Obesity Risk Prediction

Team 3

Team Members:
Tammy(Ting-Yu) Hsu, Sweekruti Palei, Linda Sun, Farin Fukunaga, Boyang Fan

Content

Introduction

Project Overview:

The objective of this study is to explore the determinants of obesity rates in Latin America, with a particular emphasis on dietary patterns, physical health, and demographic variables. The dataset contains 17 attributes and 2111 entries, offering extensive insight into the behaviors and lifestyles of participants. Originally utilized by The Coast University (University in Columbia) for academic research, the data was gathered through a web-based survey platform, with respondents remaining anonymous as they answered each questionnaire item.

Significance:

1. **Public Health Impact:** Inform public health initiatives and interventions in Latin American countries by addressing specific challenges faced by diverse populations.
2. **Personalized Health Recommendations:** Develop personalized health recommendations, considering individual eating habits, physical activity, and demographics.
3. **Policy Formulation:** Create targeted policies addressing obesity-related issues at a regional level.

Research Purpose and Questions

Research Questions:

Queston 1:	How does method of transportation affect obesity?
Question 2:	How does diet affect obesity?
Question 3:	How do alcohol and smoking affect obesity?

Intended Audience:

Considering our research questions, we aim to discover the root causes of obesity, given its potential to precipitate severe health complications regardless of age. By understanding the typical progression of obesity across individuals' lifespans, we can offer constructive guidance to help prevent excessive weight gain and promote healthier lifestyles.

Who will benefit from this research?

- People of all ages and genders
- People who are overweight or have poor lifestyles (smokers and frequent drinkers)
- Government policy makers and researchers

Dataset Overview

About Dataset - Obesity and CDV Risk

The dataset comprises estimates of obesity levels in individuals **aged 14 to 61** from **Mexico, Peru, and Colombia**, with diverse eating habits and physical conditions. Data was collected **via a web survey from anonymous users**.
It includes variables like FAVC (Frequent consumption of high-caloric food), NCP (Number of main meals), CAEC (Consumption of food between meals), FAF (Physical activity frequency), and MTRANS (Transportation used), which could be utilized for predicting and analyzing obesity or CVD risk.

Data source: Kaggle
Data scope: Mexico, Peru and Colombia (Latin America)
Timespan: 2019 Aug - 2019 Sep

Dataset Details - Variables

Obesity Categories:

- Underweight (BMI less than 18.5)
- Normal (18.5 to 24.9)
- Overweight (25.0 to 29.9)
- Obesity I (30.0 to 34.9)
- Obesity II (35.0 to 39.9)
- Obesity III (Higher than 40)

Attributes for Analysis:

- **Eating Habits:**

Frequent consumption of high-caloric food (FAVC)
Frequency of consumption of vegetables (FCVC)
Number of main meals (NCP)
Consumption of food between meals (CAEC)
Consumption of water daily (CH20)
Consumption of alcohol (CALC)

- **Physical Condition:**

Calories consumption monitoring (SCC)
Physical activity frequency (FAF)
Time using technology devices (TUE)
Transportation used (MTRANS)

