Final Project

Obesity Data Analysis

Ilseop Lee, Ramil Mammadov, Tursunai Turumbekova, Yirang Liu

Abstract

Understanding the complex interactions between genetic predispositions and lifestyle factors is essential for developing effective and personalized obesity prevention strategies. This study examines two aspects of obesity: (1) the interaction between family history of overweight and physical activity frequency on Body Mass Index (BMI), and (2) the combined effects of meal frequency and alcohol consumption on obesity levels. Using a dataset from Mexico, Peru, and Colombia, multiple linear regression (MLR) and multinomial logistic regression models were applied. The results reveal that physical activity significantly reduces BMI for individuals with a family history of overweight, highlighting the importance of personalized weight management strategies. Additionally, frequent meals combined with alcohol consumption were strongly associated with higher obesity levels, particularly severe obesity, offering actionable insights for targeted interventions.

Introduction

Obesity, a global health issue, has nearly tripled since 1975, with over 650 million adults classified as obese in 2016 (WHO). This epidemic not only imposes significant economic burdens through rising healthcare costs but also highlights the critical role of lifestyle behaviors, such as diet and physical activity, as primary drivers. These factors often interact with genetic predispositions like family history, while additional behaviors, such as alcohol consumption, can exacerbate dietary effects on weight gain.

Given the complex interplay of genetic and lifestyle factors, this study investigates two key aspects of obesity: (1) the interaction between family history of overweight and physical activity frequency on Body Mass Index (BMI), and (2) the combined effects of meal frequency and alcohol consumption on obesity levels. By addressing these questions, the study aims to uncover insights that can inform personalized obesity prevention and intervention strategies tailored to individual genetic and lifestyle factors.

Methods

Setting and Study Population

The dataset consists of 2,111 records from individuals in Mexico, Peru, and Colombia, representing demographic, dietary, and lifestyle factors. Of these, 23% were collected via a web platform, while 77% were synthetically generated using the SMOTE algorithm to balance obesity levels. The target variable categorizes individuals into seven obesity levels. Initial inspection revealed **no missing values** for critical variables, and checks for outliers and influential points were conducted to ensure robust results.

Conceptual Model

This study investigates two key research questions: (1) How does the interaction between family history of overweight and physical activity frequency influence BMI? (2) How do meal frequency and alcohol consumption interact to affect obesity levels? For Objective 1, BMI is the outcome variable, with family history of overweight and physical activity frequency (FAF) as explanatory variables, including their interaction term. For Objective 2, NObeyesdad (seven obesity levels) is the outcome, with alcohol consumption frequency (CALC), meal frequency (NCP), and their interaction term as predictors.

Statistical Analysis

To address the research objectives, multiple linear regression (MLR) and multinomial logistic regression models were applied. **For Objective 1**, MLR was used to predict BMI based on family history of overweight, frequency of physical activity, and their interaction term. **For Objective 2**, multinomial logistic regression modeled obesity levels (categorized into seven stages) using alcohol consumption frequency, daily meal frequency, and their interaction term as predictors.

Exploratory Data Analysis (EDA) was performed to summarize variable distributions and relationships. Continuous variables (e.g., BMI, frequency of physical activity, daily meal frequency) were visualized with histograms and boxplots, while categorical variables (e.g., family history of overweight, alcohol consumption frequency, obesity levels) were summarized using bar charts. Interaction effects were analyzed using scatterplots and interaction plots.

Model diagnostics included residual analysis for MLR to assess linearity, normality, and homoscedasticity, with multicollinearity addressed by excluding predictors with Variance Inflation Factor (VIF) values greater than 10. For multinomial logistic regression, model fit was evaluated using the Proportional Odds Assumption (POA), Akaike Information Criterion (AIC), and accuracy scores. Due to insufficient observations in certain alcohol consumption categories, the original four levels (Always, Frequently, Sometimes, Never) were consolidated into two groups: “Yes” (Always, Frequently, Sometimes) and “No” (Never) to enable valid analysis of interaction effects.

Results

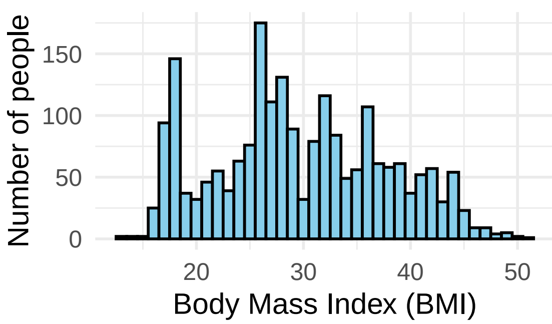
First we will present results and discussion of research question 1 and then for the research question 2.

1. Research Question 1: How family history of overweight influence the relationship between physical activity and BMI?

BMI Distribution

The distribution of BMI is typically right-skewed, as most individuals in the dataset are likely overweight or obese. The median BMI is around 29.7, which is consistent with a population that has a higher proportion of overweight and obese individuals. The distribution of BMI shows a peak around 25–30, which is categorized as overweight. This supports the general notion that the population has a significant proportion of individuals at risk for obesity-related health conditions. The histogram illustrates the distribution of BMI, highlighting the concentration of individuals with BMI values above 25.

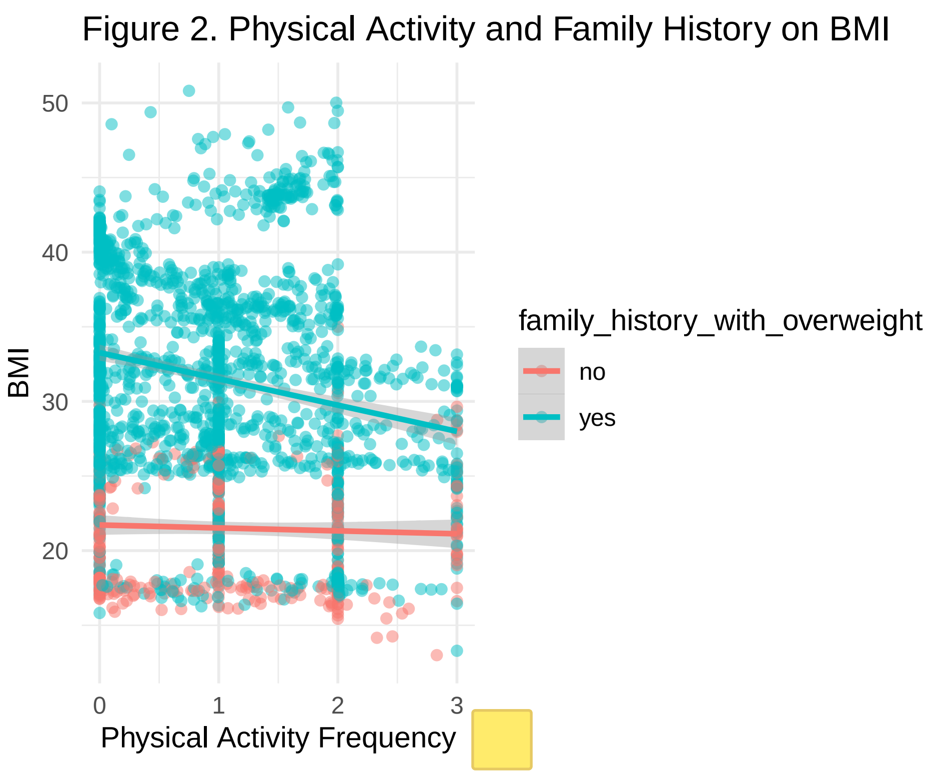
Figure 1. BMI Distribution



Physical Activity and BMI

The scatter plot explores how physical activity and family history interact to affect BMI. The relationship suggests that individuals with a family history of overweight tend to have higher BMIs even with increased physical activity. The interaction between family history and physical activity suggests that genetic factors may limit the effectiveness of physical activity on BMI reduction. Individuals without a family history of overweight show a more significant decrease in BMI with higher physical activity.

```
`geom_smooth()` using formula = 'y ~ x'
```



