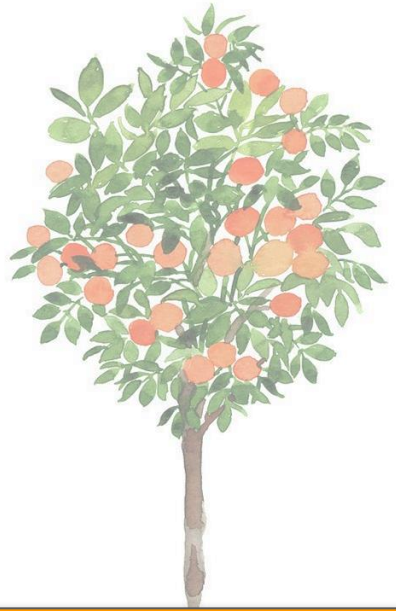


# Demonstrated Interest - Exploration

- Our Admission Officers have been looking at Demonstrated Interest in the Application Review Process
- I came up with an initial draft of behaviors and point scores to start the collaboration. Sent a Slate Form to collect their thoughts
- They suggested new behaviors and scoring
- I rebuilt my Query and scoring rubric
- I rebalanced and validated the score to the internal score weighting



# Demonstrated Interest - Exploration



## Demonstrated Interest

What does this mean to you?
Someone who is very engaged and taking actions that show interest?
Am I missing an element to this? Is there a 'good DI' that you want to see versus 'bad DI'

What actions should be added to this score? (They opened an email, they came for a campus tour, etc.?) How important are they?

How high should the scale go?
- 100-pt scale (more standardized, but would require re-balancing values of behaviors if we change what we include)
- Open-Ended with no "max score" (easier to build, but hard to see what is a "high" score by just looking at a score)
- Other

How high should the scale go? - Other

How heavily should the following behaviors be weighted?
I will treat your number like a % out of 100 (e.g. In-Person: 80%, Emails Opened: 20% = 100%)
Please put a '0' if it should not be in the score
You can view my current scoring here - https://connect.pitzer.edu/manage/query/build?id=36169d87-dfda-4903-93a8-0dd182dcd659

**Multiple Opportunities**

In-Person Event Max Points

- In-Person Event: How many points per event?




Hi all,

A quick update. I have started work on the Demonstrated Interest Score
You can view the work in progress here (but please do not edit) -

I have added questions to my **questionnaire** including how much "weight" the behaviors should have and how high the score should go

---

To continue our discussion about a Demonstrated Interest Score (aka Engagement Score/Affinity Score)

My thought is 3 new fields in Slate:

- Demonstrated Interest
  - Scored based on values per action
- Likelihood to Apply
  - Evaluated with Regression
- Likelihood to Commit
  - Evaluated with Regression

Are there any particular variables you would like added to the score? They opened an email, they came for a campus tour, etc.? Please complete **this survey** to provide your thoughts. I will follow up in late July with additional details and questions once I chat with some other institutions.

Thank you so much!

**David Dysart**
Associate Director
Enrollment Systems
Pitzer College | Office of Admission

# Demonstrated Interest - Exploration

- 12 Behaviors
- 3 Bonuses
  - to replace 3 Behaviors that some records may not perform
  - Bonuses take Total and reduce it to appropriate “weight”
- 100 pt Score for context

| Factor | Max | Formula |
|---|---|---|
| [illegible] | 24 | 6 points per |
| [illegible] | 20 | 5 points per |
| [illegible] | 10 | 5 points per |
| [illegible] | 3 | 1 points per |
| [illegible] | 8 | 4 points per |
| [illegible] | 3 | 3 points yes/no |
| [illegible] | 7 | 1 point per |
| [illegible] | 5 | 1 point per |
| [illegible] | 3 | 2 points yes/no |
| [illegible] | 3 | 2 points yes/no |
| [illegible] | 10 | Rating: 1 = 10 pts, 2 = 8, etc. |
| [illegible] | 4 | 4 points yes/no |
| [illegible] | 10 | If no sum of other factors divided by 9 |
| [illegible] | 4 | If no sum of other factors divided by 24 |
| [illegible] | 8 | If no, 4/8 points if |

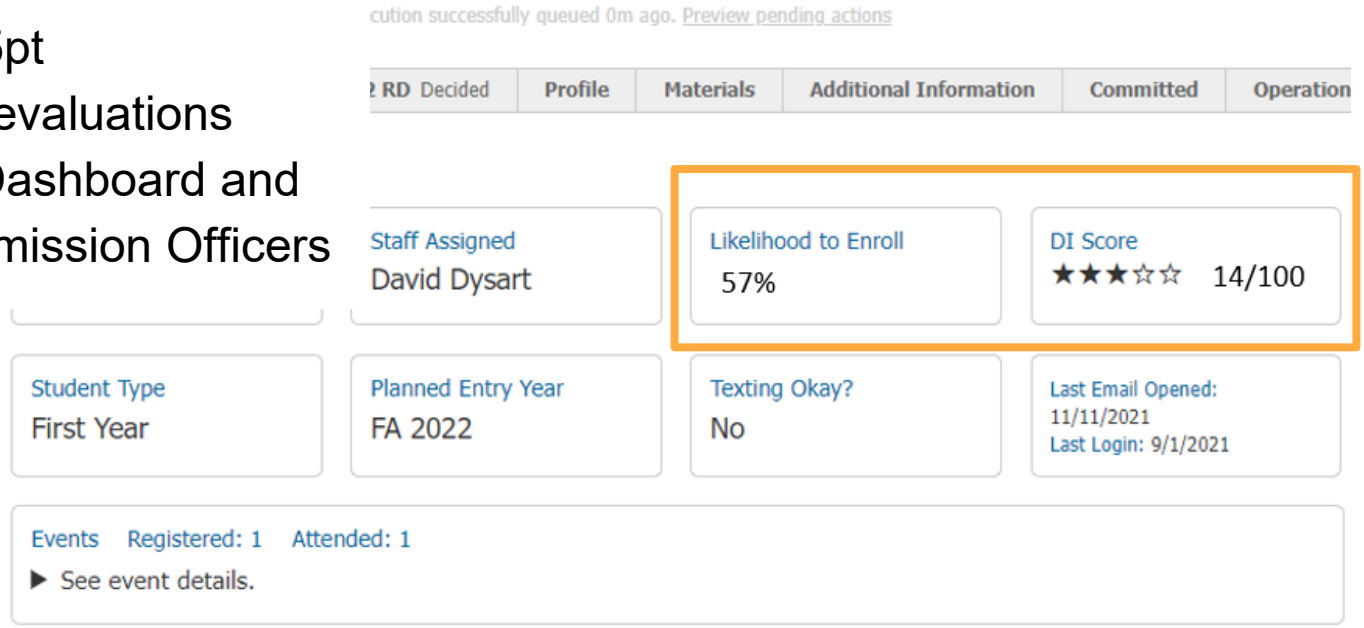
# Demonstrated Interest – Slate Translation

- Average DI will increase through cycle
- It would also be hard to score 100

| Min Value | Max Value | Export Value |
|---|---|---|
| New Code | | |
| 0 | 9.5 | 1 |
| 09.501 | 13.199 | 2 |
| 13.2 | 21.799 | 3 |
| 21.8 | 35.4 | 4 |
| 35.401 | 120 | 5 |

# Demonstrated Interest – Slate Translation

- To further contextualize –
- Used Translations
- Convert 100pt to 5pt
- Reflects our Staff evaluations
- Built into Person Dashboard and Report for Admission Officers



