Untitled

Ron Cordell, Lei Yang, Subhashini Raghunathan

March 10, 2016

Question 1

1. Install the library "astsa" using the function: install.packages("astsa")
2. Load the library: library(astsa)
3. Use the function str() to see the information of a particular data series, such as str(EQ5) for the Seismic Trace of Earthquake number 5 series
4. Plot the time series plots and histograms of the following 3 series. Feel free to use the codes provided in the R scripts. Make sure that each of your graph has a title, the axis ticks are clear, the axes are well-labelled, and use color intelligently.
5. Write a few sentences to describe each of the series.

- EQ5
- flu
- gas

```
library(astsa)
```

```
## Warning: package 'astsa' was built under R version 3.2.3
```

```
str(EQ5)
```

```
##  Time-Series [1:2048] from 1 to 2048: 0.01749 0.01139 0.01512 0.01477 0.00651 ...
```

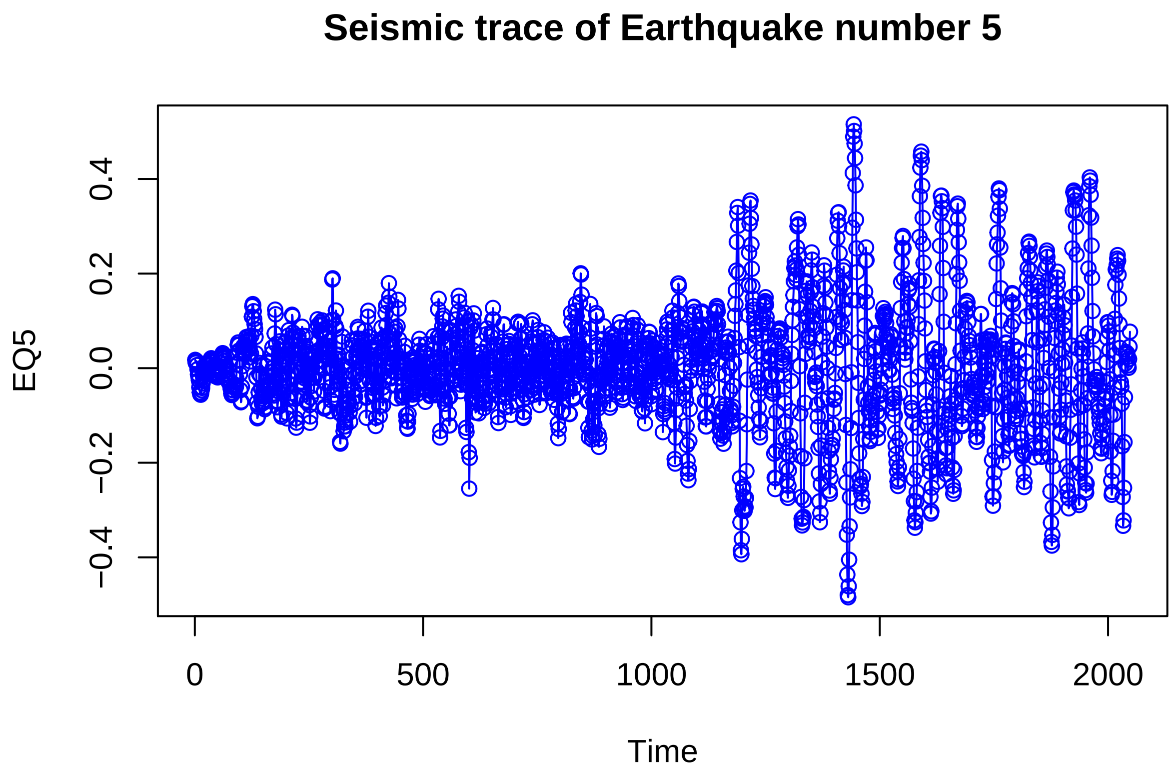
```
str(flu)
```

```
##  Time-Series [1:132] from 1968 to 1979: 0.811 0.446 0.342 0.277 0.248 ...
```

