Project Movat

A guide to understanding the social factors that may contribute to obesity

Presenters: Alexei Flores, Leonard Zhao, Katrina Nguyen

KALogic

PowerData Insights

Objective

What socioeconomic and economic factors has the highest correlation with obesity rates? The objective is to find if there is correlation between obesity rate and the following social factors:

- **Median household income**
- **% Exercise opportunities**
- **% Physically inactive**
- **% Access to primary care physicians**
- **High school graduation**
- **% Unemployed**

Data Collection

Socrata

United States Census Bureau

KALogic

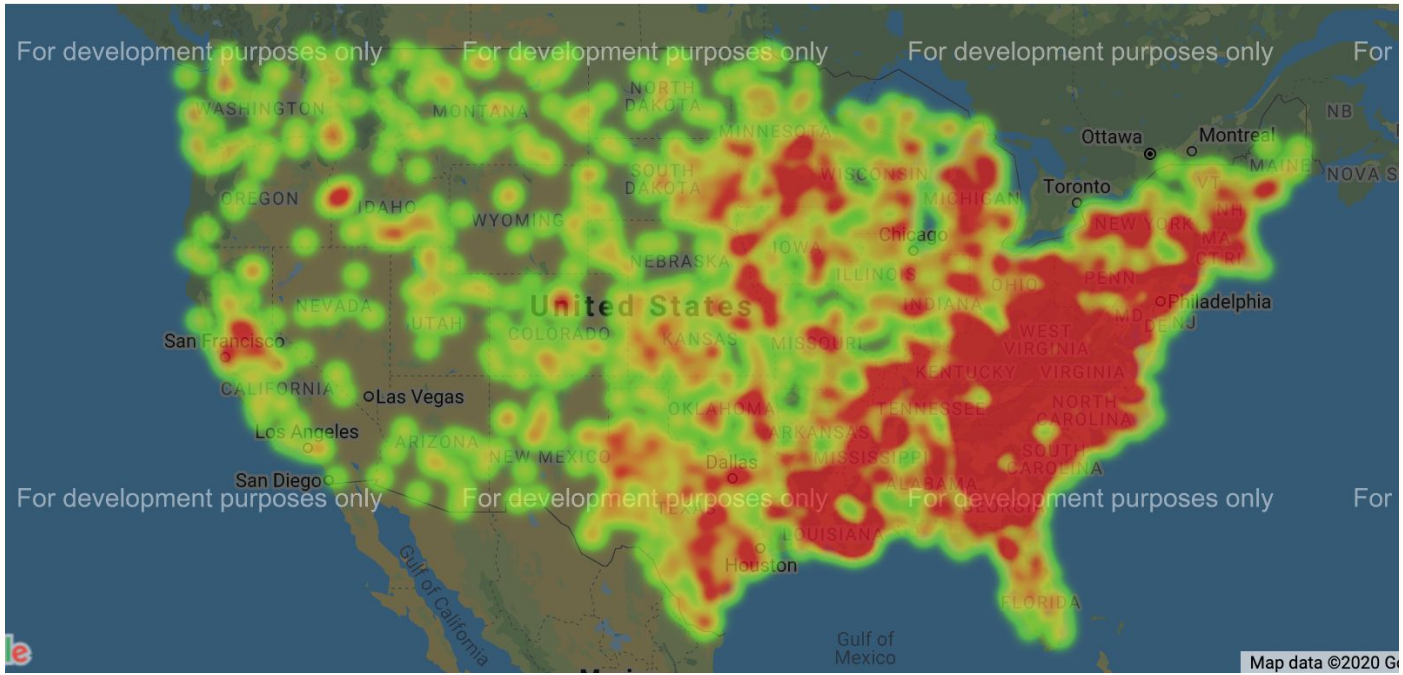
PowerData Insights

```
path = "./CountyHealthRankings"
all_files = glob.glob(path + "/*.csv")

# Read all csv files in folder and append to create one dataframe

all_states = []
for filename in all_files:
    df = pd.read_csv(filename, index_col=None, header=0)
    all_states.append(df)
frame = pd.concat(all_states, axis=0, ignore_index=True)
frame.head()
states_health_df = frame[[
                          "FIPS",
                          "% Adults with Obesity",
                          "% Physically Inactive",
                          "% With Access to Exercise Opportunities",
                          "Primary Care Physicians Rate",
                          "High School Graduation Rate",
                          "% Unemployed"
                         ]]
states_health_df.head()
```

Obesity Rate Heatmap for the Lower 48 states
Where is obesity concentrated in the US?



