Public Opinion Polling Practicum: Data/Writing Assignment 2

Directions: This assignment requires you to analyze results from a poll conducted by the Cooperative Election Study in 2020. This link contains the data you will examine for this assignment, and this link provides a list of variables you will need to complete it. For this assignment, you will analyze a subset of the CES data that includes just 11 variables. You will write a codebook for these variables that resembles this one. And you will write topline results for these variables that resemble this one. Your work is due via email no later than 11:50:59 pm on Friday, April 15, 2022. Don't forget to attach your R code (not the contents of your R console) to the end of your work.

1. Using the supplied data, create a codebook like the one linked to above that includes each variable name, its name in the dataset, and how it is coded.
    a. Age: This variable should be coded "1" for 18-24 year olds, "2" for 25-34, "3" for 35-44, "4" for 45-54, "5" for 55-64, "6" for 65-74, and "7" for 75 and older.
    b. Gender: A dichotomous variable for the respondent's gender.
    c. Education: An ordinal variable for the respondent's level of education.
    d. Race: A nominal variable for the respondent's race.
    e. Region: A four-point, nominal scale indicating whether the respondent is from a western, northeastern, midwestern, or southern state.
    f. Ideology: An ordinal variable measuring the respondent's ideology.
    g. Party Id: A categorical variable indicating a respondent's partisanship.
    h. Sexuality: A categorical variable indicating a respondent's sexuality.
    i. Born Again Evangelical: A dichotomous variable indicating whether the respondent is a "born again evangelical."
    j. Income: An ordinal variable indicating respondent's annual family income.
    k. Presidential preference: The variable, "cc20_364b" in the CES dataset.
2. Using the supplied data, create a set of topline results similar to the linked example above.
    a. First, provide tables of weighted percentages, using the weighting variable, "commonweight," of the frequency of respondents who fall into each category of each variable.
    b. Then, create weighted crosstabs that include the candidate for whom respondents intend to vote in the columns, and a separate crosstab for each of the other ten variables. Thus, each cell is the weighted percentage of respondents within a given category of a variable (say, men, in the "gender" variable) who intend to vote for a given presidential candidate in 2020.
    c. You don't need to interpret the results quite yet. We'll get to that next time. Just focus on getting the