Time Series Analysis

Course Project

- Sireesh Limbu

May 2019


WINTER IS COMING
STARK


FAMILY DUTY HONOR
TULLY


OURS IS THE FURY
BARATHEON


AS HIGH AS HONOR
ARRYN


HEAR ME ROAR
LANNISTER


GROWING STRONG
TYRELL


GAME OF THRONES


FIRE AND BLOOD
TARGARYEN


HERE WE STAND
MORMONT


WE DO NOT SOW
GREYJOY


SWORN TO STARK
BOLTON


FREY


UNBOWED UNBENT UNBROKEN
MARTELL

Introduction

Based on George R.R. Martin's best-selling book series "A Song of Ice and Fire" - Game of Thrones is an American fantasy-fiction television drama series created by David Benioff and D.B. Weiss for the popular satellite television and cable network HBO.

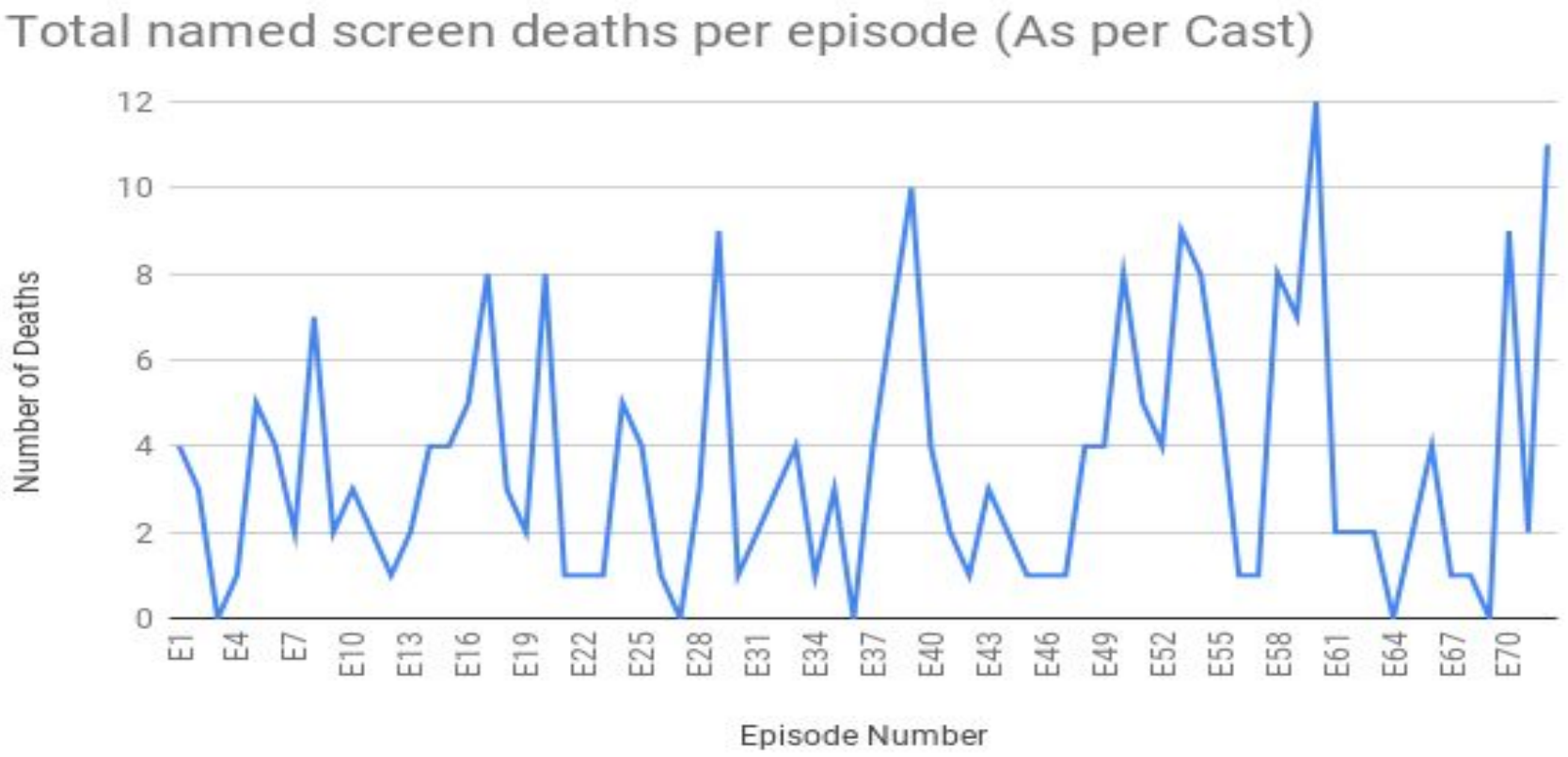
Filmed in Belfast, US,Canada and scenic parts of Europe based out of countries like Croatia, Morocco, Malta,Scotland, Northern Ireland, Iceland and Spain, the whole series has been concluded in eight seasons with a total of 72 episodes with the last episode to be aired on Monday - 20th May in Sweden.