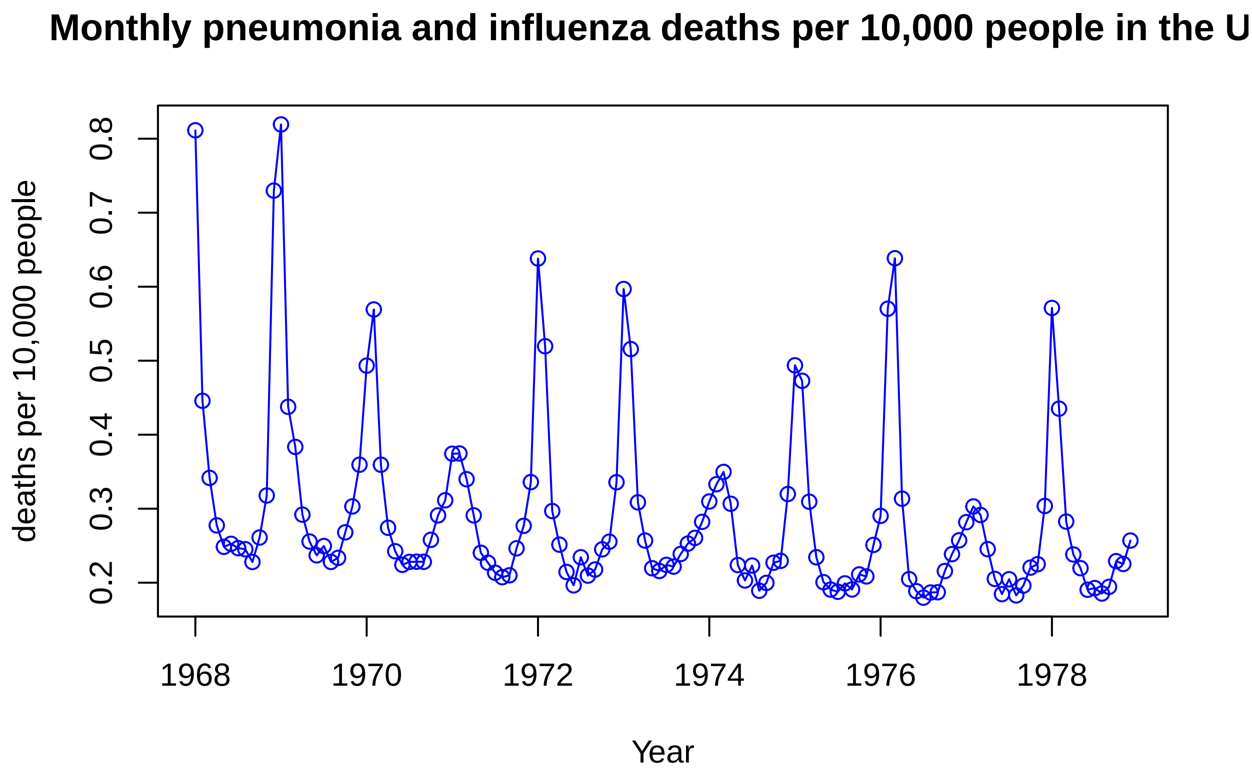
```
str(gas)
```

```
##  Time-Series [1:545] from 2000 to 2010: 70.6 71 68.5 65.1 67.9 ...
```

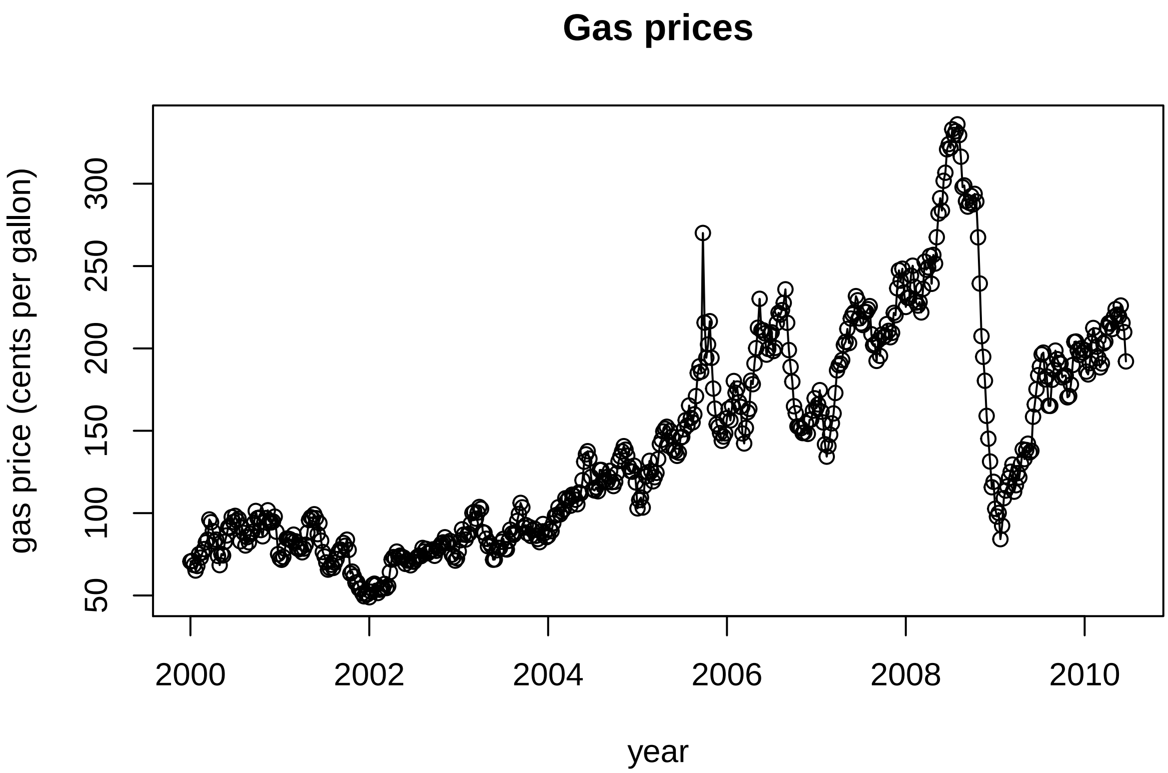
```
plot(EQ5,type="o", main="Seismic trace of Earthquake number 5", col="blue")
```




```r
plot(flu, type="o", main="Monthly pneumonia and influenza deaths per 10,000 people in the US",
     xlab = "Year", ylab ="deaths per 10,000 people", col="blue")
```




```r
plot(gas, type="o", main="Gas prices", xlab ="year", ylab="gas price (cents per gallon)")
```



EQ5: The earthquake plot follows the pattern of variation around a stable mean - in this case, the mean is 0 indicating no quake activity. The variation around the mean becomes larger as time increases. The earthquake data is desribed as 2 waves - primary and shear wave. We see a lot more variation in the shear wave.

flu: There is a lot of seasonal fluctuation in the data, and a slight downward trend. There also seems to be a cyclical pattern of 3 years - a two year period of high fatalities followed by a one year period of lower fatalities.

gas: The gas prices exhibit a general upward trend with a change in structure. From 2000-2002 the prices stay relatively flat. From 2002-20069 there is an upward trend. Then a structural change with a large drop in gas price in 2009, with another continuing upward trend.

