THE HOUSING GAMES

A CS1951A FINAL PROJECT BY

MATT COOPER, ALI PAUL, MICHAEL GILLETT AND SAM DUNIETZ

THE PROBLEM

Traditionally, Brown's housing lottery has been a notoriously messy, some might even say traumatic, process. Some have even gone as far as to call the housing lottery The Housing Games, making light of the stressful process by equating it with the fictional dystopia of The Hunger Games. Our goal was to improve this process using data. We created a variety of tools that enable students to make more informed decisions and plans about their future housing by using results from previous housing lotteries.

THE DATASET

- Housing lottery data collected from 2006-2013 on previous lottery results.
- Data collected in 2015 is formatted differently than that collected from 2006 - 2013, 2014 was missing.
- Therefore, we used the data collected over the longest consecutive period of time

METHODOLOGY

- The data was available on the reslife website as excel spreadsheets
- We converted the spreadsheets into csv files, removed any no show, pass, and dropped entries and capitalized all of the dorm names in the same format. We then used this cleaned data for visualizations and analysis.

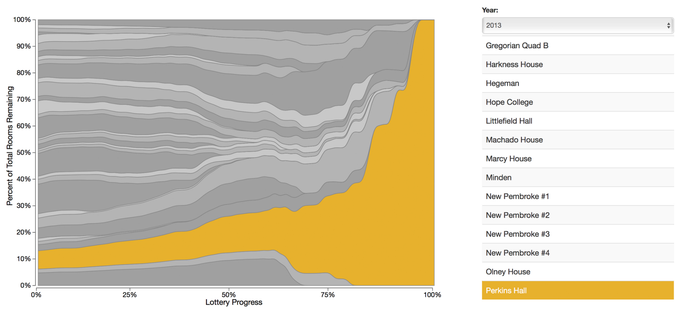
CHALLENGES

- We found some items we wanted to incorporate would take longer than we had for this project.
- For example, we wanted to collect data on the square footage of each room; however, we found that this data was only available as images in a PDF and would have taken over 30 hours to collect by hand.

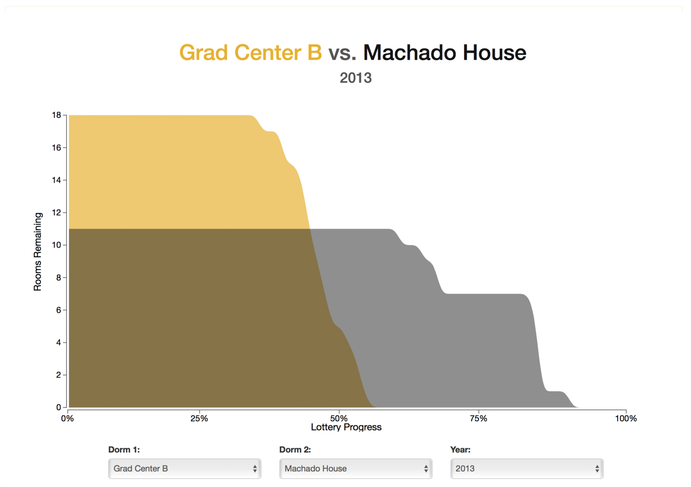
OUR RESULTS

See interactive visualizations and explanations of our analyseson our website at bit.ly/housing-games

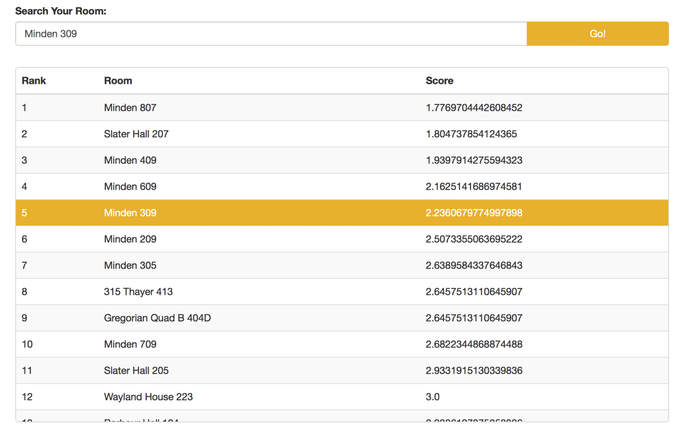
% TOTAL ROOMS BY DORMITORY



ROOMS REMAINING COMPARISON

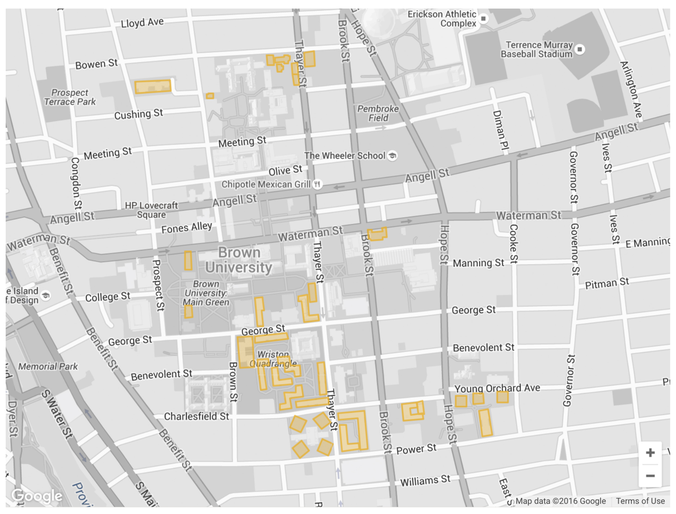


ROOM RANKING + SUGGESTIONS



Rank	Room	Score
1	Minden 807	1.7769704442608452
2	Slater Hall 207	1.804737854124365
3	Minden 409	1.9397914275594323
4	Minden 609	2.1625141686974581
5	Minden 309	2.2360679774997898
6	Minden 209	2.5073355063695222
7	Minden 305	2.6389584337646843
8	315 Thayer 413	2.6457513110645907
9	Gregorian Quad B 404D	2.6457513110645907
10	Minden 709	2.6822344868874488
11	Slater Hall 205	2.9331915130339836
12	Wayland House 223	3.0

MAP VISUALIZATION



MOST IMPORTANT FEATURES

For the machine learning aspect of this project, we ran linear regression to find out which features most strongly predict students' dorm preference. (i.e. What makes a dorm most desirable?) To do this, we used our calculated room ranking scores as predicted values and ran gradient descent to determine the strength of each feature.

Features and Weight

Distance from Nelson Fitness Center	-250.307
Distance from the Main Green	209.214
Distance from the Sharpe Refectory	101.161
Occupancy of Room	0.995