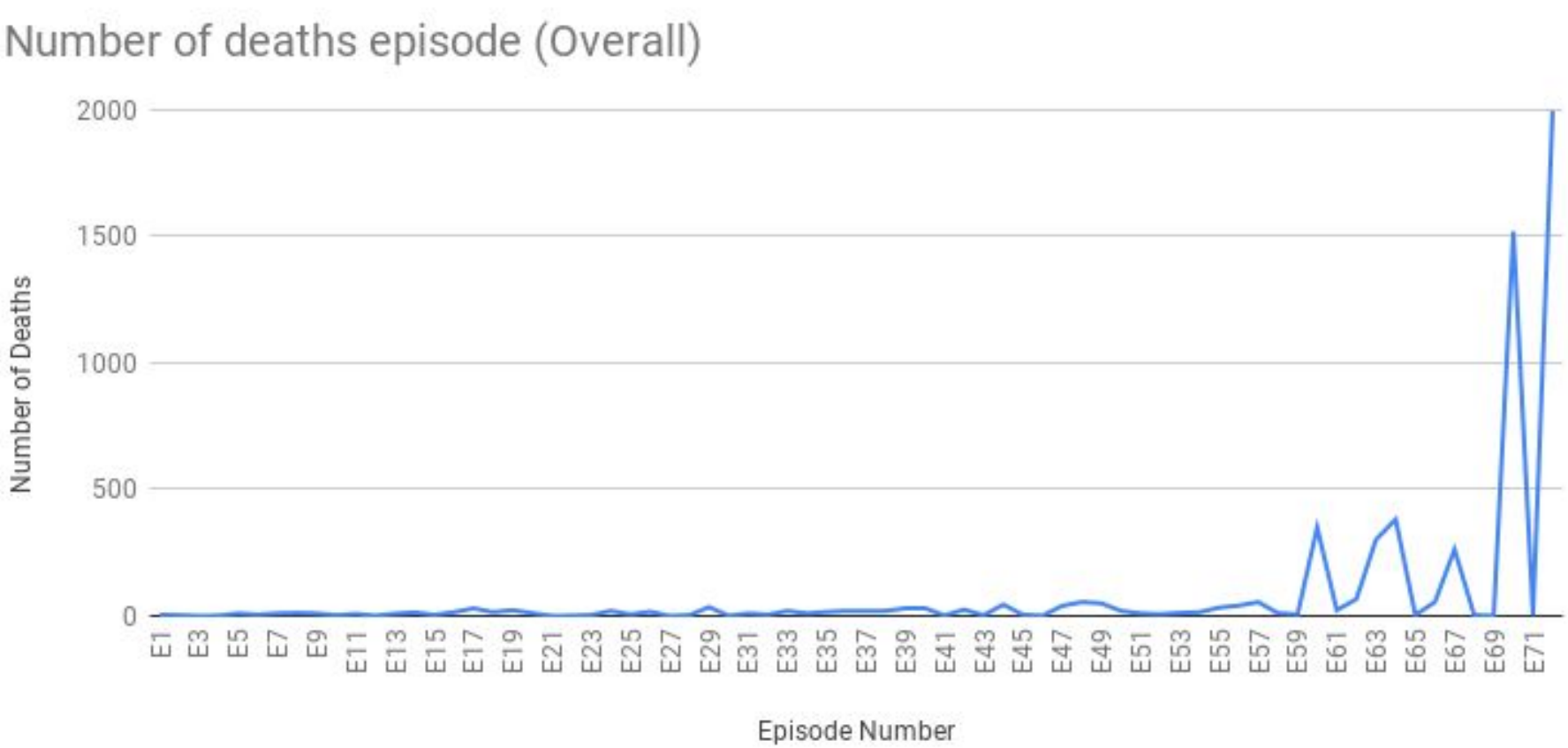
Goal

Game of Thrones, apart from being famous for various ingenious factors such as story-plot, CGI and well-drawn characters is also known for it's violence in the form of battles, wars or just random merciless killings. This has amounted to an average of 81 killings per episode and a total of 5863 on-screen killings spanning over 8 seasons until the second-last episode. The goal of this project is to analyze the data of total on-screen deaths per episode for both 'Named' and 'Overall' characters and eventually forecast the total deaths in the last episode.