Multiple linear regression model: BMI

The multiple linear regression (MLR) model was fitted to predict BMI, including the main effects of family history and physical activity frequency, as well as their interaction term. The model included 2,111 individuals after exclusions, results are summarized below:

Predictor	Estimate	Std. Error	t-value	p-value
Intercept	21.7192	0.5489	39.572	<0.001
Family History (Yes)	11.5455	0.6063	19.041	<0.001
Physical Activity Frequency (FAF)	-0.1966	0.3792	-0.519	0.604
Interaction (Family History * FAF)	-1.5608	0.4285	-3.642	<0.001

Interpretation of Results

The regression analysis provided several insights. The predicted BMI for individuals with no family history of overweight and no physical activity was 21.72. Individuals with a family history of overweight had a BMI that was, on average, 11.55 units higher than those without a family history when physical activity was absent. For individuals without a family history, the coefficient for physical activity frequency was not statistically significant ($p = 0.604$), suggesting that physical activity frequency alone does not have a strong effect on BMI.

The interaction between family history of overweight and physical activity frequency indicates that the effect of physical activity on BMI is not consistent across groups. For individuals without a family history of overweight, the effect of physical activity on BMI is small and not statistically significant. However, for individuals with a family history of overweight, physical activity has a stronger, statistically significant effect, reducing BMI by 1.56 units per additional day of physical activity per week. This suggests that physical activity is particularly effective in mitigating the influence of genetic or familial predispositions on BMI.

Model Fit and Diagnostics

The model explained 26.1% of the variability in BMI (Multiple R-squared = 0.261), suggesting that other factors influence BMI. The adjusted R-squared of 0.26 means the model explains 26.1% of the variability in BMI, which is moderate but suggests there are other factors affecting BMI not included in this model. This aligns with the result we got from the comprehensive model with all the predictors included. This is also increased compared with the model not with the interaction term, which means the interaction term is contributing to the overall model.Adjusted R-squared accounts for the number of predictors. It is very close to the Multiple R-squared, indicating the model isn't overfitting. This is also increased compared with the model not with the interaction term, which means the interaction term is contributing to the overall model. The overall model was statistically significant (F-statistic = 248.1, $p < 2.2e\text{-}16$), meaning at least one predictor or the interaction term was significantly associated with BMI.

Discussion

The analysis demonstrates that family history is a significant predictor of BMI, with individuals who have a family history of overweight having substantially higher BMIs on average. Physical activity has a stronger BMI-reducing effect for individuals with a family history of overweight, suggesting that targeted interventions could be particularly effective for this group. However, physical activity alone does not appear to have a significant effect on BMI for individuals without a family history of overweight.

Family History of Overweight is the strongest predictor in the model, showing a large positive association with BMI. By itself, physical activity frequency has no significant effect on BMI for those without a family history of overweight. However, when combined with family history, physical activity is significantly associated with reducing BMI. The results suggest that tailored interventions promoting physical activity for individuals with a family history of overweight could be more effective in reducing BMI. These findings emphasize the need for personalized approaches to weight management based on genetic and familial predispositions.

Limitations

The model explains only 26.1% of the variability in BMI, suggesting that other factors such as diet, metabolic rate, or socioeconomic status may play significant roles. The dataset's cross-sectional nature limits causal interpretations of the relationships. As the dataset is synthetic, findings may not generalize to real-world populations without further validation.

Findings

This analysis highlights the importance of family history as a predictor of BMI and the potential for physical activity to mitigate its effects. While the model captures meaningful relationships, future research should explore additional predictors and validate findings in real-world datasets. Tailored interventions addressing genetic predispositions and promoting physical activity could significantly reduce BMI and associated health risks.

2. Research Question 2: How meal frequency and alcohol consumption interact to influence obesity levels?

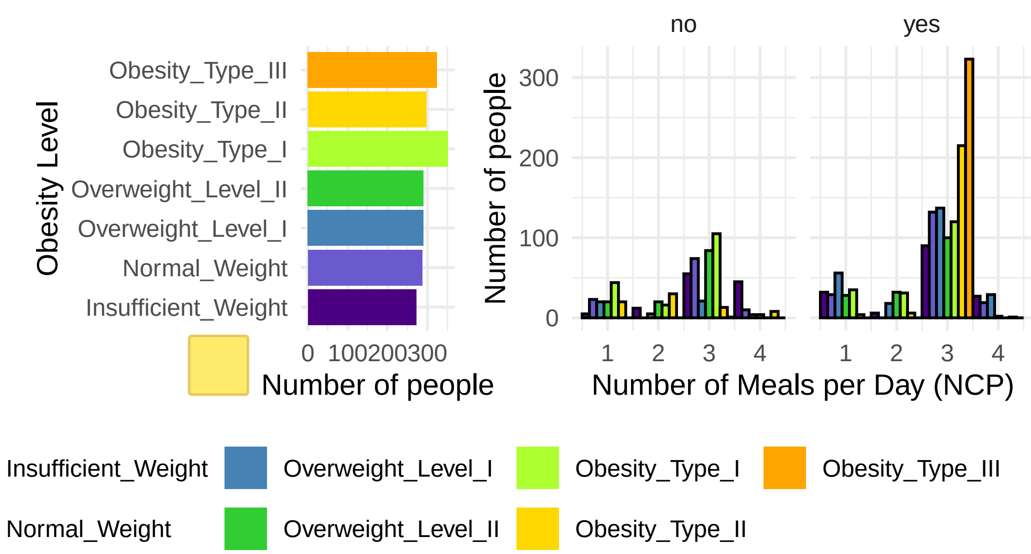
Obesity Levels Distribution and Influencing Factors

This chart highlights the relationship between obesity levels, meal frequency, and alcohol consumption. The first plot shows a high concentration of individuals in Obesity Type I and Obesity Type III, while Normal Weight also makes up a notable portion, reflecting population diversity.

The second plot reveals that consuming three meals per day is the most common pattern across both alcohol consumption groups, indicating that meal frequency plays a stronger role in obesity than alcohol use. However, alcohol appears to amplify obesity levels, as more individuals in the alcohol consumption group fall into Obesity Type II and Obesity Type III. To address the small sample size in certain alcohol consumption categories (e.g., n = 1 for "Always"), these groups were consolidated into two: Yes (Always, Frequently, Sometimes) and No (Never). This adjustment improved analysis robustness, with 1,472 individuals in the "Yes" group and 639 in the "No" group.