- **Demographic Information:** Gender, Age, Height, Weight

Data Exploration

Descriptive Statistics

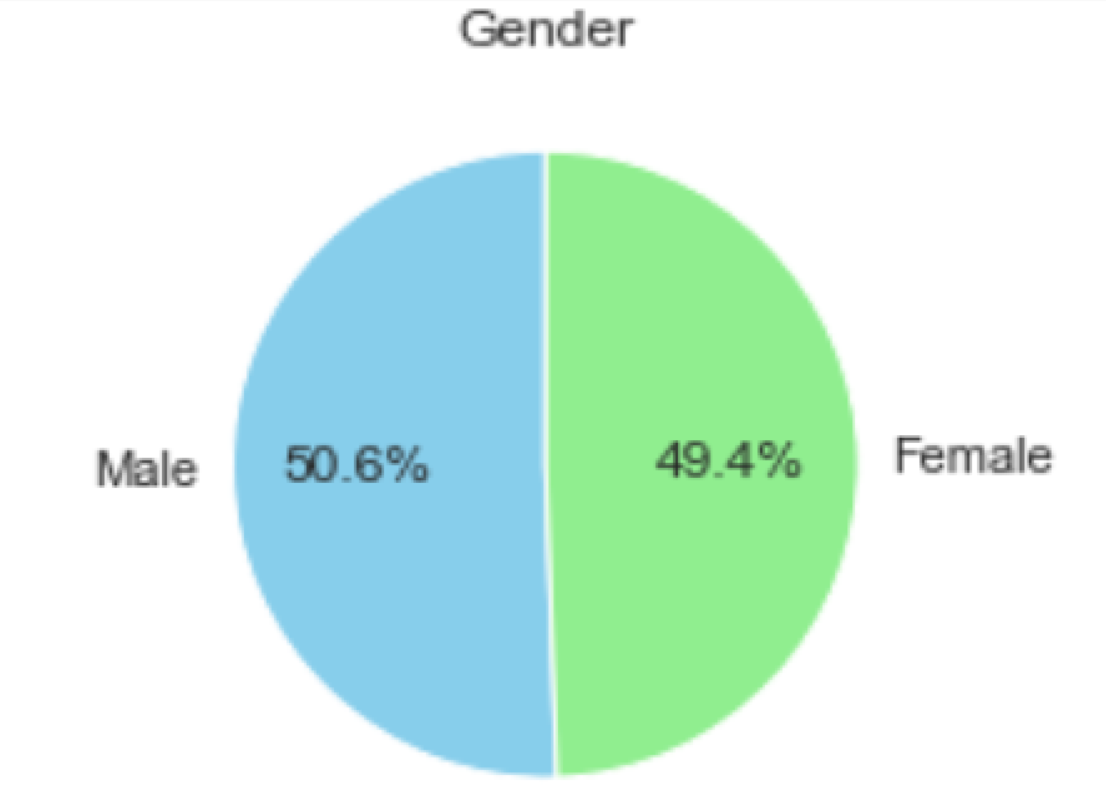
Numerical variables

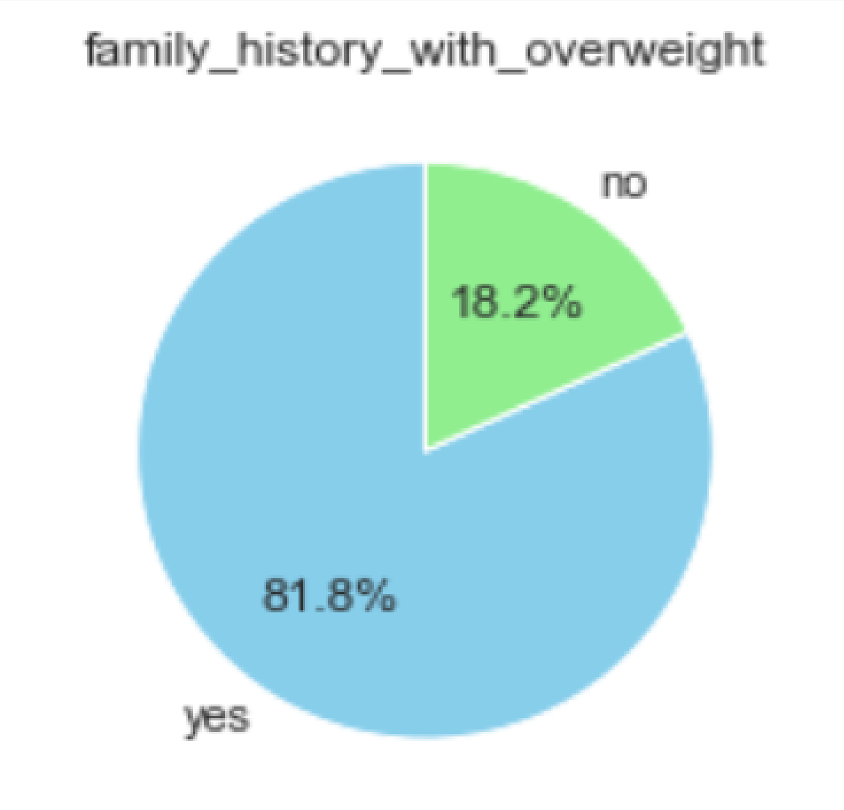
	Age	Height	Weight	Frequency of Consumption of Vegetables (FCVC)	Number of Main Meals (NCP)	Consumption of Water Daily (CH2O)	Physical Activity Frequency (FAF)	Time Using Technology Devices (TUE)
count	2111.00	2111.00	2111.00	2111.00	2111.00	2111.00	2111.00	2111.00
mean	24.31	1.70	86.59	2.42	2.69	2.01	1.01	0.66
std	6.35	0.09	26.19	0.53	0.78	0.61	0.85	0.61
min	14.00	1.45	39.00	1.00	1.00	1.00	0.00	0.00
25%	19.95	1.63	65.47	2.00	2.66	1.58	0.12	0.00
50%	22.78	1.70	83.00	2.39	3.00	2.00	1.00	0.63
75%	26.00	1.77	107.43	3.00	3.00	2.48	1.67	1.00
max	61.00	1.98	173.00	3.00	4.00	3.00	3.00	2.00

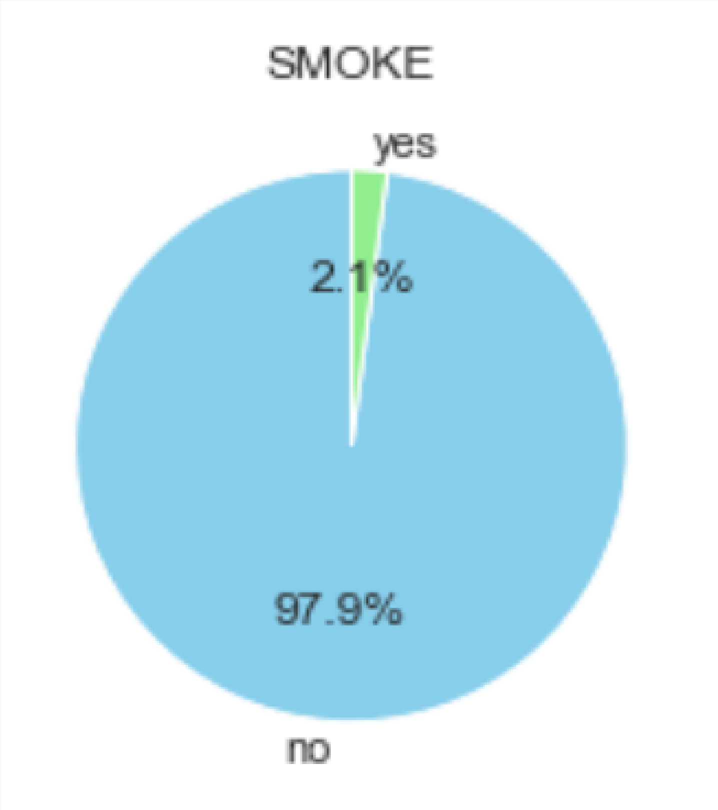
Age, height, weight, and their corresponding behaviors are crucial indicators

Descriptive Statistics

Categorical variables







Gender, family history of overweight, smoking status, and obesity type are crucial indicators

Data Preparation

Data Processing

Preprocessing

1.Check missing data and outliers
2.Standardized numerical features for better model performance

Transformations

Transform height and weight to BMI

Recoding

Apply recoding to convert labels into numerical values(yes:1, no:0)

