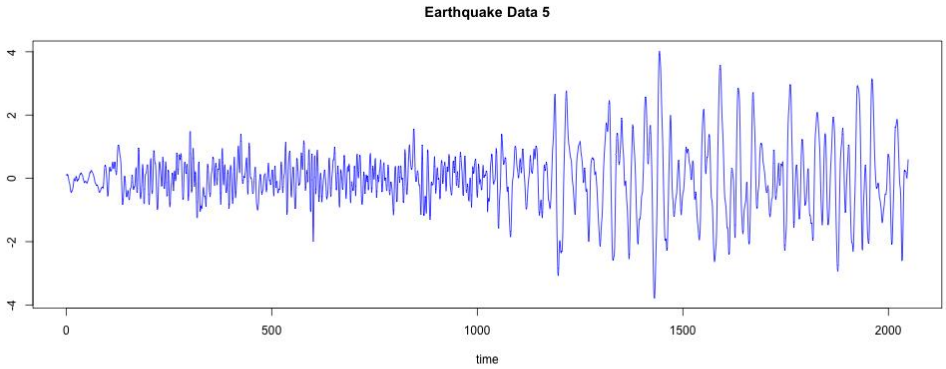
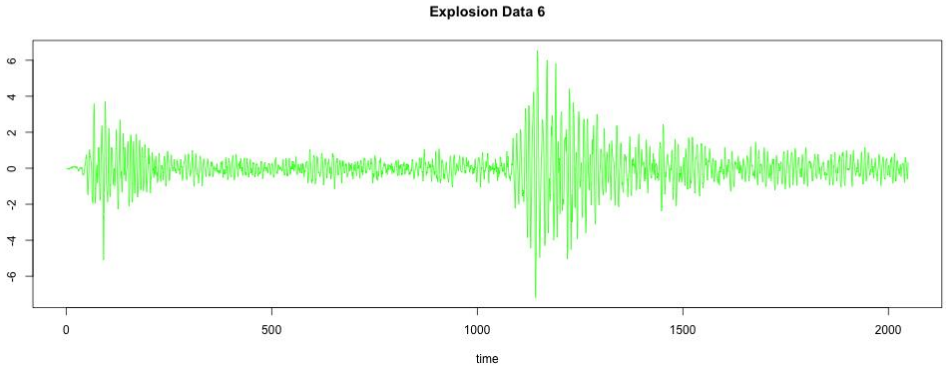
STAT 5350/7110 Assignment-#1

Name: Mohammed Raza Syed
PennID: 37486255
Degree: MSE Data Science (SEAS)

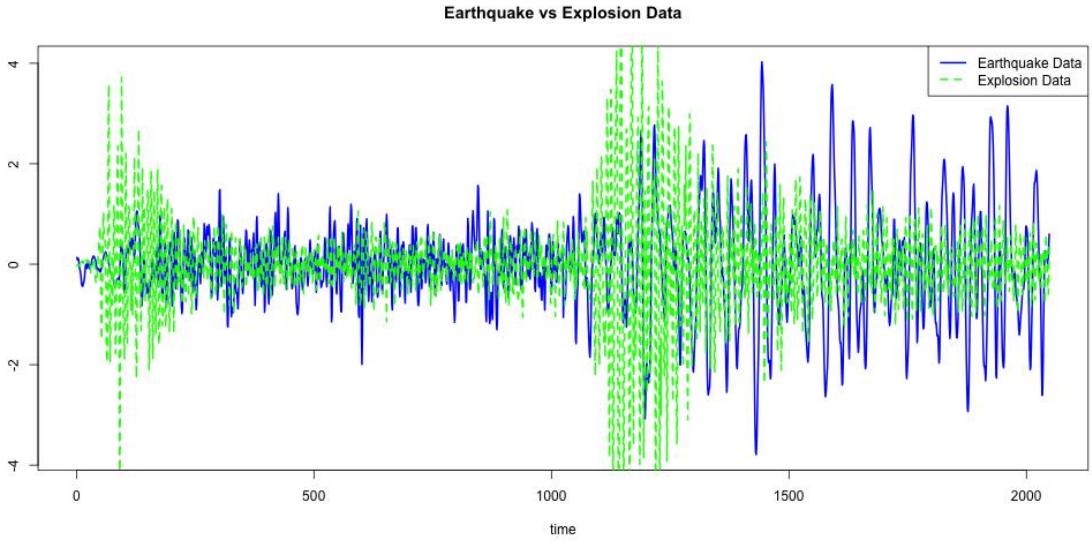
Q1)

a.





b.



c.

