

Hypothetical Case

- In 2016, the World Health Organization (WHO) reported that Obesity, Diabetes, and Blood Pressure are the three primary global health concerns.
- In response to this information, the UK's National Health Service (NHS) decided to sort an analysis of the data to understand the situation of the country.

What? (Details of the dataset)

- Dataset type: Table
- Data type: Items and attributes
- Attribute type:
 - i. Country/Region/World – Nominal (Categorical)
 - ii. Sex – Nominal (Categorical)
 - iii. Year – Ordinal (Interval)
 - iv. Prevalence of BMI>=30 kg/m² (obesity)– Ratio (Numerical)
 - v. Prevalence of raised blood pressure– Ratio (Numerical)
 - vi. Age-standardized diabetes prevalence– Ratio (Numerical)

Data Preprocessing

- The data regarding 'Prevalence of BMI \geq 30 kg/m \leq (obesity)', 'Prevalence of raised blood pressure', and 'Age-standardized diabetes prevalence' was present in three different sheets. A 'Union' operation was done on the sheets to be able to use them for analysis.
- The attributes of 'Prevalence of BMI \geq 30 kg/m \leq (obesity)', 'Prevalence of raised blood pressure', and 'Age-standardized diabetes prevalence' were converted to percentages and renamed as 'BMI in %', 'BP in %', and 'Diabetes in %' respectively.
- An extra column of 'Continents' was added for refined analysis.

Visualization tool used

Tableau Desktop 2023.2.0 was used for analysis and visualization.

Report structure

- The report is divided into 2 parts. The first part analyses the pattern for one measurement which is 'Prevalence of BMI \geq 30 kg/m \leq (obesity)' as it is considered as the root cause of other diseases.
- The second part explores the relationship between all three measurements namely 'Prevalence of BMI \geq 30 kg/m \leq (obesity)', 'Prevalence of raised blood pressure', and 'Age-standardized diabetes prevalence'. The report aims to explain the findings and aim of each visualization while explaining the marks and channels used along with justifying the fitness of the visual.

Analysis 1 – Analysis with single measurement: Prevalence of BMI ≥ 30 kg/m² (obesity)

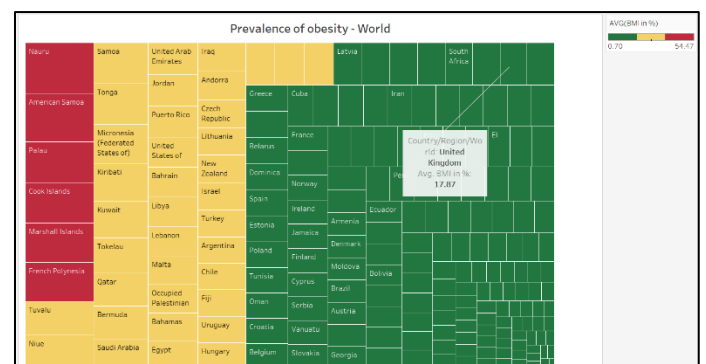
Visual 1: Tree map

1. Finding 1

- A total of 6 countries falls under red zone.
- Majority of the countries fall under green zone and hence have least prevalence of obesity.
- UK falls under green zone. Hence, we can conclude that UK is safer, and the prevalence of Obesity is much lower in the UK when compared to other countries.

2. Why?

- a. **Aim of the visualization:** The aim of the visualization is to show the global prevalence of obesity over the years. The visual divides the countries into 3 zones namely red (maximum prevalence), yellow (medium prevalence) and green (least prevalence) zones to demonstrate the severity of the prevalence of BMI>=30. The United Kingdom was annotated to show that it falls under green zone.
- b. **Actions**
- Analyze: To **discover** how obesity rates vary across countries. To **annotate** the position of UK.
 - Search: To **locate** the United Kingdom.



- iii. Query: To **summarize** the most and least impacted countries (by obesity).

c. Targets

- i. Attribute: This visual deals with 'Year', 'BMI in %' and 'Country/Region/World' to **similarities** between the countries on the basis of prevalence of obesity.

