Plastic waste inputs from Land into the Ocean

Discussion Points

1. Introduction
2. Problem statement
3. About the Data
4. Exploratory Data Analysis
5. Data Preprocessing
6. Data Analysis
7. Solution
8. Time series analysis
9. Results
10. Conclusion

Introduction

- **Plastic** is something we come across on a daily basis, from the bottled water we drink, the packaged food we eat to the computers we work with, it's a major part of our lives.
- Most of the plastic is used only once and thrown away, and the **inefficiently managed** plastic end up in the **ocean.**
- Currently **150 million metric tons** of plastic circulate our oceans, on top of which **8 million metric tons** is added every year.
- **80%** of the plastic in ocean comes from **land.**

Problem statement

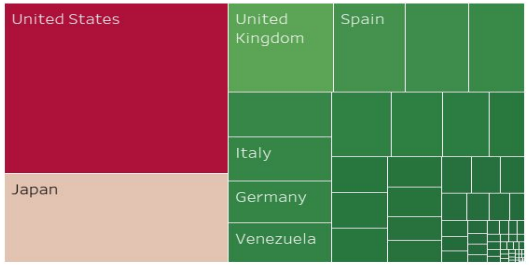
The aim of our project is to determine the factors responsible for generating tons of mismanaged plastic on land that eventually end up in the ocean and also to determine the global plastic production for coming years.

About the Data

- The dataset for **‘Mismanaged plastic estimation’** has been taken from Jambeck et al.
- The data is comprised of variables like **Economic status, Coastal population**(within 50km of coast), **plastic waste generation**, plastic waste in stream, plastic waste littered, inadequately managed plastic waste and **mismanaged plastic waste.**
- The data is categorized by country.
- The dataset for **‘Global plastic production’** has been take from Our world in data.
- The data gives the global plastic production by year.

Exploratory Data Analysis (Economic status)

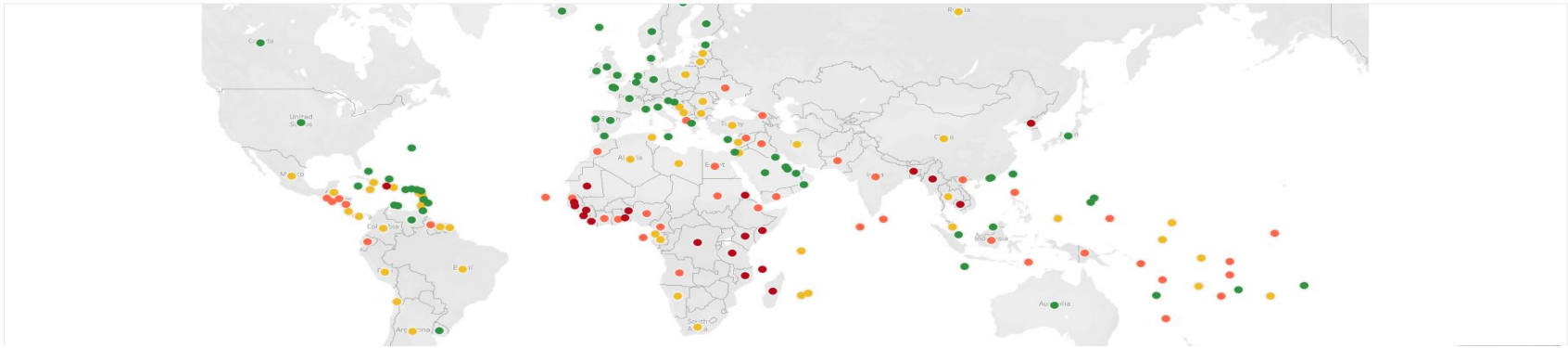
Plastic Waste generation in High Economic Countries



