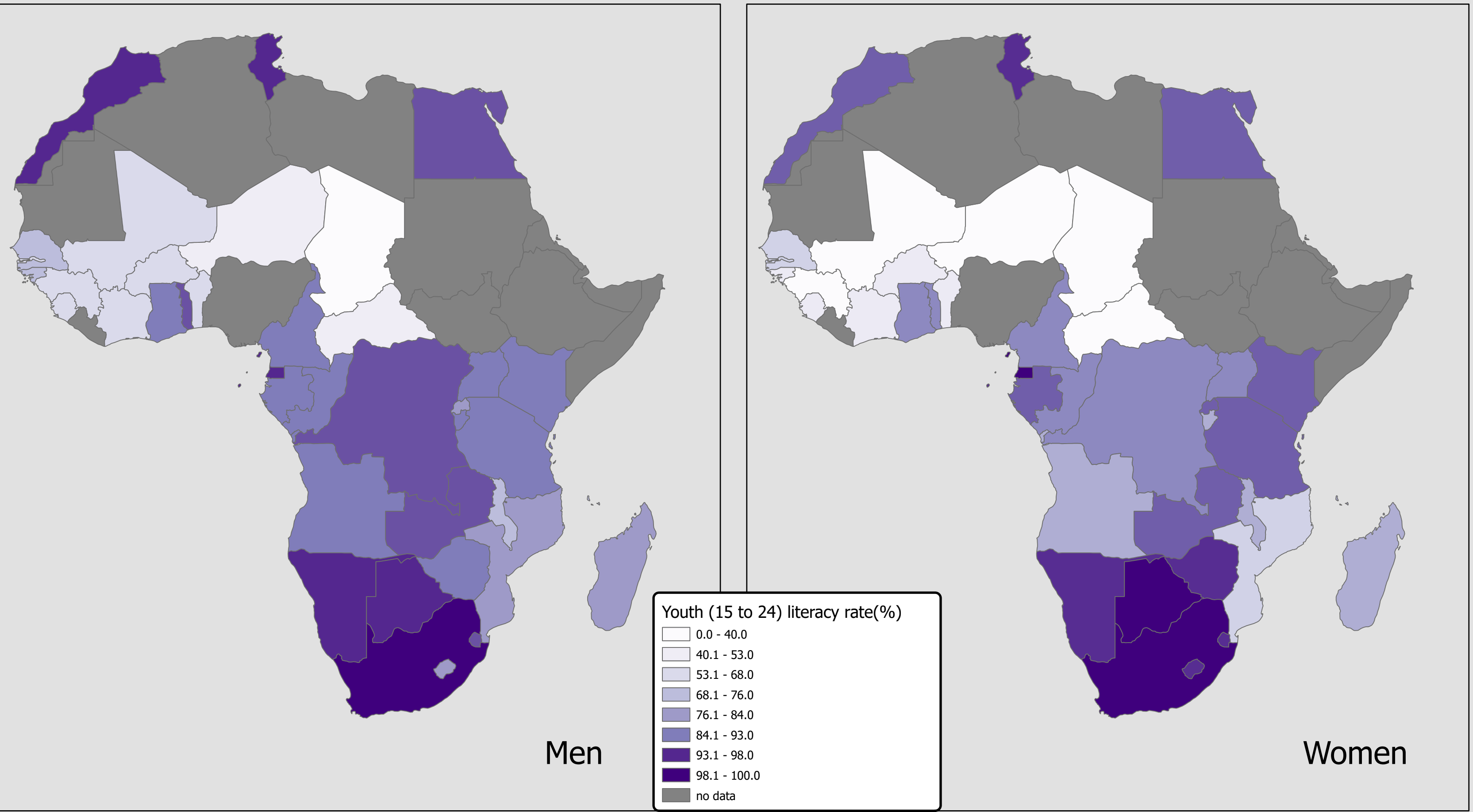


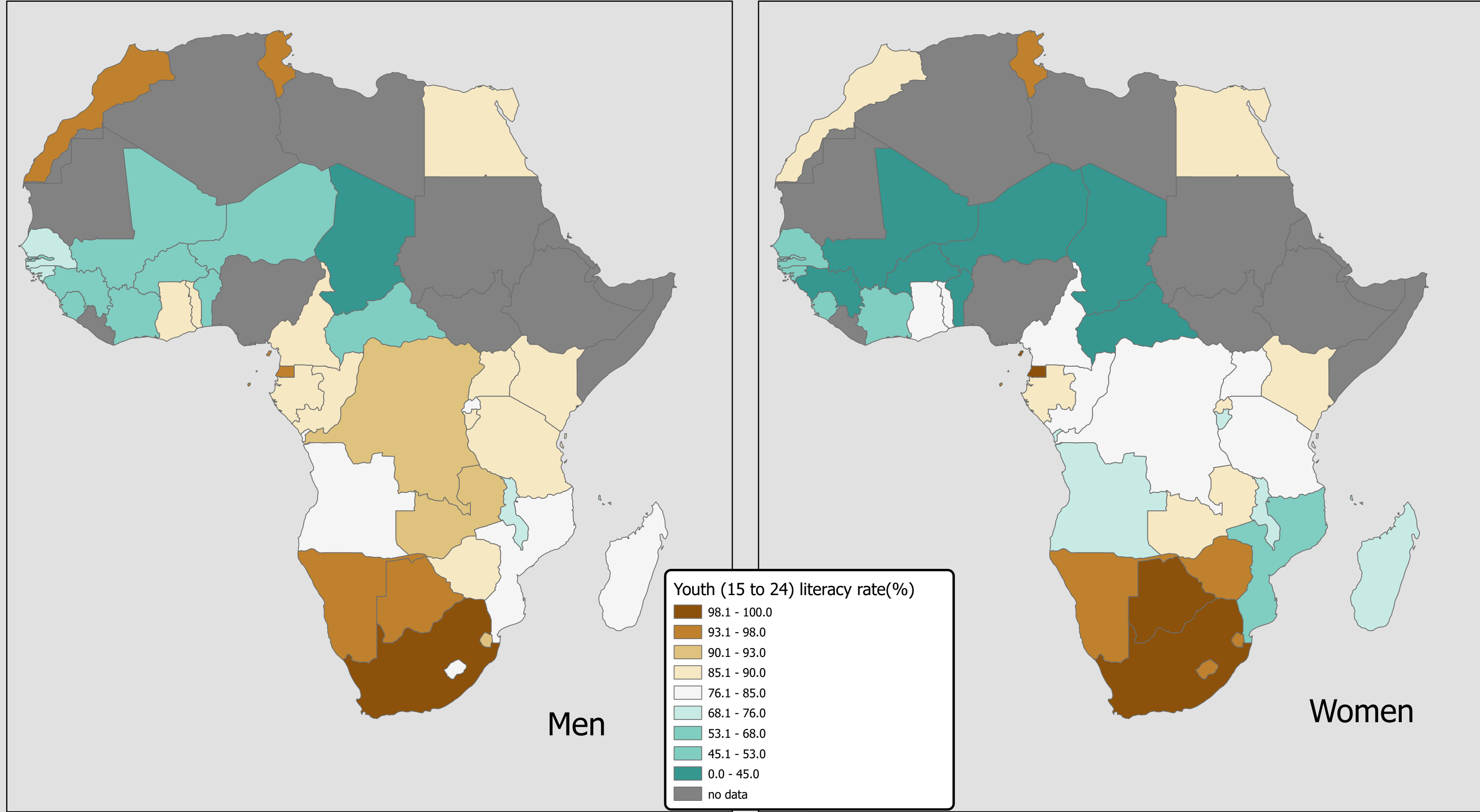
Africa's literate rate in different countries for youth (15-24) between men and women

— 2010-2018 report involves different methods to interpret the literate rate for men and women



