**Data visualization project report**

**Made by:** Osama Mohammed Abuzraiq

**Student ID:** 202010646

|  |  |
| --- | --- |
| **Title of the Dataset** | Suicide rates overview 1985 to 2016 |
| **Link to the Dataset** | https://www.kaggle.com/datasets/russellyates88/suicide-rates-overview-1985-to-2016 |
| **Number of Samples** | 27820 |
| **Number of Attributes** | 12 |

* Total number of sheets: **6**
* Total number of visualizations: **5**
* Total number of tables: **1**

Table of Contents

[**Sheet 1,** Suicide by country: 2](#_Toc155307639)

[**Sheet 2,** Suicide over the years: 3](#_Toc155307640)

[**Sheet 3,** Suicide by age group: 4](#_Toc155307641)

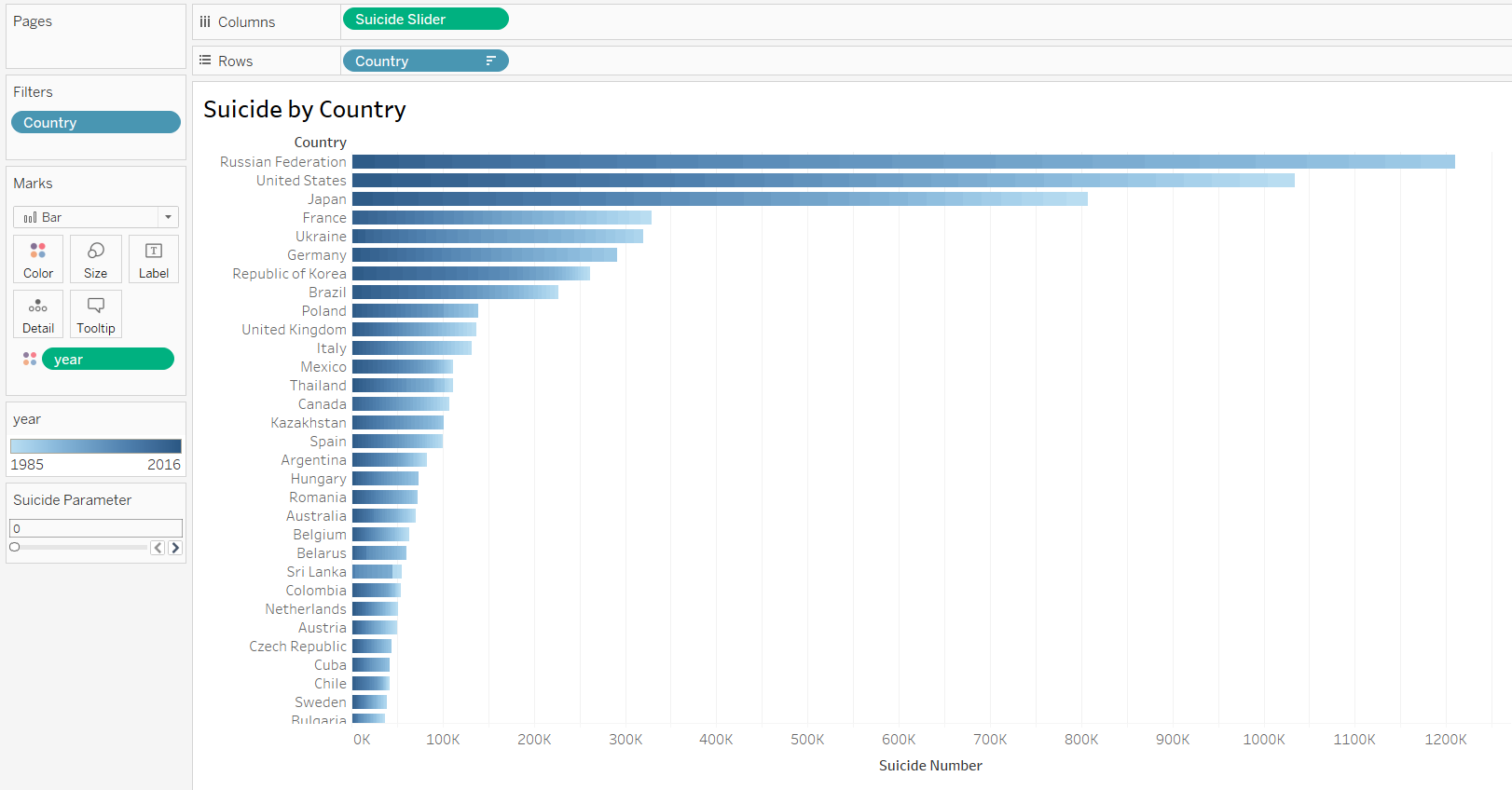
[**Sheet 4,** Suicide by generation and gender: 5](#_Toc155307642)

