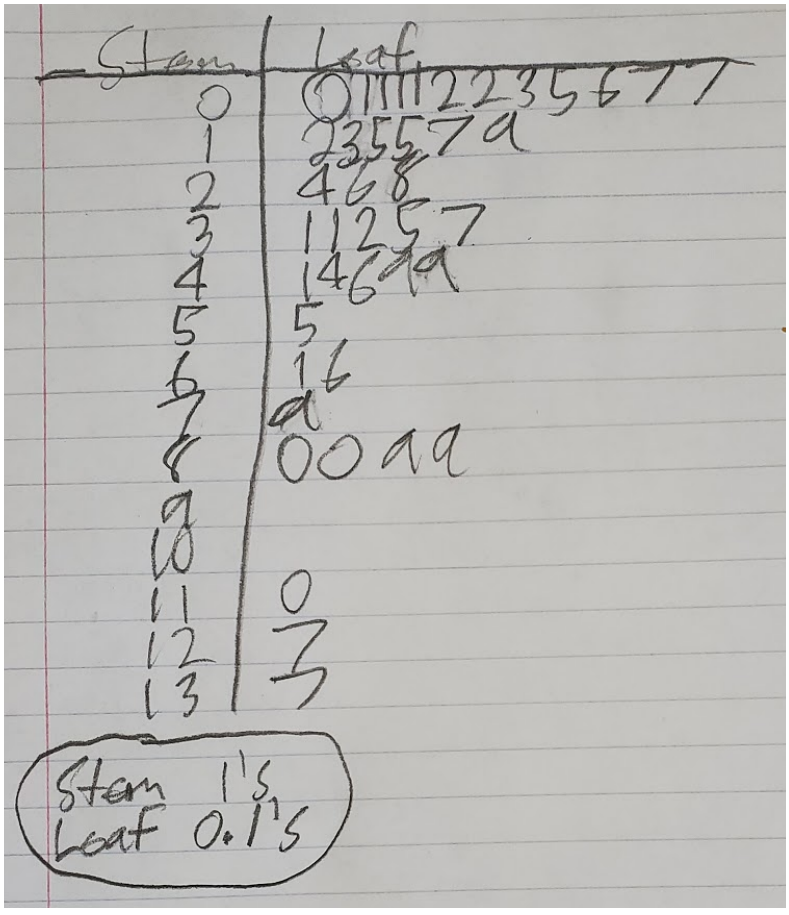
STAT 50 HW #2 Section 1.3

1.

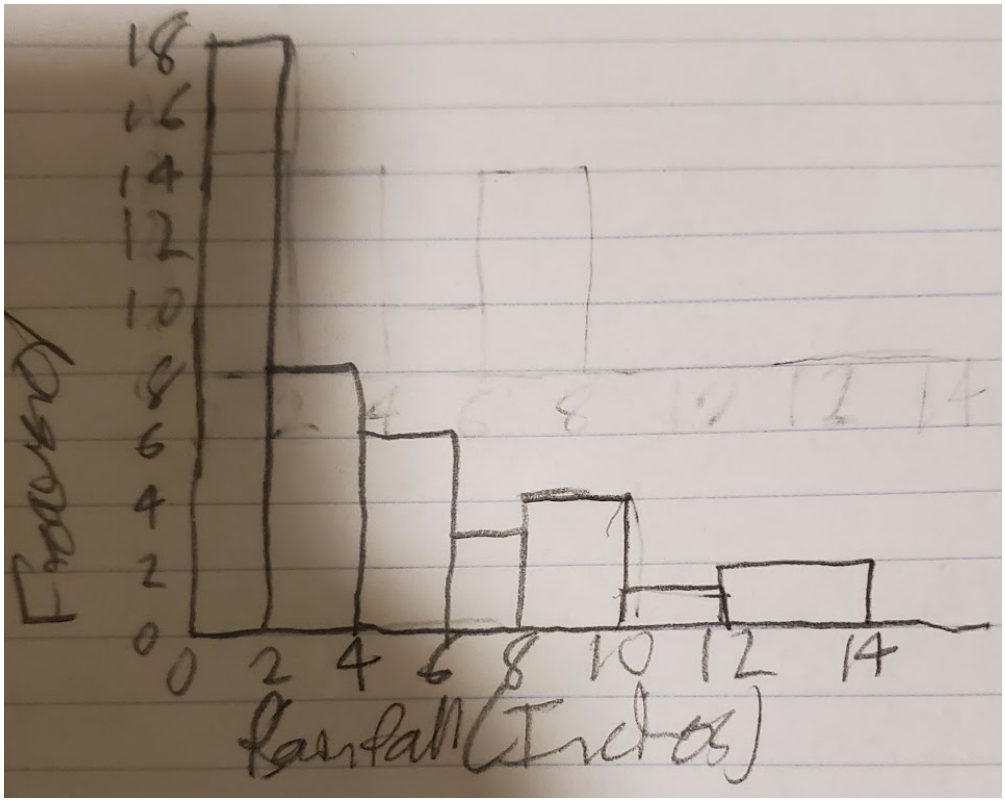
The weather in Los Angeles is dry most of the time, but it can be quite rainy in the winter. The rainiest month of the year is February. The following table presents the annual rainfall in Los Angeles, in inches, for each February from 1965 to 2006.

