1) Generate random samples using a normal distributions
We are going to generate random samples from a number of different distributions in this laboratory. The following code is for the normal distribution which is the only one that we have discussed so far in class. I will also be providing a similar code for the other distributions that we will be using in part C. The function that is used in SAS is RAND('Normal',mu,sigma).

a) Generate 20 random numbers from a normal distribution with \( \mu = 572 \) and \( \sigma = 51 \) and calculate the mean and standard deviation of the data set.

**Solution:**
*random number generation for normal is RAND('Normal',mu,sigma):

/*The %Let statement allows you to define a variable at the beginning of the code. This is useful if you want to change some of the values when you are making multiple runs because you don't have to search through the whole code to change them. To access this variable later on, put an & in front of the variable name. The first line is %Let points = 20, to access this later, use &points */

%Let points = 20; *this is the number of data points in the sample;
%Let mu = 572;
%Let sigma = 51;
%Let norm = rand('normal',&mu, &sigma);
data RandomNormal;
do x=1 to &points; *When I use &points, I don’t need to search to change the number of points. The change is now only done in the beginning of the code;
   answer=&norm;
   output;
end;
/* Only print out the randomly generated data sets, do not print out the data from the cleaned airline data set*/
proc print data=RandomNormal; run;

/*process means outputs just the number of data points, mean, standard deviation, minimum and maximum. This outputs much less material than proc univariate */
proc means data=RandomNormal;
   var answer; *I want the information just for the variable answer;
run;
Note: Each time that the program is run, you will get different values, and means, and standard deviations.

2) Determine if a distribution is normal

b) Make an appropriate histogram of the data in part (a) and visually assess if the normal density curve and the histogram density estimate are similar.

c) Make a normal probability plot of the data in part (a) and visually assess if the sample quantiles are randomly scattered below and above the line without a discernable pattern.
Solution:
I am doing the problem with the data from part (a), but it doesn’t matter what data set is used.

```sas
proc sgplot data=RandomNormal;
title 'Random Normal';
histogram answer;
density answer; /* this adds the normal density curve of the appropriate mean and standard deviation;*/
density answer/type=kernel; /* this adds the smoothed density curve;*/
run;
```

```sas
proc univariate data=RandomNormal;
qqplot answer/normal (mu=est sigma=est); /* keywords after the / I am comparing this to a normal distribution (so SAS is looking up the z values) the line will be drawn using the data points to estimate mu and sigma;*/
run;
```

b) Make an appropriate histogram of the data in part (a) and visually assess if the normal density curve and the histogram density estimate are similar.

Since the blue normal curve is close to the red smoothed curve, the randomly generated normal data appears to be a normal distribution. When you are using a histogram to determine normality, please always include the two extra lines.
c) Make a normal probability plot of the data in part (a) and visually assess if the sample quantiles are randomly scattered below and above the line without a discernable pattern.

Since the sample quantiles are randomly scattered below and above the line without a pattern, the randomly generated normal data appears to be a normal distribution.

3) Generate random samples for right skewed, left skewed, short tailed, long tailed distributions

The specific distributions used are:
right skewed: exponential distribution ($\lambda = 2$) with $\mu = 0.5$ and $\sigma = 0.5$
left skewed: Beta distribution (on $[0,1]$), $\alpha = 2$, $\beta = 0.5$ with $\mu = 0.8$ and $\sigma = 0.0457$
short tailed: Uniform (on $[a = 0, b = 2]$) with $\mu = 1$ and $\sigma = 0.3333$
long tailed: Standard Cauchy with median = 0 and $\sigma$ is not defined.

Note: The Cauchy distribution has extremely straggly long tails, so the median is a better descriptor of the center than the mean.

The following commands are used for the above distributions. You just need to insert them into the code at the beginning and then replace &norm with the appropriate distribution.

```sas
%Let right = rand('Exponential',0.5);
%Let left = rand('Beta',2,0.5);
%Let short = 2*rand('Uniform');
%Let long = rand('Cauchy');
```

No output is provided. Every time the code is run, different output will be produced.