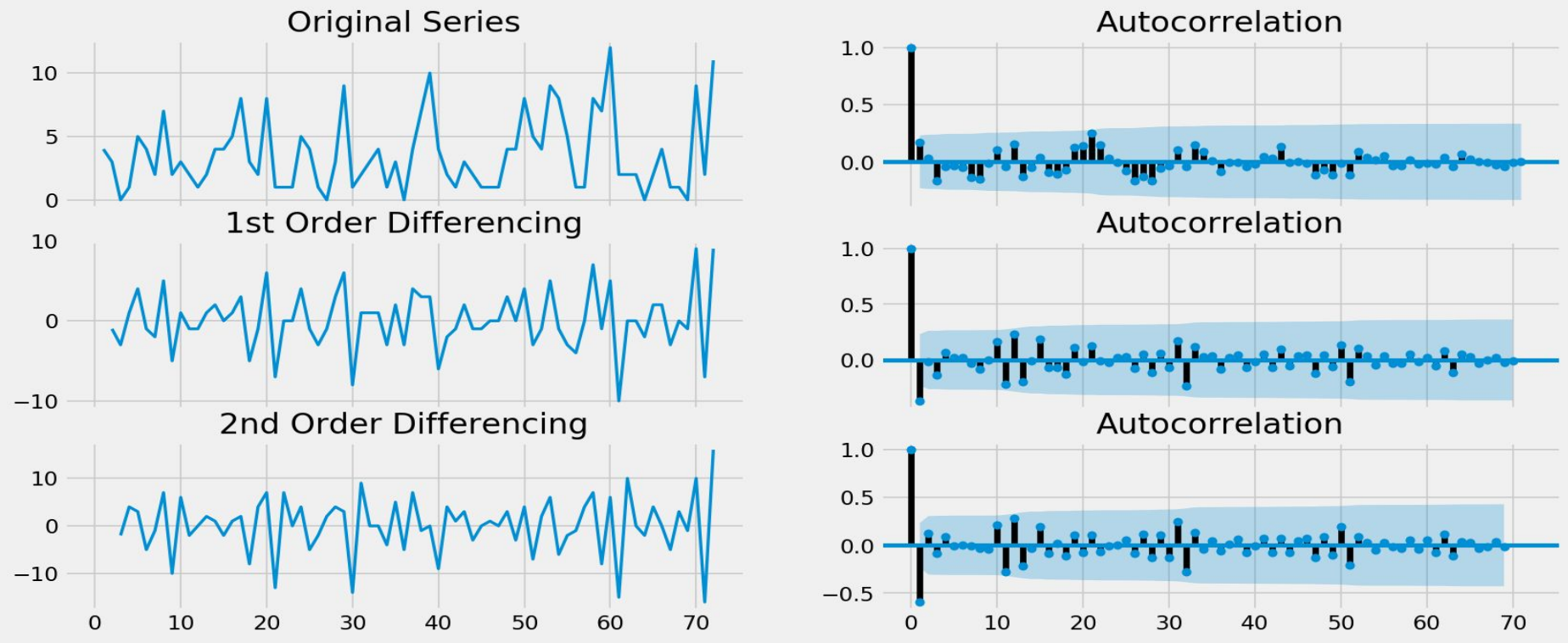
Data Collection and Initial Analysis

The total deaths statistics survey from Washington post and fandom wiki solely dedicated to Game of thrones has been the two major online sources of data for this project. The collection of total deaths for both named and un-named characters together per episode due to lack of reliable sources has been modified to add up to total deaths per season which in turn has been verified by the Washington Post survey. But, the fandom website provides reliable source of the number of on-screen deaths of named characters per episode. Due to this, the project has been divided into two parts: first to analyze, fit the model and forecast deaths for named characters and finally do the same for all characters. Following are the Number of Deaths Vs Episode Number plots for overall and 'Named' and 'Overall' deaths.


Total named screen deaths per episode (As per Cast)
Number of Deaths
0
2
4
6
8
10
12
E1
E4
E7
E10
E13
E16
E19
E22
E25
E28
E31
E34
E37
E40
E43
E46
E49
E52
E55
E58
E61
E64
E67
E70
Episode Number


Number of deaths episode (Overall)
2000
1500
1000
500
0