# Query

- Person Reference ID
- DI Score Upload
- Application
- Variable 1 Score
- Variable 2 Score
- Variable 3 Score
- DELETED
- Variable 4 Score
- Variable 5 Score
- Variable 6 Flag
- Variable 7 Score
- Variable 8 Score
- Variable 9 Score Bonus
- DELETED
- DELETED
- Variable 10 Score
- Variable 11 Score
- Variable 12 Score
- Variable 13 Score

Active

In-Person Event Score

Subquery Export

Dependent subquery

Formula

1 Row Limit

```
Case when @In-Person-Event-Count > 'maximum event' then 'maximum points' when @In-Person-Event-Count > '0' then (@In-Person-Event-Count * 'weight')
```

Export

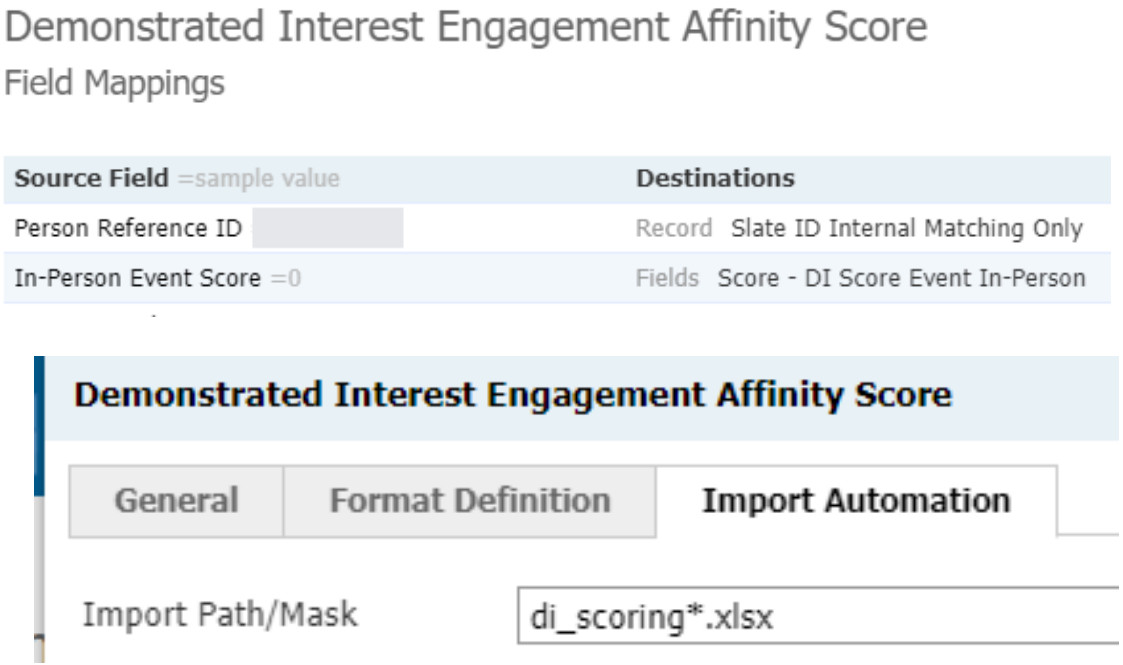
In-Person Event Count

**Schedule Export**

| | |
|---|---|
| Status | Active |
| Destination | Technolutions SFTP |
| Path | sftp://ft.technolutions.net/outgoing/ ../incoming/di_scoring Vars |
| Encryption | Secure Transfer |
| Notification | Failures and late deliveries |
| Notify Email | david_dysart@pitzer.edu |
| Format | Excel Spreadsheet |
| Suppress Empty | Allow empty files |
| Requested Delivery Window (Eastern Time) | ☑ Overnight: 2:00am–4:00am (preferred window)<br>☐ Morning: 10:00am–12:00pm<br>☐ Midday: 12:00pm–2:00pm<br>☐ Afternoon: 2:00pm–4:00pm<br>☐ Evening: 8:00pm–10:00pm |
| Requested Weekdays | ☐ Monday<br>☐ Tuesday<br>☐ Wednesday<br>☐ Thursday<br>☐ Friday<br>☐ Saturday<br>☑ Sunday |

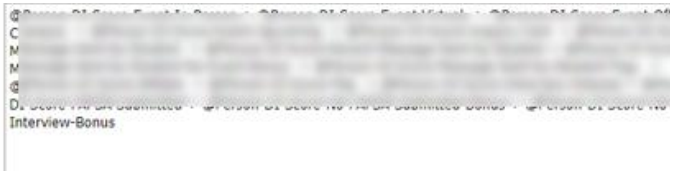
# Uploading the Data and Calculating Scores

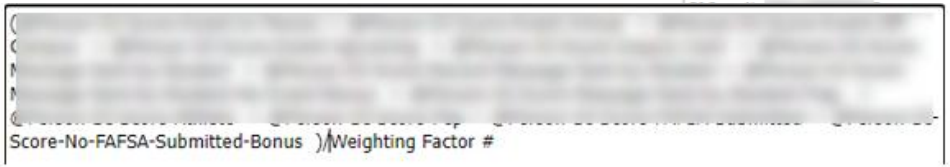
- Create fields for each Sub Score
- Build Source Formats to process the scheduled exports of SFTP server
- Files run over the weekend during slow times

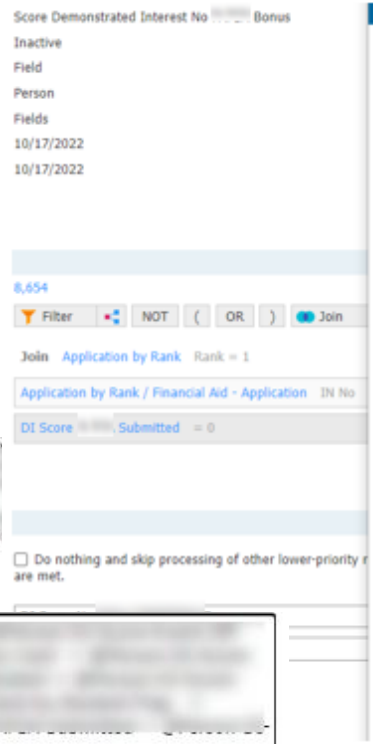


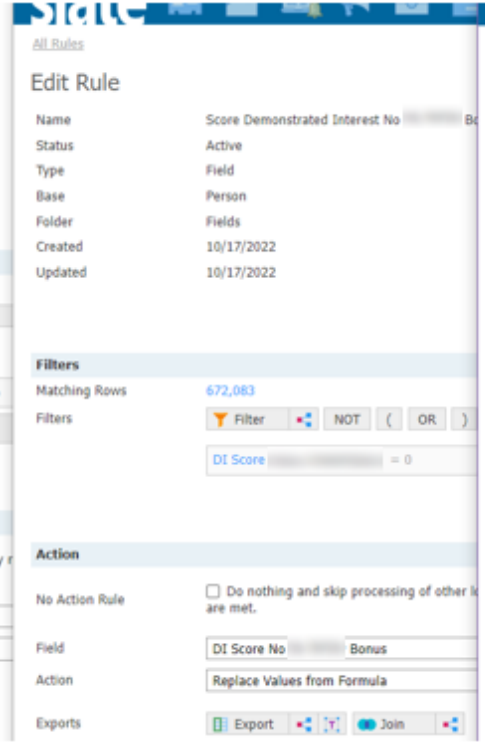
# Uploading the Data and Calculating Scores

