Dataset 3

The first step was to load the dataset and explore it in order to familiarize ourselves with the dataset. The attributes of the dataset which were used to answer the questions were:

- CNAME
- TOTCHUR
- Religions= (234 - 8)/2= 113
- ARAPO_M, GRKAD_M, ACROC_M, BEOC_M for the Orthodox Members question

In order to summarize the data, we used the function *describe*, which shows us the needed information.

As for the question “Which are the countries with the highest per person ratio of Orthodox Christian members?”, we found out that there were four Orthodox churches which create the Orthodox dogma. Those attributes were ARAPO, ACROM, GRKAD and BEOC. In order to answer the question, we summarized the members of those religions and we divided the result with the TOTMEMB (Total number of members). To get the ratio, we multiplied that number with 100 and sorted the dataset with descending order.

For the question “Can you find the 3 most extreme (outlier) counties with respect to the distribution of their churches across religions?”, we used two methods to find the extreme values. In the first method, we found the number of churches for every county and sorted the dataset with descending order. Moreover, as a second method, we visualized the result using boxplot to show the extreme values.

For the question “Where would you create a cross-religion center of discussion between religions to maximize its impact? Support the proposal based on data analysis results.”, we sorted the dataset according to the Total number of churches and we concluded that the best place to build a cross-religion center of discussion was the province with the highest number of churches.