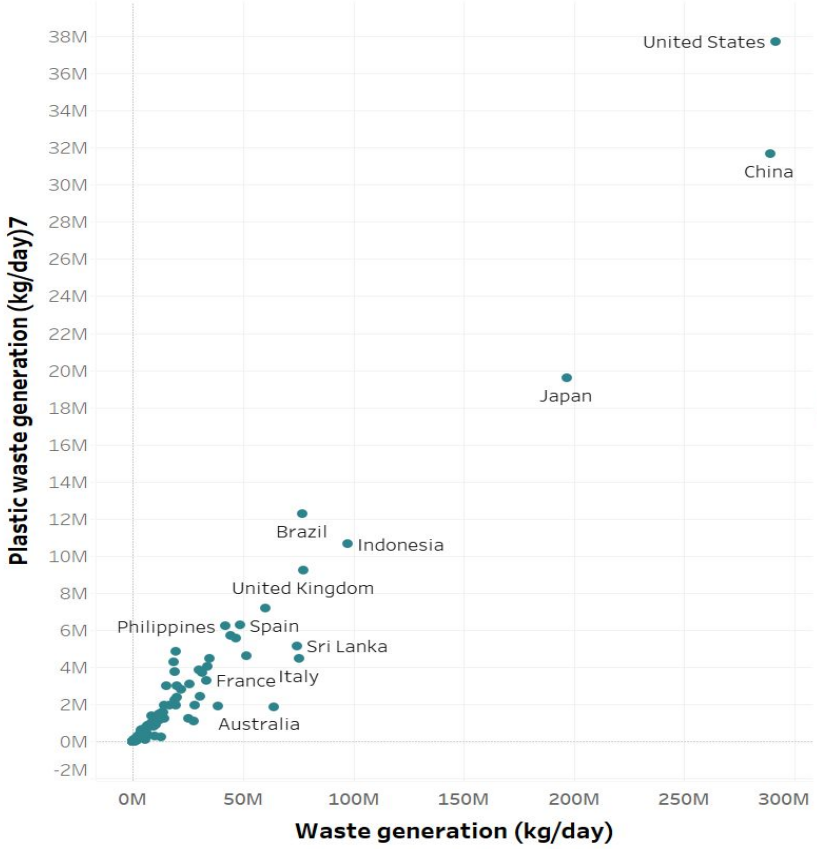
Plastic waste in comparison with number of countries