3. How?

a. Visual mapping: Marks, channels and attributes mapped (filters used if any)

- i. Marks: Area
- ii. Channels:
 - Size and position – severity of prevalence of obesity. The boxes are sorted in descending order of prevalence of obesity.
 - Color – Categorizes the countries
 - Label – To show the countries' names

b. Effectiveness of visual mapping and design: Fit of the chosen chart type and visual channels (filters of any)

- i. The visual deals with 'BMI in %' (ratio) and 'Country/Region/World' (categorical). The visuals suitable for these data types are: square pie, tree map, bubble hierarchy and pie chart. Tree map was considered as the most effective visual because it uses area, position, color hue as visual variables. For categorical and quantitative data types, position is the most effective followed by color hue and area. Bubble hierarchy is used for hierarchical display which was not the aim of this visual. Square pie uses symbol as one of its visual variables which holds the least effectiveness. We avoid pie charts as they are most contentious chart type.
- ii. The annotation 'Mark' was used to highlight the United Kingdom.

Visual 2: Line graph

1. Finding 2

- Since 1975, women have shown higher obesity rates than men.
- The prevalence of obesity in both the genders have risen constantly over the years.
- The obesity rate in women went above the average from 1994 and that in men from 2000.

2. Why?

- a. **Aim of the visualization:** The aim of the visualization is to observe the trend of prevalence of obesity amongst the sexes from 1975 to 2016 in the UK.

b. Actions

- i. Analyze: (a) To **discover** patterns, distinctions, or resemblances in the occurrence of obesity across various years in the United Kingdom. (b) To **annotate** the start year, end year and the year from when the obesity started going beyond average.
- ii. Search: To **explore** if there are differences in the way obesity has changed over the years between the sexes in UK. (b) To **locate** the year when the obesity% went above the average for the sexes.
- iii. Query: To **compare** the prevalence of obesity amongst men and women from 1975 to 2016.

c. Targets

- i. Attribute: The goal is to retrieve data on obesity across various years for each gender. The objective centers on the attributes 'Year,' 'Sex,' and 'BMI in %' to comprehend the **similarities** between men and women concerning fluctuations in obesity.

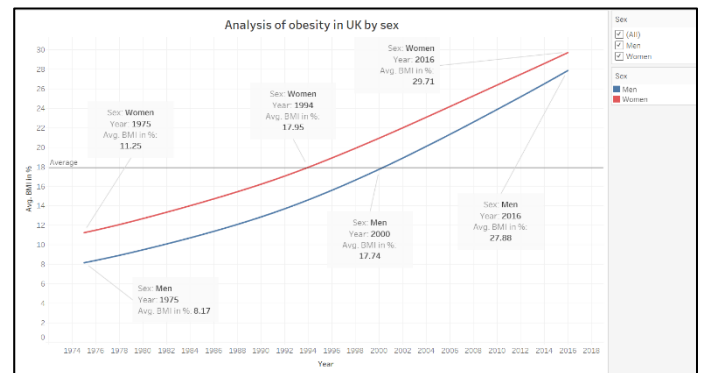
3. How?

a. Visual mapping: Marks, channels and attributes mapped (filters used if any)

- i. Marks: Line
- ii. Channels:
 - Colour – Distinguishes between the genders
 - Position (vertical and horizontal) – Avg BMI in % vs Year

b. Effectiveness of visual mapping and design: Fit of the chosen chart type and visual channels (filters of any)

- i. The visual deals with 'BMI in %' (ratio), 'Year' (interval) and 'Sex' (categorical). The most appropriate visuals would have been line chart, area chart, stacked area chart, and stream graph. Line graph is used because the visual variables are position, slope and color- hue which are more effective as compared to other variables used in other graphs. We dropped other types of charts because they use area as one of the visual variables which is least desired in case of categorical data types.
- ii. The annotation 'Mark' was used to highlight the start year, end year and the year from when the obesity started going beyond average are annotated.



Visual 3: Forecast chart

1. Finding 3

- It is estimated that the prevalence of obesity in the UK across sexes would increase to 33.72 % in the year 2023.
- It is estimated that the prevalence of obesity in the UK for women would increase to 34.10% in the year 2023.
- It is estimated that the prevalence of obesity in the UK for men would increase to 33.72 % in the year 2023.

2. Why?

- Aim of the visualization:** The aim of the visualization is to predicting the future of prevalence of obesity in UK across sexes. The predicted prevalence of obesity for the year 2023 was annotated.

b. Actions

- Analyze: To **discover** future stand of the prevalence of obesity in the UK. To **annotate** the predicted prevalence of obesity in the UK for year 2023
- Search: To **browse** to the future state of obesity in UK.
- Query: To **identify** the current and future trend of obesity amongst men and women in UK.

c. Targets

- Data: The goal is to **predict** the future **trend** of obesity in UK. The objective centers on all data over the years.