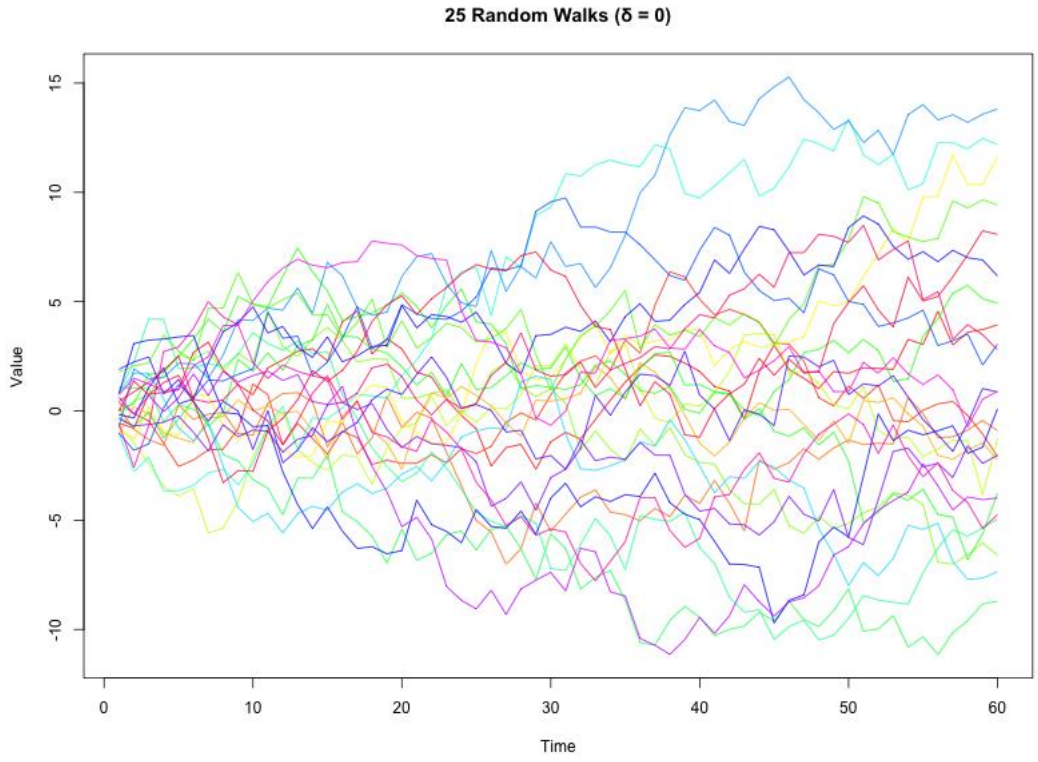
The earthquake and explosion series differ in the following ways:

i. With abrupt increases and progressively declining tendencies, the Earthquake series appears to be more stable than the Explosion series.