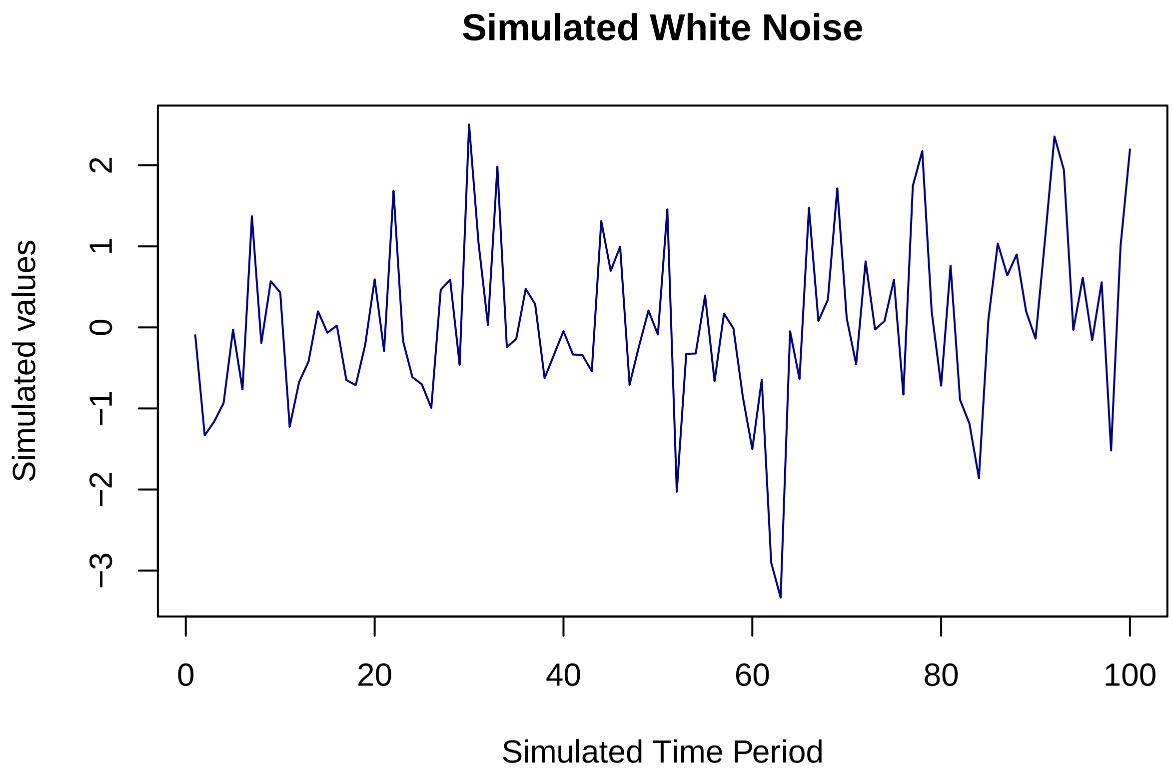
Question 2:

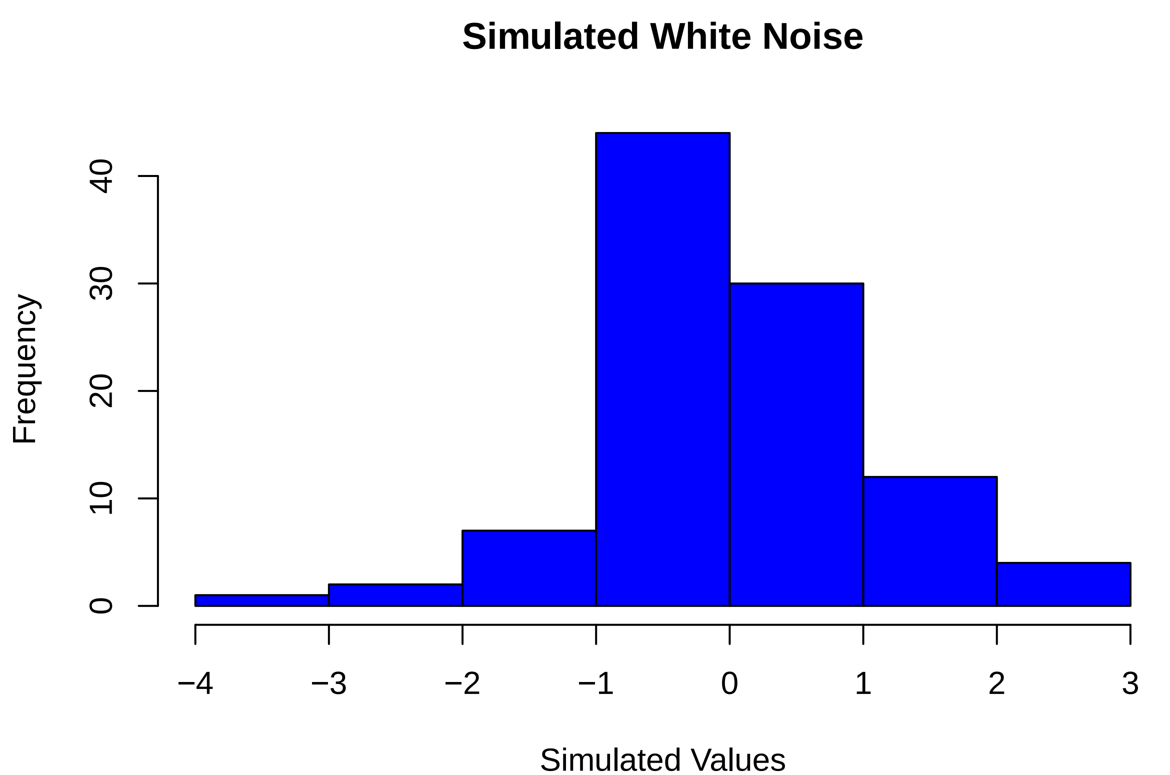
Describe 3 examples you have used in your work or encounter in real life. Ideally, you can even load at least one of these time series, plot it, and the write a few statements to describe its characteristics.

