Team 2. Project

Exploring

Meet Our Fabulous Team Members

Kevin Ngala

Andrea Nimako

FOCUS ON OBESITY

itv NEWS

Kavita Lakhani

Lynda Sempele

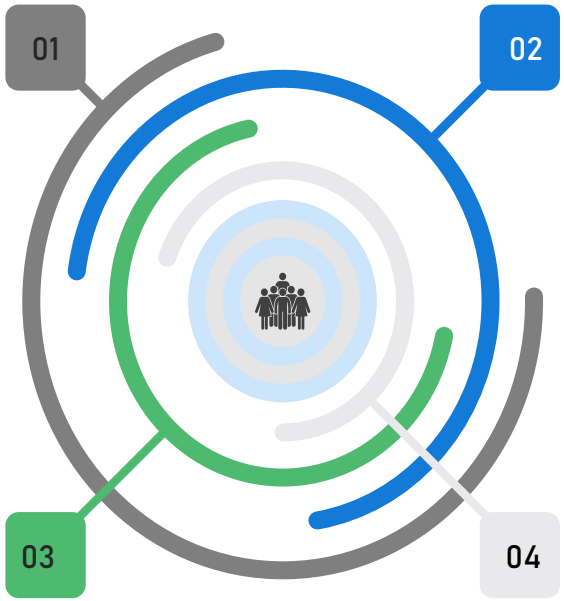
Monique Reid

Project Aim and Scope

- Identify obesity risk factors through dataset analysis.
- Explore demographics, diet, lifestyle for insights.
- Inform prevention and health policy initiatives.

Agenda

Data Collection and Preprocessing:
Exploratory, Statistical Analysis, Visualization, Interpretation of Findings & Reporting



Analysis Process with Jupyter Notebooks:
Analyzing, documenting data exploration

Pandas Data Cleaning:
Use Pandas to clean and format the obesity dataset.

Matplotlib Visualizations:
Create 6 to 8 visualizations using Matplotlib, with at least 2 per research question.

Exploring Obesity Risk Factors

Obesity Dataset Description:

1. Covers individuals aged 14 to 61 from Mexico, Peru, and Colombia.
2. Derived from a web survey, totaling 17 attributes and 2111 records.
3. Contains numeric, continuous, and categorical data for analysis.

Data Exploration:

- Cleaning, reclassification, etc.
- Jupyter Notebook
- pandas

- Descriptive statistics
- pandas
- Statistical libraries

Data Analysis:

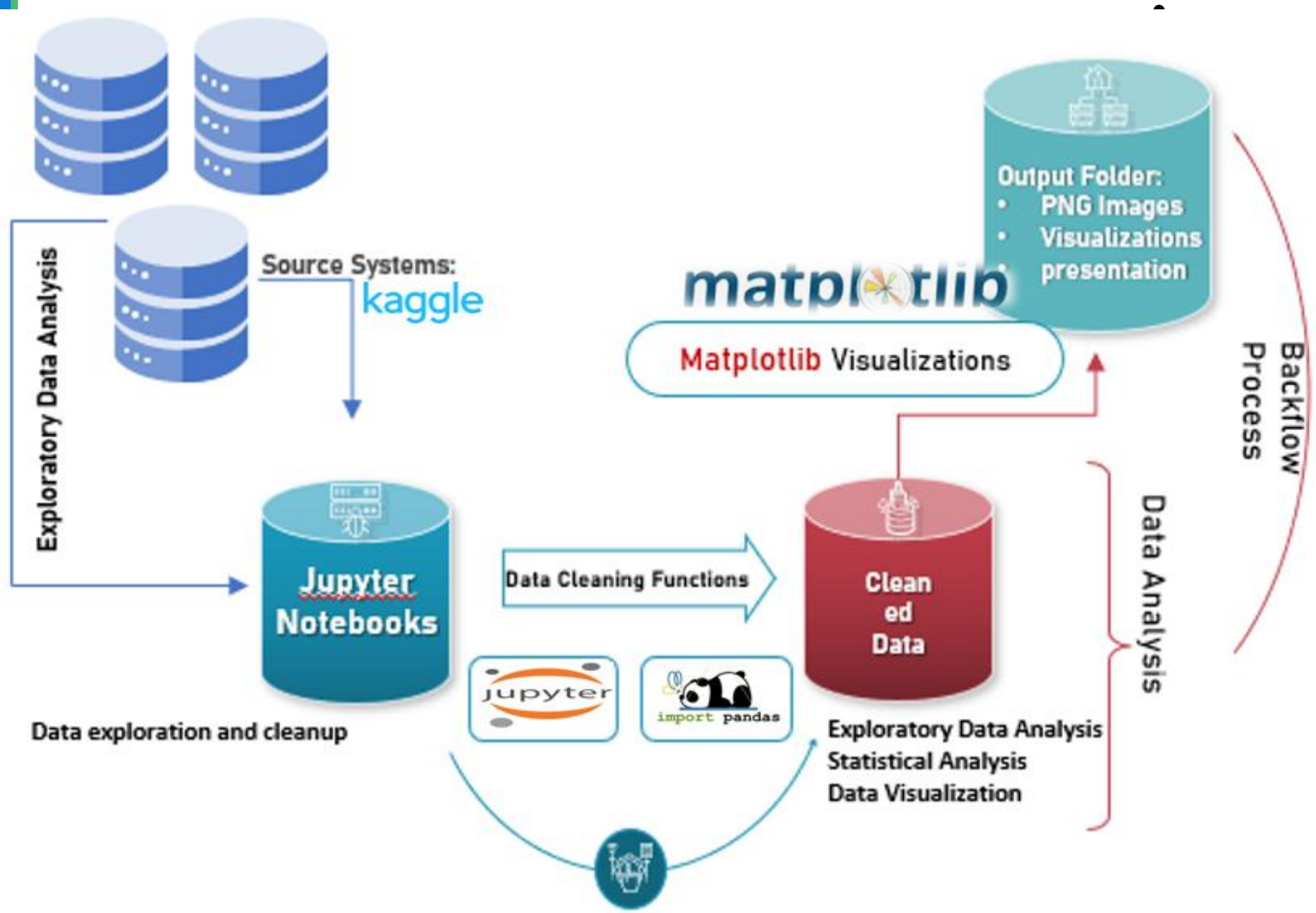
- Matplotlib
- Plotly (optional)

Data Visualization:

Obesity Dataset: *Obesity or CVD risk (Classify/Regressor/Cluster) Dataset.*

Author/ Source: *Fabio Mendoza Palechor, and Alexis Manotas. (2023). Obesity or CVD risk (Classify/Regressor/Cluster) [Data set]. Kaggle.*