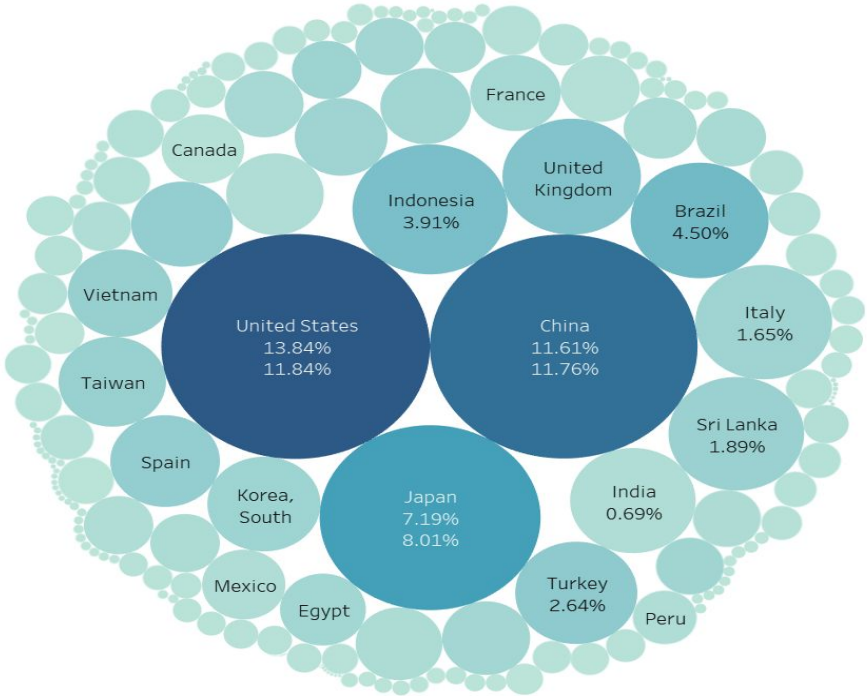
Count of countries

Economic Status of the Countries

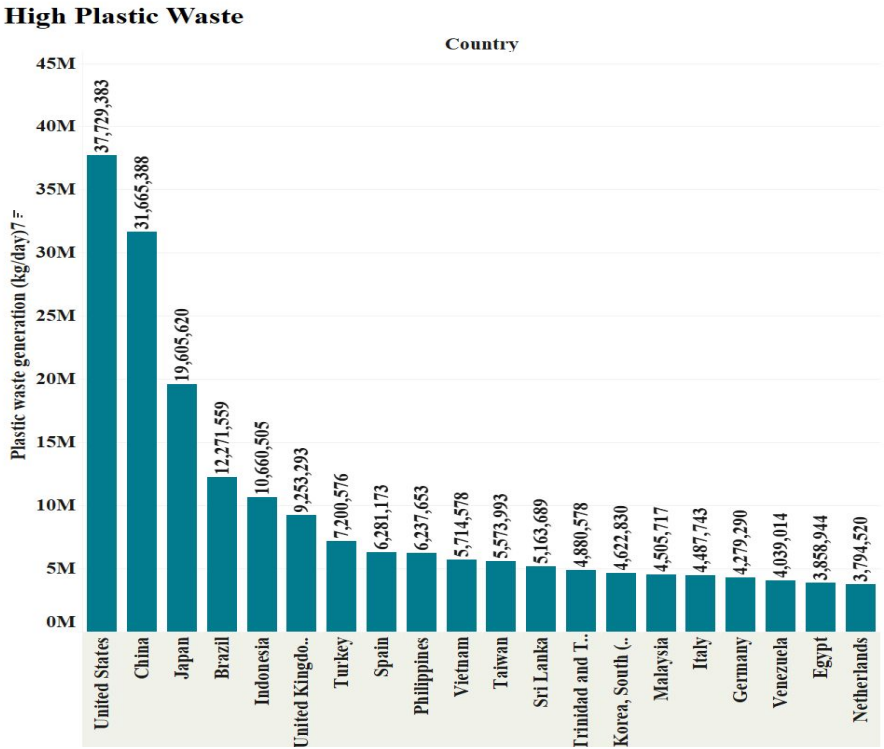


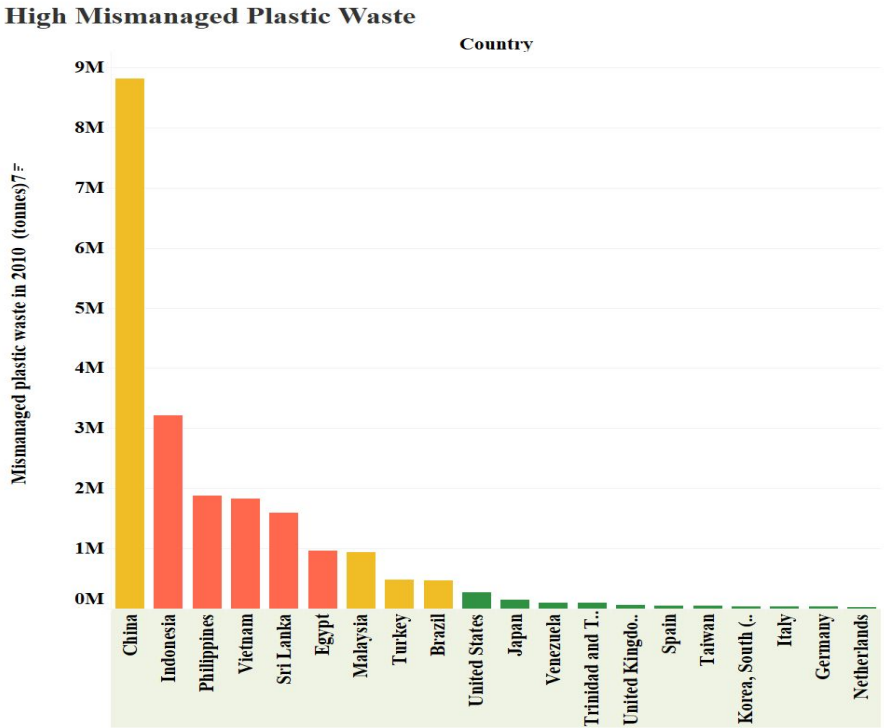
Waste Generation vs Plastic Waste Generation



