

Exercises.

1. From the movies data frame.
 - a. Count the missing values in the 'content_rating' column.
 - b. Return the rows containing any missing values, how many are they?
 - c. Return the unique values in the 'content_rating' column
 - d. Return the movies that are not rated and what is there number.
 - e. Return the movies with a higher rating than 9. How many are they?
2. Which country has the highest sum of beer_servings and spirit_servings.
3. Which continent has more beer servings, beer_servings spirit_servings wine_servings.
4. From every continent find the country with the highest beer_servings, spirit_servings, wine_servings.
5. In West Africa which country has the lowest total liters of alcohol?
6. In East Africa which country has the highest intake of alcohol?
7. From the titanic data-set what is the average age of the survivors, what was the most common Pclass and Sex.
8. Which sex recorded the most survivors according to the titanic data-set?
9. In North Africa which country drinks most?
10. From the SADC block which country is the third most consumer of Alcohol?
11. From the UFO data-set what is the overall memory usage? Which column has the highest share of memory? How can u reduce the memory usage of some of the columns for this data set? How much have you reduced the memory usage for the data-frame? Explain your findings with code.
12. Which country from the G7 has the lowest total intake of Alcohol?
13. Which movies have the highest percentage via content_rating and genre.
14. Split your movies data-set into 2 parts called train with 70 percent of the data and test with 30 percent of the data. How many records does each of the splits have?
15. From the dataset “daily-total-female-births.csv” which month , day of the week, day of the month and number of day of the year has the highest number of births?