Number of Deaths
E1
E3
E5
E7
E9
E11
E13
E15
E17
E19
E21
E23
E25
E27
E29
E31
E33
E35
E37
E39
E41
E43
E45
E47
E49
E51
E53
E55
E57
E59
E61
E63
E65
E67
E69
E71
Episode Number

Analysis (Named Characters)



Analysis (Named Characters)

As no proper stationarity could be analyzed from the plots we perform the Dickey Fuller test for stationarity to the named character deaths data and get the following output:-

In [81]:
```
#Dickey-Fuller Test for Stationarity
from statsmodels.tsa.stattools import adfuller
adf_test = adfuller(df.value)

print ("ADF = " + str(adf_test[0]))
print ("p-value = " +str(adf_test[1]))
```

```
ADF = -6.518858048374907
p-value = 1.0544759906810726e-08
```

The small p value suggests that the series is indeed stationary. Hence, no further differencing required.

Model Fitting and Forecasting (Named Characters)

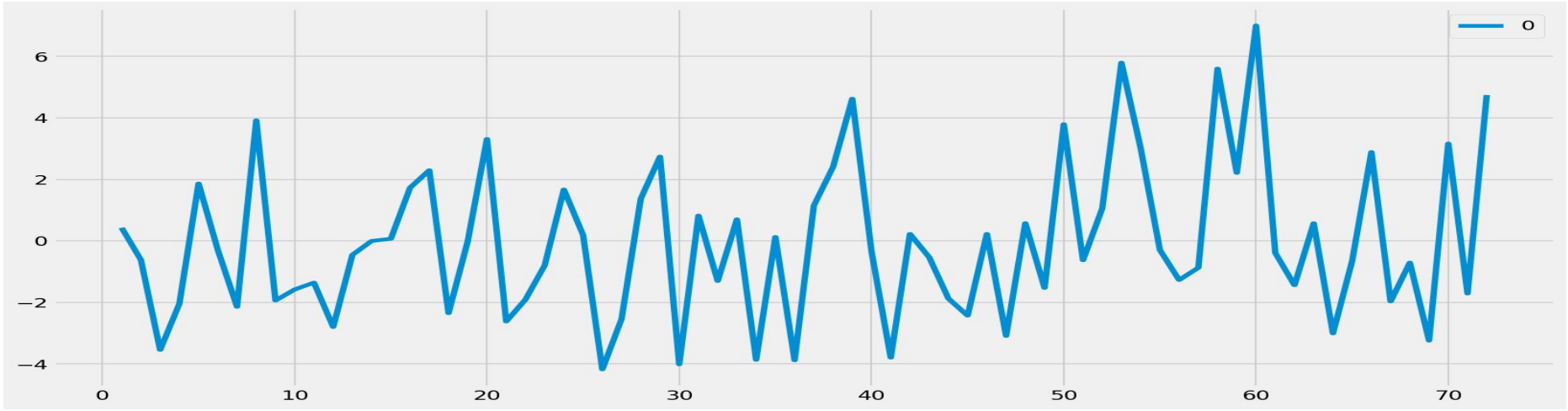
As the available data for this forecasting is limited we use a simple AR model by calling the ARMA model with zero moving-average window and 14 lag order for auto-regression for analyzing. A simple model fitting and forecast gives us the following output:-