This visualization clearly demonstrates how lifestyle factors influence obesity levels while addressing data limitations.

Figures 3 & 4: Obesity Levels and Meal Frequency by Alcohol Use



Multinomial Logistic Regression model: Obesity level

Based on a comprehensive comparison of AIC, POC, and Accuracy, the Multinomial Logistic Regression model was determined to be the most suitable for this study (refer to the Appendix for details). The model was evaluated to explain factors influencing obesity levels (1 to 7 ordinal scale), focusing on meal frequency and alcohol consumption. Consequently, Multinomial Logistic Regression was adopted.

The confusion matrix analysis demonstrated that the model performs well overall, accurately predicting many cases within each obesity category. While there is some overlap between adjacent categories such as **Overweight_Level_I** and **Overweight_Level_II**, this is expected due to the inherent similarities between these levels.

```
# weights:  35 (24 variable)
initial  value 4107.816325
iter  10 value 3905.938324
iter  20 value 3824.810564
iter  30 value 3812.763017
final  value 3812.612859
converged
```

[Table : Multinomial Logistic Regression Confusion Matrix: Actual vs. Predicted Obesity Categories]

Actual / Predicted	Insufficient _Weight	Normal _Weight	Overweight _Level_I	Overweight _Level_II	Obesity _Type_I	Obesity _Type_II	Obesity _Type_III
Insufficient_Weight	57	0	0	0	0	0	38
Normal_Weight	0	97	10	0	0	0	29
Overweight_Level_I	0	0	151	5	0	0	74
Overweight_Level_II	0	0	0	166	10	0	61
Obesity_Type_I	0	0	0	0	120	57	45
Obesity_Type_II	0	0	0	0	0	214	12
Obesity_Type_III	0	0	0	0	0	0	323

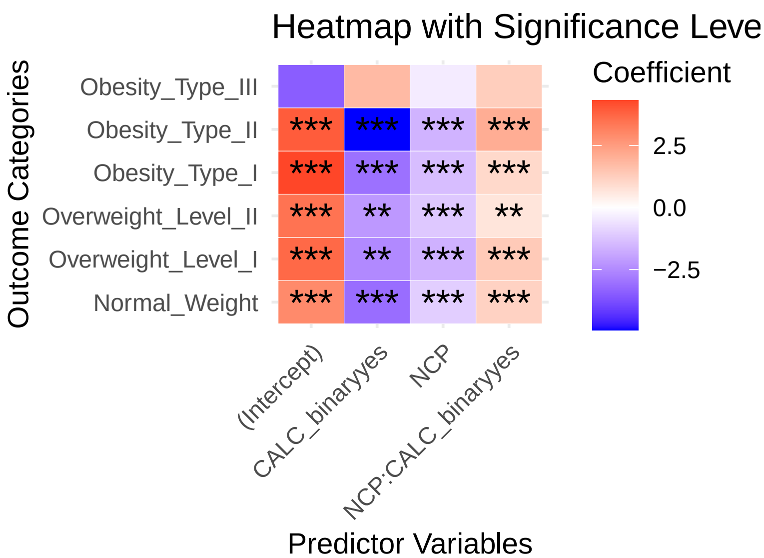
Discussion

In most categories, an increase in meal frequency generally led to a decrease in obesity probabilities. This may be attributed to the idea that those who are health-conscious tend to consume smaller meals more frequently, potentially preventing overeating. Moreover, for individuals who consume alcohol with meals, a statistically significant positive relationship was observed across most categories. Notably, the likelihood of falling into Obesity Level II increased significantly ($P < 0.01$) with an estimate of 2.093.

[Table : Obesity Outcomes by Alcohol and Meal Frequency Interaction : Reference category = Insufficient Weight]

Variable	Outcome	Estimate	StdError	ZValue	PValue
(Intercept)	Normal_Weight	2.9385147	0.6399728	4.5916245	4.40E-06
(Intercept)	Overweight_Level_I	3.6820697	0.675704	5.4492349	5.06E-08
(Intercept)	Overweight_Level_II	3.456452	0.6221582	5.555584	2.77E-08
(Intercept)	Obesity_Type_I	4.3182074	0.6054174	7.1326117	9.85E-13
(Intercept)	Obesity_Type_II	3.9410635	0.6477044	6.0846639	1.17E-09
(Intercept)	Obesity_Type_III	-3.4728078	4.6596406	-0.7452952	4.56E-01
CALC_yes	Normal_Weight	-3.1275223	0.7878578	-3.9696533	7.20E-05
CALC_yes	Overweight_Level_I	-2.5146242	0.7803434	-3.2224584	1.27E-03
CALC_yes	Overweight_Level_II	-2.1851581	0.7400744	-2.9526197	3.15E-03
CALC_yes	Obesity_Type_I	-3.0461263	0.7232943	-4.2114618	2.54E-05
CALC_yes	Obesity_Type_II	-4.9392017	0.813621	-6.0706419	1.27E-09
CALC_yes	Obesity_Type_III	1.763997	4.6872293	0.3763411	7.07E-01
NCP	Normal_Weight	-1.0293405	0.2093586	-4.9166388	8.80E-07
NCP	Overweight_Level_I	-1.6484233	0.2381772	-6.9209953	4.48E-12
NCP	Overweight_Level_II	-1.1586973	0.2037731	-5.6862127	1.30E-08
NCP	Obesity_Type_I	-1.4038596	0.1990448	-7.0529823	1.75E-12
NCP	Obesity_Type_II	-1.6066491	0.2228619	-7.2091688	5.63E-13

Variable	Outcome	Estimate	StdError	ZValue	PValue
NCP	Obesity_Type_III	-0.4179549	1.5158376	-0.2757254	7.83E-01
NCP:CALC_yes	Normal_Weight	1.1521824	0.2646998	4.3527891	1.34E-05
NCP:CALC_yes	Overweight_Level_I	1.3731994	0.2765993	4.9645794	6.89E-07
NCP:CALC_yes	Overweight_Level_II	0.6851342	0.2512175	2.727255	6
NCP:CALC_yes	Obesity_Type_I	0.9862428	0.2457876	4.0125822	6.01E-05
NCP:CALC_yes	Obesity_Type_II	2.0930359	0.2804061	7.4643021	8.37E-14
NCP:CALC_yes	Obesity_Type_III	1.2653093	1.5257052	0.8293275	4.07E-01



Limitation

If a sufficient number of observations were available for the detailed subcategories of CALC, a more granular analysis could have been conducted, providing deeper insights into the relationship between alcohol consumption and weight categories. Additionally, incorporating a measure of the quantity of food consumed would allow for a more precise investigation of meal frequency's impact on weight outcomes. Furthermore, changing the reference category in the multinomial logistic regression model to Normal_Weight could offer a clearer interpretation of the relative effects of other weight categories, although this approach has not yet been attempted.

