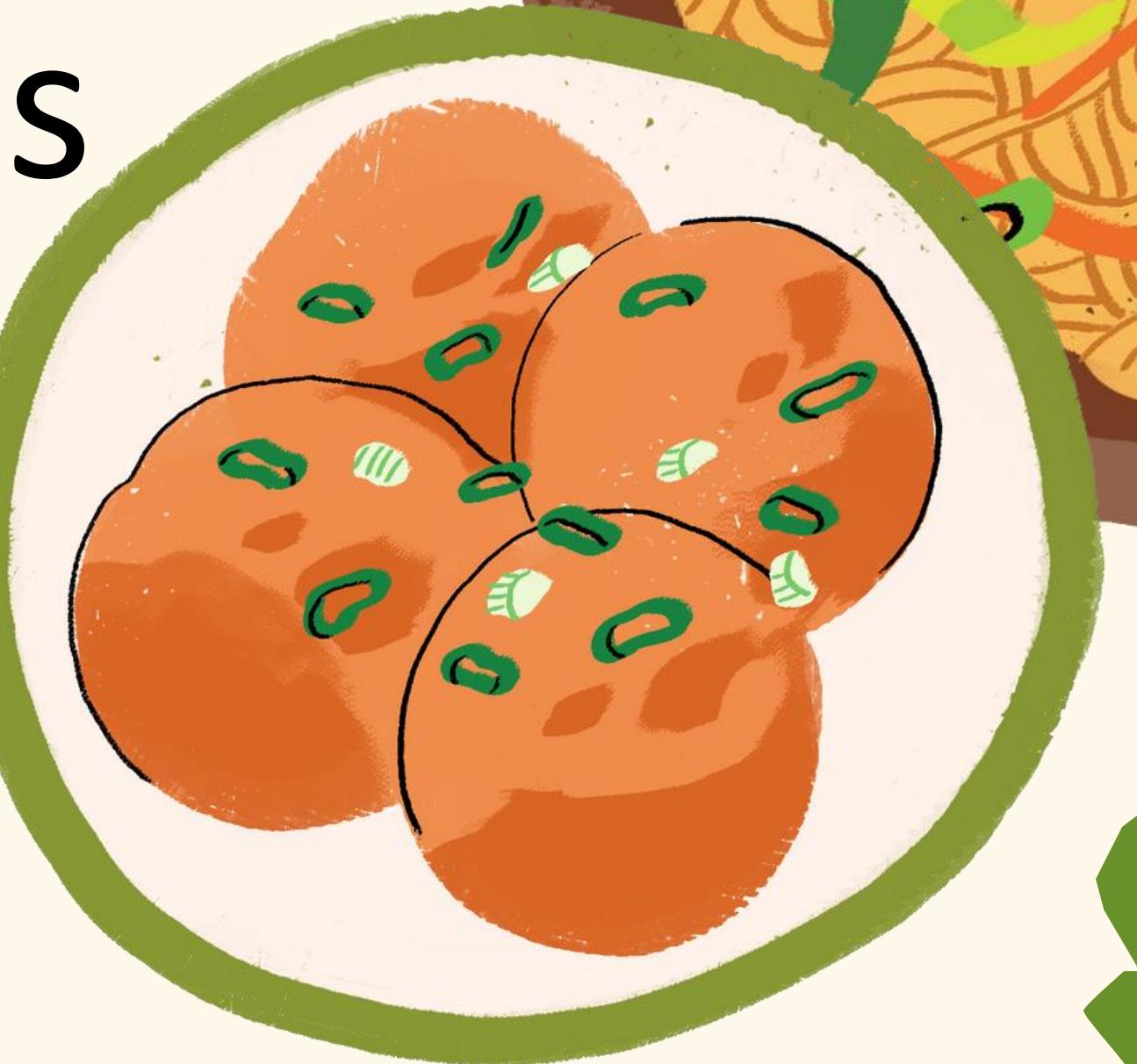


BIG DATA PROJECT

# Data-Driven Analysis of Global Obesity Trends

Done by Willy Hirwa

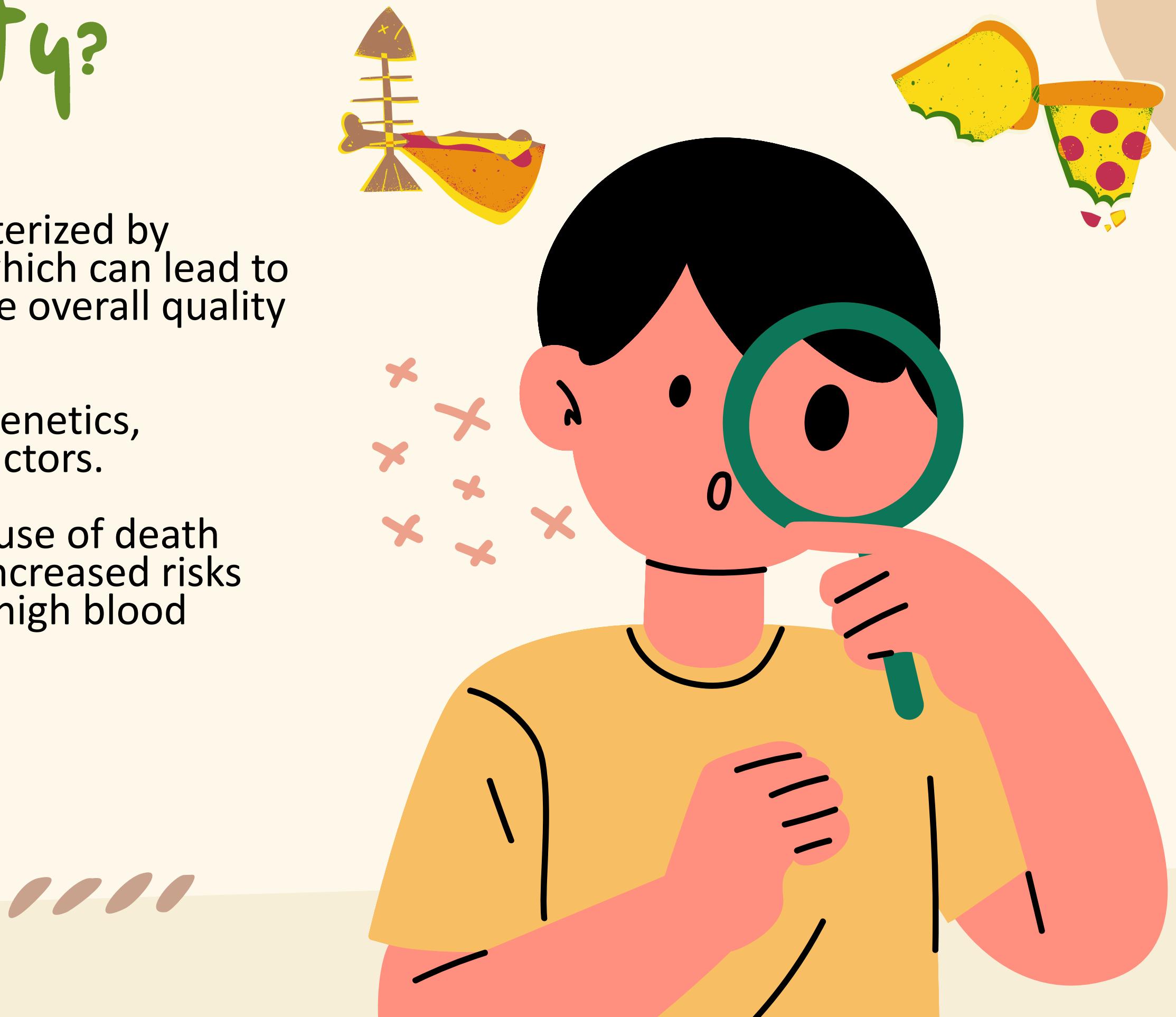


# WHAT is obesity?

Obesity is a chronic disease characterized by excessive body fat accumulation, which can lead to various health problems and reduce overall quality of life.

It's a complex issue influenced by genetics, environment, lifestyle, and other factors.

Obesity is a leading preventable cause of death worldwide and is associated with increased risks of type 2 diabetes, heart diseases, high blood pressure, and certain cancer



# Project introduction

## **Problem statement :**

- Global obesity rates have been rising, posing serious public health challenges.
- Objective: Analyze historical trends, identify high-risk countries, and forecast future patterns.
- This would help know the cause ,study it and know how to stop it

## **•Dataset Source:**

- WHO Global Health Observatory (GHO) — CSV dataset of adult obesity prevalence (%).

## **•Goal:**

- Provide insights through Python analysis and Power BI dashboard for policy and health planning.

