Project 3 Proposal
Predicting BMI Classification

Bianca Orozco

Domain

I will be using two datasets from Kaggle. The first dataset provides fast food restaurant data from all across America. This includes the state, city, postal code, and restaurant name. The second dataset provides obesity statistics across the US. This includes the state, BMI value, and sample size.

MVP

The analysis will predict which BMI class someone would be in.

Fast Food Restaurant Data

	Variable	Type	Description	Used for Model
1.	Date Added	Str	Year-month-day	N
2.	Date Updated	Str	Year-month-day	N
3.	Province (State)	Str	State Abbreviation (CA, WA)	Y
4.	City	Str	City name	Y
5.	Postal Code	Int	5 digit postal code	?
6.	Address	Str	Restaurant address	?
7.	Latitude	Float	…	N
8.	Longitude	Float	…	N
9.	Restaurant Name	Str	(McDonald's, Taco Bell)	Y

Obesity Data

	Variable	Type	Description	Used for Model
1.	Year	Str	(2011 – 2016)	N
2.	Location (State)	Str	State Abbreviation	Y
3.	Topic	Str	Topic being measured	N
4.	Question	Str	Question being answered	N
5.	Data Value	Int	BMI Score	Y
6.	BMI Class	Str	BMI Classification	Y
7.	Low Confidence Limit	Int	…	N
8.	High Confidence Limit	Int	…	N
9.	Sample Size	Int	…	N