Some examples of time series: - Monthly precipitation (rain/snow) in a geographic area - Enrollment in a school district, measured yearly - Number of cases handled by the emregency unit of a hospital, measured monthly

Question 3:

Simulate a white noise series with 1000 random draws and plot (1) a time series plot and (2) a histogram. The usual reqirements on graphics (described) in Question 1) applied.


Simulated White Noise
Simulated values
2
1
0
−1
−2
−3
0
20
40
60
80
100
Simulated Time Period

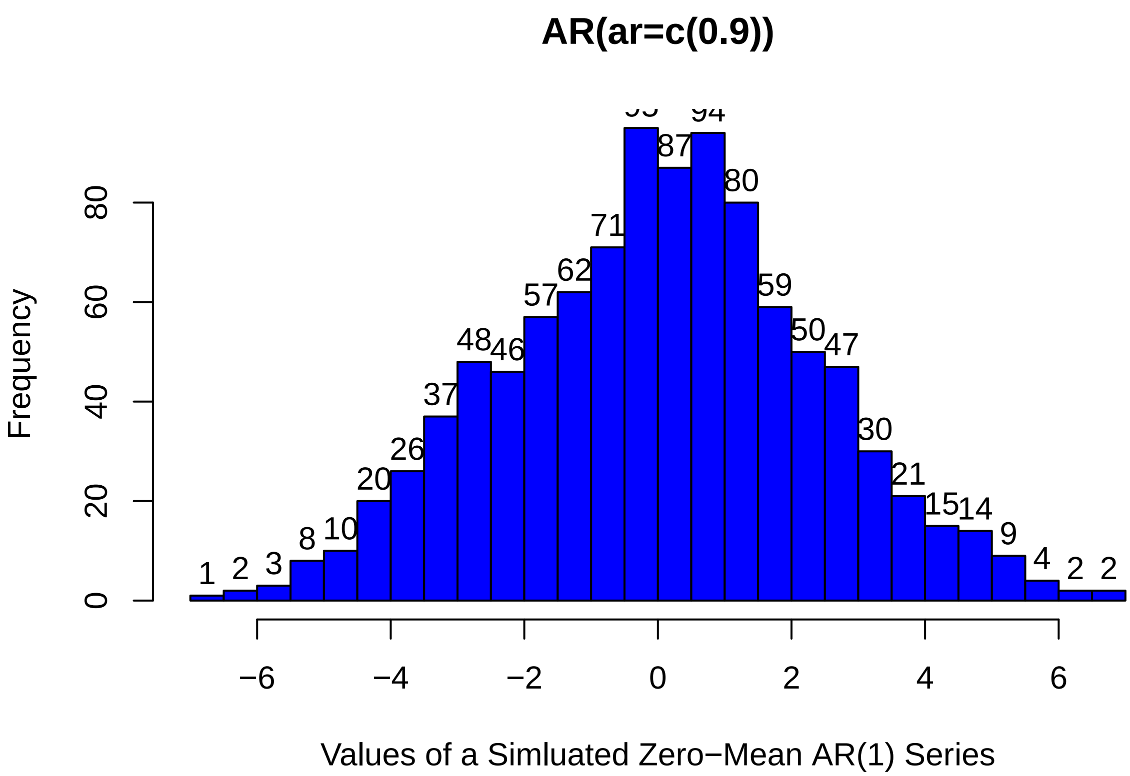

Simulated White Noise
Frequency
40
30
20
10
0
−4
−3
−2
−1
0
1
2
3
Simulated Values

Question 4:

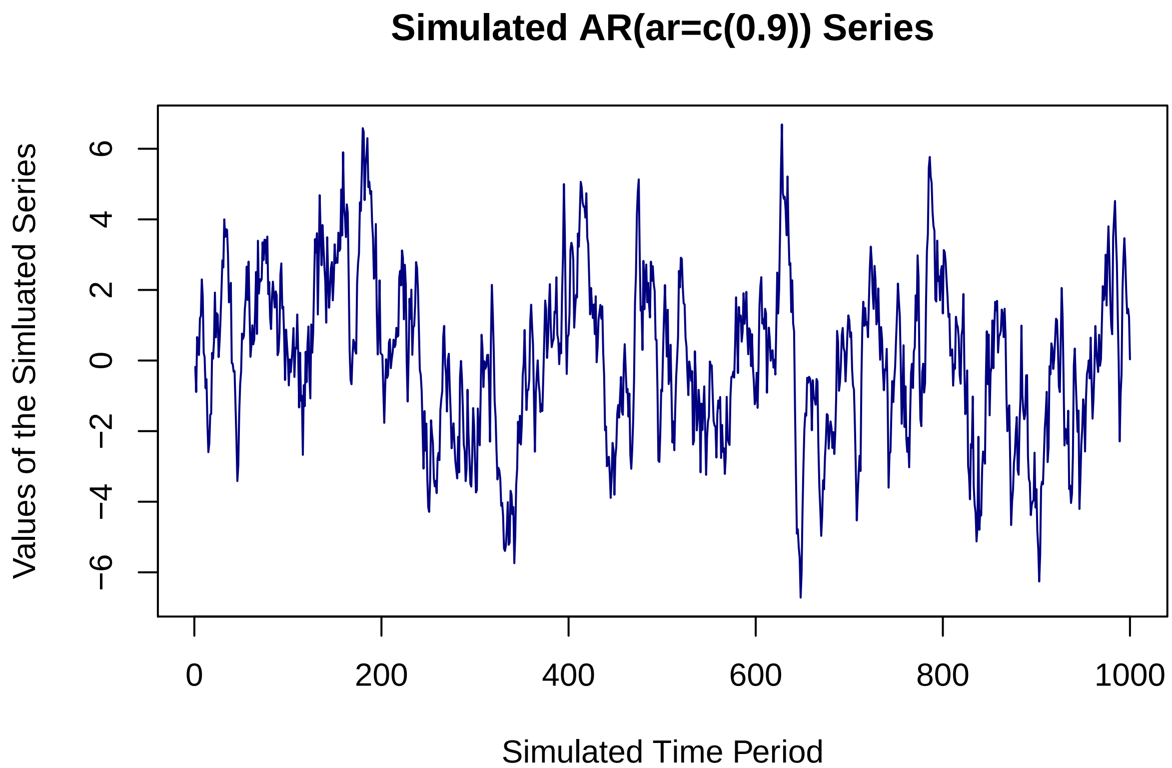
Simulate (with 1000 random draws) two the following two zero-mean autoregressive model with order 1 (i.e. AR(1)) models: yt = 0.9yt1 + w yt = 0.2yt1 + w

Plot a time plot for each of the simulated series. Graph a histogram for each of thes simulated series. Write a few statements to compare the two series.

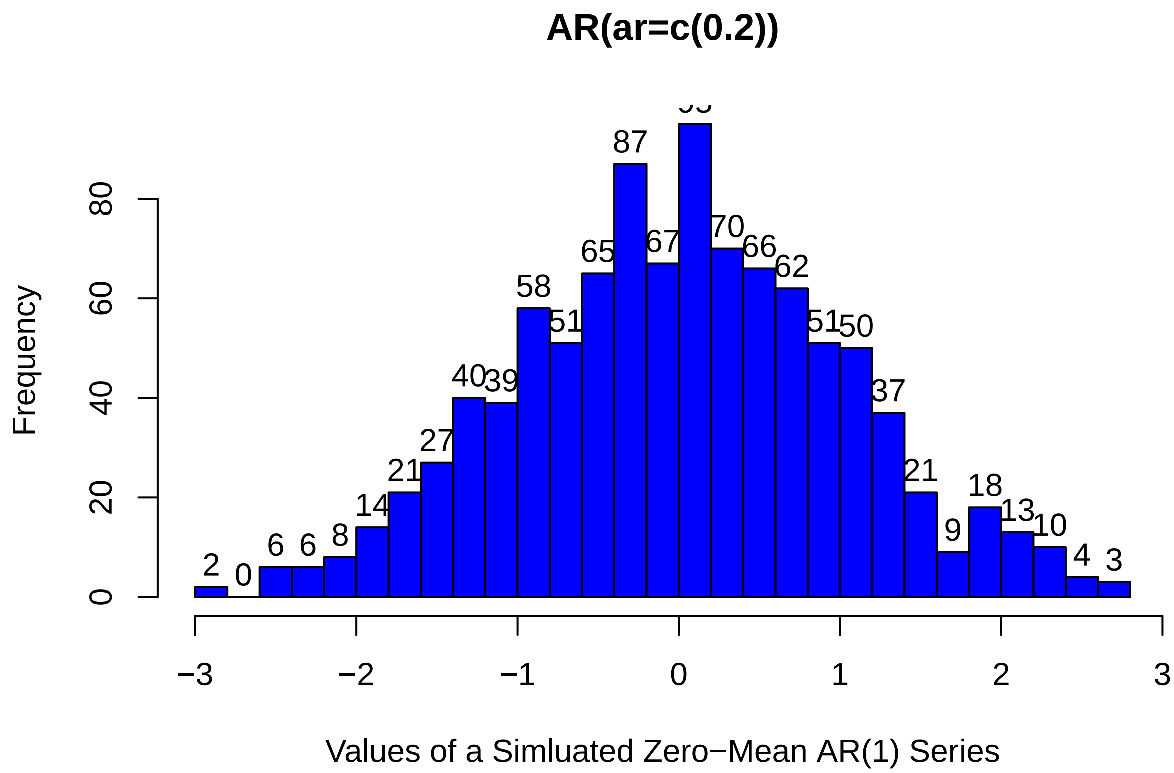
```
w=rnorm(1000,0,1)

z = w
for (t in 2:length(w)){
  z[t] <- 0.9*z[t-1] + w[t] # use the same random normal sequence generated above
}

hist(z, breaks="FD",
     main="AR(ar=c(0.9))",
     xlab="Values of a Simluated Zero-Mean AR(1) Series",
     col="blue", labels=TRUE)
```