Findings

The alcohol consumption appears to influence weight in complex ways. Individuals who consume alcohol tend to have a lower likelihood of maintaining a normal weight or being slightly overweight. However, for those with higher levels of obesity, alcohol consumption may exacerbate weight gain. Meal frequency also plays a significant role in weight management. Eating fewer meals per day is associated with a higher likelihood of maintaining a normal weight and a reduced risk of becoming overweight. Conversely, frequent meals significantly increase the risk of obesity. The combined effect of alcohol consumption and frequent meals is particularly impactful, dramatically increasing the risk of severe obesity. Together, these two factors have a synergistic effect, contributing more strongly to weight gain than either factor alone.

Conclusion

This study underscores the multifaceted nature of obesity, highlighting the critical interactions between genetic predisposition and lifestyle behaviors. The findings reveal that family history of overweight is a significant predictor of BMI, with physical activity playing a vital role in mitigating its impact. Specifically, individuals with a family history of overweight benefit significantly from increased physical activity, demonstrating the value of tailored interventions that account for genetic factors. However, for those without such a family history, the relationship between physical activity and BMI is less pronounced, suggesting that other factors may drive weight outcomes in this group.

Additionally, the interaction between meal frequency and alcohol consumption was found to significantly influence obesity levels. While consuming fewer meals per day is associated with maintaining a normal weight, frequent meals combined with alcohol consumption substantially increase the risk of severe obesity. This highlights the compounded effects of certain lifestyle behaviors and the need for integrated strategies that address multiple contributing factors simultaneously.

The findings emphasize the importance of personalized approaches to obesity prevention and management. Public health strategies should focus on promoting physical activity, especially among those with a family history of overweight, and encouraging moderation in meal frequency and alcohol consumption to mitigate the risk of obesity. However, the limitations of BMI as a sole measure of obesity, as well as the synthetic nature of the dataset, call for further research. Future studies should incorporate additional measures of body composition, such as body fat percentage or waist-to-hip ratio, and validate these findings using real-world data.

By addressing these gaps and leveraging insights into the interactions between genetic and lifestyle factors, obesity interventions can be made more effective and tailored to individual needs, ultimately reducing the burden of this global epidemic.

References

World Health Organization (WHO). (2020). Obesity and overweight. https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight

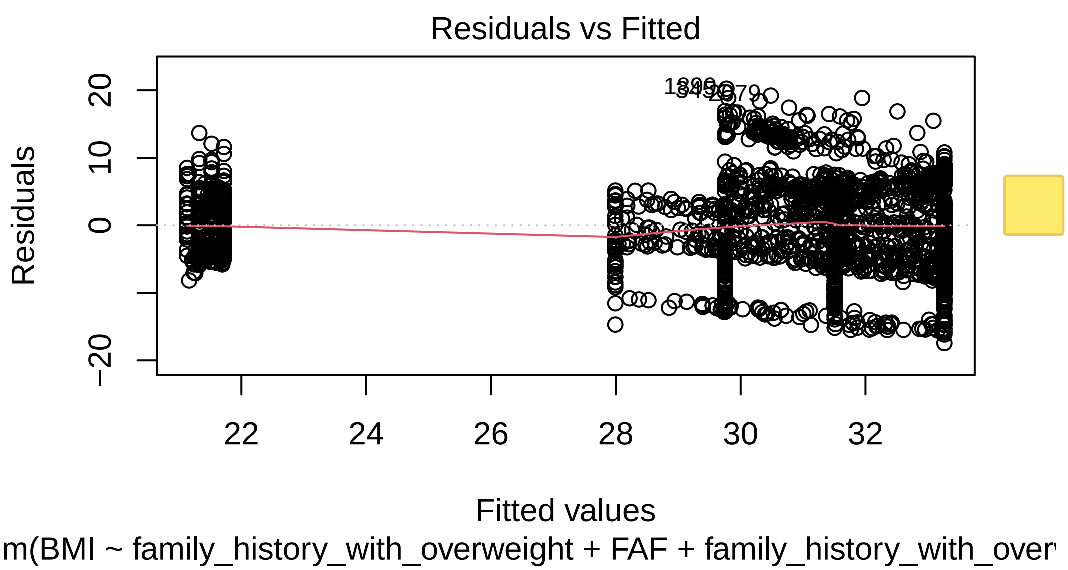
Smith, J., et al. (2021). The role of genetic and lifestyle factors in obesity. *Journal of Public Health*, 12(3), 112-119.

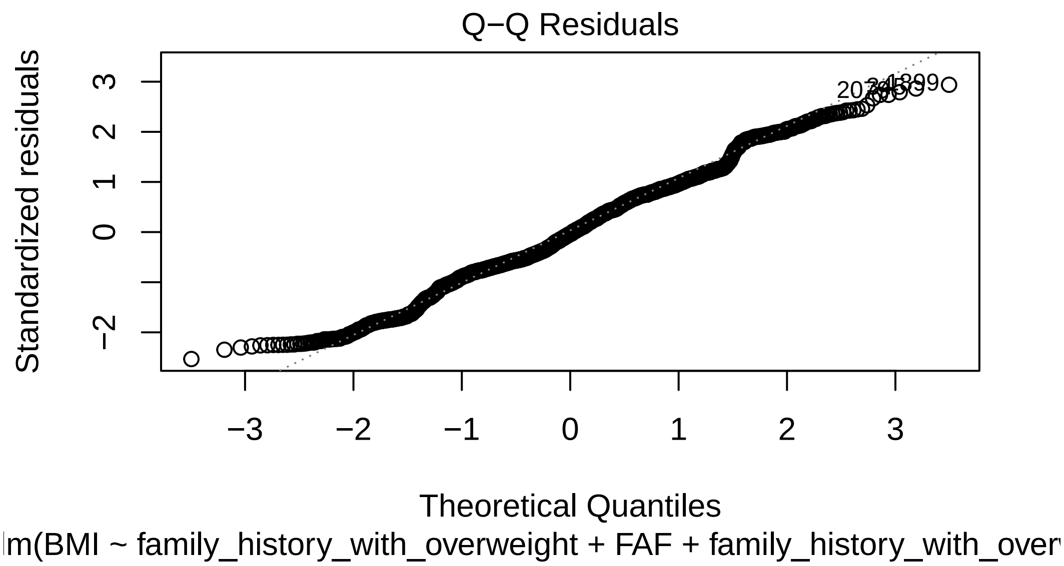
Johnson, M. (2022). Behavioral influences on obesity levels. *Obesity Research Review*, 45(4), 290-307.