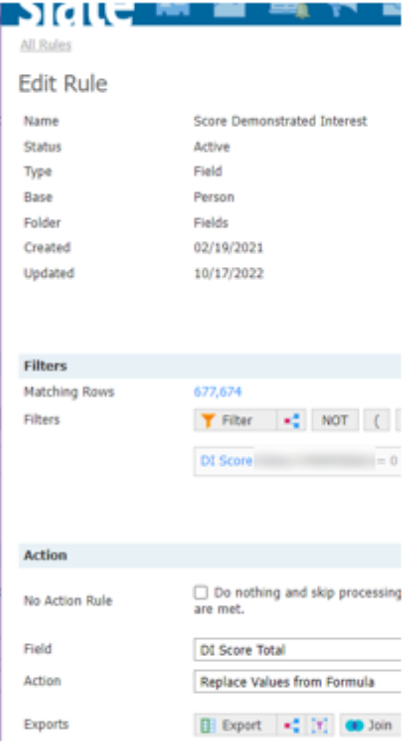
- 3 Rules create the score
  - 2 Bonuses
  - 1 Final Score rule
- 10-pt Bonus example
  - Sum other behaviors
  - Divide by 9
  - Add to Total



# Visualizing the Validation

| Variable 3 | | | | |
|---|---|---|---|---|
| Possible Scores | Average | | Weight | |
| | Manual Evaluation | Calculated Evaluation | Manual | Calc |
| 0 | 1.7 | 9.3 | - | - |
| 5 | 3.2 | 22.8 | 0.6 | 4.6 |
| 10 | 4.0 | 43.9 | 0.4 | 4.4 |

| Variable 7 | | | | |
|---|---|---|---|---|
| Possible Scores | Average | | Weight | |
| | Manual Evaluation | Calculated Evaluation | Manual | Calc |
| 0 | 1.9 | 8.0 | - | - |
| 1 | 2.3 | 16.3 | 2.3 | 16.3 |
| 2 | 2.5 | 20.7 | 1.3 | 10.4 |
| 3 | 2.6 | 23.2 | 0.9 | 7.7 |
| 4 | 2.7 | 26.9 | 0.7 | 6.7 |
| 5 | 2.8 | 35.1 | 0.6 | 7.0 |

## Variable 3

- 3 possible scores: 0, 5, 10
- Evaluations correlate with variable
- Weight: .5 Manual, 4.5 Calculated

## Variable 7

- 1 Point per behavior
- Evaluations correlate with variable
- Weight: .9 Manual, 9 Calculated but more variance

| Manual Evaluation | Calculated Evaluation | Weight |
|---|---|---|
| 1 | 5.4 | 5.4 |
| 2 | 8.9 | 4.5 |
| 3 | 17.3 | 5.8 |
| 4 | 26.0 | 6.5 |
| 5 | 44.7 | 8.9 |
| **Average** | **10.5** | **6.2** |

# Demonstrated Interest - Results

- Promising Results
  - Manual and Calculated Scores highly correlated
  - Both on data used to build score and to predict FA 23 scores
- Next Steps
  - Incorporate FA 23 data and emerging ways to demonstrate interest
  - Finetune scoring
  - Hold Feedback Sessions with Admission Officers
  - Create similar scores to aid in holistic admission

| Manual | Calculated Score | | |
|---|---|---|---|
| Score | FA 23 | FA 22 | FA 21 |
| 1 | 28.9 | 47.3 | 49.5 |
| 2 | 22.6 | 31.2 | 31.8 |
| 3 | 16.4 | 20.2 | 20.4 |
| 4 | 9.5 | 11.6 | 11.2 |
| 5 | 7 | 9.3 | 8.3 |

# Dashboard (WIP)

- X
- X
  - x

Dysart, Dahveed  Rule execution successfully queued 0m ago. Preview pending actions

Dashboard | Timeline | 2022 RD Decided | Profile | Materials | Additional Information | Committed | Operation

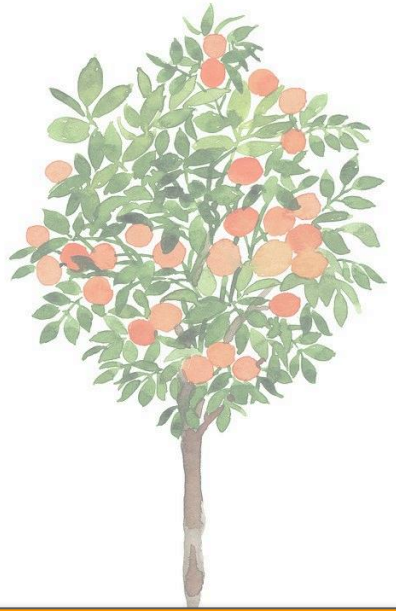
Dashboard

| Chosen Name | Staff Assigned | Likelihood to Enroll | DI Score |
|---|---|---|---|
| Dahveed | David Dysart | 57% | ★★★☆☆ 14/100 |

| Student Type | Planned Entry Year | Texting Okay? | Last Email Opened: 11/11/2021 Last Login: 9/1/2021 |
|---|---|---|---|
| First Year | FA 2022 | No | |

Events Registered: 1 Attended: 1

▶ See event details.

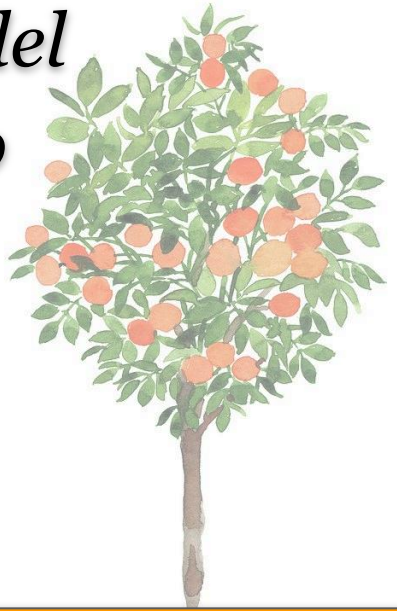


# Likelihood to Apply

*Use Slate data in a statistical model*

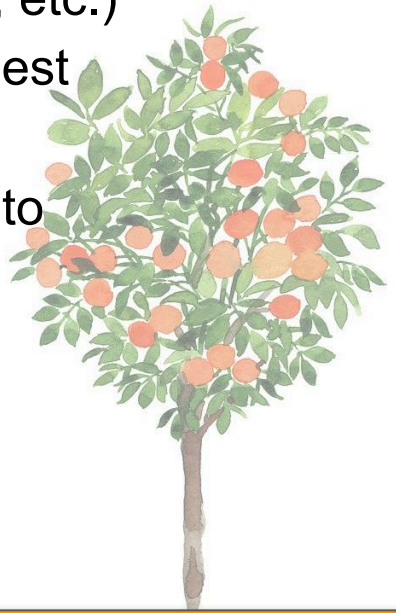
*Predict how likely an Inquiry is to apply*