KALogic
PowerData Insights

Obesity Rate vs. % Unemployed

California

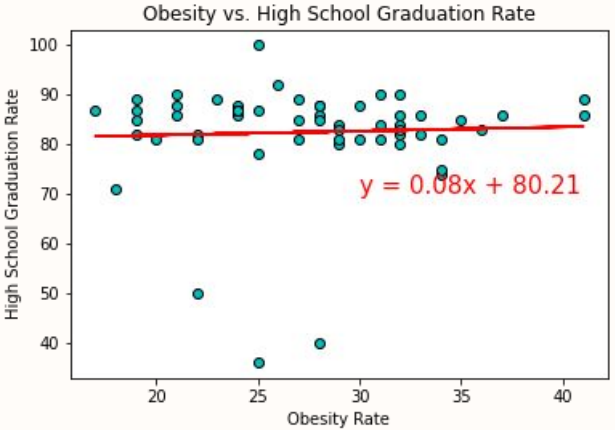
Texas

Florida

KALogic

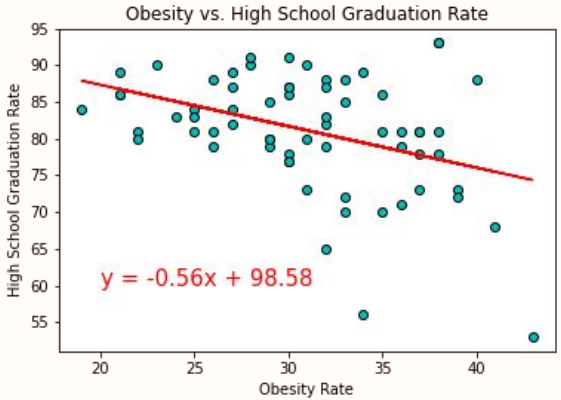
PowerData Insights

Obesity Rate vs. % High School Graduation



California

Texas

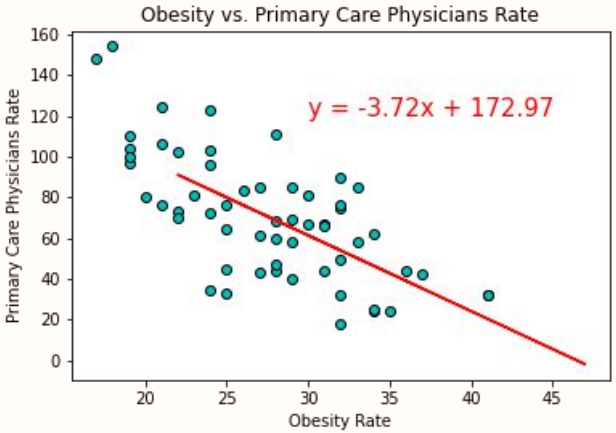


Florida

KALogic

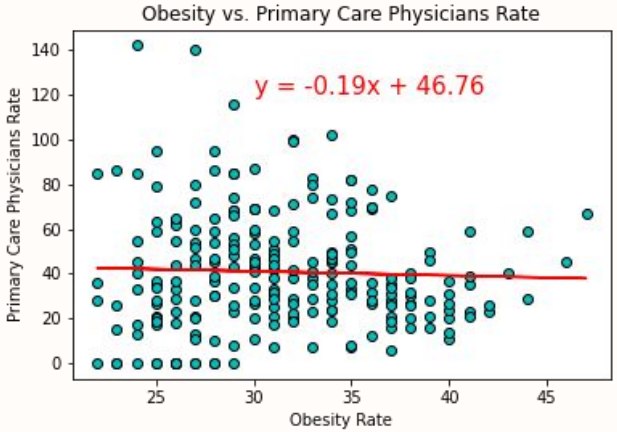
PowerData Insights

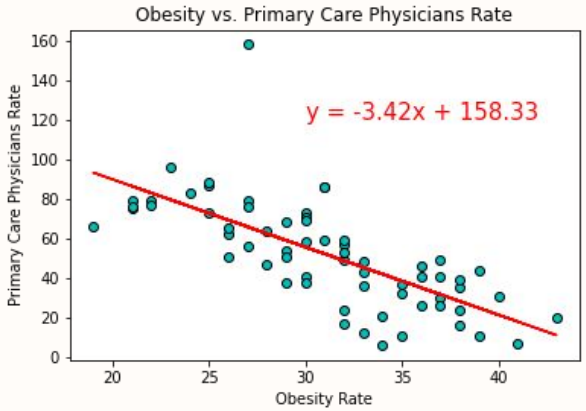
Obesity Rate vs. % Access to Primary Care Physicians



