Statistics Department,

Fergusson College,

(Established 1885)

# TIME SERIES:

# OIL PRICES IN INDIA

## Project Under

## CPE-UG

**PROJECT BY:**

**PARISEE SHIRKE**

**AJMAL ANSARI**

**ADITI IYER**

**SHUBHANGI ZANAN**

**Year 2014-15**

**S.Y.B.Sc. Statistics**

Semester 1

# CERTIFICATE

This is to certify that the project report entitled

"**TIME SERIES: OIL PRICES IN INDIA**" is an original work done by

**Parisee Shirke, Ajmal Ansari, Adithi Iyer,** and **Shubhangi Zanan**

under my supervision and able guidance.


| | |
|---|---|
| **Prof. S.S.SHENDE,** | **Prof. S.S.SHENDE,** |
| **Project Guide**, | **Head of Department**, |
| Dept. of Statistics, | Dept. of Statistics, |
| Fergusson College, | Fergusson College, |
| Pune-04. | Pune-04. |

# INDEX:

# ACKNOWLEDGEMENT

# ACKNOWLEDGEMENT

We would like to take this opportunity to thank everyone who has been a direct or indirect part of our project report entitled “TIME SERIES: OIL PRICES IN INDIA”.

We would like to thank our Project Guide Prof. S.S. Shende, HOD, Statistics Department for his valuable support and guidance, which helped us in completing this task through various stages.

We are also grateful to Dr. Mrs. Sharmishtha L. Matkar, HOD, Economics Department, who provided us with the information on the background of the subject which enhanced the quality of our project report.

We would also like to express a sense of gratitude to our teachers for their support and efforts that encouraged us to undertake this project work because of which we are able to get a practical experience of handling the data using the statistical tools which we have been studying from the last few years.

Also, a thanks to the non-teaching staff of the Statistics Department for their cooperation in making this project a success.

# ABSTRACT AND KEYWORDS

# ABSTRACT

This project is based on studying the changes in Oil Prices over the years. We study the trends in variations of Oil Prices over the years 1989 – 2010.

We use Time Series Analysis to process the data and observe the changes in the prices of Diesel, Petrol, Kerosene and LPG through the time period.

Using the patterns, hence, observed we have estimated and forecasted oil prices for the successive years from 2011 – 2020 for Diesel, Petrol, Kerosene and LPG each.

# KEYWORDS

- Time Series
- Fitting of Trend Line- Linear, Logarithmic, Exponential, Polynomial Trend Lines.
- R- Squared Value
- Residual Sum of Squares, Residual Plots.
- Exponential Smoothing
- Double Exponential Smoothing

# INTRODUCTION

# CHANGES IN OIL PRICES AND HOW THEY EFFECT OUR ECONOMY

This project report is basically focused on the changes in oil prices and how they affect our economy. Oil is an important resource in every walk of our life, may it be our households or our vehicles, and they all are in some way or the other dependent on oil.

The project is focused on the oil prices in India. A major being oil bed in India is situated in Digboi, Assam & Bombay High, Mumbai. Most of our oil is imported from the Arab countries and as the oil prices are affected there, the prices here in India also get affected.

Our industries especially the transportation industry is deeply influenced by the oil prices due to which transportation costs may it be the railway fare, the bus fare or even the local public transport fare is affected by the changes in oil prices. Other than this the chemical industries, import-export business is also affected.

So, in order to take a close look into the topic, the oil prices from the year 1989-2010 have been taken as the data. Analysis has been done using different statistical methods.

# DATA SOURCE AND SOFTWARES USED

## DATA SOURCE

The type of data used is Secondary Data.

The address of the website from which the data was obtained is:

in.reuters.com

## SOFTWARES USED

- Microsoft Excel
- Microsoft Word

# TIME SERIES ANALYSIS: DEFINITIONS AND CONCEPTS

# Time Series Analysis

A **time series or a stochastic process** (stochastic means random) is a sequence of data points, measured typically at successive points in time spaced at uniform time intervals. Time series are very frequently plotted via line charts. This involves analysis of data over time with the intention of finding some pattern in order to construct short term forecasting.

Forecasting is the use of a model to predict future values based on previously observed values. If history had been different, we would observe a different outcome, thus we can think of time series as the outcome of a random variable. The difference is that instead of dealing with individuals as units; the unit of interest here is time.

Allowing one to answer question such as what is the causal effect on a variable Y of a change in variable x

## Types of Time Series

### Continuous Time Series

A time series is said to be continuous when observations are made continuously in time. The term continuous is used for series of this type even when the measured variable can take only discrete set of values.

### Discrete Time Series

A time series is said to be discrete when observations are taken at specific time, usually equally spaced. The term discrete is used for series of this type even when the measured variable is a continuous variable.

## The Simple Moving Average

The simplest way to smooth a time series is to calculate a simple, or unweighted, moving average. The smoothed statistics is then just the mean of the last *k* observations:

$$s_t = \frac{1}{k}\sum_{n=0}^{k-1} x_{t-n} = \frac{x_t + x_{t-1} + x_{t-2} + \cdots + x_{t-k+1}}{k} = s_{t-1} + \frac{x_t - x_{t-k}}{k},$$

where the choice of an integer $k > 1$ is arbitrary.

A small value of *k* will have less of a Smoothing effect and be more responsive to recent changes in the data, while a larger *k* will have a greater Smoothing effect, and produce a more pronounced lag in the smoothed sequence. This means effectively extrapolating outside the existing data, and the validity of this section would therefore be questionable and not a direct representation of the data.

A major drawback with the SMA is that it lets through a significant amount of the signal shorter than the window length. Worse, it actually inverts it. This can lead to unexpected artefacts, such as peaks in the "smoothed" result appearing where there were troughs in the data. It also leads to the result being less "smooth" than expected since some of the higher frequencies are not properly removed.

The observed phenomenon may be an essentially random process, or it may be an orderly, but noisy, process. Whereas in the simple moving average the past observations are weighted equally, Exponential Smoothing assigns Exponentially decreasing weights over time.

# Weighted Moving Average

A slightly more intricate method for Smoothing a raw time series $\{x_t\}$ is to calculate a weighted moving average by first choosing a set of weighting factors

$\{w_1, w_2, \ldots, w_k\}$ Such that $\sum_{n=1}^{k} w_n = 1$

and then using these weights to calculate the smoothed statistics $\{s_t\}$:

$$s_t = \sum_{n=1}^{k} w_n x_{t+1-n} = w_1 x_t + w_2 x_{t-1} + \cdots + w_k x_{t-k+1}.$$

In practice the weighting factors are often chosen to give more weight to the most recent terms in the time series and less weight to older data. Notice that this technique has the same disadvantage as the simple moving average technique (i.e., it cannot be used until at least $k$ observations have been made), and that it entails a more complicated calculation at each step of the Smoothing procedure. In addition to this disadvantage, if the data from each stage of the averaging is not available for analysis, it may be difficult if not impossible to reconstruct a changing signal accurately (because older samples may be given less weight). If the number of stages missed is known however, the weighting of values in the average can be adjusted to give equal weight to all missed samples to avoid this issue.

## What is Exponential Smoothing?

This is a very popular scheme to produce a smoothed Time Series. Whereas in Single Moving Averages the past observations are weighted equally, Exponential Smoothing assigns Exponentially decreasing weights as the observation get older.

In other words, recent observations are given relatively more weight in forecasting than the older observations.

In the case of moving averages, the weights assigned to the observations are the same and are equal to 1/N. In Exponential Smoothing, however, there are one or more Smoothing parameters to be determined (or estimated) and these choices determine the weights assigned to the observations.

# <u>Simple Exponential Smoothening</u>

The Equation used in Simple Exponential Smoothing is

Current Forecast = α*previous value + (1-α)*previous Forecast value

I.e. $\hat{Y}_{t+1} = \alpha Y_t + (1-\alpha)\hat{Y}_t$

Thus $\hat{Y}_{t+1}$ is a Weighted arithmetic mean of $Y_t$ and $\hat{Y}_t$ . In this procedure, the recent past has assigned more weights whereas distant past has assigned discounted weights which decrease exponentially.

Where α is the *Smoothing factor*, and $0 < \alpha < 1$. In other words, the smoothed statistic $s_t$ is a simple weighted average of the previous observation $x_{t-1}$ and the previous smoothed statistic $s_{t-1}$. The term *Smoothing factor* applied to α here is something of a misnomer, as larger values of α actually reduce the level of Smoothing, and in the limiting case with α = 1 the output series is just the same as the original series (with lag of one time unit). Simple Exponential Smoothing is easily applied, and it produces a smoothed statistic as soon as two observations are available.

<u>Values of α close to one have less of a Smoothing effect</u> and <u>give greater weight to recent changes in the data</u>, while <u>values of α closer to zero have a greater Smoothing effect and are less responsive to recent changes</u>. There is no formally correct procedure for choosing α. Sometimes the statistician's judgment is used to choose an appropriate factor. Alternatively, a statistical technique may be used to <u>*optimize* the value of α</u>. For example, the <u>method of least squares</u> might be used to determine the value of α for which the sum of the quantities $(s_{n-1} - x_{n-1})^2$ is minimized.

## Why is Exponential Smoothening, Exponential?

By direct substitution of the defining equation for simple Exponential Smoothing back into itself we find that

$$\begin{aligned} s_t &= \alpha x_{t-1} + (1-\alpha)s_{t-1} \\ &= \alpha x_{t-1} + \alpha(1-\alpha)x_{t-2} + (1-\alpha)^2 s_{t-2} \\ &= \alpha\left[x_{t-1} + (1-\alpha)x_{t-2} + (1-\alpha)^2 x_{t-3} + (1-\alpha)^3 x_{t-4} + \cdots + (1-\alpha)^{t-1}x_0\right] + (1-\alpha)^t s_0. \end{aligned}$$

In other words, as time passes the smoothed statistic $s_t$ becomes the weighted average of a greater and greater number of the past observations $x_{t-n}$, and the weights assigned to previous observations are in general proportional to the terms of the geometric progression $\{1, (1-\alpha), (1-\alpha)^2, (1-\alpha)^3, \ldots\}$. A geometric progression is the discrete version of an exponential function, so this is where the name for this Smoothing method originated.

# Double Exponential Smoothening

Simple Exponential Smoothing does not do well when there is a trend in the data. In such situations, several methods were devised under the name "Double Exponential Smoothing" or "second-order Exponential Smoothing." The basic idea behind Double Exponential Smoothing is to introduce a term to take into account the possibility of a series exhibiting some form of trend. This slope component is itself updated via Exponential Smoothing.

One method sometimes referred to as "Holt-Winters Double Exponential Smoothing" works as follows:

Again, the raw data sequence of observations is represented by $\{x_t\}$, beginning at time $t = 0$. We use $\{s_t\}$ to represent the smoothed value for time $t$, and $\{b_t\}$ is our best estimate of the trend at time $t$. The output of the algorithm is now written as $F_{t+m}$, an estimate of the value of $x$ at time $t+m$, $m>0$ based on the raw data up to time $t$. Double Exponential Smoothing is given by the formulas

$$s_1 = x_1$$
$$b_1 = x_1 - x_0$$

And for t > 1 by

$$s_t = \alpha x_t + (1-\alpha)(s_{t-1} + b_{t-1})$$
$$b_t = \beta(s_t - s_{t-1}) + (1-\beta)b_{t-1}$$

where α is the *data Smoothing factor*, $0 < \alpha < 1$, and β is the *trend Smoothing factor*, $0 < \beta < 1$.

To forecast beyond $x_t$

$$F_{t+m} = s_t + m b_t$$

Where $a_t$ is level or intercept & $b_t$ is slope of the trend. Single Exponential Smoothing smoothes just the level. We need to smooth out slope ($b_t$) also since it undergoes frequent ups & downs, $F_{t+1}$ = Current forecast, $s_t$ = previous year's smoothed value of intercept & $b_t$ = previous year's smoothed value of slope of the trend.

We have two smoothing constants ($\alpha_1$, $\alpha_2$).The constant $\alpha_1$ is used for smoothing the level while $\alpha_2$ is used for smoothing the slope.

Setting the initial value $b_0$ is a matter of preference. An option other than the one listed above is $(x_n - x_0)/n$ for some $n > 1$.

Note that $F_0$ is undefined (there is no estimation for time 0), and according to the definition $F_1=s_0+b_0$, which is well defined, thus further values can be evaluated.

A second method, referred to as either Brown's Linear Exponential Smoothing (LES) or Brown's Double Exponential Smoothing works as follows.

$$\begin{aligned} s'_0 &= x_0 \\ s''_0 &= x_0 \\ s'_t &= \alpha x_t + (1-\alpha)s'_{t-1} \\ s''_t &= \alpha s'_t + (1-\alpha)s''_{t-1} \\ F_{t+m} &= a_t + m b_t, \end{aligned}$$

where $a_t$, the estimated level at time $t$ and $b_t$, the estimated trend at time $t$ are:

$$a_t = 2s'_t - s''_t$$

$$b_t = \frac{\alpha}{1-\alpha}(s'_t - s''_t).$$

# EXAMPLES OF TIME SERIES

## RICE PRODUCTION

**DATA:**

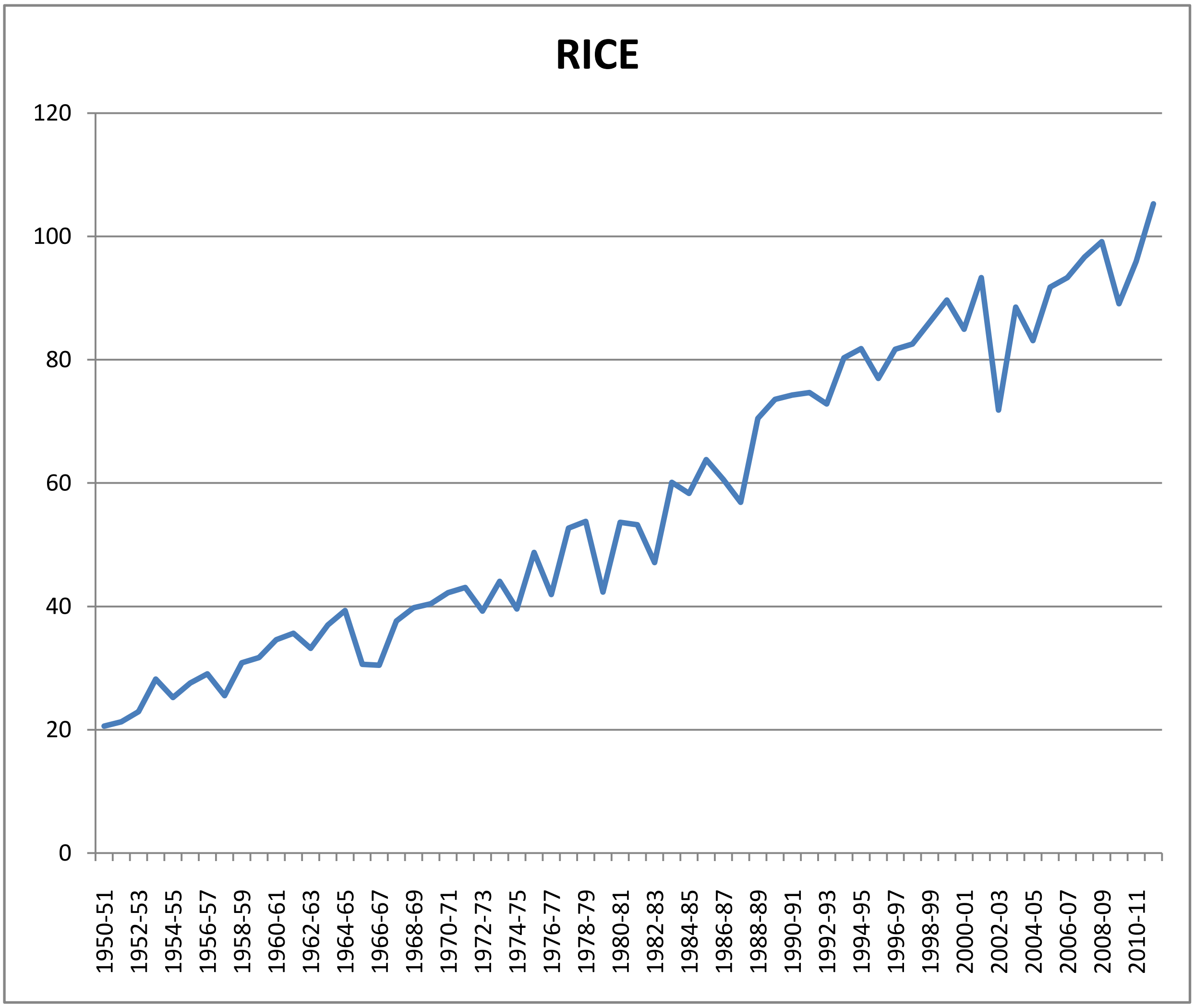
The following table gives the average annual Production of Rice in India in million tonnes:

| YEAR | RICE PROD | YEAR | RICE PROD | YEAR | RICE PROD | YEAR | RICE PROD |
|---|---|---|---|---|---|---|---|
| **1950-51** | 20.58 | **1965-66** | 30.59 | **1980-81** | 53.63 | **1995-96** | 76.98 |
| **1951-52** | 21.3 | **1966-67** | 30.44 | **1981-82** | 53.25 | **1996-97** | 81.73 |
| **1952-53** | 22.9 | **1967-68** | 37.61 | **1982-83** | 47.12 | **1997-98** | 82.54 |
| **1953-54** | 28.21 | **1968-69** | 39.76 | **1983-84** | 60.1 | **1998-99** | 86.08 |
| **1954-55** | 25.22 | **1969-70** | 40.43 | **1984-85** | 58.34 | **1999-00** | 89.68 |
| **1955-56** | 27.56 | **1970-71** | 42.22 | **1985-86** | 63.83 | **2000-01** | 84.98 |
| **1956-57** | 29.04 | **1971-72** | 43.07 | **1986-87** | 60.56 | **2001-02** | 93.34 |
| **1957-58** | 25.53 | **1972-73** | 39.24 | **1987-88** | 56.86 | **2002-03** | 71.82 |
| **1958-59** | 30.85 | **1973-74** | 44.05 | **1988-89** | 70.49 | **2003-04** | 88.53 |
| **1959-60** | 31.68 | **1974-75** | 39.58 | **1989-90** | 73.57 | **2004-05** | 83.13 |
| **1960-61** | 34.58 | **1975-76** | 48.74 | **1990-91** | 74.29 | **2005-06** | 91.79 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **1961-62** | 35.66 | **1976-77** | 41.92 | **1991-92** | 74.68 | **2006-07** | 93.35 |
| **1962-63** | 33.21 | **1977-78** | 52.67 | **1992-93** | 72.86 | **2007-08** | 96.69 |
| **1963-64** | 37 | **1978-79** | 53.77 | **1993-94** | 80.3 | **2008-09** | 99.18 |
| **1964-65** | 39.31 | **1979-80** | 42.33 | **1994-95** | 81.81 | **2009-10** | 89.09 |
| | | | | | | **2010-11** | 95.98 |
| | | | | | | **2011-12** | 105.3 |

**TIME SERIES:**

The Time Series, thus, obtained for Prices of Rice grain is as follows:




**CONCLUSION:**

The Trendline of the Production of Rice from the year 1951-2012 shows that the production increases throughout. Initially, we see that the production in the beginning was increasing slowly but after the **GREEN REVOLUTION** in **1975-76** there was a rapid increase in the production of rice, and by 2012 the production rose to five times that of the production in 1950's.

# GOLD PRICES

**DATA:**

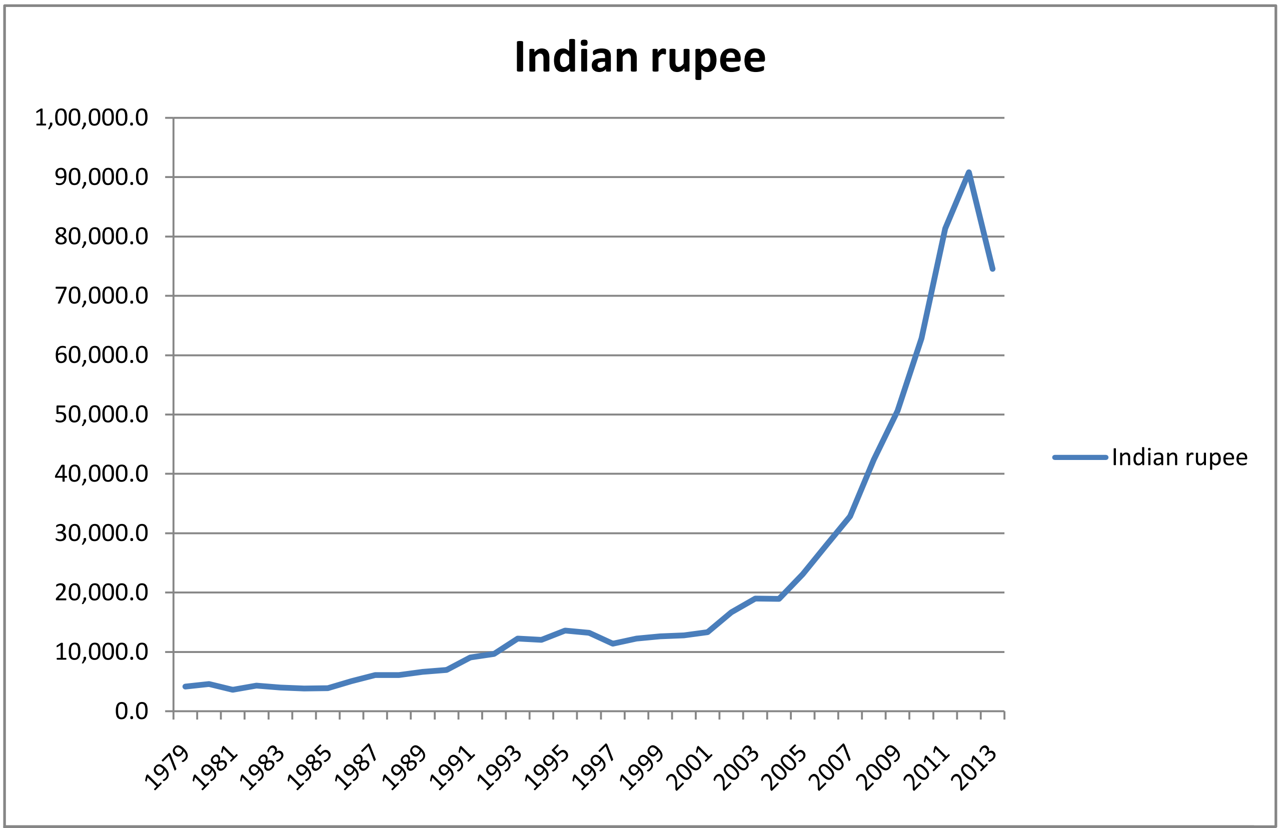
The following table gives the average annual Gold Prices in National Currency per unit Troy ounce.

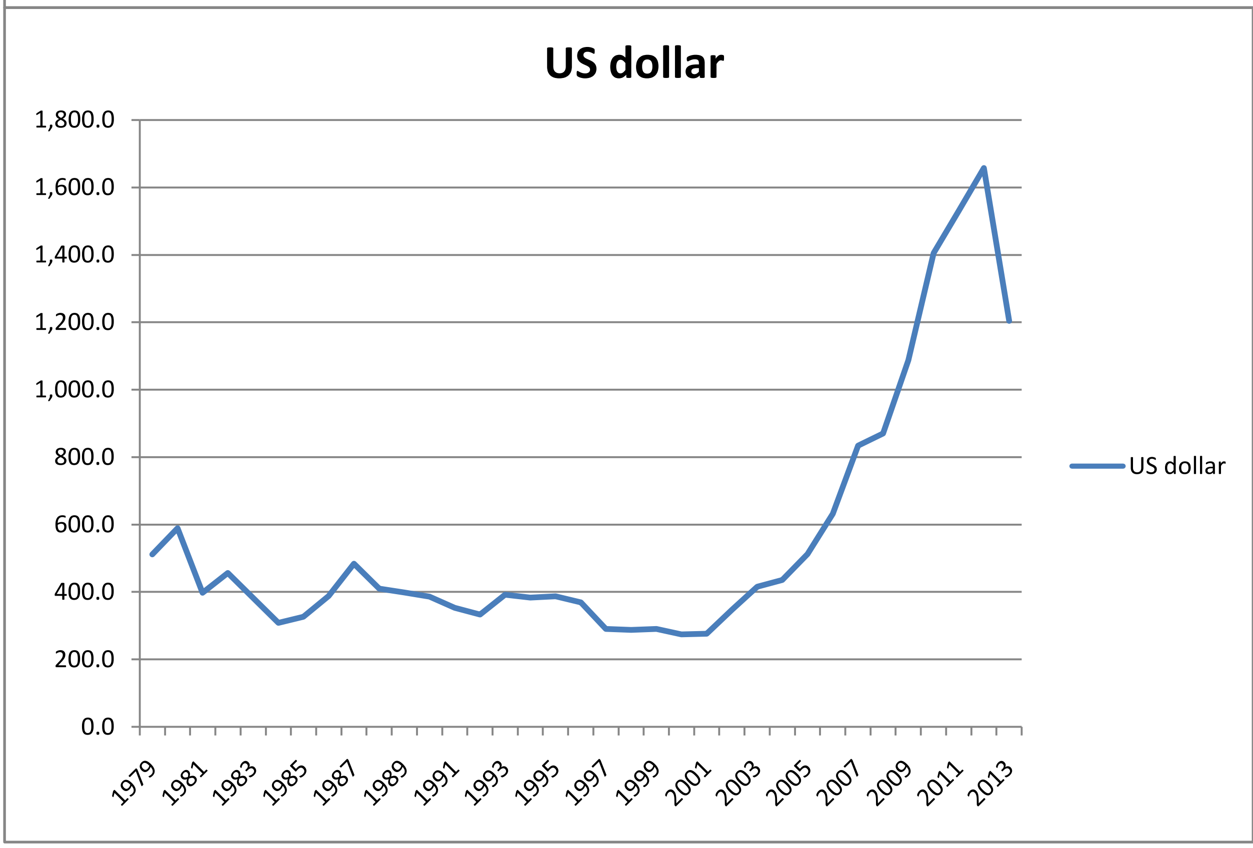
(1 Troy ounce= 31.1 grams)

| Year | Indian rupee | US dollar | Year | Indian rupee | US dollar |
|---|---|---|---|---|---|
| **1979** | 4,189.60 | 512 | **1996** | 13,228.40 | 369.3 |
| **1980** | 4,620.30 | 589.8 | **1997** | 11,378.70 | 290.2 |
| **1981** | 3,625.10 | 397.5 | **1998** | 12,230.40 | 287.8 |
| **1982** | 4,320.00 | 456.9 | **1999** | 12,631.00 | 290.3 |
| **1983** | 3,995.00 | 382.4 | **2000** | 12,811.30 | 274.5 |
| **1984** | 3,825.80 | 308.3 | **2001** | 13,332.80 | 276.5 |
| **1985** | 3,920.60 | 326.8 | **2002** | 16,648.20 | 347.2 |
| **1986** | 5,061.00 | 388.8 | **2003** | 18,991.40 | 416.3 |
| **1987** | 6,116.80 | 484.1 | **2004** | 18,935.50 | 435.6 |
| **1988** | 6,123.10 | 410.3 | **2005** | 23,090.10 | 513 |
| **1989** | 6,624.80 | 398.6 | **2006** | 27,972.30 | 632 |
| **1990** | 6,948.90 | 386.2 | **2007** | 32,862.30 | 833.8 |
| **1991** | 9,080.60 | 353.2 | **2008** | 42,374.20 | 869.8 |
| **1992** | 9,639.30 | 332.9 | **2009** | 50,606.80 | 1,087.50 |
| **1993** | 12,252.60 | 391.8 | **2010** | 62,846.90 | 1,405.50 |
| **1994** | 12,022.10 | 383.3 | **2011** | 81,303.70 | 1,531.00 |
| **1995** | 13,608.80 | 387 | **2012** | 90,814.40 | 1,657.50 |
| | | | **2013** | 74,504.30 | 1,204.50 |

**TIME SERIES:**

The Time Series for Gold Prices in Indian Rupee and US Dollar are as obtained below:



**CONCLUSION:**

Here, we have plotted the trendline for the Gold Prices (in Indian rupee& US dollar/unit Troy ounce) for the years 1979-2013. In the Indian market, we see that the prices are following an increasing trend, but initially the rise in the prices was a lower and after 1999-00, there was a price hike in the gold prices. In the US market, till 2001 the trend was approximately constant but suddenly after 2002, there was a steep increase in the prices. This also gives a fair idea of the Deterioration of the Indian Rupee in comparison with the US Dollar.

# METHODOLOGY

# METHODOLOGY

Our project group consisted of 4 people and our study was about the changes in oil prices and their effect on the economy. We contacted both the Economics and Statistics departments of our college whose guidelines throughout the project proved to be of great help.

