1. An offensive line coach for an American football team is running some analyses on his players. Their weights (in lbs) are given in the stem-and-leaf plot below:

```
Offensive Lineman Weights

30 | 7
31 | 113788
32 | 4
33 | 68
34 |
35 | 3

Key: 30|7 means a lineman with weight 307 lbs
```

(a) Calculate the mean, median, first quartile, and third quartile for this data. *The mean is* $(307+311+311+313+317+318+318+324+336+338+353)/11 = 322.4$ *lbs. The median is the middle number in the sorted list. Since there are 11 data points, this would be the 6th number, or* 318 *lbs. The first quartile is the median of the lower half of the data omitting the median, which is the 3rd number, or* 311 *lbs, and the third quartile is the median of the upper half of the data omitting the median, which is the 8th number, or* 336.

(b) Calculate the population SD, range, and interquartile range (IQR) for this data. Why might the IQR be a more useful statistic than range or population SD for this data?

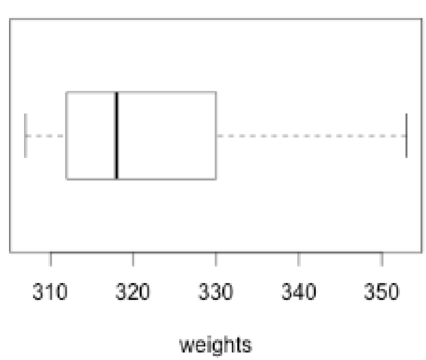
The population SD is $\sqrt{\frac{(307-322.4)^2+(311-322.4)^2+...+(353-322.4)^2}{11}} = 13.6$. *(Divided by n for population SD and divided by n-1 for sample SD.)*

The range is the difference between the largest and smallest value, or $353-307 = 46$.

The IQR is the difference between the third and first quartiles, computed in part (a), or $336-311 = 25$.

The shape of the data is somewhat skewed right, and there is one extreme value. Since the IQR is less sensitive to extreme values, it is perhaps a better measure for this data.

(c) Draw a boxplot of the data.



(d) Compare and contrast the information you can glean from the stem-and-leaf plot vs the boxplot. For example, is there information you gain/lose when you use the boxplot vs stem-and-leaf plot? *Both graphs show the shape well - the data is right skewed. The stem-and-leaf shows actual data values (and their frequency), but no summary information. The boxplot show some summary information, like the median, range, and IQR very easily, but you can't recover the original data or even the number of data points.*

2. Assume the heights of the population of sixteen-year-old girls in Great Britain are approximately normally distributed with a mean of 66 inches (5 feet 6 inches) and a standard deviation of 2.5 inches.

(a) A particular sixteen-year-old girl in the population is 62 inches tall. Calculate her z-score, and interpret this z-score in context. What percentage of sixteen-year-old girls do we expect to be taller than her? Shorter than her? What is her percentile height?
Recall that $z - score = \frac{observation - mean}{standard deviation}$. *Her z-score is* $\frac{62-66}{2.5} = -1.6$.
In context, this means that this girl has a height which is 1.6 standard deviations below the mean.
From the normal table, if we look up 1.6 *along the left side and* 0.00 *along the top, the value in the table is* 0.9452, *or* 94.52%. *This is the area to the left of* 1.60. *But by the symmetry of the normal, the area to the left of* 1.60 *is the same as the area to the right of* -1.60. *So about* 94.52% *of girls are taller than this particular girl. This means that* $100 - 94.52 = 5.48\%$ *of girls are shorter than her. So her height falls between the 5th and 6th percentile.*

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998

(b) Approximately what percentage of all sixteen-year-old girls in the population do we expect to have heights between 63.5 and 68.5 inches? *First find the two z-scores:* $\frac{63.5-66}{2.5} = -1.00$ *and* $\frac{68.5-66}{2.5} = 1.00$. *From the z-table, we calculate that the area to the left of 1.00 is* 84.13%, *which means that the area to the left of* -1.00 *is* $100 - 84.13 = 15.87\%$. *The area between them is* $84.13 - 15.87 = 68.26\%$. *So about 68% of girls have heights between 63.5 and 68.5 inches.*