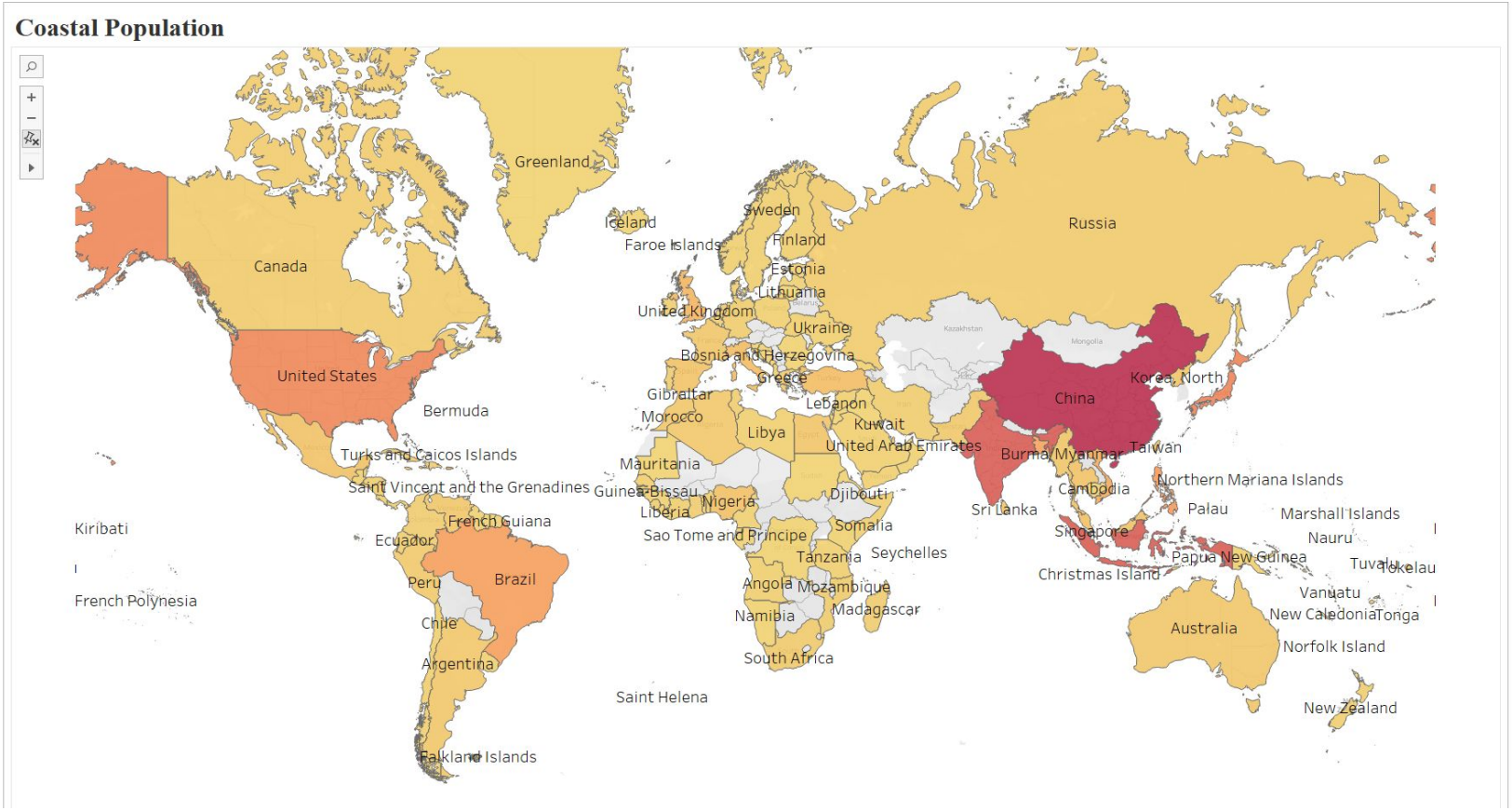


Exploratory Data Analysis (Comparison Between Total & Mismanaged plastic)