[**Sheet 5,** Suicides per 100k population: 6](#_Toc155307643)

[**Sheet 6,** Table view of Suicide No: 7](#_Toc155307644)

**January, 2024**

# **Sheet 1,** Suicide by country:

****

This sheet visualizes the ‘Suicide No for each ‘Country’ on a horizontal bar chart, and also shows the years in gradient color.

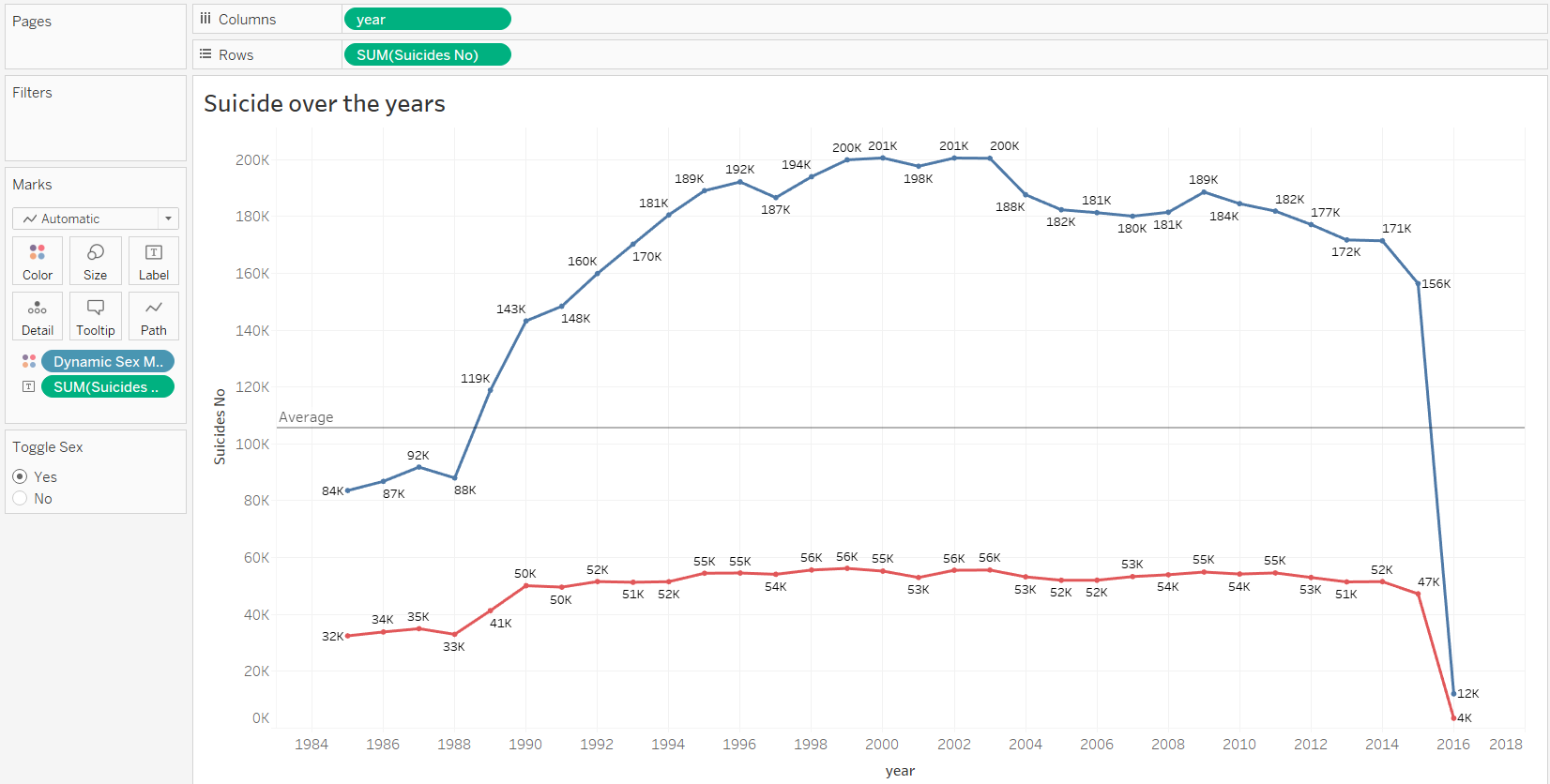
There is also a parameter named ‘Suicide Parameter’ of data type Integer that uses a range slider for changing values.