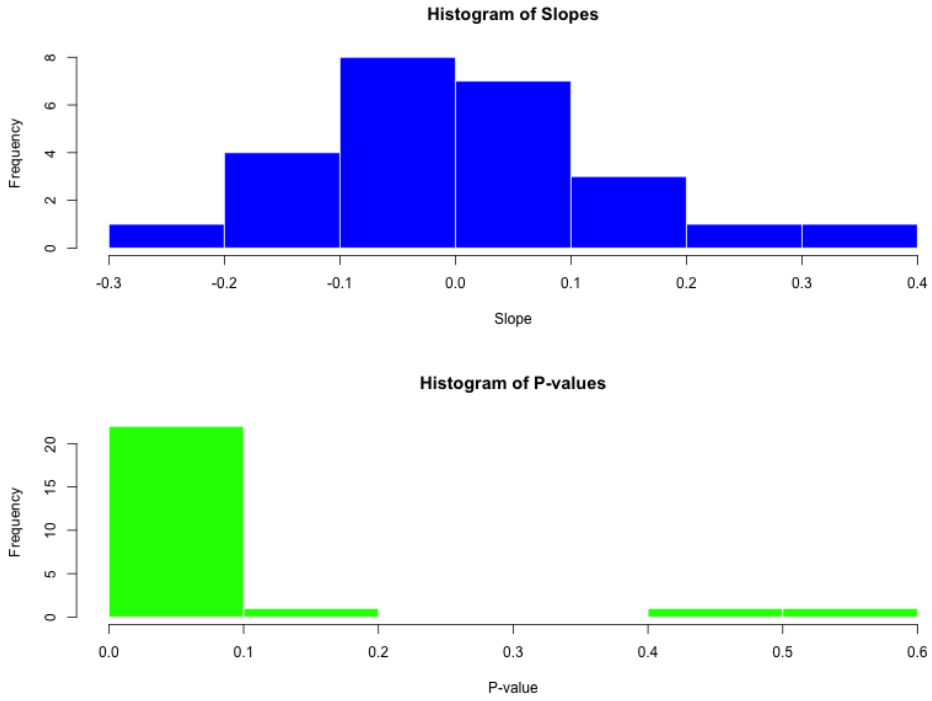
ii. Explosion reaches a peak of -6 to 6, whereas earthquake reaches a peak value of -4 to 4.

iii. The pattern of an earthquake series is characterized by constant variations, while an explosion series shows a quick decline following a spike.

Q3)

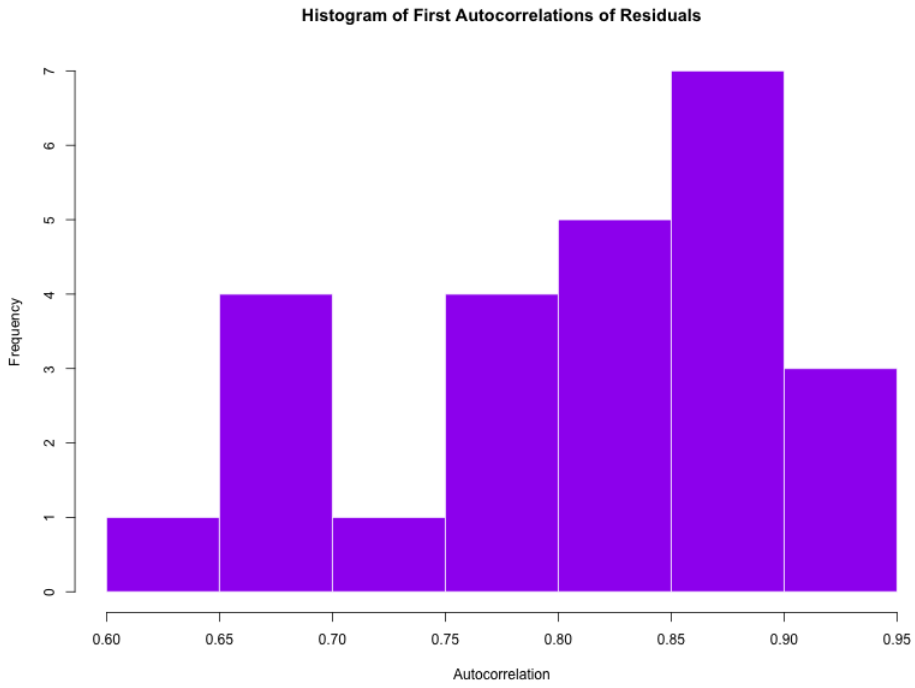


Q4)



The primary issue with determining a random walk's slope is non-stationarity of the data. This means that even in cases where there is no relationship, normal regression techniques may nevertheless reveal one (such as a significant slope). When p-values are deemed untrustworthy, inaccurate inferences regarding the slope's importance are made.