In this project, we have made use of Secondary Data. For data collection we have gone through different sites and finally collected the data which consisted of Oil Prices from the year 1989 to 2010 in Rupee per Litre. The data consisted of four categories namely Kerosene, Petrol, Diesel and LPG (liquefied Petroleum gas).

The data was the revision of the prices of the above four categories, so we averaged out the data on monthly basis and did the further analysis. The main focus in this project has been given on time series analysis and its applications.

Firstly, we have plotted the time series for the data then we have fitted different models for the different categories, the models being Linear, Exponential, Logarithmic, Polynomial (with different degrees). Thereby we found out the best fit on the basis of the graphs (Here we found out that though the value of R^2 value was the largest still the fit was not the best fit). Thereafter estimates were calculated from the best fitted curves using the given equation and the Residual Plots plotted using the estimates. Finally we did Exponential Smoothening of the data to reduce the error factor.

# MOTIVATION

# MOTIVATION

Today a common man has a lot of issues regarding his day-to-day life. Life is becoming so fast that to compete with it one has to be fast, vehicles at every house appears as a basic need today. There is at least one vehicle at everyone's place these days. If there is no gas connection you'll be frustrated cooking on coal or on a heater.

Even though the standard of living is improving, the prices of goods are pushing it down; especially the oil prices and you can see that they are soaring these days.

When our parents were at high school oil prices at that time was merely a big issue. However days are not the same if we look back the prices of oil was the same as that of the prices of chocolates today. So, you see how days change, values change, affecting the common man. How fascinating would it be if we knew about what will be the prices of oil 10 years from now? There will be a boost in the share market, but it is not possible to tell what will the exact prices of oil be after 10 years but we could try to predict the Prices by carrying a bit of statistical analysis, Forecasting and getting an Estimated value. These things in some way or the other motivated us for doing this project.

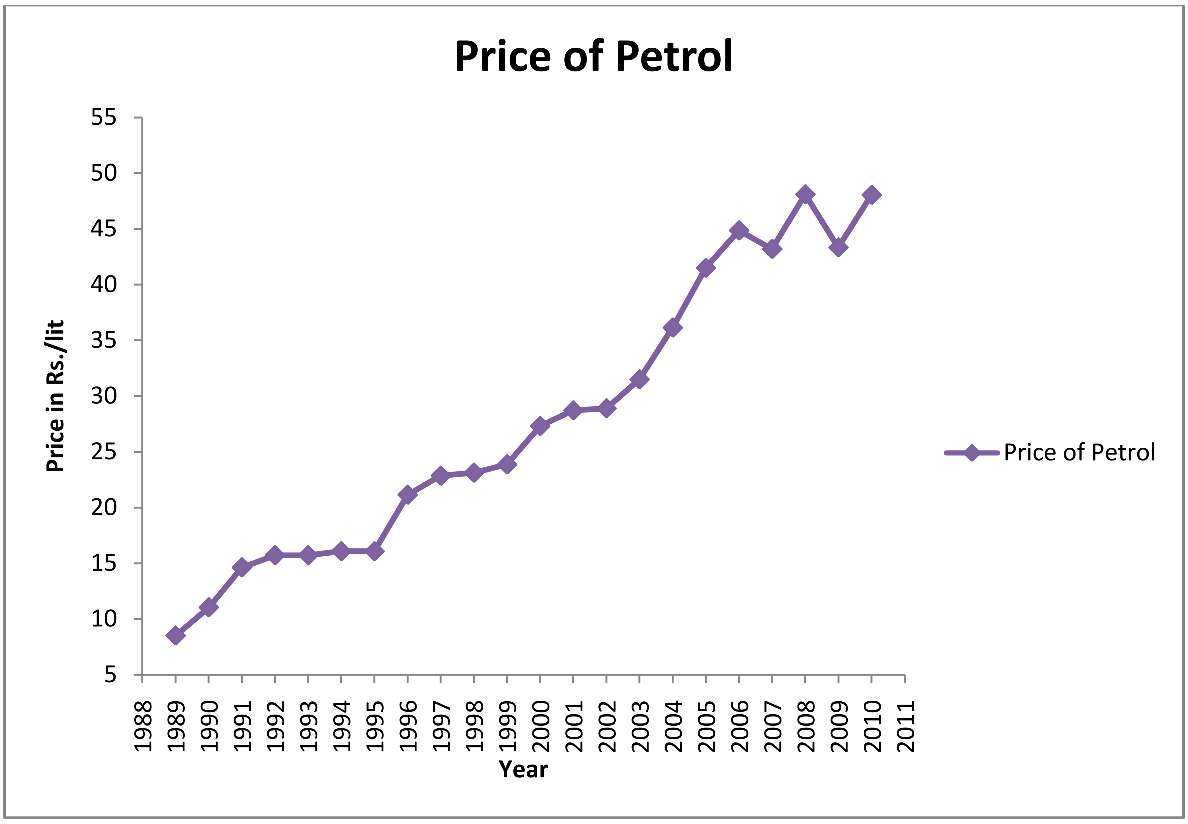
# DATA

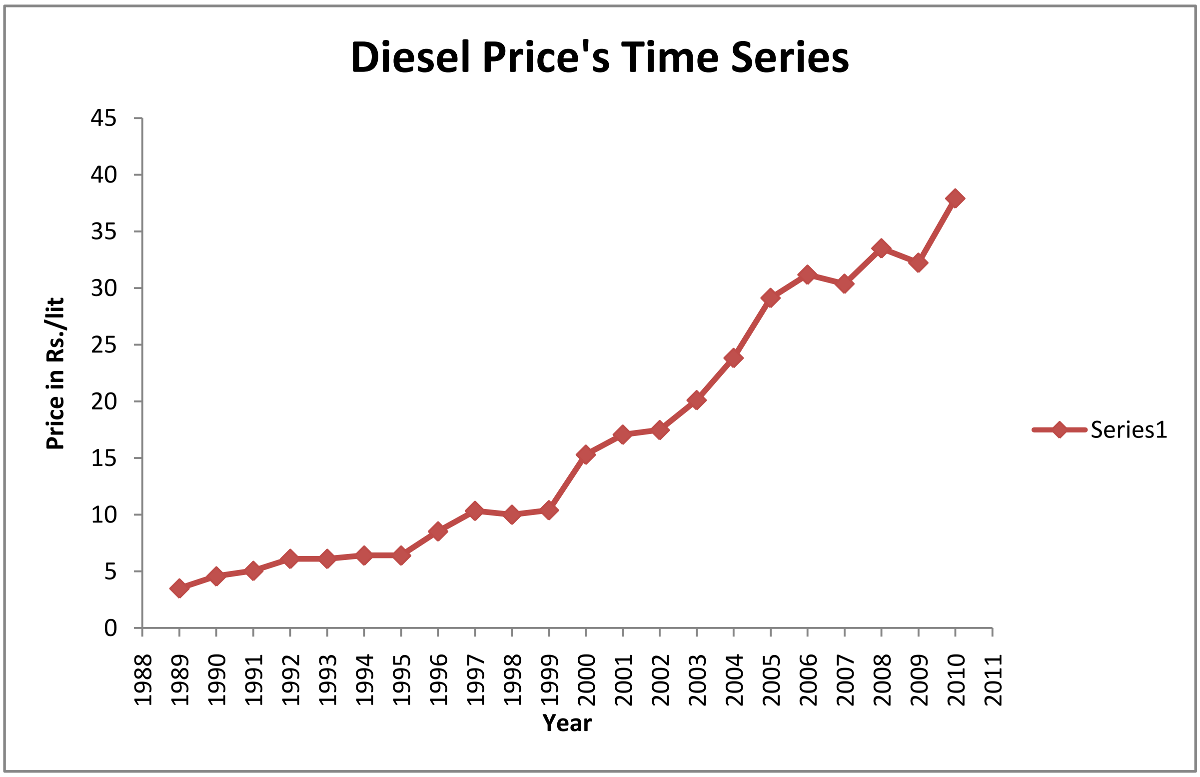
# DATA OF OIL PRICES

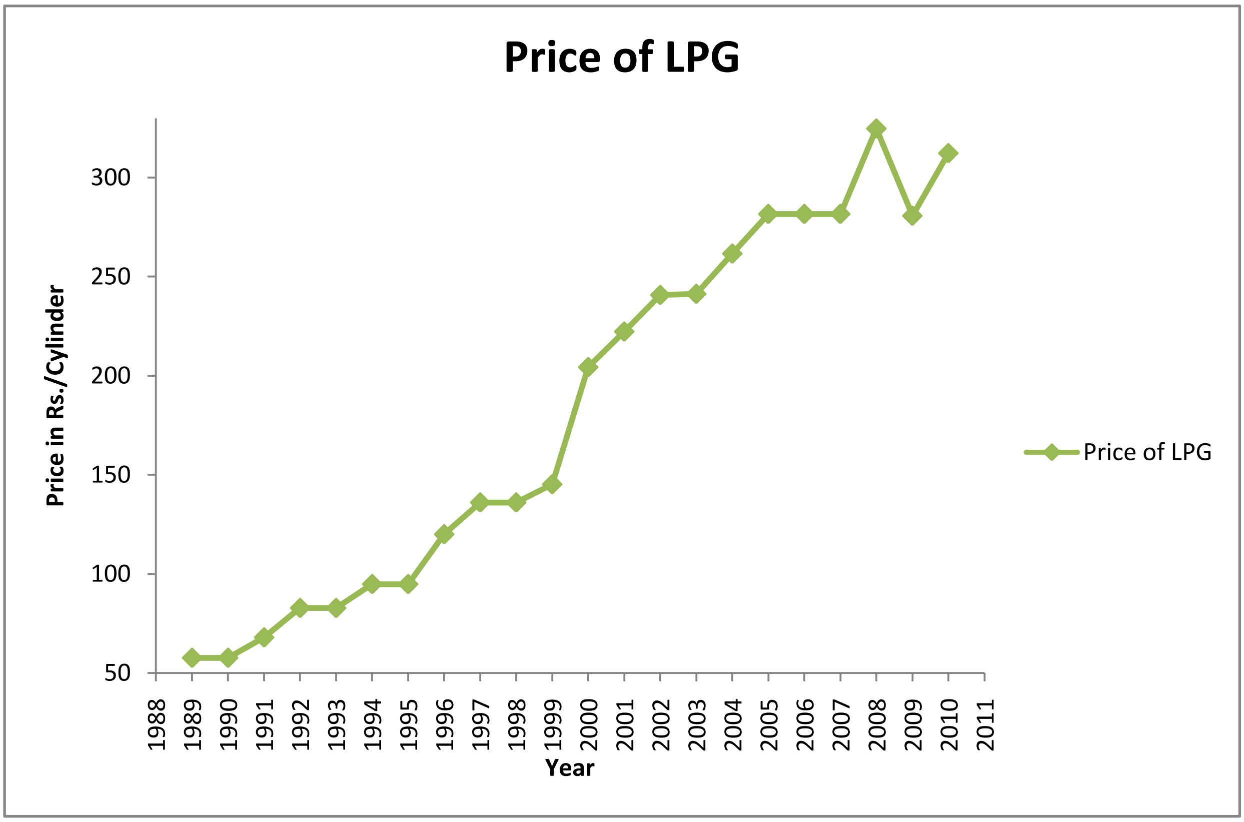
| Year | Price of Diesel (Rs./Litre) | Price of Petrol (Rs./Litre) | Price of Kerosene (Rs./Litre) | Price of LPG (Rs./Litre) |
|---|---|---|---|---|
| **1989** | 3.5 | 8.5 | 2.25 | 57.6 |
| **1990** | 4.565 | 11.035 | 2.51 | 57.6 |
| **1991** | 5.05 | 14.62 | 2.52 | 67.9 |
| **1992** | 6.11 | 15.71 | 2.52 | 82.75 |
| **1993** | 6.11 | 15.71 | 2.52 | 82.75 |
| **1994** | 6.4 | 16.0666667 | 2.52 | 94.8 |
| **1995** | 6.4 | 16.0666667 | 2.52 | 94.8 |
| **1996** | 8.53 | 21.13 | 2.52 | 119.95 |
| **1997** | 10.34 | 22.84 | 2.52 | 136 |
| **1998** | 10 | 23.115 | 2.52 | 136 |
| **1999** | 10.4 | 23.856 | 2.52 | 145.2 |
| **2000** | 15.295 | 27.2983333 | 6.27833333 | 204.308333 |
| **2001** | 17.06 | 28.7 | 7.35 | 222.25 |
| **2002** | 17.463846 | 28.8838462 | 8.85461538 | 240.723077 |
| **2003** | 20.102667 | 31.4826667 | 8.99 | 241.28 |
| **2004** | 23.821667 | 36.1283333 | 9.01 | 261.6 |
| **2005** | 29.116667 | 41.49 | 9.05666667 | 281.6 |
| **2006** | 31.165 | 44.845 | 9.0875 | 281.6 |
| **2007** | 30.365 | 43.185 | 9.1075 | 281.6 |
| **2008** | 33.49 | 48.0833333 | 9.18 | 324.733333 |
| **2009** | 32.216667 | 43.3233333 | 9.22 | 280.7 |
| **2010** | 37.89 | 48.03 | 10.22 | 312.3 |

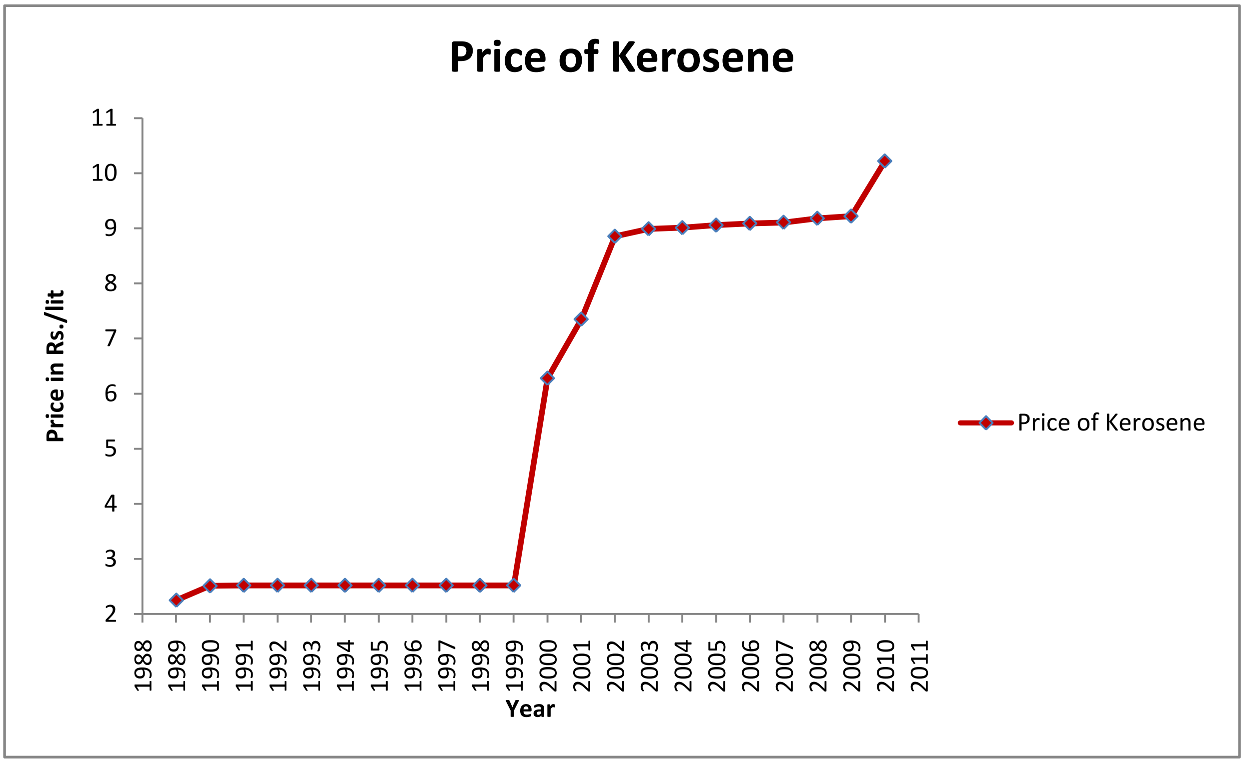
# DATA ANALYSIS

# TIME SERIES OF OIL PRICES






## CONCLUSION:

All The above graphs show an increasing trend.