* **Rows:** Contains the ‘Country’ dimension from the dataset. while the columns field.
* **Columns:** Contains a measure calculated field named ‘Suicide Slider’, this calculated field is linked to the ‘Suicide Parameter’ which we can use to show only the countries with **suicides > Suicide Parameter.**
* **Marks:** Contains the ‘Year’ in the color mark.

**Suicide Slider Calculated Field code:**

IF [Suicides No] >= [Suicide Parameter] THEN [Suicides No] END

# **Sheet 2,** Suicide over the years:



In this sheet we have a continuous line visualization of ‘Suicide No’ over the years, a reference line has been added to show the average value across all years and there is a parameter to toggle a color mark for ‘Sex’ on or off named ‘Toggle Sex’ which is of Boolean datatype.

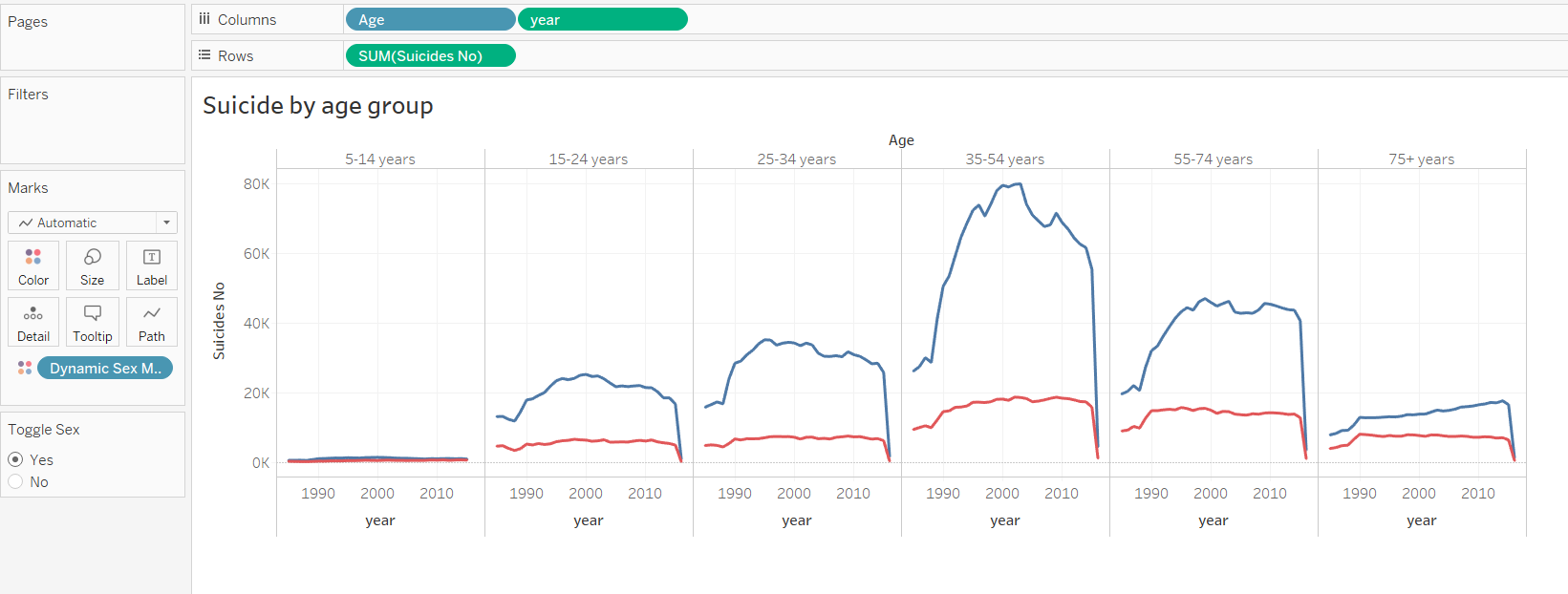
* **Rows:** Contains the sum of the ‘Suicide No’ measure.
* **Columns:** Contains the ‘Year’ measure.
* **Marks:** Contains the ‘Dynamic Sex Mark’ in the color mark which can be toggled on or off using the ‘Toggle Sex’ parameter.