0.2	3.7	1.2	13.7	1.5	0.2	1.7
0.6	0.1	8.9	1.9	5.5	0.5	3.1
3.1	8.9	8.0	12.7	4.1	0.3	2.6
1.5	8.0	4.6	0.7	0.7	6.6	4.9
0.1	4.4	3.2	11.0	7.9	0.0	1.3
2.4	0.1	2.8	4.9	3.5	6.1	0.1

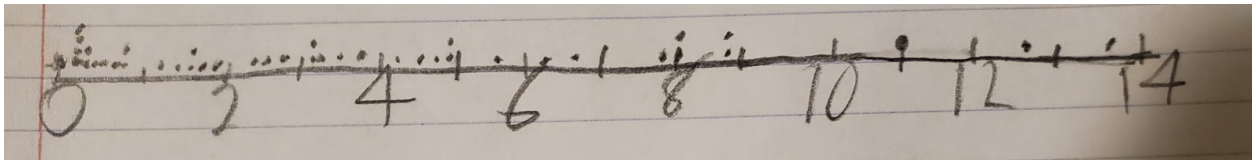
a. **Construct a stem-and-leaf plot for these data.**



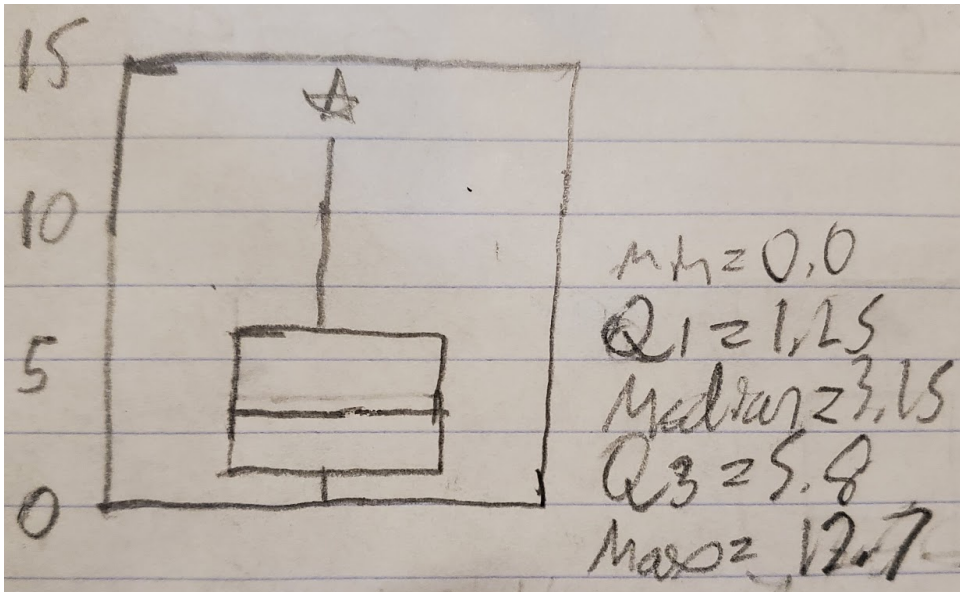
b. **Construct a histogram for these data.**



c. **Construct a dotplot for these data.**



d. **Construct a boxplot for these data. Does the box-plot show any outliers?**



The box-plot does show one outlier, which is 13.7.

5.