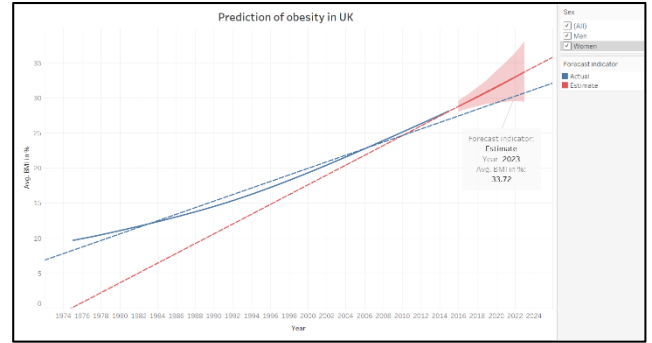
3. How?

a. Visual mapping: Marks, channels and attributes mapped (filters used if any)

- Marks: Line
- Channels:
 - Colour – Distinguishes between the actual and estimated values of obesity
 - Position (vertical and horizontal) – Avg BMI in % vs Year
 - Filters – ‘Sex’

b. Effectiveness of visual mapping and design: Fit of the chosen chart type and visual channels (filters of any)

- The visual deals with ‘BMI in %’ (ratio), ‘Year’ (interval). The aim is to predict the future value of the prevalence of obesity and hence line chart is used as it gives distinctive view of the trends. Also, the visual variables are position, slope and color- hue which are more effective.
- Filter on ‘Sex’ was used to understand the future predictions separately.
- The annotation ‘Mark’ was used to highlight predicted value of obesity for the year 2023.



Visual 4: Bar graph

1. Finding 4

- Countries like South Africa, Syrian Arab Republic, Russian Federation, Bulgaria, Greece, Ukraine, Suriname, and Venezuela have the approximately 18% (rounded-off) of prevalence of obesity which is same as the UK.

2. Why?

- Aim of the visualization:** The aim of the visualization is to compare and see the countries which have approximately same (18%) prevalence of obesity as the UK.

b. Actions

- Analyze: To **discover** the countries having same obesity rates as that of the UK.
- Search: To **browse** for the countries with obesity rates equivalent to the UK.
- Query: To **compare** the prevalence of obesity amongst countries, **summarize** the countries with same obesity rates as that of the UK.

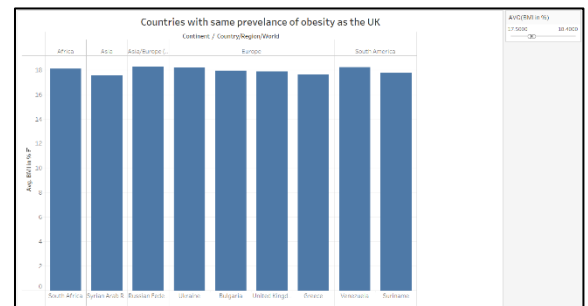
c. Targets

- Attributes: The goal is to find **similarity** between countries on the basis of prevalence of obesity over the years.

3. How?

a. Visual mapping: Marks, channels and attributes mapped (filters used if any)

- Marks: Line
- Channels:



- Position (Both horizontal and vertical): To depict the percentages
 - Filter – On 'BMI in %' to ensure that countries having similar obesity rates are compared.
- b. **Effectiveness of visual mapping and design: Fit of the chosen chart type and visual channels (filters of any)**
- The visual deals with 'Country/Region/World', 'Continents' and 'BMI in %'. A table is well suited for displaying an organized and structured data. It is easily comprehensible as the table has grouped the information and has created a sense of proximity.
 - Filter on 'BMI in %' was used to compare and list the countries having similar obesity rates only.