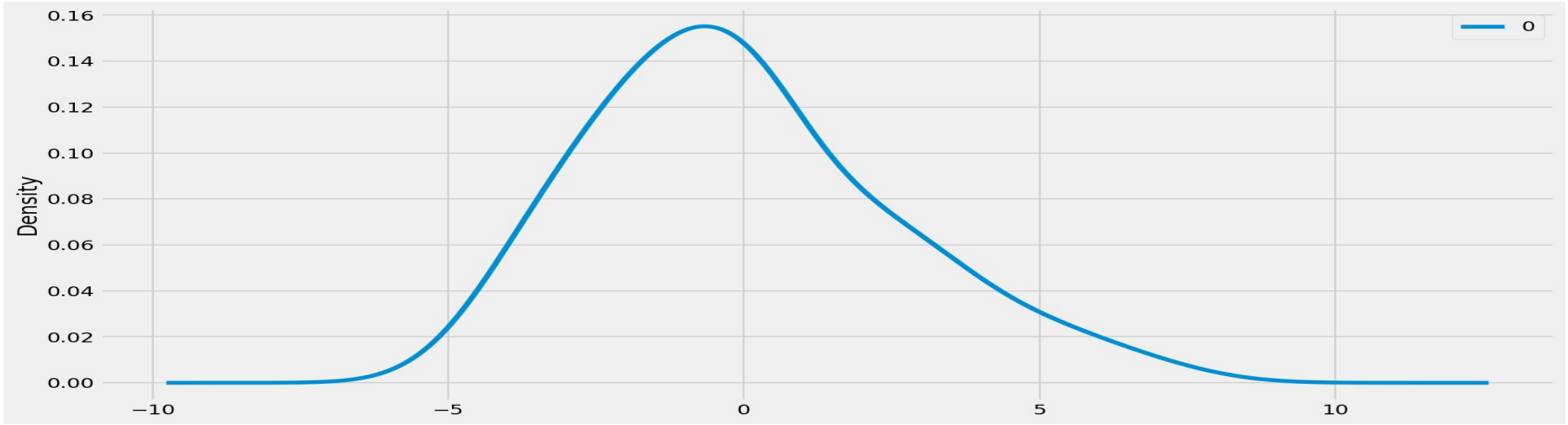
In [30]:
```
from statsmodels.tsa.arima_model import ARMA
# fit model
model2 = ARMA(df, order=(14, 0))
model_fit2 = model2.fit(disp=False)
# make prediction
yhat2 = model_fit2.predict(len(ov), len(ov))
print(yhat2)
```

```
/home/limbu/anaconda3/lib/python3.7/site-packages/statsmodels/tsa/base/tsa_model.py:221: ValueWarning: An unsuppor
ted index was provided and will be ignored when e.g. forecasting.
  ' ignored when e.g. forecasting.', ValueWarning)
```

