**MSDS 7374 Business Analytics  
Project Framing - FACE**

**Chris Ballenger  
Chad Madding**

**Framing the problem**

Real estate tries to connect clients with the right situation, but may wait for clients to come to them. We want to build an algorithm that will place clients with the right building. For our problem, we have a brewpub that recently closed and we would like to find a tenant that will work well in its new location. Our modeling approach will consist of a mix of numerical and quantitative data that will allow for both prediction and interpretation of the results. Our company possess the technological and practical expertise to implement all the Python code and ArcGIS data gathering that will be needed for this project.

**Analysis/solving the problem**

The data used in the project include the most current US Census data consisting median income and disposable income by track and ZIP code, traffic count and a CSV file entitled “Texas Brewery Hours” from [www.craftbeeraustin.com](http://www.craftbeeraustin.com). “Texas Brewery Hours” included a list of all Texas breweries and brewpubs along with their addresses and status of being opened or closed. This will allow for the breweries and brewpubs to be plotted in ArcGIS. We will also collect social rating \ check-ins data ffrom Untapped; using their API to capture information about local beers, breweries and brewpubs along with “check-in” data that will be used for our prediction model. Due to the nature of this project, our models will be more stochastic than deterministic, consisting mostly regression analysis. Our data analysis will consist of descriptive and inferential statistics along with visualizations mostly in the form of location maps with census data. After collecting data, we will merge Untapped details (See Data Definitions in appendix) along with zip code demographic and 2 neighboring zip code demographics. This will allow us to create a regression model that estimates how many monthly check-ins a zip code can handle. This data will be used in evaluating the sustainability of a new brewpub opening in the now empty Humperdinks building. We would also like to find reasons for Humperdinks’ failure and compare that to successful brewpubs in the area; Dallas’ White Rock Ale House and Ft. Worth’s Funkytown Fermatorium.

**Communicating and acting on results**

The data above will be collected and based on our results, we hope to find an optimal client who has an established brewery and is looking to expand operations into Dallas market. The reports will be compiled and presented with two separate audiences in mind. The first being tailored to the CEO of the brewery. This report will consist an explanation of the need for expansion, increase customers and more important, loyalty, by creating a brewpub footprint in Dallas it will contain local demographics and a general budget needed for the purchase of the building and the build out. This report will be mostly non-analytical in delivery. The second audience will be towards a CIO or the breweries’ chief marketing agent. This report will be more tactical in nature. There may also be some resistance from this team so there will be a need to address some issues to allow for buy in from this department. This report will contain the results from the model above and show the finding from all of our research. The audience to this report has backgrounds in statistics and data analytics. Both stories will address the rewards and the risk involved. If the project is accepted as viable for the company progress forward will be managed carefully to unify all stakeholders. This will be a new venture for an existing brewery and the actions needed to move this plan into action will need to be gradual and well thought out.

**Embedding final models and methods in** **enterprise business processes and systems**

This project will be embedded into the business processes by implementing ArcGIS simulations to examine different breweries and brewpubs in the model to assess different outcomes. We will also start predicting monthly check-ins to help clients find their next brewpub or brewery location.

**Data Descriptions**

Texas Brewery Hours

|  |  |
| --- | --- |
| Name | Used to search untappd API  Removed the following words to help search   * Company * Inc * Co * Cooperative * Co-Op * Llc |
| Street & Address |  |
| City |  |
| State |  |
| Brewery License | Brewery or Brewpub |
| License Date |  |
| Status | Active or In the workds |

**Untappd**

|  |  |
| --- | --- |
| Brewery Name |  |
| brewery\_in\_production | 0 = Yes; 1 = No |
| brewery\_id | Untappd Unique Id |
| brewery\_address, brewery\_city, brewery\_state |  |
| Brewery\_lat & brewery\_lng | Lattitude & Longitude used to find zipcode |
| beer\_count | Total Unique Beers produced (active & inactive) |
| rating\_count |  |
| rating\_score |  |
| age\_on\_service | # of Days on Untappd Application |
| brewery\_zipcode | Derived from lat\lng |
| neighboring\_zips | 2 nearest neighboring zip codes |
| checkins\_total\_count | Check-ins for life of brewery |
| checkins\_unique\_count | Unique number of users |
| checkins\_monthly\_count | Monthly checkins – **Response variable** |
| checkins\_weekly\_count | Weekly check-ins |