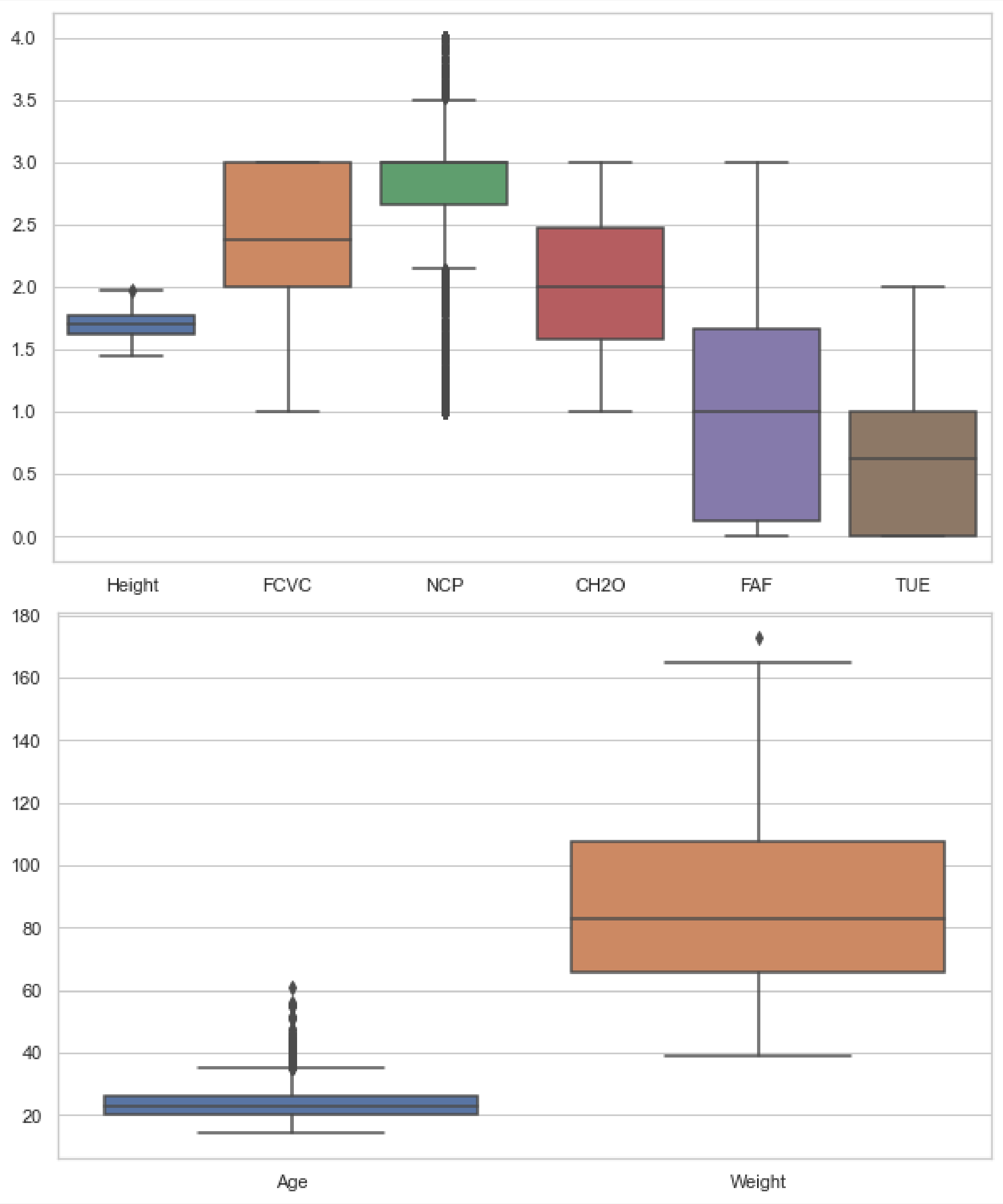
Missing Data

```
print(df.isnull().sum())

Gender                            0
Age                               0
Height                            0
Weight                            0
family_history_with_overweight    0
FAVC                              0
FCVC                              0
NCP                               0
CAEC                              0
SMOKE                             0
CH2O                              0
SCC                               0
FAF                               0
TUE                               0
CALC                              0
MTRANS                            0
NObeyesdad                        0
dtype: int64
```

No missing values

Outliers



No Outliers

Linear Regression

Research Question #1

- **How does method of transportation affect obesity?**

Linear Regression: Methodology

- Our goal: predict an (18-25 year old) individual's Body Mass Index (BMI) based on various predictors that include dietary habits, modes of transportation used, and gender.
- Methodology Steps:
  - Filter to individuals between 18 and 25 years old.
  - Creating dummy variables (diet frequency, high calories food, vegetables).
  - **Dependent variable**: BMI
  - **Independent variables**: Gender, dietary habits (FAVC, FCVC, CAEC, CALC), and transportation mode (MTRANS).
  - **Interaction terms**: Gender -- capture any differing effects these factors have on BMI by gender.

Linear Regression

Without Interaction term

```
Call:
lm(formula = BMI ~ Gender + FAVC + FCVC + CAEC + CALC + MTRANS,
    data = data_18_25)

Residuals:
     Min       1Q   Median       3Q      Max
-18.1656  -4.7749   0.1686   4.6822  25.7046

Coefficients:
                            Estimate Std. Error t value Pr(>|t|)
(Intercept)                  15.0414     7.2864   2.064 0.039195 *
GenderMale                   -0.8147     0.4137  -1.969 0.049127 *
FAVCyes                       3.9305     0.5790   6.788 1.76e-11 ***
FCVC                          3.9083     0.3694  10.581  < 2e-16 ***
CAECFrequently               -5.4390     1.2790  -4.253 2.27e-05 ***
CAECno                        0.2050     1.5639   0.131 0.895724
CAECSometimes                 4.0035     1.2087   3.312 0.000952 ***
CALCFrequently               -5.6041     7.1227  -0.787 0.431551
CALCno                       -6.6481     7.0408  -0.944 0.345237
CALCSometimes                -5.7274     7.0406  -0.813 0.416095
MTRANSBike                    1.1647     3.2136   0.362 0.717093
MTRANSMotorbike               2.6138     2.5829   1.012 0.311760
MTRANSPublic_Transportation   5.0520     0.8631   5.853 6.17e-09 ***
MTRANSWalking                 1.9565     1.5239   1.284 0.199434
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 6.895 on 1242 degrees of freedom
Multiple R-squared:  0.323,     Adjusted R-squared:  0.3159
F-statistic: 45.59 on 13 and 1242 DF,  p-value: < 2.2e-16
```

With Interaction term

```
Call:
lm(formula = Y_train ~ X_train)

Residuals:
     Min       1Q   Median       3Q      Max
-20.9344  -3.6878   0.2412   4.9678  24.9556

Coefficients: (5 not defined because of singularities)
                                            Estimate Std. Error t value Pr(>|t|)
(Intercept)                                  4.47131    2.64652   1.690 0.091439 .
X_train(Intercept)                                NA         NA      NA       NA
X_trainGenderMale                           11.79544    3.67922   3.206 0.001389 **
X_trainGenderFemale:FAVCyes                  4.73876    0.77532   6.112 1.42e-09 ***
X_trainGenderMale:FAVCyes                    3.01837    0.98349   3.069 0.002206 **
X_trainGenderFemale:FCVC                     6.40548    0.51213  12.508  < 2e-16 ***
X_trainGenderMale:FCVC                       0.73230    0.59814   1.224 0.221134
X_trainGenderFemale:CAECFrequently          -8.28890    1.93600  -4.281 2.04e-05 ***
X_trainGenderMale:CAECFrequently            -2.57105    1.88136  -1.367 0.172064
X_trainGenderFemale:CAECno                  -1.10405    2.69132  -0.410 0.681730
X_trainGenderMale:CAECno                    -0.04108    2.08764  -0.020 0.984305
X_trainGenderFemale:CAECSometimes            3.92216    1.88661   2.079 0.037881 *
X_trainGenderMale:CAECSometimes              1.86355    1.68150   1.108 0.268018
X_trainGenderFemale:CALCFrequently          -3.79178    1.71744  -2.208 0.027487 *
X_trainGenderMale:CALCFrequently             1.38713    1.62294   0.855 0.392926
X_trainGenderFemale:CALCno                  -2.63344    0.64106  -4.108 4.32e-05 ***
X_trainGenderMale:CALCno                     2.29682    0.63992   3.589 0.000348 ***
X_trainGenderFemale:CALCSometimes                 NA         NA      NA       NA
X_trainGenderMale:CALCSometimes                   NA         NA      NA       NA
X_trainGenderFemale:MTRANSBike                    NA         NA      NA       NA
X_trainGenderMale:MTRANSBike                 1.31382    3.92547   0.335 0.737930
X_trainGenderFemale:MTRANSMotorbike               NA         NA      NA       NA
X_trainGenderMale:MTRANSMotorbike            2.81164    3.42444   0.821 0.411816
X_trainGenderFemale:MTRANSPublic_Transportation  4.44506    1.67992   2.646 0.008275 **
X_trainGenderMale:MTRANSPublic_Transportation    5.60587    1.09131   5.137 3.37e-07 ***
X_trainGenderFemale:MTRANSWalking            0.52956    2.64802   0.200 0.841535
X_trainGenderMale:MTRANSWalking              2.00827    2.05278   0.978 0.328159
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 6.464 on 985 degrees of freedom
Multiple R-squared:  0.4193,    Adjusted R-squared:  0.4069
F-statistic: 33.86 on 21 and 985 DF,  p-value: < 2.2e-16
```