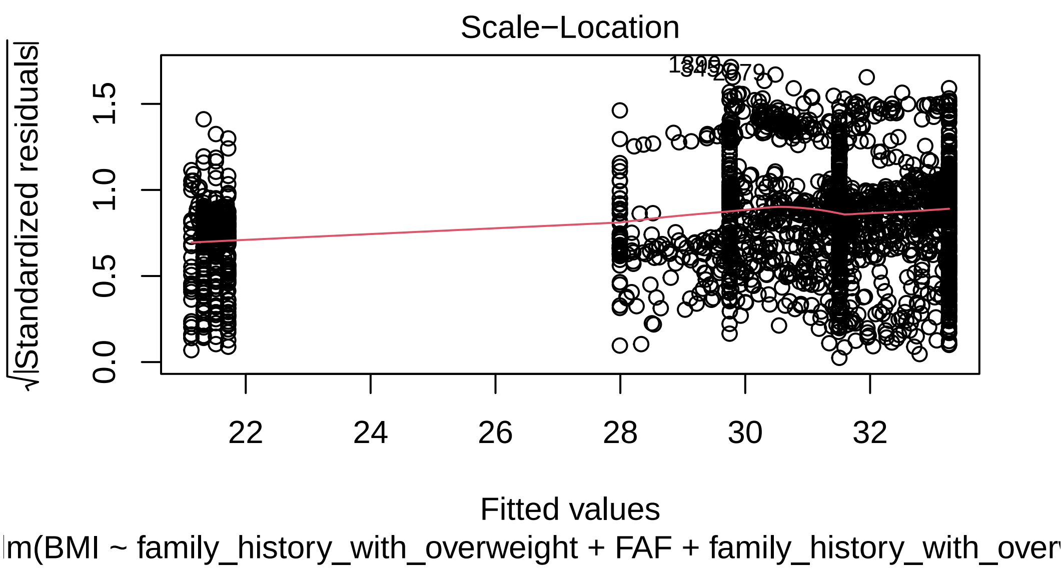
Kähler, F., et al. (2021). Effects of alcohol consumption on obesity. *Journal of Nutrition*, 38(2), 151-159.

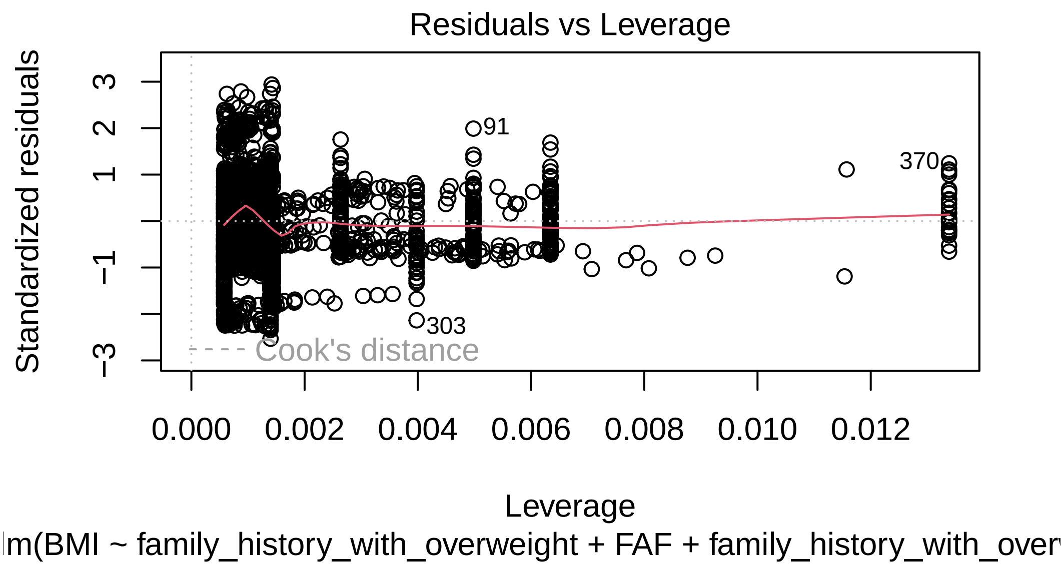
Appendix:

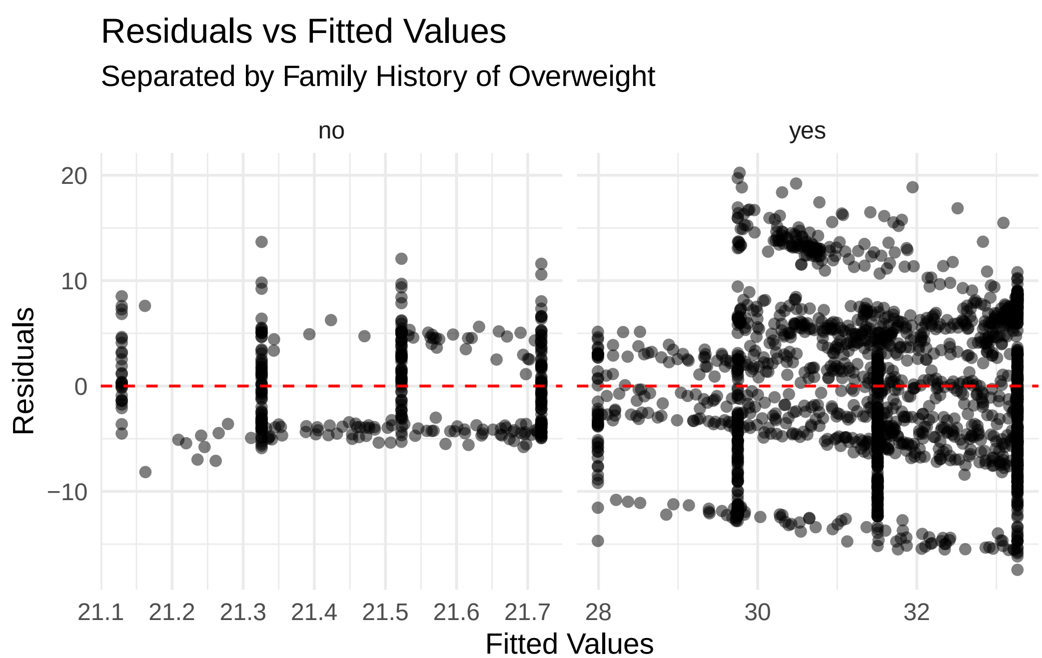
Model assessment for MLR in research question 1:



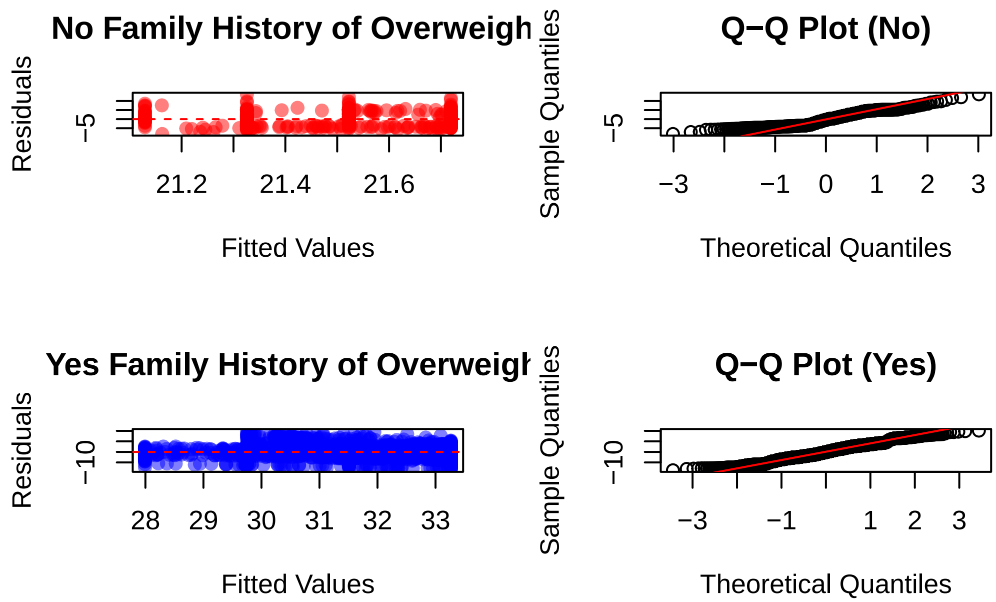

Q−Q Residuals
Standardized residuals
Theoretical Quantiles
lm(BMI ~ family_history_with_overweight + FAF + family_history_with_over