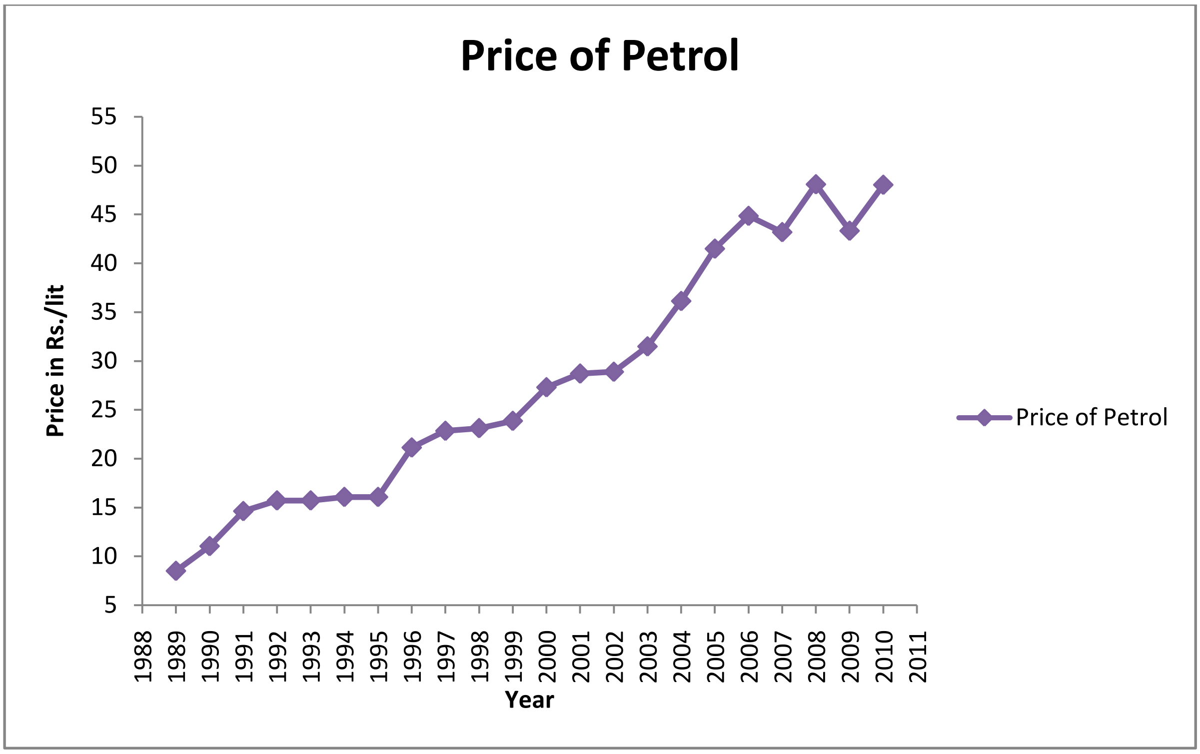
For each fuel data we could fit more than one Trend Lines, only one is selected for the analysis and estimation purpose. The selection of the graph depends on its R-squared value and whether it gives Realistic values or not.

For each fuel, Trend Line fits and their analyses are as follows-

# PETROL:

## DATA:

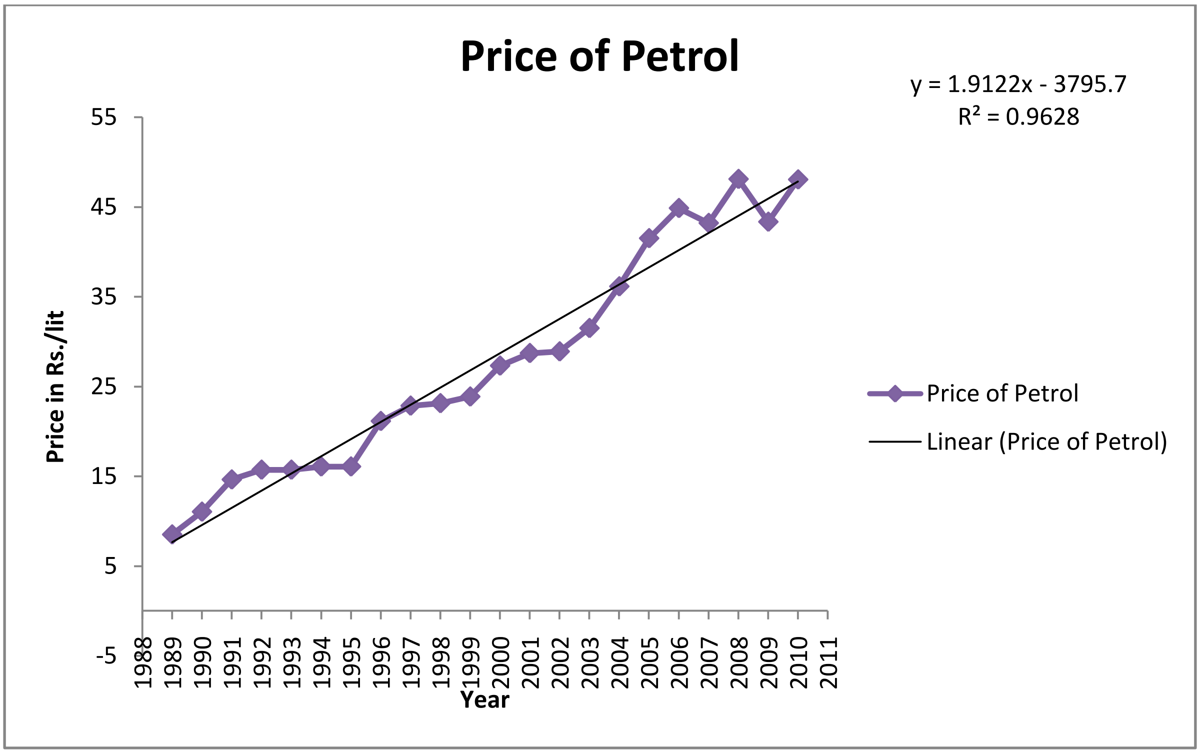
| Year | Price of Petrol (Rs./Litre) |
|---|---|
| **1989** | 8.5 |
| **1990** | 11.035 |
| **1991** | 14.62 |
| **1992** | 15.71 |
| **1993** | 15.71 |
| **1994** | 16.06666667 |
| **1995** | 16.06666667 |
| **1996** | 21.13 |
| **1997** | 22.84 |
| **1998** | 23.115 |
| **1999** | 23.856 |
| **2000** | 27.29833333 |
| **2001** | 28.7 |
| **2002** | 28.88384615 |
| **2003** | 31.48266667 |
| **2004** | 36.12833333 |
| **2005** | 41.49 |
| **2006** | 44.845 |
| **2007** | 43.185 |
| **2008** | 48.08333333 |
| **2009** | 43.32333333 |
| **2010** | 48.03 |



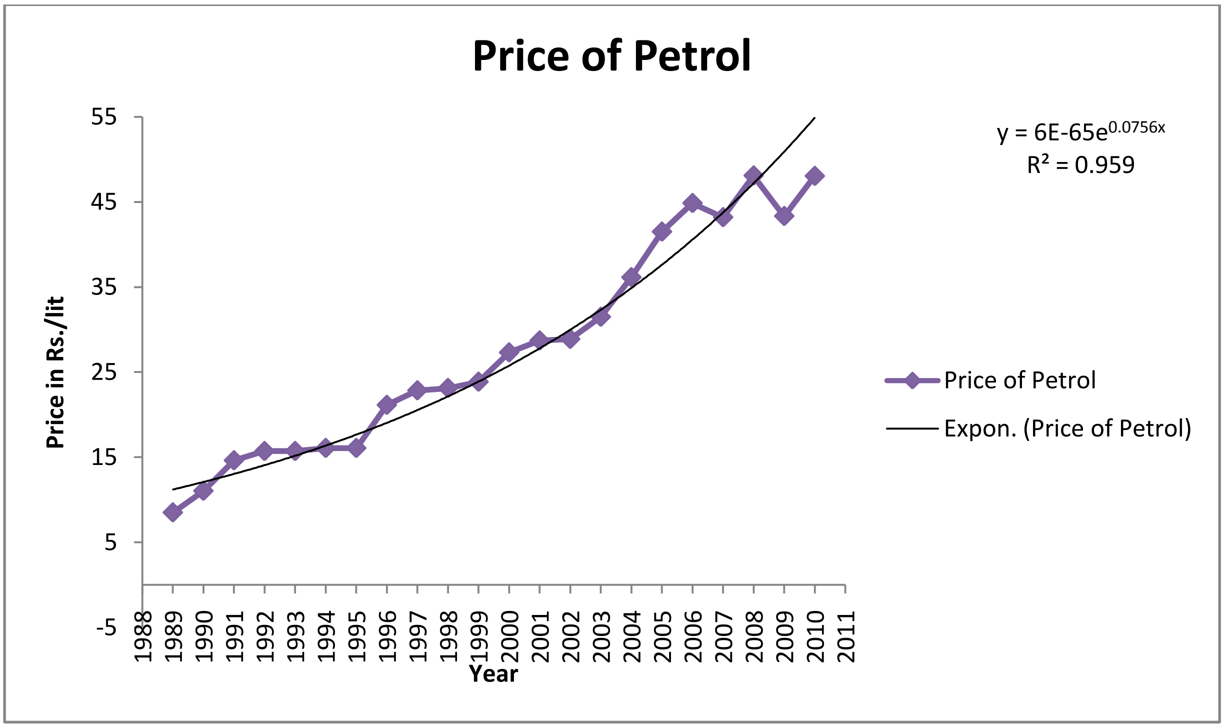

## FITTING OF TREND LINES:

We fit the following Trend Lines to the the above data:

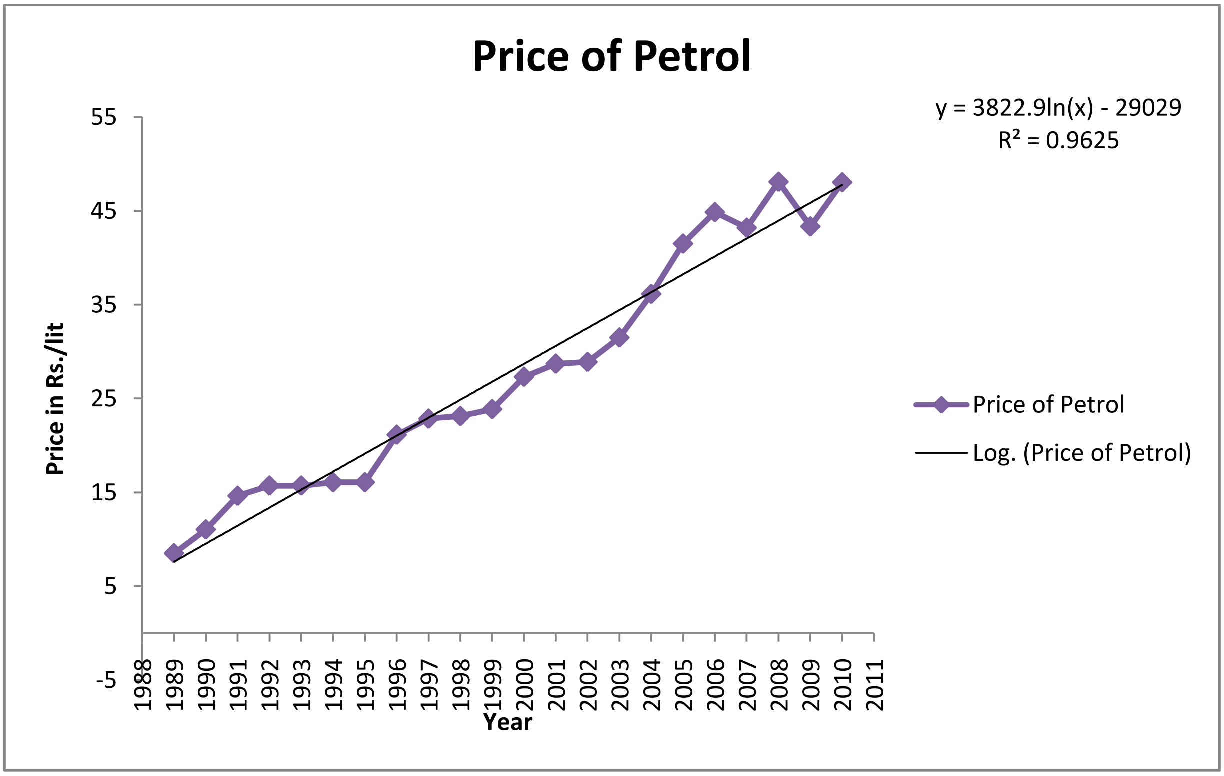
1. Linear-



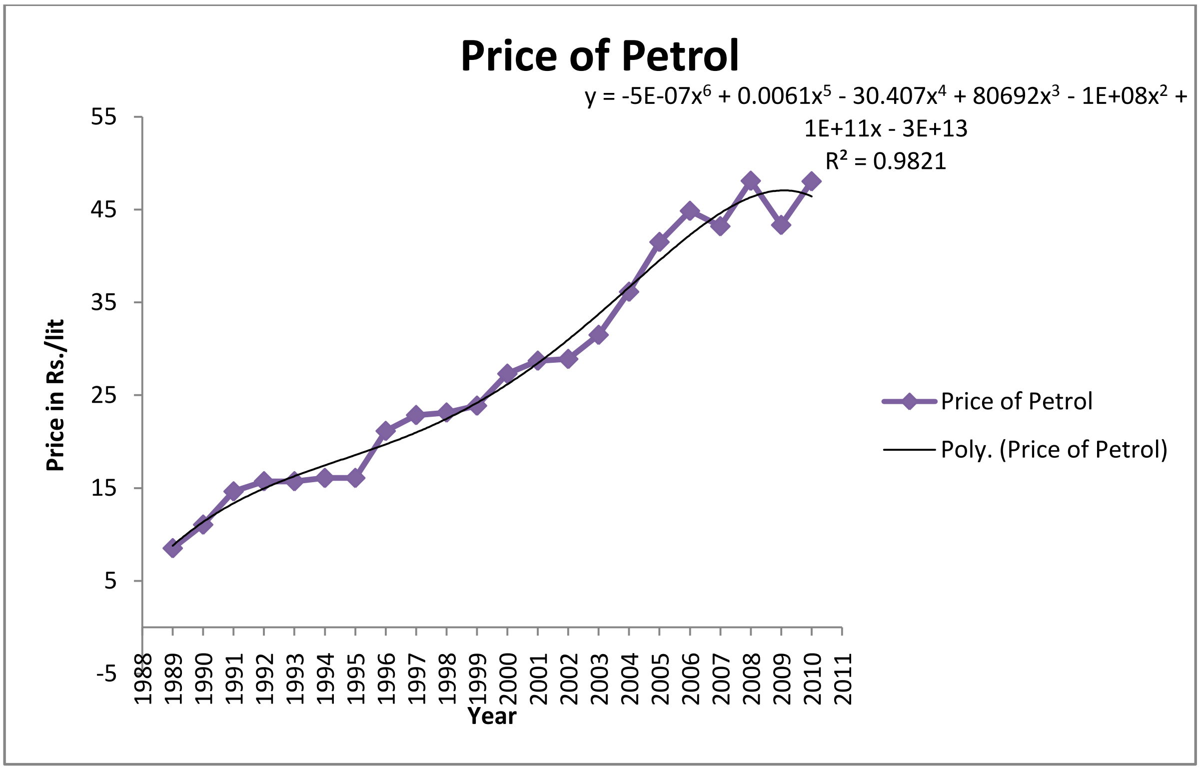
## 1. Exponential -




## 3. Logarithmic -



4. Polynomial of degree six-




## EQUATIONS:

Hence we obtain the equations of the fit from the graphs as follows:

| Trend Line | Equation |
|---|---|
| Exponential | $y = 6E\text{-}65e^{0.075x}$ |
| Linear | $y = 1.912x - 3795$ |
| Logarithmic | $y = 3822.\ln(x) - 29029$ |
| Polynomial | $y = -5E\text{-}07x^6 + 0.006x^5 - 30.40x^4 + 80692x^3 - 1E{+}08x^2 + 1E{+}11x - 3E{+}13$ |

**ESTIMATES:**

Using these equations the Petrol prices were estimated as follows:

| | | Estimates | | | |
|---|---|---|---|---|---|
| **Year** | **Price of Petrol** | **Polynomial** | **Linear** | **Logarithmic** | **Exponential** |
| | | | | | |
| **1989** | 8.5 | -108002 | 7.6658 | 7.406028 | 12.08697 |
| **1990** | 11.035 | -108001 | 9.578 | 9.327566 | 13.03618 |
| **1991** | 14.62 | -107999 | 11.4902 | 11.24814 | 14.05992 |
| **1992** | 15.71 | -107998 | 13.4024 | 13.16775 | 15.16406 |
| **1993** | 15.71 | -107996 | 15.3146 | 15.08639 | 16.35491 |
| **1994** | 16.06666667 | -107995 | 17.2268 | 17.00408 | 17.63928 |
| **1995** | 16.06666667 | -107993 | 19.139 | 18.9208 | 19.02451 |
| **1996** | 21.13 | -107992 | 21.0512 | 20.83656 | 20.51853 |
| **1997** | 22.84 | -107990 | 22.9634 | 22.75136 | 22.12987 |
| **1998** | 23.115 | -107988 | 24.8756 | 24.6652 | 23.86776 |
| **1999** | 23.856 | -107986 | 26.7878 | 26.57808 | 25.74212 |
| **2000** | 27.29833333 | -107984 | 28.7 | 28.49001 | 27.76367 |
| **2001** | 28.7 | -107982 | 30.6122 | 30.40098 | 29.94398 |
| **2002** | 28.88384615 | -107980 | 32.5244 | 32.311 | 32.29552 |
| **2003** | 31.48266667 | -107978 | 34.4366 | 34.22007 | 34.83172 |
| **2004** | 36.12833333 | -107976 | 36.3488 | 36.12818 | 37.56709 |
| **2005** | 41.49 | -107973 | 38.261 | 38.03534 | 40.51727 |
| **2006** | 44.845 | -107971 | 40.1732 | 39.94154 | 43.69914 |
| **2007** | 43.185 | -107968 | 42.0854 | 41.8468 | 47.13088 |
| **2008** | 48.08333333 | -107966 | 43.9976 | 43.75111 | 50.83212 |
| **2009** | 43.32333333 | -107963 | 45.9098 | 45.65447 | 54.82402 |
| **2010** | 48.03 | -107961 | 47.822 | 47.55689 | 59.12941 |

Here we observe that prices given by the Polynomial fit are not realistic hence we do not consider it.

## SUM OF SQUARE OF RESIDUALS (SSR):

For the other fits we calculate the residuals and their SSR. The one with least SSR is considered as the best fit.

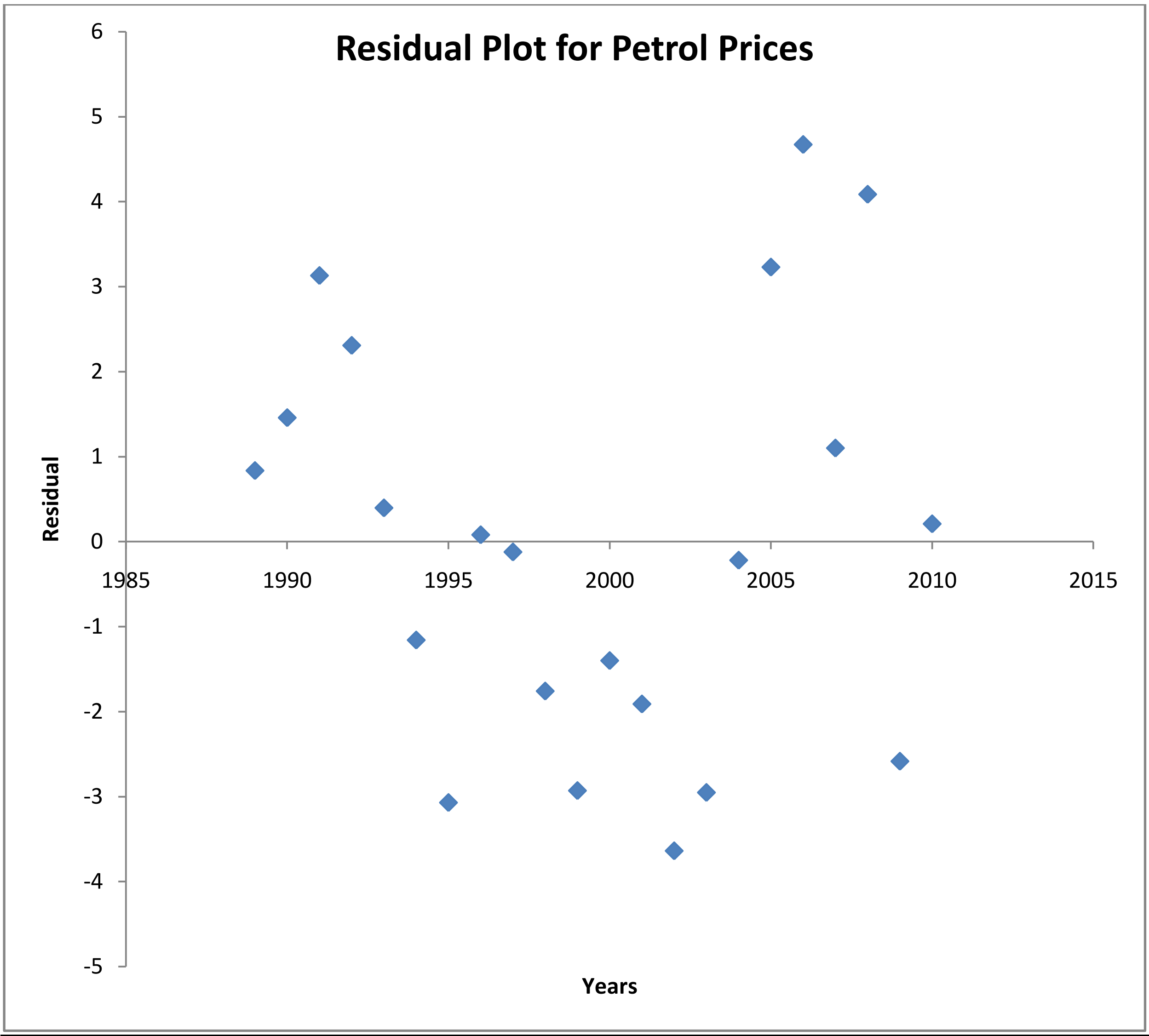
| Year | Price of Petrol | Linear | | Logarithmic | | Exponential | |
|---|---|---|---|---|---|---|---|
| | | Residual | SSR | Residual | SSR | Residual | SSR |
| 1989 | 8.5 | 0.8342 | 0.69589 | 1.093972 | 1.196774 | 11.25277 | 12.86639 |
| 1990 | 11.035 | 1.457 | 2.122849 | 1.707434 | 2.915329 | 11.57918 | 4.004713 |
| 1991 | 14.62 | 3.1298 | 9.795648 | 3.371861 | 11.36945 | 10.93012 | 0.313686 |
| 1992 | 15.71 | 2.3076 | 5.325018 | 2.542252 | 6.463047 | 12.85646 | 0.298046 |
| 1993 | 15.71 | 0.3954 | 0.156341 | 0.623607 | 0.388886 | 15.95951 | 0.415914 |
| 1994 | 16.06666667 | -1.16013 | 1.345909 | -0.93741 | 0.878734 | 18.79942 | 2.473122 |
| 1995 | 16.06666667 | -3.07233 | 9.439232 | -2.85413 | 8.146055 | 22.09685 | 8.748865 |
| 1996 | 21.13 | 0.0788 | 0.006209 | 0.293443 | 0.086109 | 20.43973 | 0.373895 |
| 1997 | 22.84 | -0.1234 | 0.015228 | 0.088642 | 0.007857 | 22.25327 | 0.504281 |
| 1998 | 23.115 | -1.7606 | 3.099712 | -1.5502 | 2.40312 | 25.62836 | 0.566641 |
| 1999 | 23.856 | -2.9318 | 8.595451 | -2.72208 | 7.409744 | 28.67392 | 3.557433 |
| 20 | 27.2983 | - | 1.9646 | - | 1.420 | 29.16 | 0.216 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **00** | 3333 | 1.40167 | 69 | 1.19168 | 099 | 534 | 54 |
| **2001** | 28.7 | -1.9122 | 3.656509 | -1.70098 | 2.89335 | 31.85618 | 1.547493 |
| **2002** | 28.88384615 | -3.64055 | 13.25363 | -3.42716 | 11.7454 | 35.93607 | 11.63949 |
| **2003** | 31.48266667 | -2.95393 | 8.725722 | -2.7374 | 7.493356 | 37.78565 | 11.21614 |
| **2004** | 36.12833333 | -0.22047 | 0.048606 | 0.000156 | 2.45E-08 | 37.78756 | 2.070021 |
| **2005** | 41.49 | 3.229 | 10.42644 | 3.454664 | 11.9347 | 37.28827 | 0.946196 |
| **2006** | 44.845 | 4.6718 | 21.82572 | 4.903456 | 24.04388 | 39.02734 | 1.312997 |
| **2007** | 43.185 | 1.0996 | 1.20912 | 1.338198 | 1.790774 | 46.03128 | 15.56996 |
| **2008** | 48.08333333 | 4.085733 | 16.69322 | 4.332223 | 18.76815 | 46.74638 | 7.555817 |
| **2009** | 43.32333333 | -2.58647 | 6.68981 | -2.33114 | 5.434205 | 57.41049 | 132.2658 |
| **2010** | 48.03 | 0.208 | 0.043264 | 0.473115 | 0.223838 | 58.92141 | 123.1969 |
| | **Total-** | | **125.1342** | | 127.0129 | | 341.6603 |

The SSR of Linear is the minimum. Hence Linear fit is the best fit for the Petrol prices.

## RESIDUAL PLOT:

After fitting the appropriate Trend Line to the oil prices, the prices were estimated for the collected data using the equations. Then the residuals were obtained as the difference of the observed prices and the estimated prices. Then a graph of the year and the corresponding residual were plotted to obtain a Residual Plot to study the goodness of the fit.