The Fitted Model

```
Call:
lm(formula = Y_train ~ X_train)

Residuals:
     Min       1Q   Median       3Q      Max 
-20.9344  -3.6878   0.2412   4.9678  24.9556 

Coefficients: (5 not defined because of singularities)
                                       Estimate Std. Error t value Pr(>|t|)    
(Intercept)                             4.47131    2.64652   1.690 0.091439 .  
X_train(Intercept)                           NA         NA      NA       NA    
X_trainGenderMale                      11.79544    3.67922   3.206 0.001389 ** 
X_trainGenderFemale:FAVCyes             4.73876    0.77532   6.112 1.42e-09 ***
X_trainGenderMale:FAVCyes               3.01837    0.98349   3.069 0.002206 ** 
X_trainGenderFemale:FCVC                6.40548    0.51213  12.508  < 2e-16 ***
X_trainGenderMale:FCVC                  0.73230    0.59814   1.224 0.221134    
X_trainGenderFemale:CAECFrequently     -8.28890    1.93600  -4.281 2.04e-05 ***
X_trainGenderMale:CAECFrequently       -2.57105    1.88136  -1.367 0.172064    
X_trainGenderFemale:CAECno             -1.10405    2.69132  -0.410 0.681730    
X_trainGenderMale:CAECno               -0.04108    2.08764  -0.020 0.984305    
X_trainGenderFemale:CAECSometimes       3.92216    1.88661   2.079 0.037881 *  
X_trainGenderMale:CAECSometimes         1.86355    1.68150   1.108 0.268018    
X_trainGenderFemale:CALCFrequently     -3.79178    1.71744  -2.208 0.027487 *  
X_trainGenderMale:CALCFrequently        1.38713    1.62294   0.855 0.392926    
X_trainGenderFemale:CALCno             -2.63344    0.64106  -4.108 4.32e-05 ***
X_trainGenderMale:CALCno                2.29682    0.63992   3.589 0.000348 ***
X_trainGenderFemale:CALCSometimes            NA         NA      NA       NA    
X_trainGenderMale:CALCSometimes              NA         NA      NA       NA    
X_trainGenderFemale:MTRANSBike               NA         NA      NA       NA    
X_trainGenderMale:MTRANSBike            1.31382    3.92547   0.335 0.737930    
X_trainGenderFemale:MTRANSMotorbike          NA         NA      NA       NA    
X_trainGenderMale:MTRANSMotorbike       2.81164    3.42444   0.821 0.411816    
X_trainGenderFemale:MTRANSPublic_Transportation  4.44506    1.67992   2.646 0.008275 ** 
X_trainGenderMale:MTRANSPublic_Transportation    5.60587    1.09131   5.137 3.37e-07 ***
X_trainGenderFemale:MTRANSWalking       0.52956    2.64802   0.200 0.841535    
X_trainGenderMale:MTRANSWalking         2.00827    2.05278   0.978 0.328159    
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 6.464 on 985 degrees of freedom
Multiple R-squared:  0.4193,	Adjusted R-squared:  0.4069 
F-statistic: 33.86 on 21 and 985 DF,  p-value: < 2.2e-16
```

- Our model is statistically significant:
  - F-statistic: 33.86
  - p-value: < 0.01, very small
- R-squared: 0.419
- Adjusted R-squared:0.4069:
  - This is a moderate amount.
  - While our model captures 41% of the variability in BMI, there is still a significant amount of variability that is not explained by the model.
- Our independent variables explain a reasonable proportion of the variability in BMI.
- There is room for improvement.