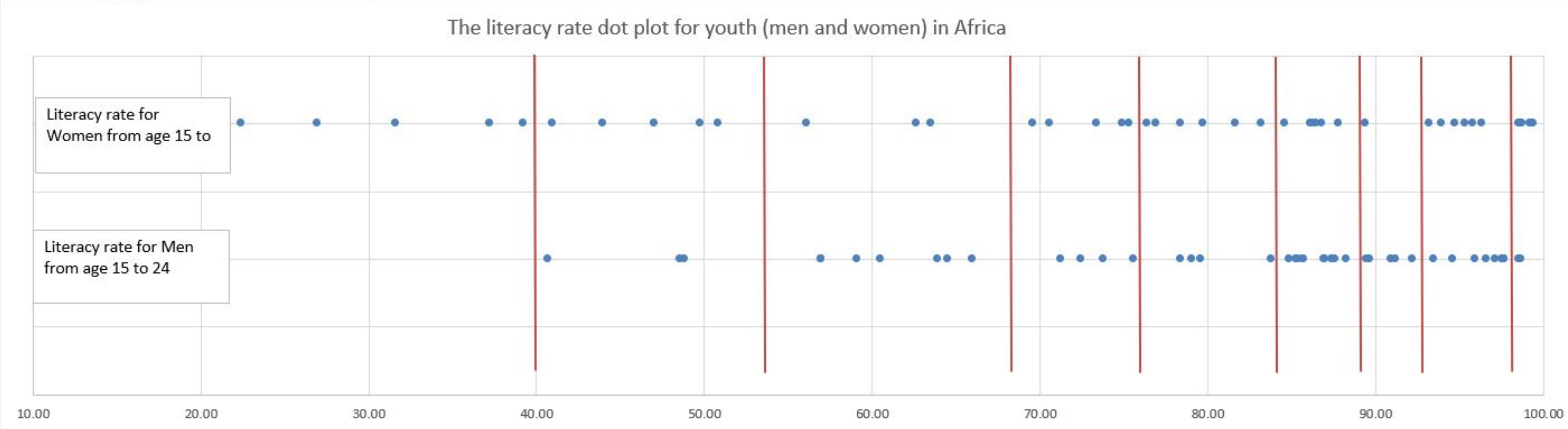
Set A: Discussion

For these two Manual Break Sequence Maps, I set the class breaks at 40, 53, 68, 76, 84, 89, 93, 98, and 100. According to the Plot, I intentionally set each class to contain 4-5 variables. The distribution of the plot is clustered at the high percent range which is around 95-100%. I set three breaks in this interval to better shows the variation on the map. If I use the equal interval in this case. The map will be in primarily dark colors which will show little variation. Furthermore, The breaks have equal spacing between the last values in the first interval and the first value for the second interval. In this way, the Audience will have a better conceptual idea of the distribution of this dataset. The weakness of this method is the audience will be misled by these maps. They will believe there are so many regions has low-literate rates without seeing the legend.