Furthermore, mark labels are enabled and they show the sum of ‘Suicide No’, and the display units were changed to Thousands (K)

**Dynamic Sex Mark calculated field code:**

IF [Toggle Sex] = TRUE THEN [Sex] END

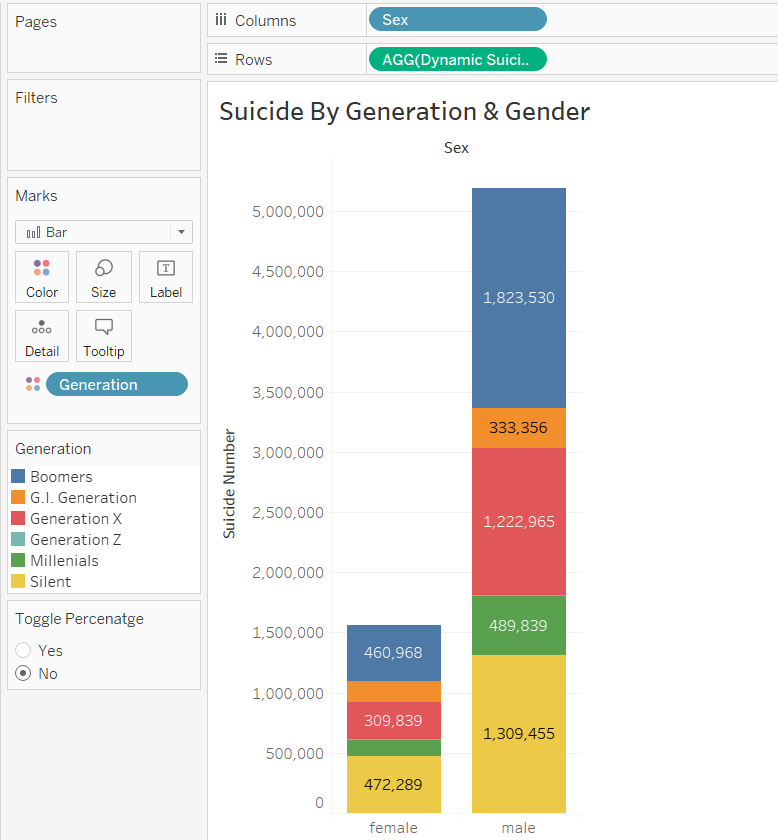
# **Sheet 3,** Suicide by age group:



This sheet visualizes the ‘Suicide No’ by ‘Years’ for each ‘Age’ group using a continuous line chart, the parameter named ‘Toggle Sex’ is also used here in the color marks to enable or disable ‘Sex’ coloring.

* **Rows:** Contains the sum of the ‘Suicide No’ measure.
* **Columns:** Contains the ‘Year’ measure and the ‘Age’ dimension.
* **Marks:** Contains the ‘Dynamic Sex Mark’ in the color mark which can be toggled on or off using the ‘Toggle Sex’ parameter.

# **Sheet 4,** Suicide by generation and gender:



Here we have a stacked bar visualization of the ‘Suicide No’ over ‘Generation’ for each gender, there is also a parameter named ‘Toggle Percentage’ that is linked to the ‘Suicide Percentage’ calculated field that turns the stacked bar into a 100% stacked bar when enabled, and back to normal stacked bar when disabled.