Exploring Obesity Risk Factors



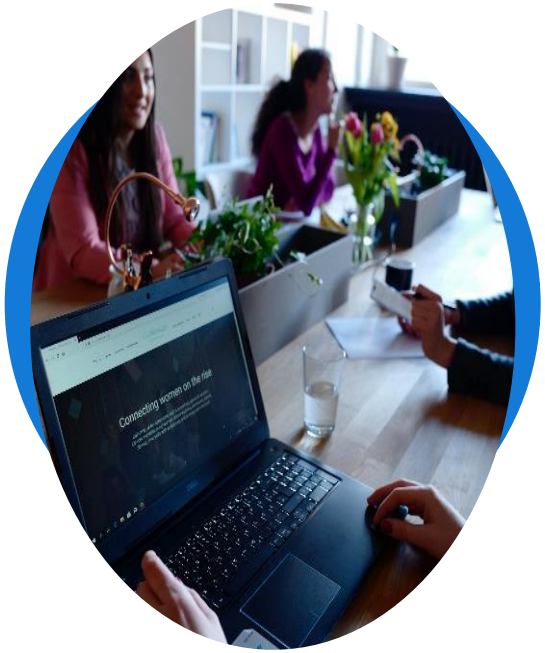
Data Cleaning

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1	Gender	Age	Height	Weight	family_hist	FAVC	FCVC	NCP	CAEC	SMOKE	CH2O	SCC	FAF	TUE	CALC	MTRANS	NObeyesdad		
2	Female	21	1.62	64	yes	no	2	3	Sometimes	no	2	no	0	1	no	Public_Tran	Normal_Weight		
3	Female	21	1.52	56	yes	no	3	3	Sometimes	yes	3	yes	3	0	Sometimes	Public_Tran	Normal_Weight		
4	Male	23	1.8	77	yes	no	2	3	Sometimes	no	2	no	2	1	Frequently	Public_Tran	Normal_Weight		
5	Male	27	1.8	87	no	no	3	3	Sometimes	no	2	no	2	0	Frequently	Walking	Overweight_Level_I		
6	Male	22	1.78	89.8	no	no	2	1	Sometimes	no	2	no	0	0	Sometimes	Public_Tran	Overweight_Level_II		
7	Male	29	1.62	53	no	yes	2	3	Sometimes	no	2	no	0	0	Sometimes	Automobile	Normal_Weight		
8	Female	23	1.5	55	yes	yes	3	3	Sometimes	no	2	no	1	0	Sometimes	Motorbike	Normal_Weight		
9	Male	22	1.64	53	no	no	2	3	Sometimes	no	2	no	3	0	Sometimes	Public_Tran	Normal_Weight		
10	Male	24	1.78	64	yes	yes	3	3	Sometimes	no	2	no	1	1	Frequently	Public_Tran	Normal_Weight		
11	Male	22	1.72	68	yes	yes	2	3	Sometimes	no	2	no	1	1	no	Public_Tran	Normal_Weight		
12	Male	26	1.85	105	yes	yes	3	3	Frequently	no	3	no	2	2	Sometimes	Public_Tran	Obesity_Type_I		
13	Female	21	1.72	80	yes	yes	2	3	Frequently	no	2	yes	2	1	Sometimes	Public_Tran	Overweight_Level_II		
14	Male	22	1.65	56	no	no	3	3	Sometimes	no	3	no	2	0	Sometimes	Public_Tran	Normal_Weight		
15	Male	41	1.8	99	no	yes	2	3	Sometimes	no	2	no	2	1	Frequently	Automobile	Obesity_Type_I		
16	Male	23	1.77	60	yes	yes	3	1	Sometimes	no	1	no	1	1	Sometimes	Public_Tran	Normal_Weight		
17	Female	22	1.7	66	yes	no	3	3	Always	no	2	yes	2	1	Sometimes	Public_Tran	Normal_Weight		
18	Male	27	1.93	102	yes	yes	2	1	Sometimes	no	1	no	1	0	Sometimes	Public_Tran	Overweight_Level_II		

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1	ID	Gender	Age	Height(m)	Weight(kg)	Fhx of Over	Freq High C	Freq vegeta	Number of	Eating inbe	Hx of smok	Water cons	Calorie Con	Freq of Phy	Time spent	Alcohol cor	Transporta	BMI	Weight classification	
2	0	Female	21	1.62	64	yes	no	2	3	Sometimes	no	2	no	0	1	no	Public_Tran	24.4	Normal	
3	1	Female	21	1.52	56	yes	no	3	3	Sometimes	yes	3	yes	3	0	Sometimes	Public_Tran	24.2	Normal	
4	2	Male	23	1.8	77	yes	no	2	3	Sometimes	no	2	no	2	1	Frequently	Public_Tran	23.8	Normal	
5	3	Male	27	1.8	87	no	no	3	3	Sometimes	no	2	no	2	0	Frequently	Walking	26.9	Overweight	
6	4	Male	22	1.78	89.8	no	no	2	1	Sometimes	no	2	no	0	0	Sometimes	Public_Tran	28.3	Overweight	
7	5	Male	29	1.62	53	no	yes	2	3	Sometimes	no	2	no	0	0	Sometimes	Automobile	20.2	Normal	
8	6	Female	23	1.5	55	yes	yes	3	3	Sometimes	no	2	no	1	0	Sometimes	Motorbike	24.4	Normal	
9	7	Male	22	1.64	53	no	no	2	3	Sometimes	no	2	no	3	0	Sometimes	Public_Tran	19.7	Normal	
10	8	Male	24	1.78	64	yes	yes	3	3	Sometimes	no	2	no	1	1	Frequently	Public_Tran	20.2	Normal	
11	9	Male	22	1.72	68	yes	yes	2	3	Sometimes	no	2	no	1	1	no	Public_Tran	23	Normal	
12	10	Male	26	1.85	105	yes	yes	3	3	Frequently	no	3	no	2	2	Sometimes	Public_Tran	30.7	Obese	
13	11	Female	21	1.72	80	yes	yes	2	3	Frequently	no	2	yes	2	1	Sometimes	Public_Tran	27	Overweight	
14	12	Male	22	1.65	56	no	no	3	3	Sometimes	no	3	no	2	0	Sometimes	Public_Tran	20.6	Normal	
15	13	Male	41	1.8	99	no	yes	2	3	Sometimes	no	2	no	2	1	Frequently	Automobile	30.6	Obese	
16	14	Male	23	1.77	60	yes	yes	3	1	Sometimes	no	1	no	1	1	Sometimes	Public_Tran	19.2	Normal	

Exploring Obesity Risk Factors

Research Questions:



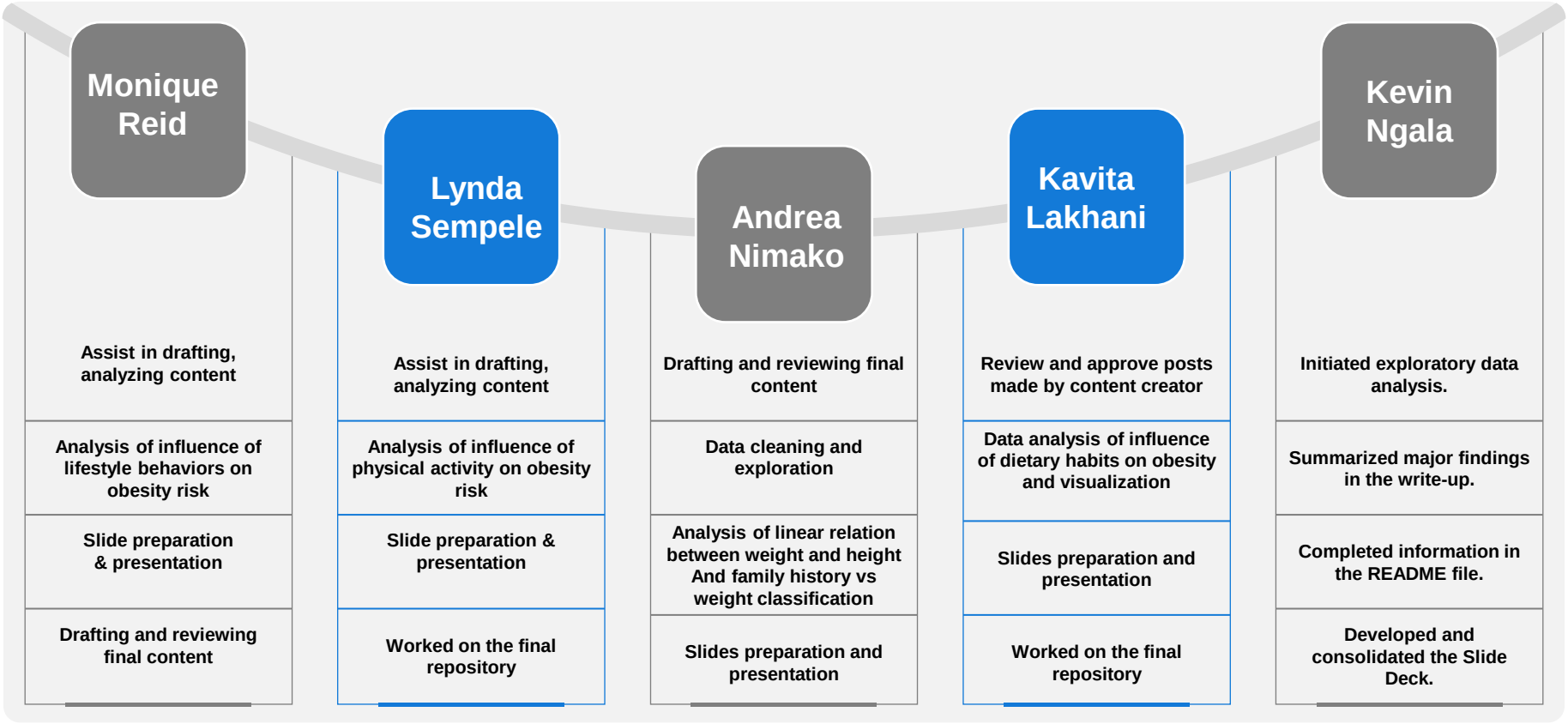
1. What is the association between dietary habits and the risk of obesity?

2. Is there a significant relationship between smoking, alcohol consumption, and the risk of obesity?

3. What is the association between physical activity levels and the risk of obesity?

Exploring Obesity Risk Factors

Roles and Responsibilities of Project Content



Completed Analysis Uploaded to GitHub

Network graph

Timeline of the most recent commits to this repository and its network ordered by most recently pushed to.

Analysis & Visualization

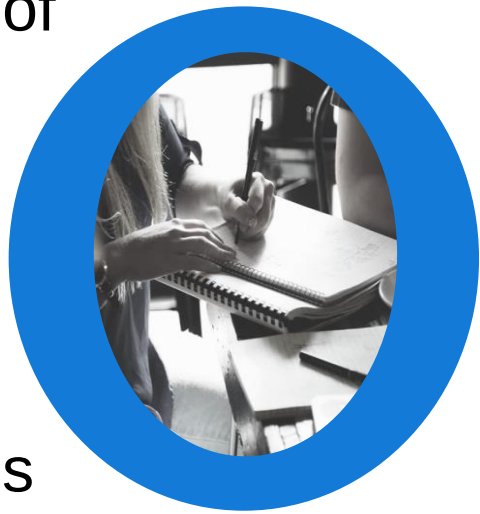
Analyses to identify correlations with obesity risk:

1. **Correlation Analysis:** Assessment of relationships between various factors

2. **Hypothesis Testing:** Evaluating significance of associations between variables and obesity.

3. **Regression Modeling:** Predicting factors influencing obesity risk using regression.

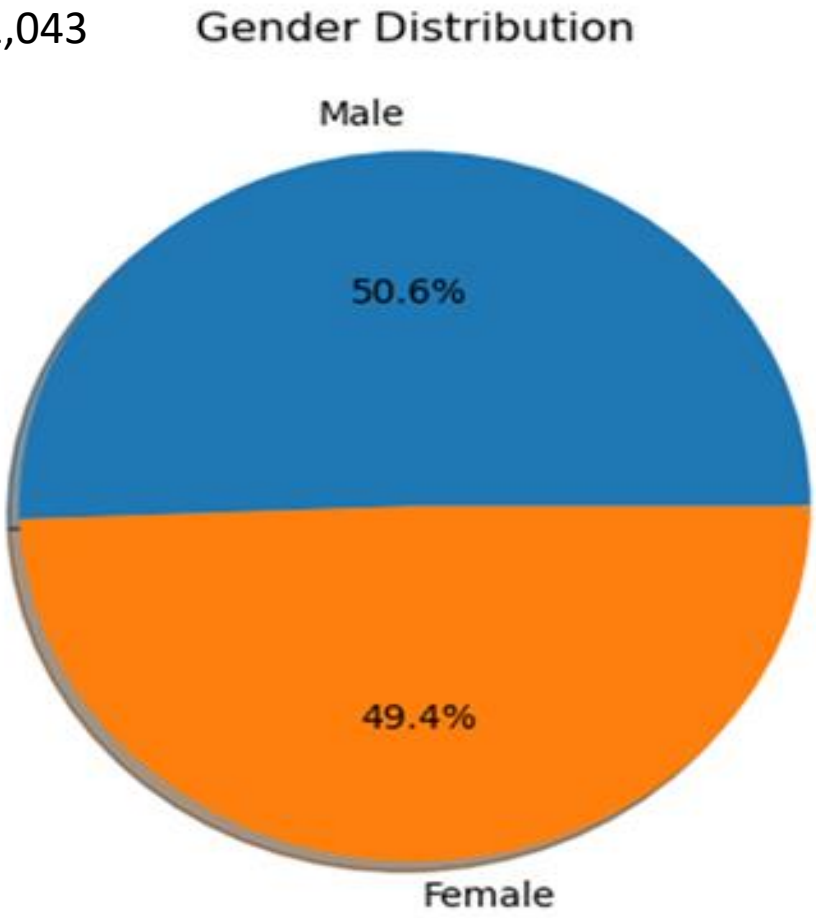
4. **Visualization:** Creating data-driven visual aids



Exploring Obesity Risk Factors

Gender Distribution:

- Males: 1,068
- Female: 1,043



Gender distribution is fairly equal in the dataset

Exploring Obesity Risk Factors

Most candidates were underweight, followed by those with normal weight.