California

Texas



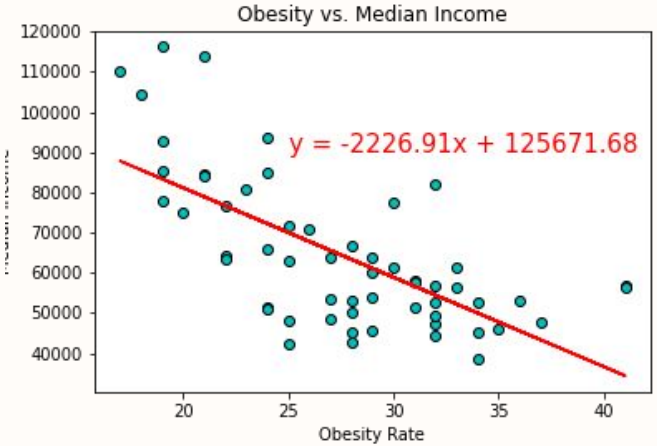


Florida

KALogic

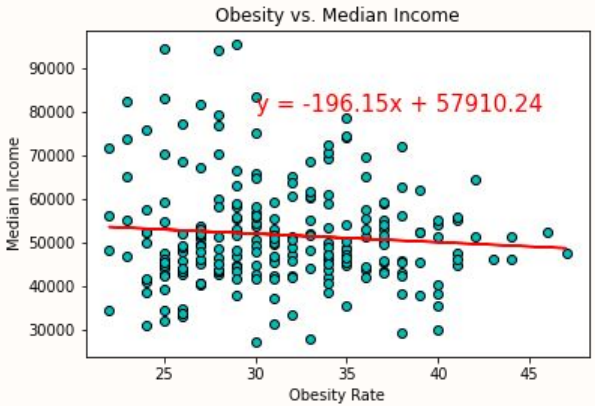
PowerData Insights

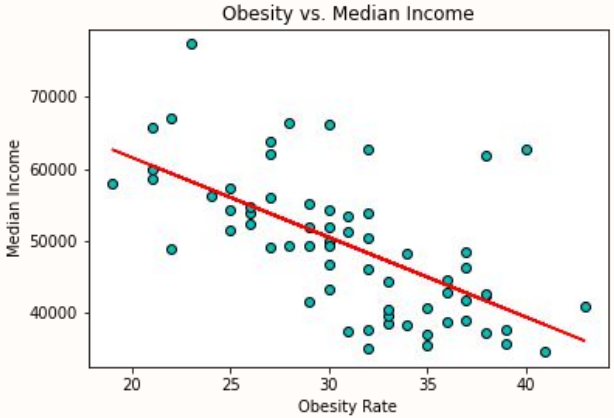
Obesity Rate vs. Median Income



California

Texas



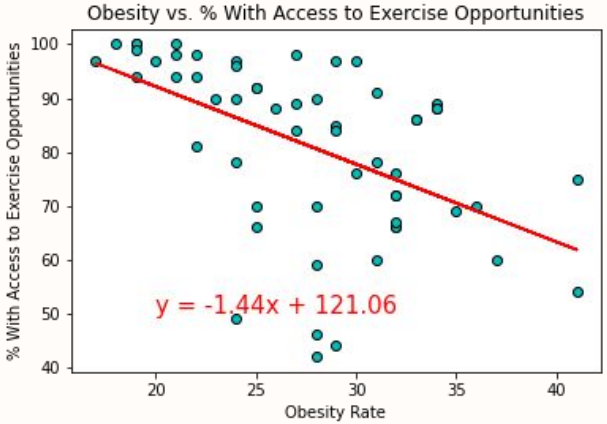


Florida

KALogic

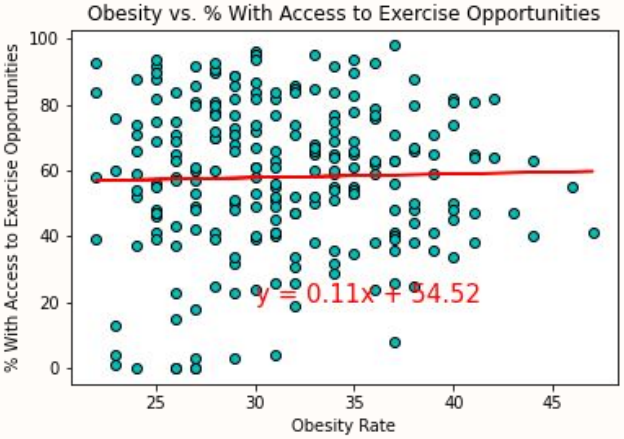
PowerData Insights

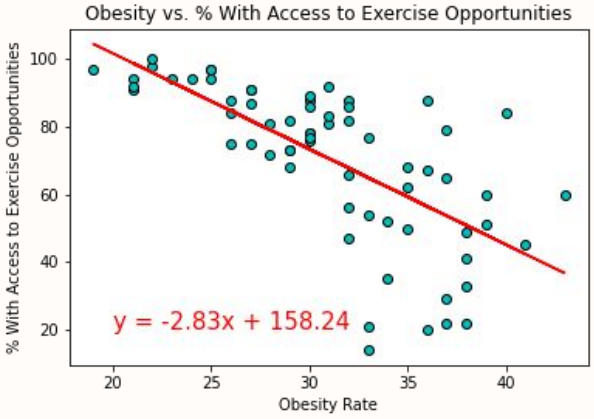
Obesity Rate vs. % Exercise Opportunities