Analyzing Results

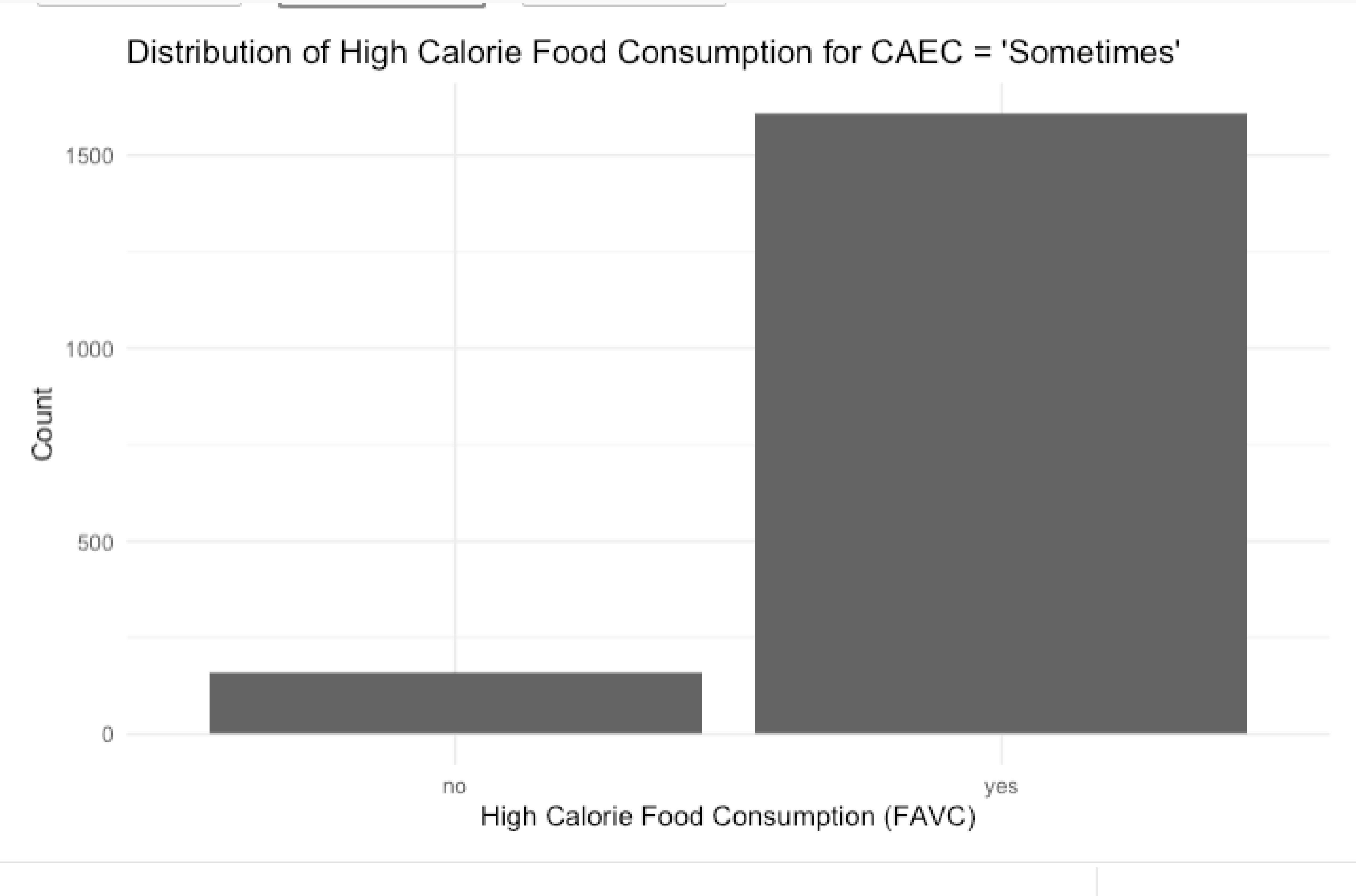
```
Call:
lm(formula = Y_train ~ X_train)

Residuals:
     Min       1Q   Median       3Q      Max 
-20.9344  -3.6878   0.2412   4.9678  24.9556 

Coefficients: (5 not defined because of singularities)
                                           Estimate Std. Error t value Pr(>|t|)    
(Intercept)                                 4.47131    2.64652   1.690 0.091439 .  
X_train(Intercept)                               NA         NA      NA       NA    
X_trainGenderMale                          11.79544    3.67922   3.206 0.001389 ** 
X_trainGenderFemale:FAVCyes                 4.73876    0.77532   6.112 1.42e-09 ***
X_trainGenderMale:FAVCyes                   3.01837    0.98349   3.069 0.002206 ** 
X_trainGenderFemale:FCVC                    6.40548    0.51213  12.508  < 2e-16 ***
X_trainGenderMale:FCVC                      0.73230    0.59814   1.224 0.221134    
X_trainGenderFemale:CAECFrequently         -8.28890    1.93600  -4.281 2.04e-05 ***
X_trainGenderMale:CAECFrequently           -2.57105    1.88136  -1.367 0.172064    
X_trainGenderFemale:CAECno                 -1.10405    2.69132  -0.410 0.681730    
X_trainGenderMale:CAECno                   -0.04108    2.08764  -0.020 0.984305    
X_trainGenderFemale:CAECSometimes           3.92216    1.88661   2.079 0.037881 *  
X_trainGenderMale:CAECSometimes             1.86355    1.68150   1.108 0.268018    
X_trainGenderFemale:CALCFrequently         -3.79178    1.71744  -2.208 0.027487 *  
X_trainGenderMale:CALCFrequently            1.38713    1.62294   0.855 0.392926    
X_trainGenderFemale:CALCno                 -2.63344    0.64106  -4.108 4.32e-05 ***
X_trainGenderMale:CALCno                    2.29682    0.63992   3.589 0.000348 ***
X_trainGenderFemale:CALCSometimes                NA         NA      NA       NA    
X_trainGenderMale:CALCSometimes                  NA         NA      NA       NA    
X_trainGenderFemale:MTRANSBike                   NA         NA      NA       NA    
X_trainGenderMale:MTRANSBike                1.31382    3.92547   0.335 0.737930    
X_trainGenderFemale:MTRANSMotorbike              NA         NA      NA       NA    
X_trainGenderMale:MTRANSMotorbike           2.81164    3.42444   0.821 0.411816    
X_trainGenderFemale:MTRANSPublic_Transportation  4.44506    1.67992   2.646 0.008275 ** 
X_trainGenderMale:MTRANSPublic_Transportation   5.60587    1.09131   5.137 3.37e-07 ***
X_trainGenderFemale:MTRANSWalking           0.52956    2.64802   0.200 0.841535    
X_trainGenderMale:MTRANSWalking             2.00827    2.05278   0.978 0.328159    
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 6.464 on 985 degrees of freedom
Multiple R-squared:  0.4193,	Adjusted R-squared:  0.4069 
F-statistic: 33.86 on 21 and 985 DF,  p-value: < 2.2e-16
```

1. **Gender**: In our data, Male have approximately -0.81units lower BMI than female, if other variable stay the same.
2. **Dietary Habits**: (High-calorie food): For both genders, high calories food lead to significant increase in BMI.
3. **Dietary Habits:** Food Between Meals (CAEC): People who frequently eat between meals have lower BMI than those who eat sometimes between meals.
4. **Alcohol Consumption** (CALC): Drinking alcohol leads to significant increases in BMI.
5. **Transportation Modes** (MTRANS): Biking and Walking leads to significant lower BMI.


Distribution of High Calorie Food Consumption for CAEC = 'Sometimes'
Count
1500
1000
500
0
no
yes
High Calorie Food Consumption (FAVC)

Linear Regression: Takeaways

Conclusions:

- The model indicates that certain behaviors, such as frequent high-calorie food consumption and alcohol use, are positively associated with higher BMI, while frequent vegetable consumption and active transportation modes like biking and walking are associated with lower BMI, particularly in females.

Shortcomings:

- **Overfitting:** We considered the potential for overfitting due to including many predictors relative to the number of observations.
- **Causality:** These results should be taken as associations that warrant further investigation rather than direct causation.

Logistic Regression

Research Question #2

- **How does diet affect obesity?**

Logistic Regression: Methodology

- Logistic Regression: Utilizing multi-class classification, we predict obesity levels based on personal habits and traits, while also assessing the correlation between specific lifestyle factors and the identified obesity categories.
- Methodology Steps:
  - Drop the variables that intuitively correlate with BMI e.g. Height, Weight.
  - Classify obesity levels into three categories: Normal Weight (1), Overweight and Obesity (2), and Insufficient Weight (3)
  - Apply standardization to the training and testing datasets.
  - Fit the model.