01

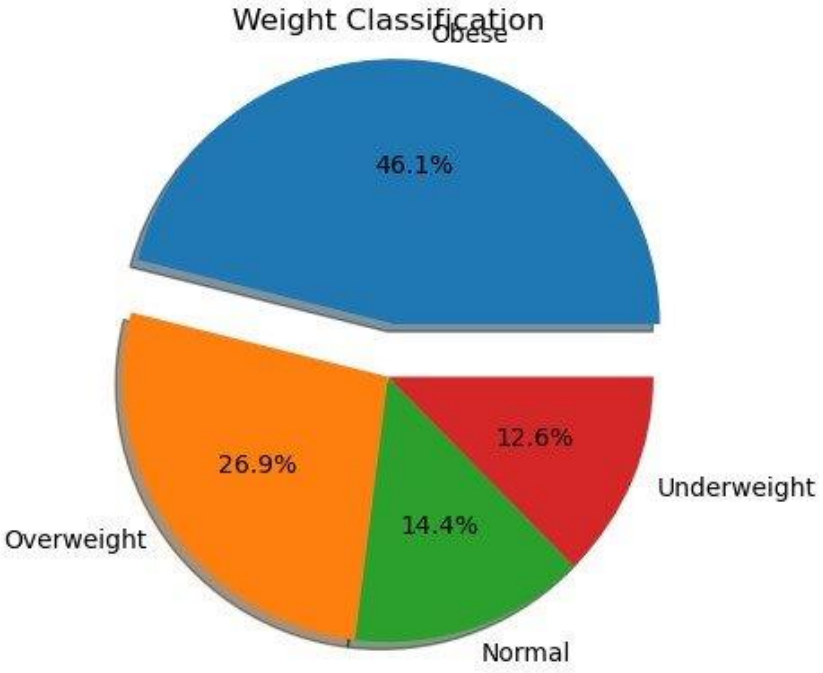
46. 1 % of the subjects were obese

02

US obesity prevalence was 41.9% in 2017 – March 2020

NHANES, 2021, CDC

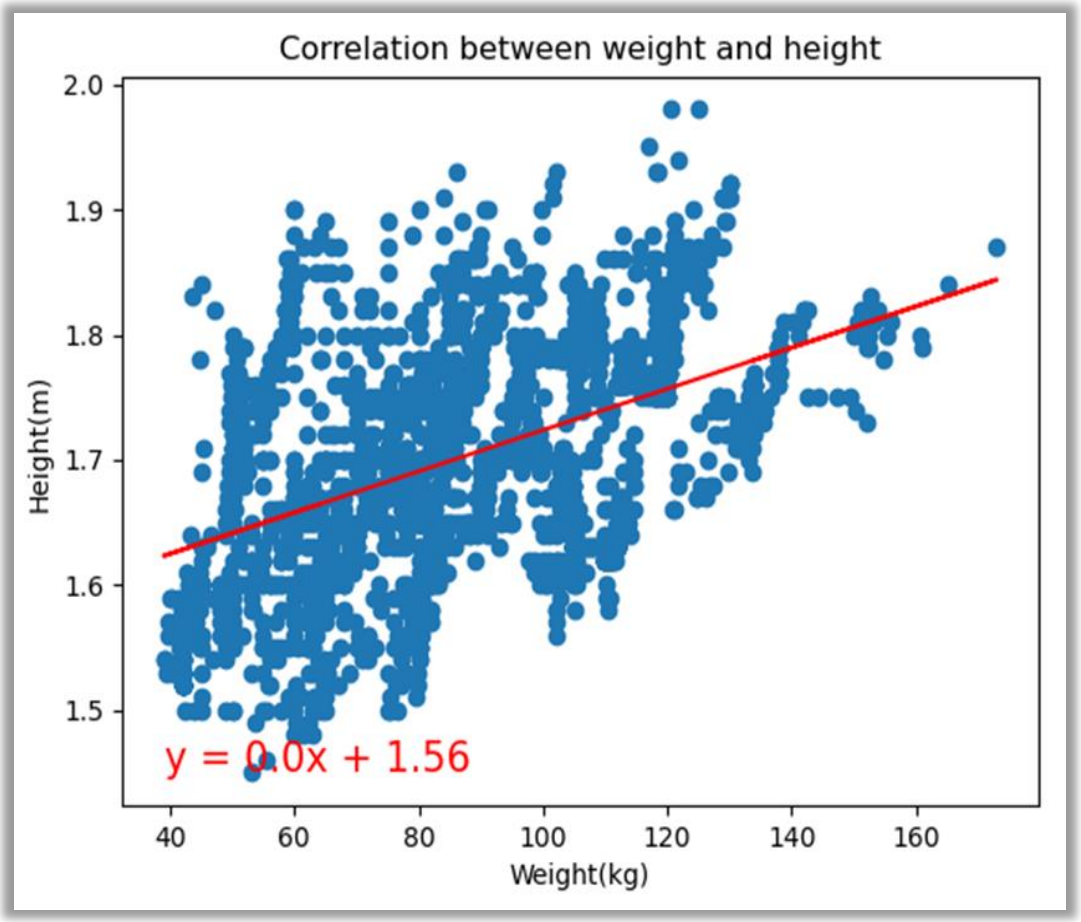
Family history across the various weight classifications