Out[30]:
```
array([1.57700512])
```

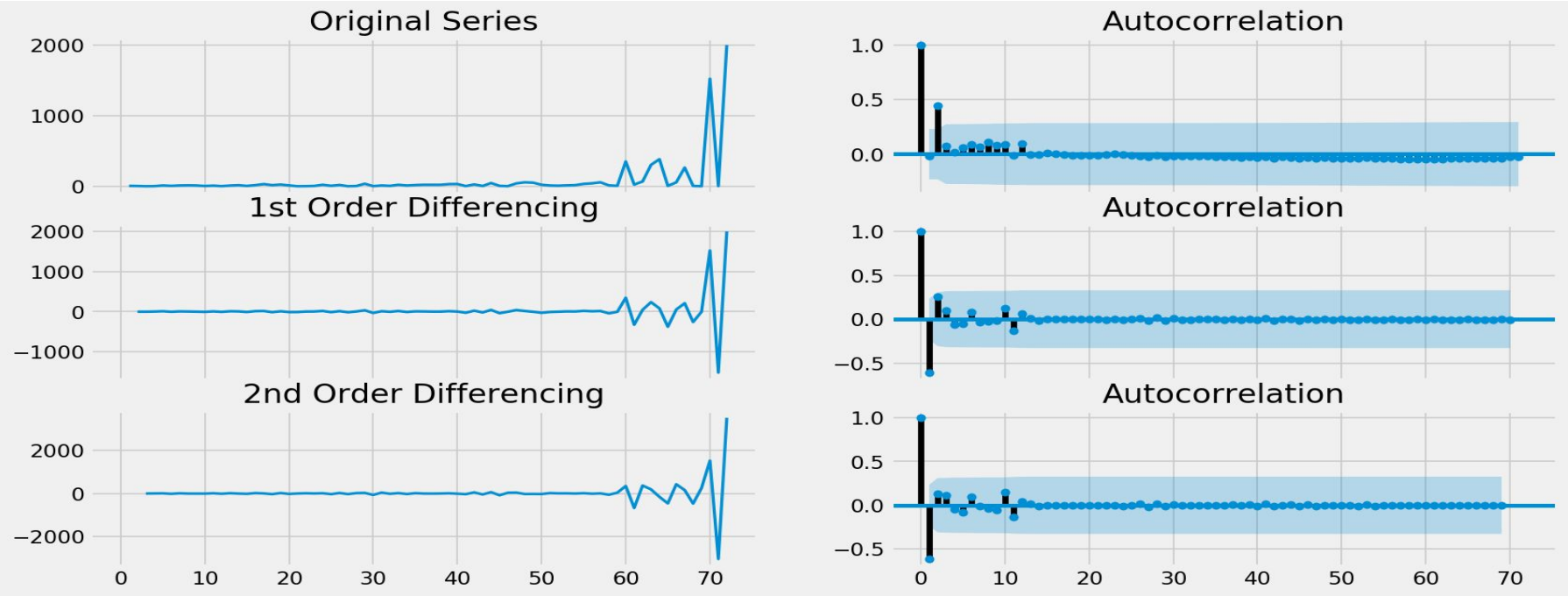
Which is approximately = 2 Deaths


0
6
4
2
0
−2
−4
0
10
20
30
40
50
60
70


0
0.16
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
Density
−10
−5
0
5
10

Analysis (Overall Deaths)

We perform a similar analysis for all on-screen deaths for each episode.



Analysis (Overall Deaths)

Clearly there is no visible seasonality or trend. We go ahead with the Dickey Fuller test for stationarity and get the following output:-

In [5]:
```
#Dickey-Fuller Test for Stationarity
from statsmodels.tsa.stattools import adfuller

adf_test = adfuller(ov.value)

print ("ADF = " + str(adf_test[0]))
print ("p-value = " +str(adf_test[1]))
```

```
ADF = 5.217848071894859
p-value = 1.0
```

The large p-value indicates that the data as suspected is non-stationary.

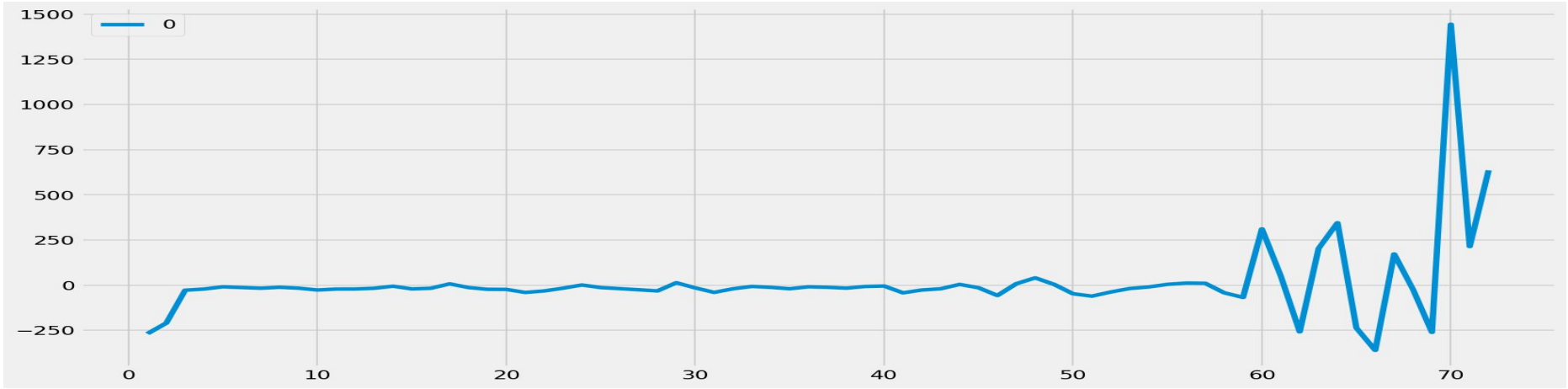
Model Fitting and Forecasting (All Characters)

