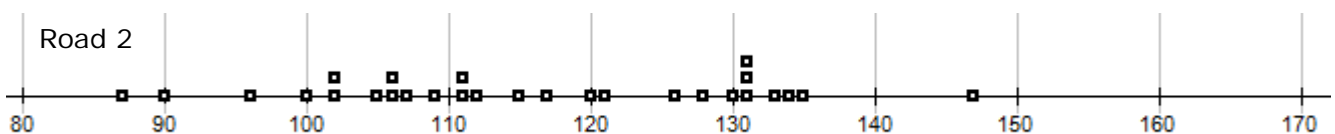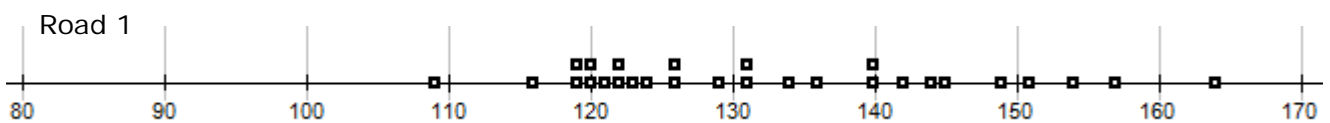# Routine Box and Whisker Practice #2

Q1   The number of cars are measured on two roads each day for four weeks (28 days).

Which road is busier, and by how much? Discuss any features in the data.

	Road 1	Road 2
Lowest	109	87
Q1	121	105.5
Median	130	113.5
Q3	143	130.5
Highest	164	147
Mean	132.6	115.8

Road 1

Road 2

Q2   The speed of the trucks on the two roads is also measured along the straights.

Describe the differences between the speeds on the two roads. (The speed limit is 100.)

	Road 1	Road 2
Lowest	85	91
Q1	96	97
Median	99	101
Q3	101	106
Highest	108	133
Mean	98.2	102.1

Road 1

Road 2

# Suggested Answers: Routine Box and Whisker Practice #2

Q1   The number of cars.
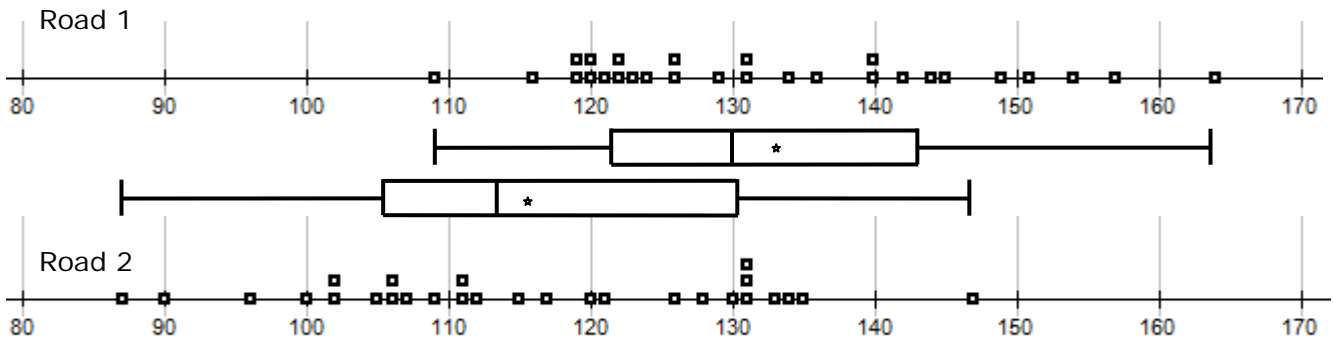
Road 1 is busier by about 15 to 20 cars a day.

1) Each point of its box and whisker is about 15 to 20 higher. In particular the typical values given by median and mean are about 17 different. But with such a small sample size, we cannot assume too much exact numerical accuracy (this is why I say 15–20).

   The difference seems significant. The median for Road 2 is well outside the IQR for Road 1 and the median for Road 1 is very nearly outside the IQR for Road 2. This suggests the difference is probably not just sample variation.

2) Both roads have a fairly even spread of values. Although there are small clumps at 120 for Road 1 and at 130 for Road 2, that may well just be random variation.

3) Both roads have very similar ranges and IQRs.

4) Both distributions have slightly longer tails to the right (skew) which is why the means are a bit higher than the medians.

Q2   The speed of trucks.
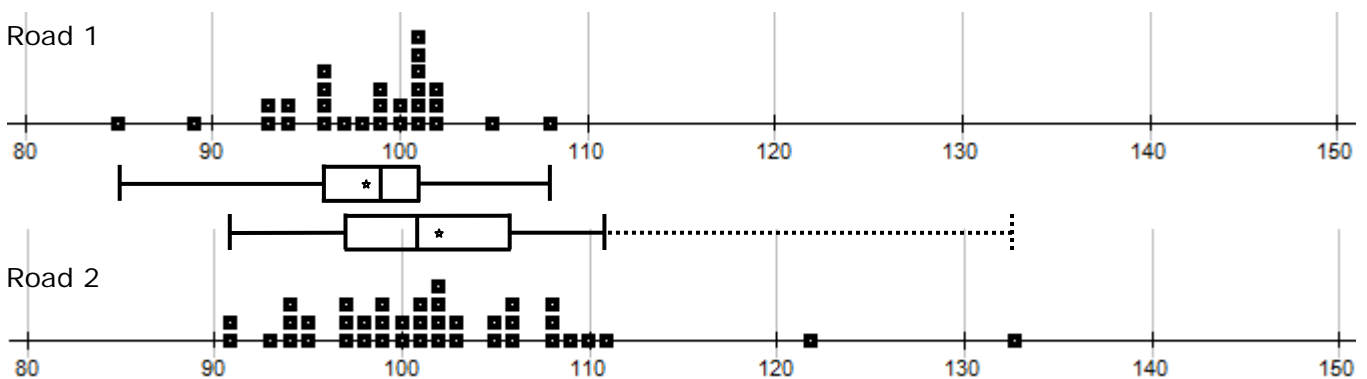
The trucks are measured as going typically faster on Road 2 but the difference is **not** significant.

1) Each point on the box and whisker for Road 2 is higher so it seems it has higher speeds. But both medians are inside the opposite IQRs. This means the difference may easy be due to sample variation alone. Thus we cannot say the difference is significant.

   The mean for Road 2 is much higher, but this is largely due to the two extreme values.

2) Road 1 has an even spread of values, with a couple high and a couple low. Road 2 has two extreme values that greatly distort what is already a less highly centred distribution.

3) Both roads have very similar ranges and IQRs, if one ignores the two extreme values.

We need to know if extreme speeds are a feature of Road 2, or whether our sample is unusual, before we can make firm answers on which has higher speeds.

Note: it is **not** a relevant feature that Road 2 has a bigger sample. That is why we use statistics!