1. In a hypothetical study of crime in Western Europe, Interpol looked at the incidence of crime by country in 2000 and 2010 as well as the incidence of crime over time in Western Europe as a whole.

(a) (3 pts) The researchers found that in Germany there were 15,043 crimes committed in 2010, while in Belgium there were 6,031 crimes committed in the same period. They concluded that Belgians are more law-abiding than Germans. **Do the numbers reported here support the conclusion, or is something missing? Explain briefly.**

The numbers do not support the claim — the population sizes of Germany and Belgium are missing from the discussion. More crimes were committed in Germany, but if the population of Germany is much bigger than the population of Belgium, then the number of crimes per capita in Germany is likely to be lower than in Belgium.

(b) (3 pts) The researchers also found that throughout Western Europe there were 123,077 crimes committed in 2000 compared to 117,210 crimes committed in 2010. They concluded that overall, Europeans have become more law-abiding in the past decade. **Do you agree? Explain why or why not.**

In this case, the numbers do support the claim. The population of Europe in 2010 was greater than it was in 2000, and the number of crimes has gone down. This means that the crime rate has gone down.

2. California is evaluating a new rehabilitation program for prisoners before their release. The goal of the program is to reduce the recidivism rate — the percentage of released prisoners who return to prison within two years of their release. The program is voluntary and involves several months of “boot camp” (military style basic training with strict discipline). According to a prison spokesman, “Those who complete boot camp are less likely to return to prison than other inmates.”

(a) (2 pts) Is this study an observational study or a controlled experiment? Justify your answer.

This is an observational study — the subjects in the study choose which group they are in, control or treatment, not the researchers.

(b) (2 pts) What is the treatment group according to the spokesman’s statement? What is the control group?

The treatment group in this study is the group of inmates who volunteered for the boot camp and completed the program. The control group includes the inmates who did not volunteer for the boot camp and the inmates who did not complete the boot camp.

(c) (2 pts) Does the study show that the program is working, or is there another explanation for the results? Justify your answer briefly.

No — there is a possible confounding variable, namely the inmates’ motivation not to return to prison. Inmates who volunteer for the boot camp in the first place are likely to be motivated to stay out of prison in the future, and those that complete the program are likely to be the most motivated of all. It is possible that these prisoners would avoid returning to prison on their own, without the boot camp.
3. The histogram below shows the distribution of blood pressure for 14,148 women participating in a certain drug study. Use the histogram to answer the following questions. Explain your answers and show your work.

![Histogram of Blood Pressure Distribution](image)

(a) (3 pts) Is the percentage of women with blood pressure above 120mm closer to 47%, 57% or 67%?

Explanation: the percentage of women with blood pressure above 120mm is equal to the area of the histogram between 120 and 160. The heights of the bars from 120 to 140 are (approximately) 2.2, 2.1, 1.8, and 1.1, all with width 5, and the heights of the last two bars are (approximately) 0.8 and 0.2, both with width 10. So the total area is (approximately)

$$5 \times (2.2 + 2.1 + 1.8 + 1.1) + 10 \times (0.8 + 0.2) = 46,$$

which is closest to 47.

(b) (3 pts) Is the percentage of women with blood pressure between 105mm and 120mm closer to 36%, 41% or 46%?

Explanation: the percentage of women with blood pressure between 105mm and 120mm is equal to the area of the histogram between 105 and 120. The heights of the bars from 105 to 120 are (approximately) 1.95, 2.9 and 3.2, all with width 5. So the total area is (approximately)

$$5 \times (1.9 + 2.9 + 3.2) = 40$$

which is closest to 41.

(c) (2 pts) In what bin will we find the median blood pressure for women in this study?

The median is the number m such that 50% of the data lies below it (and the other 50% lies above it). About 40% of the data lies between 105 and 120 and about $10 \times 0.5 + 5 \times 1.6 = 13\%$

of the data lies between 90 and 105, so about 53% of the data lies below 120. Since $5 \times 3.2 = 16\%$ of the data lies between 115 and 120, it follows that the median lies in that bin—115 to 120—because 37% of the data lies below 115 and 53% of data lies below 120.