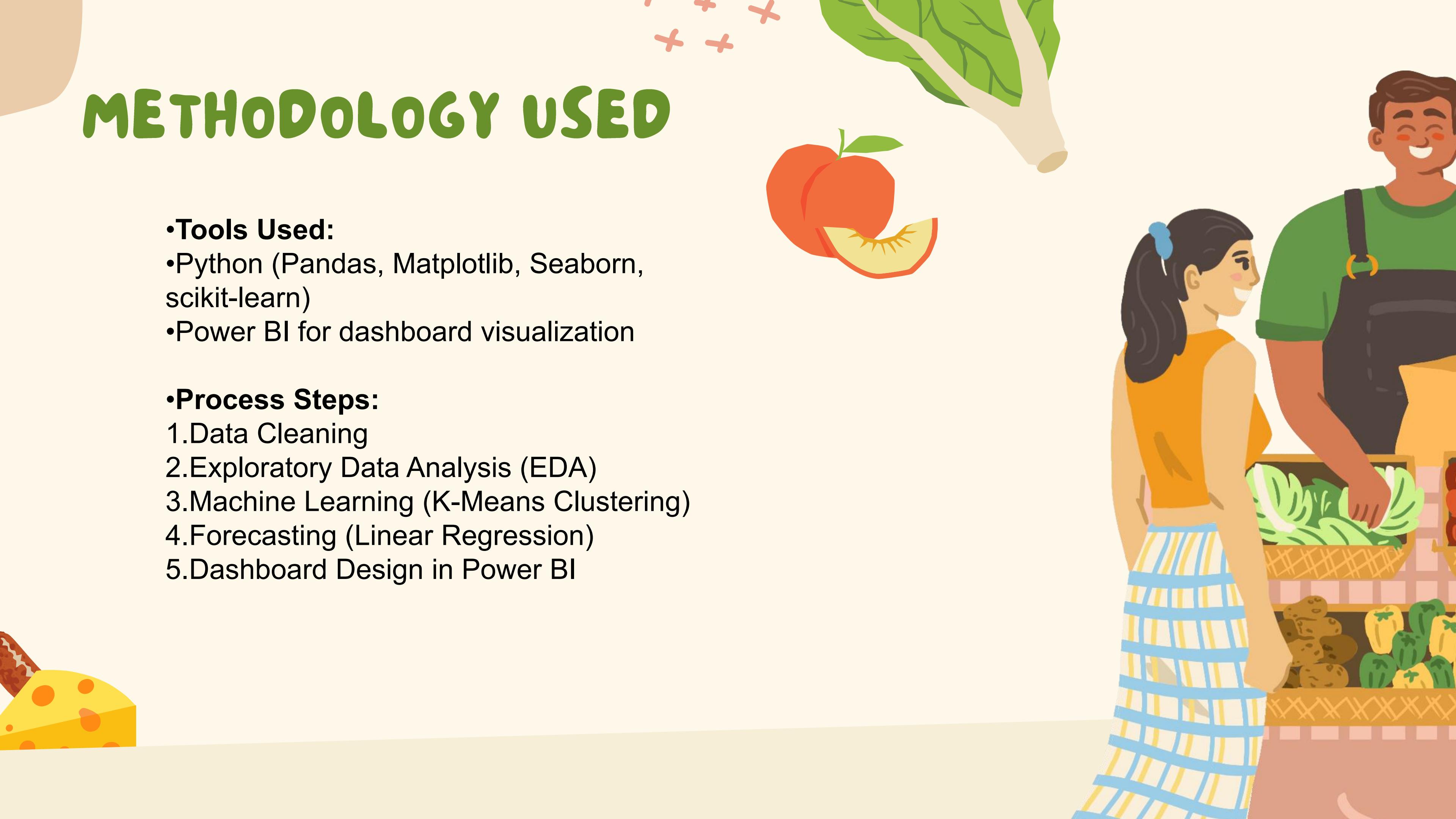
# METHODOLOGY USED

- **Tools Used:**

- Python (Pandas, Matplotlib, Seaborn, scikit-learn)
- Power BI for dashboard visualization

- **Process Steps:**

1. Data Cleaning
2. Exploratory Data Analysis (EDA)
3. Machine Learning (K-Means Clustering)
4. Forecasting (Linear Regression)
5. Dashboard Design in Power BI