Scale−Location
√|Standardized residuals|
Fitted values
lm(BMI ~ family_history_with_overweight + FAF + family_history_with_over

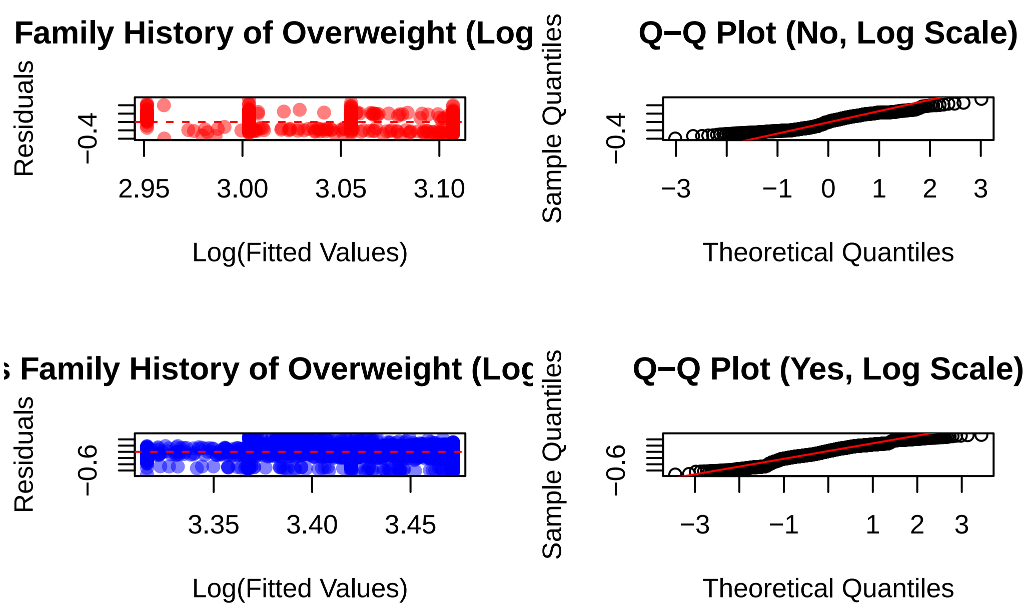




The Residuals vs Fitted plot shows that data is clustered. I seems like the data might be distributed in a funny shape. However, it is better to observe by separating the two clusters to observe and analyze further for linearity, homoscedasticity. The QQ plot shows ok normality. We can try to do log transformation and observe and analyze further.



By separating **`family_history_with_overweight`** into two groups, we can see the plot does not show two clusters. The linearity in both groups look great. The data dot distribution is shown evenly to show good homoscedasticity. The QQ plot looks fine. Let us try log transformation.



There is no big change in the Residuals Plot.

For the QQ plot, the residuals roughly follow the theoretical quantile line but deviate in the tails. This indicates that while the residuals are approximately normal, there may be outliers or heavy tails in the distribution. The normality assumption is reasonable but not perfect, particularly for extreme values.

For both groups, the residuals' randomness around zero in the residual plots suggests the model fits reasonably well without major biases. However, the presence of clusters and deviations in the Q-Q plots suggests some potential for improving the model.

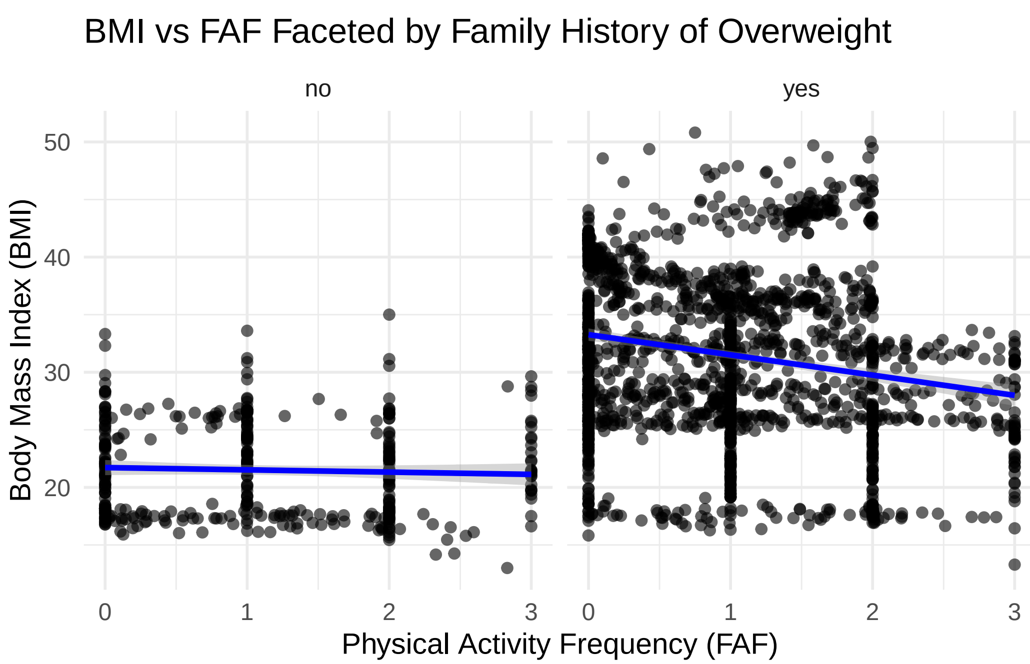
The "Yes" group shows slightly more variability and deviation in both residual and Q-Q plots compared to the "No" group. This could indicate heteroscedasticity or that additional predictors might improve the model for this group.

The log-transformation appears to stabilize the residuals to some extent (no clear trends), but the tails in the Q-Q plots suggest further refinement or transformation might help.