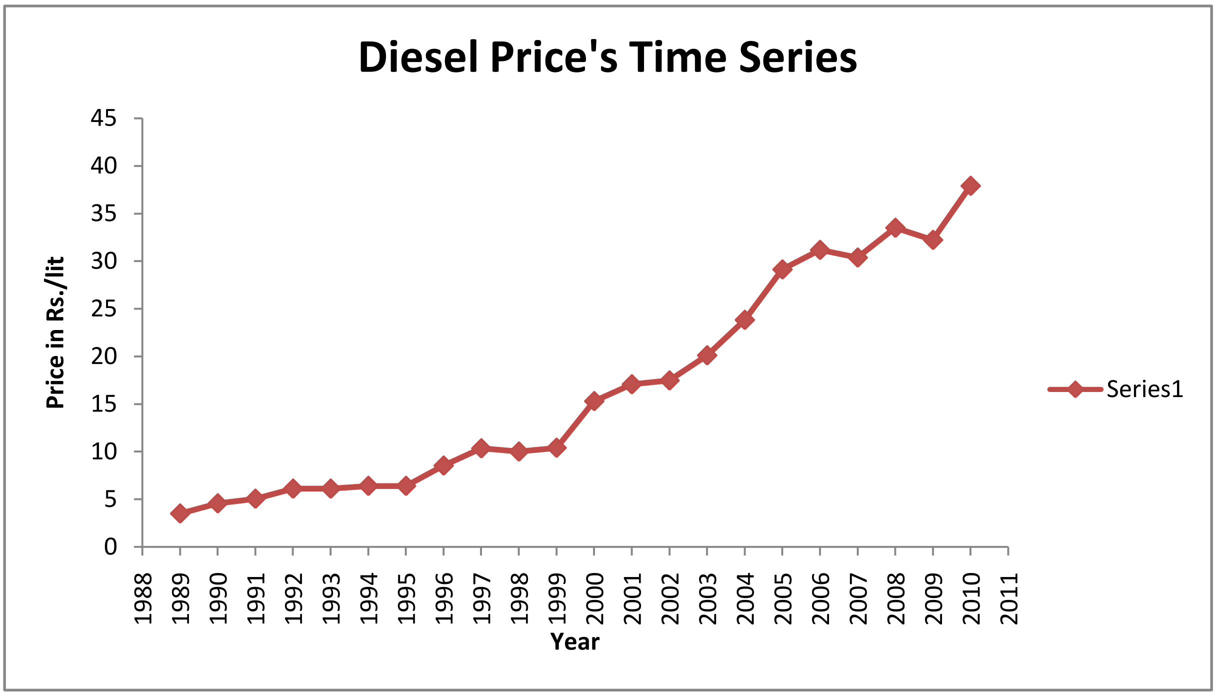

## CONCLUSION:

From the graphs we can conclude that the Linear fit for Petrol is good as the Residual Plot does not show any pattern

# DIESEL:

**DATA:**

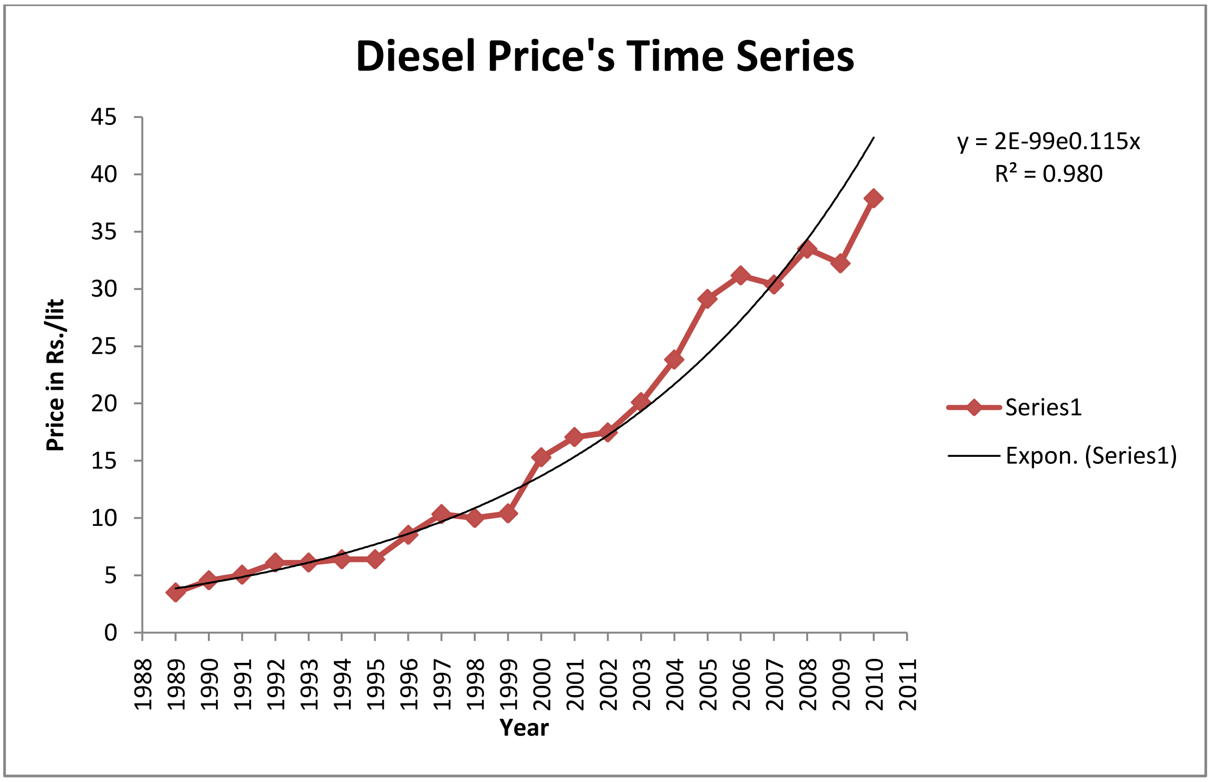
| Year | Price of Diesel (Rs./Litre) |
|---|---|
| **1989** | 3.5 |
| **1990** | 4.565 |
| **1991** | 5.05 |
| **1992** | 6.11 |
| **1993** | 6.11 |
| **1994** | 6.4 |
| **1995** | 6.4 |
| **1996** | 8.53 |
| **1997** | 10.34 |
| **1998** | 10 |
| **1999** | 10.4 |
| **2000** | 15.295 |
| **2001** | 17.06 |
| **2002** | 17.463846 |
| **2003** | 20.102667 |
| **2004** | 23.821667 |
| **2005** | 29.116667 |
| **2006** | 31.165 |
| **2007** | 30.365 |
| **2008** | 33.49 |
| **2009** | 32.216667 |
| **2010** | 37.89 |



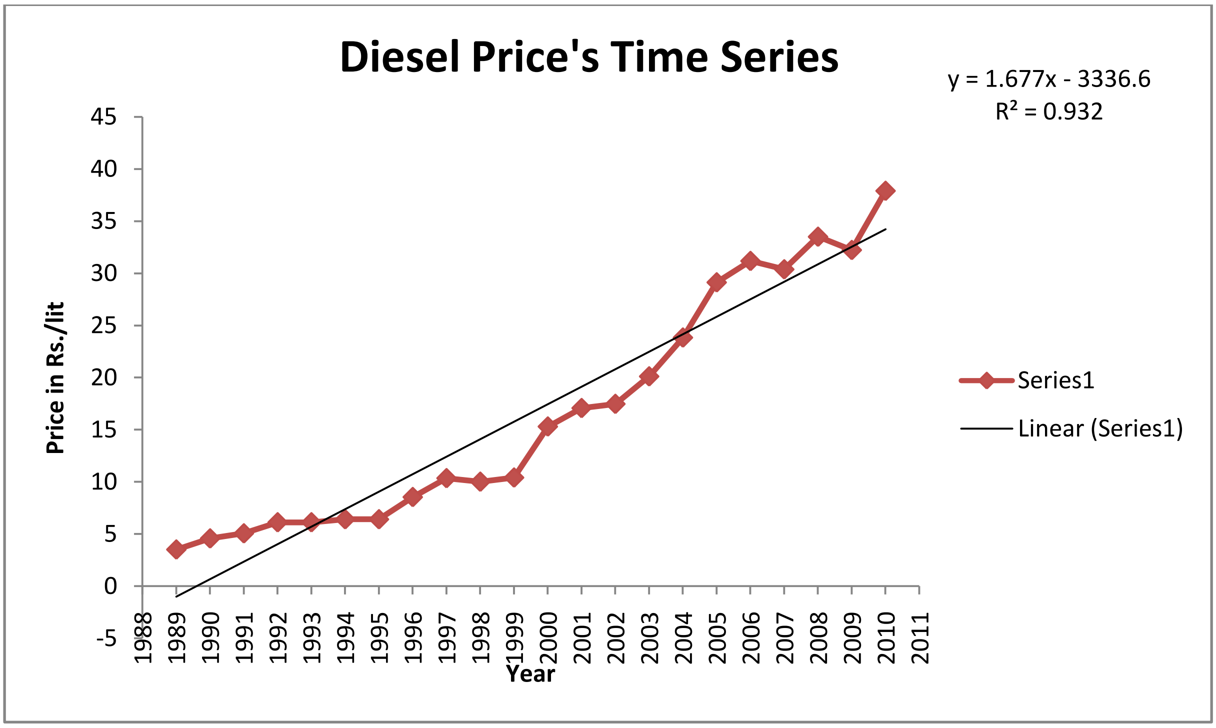

## TREND LINE:

We fit the following Trend Lines to the above data:

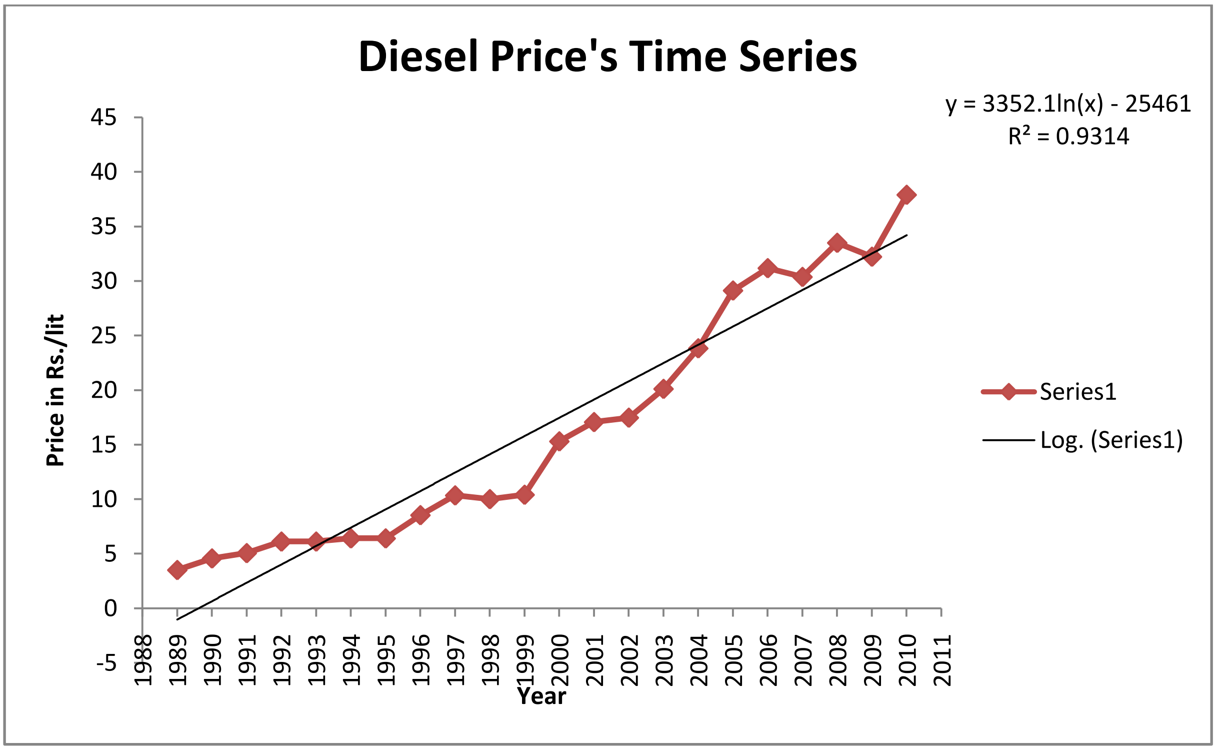
1. Exponential-



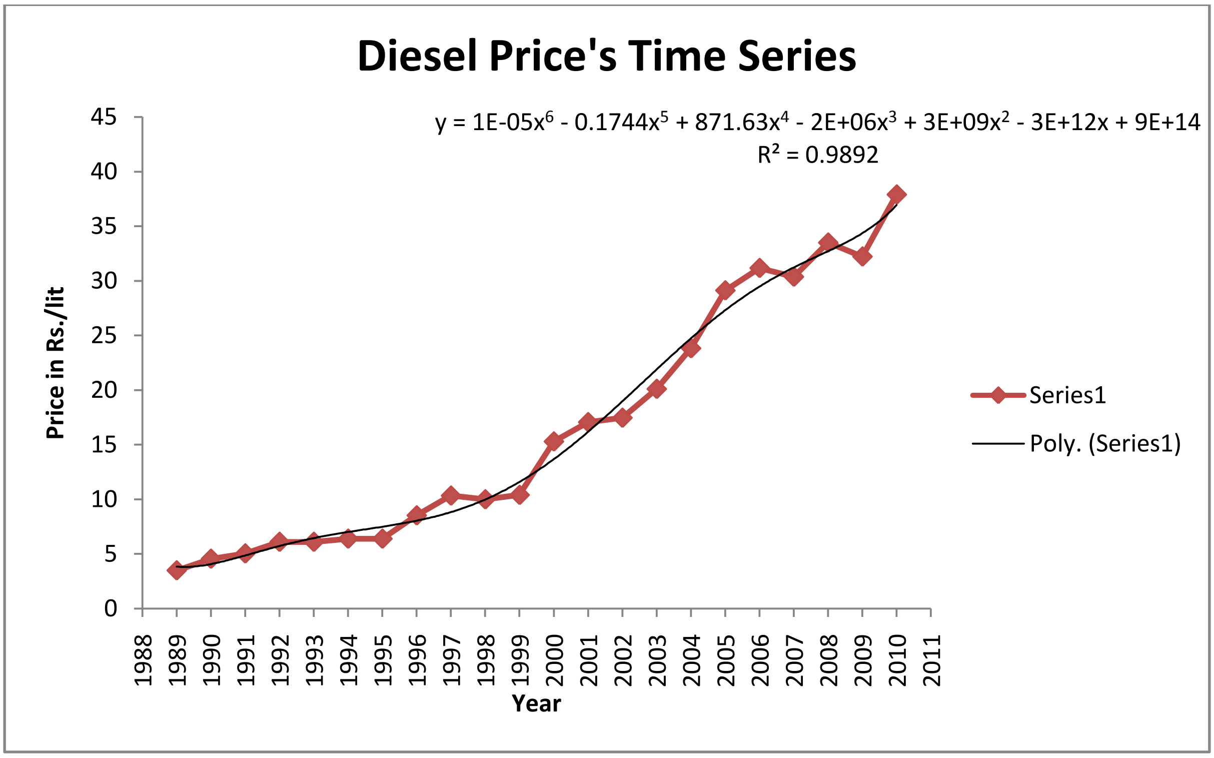
## 2. Linear-




## 3. Logarithmic-



4. Polynomial-




## EQUATIONS:

Hence we obtain the equations of the fit from the graphs as follows:

| ***Trend Line*** | ***Equation*** |
|---|---|
| Exponential | $y = 2E\text{-}99e^{0.115x}$ |
| Linear | $y = 1.677x - 3336$ |
| Logarithmic | $y = 3352.1\ln(x) - 25461$ |
| Polynomial | $y = 1E\text{-}05x^6 - 0.174x^5 + 871.6x^4 - 2E{+}06x^3 + 3E{+}09x^2 - 3E{+}12x + 9E{+}14$ |