**ArcGIS Demographic**

|  |  |  |
| --- | --- | --- |
| OBJECTID | Include? | next 2 zips? |
| ID |  |  |
| NAME |  |  |
| STATE\_NAME |  |  |
| ST\_ABBREV |  |  |
| 2018 Total Population (Esri) | Y | Sum |
| 2018 Household Population (Esri) | Y | Sum |
| 2018 Population Density (Pop per Square Mile) (Esri) |  |  |
| 2018 Total Households (Esri) | Y | Sum |
| 2018 Average Household Size (Esri) |  |  |
| 2010-2018 Population: Annual Growth Rate (Esri) |  |  |
| 2010-2018 Households: Annual Growth Rate (Esri) |  |  |
| 2018 Generation Alpha Population (Born 2017 or Later) |  |  |
| 2018 Generation Z Population (Born 1999 to 2016) | Y | Sum |
| 2018 Millennial Population (Born 1981 to 1998) | Y | Sum |
| 2018 Generation X Population (Born 1965 to 1980) | Y | Sum |
| 2018 Baby Boomer Population (Born 1946 to 1964) | Y | Sum |
| 2018 Silent & Greatest Generations Population (Born 1945/Earlier) |  |  |
| 2018 Population by Generation Base |  |  |
| 2018 Total Population Age 0-4 (Esri) |  |  |
| 2018 Total Population Age 5-9 (Esri) |  |  |
| 2018 Total Population Age 10-14 (Esri) |  |  |
| 2018 Total Population Age 15-19 (Esri) |  |  |
| 2018 Total Population Age 20-24 (Esri) | Y | Sum |
| 2018 Total Population Age 25-29 (Esri) | Y | Sum |
| 2018 Total Population Age 30-34 (Esri) | Y | sum |
| 2018 Total Population Age 35-39 (Esri) | Y | sum |
| 2018 Total Population Age 40-44 (Esri) | Y | sum |
| 2018 Total Population Age 45-49 (Esri) | Y | sum |
| 2018 Total Population Age 50-54 (Esri) | Y | sum |
| 2018 Total Population Age 55-59 (Esri) | Y | sum |
| 2018 Total Population Age 60-64 (Esri) | Y | sum |
| 2018 Total Population Age 65-69 (Esri) | Y | sum |
| 2018 Total Population Age 70-74 (Esri) | Y | sum |
| 2018 Total Population Age 75-79 (Esri) | Y | sum |
| 2018 Total Population Age 80-84 (Esri) | Y | sum |
| 2018 Total Population Age 85+ (Esri) | Y | sum |
| 2018 Median Age (Esri) | Y | Avg |
| 2018 Male Population (Esri) | Y | Sum |
| 2018 Median Male Age (Esri) | Y | Avg |
| 2018 Female Population (Esri) | Y | Sum |
| 2018 Median Female Age (Esri) | Y | Avg |
| 2018 Total Population by Five-Year Age Base (Esri) |  |  |
| 2018 Civilian Population Age 16+ in Labor Force (Esri) |  |  |
| 2018 Employed Civilian Population Age 16+ (Esri) |  |  |
| 2018 Unemployed Population Age 16+ (Esri) |  |  |
| 2018 Unemployment Rate (Esri) |  |  |
| 2018 Hispanic Population (Esri) | Y | sum |
| 2018 White Non-Hispanic Population (Esri) | Y | sum |
| 2018 Black/African American Non-Hispanic Population (Esri) | Y | sum |
| 2018 American Indian/Alaska Native Non-Hispanic Population (Esri) | Y | sum |
| 2018 Asian Non-Hispanic Population (Esri) | Y | sum |
| 2018 Pacific Islander Non-Hispanic Population (Esri) | Y | sum |
| 2018 Other Race Non-Hispanic Population (Esri) | Y | sum |
| 2018 Multiple Races Non-Hispanic Population (Esri) | Y | sum |
| 2018 Diversity Index (Esri) |  |  |
| 2018 Population by Race Base (Esri) |  |  |
| 2018 Household Income less than $15,000 (Esri) | Y | sum |
| 2018 Household Income $15,000-$24,999 (Esri) | Y | sum |
| 2018 Household Income $25,000-$34,999 (Esri) | Y | sum |
| 2018 Household Income $35,000-$49,999 (Esri) | Y | sum |
| 2018 Household Income $50,000-$74,999 (Esri) | Y | sum |
| 2018 Household Income $75,000-$99,999 (Esri) | Y | sum |
| 2018 Household Income $100,000-$149,999 (Esri) | Y | sum |
| 2018 Household Income $150,000-$199,999 (Esri) | Y | sum |
| 2018 Household Income $200,000 or greater (Esri) | Y | sum |
| 2018 Median Household Income (Esri) | Y | avg |
| 2018 Average Household Income (Esri) | Y | avg |
| 2018 Per Capita Income (Esri) | Y | avg |
| 2018 Households by Income Base (Esri) | Y | avg |
| 2018 Median Home Value (Esri) | Y | avg |
| 2018 Average Home Value (Esri) | Y | avg |
| 2023 Total Population (Esri) |  |  |
| 2023 Household Population (Esri) |  |  |
| 2023 Population Density (Pop per Square Mile) (Esri) |  |  |
| 2023 Total Households (Esri) |  |  |
| 2023 Average Household Size (Esri) |  |  |
| 2018-2023 Population: Annual Growth Rate (Esri) |  |  |
| 2018-2023 Households: Annual Growth Rate (Esri) |  |  |
| 2018-2023 Per Capita Income: Annual Growth Rate (Esri) |  |  |
| 2018-2023 Median Household Income: Annual Growth Rate (Esri) |  |  |
| 2023 Diversity Index (Esri) |  |  |
| 2023 Median Household Income (Esri) |  |  |
| 2023 Average Household Income (Esri) |  |  |
| 2023 Per Capita Income (Esri) |  |  |
| 2010 Total Population (U.S. Census) |  |  |
| Shape\_\_Area |  |  |
| Shape\_\_Length |  |  |
| 2018 Total Housing Units (Esri) | Y | sum |
| 2018 Owner Occupied Housing Units (Esri) | Y | sum |
| 2018 Renter Occupied Housing Units (Esri) | Y | sum |
| 2010 Owner-occupied Housing Units with a Mortgage/loan (U.S. Census) |  |  |
| 2010 Owner-occupied Housing Units Owned Free and Clear (U.S. Census) |  |  |
| 2010 Population Living in Owner-occupied Housing Units (U.S. Census) |  |  |
| 2010 Population Living in Renter-occupied Housing Units (U.S. Census) |  |  |
| 2018 Vacant Housing Units (Esri) |  |  |