Problem 1

(a) A case is a collision between aircraft in wildlife.

(b) ac_mass is a continuous numerical variable

effect is a nominal categorical variable

num_engs is a discrete numerical variable

height is a continuous numerical variable

birds_struck is a discrete numerical variable

Problem 2

(a) It's not a variable, because it corresponds to more than one cases

(b) It's a categorical variable

(c) It's not a variable, because it corresponds to more than one cases

Problem 3

(a) This affects the mean. It increases the mean by

$$(98600 - 98600)/1000 = 887.4$$

(b) This doesn't affect the median.

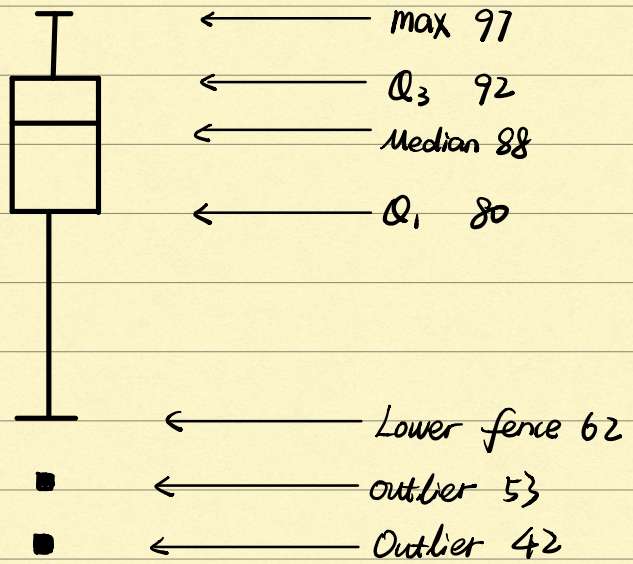
Problem 4

(a) $IQR = Q_3 - Q_1 = 12$

$Q_3 + 1.5\,IQR = 110$ $\quad$ $Q_1 - 1.5\,IQR = 62$

There are two outliers : 42, 53

(b)



Problem 5

(a) The distribution is expected to be right skewed. Because most of the data are on the left side of the histogram, the few larger values (> $6 million) bring the mean upwards and larger than median, so the data are right skewed.

(b) The distribution is expected to be right skewed. Because most of the data are on the left side of the histogram, the few larger values (drink excessively) bring the mean upwards and larger than median, so the data are right skewed.

(c) Measures of center and spread don't tell about skewness, but the distribution is highly possible to be left skewed. Because values within 1SD lies roughly 68%, which means 34% students have scores >85 and 66% have scores <85 (roughly). The tail lies at low score.