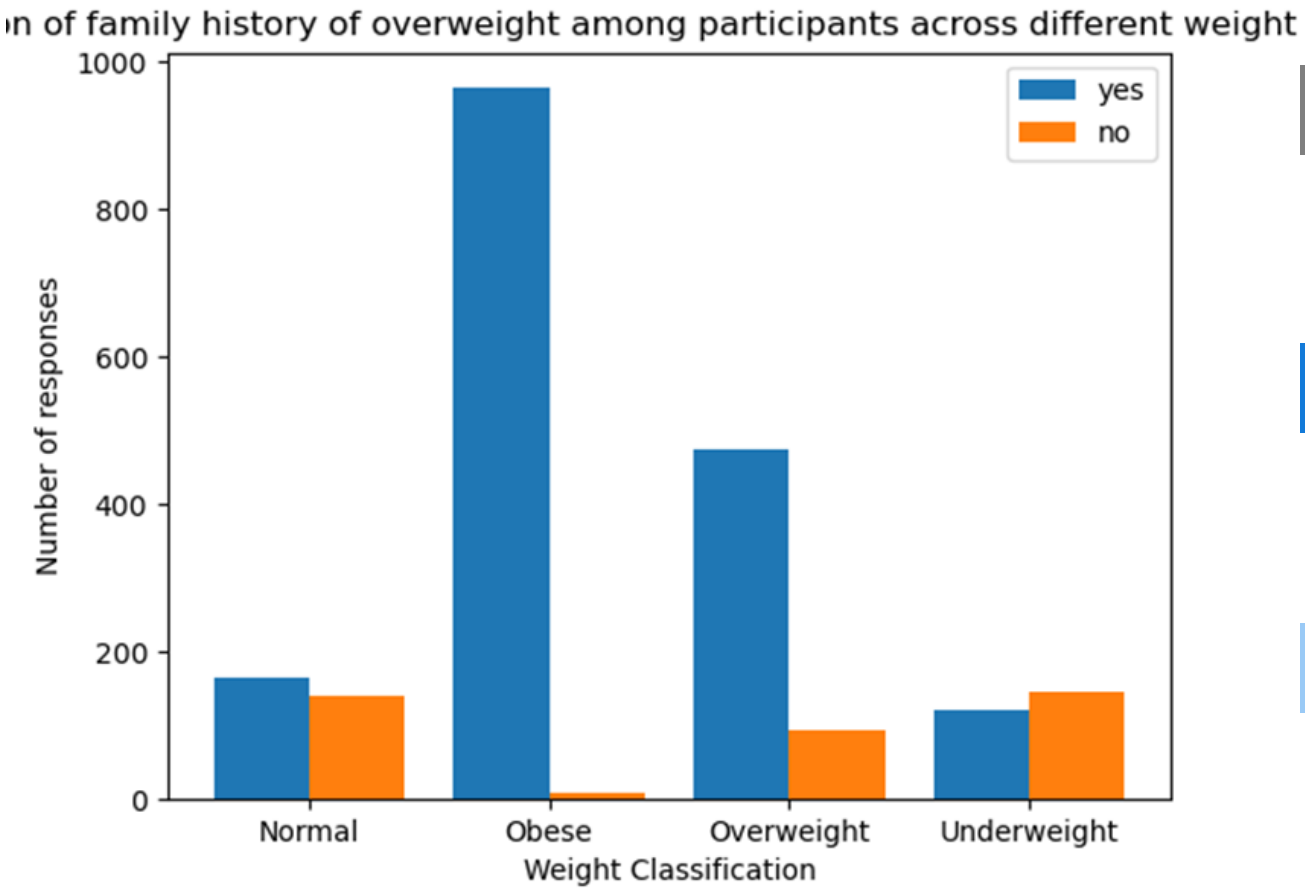
Exploring Obesity Risk Factors

01 Moderate correlation (r-value = 0.46.

02 Factors such as age, genetics, diet and physical activity can influence a person's weight and height.



Family history across the various weight classifications



01 Those with normal weight or underweight had an almost equal number of yes and no responses.

02 However, the majority of respondents classified as overweight or obese answered "yes."

03 Having a positive family history of overweight may be a significant factor to one's risk of become obese.

Exploring Obesity Risk Factors

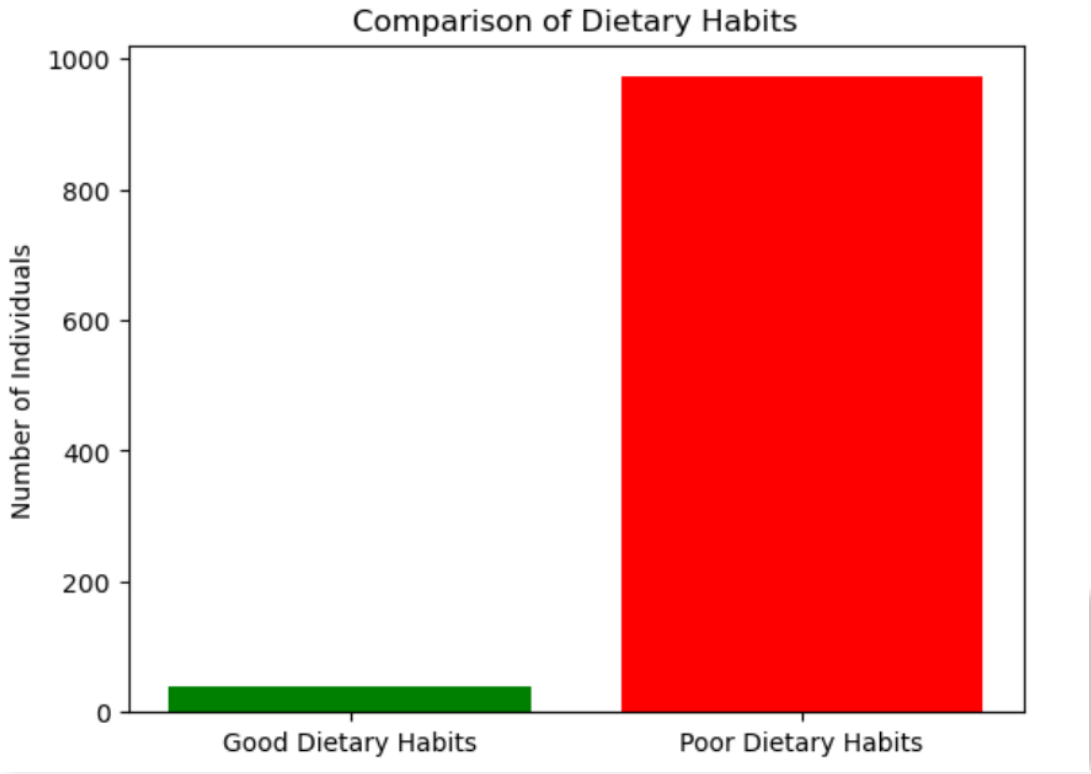
Research Question 1.

1

What is the association between dietary habits and the risk of obesity?

Exploring Obesity Risk Factors

Comparison of Dietary Habits in Obese population



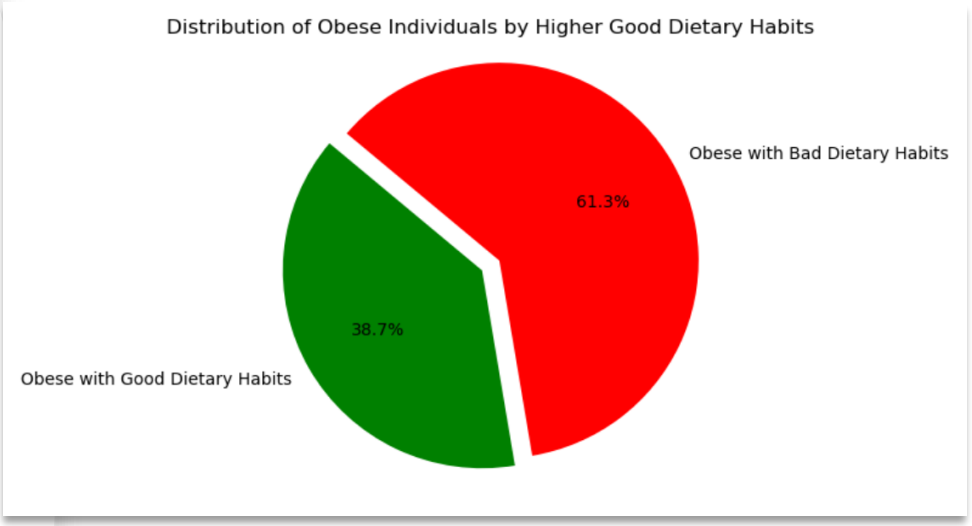
Data reveals a concerning trend:

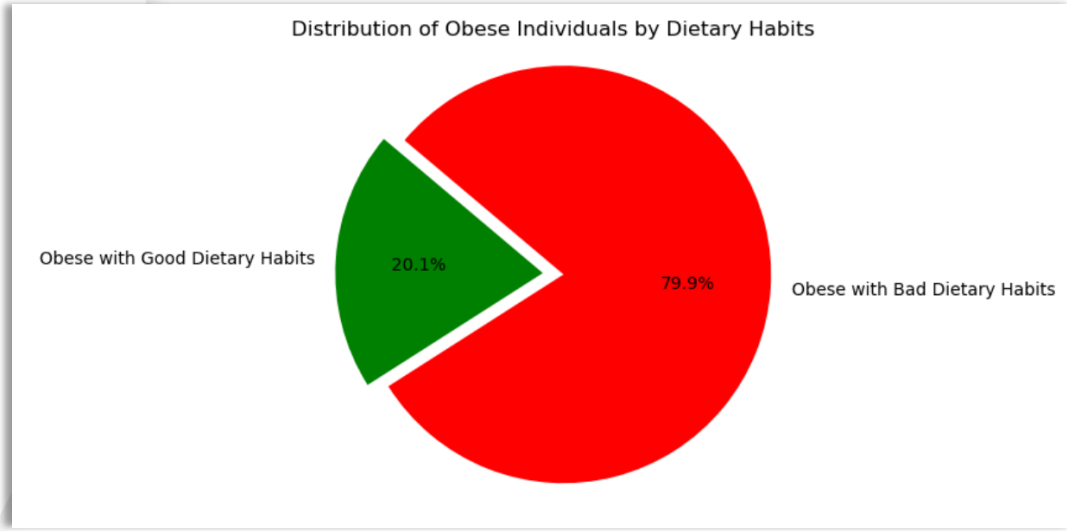
- poor dietary habits are more common among overweight and obese individuals compared to those with healthier habits.

