Date: 9/9/2025



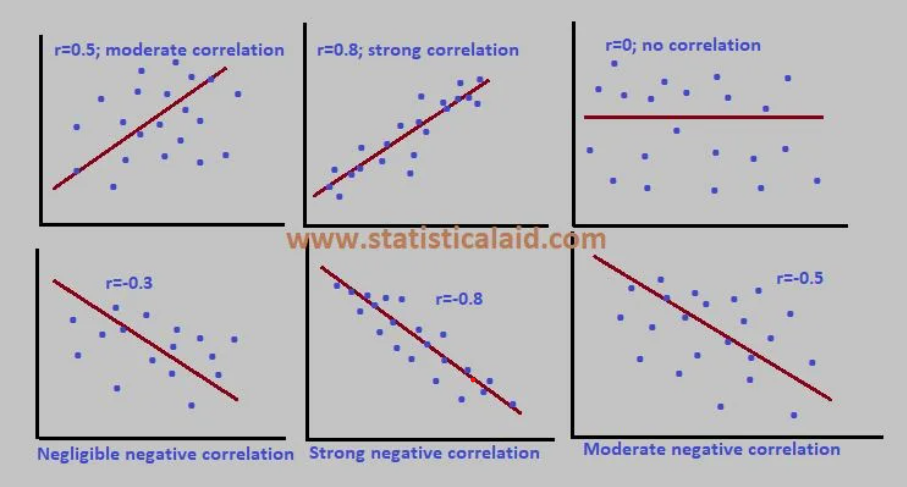
This is the screen shot of World happiness report. Here is the explanation of each column:

* **Country or region**  
  → The name of the country or region being evaluated in the report.
* **Dystopia + residual**  
  → "Dystopia" is a hypothetical country that has the world's lowest values for each happiness factor. The "residual" part is the difference between the actual country’s happiness score and the sum of the measured factors. It captures unexplained influences (things not directly measured).
* **Freedom to make life choices**  
  → Measures how much people feel free to make key life decisions (e.g., career, marriage, residence). It’s based on survey responses about freedom and autonomy.
* **Generosity**  
  → Reflects the tendency of people to give (time, money, or help) to others in need. It comes from survey questions about donations and helping behavior.
* **Healthy life expectancy**  
  → Average number of years a person is expected to live in good health, based on WHO data.
* **Log GDP per capita**  
  → The natural logarithm of Gross Domestic Product per person. It measures the economic strength of a country, adjusted for population size.
* **Overall rank**  
  → The country’s ranking in the World Happiness Report based on its overall happiness score (1 = happiest).
* **Perceptions of corruption**  
  → Survey-based measure of how corrupt people think their government and businesses are. Lower corruption → higher trust → higher happiness.
* **Score**  
  → The overall happiness score of the country, which is the main number used to rank nations. It’s a weighted combination of all the factors above.
* **Social support**  
  → Measures the strength of social networks—whether people have someone to rely on in times of trouble. It’s based on survey responses.

Note:

* Scatter chart = **best when your goal is correlation, pattern detection, or spotting outliers**, not when you just need totals/rankings.
* When there are too many numerical column and very less categorical data (may be 1) we can use scatter chart.

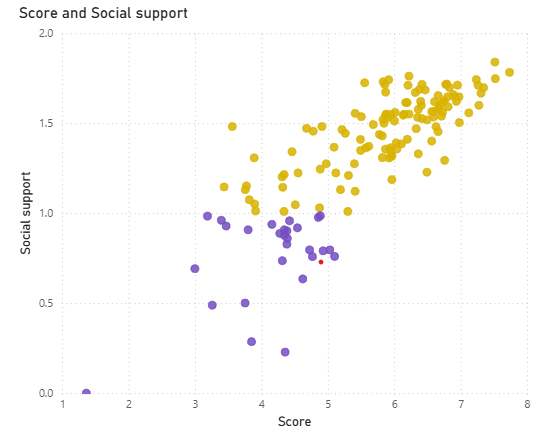
**Correlation:**



* correlation is a measure of the extent to which two or more variables are related.
* The strength and direction of the relationship are measured by the correlation coefficient, a value that ranges from -1 to +1.