California

Texas





Florida

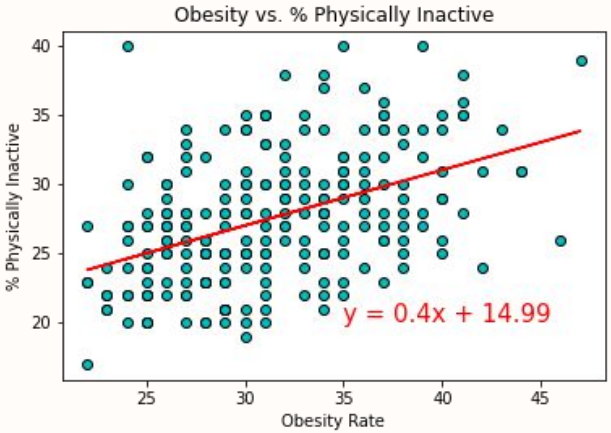
KALogic

PowerData Insights

Obesity Rate vs. % Of Physically Inactive

California

Texas



Florida

KALogic

PowerData Insights

Omnibus Model testing 6predictors vs. Obesity Rate

OLS Regression Results

Dep. Variable:	% Adults with Obesity	R-squared (uncentered):	0.980
Model:	OLS	Adj. R-squared (uncentered):	0.980
Method:	Least Squares	F-statistic:	2886.
Date:	Fri, 27 Nov 2020	Prob (F-statistic):	2.58e-293
Time:	14:26:43	Log-Likelihood:	-1025.5
No. Observations:	354	AIC:	2063.
Df Residuals:	348	BIC:	2086.
Df Model:	6		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
% With Access to Exercise Opportunities	0.0211	0.012	1.731	0.084	-0.003	0.045
% Physically Inactive	0.6381	0.040	15.968	0.000	0.560	0.717
High School Graduation Rate	0.0873	0.018	4.773	0.000	0.051	0.123
Median Income	2.907e-05	2e-05	1.455	0.147	-1.02e-05	6.84e-05
Primary Care Physicians Rate	-0.0167	0.011	-1.587	0.113	-0.037	0.004
% Unemployed	0.8521	0.149	5.722	0.000	0.559	1.145

Omnibus:	7.883	Durbin-Watson:	1.942
Prob(Omnibus):	0.019	Jarque-Bera (JB):	8.571
Skew:	0.267	Prob(JB):	0.0138
Kurtosis:	3.543	Cond. No.	3.50e+04

Reduced model testing 3 predictors vs. Obesity Rate

OLS Regression Results

Dep. Variable:	% Adults with Obesity	R-squared:	0.399
Model:	OLS	Adj. R-squared:	0.394
Method:	Least Squares	F-statistic:	77.39
Date:	Fri, 27 Nov 2020	Prob (F-statistic):	2.02e-38
Time:	20:33:06	Log-Likelihood:	-1020.0
No. Observations:	354	AIC:	2048.
Df Residuals:	350	BIC:	2063.
Df Model:	3		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	13.8525	2.193	6.316	0.000	9.539	18.166
% Physically Inactive	0.5491	0.046	11.996	0.000	0.459	0.639
Median Income	-1.219e-05	1.98e-05	-0.616	0.538	-5.11e-05	2.67e-05
% Unemployed	0.6533	0.154	4.252	0.000	0.351	0.956

Omnibus:	9.467	Durbin-Watson:	1.888
Prob(Omnibus):	0.009	Jarque-Bera (JB):	9.480
Skew:	0.392	Prob(JB):	0.00874
Kurtosis:	3.171	Cond. No.	5.24e+05