**ESTIMATES:**

Using these equations the Diesel prices were estimated as follows:

| Year | Price of Diesel | Estimation | | | |
|---|---|---|---|---|---|
| | | Linear | Logarithmic | Exponential | Polynomial |
| **1989** | 3.5 | -0.447 | -1.26184128 | 4.35891396 | -9.21545E+13 |
| **1990** | 4.565 | 1.23 | 0.42300419 | 4.89014979 | -9.12718E+13 |
| **1991** | 5.05 | 2.907 | 2.10700322 | 5.48612916 | -9.03882E+13 |
| **1992** | 6.11 | 4.584 | 3.79015666 | 6.15474258 | -8.95039E+13 |
| **1993** | 6.11 | 6.261 | 5.47246535 | 6.90484221 | -8.86187E+13 |
| **1994** | 6.4 | 7.938 | 7.15393015 | 7.74635907 | -8.77328E+13 |
| **1995** | 6.4 | 9.615 | 8.83455189 | 8.69043448 | -8.6846E+13 |
| **1996** | 8.53 | 11.292 | 10.5143314 | 9.7495676 | -8.59584E+13 |
| **1997** | 10.34 | 12.969 | 12.1932696 | 10.9377809 | -8.507E+13 |
| **1998** | 10 | 14.646 | 13.8713673 | 12.2708059 | -8.41808E+ |

| | | | | | |
|---|---|---|---|---|---|
| | | | | | 13 |
| **1999** | 10.4 | 16.323 | 15.5486252 | 13.7662912 | -8.32908E+13 |
| **2000** | 15.295 | 18 | 17.2250444 | 15.4440364 | -8.24E+13 |
| **2001** | 17.06 | 19.677 | 18.9006255 | 17.3262542 | -8.15084E+13 |
| **2002** | 17.46384615 | 21.354 | 20.5753695 | 19.4378644 | -8.0616E+13 |
| **2003** | 20.10266667 | 23.031 | 22.2492772 | 21.8068237 | -7.97228E+13 |
| **2004** | 23.82166667 | 24.708 | 23.9223493 | 24.4644963 | -7.88289E+13 |
| **2005** | 29.11666667 | 26.385 | 25.5945868 | 27.4460685 | -7.79341E+13 |
| **2006** | 31.165 | 28.062 | 27.2659905 | 30.7910153 | -7.70385E+13 |
| **2007** | 30.365 | 29.739 | 28.9365612 | 34.5436221 | -7.61422E+13 |
| **2008** | 33.49 | 31.416 | 30.6062997 | 38.7535721 | -7.52451E+13 |
| **2009** | 32.21666667 | 33.093 | 32.2752069 | 43.4766032 | -7.43472E+13 |
| **2010** | 37.89 | 34.77 | 33.9432835 | 48.7752462 | -7.34485E+13 |

Here we observe that prices given by the Polynomial line are not realistic hence we do not consider it.

**SUM OF SQUARES OF RESIDUALS (SSR):**

For the other fits we calculate the residuals and their SSR. The one with least SSR is considered as the best fit.

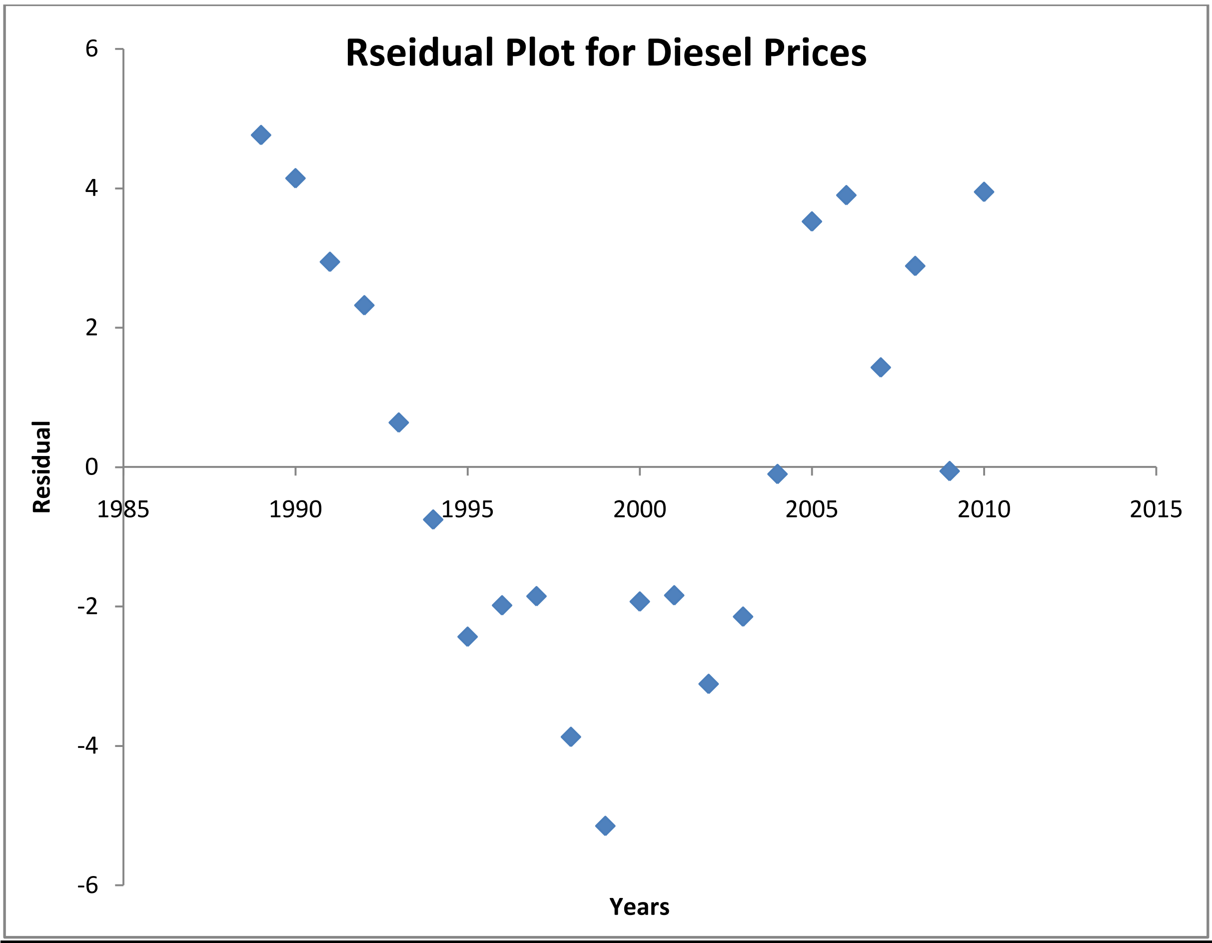
| Year | Price of Diesel | Linear | | Logarithmic | | Exponential | |
|---|---|---|---|---|---|---|---|
| | | Residual | SSR | Residual | SSR | Residual | SSR |
| 1989 | 3.5 | 3.947 | 15.57881 | 4.761841 | 22.67513 | -0.85891 | 0.737733 |
| 1990 | 4.565 | 3.335 | 11.12222 | 4.141996 | 17.15613 | -0.32515 | 0.105722 |
| 1991 | 5.05 | 2.143 | 4.592449 | 2.942997 | 8.66123 | -0.43613 | 0.190209 |
| 1992 | 6.11 | 1.526 | 2.328676 | 2.319843 | 5.381673 | -0.04474 | 0.002002 |
| 1993 | 6.11 | -0.151 | 0.022801 | 0.637535 | 0.40645 | -0.79484 | 0.631774 |
| 1994 | 6.4 | -1.538 | 2.365444 | -0.75393 | 0.568411 | -1.34636 | 1.812683 |
| 1995 | 6.4 | -3.215 | 10.33623 | -2.43455 | 5.927043 | -2.29043 | 5.24609 |
| 1996 | 8.53 | -2.762 | 7.628644 | -1.98433 | 3.937571 | -1.21957 | 1.487345 |
| 1997 | 10.34 | -2.629 | 6.911641 | -1.85327 | 3.434608 | -0.59778 | 0.357342 |
| 1998 | 10 | -4.646 | 21.58532 | -3.87137 | 14.98748 | -2.27081 | 5.156559 |
| 1999 | 10.4 | -5.923 | 35.08193 | -5.14863 | 26.50834 | -3.36629 | 11.33192 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **2000** | 15.295 | -2.705 | 7.317025 | -1.93004 | 3.725071 | -0.14904 | 0.022212 |
| **2001** | 17.06 | -2.617 | 6.848689 | -1.84063 | 3.387902 | -0.26625 | 0.070891 |
| **2002** | 17.46384615 | -3.89015 | 15.1333 | -3.11152 | 9.681578 | -1.97402 | 3.896748 |
| **2003** | 20.10266667 | -2.92833 | 8.575136 | -2.14661 | 4.607937 | -1.70416 | 2.904151 |
| **2004** | 23.82166667 | -0.88633 | 0.785587 | -0.10068 | 0.010137 | -0.64283 | 0.41323 |
| **2005** | 29.11666667 | 2.731667 | 7.462003 | 3.52208 | 12.40505 | 1.670598 | 2.790898 |
| **2006** | 31.165 | 3.103 | 9.628609 | 3.89901 | 15.20228 | 0.373985 | 0.139865 |
| **2007** | 30.365 | 0.626 | 0.391876 | 1.428439 | 2.040438 | -4.17862 | 17.46088 |
| **2008** | 33.49 | 2.074 | 4.301476 | 2.8837 | 8.315728 | -5.26357 | 27.70519 |
| **2009** | 32.21666667 | -0.87633 | 0.76796 | -0.05854 | 0.003427 | -11.2599 | 126.7862 |
| **2010** | 37.89 | 3.12 | 9.7344 | 3.946716 | 15.57657 | -10.8852 | 118.4886 |
| | **Total-** | | 188.5002 | | **184.6002** | | 327.7382 |

The SSR of Logarithmic is the minimum. Hence <u>Logarithmic fit is the best fit for the Diesel prices.</u>

## RESIDUAL PLOT:

After fitting the appropriate Trend Line to the oil prices, the prices were estimated for the collected data using the equations. Then the residuals were obtained as the difference of the observed prices and the estimated prices. Then a graph of the year and the corresponding residual were plotted to obtain a Residual Plot to study the goodness of the fit.



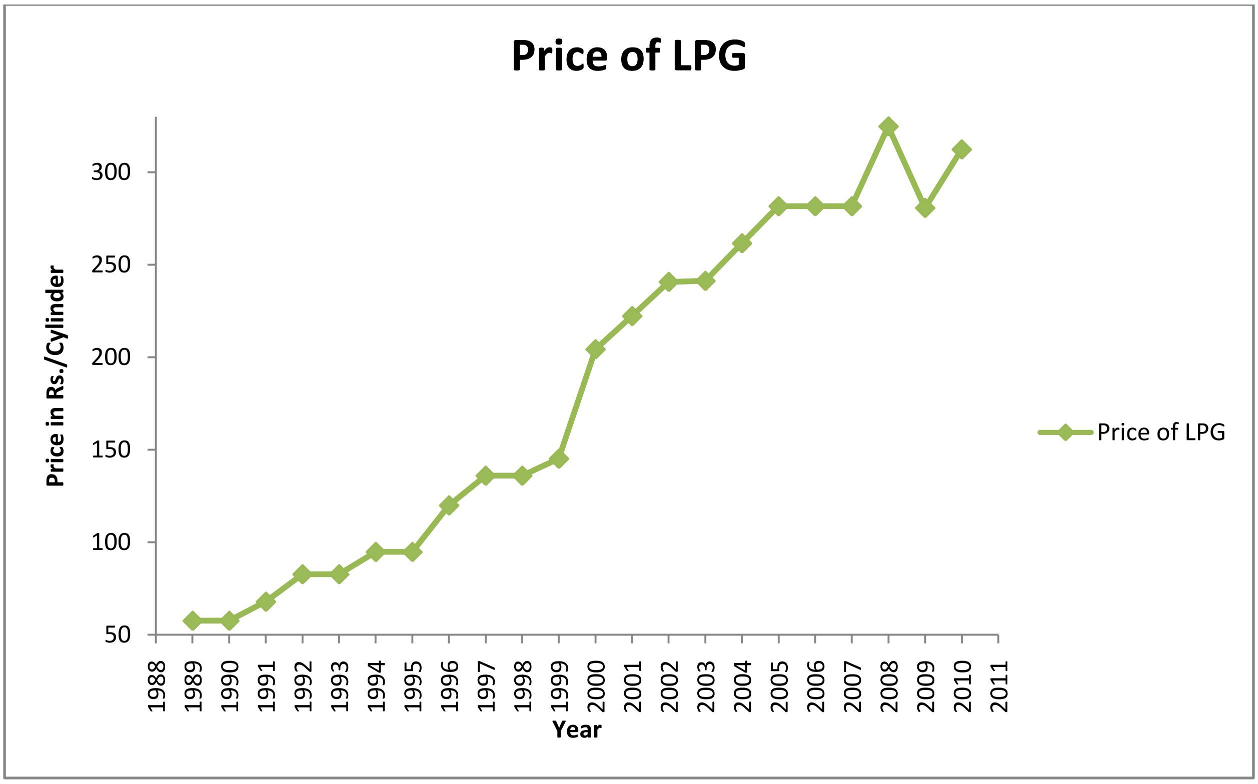

## CONCLUSION:

The Residual Plot of Diesel shows a certain pattern, which means that the fit is not very good. But due to the limitations of Microsoft Excel other fits, which could be better, cannot be used for the estimation. Hence we proceed with Logarithmic Fit as the best fit for Diesel Prices.

