Q1 :Import the CreditCard.csv file.

In this dataset, you will find information related to credit card use for a sample of 300 credit card holders. Answer the following questions:

[6 Points] Create a machine learning model that generates a prediction for the outstanding balance of a particular card holder. Show all of your steps and provide evidence regarding the performance of your model.

[4 Points] Devise some rule that will allow you to choose who should or should not be approved for a new credit card based on what you have observed in your analysis. Be sure that your rule includes the individuals' balance.

Import the CreditCardApplications.csv file.

In this dataset, you will find individual characteristcs but you will not find data on their outstanding balance. This dataset represents information for 50 people who are applying for a credit from a company with which they have not had a credit card before.

[5 Points] Deploy your model to generate predicted card balance. Apply your rule that you gave in question 3.2 and label who would be approved for a new credit card and who wouldnt. Report a dataset with these labels. What percentage of the applicants will you approve?

[10 Points] Reflect back on your analysis so far and the labels that you have generated. You have done some impressive work creating a tool that would be valuable for a credit card company. The model that you have developed would likely help them to determine the best candidates for new credit cards. However, the models that we develop using machine learning techniques can have pitfalls if not avoided. People often think that data analysis, and strickly quantitative procedures or algorithsms used during data analysis, are impartial or neutral. Perhaps many will think that decisions made based on data analysis can be considered as objective and "based on the data" or "based by the science"; that we do not need humans exercising virtue when implementing these techniques or the decisions that are based on them. But this is not necessarily the case.

Look back at your model and determine if racial discrimination could have take place in your analysis. Is it possible? Did it occur in your model? How would you know? Explain your answer, showing any evidence you need including tables, dataframes, or graphs. How can someone avoid it?

If you are concerned that your model may have set you up for a decision rule that implicitly excluded particular groups on the basis of race, go a head and go back to revise your model accordingly. Note here what you did, if anything, to modify your model in this step.

Do you think that it could be important for humans to play a supervisory role over the machine learning and artificial intelligence algorithms that are used to make important decisions in our increasingly technolical world? If so, what role should they play? Do you think that humans exercising virtue (as opposed to vice, malice, or prejiduce) are important for data analytics? Why or why not?