Hypothesis Testing

What are Type I and Type II error?

1) State what Type I and Type II error are in 1-2 sentences.
2) Elaborate in your own words.
3) Give concrete example(s) of Type I and Type II error.
4) Illustrate Type I and Type II with a picture, diagram, metaphor, or analogy.

What is a p-value?

1) State what a p-value is in 1-2 sentences.
2) Elaborate in your own words.
3) Give concrete example(s) of p-value.
4) Illustrate p-value with a picture, diagram, metaphor, or analogy.

Pair Roleplaying

Employee Role:
You work at a social media site. You ran an experiment to change the color of the signup button from blue to green. Your null hypothesis was that the green button would not increase the number of signups. You got a p-value of 0.051.

Manager Role:
You work at a social media site. You are a manager & not a data scientist.
You asked your data scientist to run an experiment to see if you should change the color of the signup button.
You see that the experiment got a p-value of 0.051, which is pretty darned close to 0.05, so you'd like to implement the change.

Pair Roleplaying

How did you agree to proceed? Implement or not?

Why did you come to that conclusion?

What was wrong with this situation?

Was there anything else that the employee could have said to make the situation more clear?

How do you feel about the following situation?

You ran an experiment to see if you get more users to join the site by using a texting campaign. You got a p-value of 0.049. It will cost $0.05 to send each text.

Do you want to implement the change?
Do you have any hesitations?

Hypothesis Testing

You live in San Francisco and your friend lives in New York. You believe that the rental price of one bedrooms is more expensive in SF than NY, but your friend is dubious. How would you design an experiment to determine if you’re correct?

You live in San Francisco and your friend lives in London. You're both playing Pokemon Go. You've found quite a few Zubats, but your friend hasn't found any. On the other hand, your friend has had better luck finding Snorlax. You want to determine if the pokemon populations are different in London and San Francisco, or if it's just been chance that you've found different Pokemon. How would you design an experiment?

A designer has made three versions of the checkout page. He wants you to test which one is better.

How would you set up this test?
How many tests do you need?
Do you see any issues?
Would you have any recommendations of what to do?