Applying the best model to all US counties

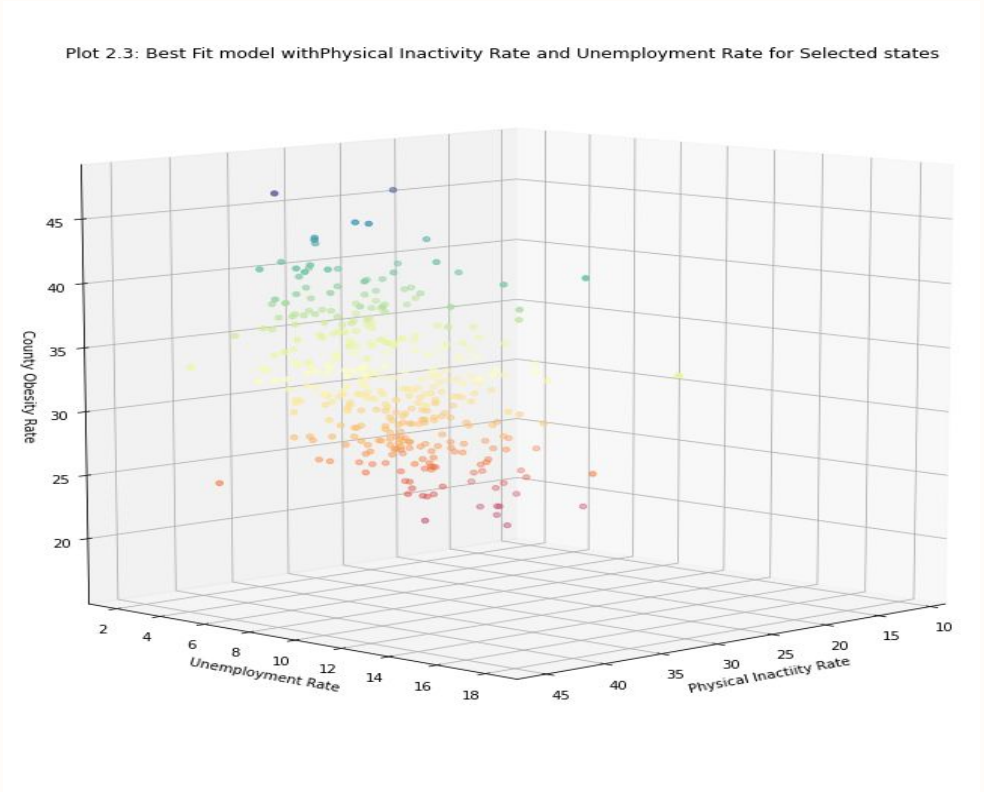
OLS Regression Results

Dep. Variable:	% Adults with Obesity	R-squared:	0.342
Model:	OLS	Adj. R-squared:	0.341
Method:	Least Squares	F-statistic:	503.9
Date:	Sat, 28 Nov 2020	Prob (F-statistic):	1.05e-263
Time:	05:00:45	Log-Likelihood:	-8451.7
No. Observations:	2914	AIC:	1.691e+04
Df Residuals:	2910	BIC:	1.694e+04
Df Model:	3		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	18.0186	1.143	15.766	0.000	15.778	20.260