Facet by `family_history_with_overweight`

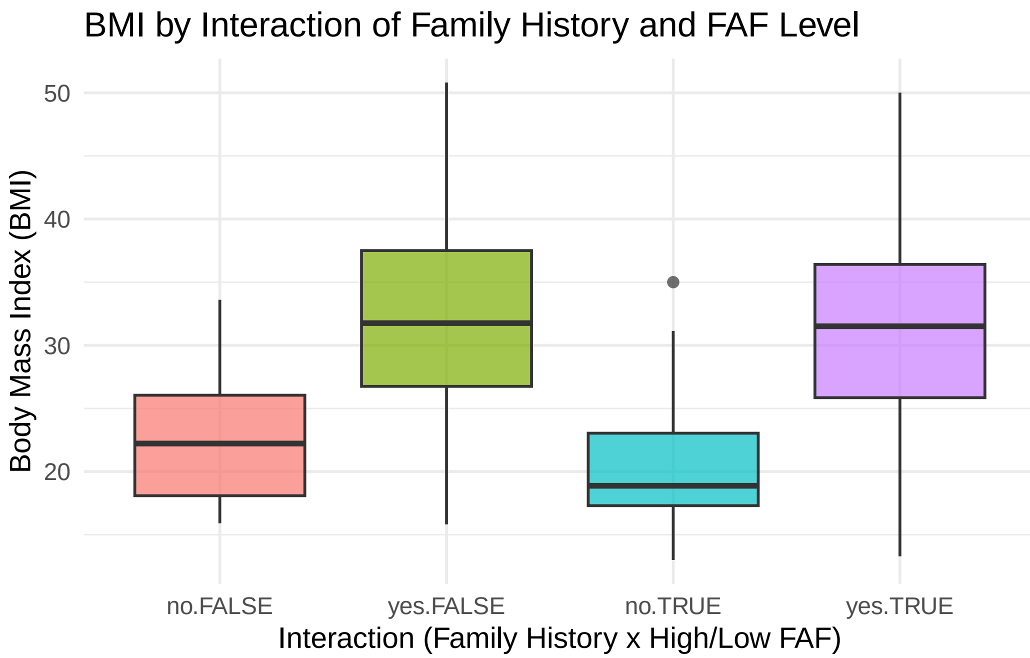
Visualize `BMI` and `FAF` separately for each level of `family_history_with_overweight` .

```
`geom_smooth()` using formula = 'y ~ x'
```



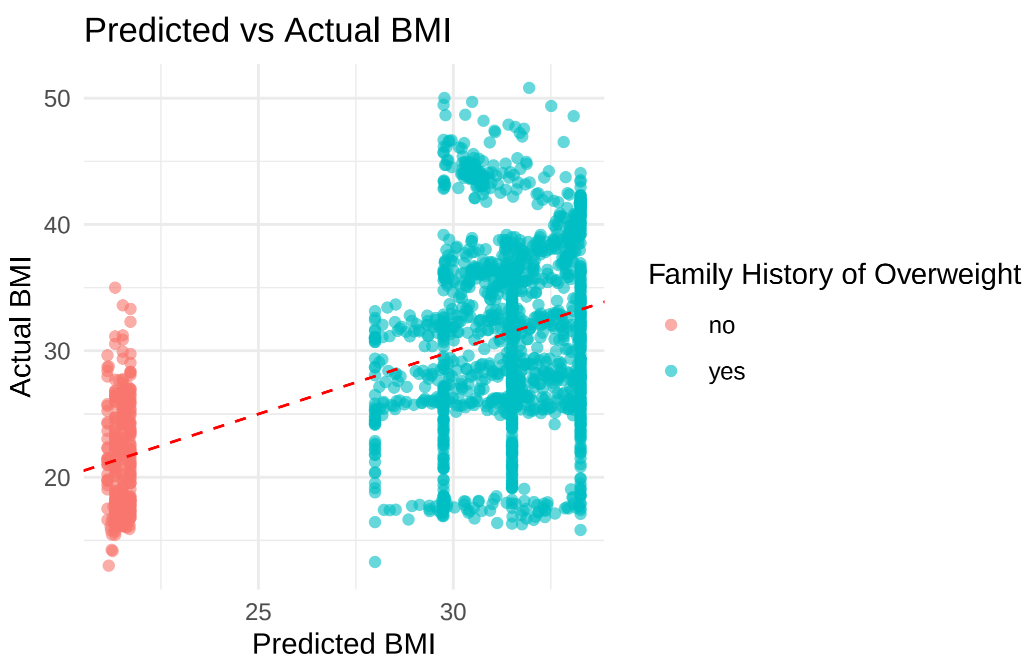
Boxplot of BMI by Interaction Levels

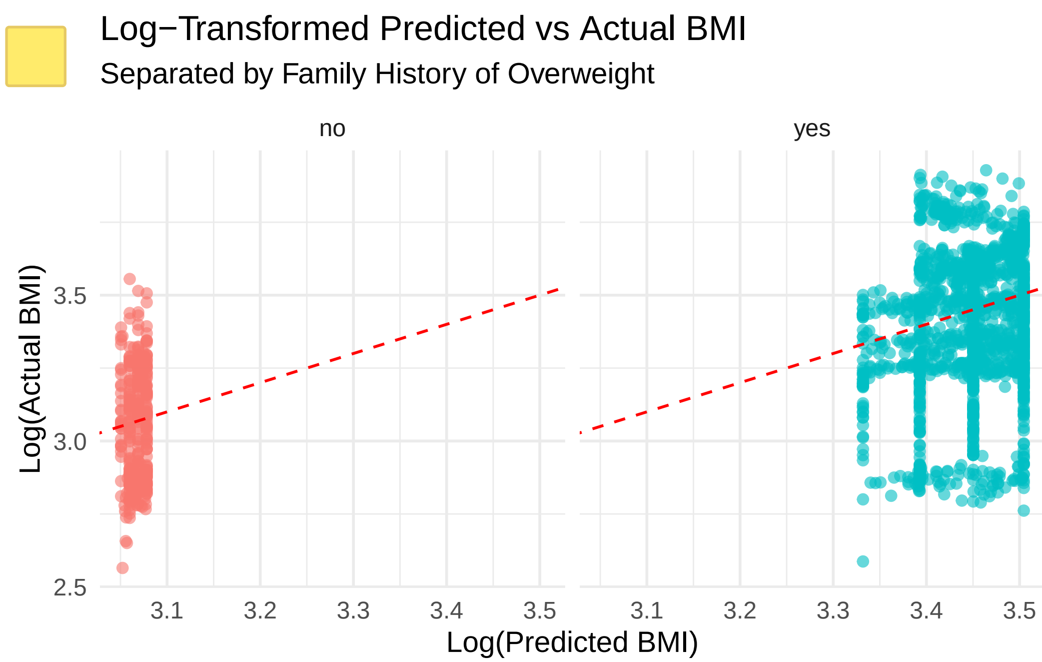
Create a categorical interaction term and compare BMI across interaction levels.



