Problem Set 5 — KEY

1) population: all students in CA enrolled in public schools
parameter: true mean income level
Sample: 21,000 students that returned survey
statistic: mean income level of the sample ; $65,000

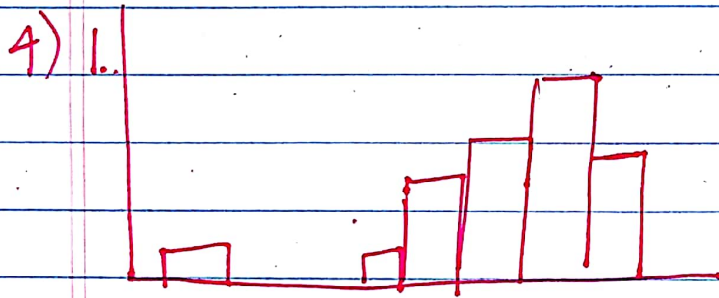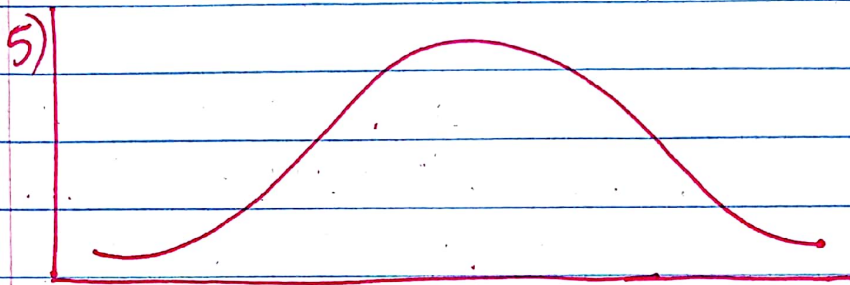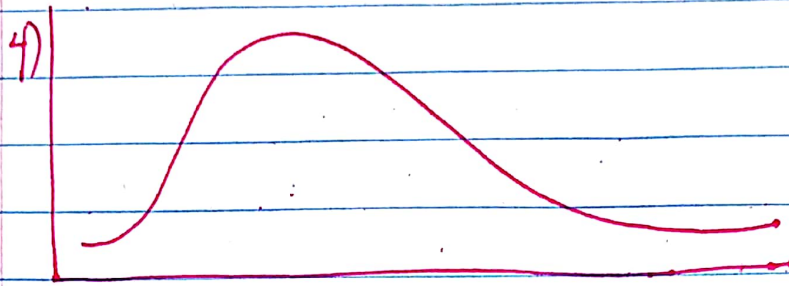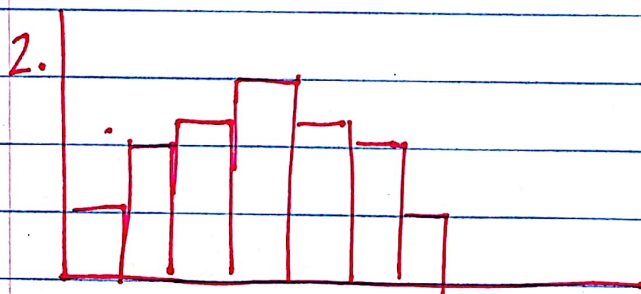2) a. Sex → categorical
b. height → quantitative (inches?)
c. handedness → categorical
d. distance → quantitative (yards)
e. sideline crossed → categorical

3) 1)

2)

3)

4)

5)

4) 1.
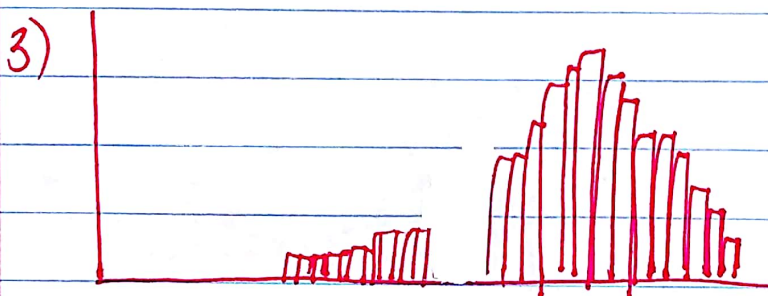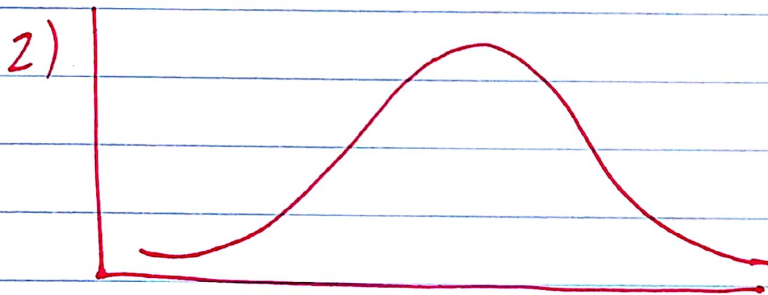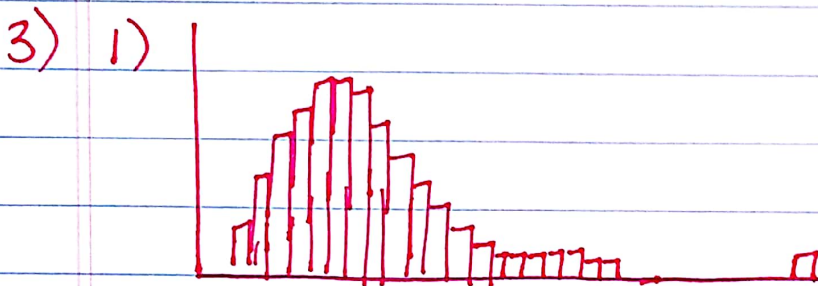
The distribution for Set A is unimodal and skewed left. The data is centered at a median of 37.5 units and has a range of 14 units. There is a possible low outlier at 26 units.

2.

The distribution of Set B is unimodal and appx. symmetric. The data is centered at a mean of 55 units and has a standard deviation of 1.70 units.

The most important thing here was to choose a
company and make a sound argument.

5) You could choose either company here. If you
chose [A], you'd need to give reasoning along the
lines of "A score of 300 shows that the curriculum
has been utilized to produce substantial gains.
Because VUSD has a highly qualified and
experienced teaching staff, they should be
confident that the staff can use the curriculum
to reproduce such results."

[B] Company's A distribution includes an observation
at 300, which by definition, is an unusual
outcome. Comparing medians, we find that
Company B's curriculum has a higher
median, 19, compared to Company A's
median of 4.5. Company B's dataset
reveals that all students made gains,
while some students using company
A's curriculum regressed. Excluding
the outlier for company A, 70% percent
of company B students outperformed
All Company A students.

Key takeaways: • Discuss outlier
                • Compare medians
                • Consider spread (if it helps your
                         compare   argument)
*Sorry. These didn't have good shapes to
                                       classify.