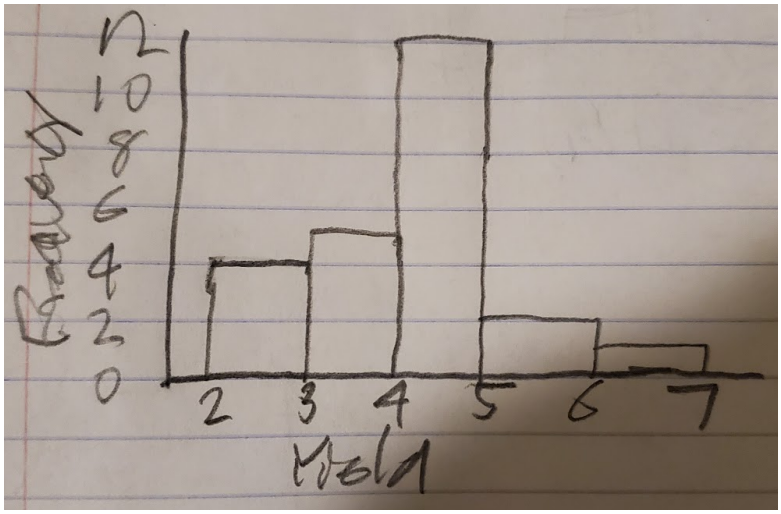
A certain reaction was run several times using each of two catalysts, A and B. The catalysts were supposed to control the yield of an undesirable side product. Results, in units of percentage yield, for 24 runs of catalyst A and 20 runs of catalyst B are as follows:

Catalyst A			
4.4	3.4	2.6	3.8
4.9	4.6	5.2	4.7
4.1	2.6	6.7	4.1
3.6	2.9	2.6	4.0
4.3	3.9	4.8	4.5
4.4	3.1	5.7	4.5

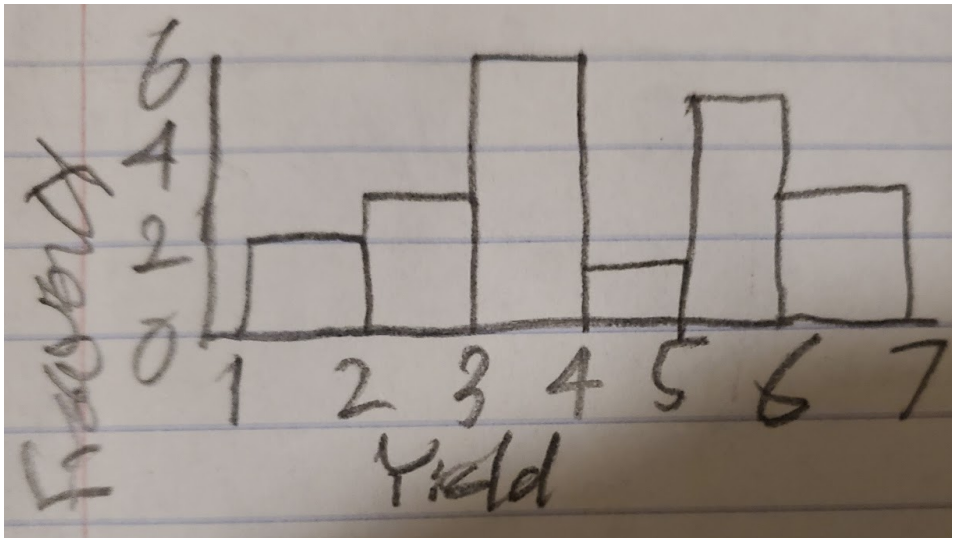
Catalyst B			
3.4	1.1	2.9	5.5
6.4	5.0	5.8	2.5
3.7	3.8	3.1	1.6
3.5	5.9	6.7	5.2
6.3	2.6	4.3	3.8

a. **Construct a histogram for the yields of each catalyst.**

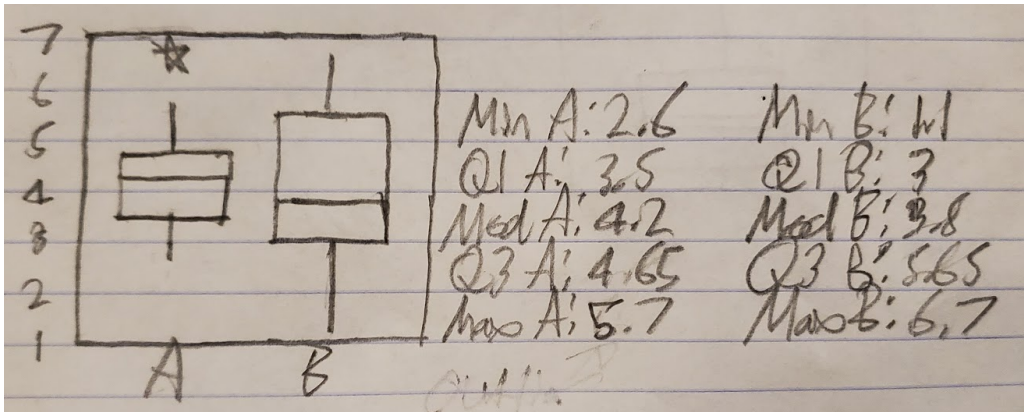
Catalyst A:



Catalyst B:



b. **Construct comparative boxplots for the yields of the two catalysts.**



c. **Using the boxplots, what differences can be seen between the results of the yields of the two catalysts?**

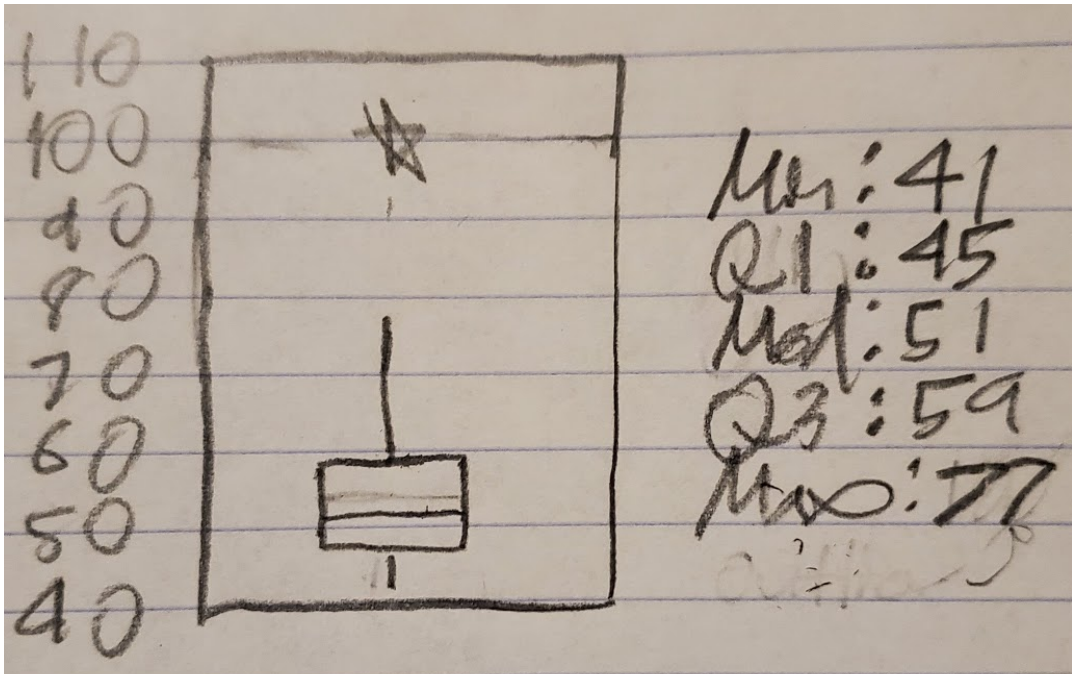
The yields for catalyst B are more spread out and have a larger interquartile range. The median for catalyst A is larger than that of catalyst B. The median for catalyst B is closer to the 1st quartile than the third, but appears to be equally distant from the largest and smallest values in the group based on the whiskers. Based on the whiskers, one could say that the yields for catalyst B are somewhat symmetric. There is an Outlier for one of the yields of catalyst A and the yields for A (not including the outlier) appear to be rather symmetric as well.

11.

The following table presents the number of students absent in a middle school in northwestern Montana for each school day in January 2008.

Date	Number Absent
Jan. 2	65
Jan. 3	67
Jan. 4	71
Jan. 7	57
Jan. 8	51
Jan. 9	49
Jan. 10	44
Jan. 11	41
Jan. 14	59
Jan. 15	49
Jan. 16	42
Jan. 17	56
Jan. 18	45
Jan. 21	77
Jan. 22	44
Jan. 23	42
Jan. 24	45
Jan. 25	46
Jan. 28	100
Jan. 29	59
Jan. 30	53
Jan. 31	51

a. **Construct a boxplot.**



b. **There was a snowstorm on January 27. Was the number of absences the next day an outlier?**

Based on the boxplot, yes. The value of 100 is an outlier.

15.

Following are summary statistics for two data sets, A and B.