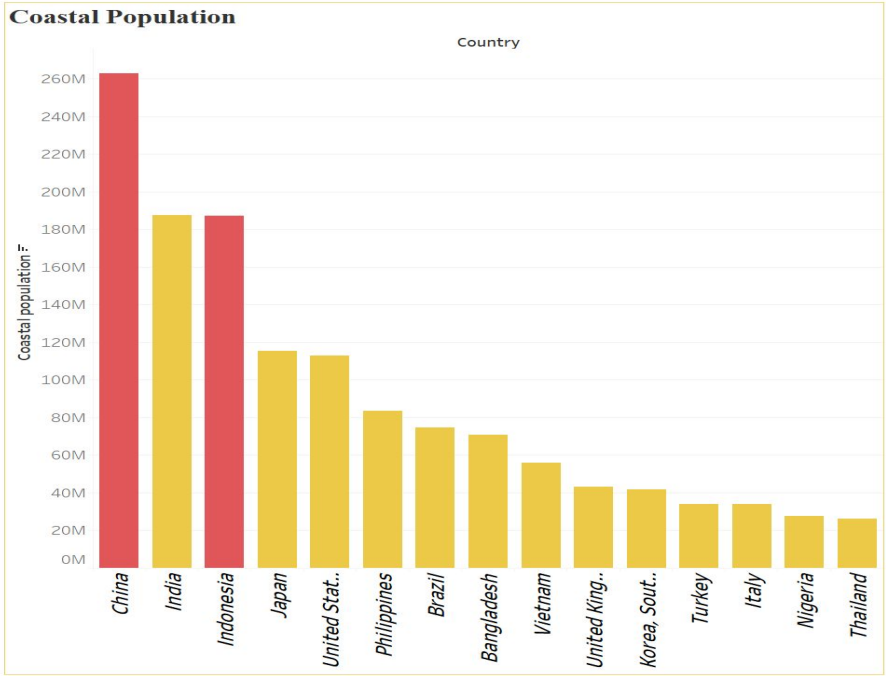


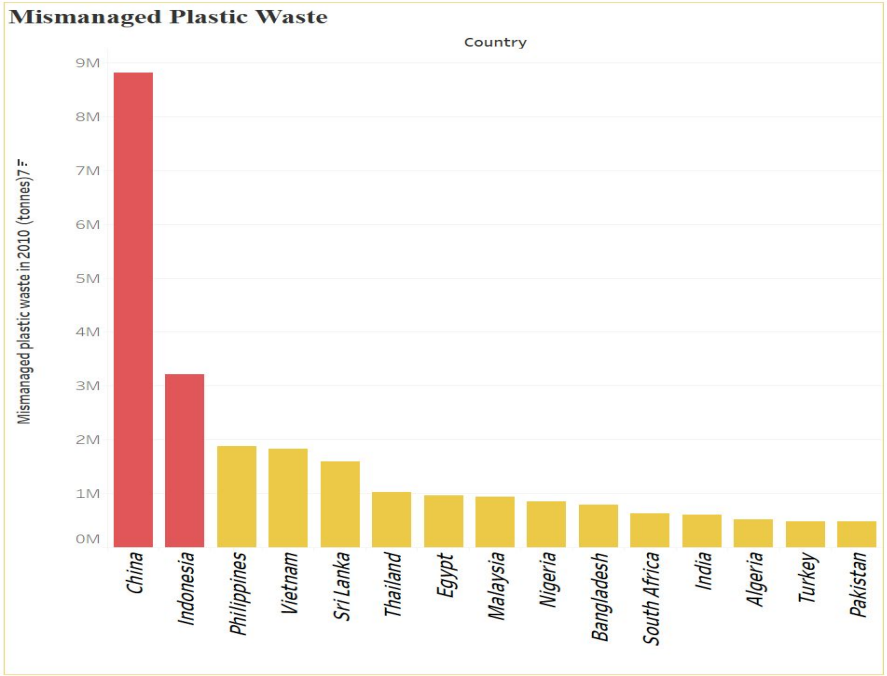


Exploratory Data Analysis (Coastal Population)



Coastal population





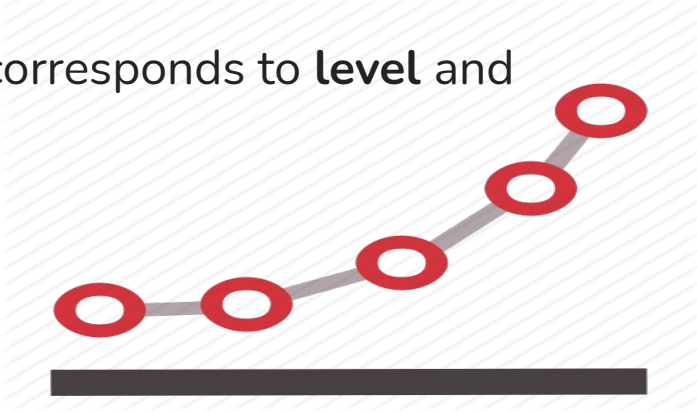
Solution

- While **enhancing the waste management system** is a long term solution, **cutting down on the plastic production** can largely help us handle the current situation effectively.
- As Erin Simon, the director of world wildlife organization says referring to the plastic waste, “**When the sink is flooding, you don’t start with the mop; you start by turning off the tap.”**

Time Series Analysis on Global plastic production using,

Holt's Exponential Smoothing method

- For forecasting, the **exponential smoothing** method gives more weight to the recent observations and lesser on older observations.
- **Holt's method** applies the exponential smoothing method capturing the trend of the data.
- It has two smoothing parameters, α and β, which corresponds to **level** and **trend** respectively

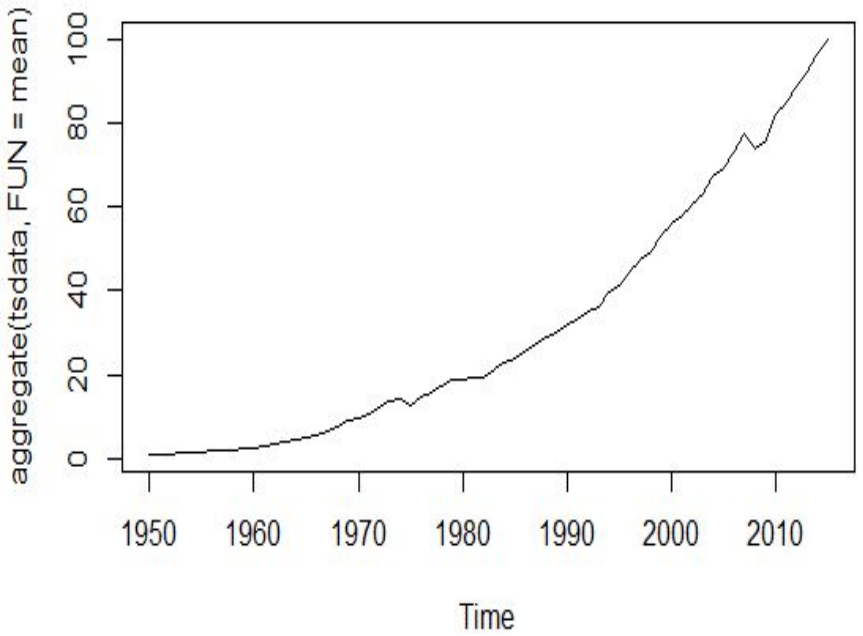


1. Data collection

```
> data <- global.plastics.production
> head(data)
  Entity     Code Year Global.plastics.production..million.tonnes...tonnes.
1  World OWID_WRL 1950                                              2000000
2  World OWID_WRL 1951                                              2000000
3  World OWID_WRL 1952                                              2000000
4  World OWID_WRL 1953                                              3000000
5  World OWID_WRL 1954                                              3000000
6  World OWID_WRL 1955                                              4000000
> summary(df)
      Year      Global.plastics.production..million.tonnes...tonnes.
 Min.   :1950   Min.   :  2000000
 1st Qu.:1966   1st Qu.: 20750000
 Median :1982   Median : 76500000
 Mean   :1982   Mean   :118530303
 3rd Qu.:1999   3rd Qu.:198500000
 Max.   :2015   Max.   :381000000
```