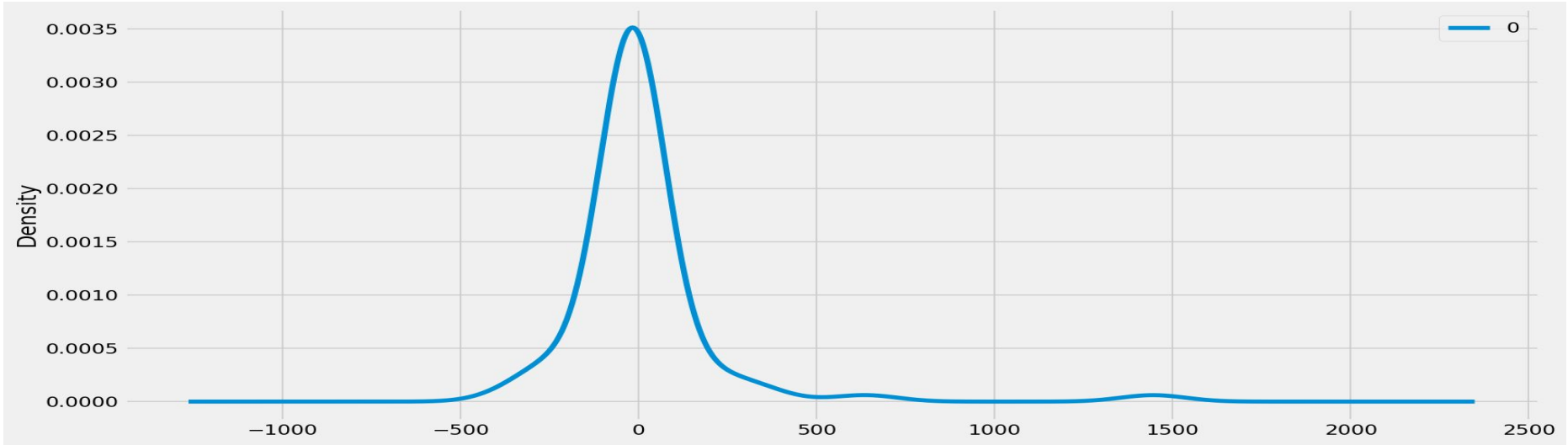
We again use the AR model by calling ARMA and plugging the Moving Average window to zero and 3rd order lag to fit and forecast the non-stationary data. We get the following forecast for the last episode:-

In [91]:
```
from statsmodels.tsa.arima_model import ARMA
# fit model
model2 = ARMA(ov, order=(3, 0))
model_fit2 = model2.fit(disp=False)
# make prediction
yhat2 = model_fit2.predict(len(ov), len(ov))
print(yhat2)
```

```
/home/limbu/anaconda3/lib/python3.7/site-packages/statsmodels/tsa/base/tsa_model.py:221: ValueWarning: An unsupported index was provided and will be ignored when e.g. forecasting.
  ' ignored when e.g. forecasting.', ValueWarning)
```

```
72    20.615011
dtype: float64
```


0
1500
1250
1000
750
500
250
0
−250
0
10
20
30
40
50
60
70


0
Density
0.0035
0.0030
0.0025
0.0020
0.0015
0.0010
0.0005
0.0000
−1000
−500
0
500
1000
1500
2000
2500

Result

After choosing simple AR models for both the data of on-screen deaths we conclude a prediction of around **2** ‘Named’ character deaths and **20** ’Overall’ deaths in the final episode of Game of Thrones.

When your aunt gets you the socks you wanted for xmas

VALAR

MORGHULIS