*Automatically apply model to all Slate records*



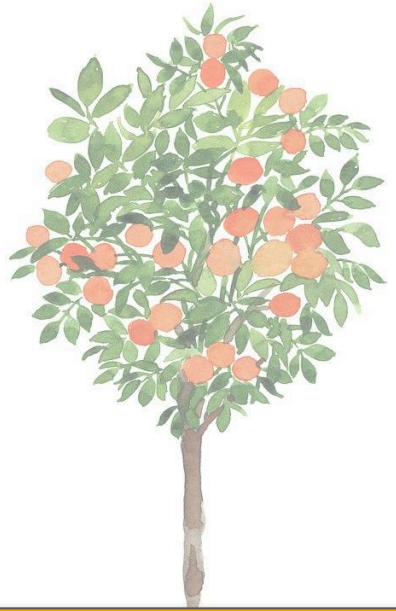
# Likelihood to Apply

- Next Step in this was to build a statistical model to predict Application Likelihood
- Turned to SPSS – Statistical Analysis Software (STATA, R, etc.)
- Exploration, Trial, and Error to find the best variables and best coding
- Run a Binary Logistic Regression. Create Formula is Slate to transform results into a % likelihood to apply
- Build the Queries and Source Formats
- Create new Fields to capture the snapshot data
- Viola



# Likelihood to Apply

- There are several uses you can leverage this for –
- Planning your cycle
- Application evaluation time/staffing,
- Strategic recruitment
- Interventions and application pushes



# Visualizing the Validation

| A | C | E | G |
|---|---|---|---|
| Variable 1 Count | | Variable Flag | |
| Values | % Apply | Values | % Apply |
| 0 | 6% | 0 | 6% |
| 1 | 25% | 1 | 30% |
| 2 | 29% | | |
| 3 | 65% | | |
| 4 | 68% | | |
| 5 | 81% | | |

**Count:** The number of behaviors
**Flag:** If they have the behavior at any level
**% Apply:** What % of the FA 20-22 Population Applied

| AG | AI | AK | AM |
|---|---|---|---|
| Variable 3 Count | | Variable 3 Flag | |
| Values | % Apply | Values | % Apply |
| 0 | 6% | 0 | 6% |
| 1 | 57% | 1 | 63% |
| 2 | 91% | | |
| 3 | 97% | | |
| 4 | 100% | | |
| 5 | 100% | | |
| 6 | 100% | | |
| 7 | - | | |
| 8 | 100% | | |
| 9 | 100% | | |

## Variable 3

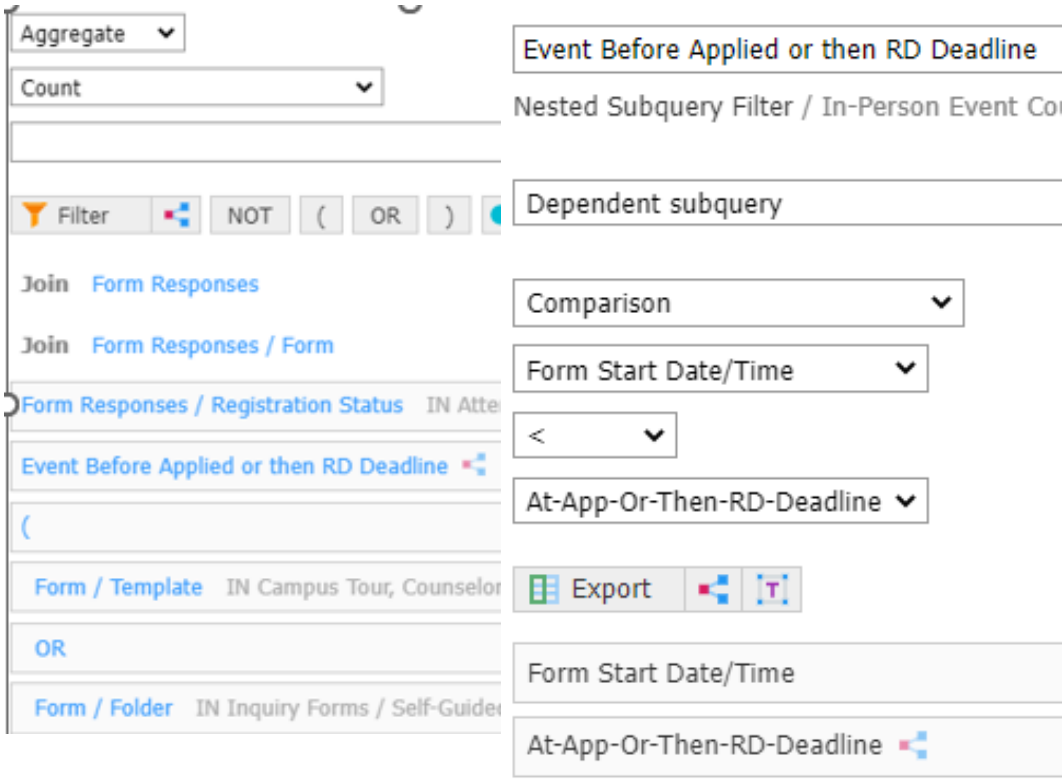
- Linear relationship between the Count and % Applying.
- Very inconsistent. Flag a better fit than the Counts

## Variable 1

- Linear relationship between the Count and % Applying.
- Not very consistent, but a better fit than the Flag

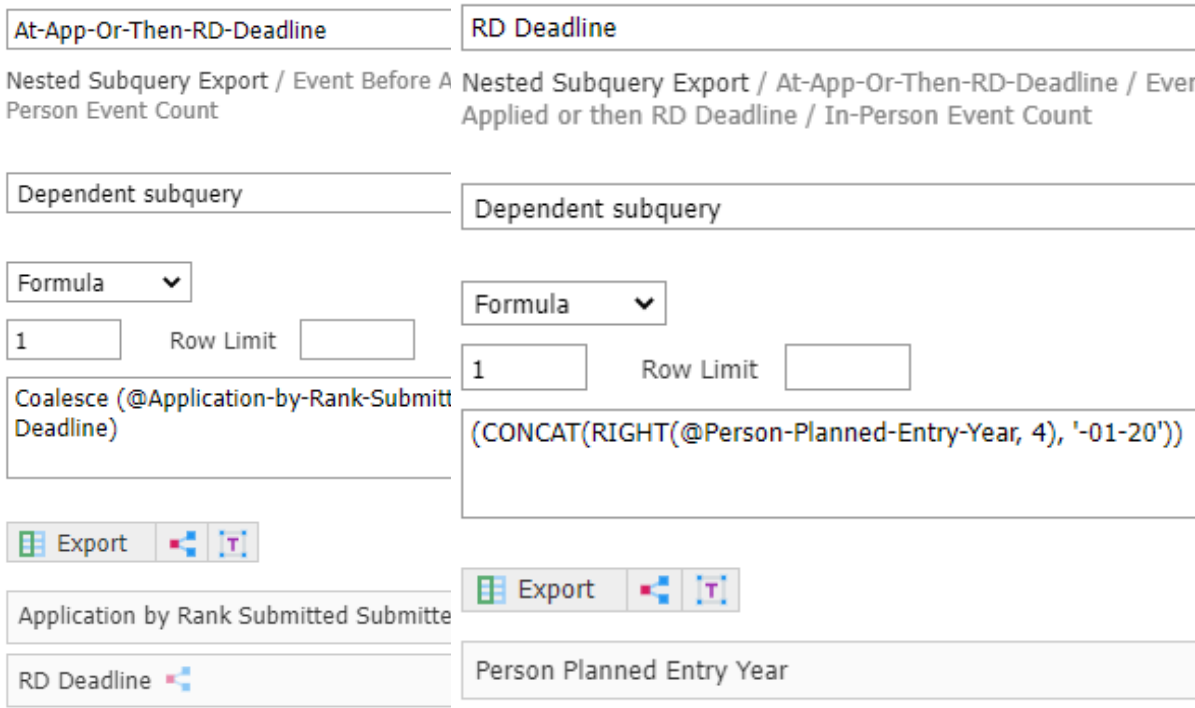
# Getting Snapshot Data

- Configurable Joins
- Decide on Counts versus Existence (Yes/No) Flags
- IMPORTANT – Only pull behaviors before they apply