2. Data Preprocessing

```
> df <- select(data, -c(Entity, Code))
> colnames(df) = c("Year", "Gpp")
> #scaling
> df$Gpp <- rescale(df$Gpp, to = c(1,100))
> head(df)
  Year      Gpp
1 1950 1.000000
2 1951 1.000000
3 1952 1.000000
4 1953 1.261214
5 1954 1.261214
6 1955 1.522427
```



3. Splitting the data into Training and Test set

```
> #training and test data
> trainingdata <- ts(df$Gpp, frequency = 1, start = 1950, end = 2005)
> testdata <- ts(df$Gpp, frequency = 1, start = 2006, end = 2015)
```

4. Holt's Exponential Smoothing model

```
> #holts exponential smoothing model
> model <- holt(trainingdata, h = 10)
```

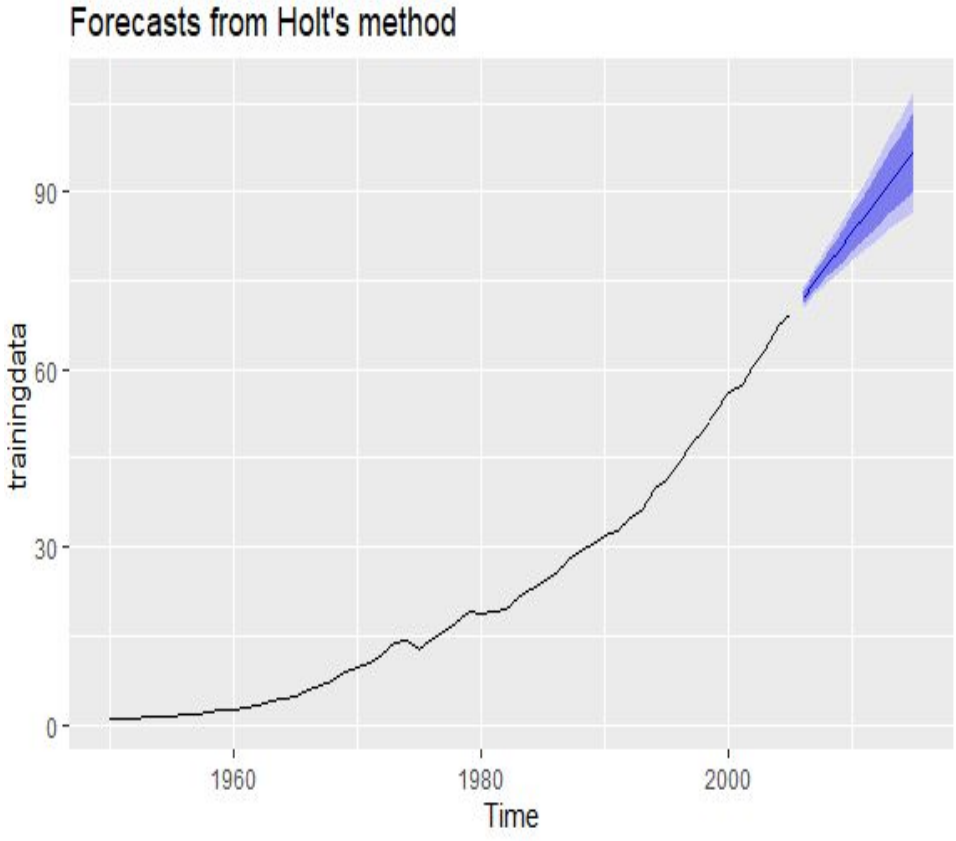
Model diagnostics

```
> summary(model)

Forecast method: Holt's method

Model Information:
Holt's method

Call:
 holt(y = trainingdata, h = 10)

  Smoothing parameters:
    alpha = 0.7886
    beta  = 0.2397

  Initial states:
    l = 0.8611
    b = 0.1007

  sigma:  0.8175

     AIC     AICc      BIC
208.7041 209.9041 218.8308

Error measures:
                     ME      RMSE       MAE      MPE     MAPE      MASE        ACF1
Training set 0.1979264 0.7877833 0.5684331 1.587652 4.584818 0.4352245 -0.0460277
```