Analysis 2 – Analysis with two or three measurements

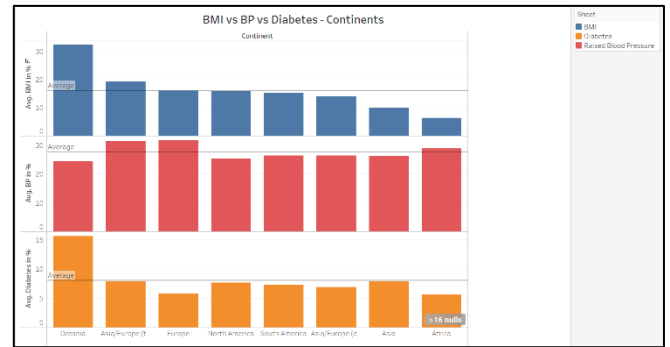
Visual 5: Bar Chart

1. Finding 5

- Oceania shows an opposite behavior as compared to the other continents in the spread of BMI, BP and diabetes.
- In Oceania, the BMI and diabetes are above the average level but the BP is below average level.

2. Why?

- Aim of the visualization:** The aim of the visualization is to understand the spread of obesity, BP and diabetes at continent level.
- Actions**
 - Analyze: To **discover** the distribution of obesity, blood pressure, and diabetes on a continental scale.
 - Search: To **browse** for the continent/s with unusual trend.
 - Query: To **compare and summarize** the prevalence of obesity, BP and diabetes amongst continents.
- Targets**
 - Attributes: The goal is to find **outlier** in the spread of obesity, BP and diabetes.



3. How?

- Visual mapping: Marks, channels and attributes mapped (filters used if any)**
 - Marks: Line
 - Channels:
 - Colour: For attributes of BMI, BP and diabetes
 - Position (Horizontal and vertical): To show the percentage spread in each continent
- Effectiveness of visual mapping and design: Fit of the chosen chart type and visual channels (filters of any)**
 - The visualization focuses on 'Continents', 'BMI in %', 'Diabetes in %', and 'BP in %'. In this scenario, both area charts and bar charts can be employed. However, a bar chart is more suitable as it enables precise comparisons, both in terms of relative and absolute values. Additionally, it utilizes position and color as visual variables, which are more effective at a higher level in the visual hierarchy.

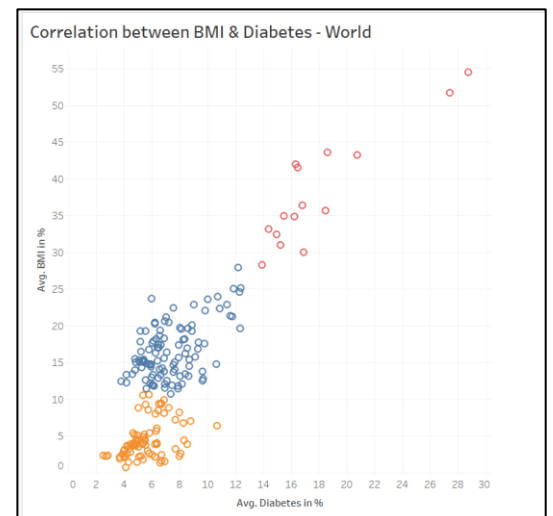
Visual 6: Scatter plot

1. Finding 6

- There is a positive correlation in prevalence of obesity (BMI) and Diabetes across the world.
- The countries were divided into 3 (using elbow method) clusters (as shown through different colors) and it was observed that the clusters so formed are distinct. This means that countries in each cluster behave similar.

2. Why?

- Aim of the visualization:** The aim of the visualization is to observe the correlation between BMI and diabetes. The visualization also divided the countries into 3 clusters to understand how similar the countries behave.
- Actions**
 - Analyze: To **discover** correlation between BMI and diabetes across the countries.
 - Search: To **explore** the correlation, patterns and clusters of countries which behave similar with respect to BMI and diabetes.
 - Query: To **identify** correlation between BMI and diabetes.



c. **Targets**

- i. Attributes: The goal is to find **similarity** between countries on the basis of **correlation** between BMI and diabetes.

3. **How?**

a. **Visual mapping: Marks, channels and attributes mapped (filters used if any)**

- i. Marks: Dot or circle
- ii. Channels:
 - Colour: To show the clusters of countries
 - Position (Both horizontal and vertical): Each dot represents the correlation value.

b. **Effectiveness of visual mapping and design: Fit of the chosen chart type and visual channels (filters of any)**

- i. The visual deals 2 quantitative variables which are 'BMI in %' and 'Diabetes in %' and the aim was to find the correlation, explore patterns and clusters. In this scenario, scatter plot is the most appropriate choice. It uses color and position as the channels which are the most effective in the visual hierarchy.