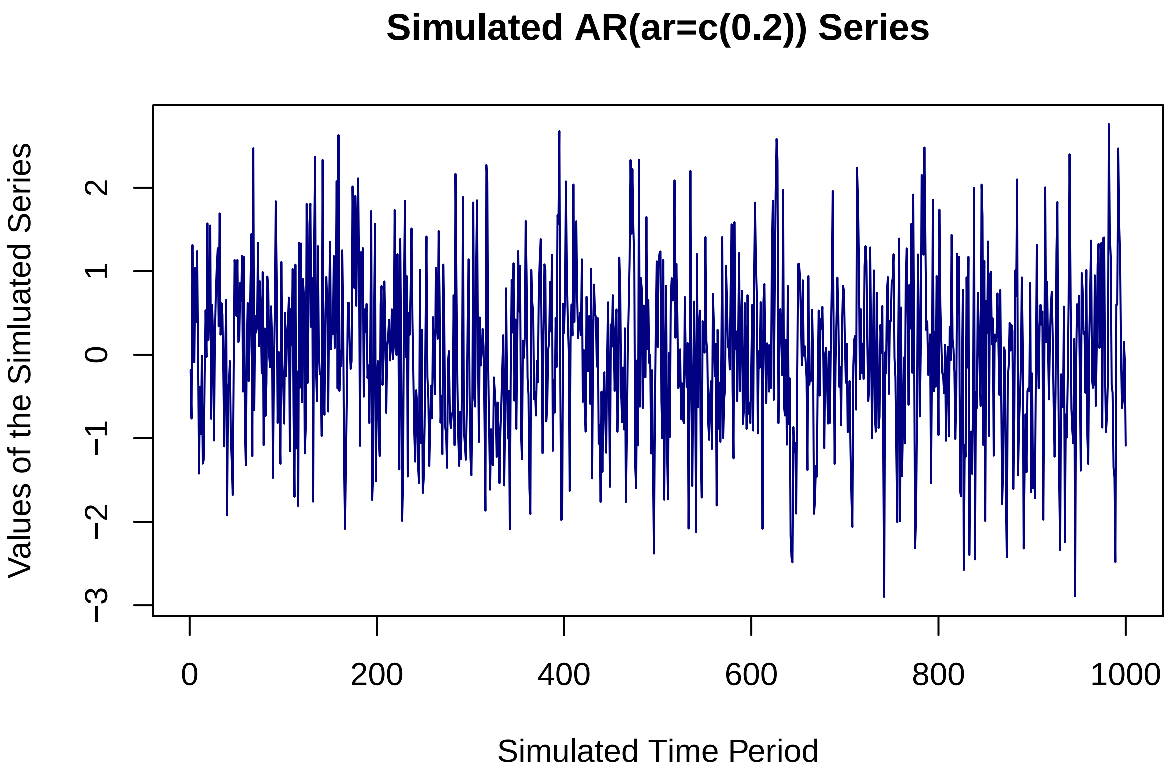

```
plot.ts(z, main="Simulated AR(ar=c(0.9)) Series", col="navy",
        ylab="Values of the Simluated Series",
        xlab="Simulated Time Period")
```




```
z = w
for (t in 2:length(w)){
  z[t] <- 0.2*z[t-1] + w[t] # use the same random normal sequence generated above
}

hist(z, breaks="FD",
     main="AR(ar=c(0.2))",
     xlab="Values of a Simluated Zero-Mean AR(1) Series",
     col="blue", labels=TRUE)
```




```
plot.ts(z, main="Simulated AR(ar=c(0.2)) Series", col="navy",
        ylab="Values of the Simluated Series",
        xlab="Simulated Time Period")
```



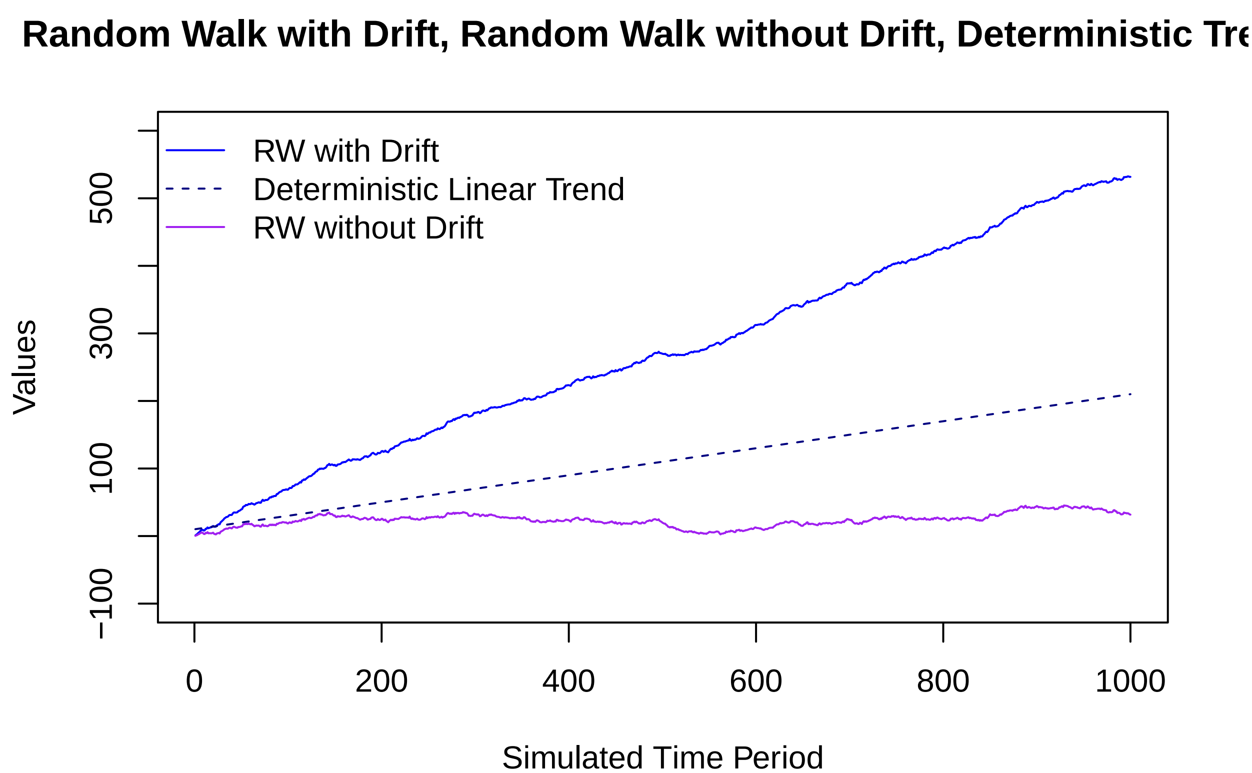
The AR series with C=0.9 is much smoother than that series with C=0.2 Both series have a variation around the mean 0, which is to be expeted since they are drawn from a standard normal distribution. The series with C=0.9 has higher variability than the one with C=0.2