* **Rows:** Contains the ‘Dynamic Suicide’ measure which is linked to the ‘Toggle Percentage’ parameter, used for changing the stacked bar chart type.
* **Columns:** Contains the ‘Sex’ dimension.
* **Marks:** Contains the ‘Generation’ dimension in the color mark, which causes a card to appear for info about the colors.

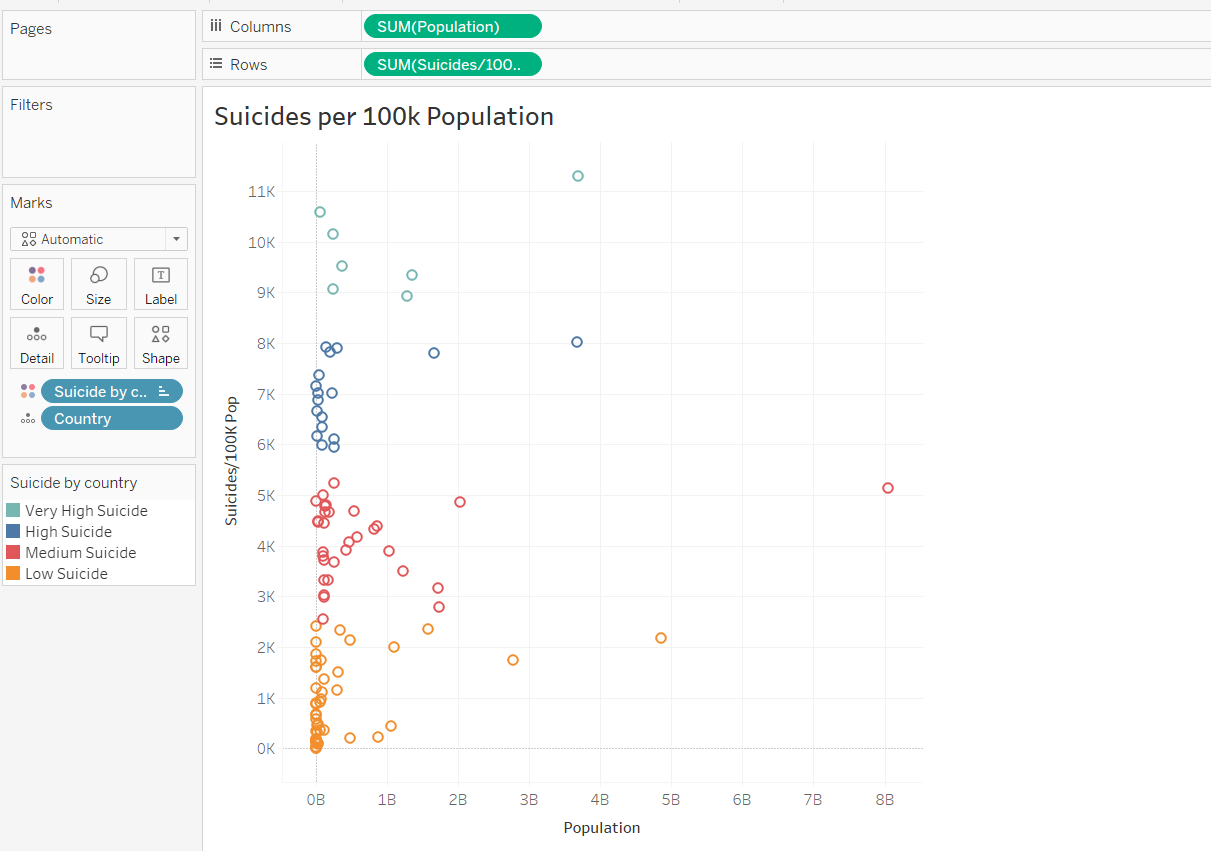
**Dynamic Suicide calculated field code:**

IF [Toggle Percentage] THEN ***[Suicide Percentage]*** ELSE SUM([Suicides No]) END