(c) Suppose a particular sixteen-year-old girl in the population is at the 93rd percentile for height. How tall is she in inches? *We start by looking in the z-table for a value near* 0.9300*. The closest we can get is* 0.9306*, which looking to the edges of the table occurs at a z-score of about* 1.48*. We convert this to inches by multiplying by the SD and adding the mean:* $1.48 * 2.5 + 66 = 69.7$*. So the girl is about 69.7 inches tall.*

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998

(d) Consider three sixteen-year-old girls, Anne, Betty, and Cathy. Anne is 64 inches tall, Betty has a z-score of -0.2, and Cathy is at the 49th percentile. List the three girls from tallest to shortest. *You could do this problem by converting each girl's height into any of the three scales: inches, z-score, or percentile. The easiest is probably z-score. Anne's z-score is* $\frac{64-66}{2.5} = -0.8$*, and Cathy's can be gotten from the normal table and is* -0.025*. Therefore, from tallest to shortest is Cathy, Betty, Anne.*

	Inches	Z-score	Percentile
Anne	64	-0.8	21st
Betty	65.9375	-0.025	49th
Cathy	65.95	-0.2	42nd

3. Six people took a survey. The times to complete the survey (in minutes) were: $7, 12, 17, 22, 12, 5$. This yields an average of 12.5 minutes.

   (a) What is the sample standard deviation of this data? $\sqrt{\frac{(7-12.5)^2+(12-12.5)^2+...+(5-12.5)^2}{5}} = 6.28$

   (b) If we multiply each person's time by 2 what happens to the average and standard deviation?
   Both the average and SD would be multiplied by 2. So the new mean is 25 minutes, and the new SD is 12.56 minutes.
   Before:
   $mean = \frac{(7+12+17+22+12.5)}{5} = 12.5$
   $SD = \sqrt{\frac{(7-12.5)^2+(12-12.5)^2+(17-12.5)^2+(22-12.5)^2+(12.5-12.5)^2}{5}} = 6.28$
   After:
   $mean = \frac{(2\times7+2\times12+2\times17+2\times22+2\times12.5)}{5} = 2 \times \frac{(7+12+17+22+12.5)}{5} = 2 \times 12.5$
   $SD = \sqrt{\frac{(2\times7-2\times12.5)^2+(2\times12-2\times12.5)^2+(2\times17-2\times12.5)^2+(2\times22-2\times12.5)^2+(2\times12.5-2\times12.5)^2}{5}}$
   $= \sqrt{4 \times (\frac{(7-12.5)^2+(12-12.5)^2+(17-12.5)^2+(22-12.5)^2+(12.5-12.5)^2)}{5}} = 2 \times 6.28$

   (c) If we add 10 minutes to each person's time, what happens to the average and standard deviation?
   The average would increase by 10, to 22.5 minutes, since all the data points would be shifted over by that amount. However, the spread of the data is not affected by adding the same number to all data points, so the SD remains 6.28 minutes.
   Before:
   $mean = \frac{(7+12+17+22+12.5)}{5} = 12.5$
   $SD = \sqrt{\frac{(7-12.5)^2+(12-12.5)^2+(17-12.5)^2+(22-12.5)^2+(12.5-12.5)^2}{5}} = 6.28$
   After:
   $mean = \frac{((7+10)+(12+10)+(17+10)+(22+10)+(12.5+10))}{5} = \frac{(7+12+17+22+12.5)+5\times10}{5} = 12.5 + 10$
   SD: For each data points, its deviation from mean doesn't change. So SD remains the same.

   (d) If we multiply each time by -1, what happens to the average and standard deviation? *Similar to (b). The average would be -1 times the original average, or -12.5 mintues. But changing the sign on the data points again does not change the spread, it simply puts all the data points on the other side of zero. So the SD stays the same, 6.28 minutes.*