Comparing the forecast with the test data

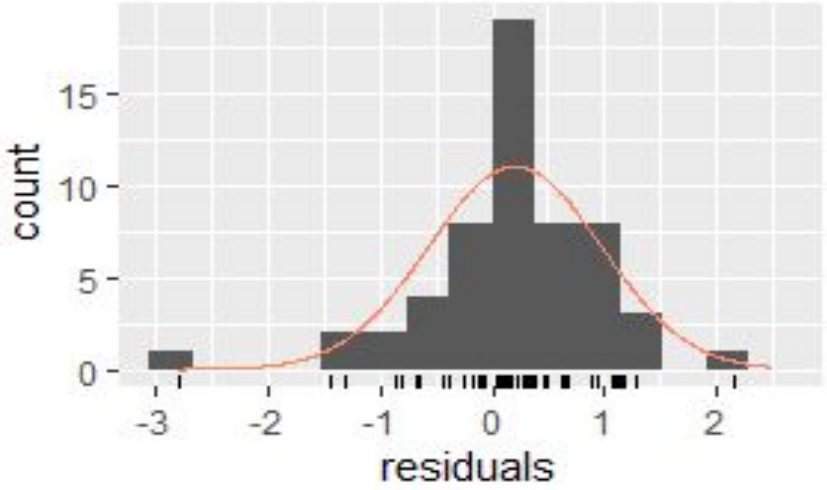
RMSE : 0.7877

```
Forecasts:
     Point Forecast    Lo 80     Hi 80    Lo 95     Hi 95
2006       72.11623 71.06853  73.16393 70.51392  73.71854
2007       74.87316 73.37040  76.37591 72.57489  77.17142
2008       77.63008 75.62436  79.63580 74.56260  80.69757
2009       80.38701 77.83403  82.93998 76.48257  84.29144
2010       83.14393 80.00250  86.28537 78.33952  87.94835
2011       85.90086 82.13232  89.66939 80.13738  91.66434
2012       88.65778 84.22569  93.08988 81.87948  95.43609
2013       91.41471 86.28446  96.54497 83.56866  99.26076
2014       94.17164 88.31023 100.03304 85.20739 103.13589
2015       96.92856 90.30442 103.55271 86.79780 107.05932
> testdata
   Year       Gpp
57 2006  73.61741
58 2007  77.53562
59 2008  73.87863
60 2009  75.70712
61 2010  82.23747
62 2011  85.37203
63 2012  88.76781
64 2013  92.42480
65 2014  96.34301
66 2015 100.00000
```



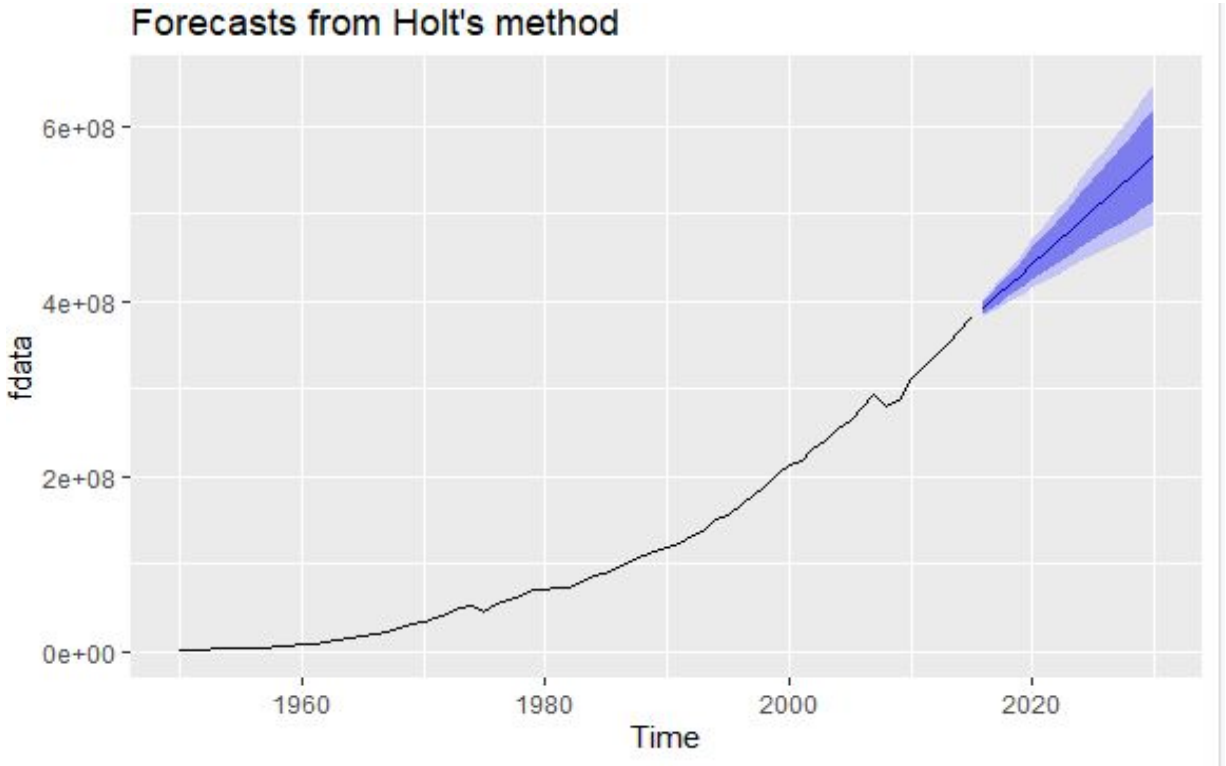
Residuals

```
> #residuals
> checkresiduals(model)

        Ljung-Box test

data:  Residuals from Holt's method
Q* = 8.2137, df = 6, p-value = 0.2229

Model df: 4.   Total lags used: 10
```