	A	B
Minimum	0.066	−2.235
1st Quartile	1.42	5.27
Median	2.60	8.03
3rd Quartile	6.02	9.13
Maximum	10.08	10.51

a. **Compute the interquartile ranges for both A and B.**

IQR for set A:

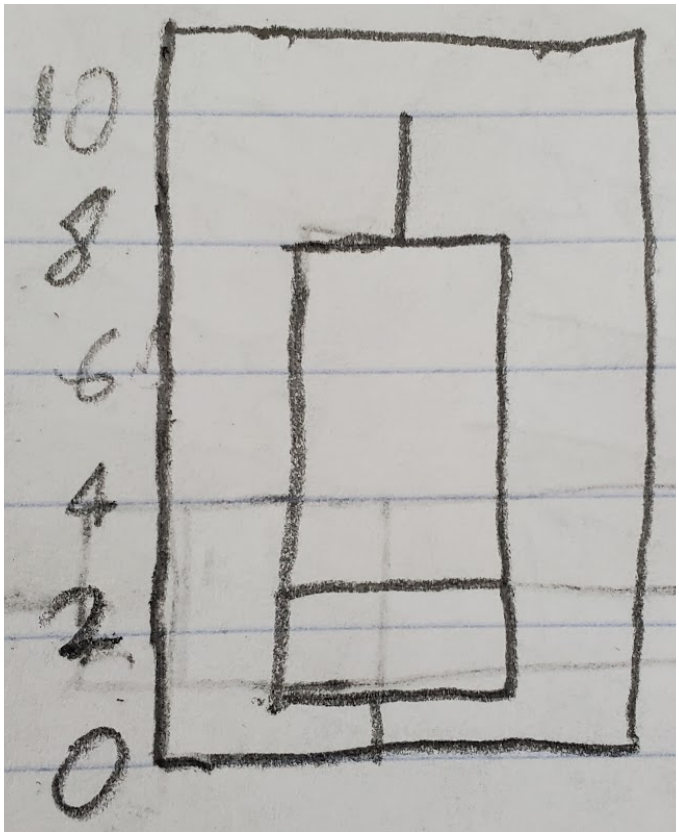
6.02 - 1.42 = 4.6

IQR for set B:

9.13 - 5.27 = 3.86

b. **Do the summary statistics for A provide enough information to construct a boxplot? If so, construct the boxplot. If not, explain why.**

Yes.

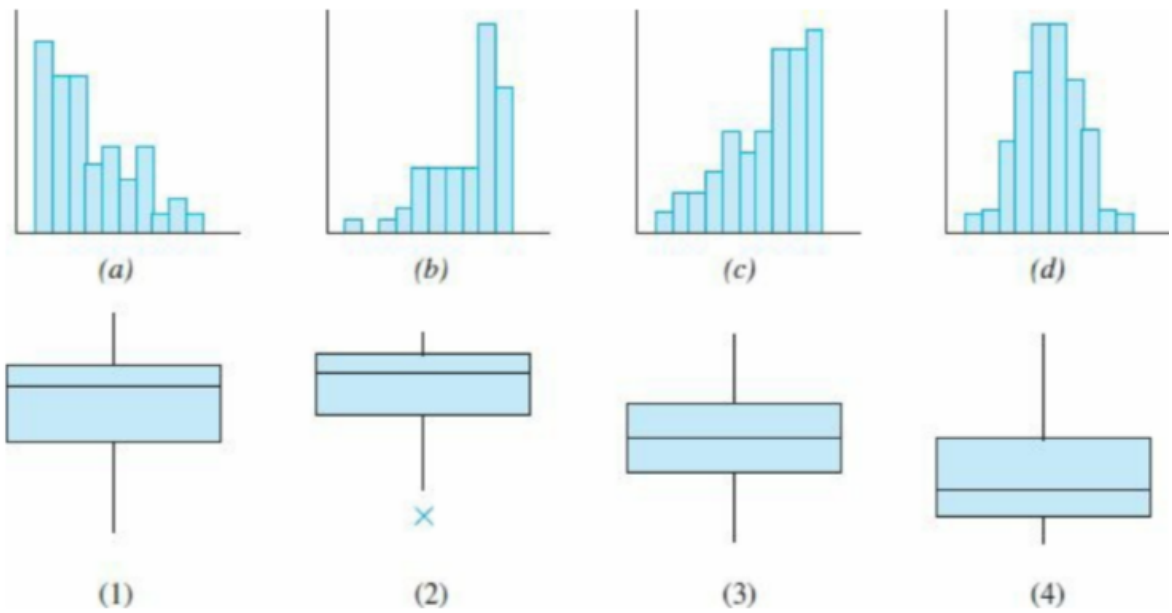


c. **Do the summary statistics for B provide enough information to construct a boxplot? If so, construct the boxplot. If not, explain why.**

No. If we multiply the IQR by 1.5 (3.86 * 1.5 = 5.79) and subtract the result from the 1st quartile (5.27 - 5.79 = -0.52), we should notice that the minimum of -2.235 is an outlier. Because this is a five number summary, we don't know what minimum value (*that isn't an outlier) our lower whisker should go to instead.

16.

Match each histogram to the boxplot that represents the same data set.



d→3

a→4

b→2

c→1