**CPC Assessment – 2024**

**Section A**

Solve the following questions in R and share the script

Load the following datasets from the openintro package: census, sowc\_child\_mortality, sowc\_demographics, sowc\_maternal\_newborn

1. How would you remove NA values from personal income variable in census data
2. Is it possible to see only the 25th to 35th observations of the census data? How?
3. How would you find the mean total income for every person in the census dataset?
4. Join sowc\_demographics and sowc\_child\_mortality in such a way that you have three variables: countries, life expectancy of 2018 and under5 mortality in 2018. Lets call this new dataset joined\_data
5. To the previous question, now do a left join with sowc\_maternal\_newborn dataset
6. From sowc\_demographics, create a variable called mean\_life\_exp that calculates the mean life expectancy of 1970, 2000 and 2018
7. Make a bar chart of top 12 countries having the highest mean\_life\_exp with appropriate formatting
8. How would you find out which countries that are present in sowc\_maternal\_newborn are not present in sowc\_child\_mortality
9. How would you find the highest personal income for each race?
10. In the sowc\_demographics dataset, is there a way by which you can fill the NA values in the under18\_pop\_2018 with a 0

**Section B**

According to you, what are some current important trends in fertility worldwide? What does past literature show? In about 500-750 words, explain your response with appropriate citations.