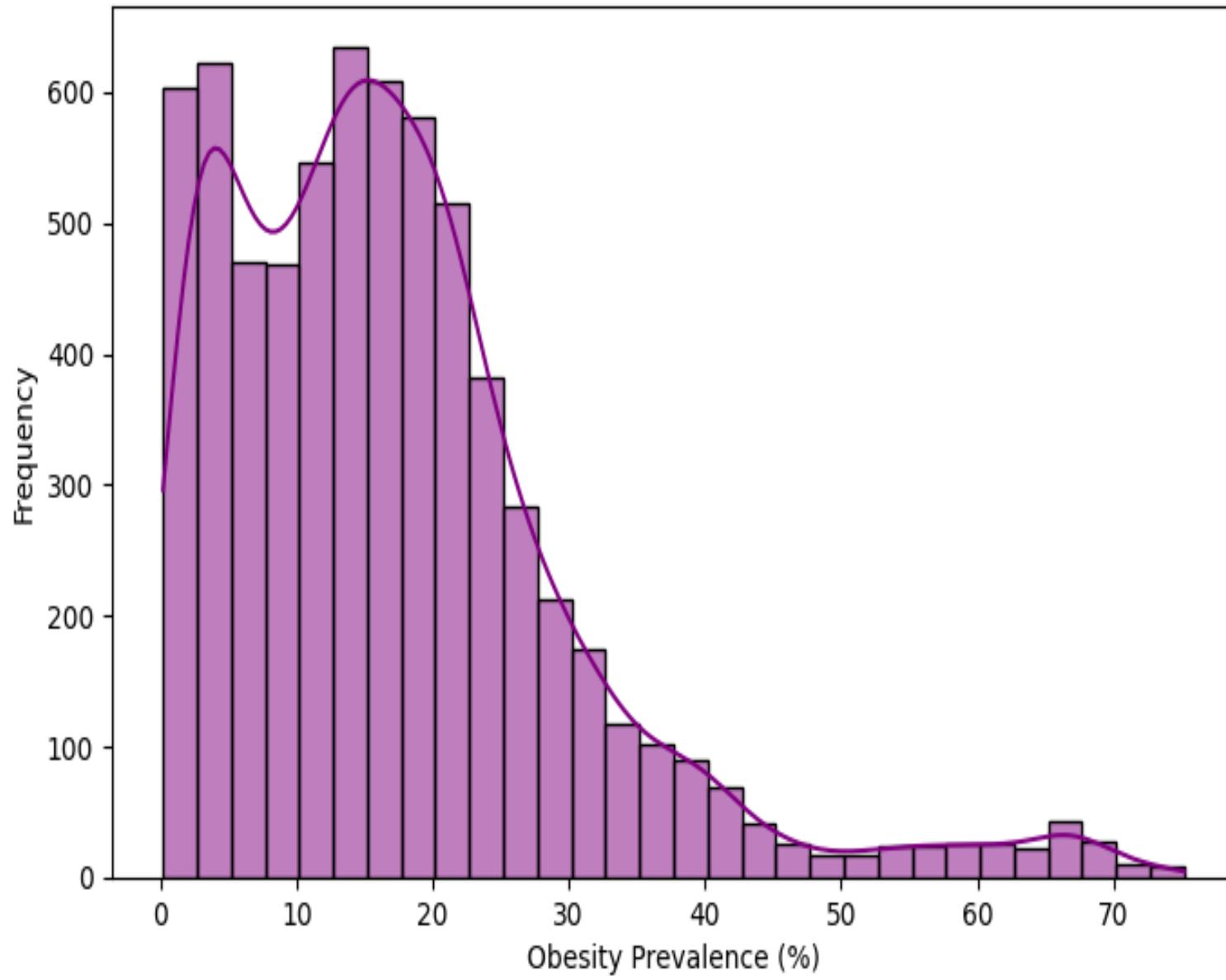


# Data cleaning

- Removed missing values, renamed columns, and ensured year formatting
- Dropped unnecessary columns (Code)
- Saved cleaned dataset as `cleaned_obesity_data.csv`
- Output: **200+ countries, 1990–2022, prevalence %**