The result of fitting model

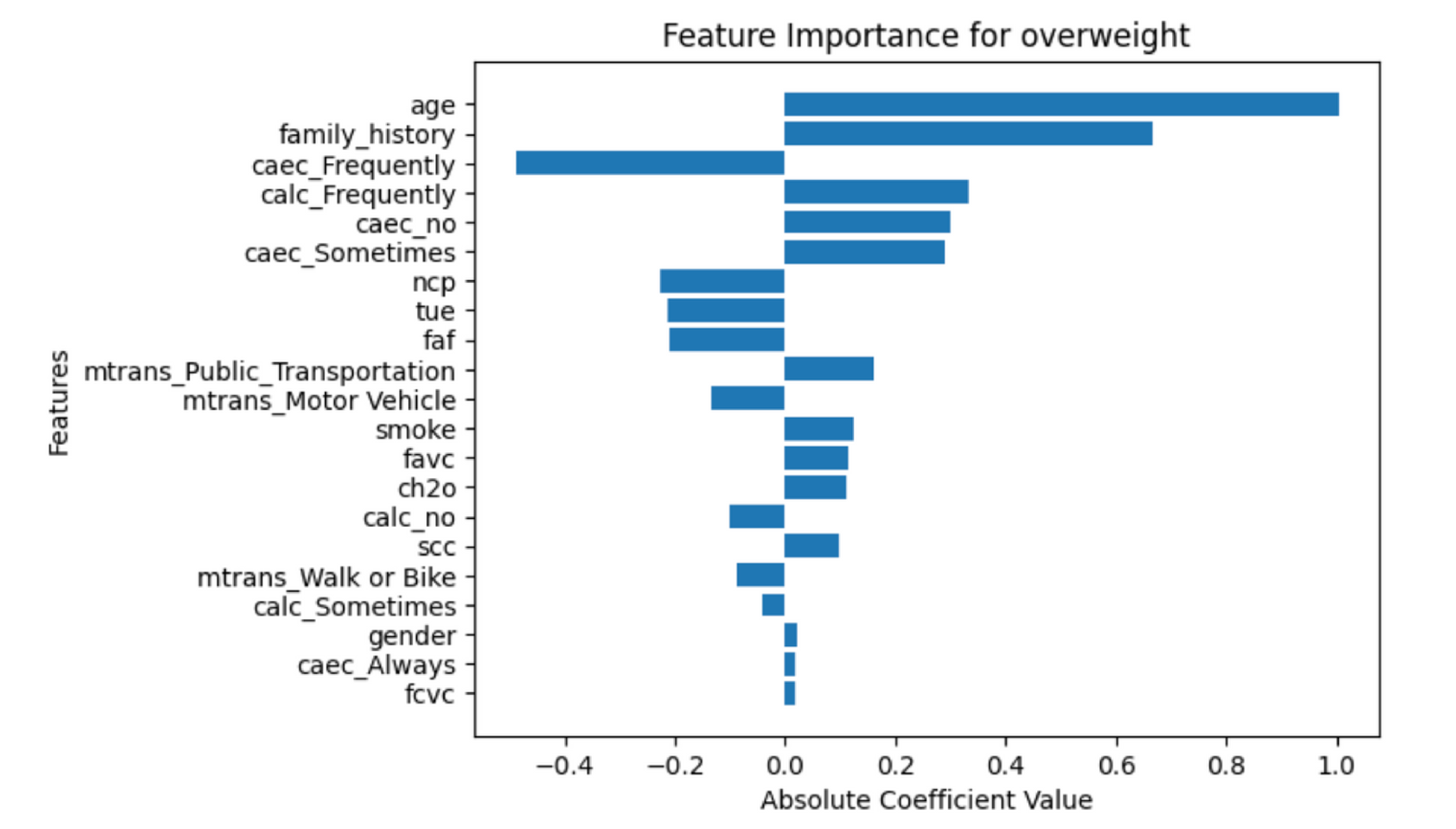
Accuracy: 0.8390532544378698

Classification Report:

	precision	recall	f1-score	support
(Normal Weight: 18.5<BMI<24.9) 1	0.64	0.37	0.46	115
(Overweight: BMI>25) 2	0.90	0.95	0.92	621
(Insufficient Weight: BMI<18.5) 3	0.62	0.72	0.67	109
accuracy			0.84	845
macro avg	0.72	0.68	0.68	845
weighted avg	0.83	0.84	0.83	845

The model performs well in predicting overweight, but shows comparatively lower performance in predicting normal weight and insufficient weight

Feature Importance - Overweight



The variables have most significant impact:

(1) Age

(2) Family History

(3) Diet Variables - Calc
(Consumption of alcohol)
Frequent consumption of alcohol is the third most positively associated feature.

(4) Diet Variables - Caec
(Consumption of food between meals)
The specific frequency of consuming food between meals plays a significant role.

Logistic Regression: Takeaways

Conclusions:

- Age and family history persist as crucial determinants of higher BMI.
- Among diet variables, the consumption of alcohol and the frequency of consuming food between meals stand out as notably impactful compared to other dietary factors.

Shortcomings:

- Difficulty in accurately predicting levels of overweight (7 levels) with an accuracy of only 0.6, leading to the decision to collapse them into 3 levels.
- Challenges in distinguishing between insufficient weight and normal weight. While overweight can be easily discerned behaviorally, accurately identifying underweight poses more challenges.