% Physically Inactive	0.5141	0.015	34.820	0.000	0.485	0.543
High School Graduation Rate	-0.0113	0.012	-0.958	0.338	-0.034	0.012
% Unemployed	0.4409	0.060	7.342	0.000	0.323	0.559

Omnibus:	35.011	Durbin-Watson:	1.708
Prob(Omnibus):	0.000	Jarque-Bera (JB):	55.623
Skew:	0.089	Prob(JB):	8.35e-13
Kurtosis:	3.653	Cond. No.	1.31e+03

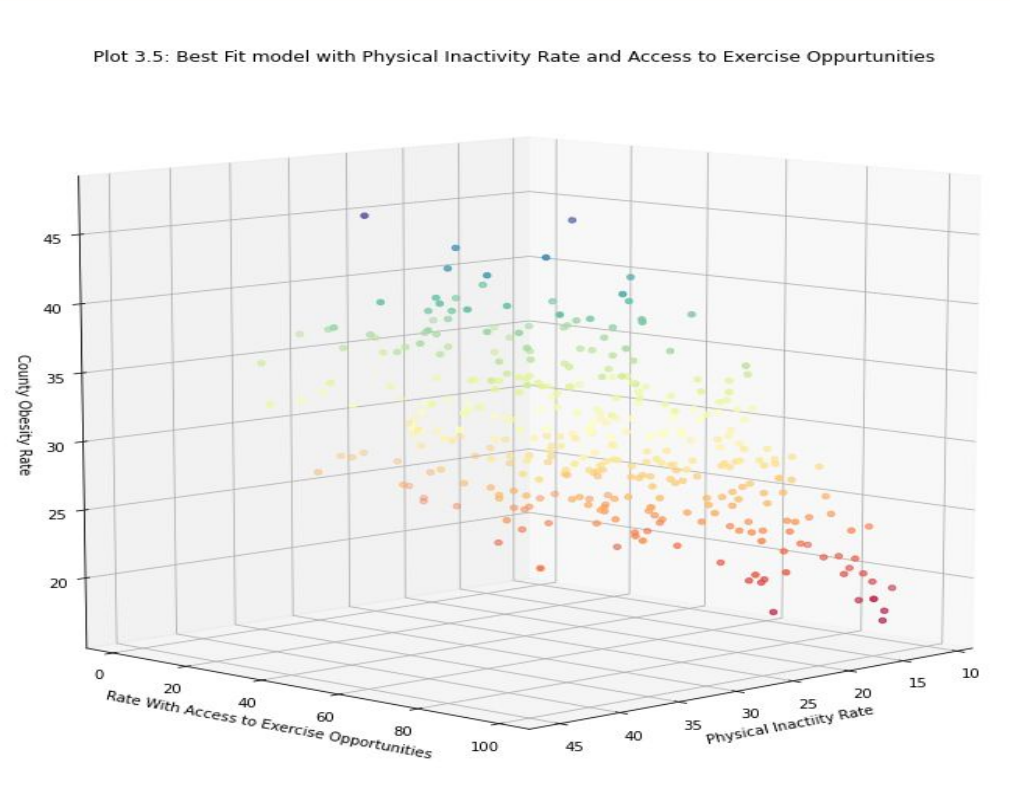
Reduced model testing 2 predictors vs. Obesity Rate