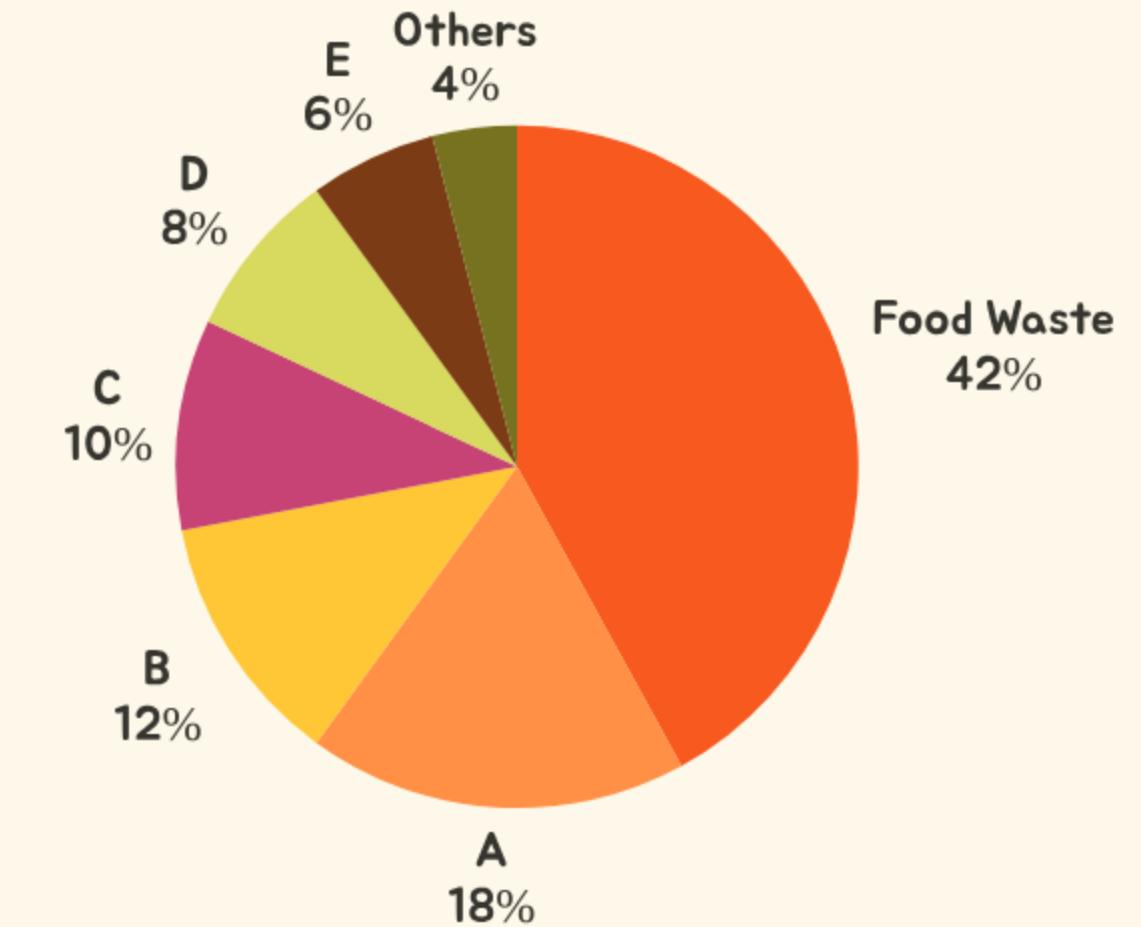
# EXPLORATORY ANALYSIS (EDA)

Distribution of Obesity Prevalence (%)