Exploring Obesity Risk Factors

Distribution of Obese Individuals by Dietary Habits

- If three serving of veggies and at least two glasses of water are consumed that increases good dietary habits and reduced obesity.
- Whereas if less veggies and water is consumed than that increases obesity. This shows that obesity is influenced by dietary habits





Exploring Obesity Risk Factors

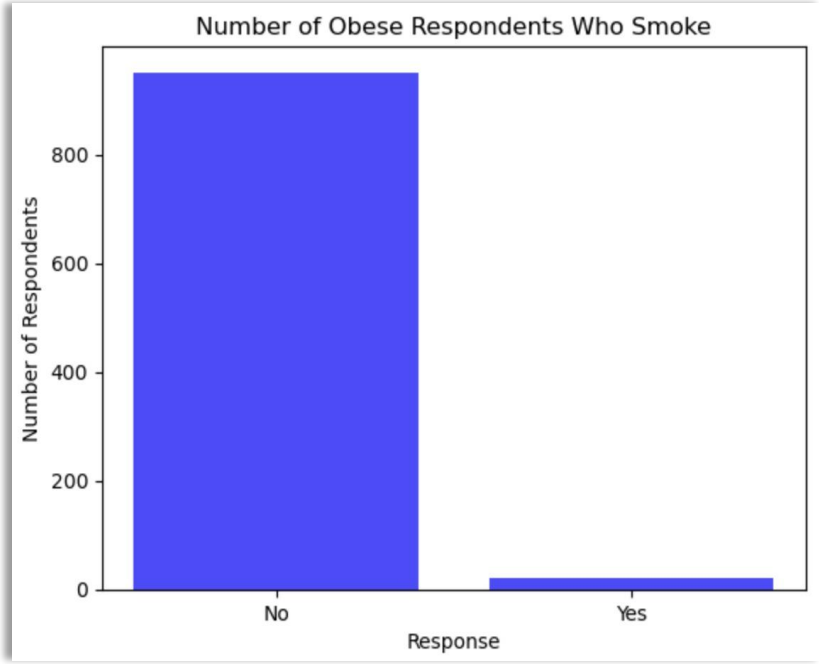
Research Question 2.

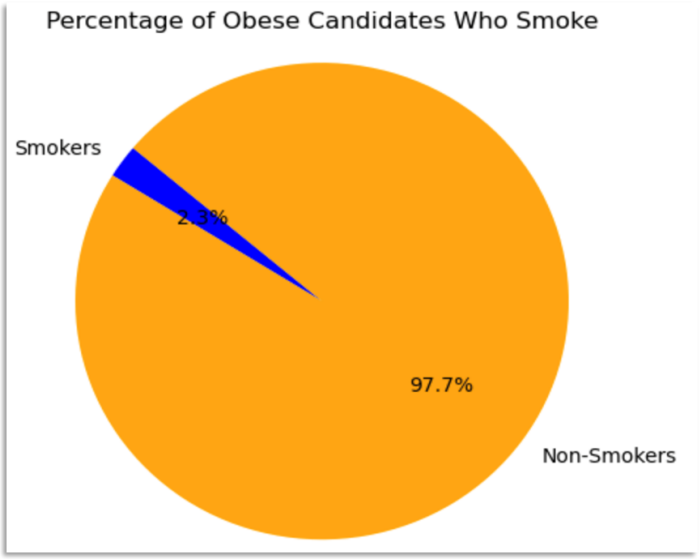
2

Is there a significant relationship between smoking, alcohol consumption, and the risk of obesity?

Exploring Obesity Risk Factors

Number of Respondents Related to Smoking and Alcohol Consumption

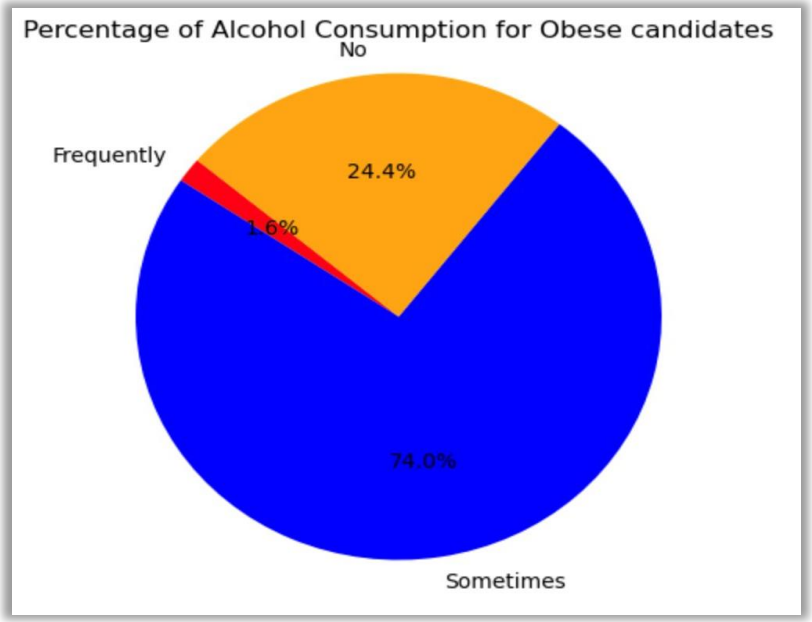




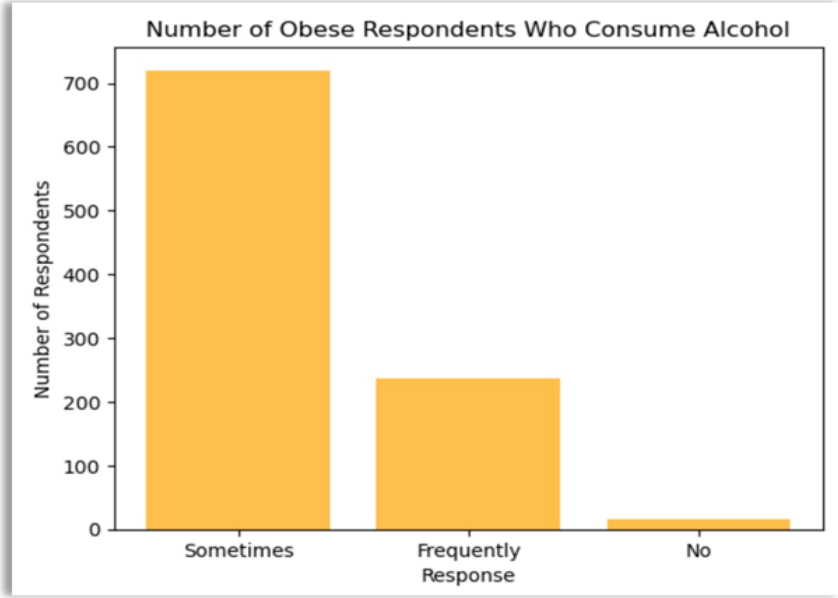
- **Majority of the non-smoking population 951**
- **A total of 957 respondents consume alcohol and are obese**

Exploring Obesity Risk Factors

Percentage of Respondent Votes for Alcohol Consumption



The percentages depicted in this pie chart shows a clear correlation between alcohol consumption and obesity.