Set B: Discussion

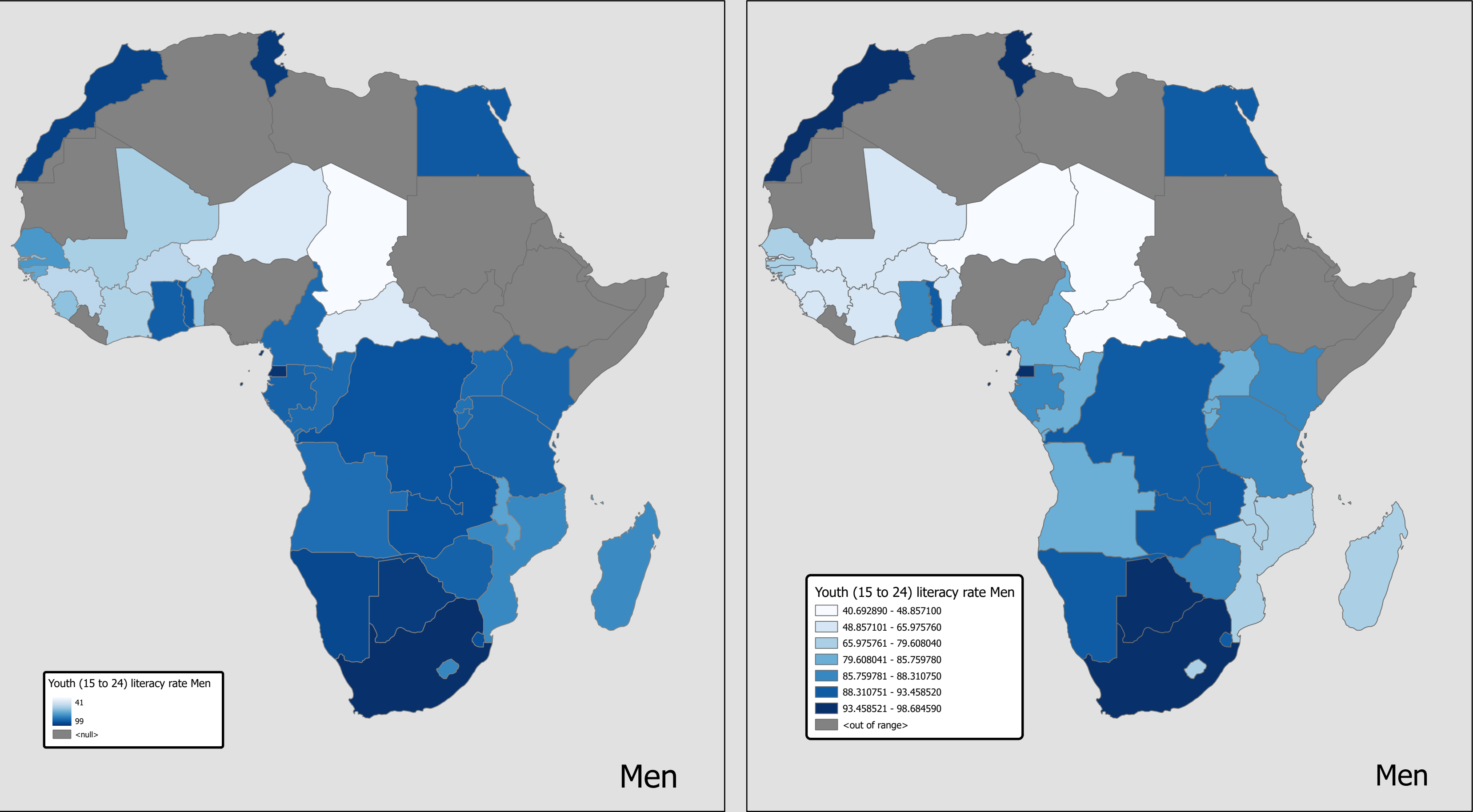
I set the critical break in this map to 85. The reason why I choose it is the distribution pattern of the literate rate data. The mean value of this dataset is 72.17. For a normal distribution dataset value, 72.17 would be a good choice. However, the value for the literate rate tends to cluster in 85-90 intervals. In this case, choosing the 50th percentile would be a better choice for diverging colors to show the variation of the data. 85 is the number close to the 50th percentile in both of the datasets which is the reason I choose it. The weakness of this method is the audience does not know which value represents the high value. The strength is each color is distributed almost equal amount of the area. The audience could easily see the difference between men and women by examining the different color



The class breaks are set at 40, 53, 68, 76, 84, 89, 93, 98, and 100. The max value for men is 98.96, the min value is 40.69, the mean is 79.30 and the standard deviation is 15.12. The max value for women is 99.40, the min value is 22.39, the mean is 72.17 and the standard deviation is 22.04. There are 13 countries has no data.