Visual 7: Scatter plot

1. **Finding 7**

- There is little to no correlation between BMI and diabetes for the countries having similar prevalence of obesity (Finding 4) as that of the UK.

2. **Why?**

- a. **Aim of the visualization:** The visual aims to find the correlation between BMI and diabetes for the countries which had similar prevalence of obesity (according to Finding 4) as that of the UK.

b. **Actions**

- i. Analyze: To **discover** correlation between BMI and diabetes for countries having similar BMI as that of the UK. To **annotate** the country with maximum correlation between BMI and diabetes.
- ii. Search: To **locate** for the country with maximum correlation between BMI and diabetes.
- iii. Query: To **compare** correlation between BMI and diabetes and identify the country with the maximum correlation between BMI and Diabetes for countries having similar BMI as that of the UK.

c. **Targets**

- i. Data: The goal is to target the **outlier** with maximum correlation between BMI and diabetes amongst the countries having similar prevalence of obesity as that of UK.

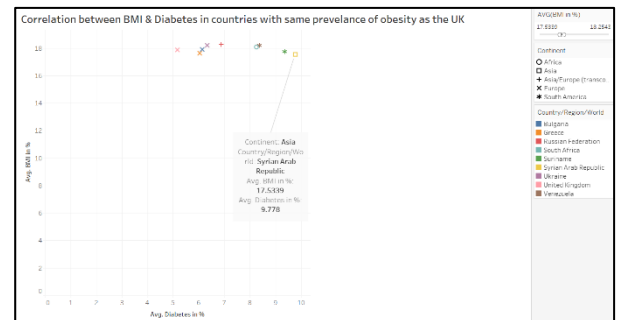
3. **How?**

a. **Visual mapping: Marks, channels and attributes mapped (filters used if any)**

- i. Marks: Point/ Dot
- ii. Channels:
 - Position (Both horizontal and vertical): Value of correlation
 - Colour: Country/ Region/ World (To distinguish between the countries)
 - Shape: Continents (To distinguish between continents as effective grouping is done)
 - Filter: On 'BMI in %' to ensure that countries having similar obesity rates are compared.

b. **Effectiveness of visual mapping and design: Fit of the chosen chart type and visual channels (filters of any)**

- i. The visual deals 2 quantitative variables which are 'BMI in %' and 'Diabetes in %' and the aim was to find the correlation, explore patterns and clusters. In this scenario, scatter plot is the most appropriate choice. It uses color and position as the channels which are the most effective in the visual hierarchy.
- ii. Filter on 'BMI in %' was used to compare and list the countries having similar obesity rates only.



Visual 8: Bar chart

1. Finding 8

- For the countries sharing the similar obesity rates as that of the UK, the obesity rates have gone beyond the average values but the diabetes rates have always been below the average values.

2. Why?

- Aim of the visualization:** The aim of the visualization is to compare the prevalence rates of obesity and diabetes amongst the countries which have approximately same (18%) prevalence of obesity as that of the UK.

b. Actions

- Analyze: To **discover** the spread of BMI and diabetes in countries having same obesity rates as that of the UK.
- Search: To **browse** trends of BMI and diabetes in countries with obesity rates equivalent to the UK.
- Query: To **compare and summarize** the prevalence of obesity and diabetes amongst countries with same obesity rates as that of the UK.

c. Targets

- Attributes: The goal is to find **similarity** between countries on the basis of prevalence of obesity and diabetes.

3. How?

a. Visual mapping: Marks, channels and attributes mapped (filters used if any)

- Marks: Line

- Channels:

- Position (Both horizontal and vertical): To depict the percentages
- Colour: To distinguish between BMI in % and Diabetes in %
- Filter – On 'BMI in %' to ensure that countries having similar obesity rates are compared.

b. Effectiveness of visual mapping and design: Fit of the chosen chart type and visual channels (filters of any)

- The visual deals with 'Country/Region/World', 'Continents', 'BMI in %' and 'Diabetes in %'. In such a case area chart, bar chart can be used. Bar chart is more appropriate because it allows us to make accurate comparisons (both relative and absolute) between categories. Also, it uses position and color as its visual variables which are at a higher level of effectiveness in the visual hierarchy.
- Filter on 'BMI in %' was used to compare the countries having similar obesity rates only.