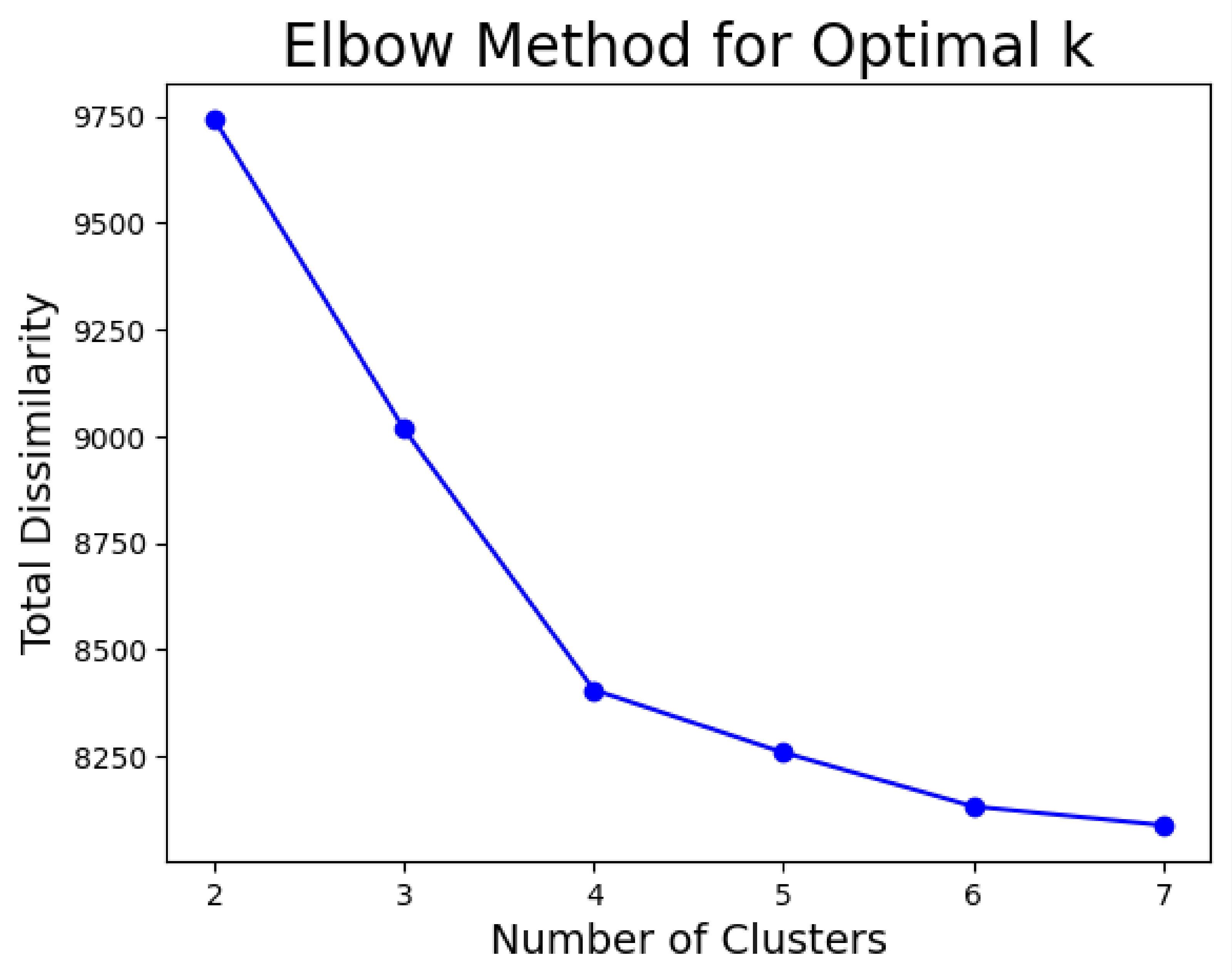
K-Modes Clustering

Research Question #3

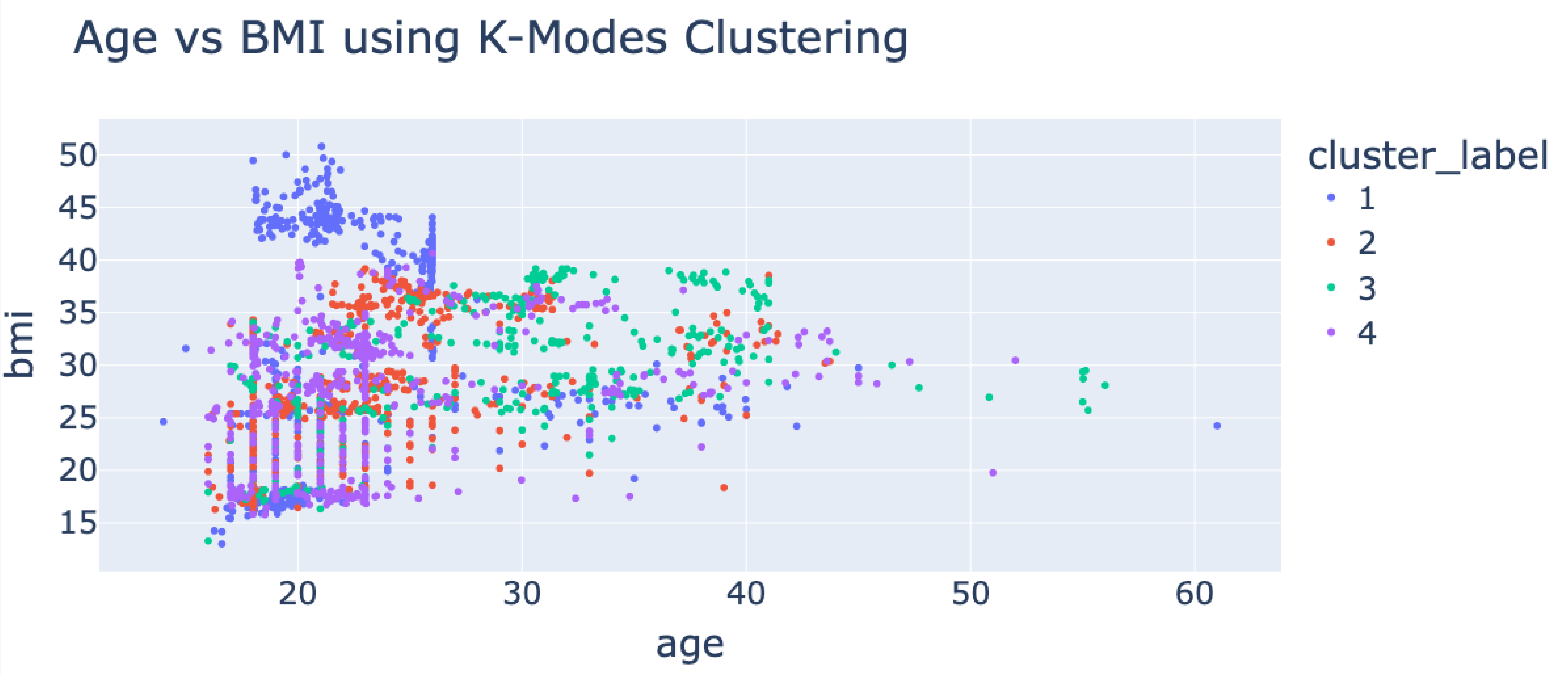
- How does alcohol and smoking affect obesity?

K-Modes: Methodology

- K-Modes: Cousin of K-Means that allows for categorial variables
- Only use behavioral variables e.g. alcohol consumption
    - Including BMI would restate the obvious
    - Exclude gender and age
- Optimal Number of Clusters: Elbow Method
    - What’s the best number of clusters?