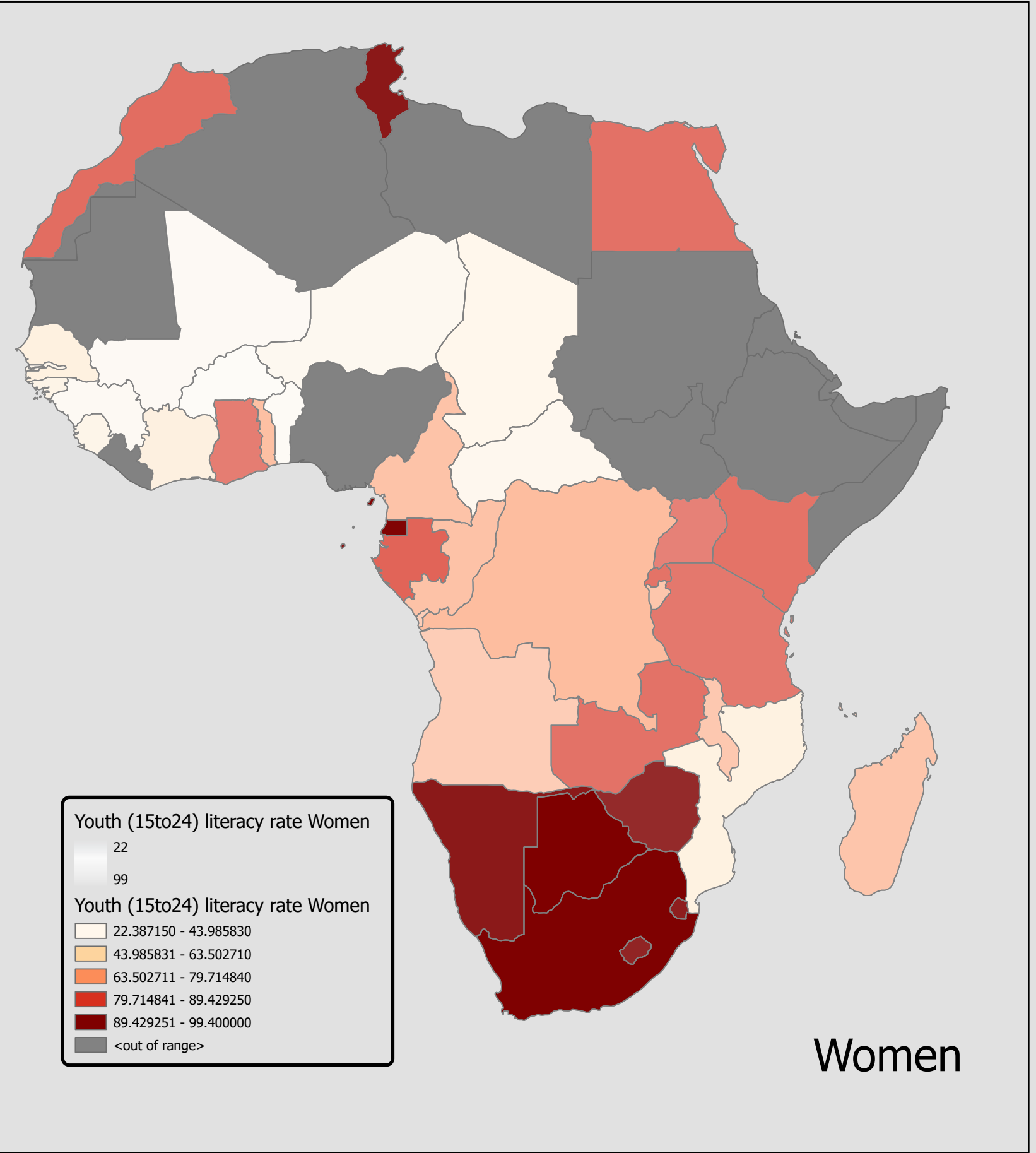
Set C: Discussion

In this map set, I choose Natural breaks for the classification method. Compare to the unclassified map, I aim to group similar values together and maximize the difference between classes. It set the breaks in the place where display large difference. According to the plot, the data is not normally distributed. Most of the values are distributed between 70- 100. Those areas are high literate rate areas. The Natural break assigns the unique break numbers for each of these classes to display the difference between high literate rates. The audience can not see the difference in the unclassified map, especially at the bottom part of Africa. Furthermore, the audience could not interpret the map for the unclassified map in the legend. Natural breaks do a better job for the legend. The weakness of this method will be audience will be misled by this classification. They will think the regions with high literate rates are only a few of them without checking the legend.



Set D: Discussion

The opacity mask is used in this section to de-emphasized the low and mid values in the map. According to the plot, the global value is near the end of the data range. Therefore two breaks are sufficient in this case. In the setting, I set 15% Transparency for white from 0 to 40 % of the data position. The color of the low literate rate is mixed with the white color. Then I set 70% transparency for white from 40% to 85% of the data position. Again, I would like to emphasize the high literate rate between 85-100%. I set 100% transparency for white from 85-100%. The Audience would process pre-attentively for the value between 85-100 positions in the dataset. The weakness of this method is the audience can not identify the difference between low-literate rate regions especially for the literate rate is between 0-63



Set E: Discussion

The bivariate color technique is really effective in interpreting the two variables in the data. First of all, I can easily know where the high literate rate and low literate rate areas for both men and women. The high-literate rate areas are at the top and the bottom of Africa. The low-literate areas are in the middle part of Africa. In addition, I could see the difference between those two variables. For example, which area has a low literate rate for men but a high-literate rate for women? By examining the map, the audience could easily answer those questions. The weakness of this method is the audience can not check the individual variable and there is no value display in the legend.