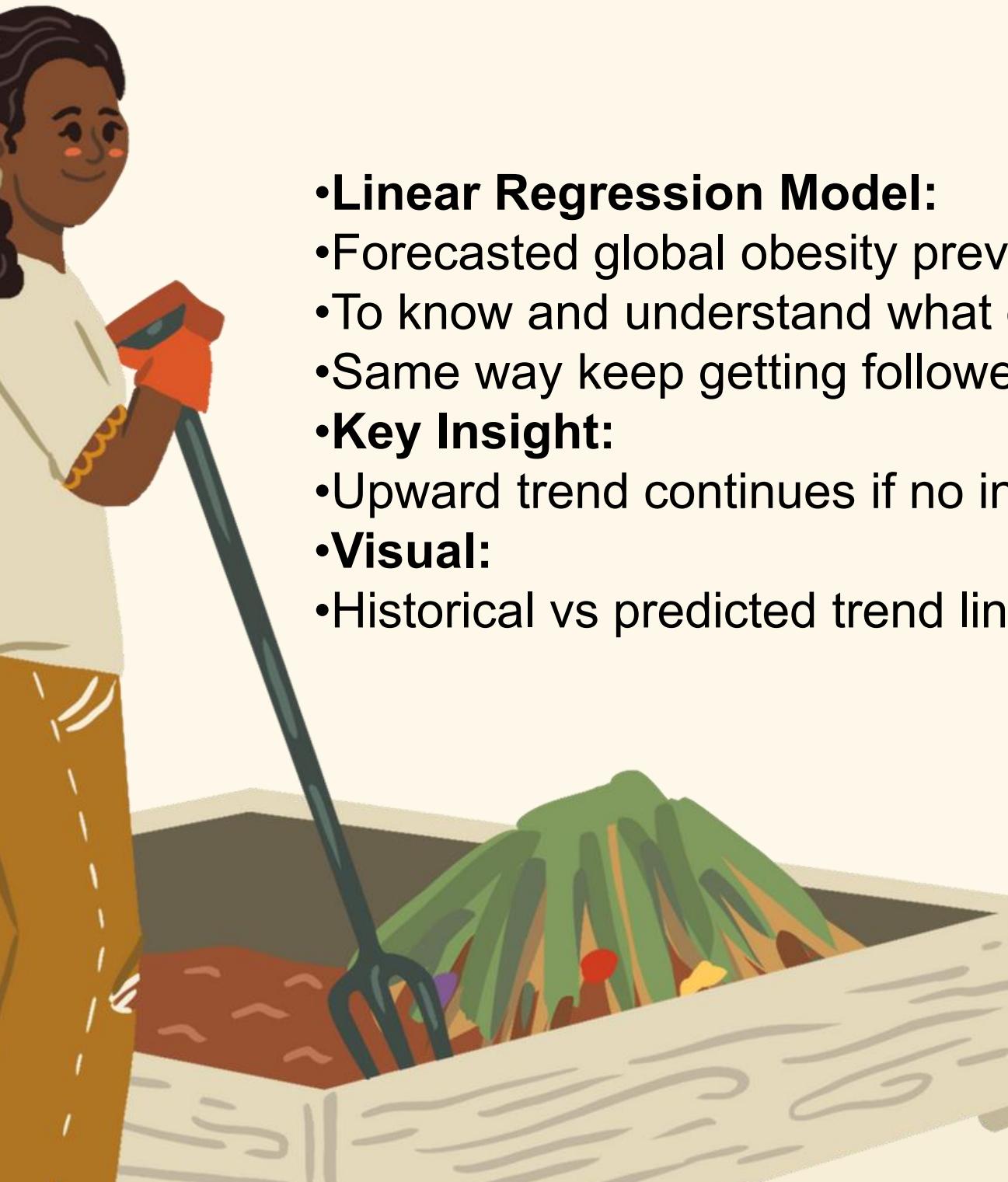
- Generated descriptive statistics
- Visuals:**
  - Global average obesity trend (1990–2022)
  - Top 10 and bottom 10 countries by prevalence (latest year)
  - Distribution histogram of prevalence
- Insights:**
  - Steady rise globally
  - Significant disparities between countries

