**Analysis of obesity levels based on eating habits and physical condition:**

**AKANKSHA SAMINDLA**

**PSID: 2302040**

**Descriptive:**

This project presents data for the estimation of obesity levels in individuals. Based on their eating habits and physical condition. The data contains 17 attributes and 2111 records, the records are labeled with the class No Obesity (Obesity Level), that allows classification of the data using the values of Insufficient Weight, Normal Weight, Overweight Level I, Overweight Level II, Obesity Type I, Obesity Type II and Obesity Type III.

**Data Source:** From UCI Repository

**https://archive.ics.uci.edu/dataset/544/estimation+of+obesity+levels+based+on+eating+habits+and+physical+condition**

**Motivation:**

Obesity is a major public health crisis, affecting millions of people worldwide. It is associated with various chronic health conditions, such as heart disease, diabetes, and certain cancers. Understanding its underlying causes and contributing factors is essential for addressing this significant health challenge.

This analysis would help to improve the quality of life for individuals, reduce healthcare costs, and develop effective strategies for prevention and treatment.

**This project can solve the below questions:**

1. What health consequences are associated with obesity in the population?
2. What are the key dietary habits of individuals in the population under study?
3. What is the physical activity level of the population?
4. What are the key lifestyle factors contributing to the individual's risk of obesity, including dietary habits, physical activity levels, and family history?
5. Is the individual aware of and actively managing their calorie intake, and are there behaviors such as smoking or excessive screen time that may impact their weight and overall health?

The context was chosen because obesity is a global public health concern associated with numerous chronic health issues, reduced quality of life, and increased healthcare costs.

**Goal:**

The objective of an analysis of obesity levels based on eating habits and physical condition is to understand and examine the relationship between the factors such as Age, height, weight, family history with overweight etc. This analysis can help raise awareness of how their own eating habits and physical condition might be contributing to their weight. It can motivate individuals to make healthier lifestyle choices.

**Key Elements in the dataset:**

* Gender
* Age
* Height
* Weight
* Family history with overweight
* High Caloric food
* Vegetables in the meal
* Meals in a day
* Any meals between foods
* Smoking
* Water consumption
* Calories
* Physical Activity
* Use of technological devices
* Alcohol consumption
* Mode of transportation

These elements are crucial for assessing an individual's obesity level because they collectively provide insights into various factors that influence weight and overall health.

**Analytical Approach:**

1. Data Preprocessing
2. Descriptive Statistics
3. Correlation Analysis and Regression Analysis.
4. Visualizations: Power BI visualization types such as Gauge Charts, Pie Charts, Line Charts, and Point Maps can be used for the overview page.
5. Recommendations/Conclusions.

**Benefits of analysis:**

* Individuals: Personalized advice to reduce obesity risk.
* Healthcare Professionals: Informed patient interventions.
* Public Health Authorities: Data for policy and campaigns.
* Researchers: Insights for further studies.
* Employers and Insurers: Tailored wellness programs and premiums.
* Fitness and Nutrition Professionals: Customized plans.
* Educational Institutions: Nutrition education.
* Community Organizations: Local health initiatives.
* Pharmaceutical and Healthcare Companies: Target populations.
* Government Policy Makers: Informed regulations and policies to combat obesity and promote healthier lifestyles.

**Analysis using Power BI Tool:**

**Data Visualizations:**

* Create various visualizations in Power BI to represent the data effectively. Common visualizations might include:
* Clustered charts or Stacked charts to show distribution.
* Pie charts to represent the percentage of individuals in different weight categories (e.g., underweight, normal weight, overweight, obese).
* Line charts to track trends over time.

**Interactive Dashboards**:

* Building interactive dashboards that allow us to explore the data. Use slicers, filters, and drill-through functionality to enable users to focus on specific aspects of the data.
* Matrix used for representation of count etc.

Calculation was made to obtain the mass body index for everyone, the results were compared with the data provided by WHO and the Mexican Normativity.