# Getting Snapshot Data

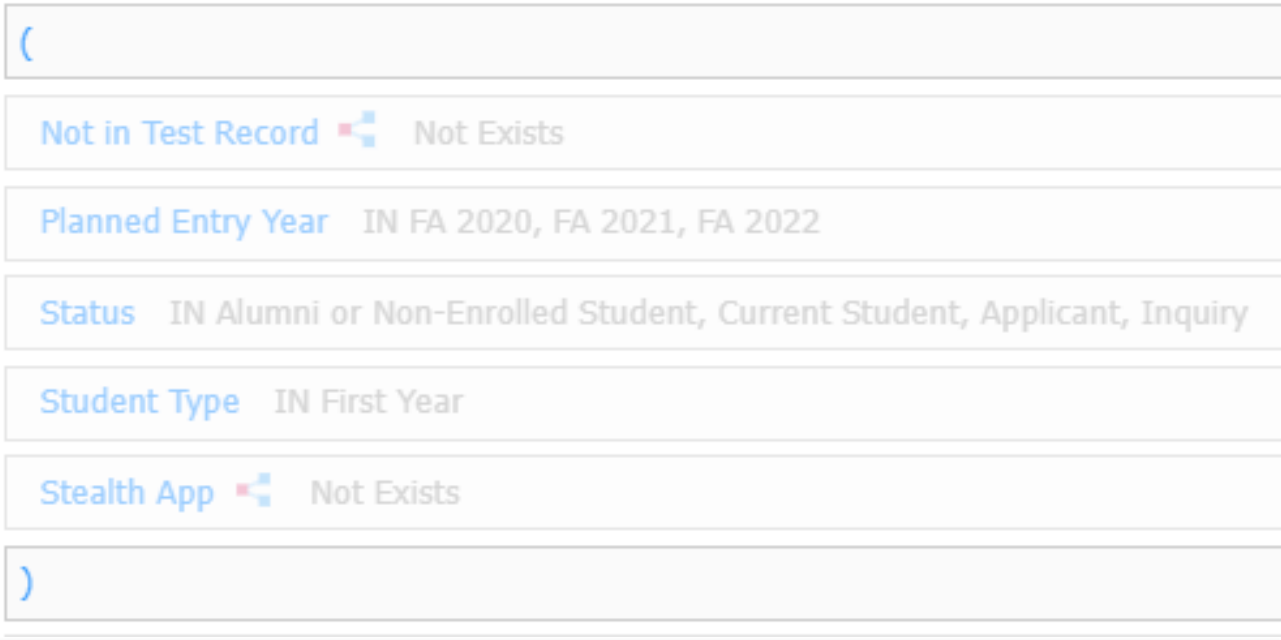
- Use Comparison subquery filters
- Compare behaviors date to Date Applied/Application Date for non-applies
- Pulling data for multiple historical years is a more complicated formula
- Watch for re-applies and other fallouts



# Data Points

- Removed Prospects
- Removed Stealth Applicants
- Missing Data

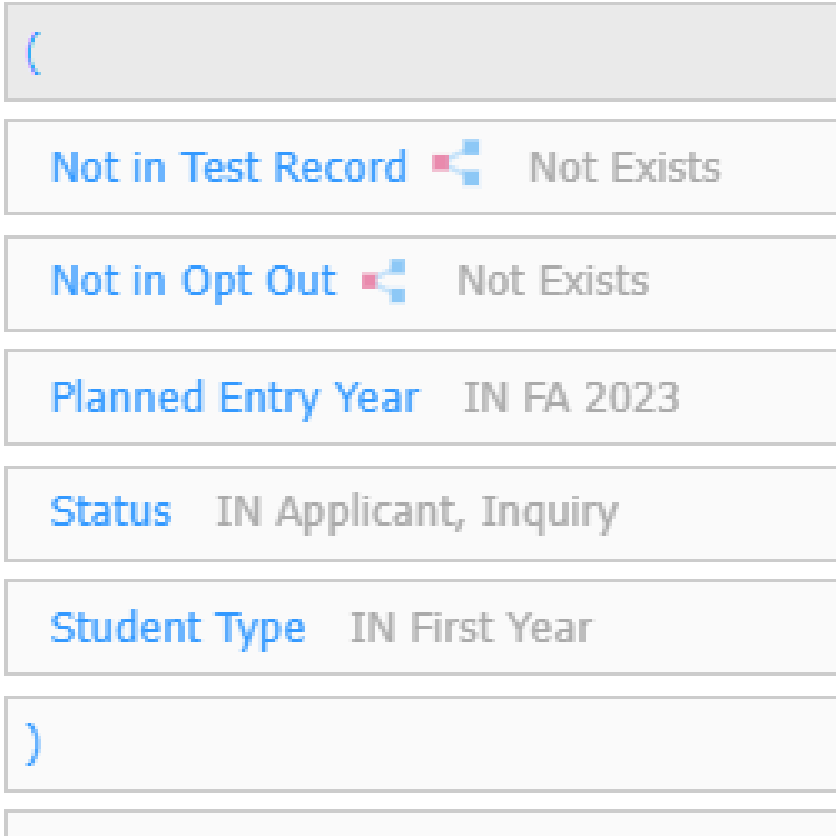
Model



# Data Points

- Removed Prospects
- Removed Stealth Applicants
- Missing Data

Predictive

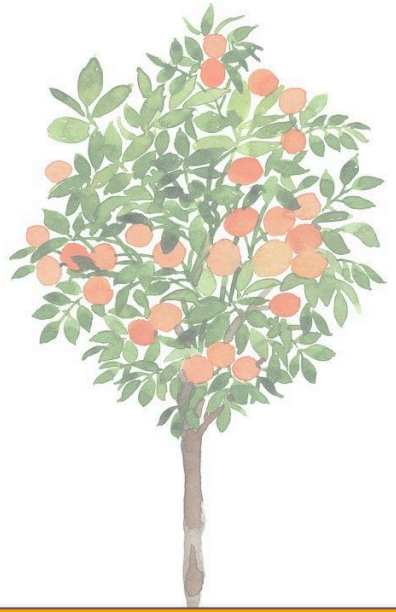


# Checking Results

- I had 3 promising models
  1. Counts with Outliers
  2. Counts excluding Outliers
  3. Flags

| AI | AK | AL | AM | AN |
|---|---|---|---|---|
| Variable 2 Flag | | | | |
| Values | % Apply | Pre - Count | Pre - Count | Predicted Ave Flags |
| 0 | 6% | 14% | 14% | 14% |
| 1 | 48% | 52% | 68% | 55% |

| AC | AE | AF | AG | AH |
|---|---|---|---|---|
| Variable 2 Count | | | | |
| Values | % Apply | Pre - Count | Pre - Count | Predicted Ave Flags |
| 0 | 6% | 14% | 14% | 14% |
| 1 | 43% | 44% | 58% | 51% |
| 2 | 47% | 52% | 73% | 52% |
| 3 | 63% | 77% | 94% | 71% |
| 4 | 79% | 86% | 98% | 78% |
| 5 | 83% | 90% | 99% | 79% |
| 6 | 80% | 91% | 99% | 78% |
| 7 | 50% | 98% | 100% | 89% |
| 8 | 100% | 100% | 100% | 97% |
| 9 | 100% | 100% | 100% | 96% |
| 10 | 100% | 99% | 100% | 86% |
| 11 | 100% | 100% | 100% | 98% |