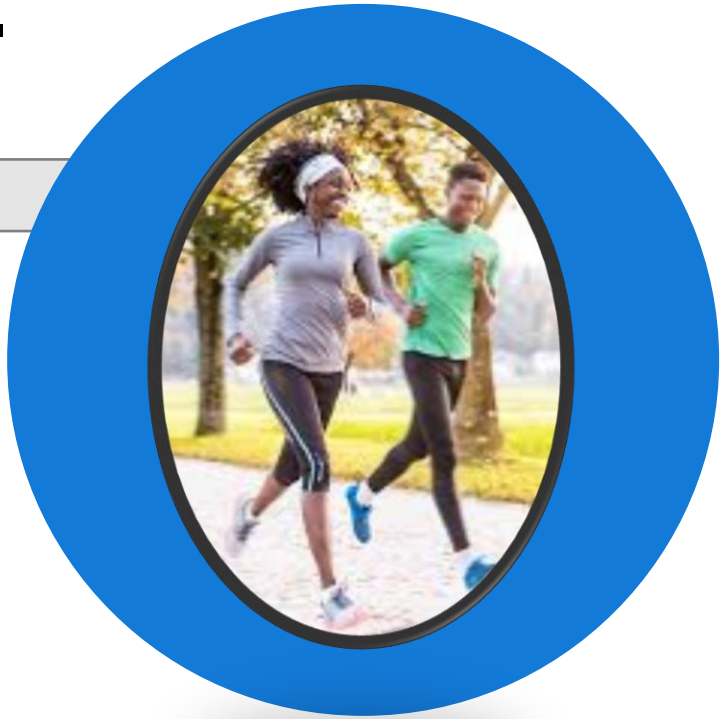


Exploring Obesity Risk Factors

Research Question 3.

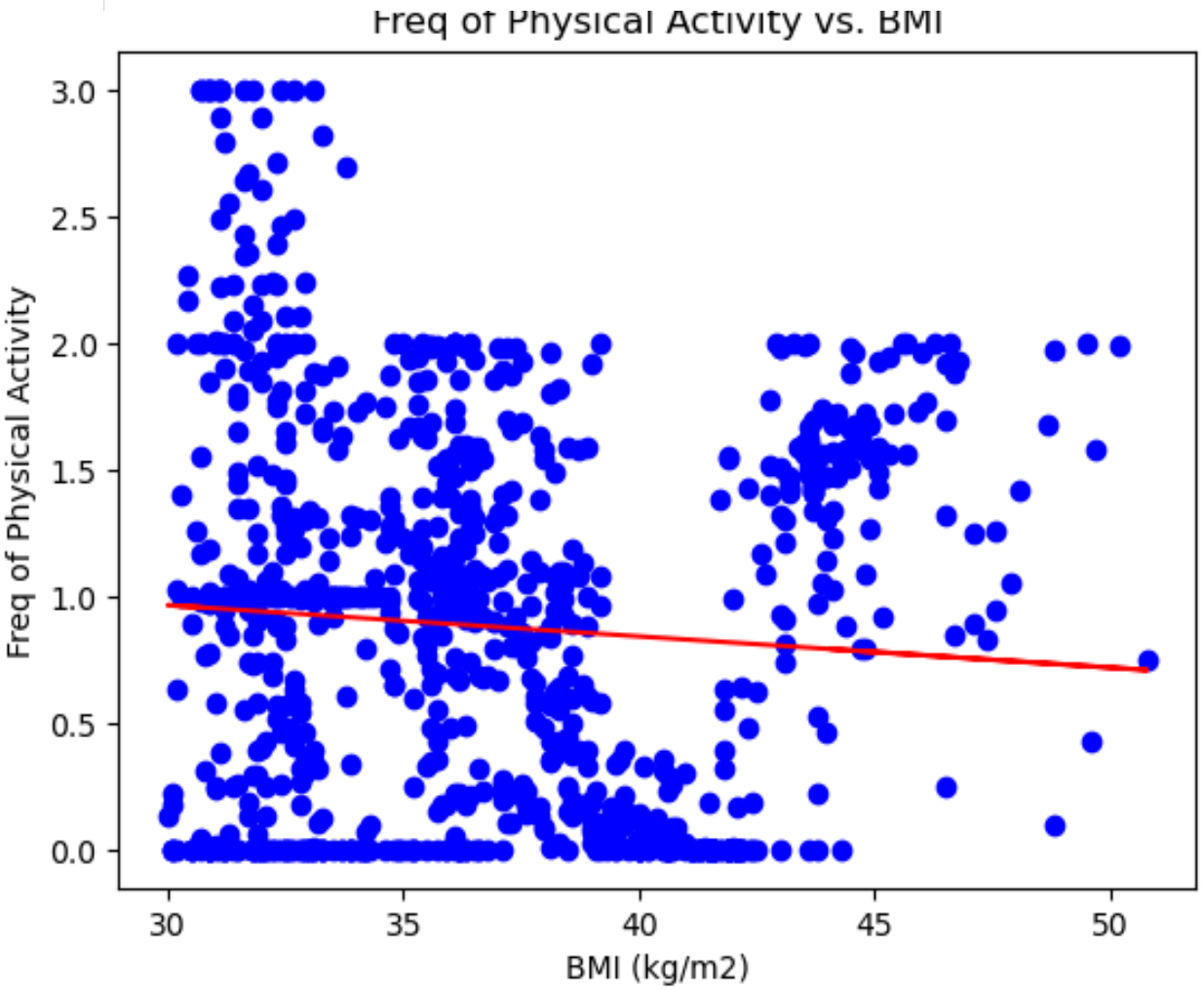
3

What is the association between physical activity levels and the risk of obesity?