**Types of correlation:**

* Positive correlation: This occurs when two variables move in the same direction. As one variable increases, the other also increases. As one decreases, the other decreases.
  + Example: The correlation between temperature and ice cream sales. As the temperature goes up, ice cream sales also tend to go up.
* Negative correlation: This occurs when two variables move in opposite directions. As one variable increases, the other decreases.
  + Example: The correlation between the price of a product and the quantity demanded. As the price of a product increases, the demand for it tends to decrease.
* Zero correlation: This indicates that there is no linear relationship between the variables. A change in one variable has no predictable effect on the other.
  + Example: The relationship between a person's height and their political preference.

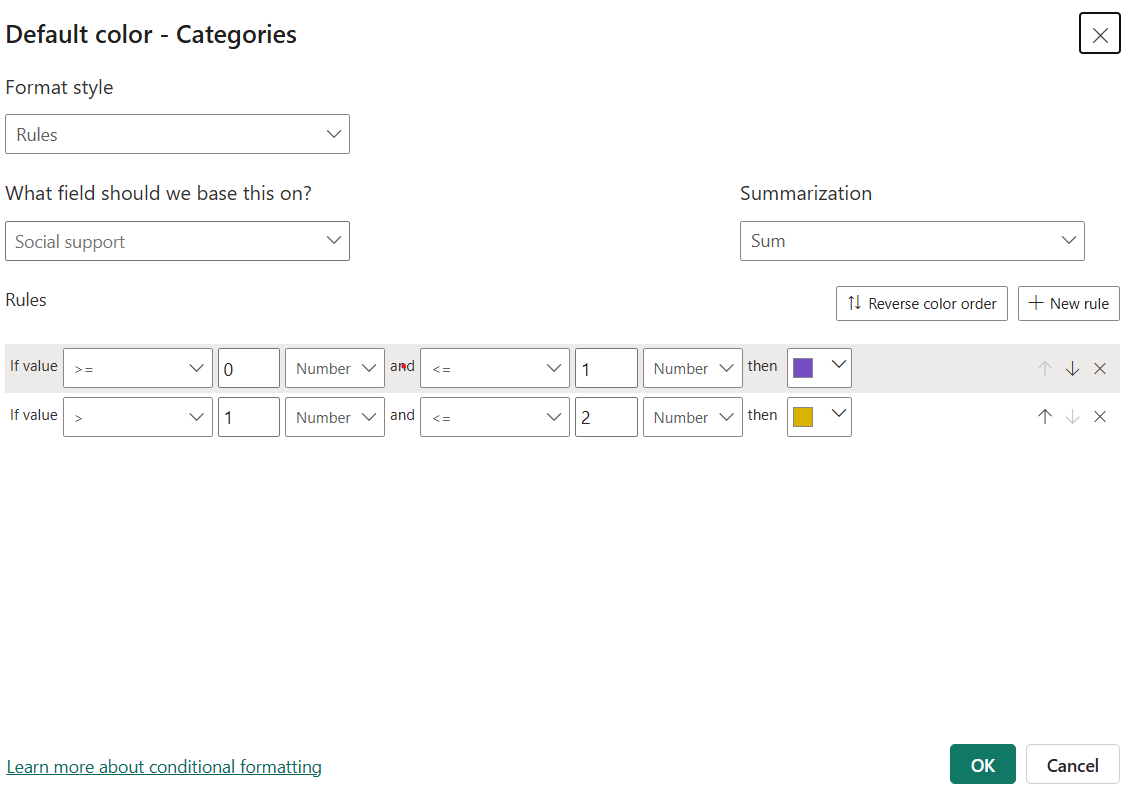


**Insights:**

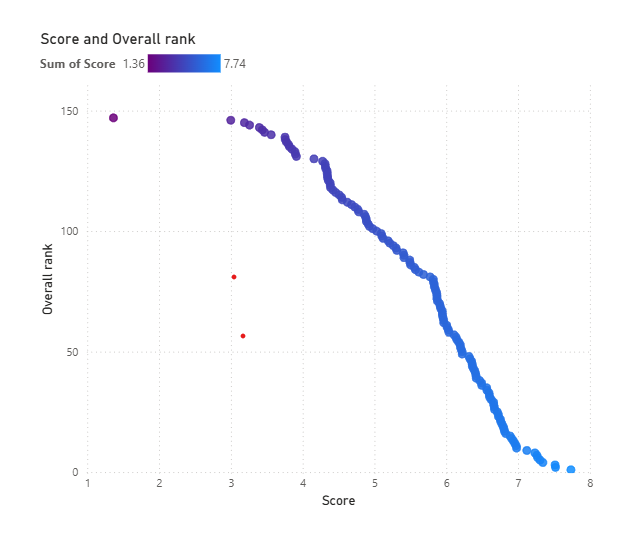
* **Positive Correlation**
* As the **Score (happiness level)** increases, **social support** also tends to increase.
* This suggests that countries where people feel they can rely on others generally report higher happiness.
* **Clusters (color groups)**
* The chart shows **two groups (yellow & purple)** – likely representing regions (e.g., developed vs developing countries).
* **Yellow cluster:** Higher scores (5–8) and higher social support (1–2).
* **Purple cluster:** Lower scores (2–5) and lower social support (0–1).
* This indicates a divide between countries with strong social systems vs weaker ones.
* **Outliers**
  + - A few countries (bottom-left, near Score ~1–3, Social support <0.5) are extreme outliers → very low happiness & weak social networks.
    - Similarly, top-right countries (Score ~7–8, Social support ~1.8–2) are leaders (likely Nordic countries).
* **Middle Zone (Score ~4–5, Support ~0.8–1.2)**
* Countries in this band show **average happiness and moderate support** → possibly developing nations where economic growth is happening but social trust/support is not fully strong.

**Steps to reproduce:**

1. Select scatter chart from build visuals.
2. Pass score to x-axis. By default, it will be taken as sum of score. Then using dropdown select don’t summarize.
3. Pass social support to y-axis 🡪 then don’t summarize.
4. Go to format visual 🡪 marker🡪 go to colour🡪 click on Fx.
5. Here you can do conditional formatting.