# APPLYING CLUSTERING WITH MACHINE LEARNING



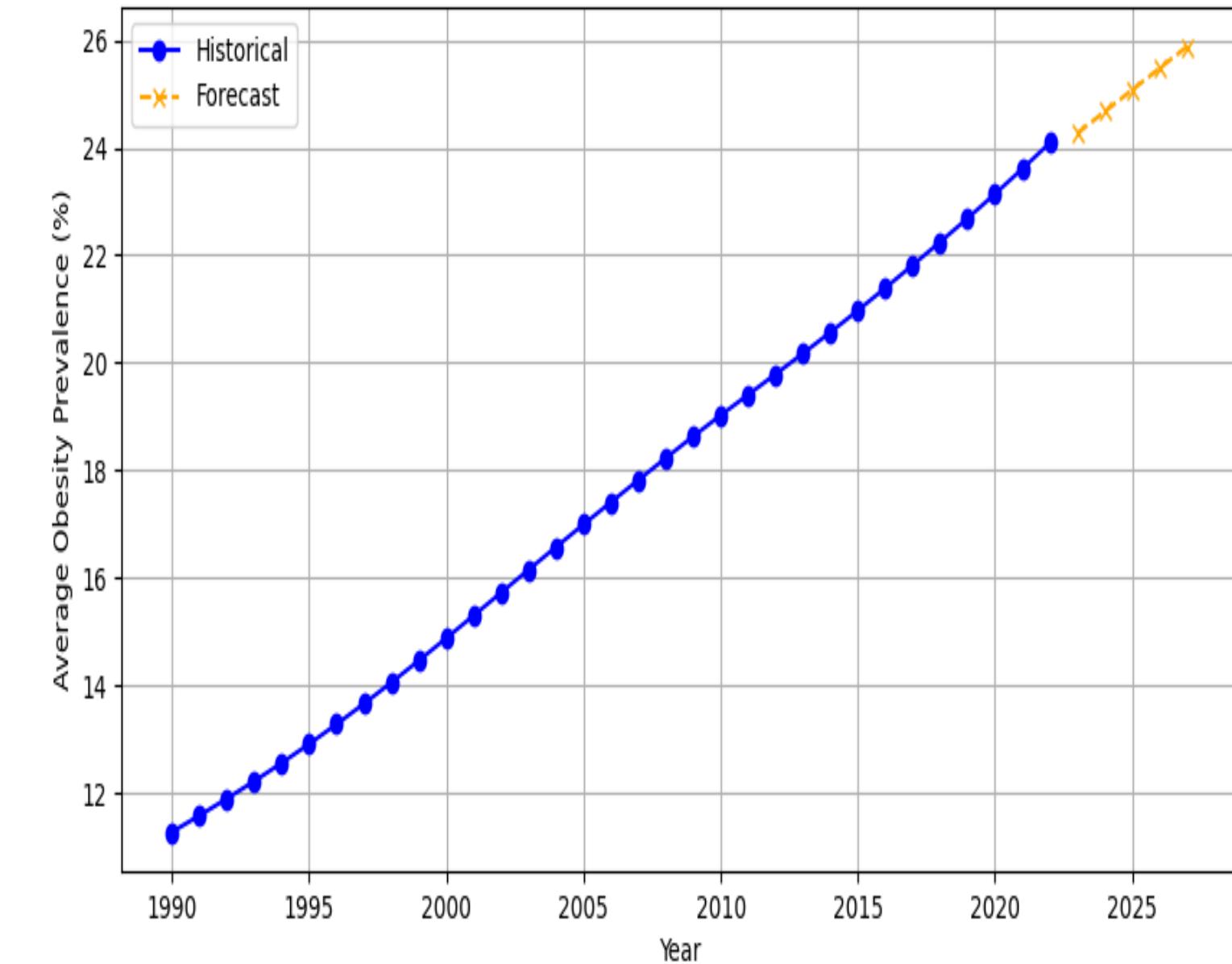
- **K-Means Clustering (3 groups):**
- Grouped countries into **Low, Medium, High prevalence clusters**
- **Evaluation:**
- Silhouette Score  $\approx$  (value you got)
- **Visual:**
- Scatter plot of clusters (colored groups)
- **Use Case:**
- Helps identify priority regions for interventions