# Likelihood to Apply

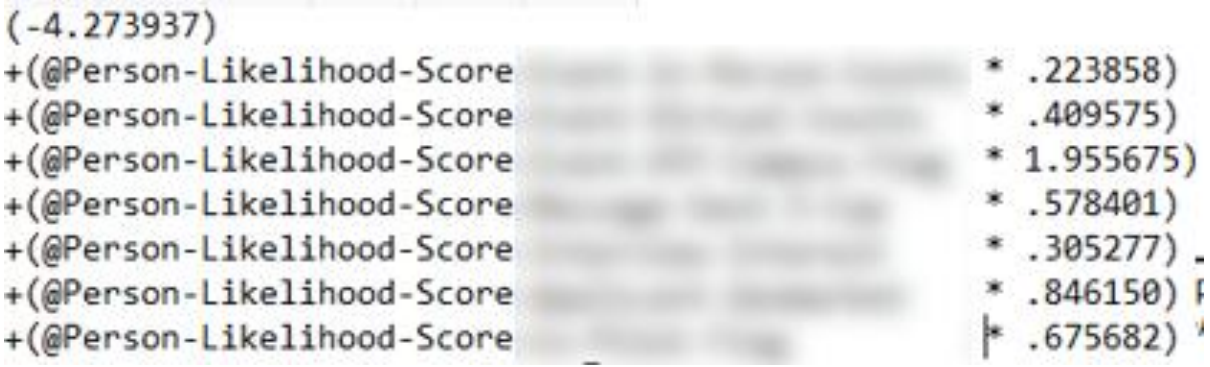
- Use Unstandardized Regression Weight (B) in your Slate Query

**Variables in the Equation**

| | | B | S.E. | Wald | df | Sig. | Exp(B) | 95% C.I.for EXP(B) Lower | Upper |
|---|---|---|---|---|---|---|---|---|---|
| Step 1[a] | [redacted] | .224 | .022 | 100.675 | 1 | <.001 | 1.251 | 1.197 | 1.307 |
| | [redacted] | .410 | .029 | 192.932 | 1 | <.001 | 1.506 | 1.422 | 1.596 |
| | [redacted] | 1.956 | .048 | 1670.357 | 1 | .000 | 7.069 | 6.436 | 7.764 |
| | [redacted] | .578 | .028 | 414.139 | 1 | <.001 | 1.783 | 1.687 | 1.885 |
| | [redacted] | .305 | .018 | 289.232 | 1 | <.001 | 1.357 | 1.310 | 1.406 |
| | [redacted] | .846 | .021 | 1633.828 | 1 | .000 | 2.331 | 2.237 | 2.428 |
| | [redacted] | .676 | .036 | 350.986 | 1 | <.001 | 1.965 | 1.831 | 2.109 |
| | [redacted] | -4.274 | .060 | 5011.029 | 1 | .000 | .014 | | |

# Likelihood to Apply

- Export Formula: Likelihood to Apply Logit




- Export Formula: Likelihood to Apply % Score

```
(POWER(2.718281828,@Person-Likelyhood-to-Apply-Logit))/
(1+POWER(2.718281828,@Person-Likelyhood-to-Apply-Logit))
```

# Likelihood to Apply - Results

- For Fall 2023
- Removing Stealth apps, Prospect to App, no Score
  - Predicted 1179 apps
  - Received 1351apps
- Underpredicted low behavior records
  - <10% likely to apply
  - 90% of all records
- If looking at just Likelihood >=10%, overpredicted
  - Predicted 803 apps
  - Received 669 apps

# Likelihood to Apply

**Variables in the Equation**

| | | B | S.E. | Wald | df | Sig. | Exp(B) | 95% C.I.for EXP(B) Lower | Upper |
|---|---|---|---|---|---|---|---|---|---|
| Step 1[a] | [redacted] | .224 | .022 | 100.675 | 1 | <.001 | 1.251 | 1.197 | 1.307 |
| | [redacted] | .410 | .029 | 192.932 | 1 | <.001 | 1.506 | 1.422 | 1.596 |
| | [redacted] | 1.956 | .048 | 1670.357 | 1 | .000 | 7.069 | 6.436 | 7.764 |
| | [redacted] | .578 | .028 | 414.139 | 1 | <.001 | 1.783 | 1.687 | 1.885 |
| | [redacted] | .305 | .018 | 289.232 | 1 | <.001 | 1.357 | 1.310 | 1.406 |
| | [redacted] | .846 | .021 | 1633.828 | 1 | .000 | 2.331 | 2.237 | 2.428 |
| | [redacted] | .676 | .036 | 350.986 | 1 | <.001 | 1.965 | 1.831 | 2.109 |
| | [redacted] | -4.274 | .060 | 5011.029 | 1 | .000 | .014 | | |

- Likelihood to Apply Logit

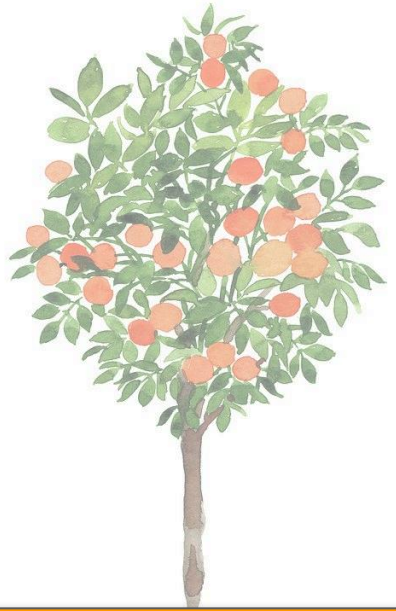
```
(-4.273937)
+(@Person-Likelihood-Score [redacted] * .223858)
+(@Person-Likelihood-Score [redacted] * .409575)
+(@Person-Likelihood-Score [redacted] * 1.955675)
+(@Person-Likelihood-Score [redacted] * .578401)
+(@Person-Likelihood-Score [redacted] * .305277)
+(@Person-Likelihood-Score [redacted] * .846150)
+(@Person-Likelihood-Score [redacted] * .675682)
```

- Likelihood to Apply Score

```
(POWER(2.718281828,@Person-Likelyhood-to-Apply-Logit))/
(1+POWER(2.718281828,@Person-Likelyhood-to-Apply-Logit))
```

# Likelihood to Apply

- There are several uses you can leverage this for –
- Planning your cycle
- Application evaluation time/staffing,
- Strategic recruitment
- Interventions and application pushes

