Multiple Regression Exercises

A clinical psychologist is interested in the effects of a treatment plan on the general wellbeing of clients. She is also interested in whether the level of severity of symptoms when entering treatment or the clients' trust in the psychologist predict wellbeing.

The severity_data.csv dataset contains 5 variables:

- client: an anonymous ID of the client
- treatment_group (control or therapy): Which treatment the client has been assigned to
- level_of_severity (1-20): A rating assigned by clinicians as to the severity of the client's difficulties when they entered treatment (a computed value based on a battery of tests)
- trust_score (0-100): A rating by the client as to the level of trust they have in their psychologist based on their interaction to this point
- wellbeing_after_3_months (0-10): Client's score on a psychometric measure of general wellbeing

To answer the research questions outlined above, conduct the following analyses:

1. Calculate if there is a difference in the means of *severity level*, *trust score* and *wellbeing after 3 months* between the 2 treatment groups (Hint: use the describeBy function in the Psych package)
2. Make a histogram for each of the variables in the previous question and assess the distribution (Hint: Use the ggplot2 package to make plots)
3. Run a heirarchical multiple regression analysis comparing the following models (in each case, *wellbeing after 3 months* is the outcome). At this stage, check main effects only, not interactions:
   - Model 0: Constant
   - Model 1: treatment_group, severity_level
   - Model 2: treatment_group, severity_level, trust_score
4. Check the assumption of multicoliniearity in Model 2 (Hint: Install and load the *mctest* package)
5. Use the plot() function to check other assumptions of Model 2
6. Run Model 2 again, including interactions this time.
7. What conclusions can you draw from these analyses in relation to your research question?
8. Run Model 2 as a stepwise analysis (Hint: Install and load the *MASS* package first)
9. Are the results of the stepwise analysis consistent with your previous conclusions?