# INOVATION WITH FORECASTING



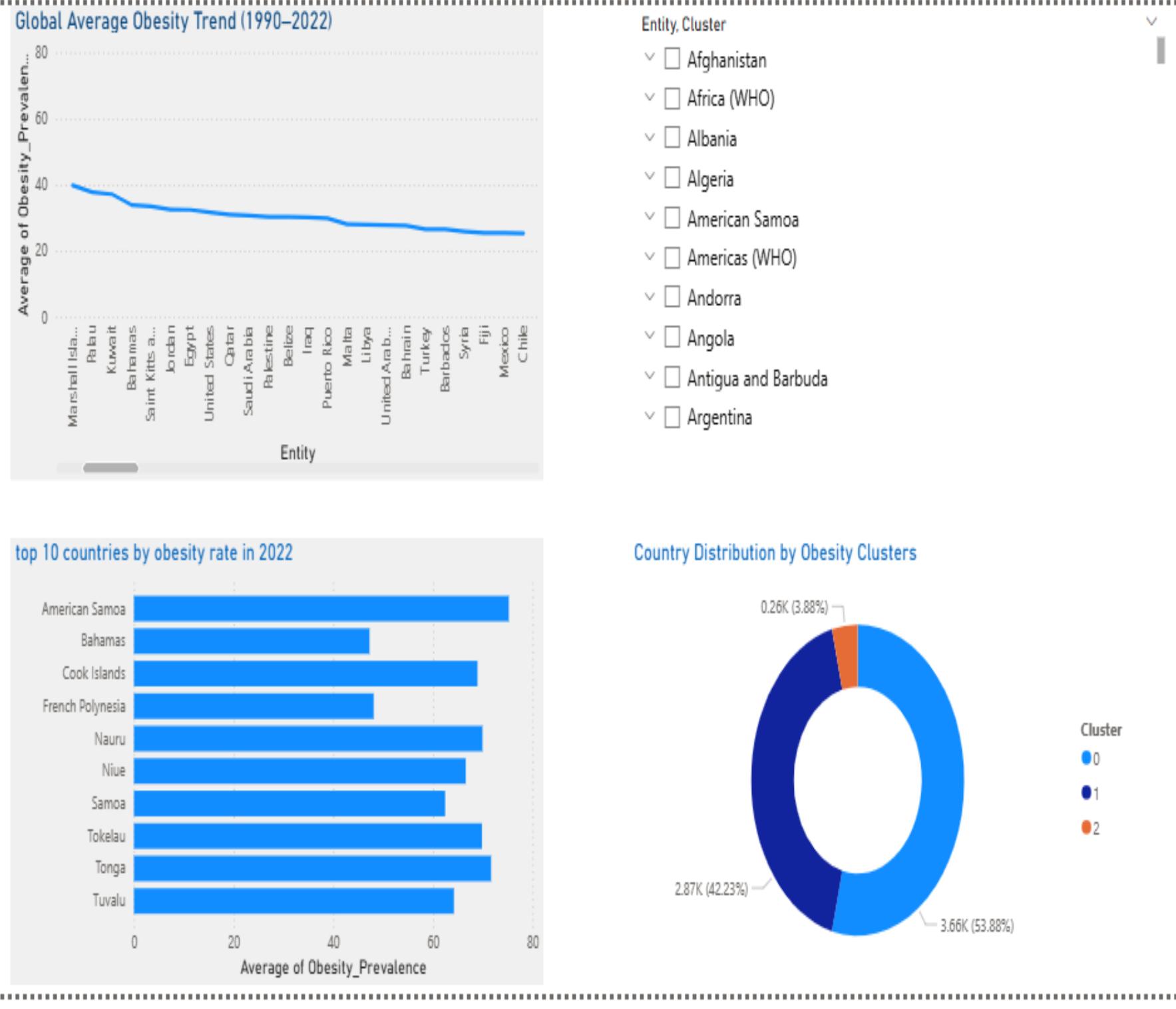
- **Linear Regression Model:**
- Forecasted global obesity prevalence (2023–2027)
- To know and understand what going to happen if the
- Same way keep getting followed
- **Key Insight:**
- Upward trend continues if no intervention
- **Visual:**
- Historical vs predicted trend line chart

Forecast of Global Obesity Trend (2023-2027)



# USING POWER BI FOR MORE CLEAR AND DESCRIPTiVЕ viSUALS

- Combined cleaned and clustered datasets
- **Visuals included:**
  - Global obesity trend (line chart)
  - Top/Bottom 10 countries (bar charts)
  - Cluster distribution (donut chart)
  - Year and country slicers for interactivity
- **Innovative Features:**
  - Forecasting insights
  - Slicers and dynamic filters
  - Color-coded clusters for risk levels



# WHAT CAN WE DO AFTER FINDINGS?

## •SUMMARY FINDINGS

- Global obesity prevalence increased steadily over last 3 decades
- Top 10 countries show >30% prevalence, bottom 10 <5%
- Clustering reveals clear risk groupings:
- High-risk (mostly developed nations)
- Medium-risk (emerging economies)
- Low-risk (developing nations)
- Forecast predicts continued rise post-2022

## •RECOMMENDATION

### •High prevalence countries:

- Implement stricter nutrition and activity policies

### •Medium prevalence countries:

- Preventive education and targeted interventions

### •Low prevalence countries:

- Maintain current health policies to prevent rise

### •Global:

- Collaborative strategies between WHO and local governments



## FUTURE WORK

- Integrate additional health indicators (BMI, diabetes rates, physical activity)
- Compare obesity trends against **GDP and urbanization** for deeper insights
- Explore **seasonal forecasting models** (ARIMA, Prophet) for better predictions
- Extend dashboard with **AI insights and drill-through analysis**



# THANK YOU!!!

Remember health is wealth  
@WILLY