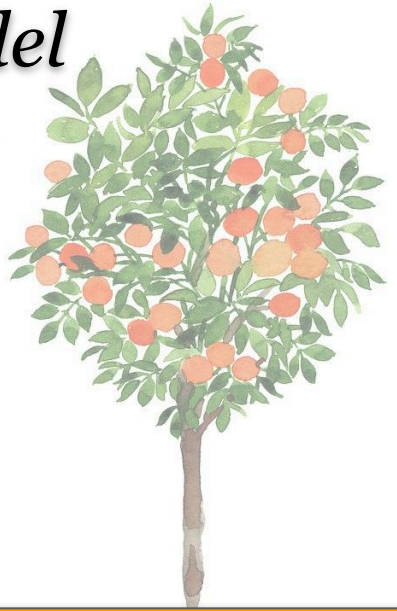


# Likelihood to Enroll

*Use Slate data in a statistical model*

*Predict how likely an Applicant is to enroll*

*Automatically apply model to all Slate records*

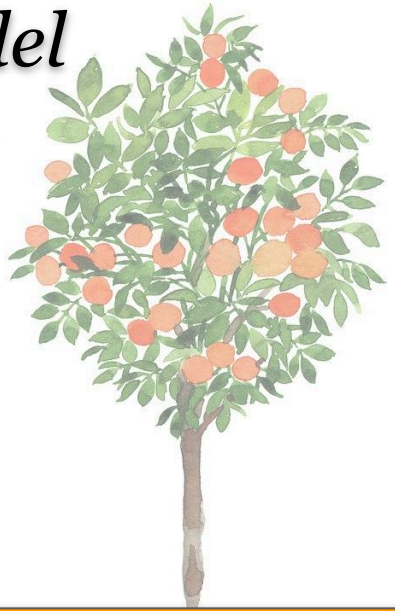


# Likelihood to Enroll (with 2nd Models)

*Use Slate data in a statistical model*

*Predict how likely an Applicant is to enroll*

*Automatically apply model to all Slate records*



# Likelihood to Enroll

- Changing Base – Person to Applications
- Model based on Admitted Regular Decision Applicants
- Snapshot of behaviors at time of Admit Decision

Model




Predictive

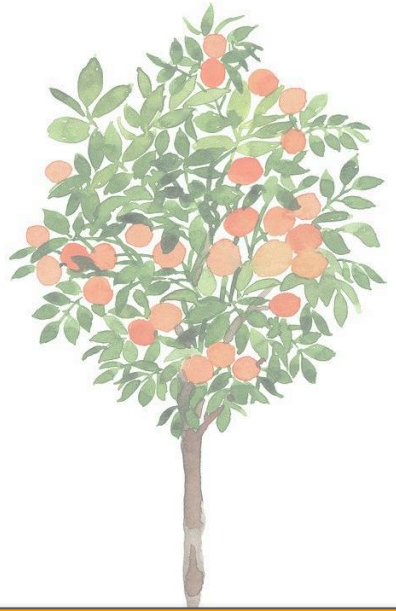


# Checking Results

- Only 1 model
- Better Fit
- Less Variance explained

| Variable 2 | | | |
|---|---|---|---|
| Values | % of Records | % Enroll | Predicted Ave |
| 0 | 58.7% | 26% | 26% |
| 1 | 19.7% | 27% | 30% |
| 2 | 9.7% | 32% | 32% |
| 3 | 5.3% | 37% | 34% |
| 4 | 2.3% | 29% | 39% |
| 5 | 4.2% | 51% | 46% |

| Variable 5 Recoded into Bins | | | |
|---|---|---|---|
| Values | % of Records | % Enroll | Predicted Ave |
| 1-2 | 0.1% | 100% | 56% |
| 2-3 | 5.7% | 39% | 41% |
| 3-4 | 48.6% | 33% | 32% |
| 4-5 | 36.4% | 24% | 25% |
| 5-6 | 8.8% | 21% | 20% |
| 6-7 | 0.3% | 20% | 18% |



# Likelihood to Enroll

**Variables in the Equation**

| | B | S.E. | Wald | df | Sig. | Exp(B) | 95% C.I.for EXP(B) Lower | Upper |
|---|---|---|---|---|---|---|---|---|
| Step 1[a] | -.517 | .095 | 29.396 | 1 | <.001 | .596 | .494 | .719 |
| | .145 | .043 | 11.583 | 1 | <.001 | 1.156 | 1.064 | 1.257 |
| | .393 | .123 | 10.163 | 1 | .001 | 1.482 | 1.164 | 1.888 |
| | .219 | .051 | 18.505 | 1 | <.001 | 1.245 | 1.126 | 1.375 |
| | -.608 | .158 | 14.748 | 1 | <.001 | .544 | .399 | .743 |
| | 2.918 | .728 | 16.077 | 1 | <.001 | 18.504 | | |

- Likelihood to Apply Logit

```
(2.917969)
+(@Application-Likelihood-Enrollment-Score  * 0.218806)
+(@Application-Likelihood-Enrollment-Score  * -0.607924)
+(@Application-Likelihood-Enrollment-Score  * 0.145264)
+(@Application-Likelihood-Enrollment-Score  * -0.517411)
+(@Application-Likelihood-Enrollment-Score  * 0.393421)
```

- Likelihood to Apply Score

```
(POWER(2.718281828,@Application-Likelihood-to-Enroll-Logit ))
/(1+POWER(2.718281828,@Application-Likelihood-to-Enroll-Logit ))
```

# Likelihood to Enroll (bonus models) (with 2nd Models)

- FA 22 Financial Aid Yes

| | B |
|---|---|
| [redacted] | -.448687 |
| [redacted] | .222622 |
| [redacted] | -.300510 |
| [redacted] | 1.267292 |
| [redacted] | .382951 |
| [redacted] | -.765054 |
| [redacted] | .469116 |
| [redacted] | -.517508 |
| [redacted] | -.516972 |
| Constant | -.534354 |

- FA 22 Financial Aid No

| | B |
|---|---|
| [redacted] | -.275253 |
| [redacted] | -.521095 |
| [redacted] | .262100 |
| [redacted] | .165391 |
| [redacted] | .117656 |
| [redacted] | -.715773 |
| [redacted] | -.693433 |
| [redacted] | .889174 |
| Constant | 1.254885 |

# Likelihood to Enroll - Results

- As of 5/1/23 for Fall 2023
- Underpredicted by 7 students
  - Model is based on Census, so summer melt may get closer
  - Fina Yes Overpredicted by 8 students
  - Fina No Underpredicted by 15 students
- Underpredicted Enrollment rate
  - Bands: 10-20%, 60-80% likely to enroll
- Overpredicted Enrollment rate
  - Bands: 50-60% likely to enroll

| Predicted Likelihood | | Textbook Yield |
|---|---|---|
| 100% | 90.0% | - |
| 90.0% | 80.0% | - |
| 80.0% | 70.0% | 100% |
| 70.0% | 60.0% | 75% |
| 60.0% | 50.0% | 36% |
| 50.0% | 40.0% | 46% |
| 40.0% | 30.0% | 33% |
| 30.0% | 20.0% | 26% |
| 20.0% | 10.0% | 22% |
| 10.0% | 0.0% | 0% |

# Likelihood to Enroll – Results (with 2nd Models)

- As of 5/1/23 for Fall 2023
- “Textbook” FA 20-22 model
  - Fina Yes Overpredicted by 8 students
  - Fina No Underpredicted by 15 students
- FA 22 Fina Yes and FINA No models
  - Fina Yes Underpredicted by 6 students
  - Fina No Overpredicted by 24 students
- Differences
  - Large differences in individual-level predictions between models (up to~50-60% differences)

| Predicted Likelihood | | Textbook | Dual Model |
|---|---|---|---|
| | | Yield | Yield |
| 100% | 90.0% | - | 100% |
| 90.0% | 80.0% | - | 67% |
| 80.0% | 70.0% | 100% | 50% |
| 70.0% | 60.0% | 75% | 47% |
| 60.0% | 50.0% | 36% | 35% |
| 50.0% | 40.0% | 46% | 40% |
| 40.0% | 30.0% | 33% | 33% |
| 30.0% | 20.0% | 26% | 24% |
| 20.0% | 10.0% | 22% | 16% |
| 10.0% | 0.0% | 0% | 16% |