KALogic

PowerData Insights

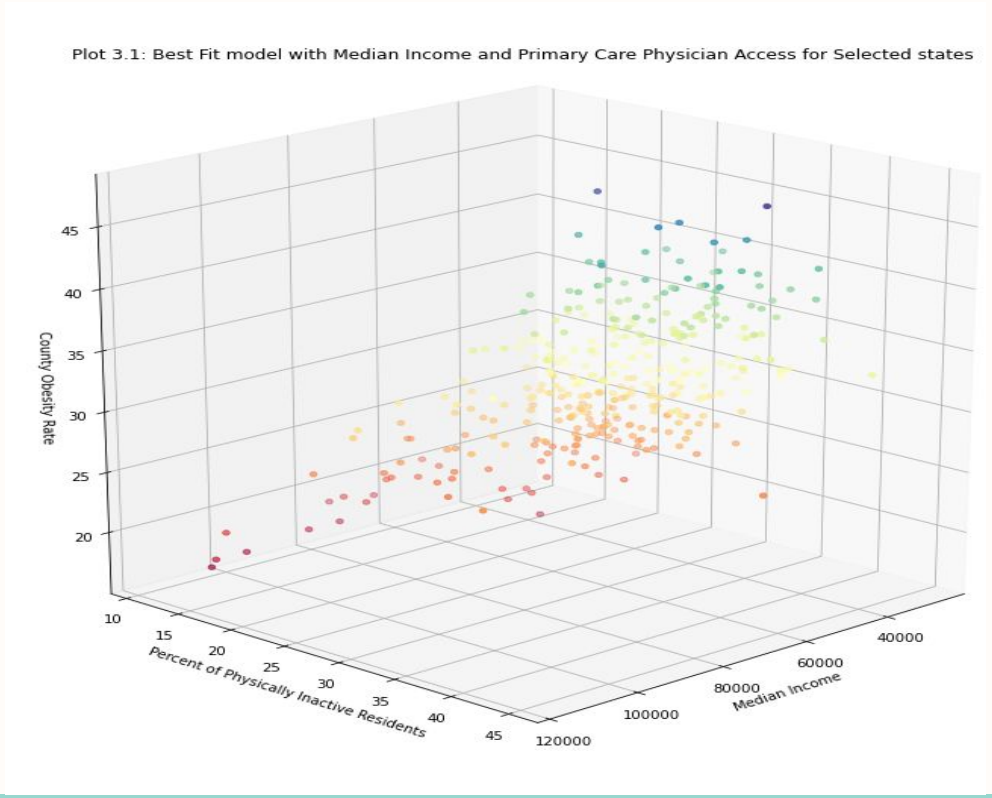
Reduced model testing 2 predictors vs. Obesity Rate



