**Question 1**

(a) This is a regression problem because the outcome variable (CEO salary) is continuous. We are interested in inference because we want to understand how the predictors (profit, number of employees, and industry) affect the outcome (CEO salary).

n (number of observations) = 500 (each firm is an observation)

p (number of features) = 4 (profit, number of employees, industry, and CEO salary). Note that CEO salary is our target variable.

(b) This is a classification problem because the outcome (whether the product is a success or failure) is binary. We are interested in prediction because we want to predict the outcome for a new product.

n (number of observations) = 20 (each product is an observation)

p (number of features) = 14 (success or failure, price charged for the product, marketing budget, competition price, and ten other variables). Note that "success or failure" is our target variable.

(c) This is a regression problem because the outcome (% change in the USD/Euro exchange rate) is a continuous variable. We are interested in prediction because we want to predict the % change in the USD/Euro exchange rate based on the % changes in different stock markets.

n (number of observations) = 52 (assuming there are 52 weeks in 2012 and we have recorded data for each week)

p (number of features) = 4 (% change in the USD/Euro, the % change in the US market, the % change in the British market, and the % change in the German market). Note that the % change in the USD/Euro is our target variable.

**Question 2**



**Question 3**  
Calculate the entropy for the 'Repeat Customer' attribute

From our dataset, we have:

'YES' : 6

'NO' : 4

So the entropy is:

*Entropy (RepeatCustomer)=−(6/10 log2 (6/10 )+4/10 log2 (4/10 ))*

*= 0.971*

Next, we calculate the entropy for 'Repeat Customer' for each unique value of each attribute, then calculate the weighted sum of these entropies and subtract it from the original entropy of 'Repeat Customer' to get the information gain.

For the 'Age' attribute, we have the following distribution:

'20..30' : 4 'YES', 2 'NO'

'31..40' : 1 'YES', 1 'NO'

'41..50' : 1 'YES', 0 'NO'

'51..60' : 0 'YES', 1 'NO'

The information gain for the 'Age' attribute is:

*IG(RepeatCustomer,Age)=Entropy(RepeatCustomer)*

*− (6/10 Entropy(Age=′20..30′)*

*+ 2/10 Entropy(Age=′31..40′)*

*+ 1/10 Entropy(Age=′41..50′)*

*+ 1/10 Entropy(Age=′51..60′))*

= 0.276

Similarly:

'City' = 0.020

'Gender' = 0.061

'Education' = 0.609

Education has highest gain, so it will be the root of the decision tree

**Question 4**

a) The correct answer is III. For a fixed value of IQ and GPA, males earn more on average than females provided that the GPA is high enough.

The coefficient for the Gender variable (X3) is positive (β3 = 35), which suggests that being female is associated with a higher starting salary, all else being equal. However, the interaction term between GPA and Gender (X5) is negative (β5 = -10). This means that the increase in salary associated with each additional point of GPA is less for females than it is for males. If GPA is high enough, this negative interaction effect can outweigh the base positive effect of being female, meaning that males with the same IQ and high GPA would have a higher salary.

b) To predict the salary of a female with IQ of 110 and a GPA of 4.0, we can substitute these values into the regression equation:

Salary = β0 + β1X1 + β2X2 + β3X3 + β4X4 + β5\*X5

where: X1 = GPA = 4.0 X2 = IQ = 110 X3 = Gender = 1 (for female) X4 = Interaction between GPA and IQ = GPA \* IQ = 4.0 \* 110 = 440 X5 = Interaction between GPA and Gender = GPA \* Gender = 4.0 \* 1 = 4.0

So:

Salary = 50 + 204.0 + 0.07110 + 351 + 0.01440 - 10\*4.0 Salary = 50 + 80 + 7.7 + 35 + 4.4 - 40 Salary = 137.1 (in thousands of dollars)

c) False. The coefficient for the interaction term indeed is small, but this doesn't necessarily mean there is little evidence of an interaction effect. The magnitude of the coefficient tells us about the strength of the interaction, but it doesn't directly indicate the statistical significance of this interaction. To make a claim about whether there's evidence of an interaction, we would need to conduct a hypothesis test or look at a confidence interval for this coefficient. If the coefficient is statistically significantly different from zero, that's evidence of an interaction effect. Without this information, we can't definitively say there's little evidence of an interaction effect.