Q5)

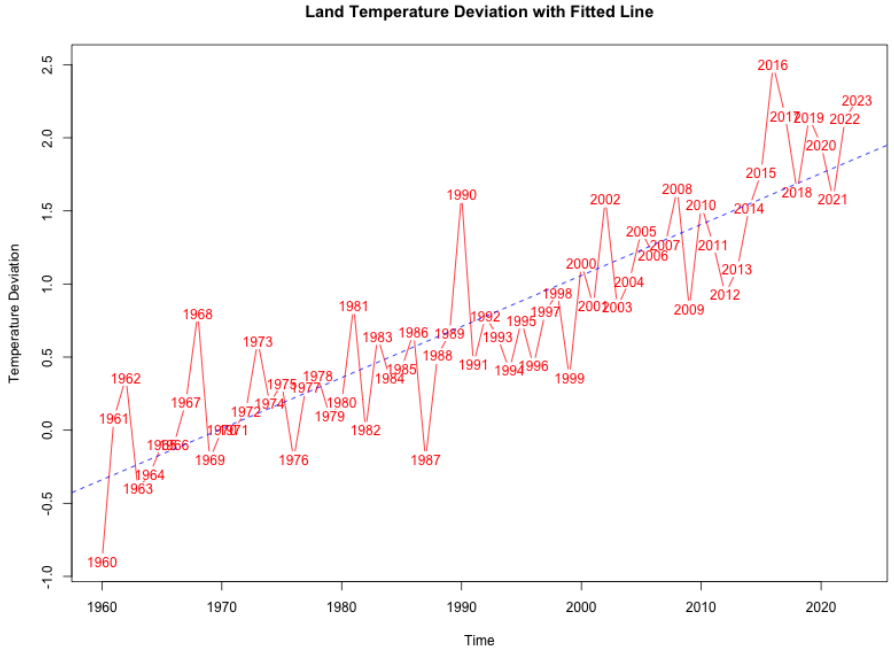


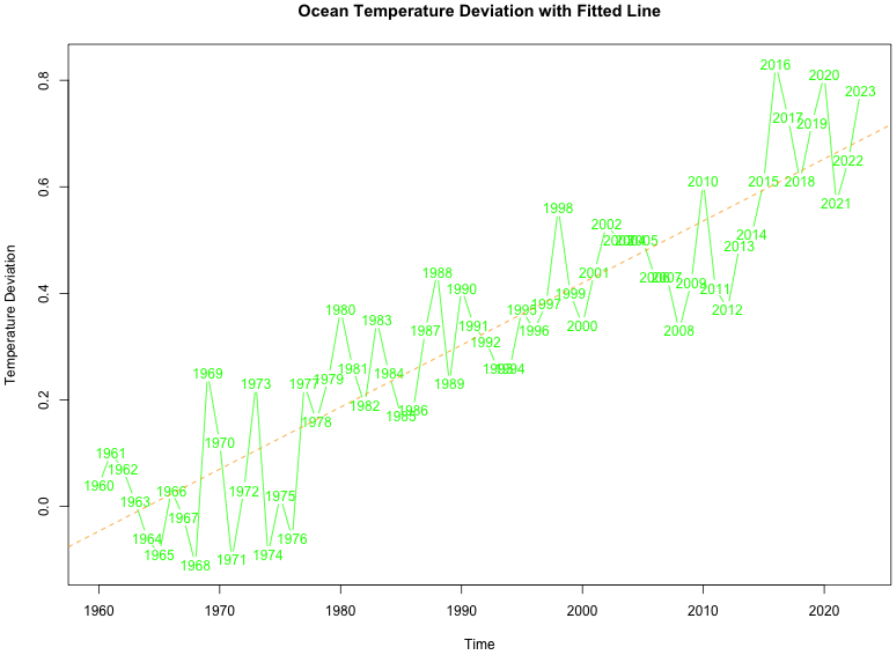
Q6)

```
> land_window <- window(gtemp_land, start=1960)
> ocean_window <- window(gtemp_ocean, start=1960)
> attributes(land_window)
$tsp
[1] 1960 2023    1

$class
[1] "ts"

> attributes(ocean_window)
$tsp
[1] 1960 2023    1

$class
[1] "ts"

> land_index <- gtemp_land[111:174]
> ocean_index <- gtemp_ocean[111:174]
> attributes(land_index)
NULL
> attributes(ocean_index)
NULL
```

Thus, it can be inferred from a comparison of the windows and index functions that the windows function extracts all the required information, including class, date, frequency, etc. The index function, on the other hand, just extracts the vectors without any regard to timing.