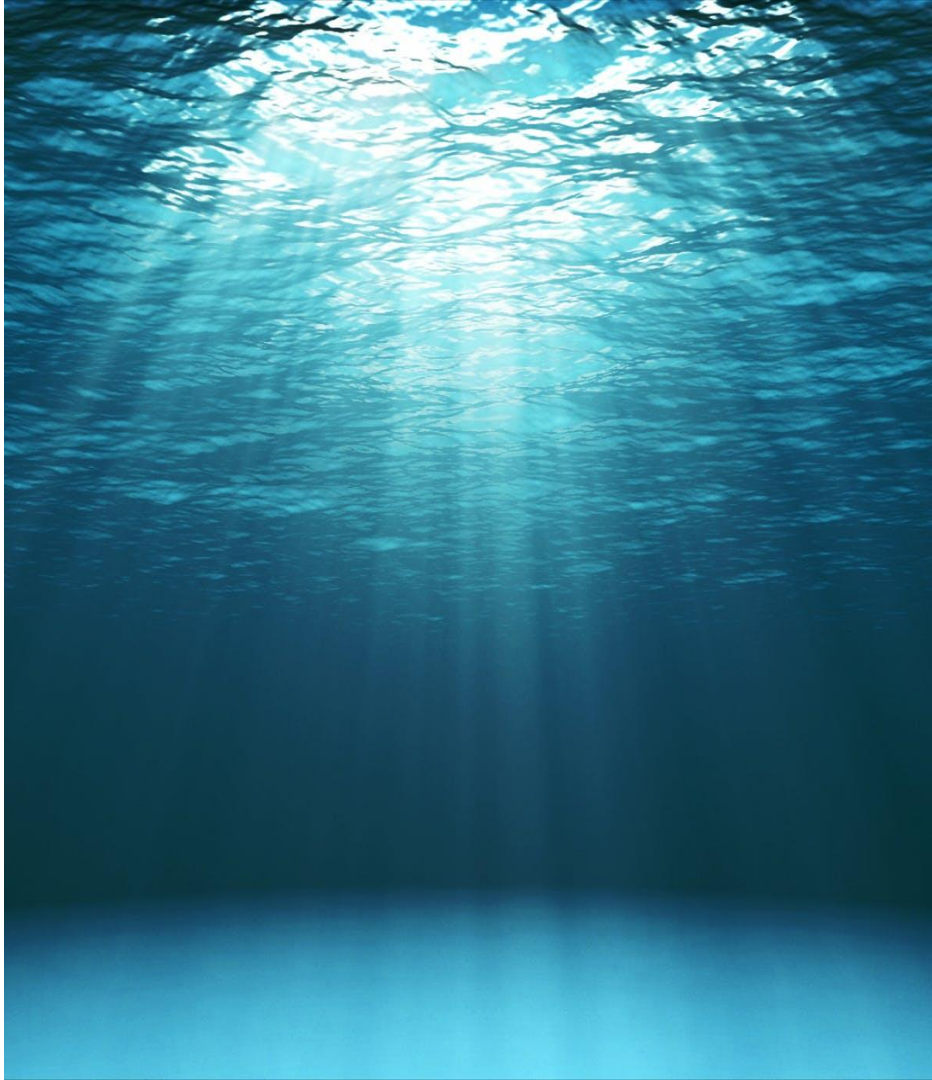
Forecasting Global plastic production for the next **15 years**

- The plastic production will be nearly **doubled** to what it is currently in the next 15 years, by 2030 it will reach 570 million metric tons approximately.



Conclusion

- From the analysis we determined the factors responsible for ineffective waste management.
- From the forecasts, we can conclude that we will only deteriorate the situation if we take no action on the amount of plastic being produced.



Any Questions?

References

1. Plastic waste inputs from land into the ocean, Jenna R Jambeck, Ronald Geyer.
2. Fighting for Trash Free Seas® ENDING THE FLOW OF TRASH AT THE SOURCE, Ocean conservancy.
3. Plastic in the ocean, 2019, world wildlife organization.
4. Exponential smoothing, UC business analysis guide.