***[Suicide Percentage]* calculated field code:**

SUM([Suicides No])/SUM({EXCLUDE [Generation]:SUM([Suicides No])})

# **Sheet 5,** Suicides per 100k population:



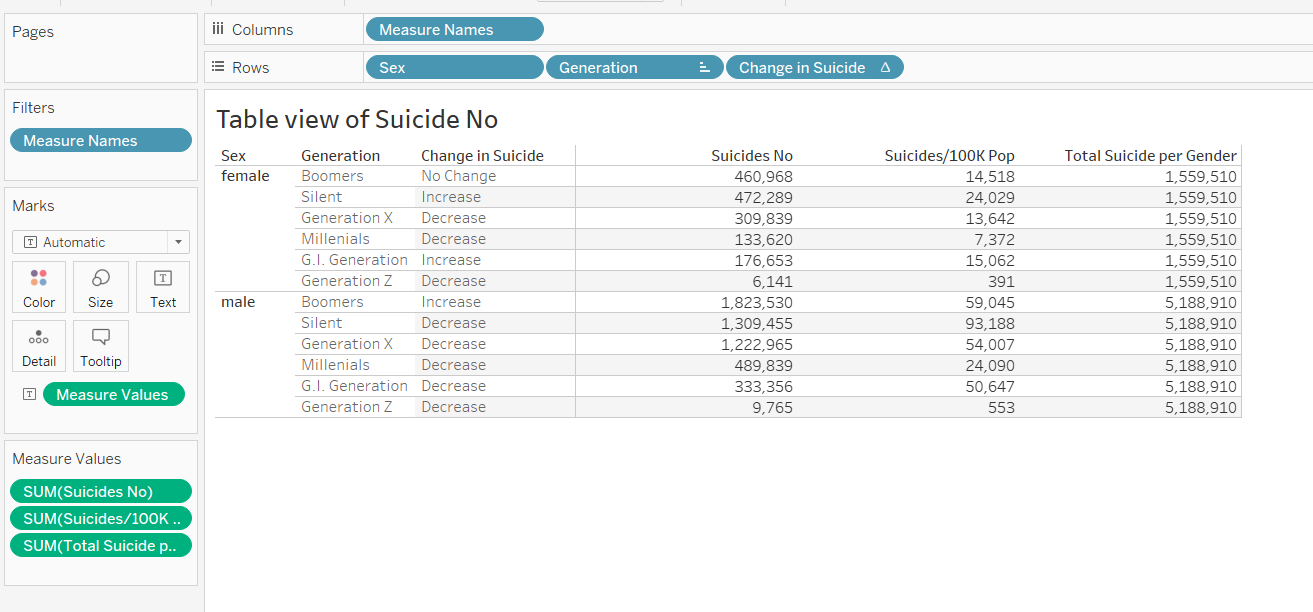
This visualization is a scatter plot of the ‘Suicides/100k pop’ and the ‘Population’ for each country, the scatter plot is clustered to 4 clusters based on the number of suicides ranging from Low to Very High.

Each point in the plot represents a country.

* **Rows:** Contains the sum of the ‘Suicides/100k’ measure
* **Columns:** Contains the sum of the ‘Population’ measure
* **Marks:** Contains the ‘Country’ in the detail mark to show countries on the plot, and contains the ‘Suicide by country’ cluster in the color mark to represent clusters by color.

**Note:** The clustering is done with the ‘Suicides/100k pop’ measure.

**Sheet 6,** Table view of Suicide No:



This sheet is a table view of the Suicide No, Suicides/100k Pop and ‘Total Suicide Per Gender’ calculated field for Sex and Generation, also there is a calculated field for ‘Change in Suicide’ which prints ‘Increase’, ‘Decrease’ or ‘No change’ based on the ‘Suicide No’

* **Rows:** Contains the Sex, Generation, and the ‘Change in Suicide’ calculated field.
* **Columns:** Contains Measure Names, which give the numerical fields in the table a column name.
* **Marks:** Contains the Measure Values, which give the numerical fields in the table values.
* **Measure Values:** Sum of Suicide No, Sum of Suicides/100k Pop, Sum of Total Suicide per Gender.