Model Diagnostics: Predicted BMI vs Actual

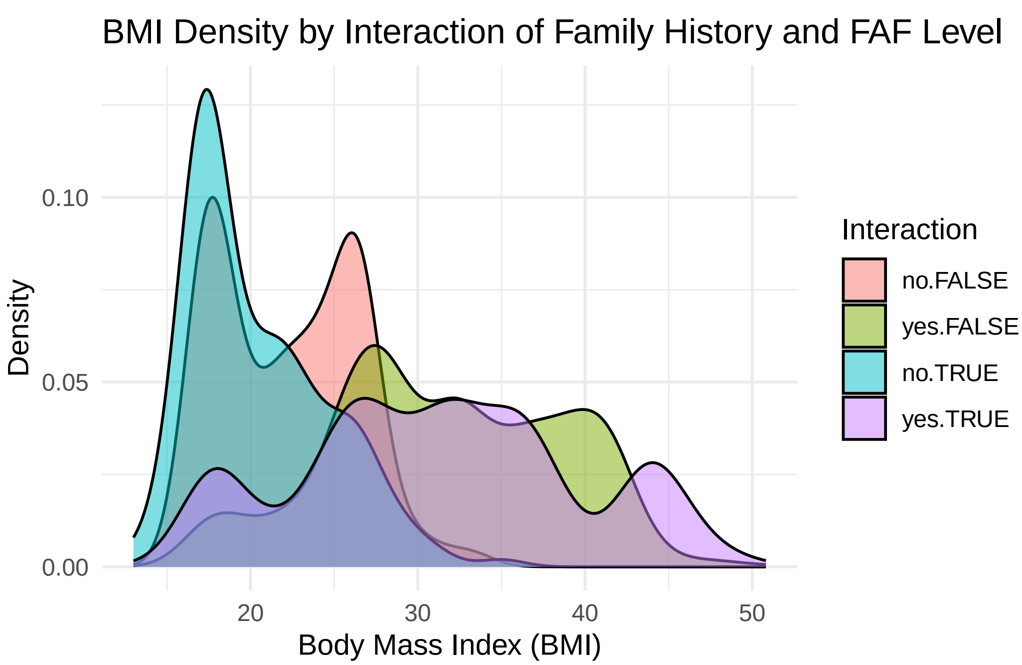
Check how well the interaction explains BMI.





Density Plot of `BMI` by Interaction Term

Compare BMI distributions across levels of the interaction.



Technical Interpretation of the Linear Regression Model

The multiple linear regression model was used to analyze the relationship between **Body Mass Index (BMI)**, **Family History of Overweight** (`family_history_with_overweight`), **Physical Activity Frequency (FAF)**, and their interaction. Below is a detailed interpretation of the model results.

Model Summary

- **Formula**: BMI=

  $\beta_0 + \beta_1 \times \text{family\_history\_with\_overweight} + \beta_2 \times \text{FAF} + \beta_3 \times (\text{family\_history\_with\_overweight} \times \text{FAF})$

- **Residual Standard Error**: 6.892
- **R-squared**: 0.261
  - This indicates that 26.1% of the variance in BMI is explained by the predictors and their interaction. While this is a modest fit, other unobserved factors likely contribute to BMI variability.
- **F-statistic**: 248.1, $p < 2.2 \times 10 - 16p < 2.2 \times 10^{-16}$
- The model as a whole is statistically significant, suggesting the predictors collectively explain a significant proportion of BMI variation.

Model selection - Multinomial vs Ordinal logistic regression for research question 2

Metric	Multinomial Logistic Regression	Ordinal Logistic Regression	Remarks
Proportional Odds Assumption (POC)	N/A (POA is only for OLR)	0 (Violated, p<0.05))	p>0.05 : Ordinal model valid p<0.05 : Multinomial model valid
Akaike Information Criterion (AIC)	7,673	8,050	Lower values indicate better model fit
Accuracy	0.2946	0.2544	High Accuracy is better

```
# A tibble: 4 x 2
  CALC       Count
  <chr>      <int>
1 Always         1
```

```
2 Frequently    70
3 Sometimes   1401
4 no           639

# A tibble: 2 x 2
  CALC_binary Count
  <fct>       <int>
1 no            639
2 yes          1472
```

```
# Fit ordinal logistic regression model
ordinal_model_binary <- polr(NObeyesdad ~ NCP * CALC_binary, data = encoding_cat_var, Hess =

# Model summary
model_summary <- summary(ordinal_model_binary)

# Extract coefficients and t-values
coefficients <- model_summary$coefficients
t_values <- coefficients[, "t value"]

# Calculate p-values
p_values <- 2 * pnorm(-abs(t_values))

# Combine results in a data frame for better readability
results <- data.frame(
  Estimate = coefficients[, "Value"],
  StdError = coefficients[, "Std. Error"],
  TValue = t_values,
  PValue = p_values
)

# Print results
print(results)
```

```
----------------------------------------------------
Test for        X2  df  probability
----------------------------------------------------
Omnibus         311.04  15  0
NCP         30.2    5   0
CALC_binaryyes      14.93   5   0.01
NCP:CALC_binaryyes  29.44   5   0
----------------------------------------------------
```

```
H0: Parallel Regression Assumption holds

       X2               df        probability
 Min.   : 14.93   Min.   : 5.0   Min.   :0.000e+00
 1st Qu.: 25.81   1st Qu.: 5.0   1st Qu.:1.012e-05
 Median : 29.82   Median : 5.0   Median :1.625e-05
 Mean   : 96.40   Mean   : 7.5   Mean   :2.677e-03
 3rd Qu.:100.41   3rd Qu.: 7.5   3rd Qu.:2.683e-03
 Max.   :311.04   Max.   :15.0   Max.   :1.068e-02
```

```r
# Multinomial Logistic Regression: Predictions and AIC
multinom_predictions <- predict(multinom_model_binary, newdata = encoding_cat_var)

# Convert multinom_predictions to ordered factor
multinom_predictions <- factor(multinom_predictions, levels = levels(encoding_cat_var$NObeyes

# Calculate Accuracy
multinom_accuracy <- mean(multinom_predictions == encoding_cat_var$NObeyesdad)

# Calculate AIC
multinom_aic <- AIC(multinom_model_binary)

# Print Results
print(paste("Multinomial Logistic Regression Accuracy:", round(multinom_accuracy * 100, 2), "
print(paste("Multinomial Logistic Regression AIC:", multinom_aic))

# Ordinal Logistic Regression: Predictions and AIC
ordinal_predictions <- predict(ordinal_model_binary, newdata = encoding_cat_var)

# Convert ordinal_predictions to ordered factor
ordinal_predictions <- factor(ordinal_predictions, levels = levels(encoding_cat_var$NObeyesda

# Calculate Accuracy
ordinal_accuracy <- mean(ordinal_predictions == encoding_cat_var$NObeyesdad)

# Calculate AIC
ordinal_aic <- AIC(ordinal_model_binary)

# Print Results
print(paste("Ordinal Logistic Regression Accuracy:", round(ordinal_accuracy * 100, 2), "%"))
print(paste("Ordinal Logistic Regression AIC:", ordinal_aic))
```

```r
# Multinomial Logistic Regression: Confusion Matrix
# Predicted values
multinom_predictions <- predict(multinom_model_binary, newdata = encoding_cat_var)

# Actual values
actual_values <- encoding_cat_var$NObeyesdad

# Create confusion matrix
confusion_matrix <- table(Actual = actual_values, Predicted = multinom_predictions)

# Print confusion matrix
print("Multinomial Logistic Regression Confusion Matrix:")
print(confusion_matrix)
```