# Next Steps - Model

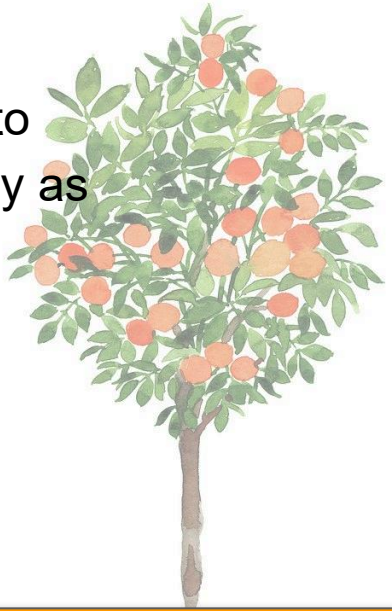
- Collaborate with Institutional Research to get their input
- Explore multi-models (FINA, International, etc.)
- Explore if there are groups I'm systematically miss-predicting
- Model Retention behaviors
- Finetune variables
  - Behaviors like emails and logins can warp models
- Use of non-Pandemic era data

# Next Steps - Use

- Plug into Net Tuition Revenue model
  - Better understand funding and headcount
- Work with OSA on support
  - Provide proactive outreach or support
- Augment yield efforts
  - Mailings/Calls/Programming to manage yield
- Measure effectiveness of use
  - Analyze the results of those receiving special yield outreach in moving yield likelihood

# Limitations

- It’s hard right now with the pandemic and changing higher education landscape to treat multiple cycles of prospective students the same way. That will make generalizing your historical data to upcoming cycles harder.
- Also, every student is different! Applying a one-size-fits-all model to tens of thousands of people will create outliers. Some people apply as their first documentable action. Some people visit campus, submit documents, email 20 times, and still don’t apply



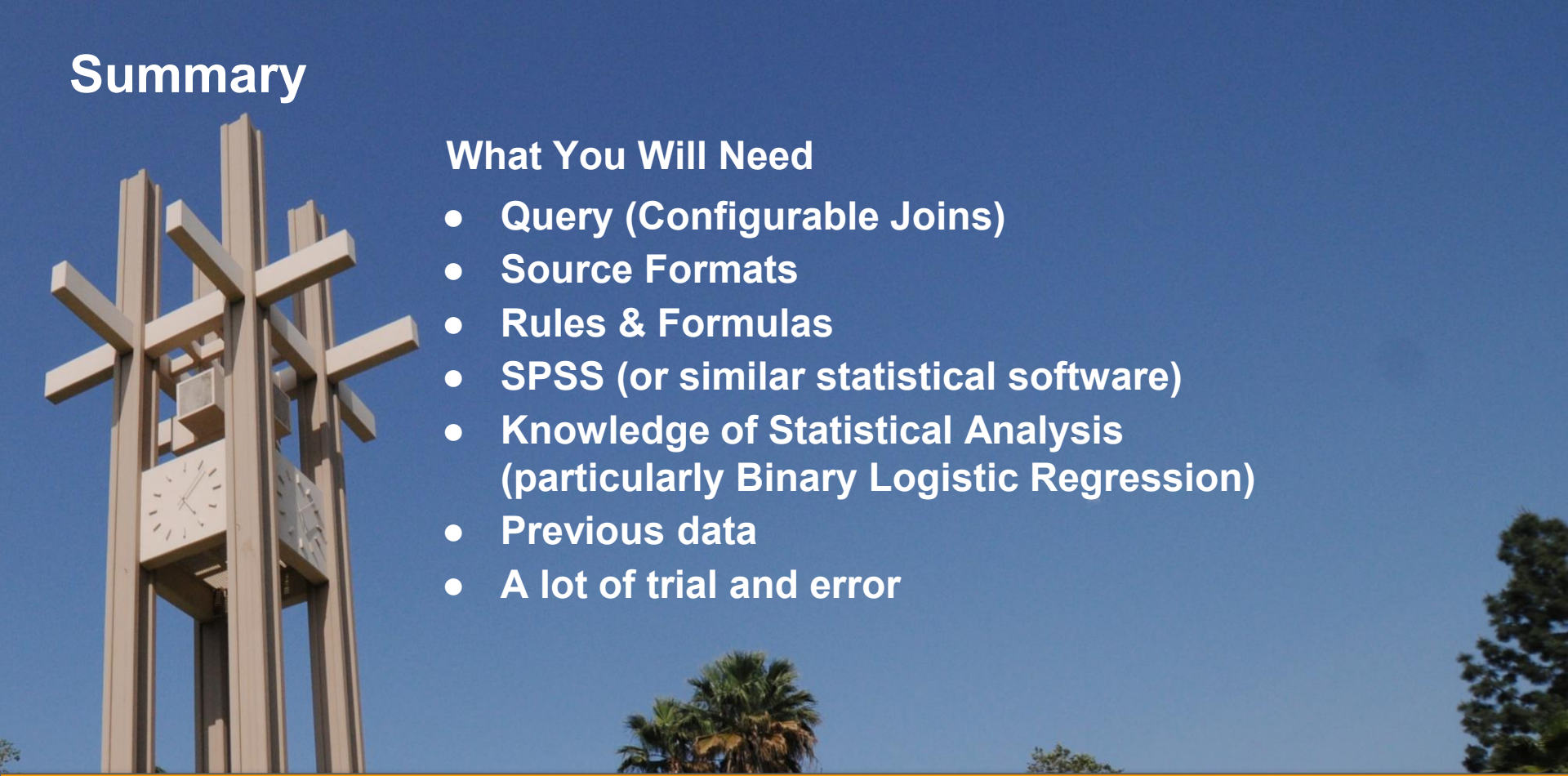
# Tips

- Reverse coding!
  - If you have a Behavior or score where 1 is better than 5, It might be easier recoding it in the Query so the highest value = the best score. That makes interpreting your results and formulas more intuitive to work with
- Overfitting your model
  - Don’t overly predict your historical data that it can’t predict future behaviors
  - Looking more times at more variables = more likely to find random significant effects

# Tips

- Changing Policies (especially pandemic-related changes in model data)
  - Be mindful if you started or stopped offering policies or ways to engage with your school
- Code/Group your variables to best suit your model (with caution)
  - Recoding your data may better predict behaviors, but there are a lot of considerations to be taken to adhere to good modeling
- Data Snapshots
  - Don’t model based on behaviors after your modeling! Applying students go to more events after applying, Enrolling students send more messages after they’ve committed

# Summary



## What You Will Need

- Query (Configurable Joins)
- Source Formats
- Rules & Formulas
- SPSS (or similar statistical software)
- Knowledge of Statistical Analysis (particularly Binary Logistic Regression)
- Previous data
- A lot of trial and error

# Briefcases and Resources

- https://www.pitzer.edu/admission/pitzer_resource_modeling/



# Briefcases and Resources

- https://www.pitzer.edu/admission/pitzer_resource_modeling/




Suitcases

08454bca-159c-4d8b-b84e-a95ba7941451:pzr

| Name |
| --- |
| Add Item |
| Pres Sample Likelihood Score Apply |
| Pres Sample Likelihood Score Apply Logit Calculation |
| Pres Sample Likelihood Score Apply Probability Calculation |
| Pres Sample Likelihood Score Enroll Logit Calculation |
| Pres Sample Likelihood Score Enroll Probability Calculation |
| Pres Sample Likelihood Score Enroll |
| Pres Sample DI Score |

# Questions?




**David Dysart**

**Sr. Associate Director**
**Admission Operations**

David_Dysart@Pitzer.edu