* Underweight 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

**Analysis using features collected during the survey:**

1. **Female and Male contribution towards obesity from the survey:**

* For finding the female count and male count which contributes to obesity and normal weights I have added new measure. In that, I used Dax function “Calculate along with count rows”.
* Then represented it using the matrix in Dashboard.
* Matrix table in Power BI is a versatile grid-like data visualization for

1. **Effect of main meals on Obesity:**

* Here I have used Donut chart for representing the effect of main meals on obesity.
* Donut charts are suitable for showing the composition of a whole, in this case, the relative contributions of different dietary components to obesity.
* Main meals 1,2,3,4 represents the below:
* 1-> 1 r 2 meals in a day
* 2-> 3 meals
* 3-> more than 3 meals
* 4-> no meals

**Observations:**

* The one who is having more than 3 main meals a day contributing for obesity compared to others.

1. **Obesity based on mode of Transportation**:

* I have used line chart for this case.
* A line chart in Power BI is a visual representation of data trends over time. It uses lines to connect data points on a graph, making it easy to see how values change. Line charts are great for tracking and analyzing time-based data like sales trends, stock prices, or any data that varies over time.

**Observations:**

* The one who are using public transportation contributes more obesity levels compared to other mode of transportations.

1. **Obesity w.r.t Family History with overweight:**

* Has a family member suffered or suffers from overweight? Yes/No
* I have used stacked Column chart for analyzing the above case
* Stacked charts are particularly useful for illustrating parts-to-whole relationships and comparisons among categories.

**Observations:**

* If the family History has overweight then, they are contributing obesity.

1. **Obesity based on physical activity**:

* How often do you have physical activity?
* No physical activity  🡪1
* 1 or 2 days. 🡪 2
* 2 or 4 days. 🡪 3
* 4 or 5 days. 🡪 4
* I have used pie chart for the above case.
* It's used to show the composition or distribution of a whole, making it easy to understand relative proportions briefly.

**Observations:**

* The one who is not having any physical activity is contributing to more obesity.

1. **Obesity based on high caloric food:**

* I have used stacked area chart for the above.
* A stacked area chart is a data visualization that shows multiple data series as layers of colored areas. Each layer represents a different category or component, and the vertical axis typically represents a quantitative value. Stacked area charts are used to display how these categories contribute to a whole over time or across categories.

**Observations:**

* The one who is consuming high caloric food is contributing to more obesity.

1. **Obesity based on water consumption:**

* I have used clustered column chart for the above.
* clustered column chart is a data visualization in which vertical columns are grouped or clustered together to represent different categories or data series. It's used to compare and display data in separate groups, making it easy to compare values within each group while also contrasting different groups.
* How much water do you drink daily?
* Less than a liter
* Between 1 and 2 L
* More than 2 L

**Observations:**

* On an average, the one who is drinking water between 1 – 2L is contributing more obesity from the data that has been collected during the survey.

1. **Obesity based on age:**

* Used line chart for this case.

**Observations:**

* Age between 18 – 38 is contributing more obesity compared to other ages using the data from survey.

1. **Alcohol versus Obesity:**

* How often do you drink alcohol?
* I do not drink.
* Sometimes
* Frequently
* Always
* I have used clustered bar char for the above.

**Observations:**

* The one who have habit of drinking of alcohol contributing more obesity.

**Conclusion:**

The following factors contribute to obesity based on the data from the survey:

1. **Physical Conditions:**

* Lack of Physical Activity.
* Age (specifically, the age group between 18 and 38).

1. **Food Habits:**

* Consumption of High-Caloric Food.
* Having more than three main meals a day.
* Drinking between 1 and 2 liters of water daily.
* Frequent or constant alcohol consumption.
* These factors are associated with higher obesity levels, but it's important to consider that obesity is a multifactorial condition influenced by various elements, including genetics and overall lifestyle choices. Further research may be needed to establish causation.