1. Make changes as per the requirement and click ok.

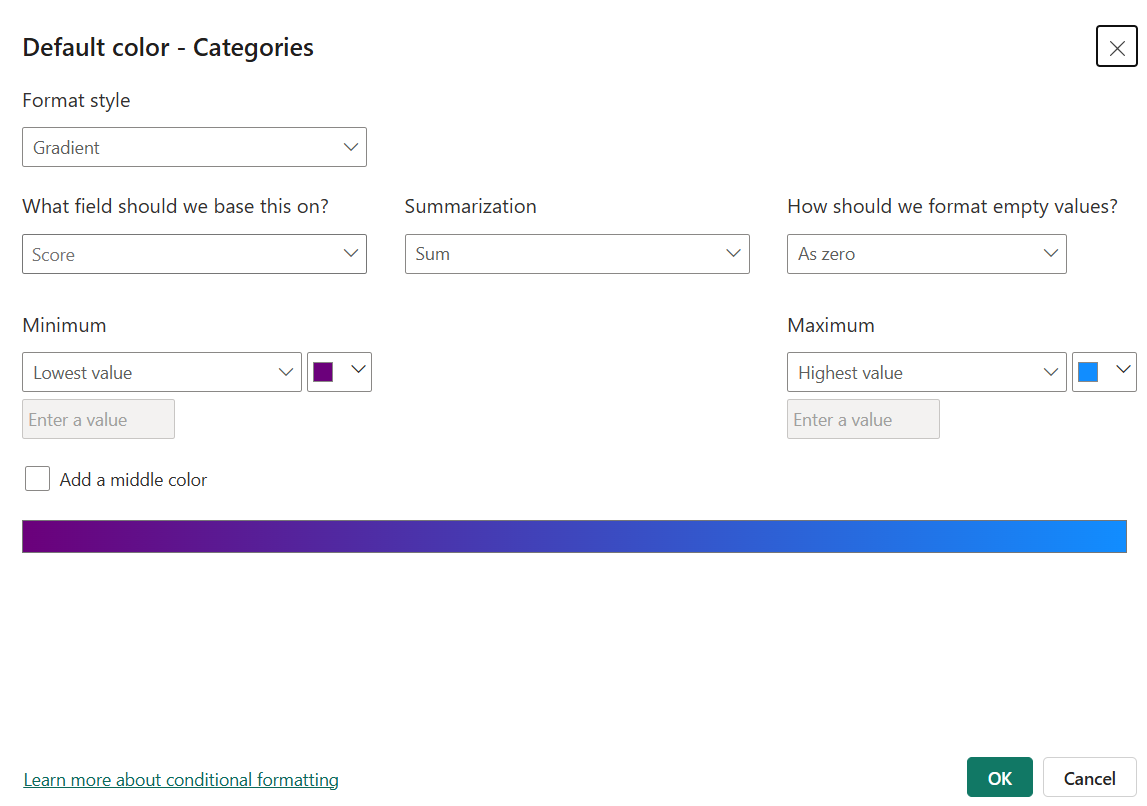


**Insights:**

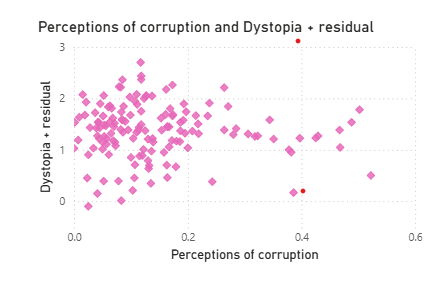
* Strong Negative Correlation
* As Score increases, Overall rank decreases (moves closer to rank
* This is expected: higher happiness scores directly lead to better ranking.
* Top performers
* Countries on the right side (Score ~6.5–8, blue dots, Rank 1–20) are the happiest nations (likely Nordic countries, Switzerland, etc.)
* They cluster tightly → meaning all these top countries consistently have both high scores and high ranks.
* Bottom performers
* Countries on the left side (Score ~1–3, purple dots, Rank 130–150) are the least happy nations.
* They show weaker clustering and more spread → some have slightly better rank despite low scores (suggesting small differences matter a lot at the bottom).
* Middle band (Score 4–6, Rank 40–110)
* This group shows the most spread.
* Even small differences in happiness score here result in noticeable rank jumps.
* Example: A score increase from 4.5 to 5.0 can push a country up by 20–30 ranks.
* Colour gradient adds clarity
* The shift from purple (low scores, poor ranks) to blue (high scores, top ranks) reinforces the strong link between score and rank.

**Steps to reproduce:**

1. Select scatter chart from build visuals.
2. Pass score to x-axis. By default, it will be taken as sum of score. Then using dropdown select don’t summarize.
3. Pass Overall rank to y-axis 🡪 then don’t summarize.
4. Go to format visual 🡪 marker🡪 go to colour🡪 click on Fx.
5. Here you can do conditional formatting.



1. Make changes as per the requirement and click ok.



**Insights:**

* Weak negative corelation
* Most countries cluster around **low corruption perception (0.1–0.2)** and **mid residual happiness (1–2.5)**.
* **Few outliers** exist with high corruption (0.4+) and low happiness residual.
* The overall **correlation is weak negative**, suggesting corruption has *some negative effect on residual happiness*, but it is not a dominant factor.

Steps to reproduce:

1. Select scatter chart from build visuals.
2. Pass perception of corruption to x- axis and dystopia + residual to y-axis.
3. Go to format visual🡪 go to marker🡪 select the shape and colour
4. Turn on the boarder button.