Heat map

Task:

Heat map describes the correlations among few variables that impact suicide rates. Correlation is a statistical metric for showing the extent of relationship between contained index. Normally, heat map use python to implement the correlation matrix, utilize gradient color on the right side to implement the number of correlation.

For tasking, we need to identify the categorical data as country, year, sex(male/female), as well as numeric data as suicides /100k, age/generation, sex, population, gdp\_for\_year($) & gdp\_per\_capita($), suicide\_no, Continent Name. In this case, we choose location data set(country) to find the main Influencing factor that determine suicide rate in different countries, additionally provide two heat maps with male and female to distinguish the most effected factors leading to suicide of two genders.

Analysis:

The reason why we use the heat map to visualize suicide rate is that heat map generally means the point data analysis by calculating Kernel Density Estimation.

Kernel Density Analysis (KDE) visualizes features to achieve the translate from discrete object models to continuous-field models by calculating the density around the features, create a smooth surface for feature pattern detection and discovery in the end. Based on KDE, dot data analysis can be used to describe any type of incident data, because each event can be abstracted into a spatial location point. We can use point data to analyze the discipline behind the data, which is called ‘point mode’. Point mode is ubiquitous in nature and economic society. Through analysis, we can make point data into point information, which can better understand the spatial point process and accurately find the discipline behind the space point.

Visualization:

Heat map use projection coordinates to analysis, so its a square graph contains (number of index)\*\*2 little squares, with density above. Gradient color chart on the right side defines correlations in diversity extents, for example, deep color represent strong connection while light color express there is raw relationship among two factors.

Implement:

Python has multi-functional libraries and functions to complete the drawing of heat map in a few line. The instances are showed below:

1. Use python-numpy-np.array to input data set, and also can use panda-dataframe function.
2. Use color parameter like vmax, vmin to ensure maximize and minimize range of value; matplotlib-colormap.
3. Apply ax.set\_title, ax.set\_xlabel, ax.set\_ylabel to build coordinate.

With python we can easily import Express to provide simply grammar for complex visualization graph, most drawings merely require to call one function. Additionally, changing currently parameter can be valuable to create extra heat map to analyze suicide rate multilayered.