Question 5:

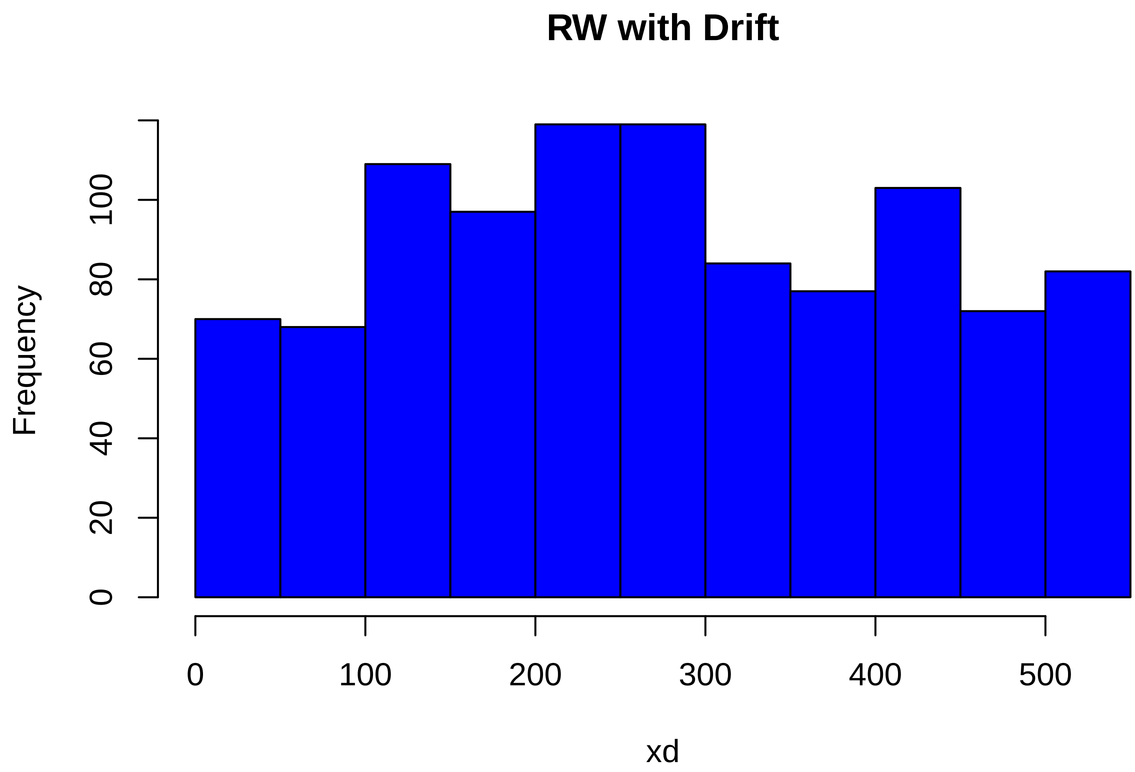
Simulate (with 1000 random draws) the following 3 models: 1. A deterministic linear (time) trend of the form: yt = 10 + 0.5t 2. Random walk without drift 2 3. Random walk with drift = 0.5 Plot a time plot for each of the simulated series. Graph a histogram for each of thes simulated series. Write a few statements to compare the two series.

```
w=rnorm(1000,0,1)

x=cumsum(w)

# Random walk with drift = 0.5
wd = 0.5 + w;
xd = cumsum(wd)


par(mfrow=c(1,1))
plot.ts(xd, main="Random Walk with Drift, Random Walk without Drift, Deterministic Trend",
        col="blue", ylab="Values", xlab="Simulated Time Period", ylim = c(-100,600), bg=38)
lines(0.2*(1:length(xd))+10, lty="dashed", col="navy")
lines(x, col="purple")
```