Q7)

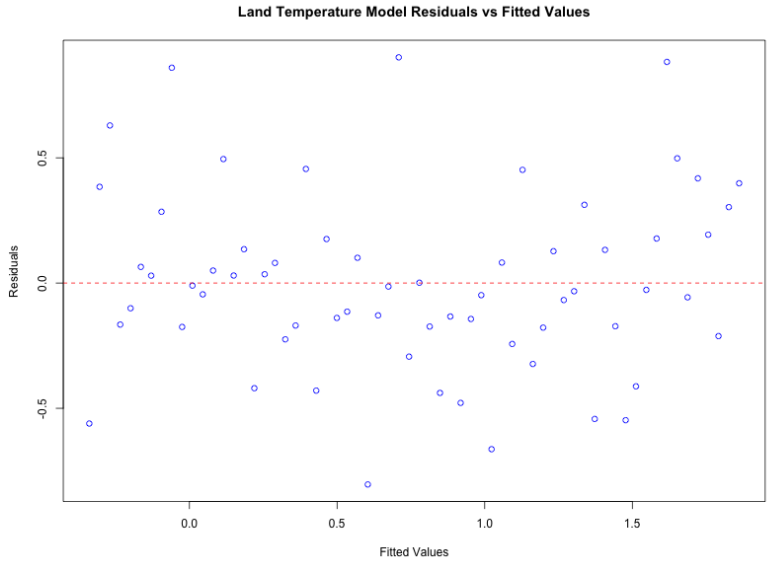

Land Temperature Deviation with Fitted Line
Temperature Deviation
Time
-1.0
-0.5
0.0
0.5
1.0
1.5
2.0
2.5
1960
1970
1980
1990
2000
2010
2020

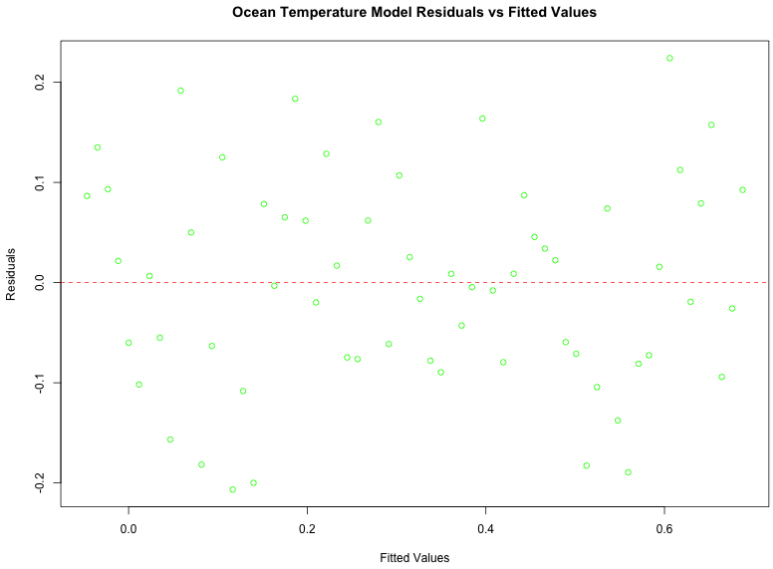

Ocean Temperature Deviation with Fitted Line
Temperature Deviation
Time
0.0
0.2
0.4
0.6
0.8
1960
1970
1980
1990
2000
2010
2020

Q8) First, there has been an increasing temperature pattern or trend throughout the years, with occasional brief swings but again increasing over time, according to observations of both the land and ocean temperature plots.

Second, as the p-values for the land and ocean plots are extremely small (<2e-16), indicating a considerable increase in temperature, the estimated slopes are significantly different from zero.

Q9) The residuals plot suggests that there isn't a clear flaw in the model. The fact that the residuals are uniformly distributed about zero indicates that the linear regression models for both ocean and land temperatures correctly represent the trend in the data, without of systematic bias and inconsistent error patterns.





Q10) The residuals for land and ocean temperatures show a weakly positive association ($r = 0.28$). The residuals from the two models appear to be mostly independent, and the Cross Correlation plot indicates that no cross-correlation values surpass the 95% significance threshold, indicating that the association is not statistically significant. Hence they are independent of each other.