Show all of your work. Please staple in the upper left hand corner.
There are problems out of the section exercises on this homework. You *must show all work* to receive credit. If answers are directly copied from the back of the book no points will be awarded. Write answers in your own words.

1. Chapter 6 - Section 1 - Question 2. Do parts (a)-(e) as the problem dictations (using large-sample normal distribution based confidence intervals).

   (a) *(3 pts)* Using normal distribution quantiles we want to use $z_{.95} = 1.645$ in our confidence interval

$$\bar{y} \pm z_{.95}\frac{s}{\sqrt{n}}$$
$$142.7 \pm 1.645\frac{98.2}{\sqrt{26}}$$
$$142.7 \pm 31.67759$$

   So our 90% confidence interval is: (111.0224, 174.3776)

   (b) *(4 pts)* For a 95% confidence interval we want to use $z_{.975} = 1.96$

$$\bar{y} \pm z_{.975}\frac{s}{\sqrt{n}}$$
$$142.7 \pm 1.96\frac{98.2}{\sqrt{26}}$$
$$142.7 \pm 37.74617$$

   So our 95% confidence interval is: (104.9538, 180.4462). This is a wider interval than the one in part (a) which makes sense since we want to be more confident that the true mean falls in our interval and to do that we need to consider more possible values for the true mean.

   (c) *(4 pts)* For a 90% upper confidence bound we want to use $z_{.9} = 1.28$ as the quantile we use to calculate the upper endpoint.
   The upper end point has the form:

$$\begin{aligned}\bar{y} + z_{.9}\frac{s}{\sqrt{n}} &= 142.7 + 1.28\frac{98.2}{\sqrt{26}} \\ &= 142.7 + 24.68089 \\ &= 167.3809\end{aligned}$$

   So our 90% upper confidence bound is 167.38. We could also say that a 90% confidence interval is $(-\infty, 167.28)$. This value is lower than the upper endpoint in part (a). This makes sense because we are including all of the values below the lower endpoint for the interval in part (a) so this allows us to bring the upper bound down a little bit.

   (d) *(4 pts)* For a 95% upper confidence bound we want to use $z_{.95} = 1.645$ as the quantile we use to calculate the upper endpoint.
   The upper end point has the form:

$$\begin{aligned}\bar{y} + z_{.95}\frac{s}{\sqrt{n}} &= 142.7 + 1.645\frac{98.2}{\sqrt{26}} \\ &= 142.7 + 31.67759 \\ &= 174.3776\end{aligned}$$

So our 95% upper confidence bound is 174.3776. This value is higher than the value in part (c) since we need to include more possible values for μ to gain more confidence that our bound truly captures the true parameter. (Side note: notice that this upper bound is the same upper bound as the one in part (a) - can you explain why that is?)

(e) *(3 pts)* We are 90% confident that the true mean aluminum content for recycled PET plastic from this recycling facility is between 111.0 and 174.3776 ppm.

2. Chapter 6 - Section 1 - Question 2. Do parts (a)-(e) using t-based confidence intervals.

(a) *(3 pts)* Using t distribution quantiles we want to use $t_{.95,25} = 1.708$ in our confidence interval

$$\begin{aligned}
&\bar{y} \pm t_{.95,25}\frac{s}{\sqrt{n}} \\
&142.7 \pm 1.708\frac{98.2}{\sqrt{26}} \\
&142.7 \pm 32.8937
\end{aligned}$$

So our 90% confidence interval is: (109.8063, 175.5937)

(b) *(4 pts)* For a 95% confidence interval we want to use $t_{.975,25} = 2.06$

$$\begin{aligned}
&\bar{y} \pm t_{.975,25}\frac{s}{\sqrt{n}} \\
&142.7 \pm 2.06\frac{98.2}{\sqrt{26}} \\
&142.7 \pm 39.67273
\end{aligned}$$

So our 95% confidence interval is: (103.0273, 182.3727). This is a wider interval than the one in part (a) which makes sense since we want to be more confident that the true mean falls in our interval and to do that we need to consider more possible values for the true mean.

(c) *(4 pts)* For a 90% upper confidence bound we want to use $t_{.9,25} = 1.316$ as the quantile we use to calculate the upper endpoint.

The upper end point has the form:

$$\begin{aligned}
\bar{y} + t_{.9,25}\frac{s}{\sqrt{n}} &= 142.7 + 1.316\frac{98.2}{\sqrt{26}} \\
&= 142.7 + 25.34432 \\
&= 168.0443
\end{aligned}$$

So our 90% upper confidence bound is 168.0443. We could also say that a 90% confidence interval is $(-\infty, 168.0443)$. This value is lower than the upper endpoint in part (a). This makes sense because we are including all of the values below the lower endpoint for the interval in part (a) so this allows us to bring the upper bound down a little bit.

(d) *(4 pts)* For a 95% upper confidence bound we want to use $t_{.95,25} = 1.708$ as the quantile we use to calculate the upper endpoint.

The upper end point has the form:

$$\begin{aligned}
\bar{y} + t_{.95,25}\frac{s}{\sqrt{n}} &= 142.7 + 1.708\frac{98.2}{\sqrt{26}} \\
&= 142.7 + 32.8937 \\
&= 175.5937
\end{aligned}$$

So our 95% upper confidence bound is 175.5937. This value is higher than the value in part (c) since we need to include more possible values for μ to gain more confidence that our bound truly captures the true parameter. (Side note: notice that this upper bound is the same upper bound as the one in part (a) - can you explain why that is?)

(e) *(3 pts)* We are 90% confident that the true mean aluminum content for recycled PET plastic from this recycling facility is between 109.8063 and 175.5937 ppm.

So using the t-distribution didn't change too much in this case but it was a more honest analysis of the data. We would want to check if the data itself was approximately normally distributed too.

3. *(5 pts)* Chapter 6 - Section 1 - Question 3.

We have the equation

$$n = \left(\frac{z_{1-\alpha/2} * s}{\Delta}\right)^2$$

where Δ is the desired (half) width of the confidence interval, $z_{1-\alpha/2}$ is the normal quantile that we would use in our $(1-\alpha)$% confidence interval and s is the standard deviation. We are given $\Delta = 20$ and for a 90% confidence interval we would use $z_{.95} = 1.645$. We need s which we can estimate using the previous problem so use $s = 98.2$. This gives

$$\begin{aligned} n &= \left(\frac{z_{1-\alpha/2} * s}{\Delta}\right)^2 \\ &= \left(\frac{1.645 * 98.2}{20}\right)^2 \\ &= 65.23712 \end{aligned}$$

Since our sample size needs to be an integer and since this is the minimum sample size to meet our requirements we need to round up to give $n = 66$.

4. *(9 pts)* Read Guinness, Gosset, Fisher, and Small Samples. William Gosset is credited as the discoverer of the t-distribution. Provide a brief summary (.5-1 pages) of the events that led to him making his discovery.

Answers will vary