1) A)

i) The parameter of interest is the employed upper year undergraduate students.

ii) The population of interest is 121. sum((body$employed!="no")*1)

B) large sample (approximate Z) test for a population proportion

2)

A) large sample (approximate Z) test for a population proportion

B) Conditions for using a large sample (approximate Z) test for a population proportion are:

It has to be SRS: Not mentioned so we assume it's not met. It's a voluntary response sample. Count (X) is described by a binomial model: We know that x is described by a binomial model since we have either employed or not employed so only 2 outcomes. The number of observations n is fixed and equals 558. Whether one is employed or not, it doesn't affect the other person's response so yes, each observation is independent. The probability of success p is the same for each outcome, and as we see in our data, whether someone is employed or not, the probability of success isn't changing since there is no relation between the two outcomes. Np_0 => 10 and nq_0 => 10: the numbers of success and failures are both greater than 10. Since sum((body$employed!="no")*1) gives us the number of successes and sum((body$employed=="no")*1) gives us the number of fails.

3) A) large sample (approximate Z) test for a population proportion

B) > prop.test(121, 558, 0.29, "greater", 0.95)
So, the value 121 is the number of students who are employed while studying and we got that number using the r code: sum((body$employed!="no")*1). 558 is the number of individuals in the sample, we got that number from the r code View(body) which shows how many people there are. 0.29 is the p value we got from what the researcher got in the study. We used greater since we have more than 2 things to look at. And finally, we are told to use .95 confidence interval.

C) The output gives us the number of "success" which comes out to be 121 out of 558 samples. The p value of the researcher is 0.29 and the 0.216 was the p value based on the data that we provided.

4) We need to use 184 people for the sample size based on the Margin of error formula. R code used was > zVal<-qnorm(.015)

> zVal

(0.015 comes from the 3 percent left from the 97 percent confidence interval given. The 3 is divided by 2 since we have 2 sides.)

which calculates the z score to be -2.17009 (We need the Z score in the margin of error formula, which we rearranged to get the population n.)