Exploring Obesity Risk Factors

Correlation between physical activity & BMI



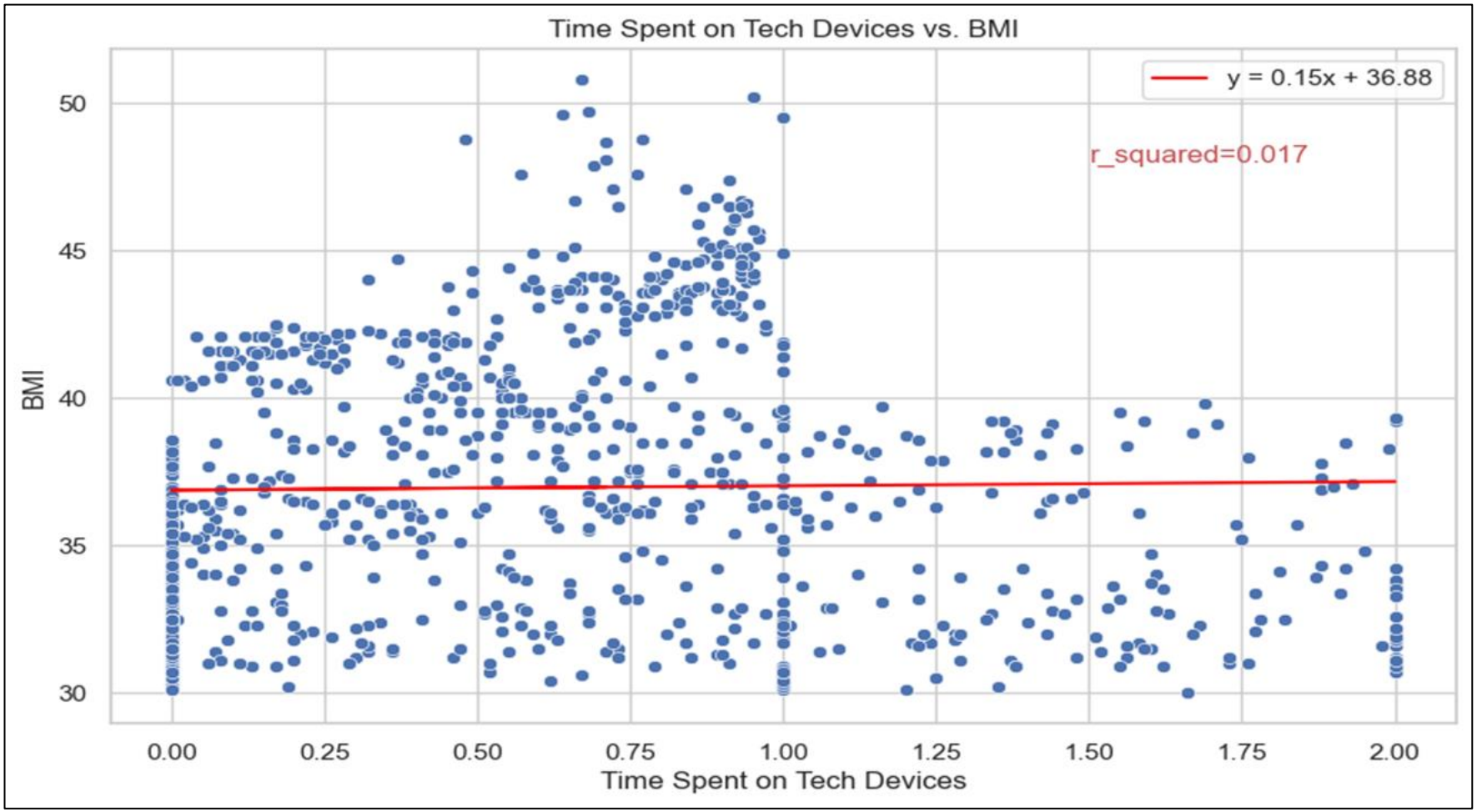
1 **R-Value**
-0.07

2 **P-Value**
0.02

Exploring Obesity Risk Factors

Time spent on tech devices vs BMI

R-Value = 0.017 , p-value = 0.59



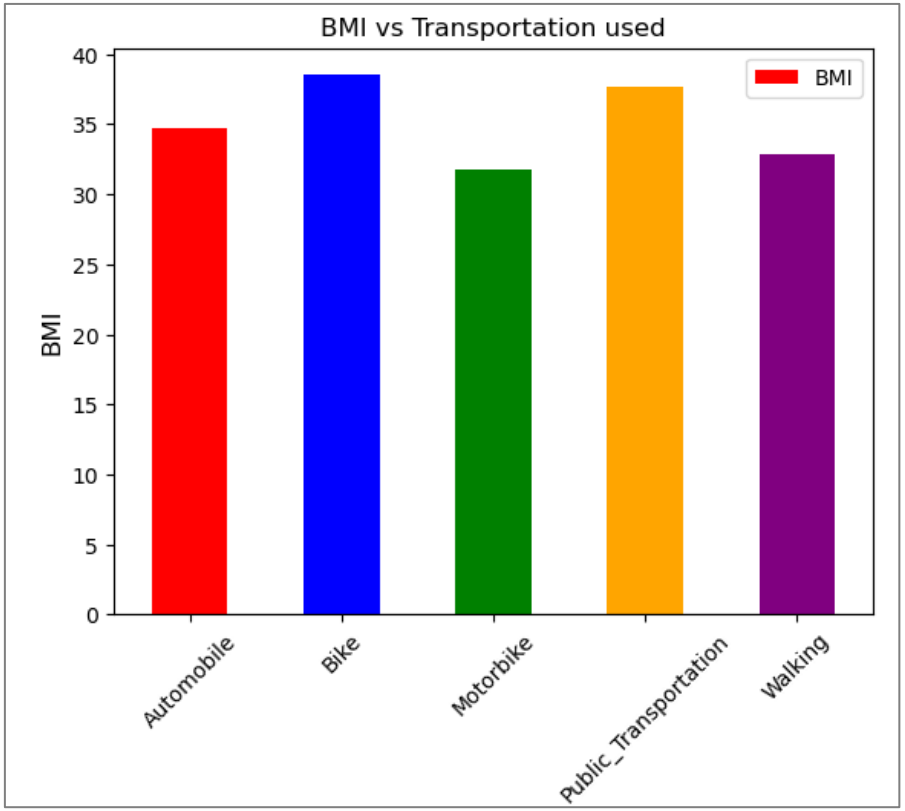
Exploring Obesity Risk Factors

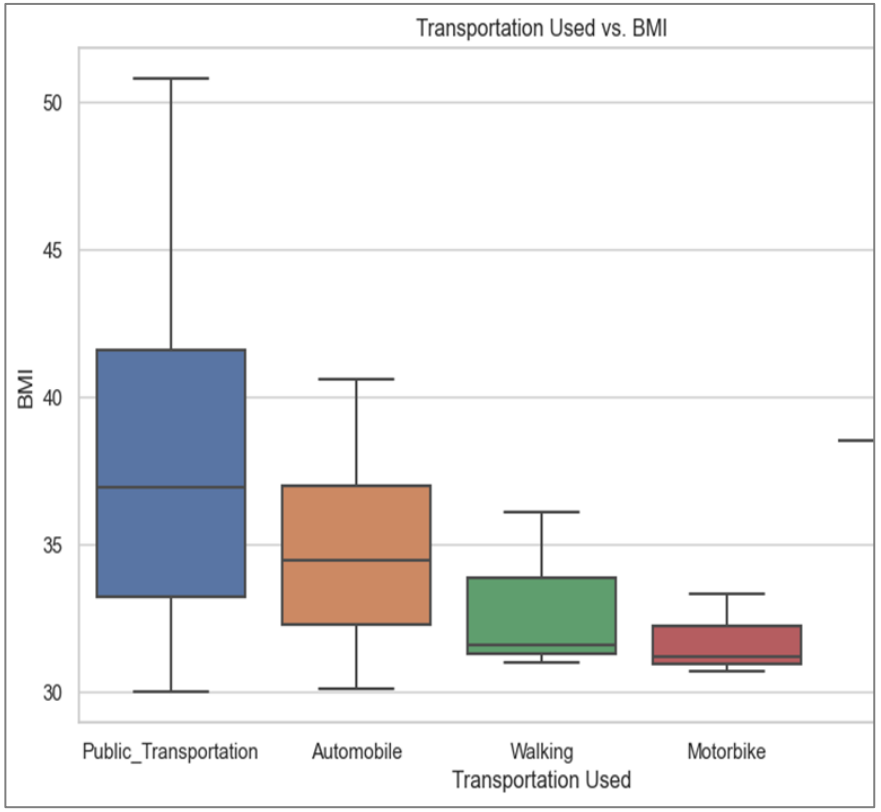
Time spent on devises& Freq of physical activity vs BMI

	Time spent on Tech devices	Freq of Physical Activity	BMI
Time spent on Tech devices	1.000000	0.136925	0.017458
Freq of Physical Activity	0.136925	1.000000	-0.072582
BMI	0.017458	-0.072582	1.000000

Exploring Obesity Risk Factors

Transportation used vs BMI





Results & Conclusion

Results:

1. ***Gender Distribution***: Nealy even split -- 1,068 (**51%**) males & 1,043 (**49%**) females.
2. ***Obesity Distribution***: Obese (46%), Overweight (27%), Normal weight (14%), and Underweight (13%)
3. ***Family History***: Overweight or obese respondents more likely to report family history of overweight.

Conclusion:

1. Dietary habits, alcohol consumption, and physical activity strongly affect obesity risk.
2. Crucial for targeted interventions and health policies.

Limitations:

1. Reliance on self-reported data may introduce biases.
2. Dataset's regional focus may limit generalizability.

References

- Fabio Mendoza Palechor, and Alexis Manotas. (2023). *Obesity or CVD risk (Classify/Regressor/Cluster) [Data set]. Kaggle.*

- NHANES. (2021). National Health and Nutrition Examination Survey. *Centers for Disease Control and Prevention. Retrieved from https://www.cdc.gov/nchs/nhanes/index.htm*

This Concludes Our Presentation

Thanks for your attention!