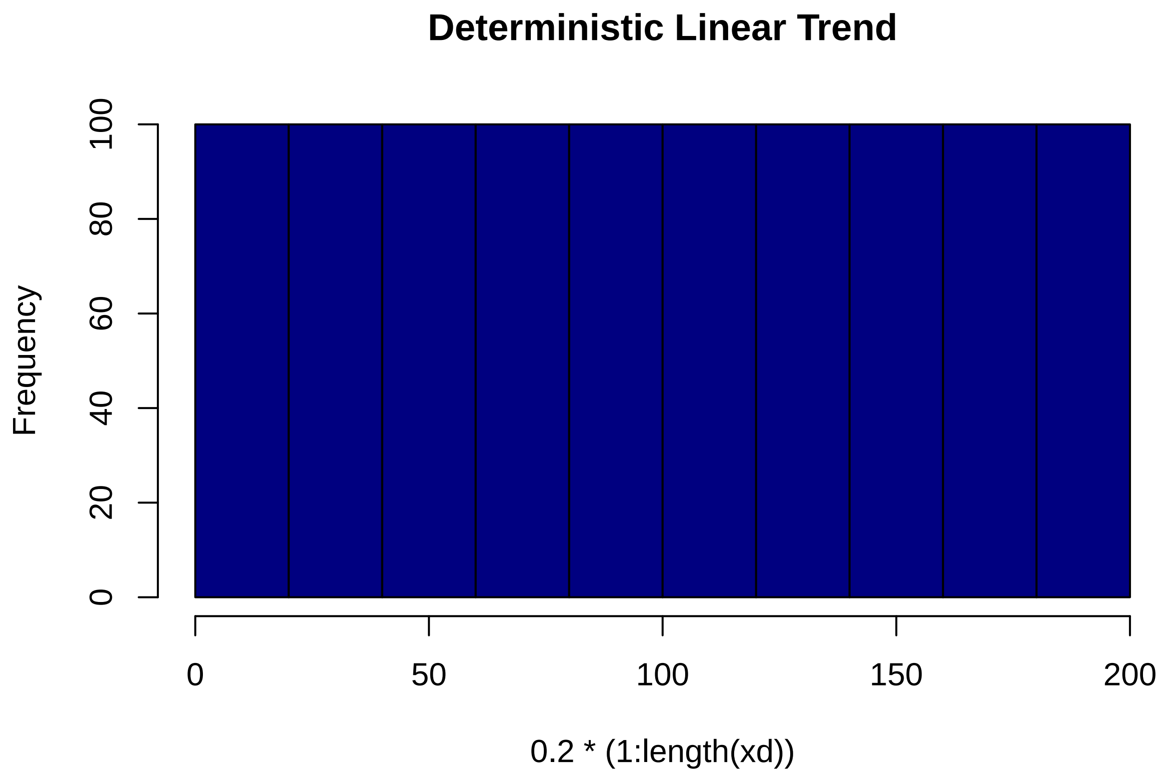
```r
# Add Legend
leg.txt <- c("RW with Drift", "Deterministic Linear Trend", "RW without Drift")
legend("topleft", legend=leg.txt, lty=c(1,2,1), col=c("blue","navy","purple"),
       bty='n', cex=1, merge = TRUE, bg=336)
```



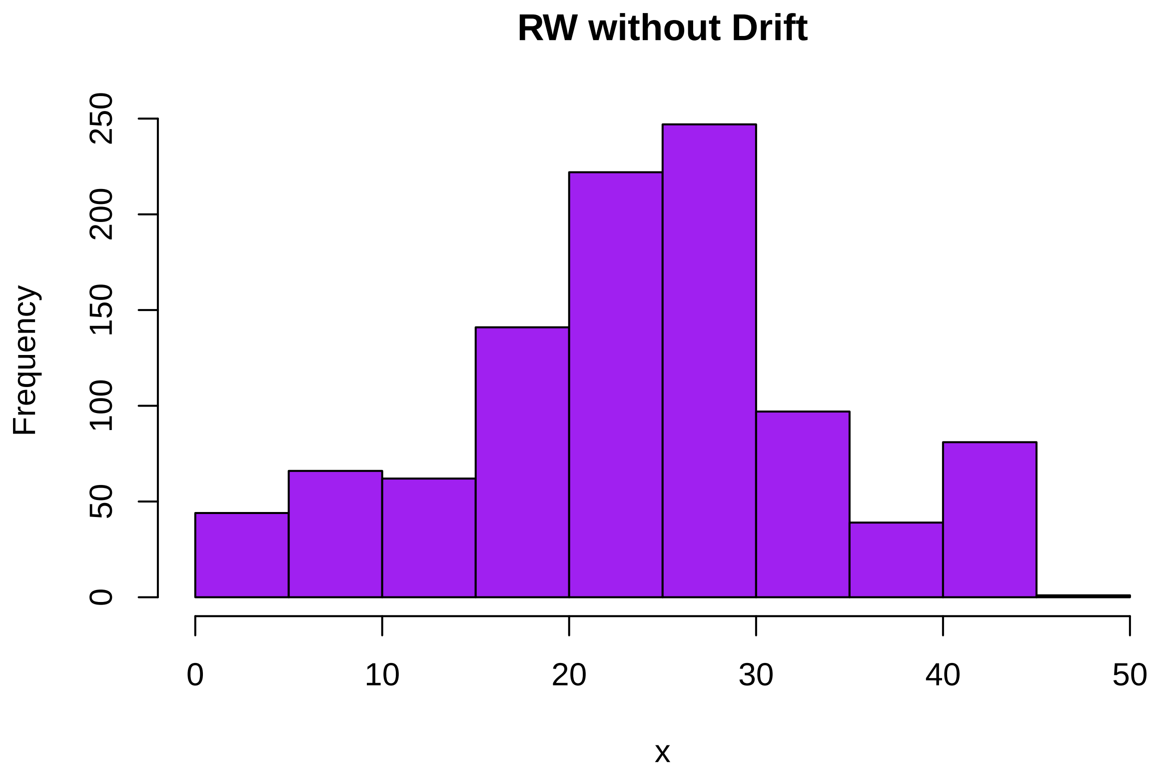

```r
hist(xd, main="RW with Drift", col="blue")
```




```
hist(0.2*(1:length(xd)), main="Deterministic Linear Trend", col="navy")
```




```r
hist(x, main="RW without Drift", col="purple")
```



The random walk with drift has a definite upward trend, this is due to the additon of the positive drift of 0.5

The random walk without drift averages out to 0, which is consistent with the standard normal distribution it is drawn from.

The deterministic linear trend is an expected straight line trending upwards.