KALogic

PowerData Insights

Reduced model testing 2 predictors vs. Obesity Rate

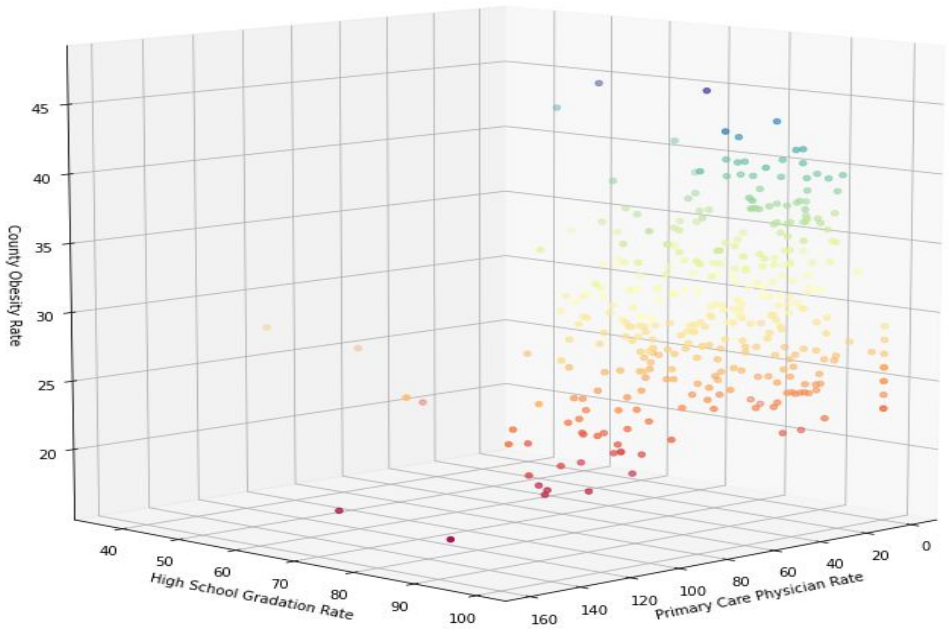


KALogic

PowerData Insights

Reduced model testing 2 predictors vs. Obesity Rate

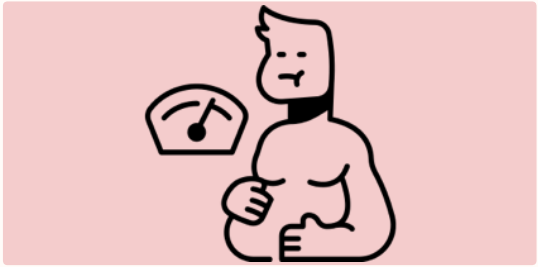
Plot 3.2: Best Fit model with High School Graduation Rate and Primary Care Physician Access for Selected states



KALogic

PowerData Insights

General Findings



33%

Adults were considered to be obese in 2016

15%

Of the top 100 counties with highest obesity rates is in Mississippi, making it the worst state

27%

Adults reported NO leisure-time physical activity

KALogic

PowerData Insights

Limitations

- Median Income data is limited to what we can get from the US Census, therefore it may not be a perfect representation of the correlation between obesity and median income for the year we looked at
- Data is at the county level; and findings may not be applicable to individuals
- Datasets only represent one year
- BMI is not the best indicator of obesity

Conclusions

- Obesity is not necessarily a personal failure. While the best predictors for obesity are wellness indicators like Physical Inactivity; social indicators appear to have significant influence on the outcome as well even when we control for wellness indicators:

- Physical inactivity matters; but substantially less in counties with high median incomes. (Plot 3.1)

- Exercise opportunities and being physically inactive have some the highest correlation, so increasing exercise opportunity can be one solution in the fight against obesity; but only if combined with campaigns to encourage less inactivity. (Plot 3.5)

Take care of your body.

- Eat well-balanced, healthy meals
- Exercise regularly
- Stress less, meditate, and take deep breaths

In other words, do the OPPOSITE of what we've all been doing since the start of the pandemic/ data bootcamp

KALogic

PowerData Insights

FOR OUR EYES ONLY: Explaining MR