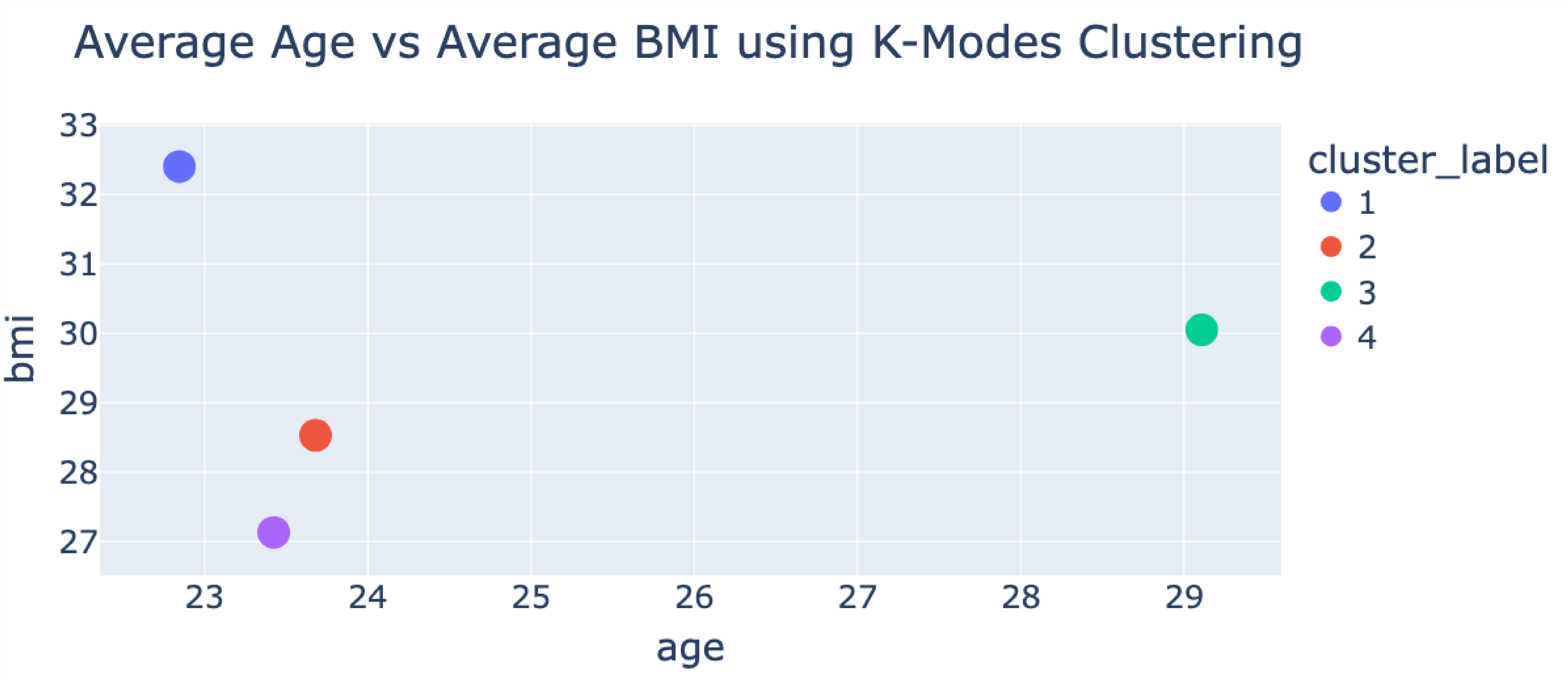


K-Modes: Age vs BMI



Not very useful...

K-Modes: Age vs BMI



K-Modes: Cluster 1&2

Cluster 1: Party Girls

- Avg. BMI: 32.25
- % BMI >30: 55%
- Avg. Age: 22
- Gender:
  - Female: 87%
  - Male: 13%
- Consumption of High Calorie Foods:
  - Yes: 90%
- Consumption of Alcohol:
  - Sometimes or More: 94% (2nd Highest)
- Smoker?:
  - Yes: 1.3%
- Frequency of Physical Activity:
  - 0.85 (Lowest of all groups)

Cluster 2: Party Bros

- Avg. BMI: 27.94
- % BMI >30: 40%
- Avg. Age: 23
- Gender:
  - Female: 21%
  - Male: 79%
- Consumption of High Calorie Foods:
  - Yes: 92%
- Consumption of Alcohol:
  - Sometimes or More: 95% (Highest)
- Smoker?
  - Yes: 3.8% (Highest)
- Frequency of Physical Activity:
  - 1.00

K-Modes: Cluster 3&4

Cluster 3: Late 20s

- Avg. BMI: 29.89
- % BMI >30: 49%
- Avg. Age: 29
- Gender:
  - Female: 10%
  - Male: 90%
- Consumption of High Calorie Foods:
  - Yes: 89%
- Consumption of Alcohol:
  - Sometimes or More: 78%
- Smoker?:
  - Yes: 2.1%
- Frequency of Physical Activity:
  - 1.19 (Highest of all groups)

Cluster 4: Sober Sally

- Avg. BMI: 27.99
- % BMI >30: 38%
- Avg. Age: 23
- Gender:
  - Female: 59%
  - Male: 41%
- Consumption of High Calorie Foods:
  - Yes: 81% (Lowest of all groups)
- Consumption of Alcohol:
  - Sometimes or More: 5% (Lowest of all)
- Smoker?
  - Yes: 1.1%
- Frequency of Physical Activity:
  - 1.08

K-Modes: Takeaways

Conclusions:

- Drinking was biggest factor separating clusters.
    - Not drinking is related with lower BMI.
- Relatively small proportion of the population smoke - not a major factor.
- Age - older people had higher BMIs even with more physical activity.
- Among clusters that did drink:
    - Males had more physical activity and lower BMI.
    - Females had less physical activity and higher BMI.

Shortcomings of this Model:

- Not causal - potential for omitted variable bias / confounding variables.
- Gender and Age bias despite them not being explicitly included.

Conclusion and Policy Recommendations

Conclusions

- Linear Regression reveals significant gender disparities in behaviors that affect BMI, with females more sensitive to diet and males more sensitive to transportation method.
- High-calorie food consumption is identified as a major contributor to increased BMI across all genders.
- Clustering analysis identified that drinking habits were a significant factor distinguishing lifestyle clusters. Non-drinkers tended to have lower BMI, emphasizing the role of alcohol consumption in obesity.
- Diet variables like alcohol consumption and the frequency of eating between meals are identified as significant contributors in determining high BMI levels.

Policy Recommendations

- **Public Awareness of Dietary Choices:**
  - Implement policies and programs that encourage and facilitate healthy food choices and access, such as reducing the availability and affordability of high-calorie foods and beverages, increasing the consumption of fruits and vegetables, and providing nutrition labeling and guidance.
- **Transportation Modes and Physical Activity:**
  - Implement policies and programs that encourage and facilitate physical activity and active transportation, such as creating safe and accessible environments for walking and cycling, providing incentives and subsidies for public transit and bike-sharing, and integrating physical activity into school and work settings.

Policy Recommendations

- **Cluster-Based Health Interventions:**
  - Implement cluster-specific health interventions considering the identified clusters' characteristics. Provide screening and counseling services for individuals at high risk of obesity, and offer personalized and tailored interventions that address their specific needs and preferences, such as dietary advice, behavioral therapy, and pharmacological or surgical treatment.
- **Alcohol Education and Regulation:**
  - Implement educational programs on the health risks associated with excessive alcohol consumption. Consider regulatory measures, such as alcohol taxation and advertising restrictions, to curb unhealthy drinking habits and reduce obesity rates.

Thank You!
Any Questions?