# LPG:

**DATA:** (1 Cylinder= 14.2 Litre)

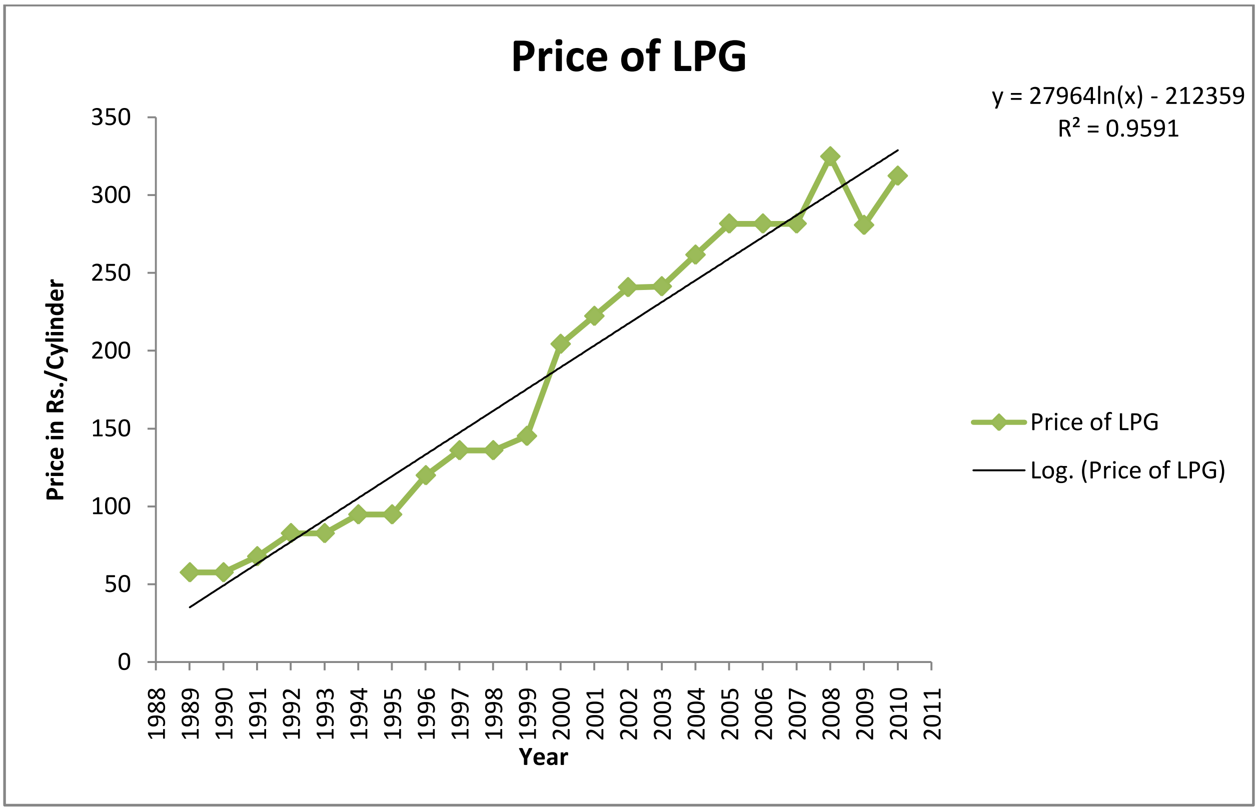
| Year | Price of LPG (Rs./Cylinder) |
|---|---|
| **1989** | 57.6 |
| **1990** | 57.6 |
| **1991** | 67.9 |
| **1992** | 82.75 |
| **1993** | 82.75 |
| **1994** | 94.8 |
| **1995** | 94.8 |
| **1996** | 119.95 |
| **1997** | 136 |
| **1998** | 136 |
| **1999** | 145.2 |
| **2000** | 204.3083 |
| **2001** | 222.25 |
| **2002** | 240.7231 |
| **2003** | 241.28 |
| **2004** | 261.6 |
| **2005** | 281.6 |
| **2006** | 281.6 |
| **2007** | 281.6 |
| **2008** | 324.7333 |
| **2009** | 280.7 |
| **2010** | 312.3 |



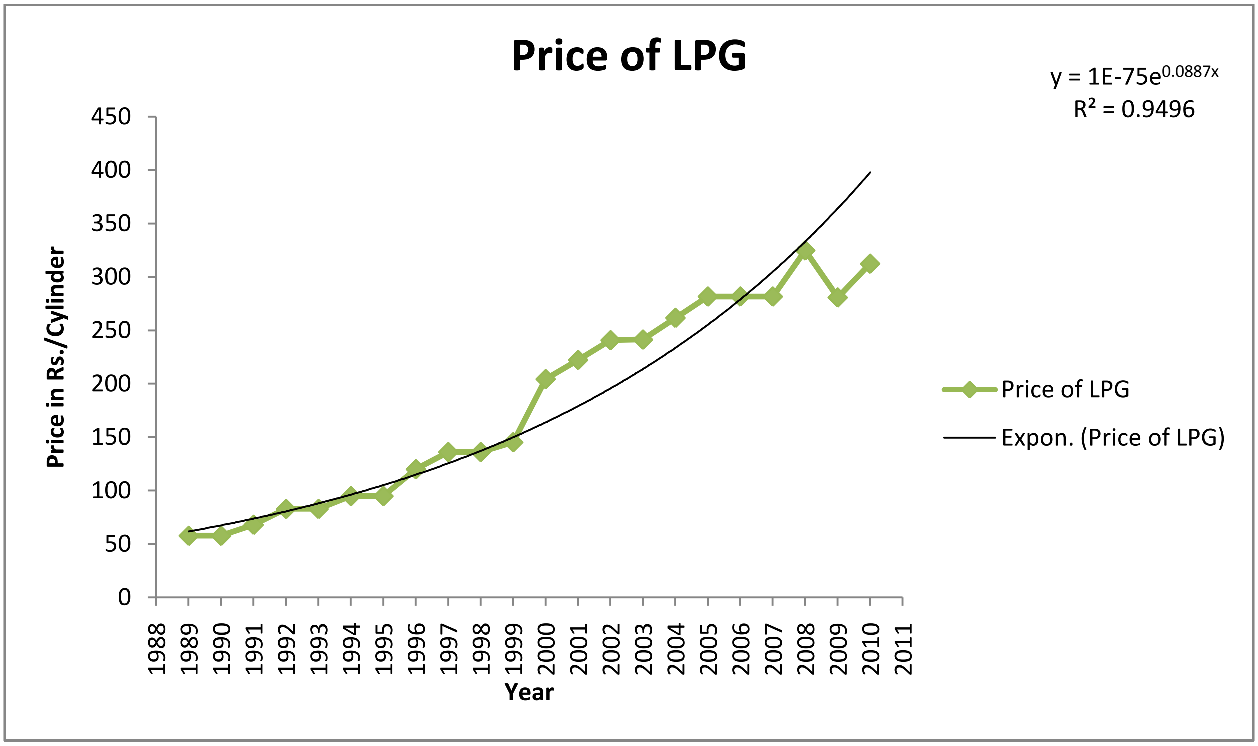

## TREND LINE:

We fit the following Trend Lines to the above data:

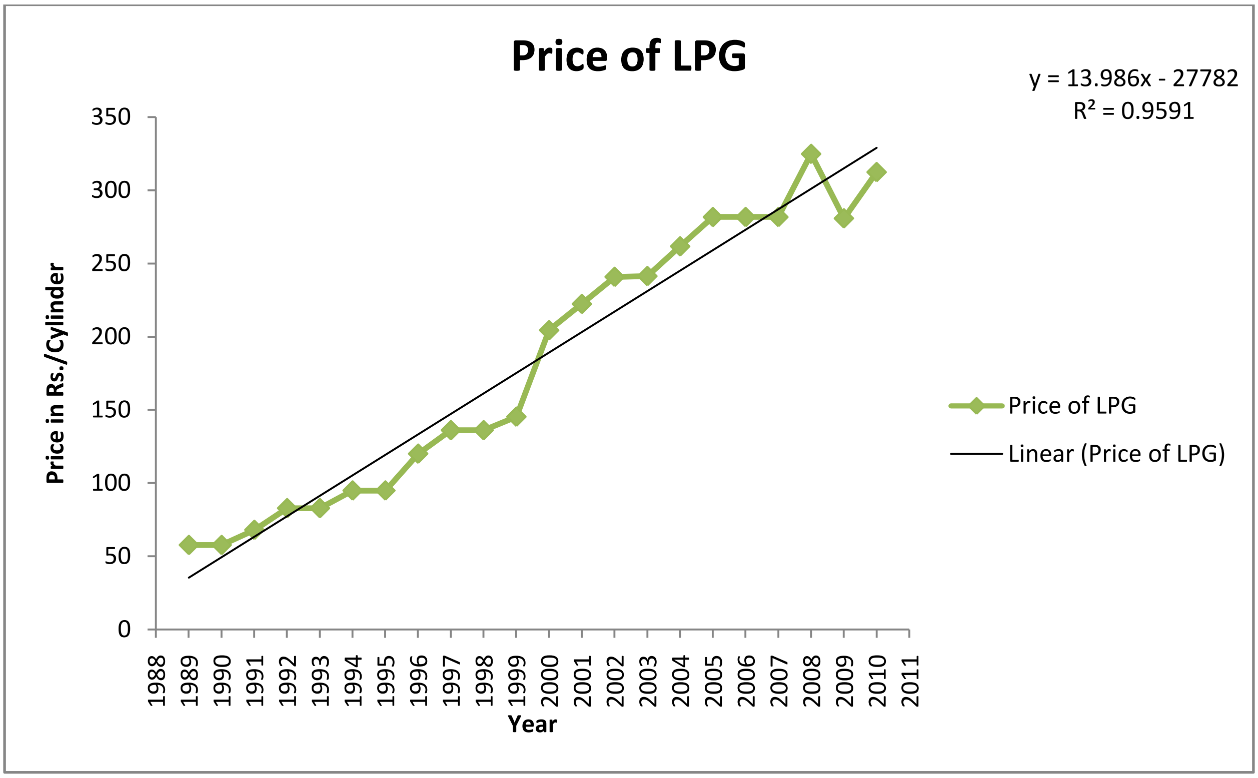
1. Logarithmic-



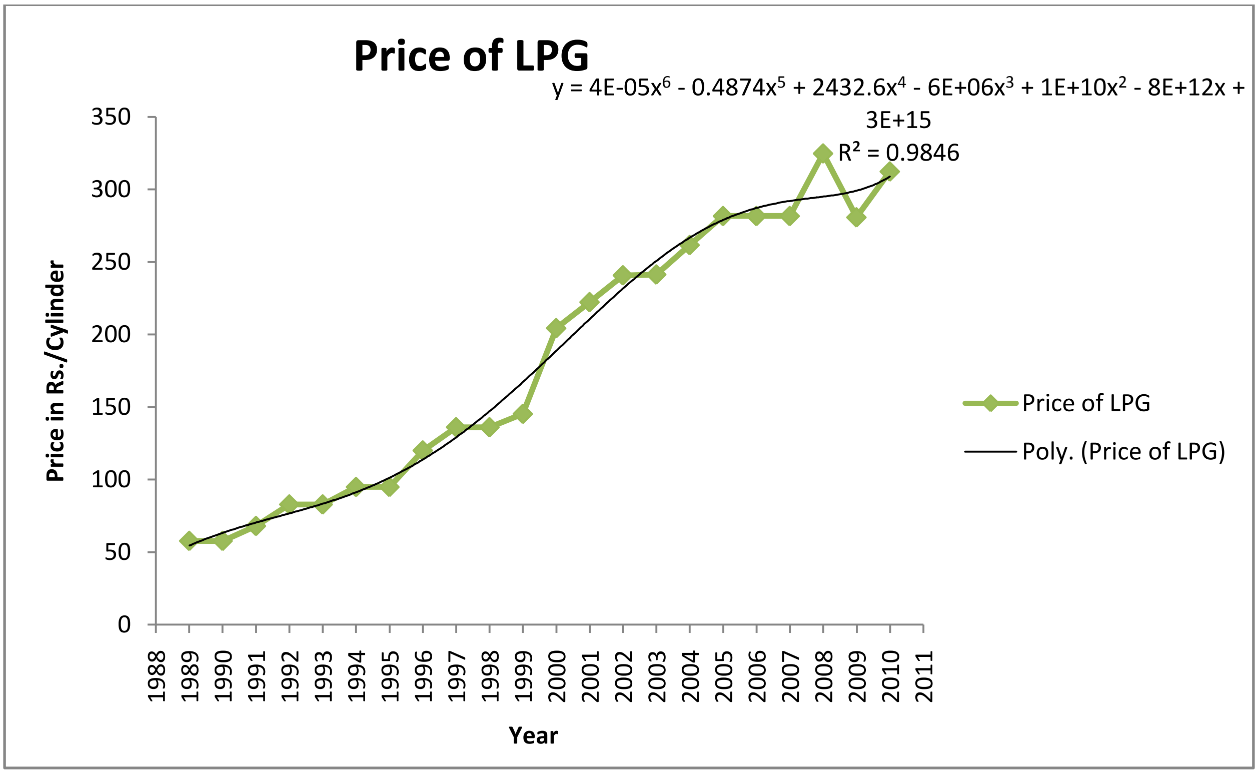
## 2. Exponential-




## 3. Linear-



4. Polynomial of degree six-




## EQUATIONS:

Hence we obtain the equations of the fit from the graphs as follows:

| Trend Line | Equation |
|---|---|
| Exponential | $y = 1E\text{-}75e^{0.088x}$ |
| Linear | y = 13.98x - 27782 |
| Logarithmic | y = 27964ln(x) - 21235 |
| Polynomial | $y = 4E\text{-}05x^6 - 0.487x^5 + 2432.x^4 - 6E{+}06x^3 + 1E{+}10x^2 - 8E{+}12x + 3E{+}15$ |

**ESTIMATES:**

Using these equations the LPG prices were estimated as follows:

| Year | Price of LPG | Estimates | | | |
|---|---|---|---|---|---|
| | | Linear | Exponential | Logarithmic | Polynomial |
| **1989** | 57.6 | 35.33592 | 53.54935556 | 191162.4099 | 4.81638E+15 |
| **1990** | 57.6 | 49.32144 | 58.47526021 | 191176.4657 | 4.82286E+15 |
| **1991** | 67.9 | 63.30696 | 63.85428958 | 191190.5144 | 4.82935E+15 |
| **1992** | 82.75 | 77.29248 | 69.72812576 | 191204.5561 | 4.83584E+15 |
| **1993** | 82.75 | 91.27799 | 76.1422851 | 191218.5907 | 4.84234E+15 |
| **1994** | 94.8 | 105.2635 | 83.14647091 | 191232.6183 | 4.84884E+15 |
| **1995** | 94.8 | 119.249 | 90.79495862 | 191246.6388 | 4.85536E+15 |
| **1996** | 119.95 | 133.2345 | 99.14701635 | 191260.6524 | 4.86187E+15 |
| **1997** | 136 | 147.2201 | 108.2673642 | 191274.6589 | 4.8684E+15 |
| **1998** | 136 | 161.2056 | 118.2266757 | 191288.6584 | 4.87492E+15 |
| **1999** | 145.2 | 175.1911 | 129.1021256 | 191302.6509 | 4.88146E+15 |
| **2000** | 204.3083 | 189.1766 | 140.9779876 | 191316.6364 | 4.888E+15 |
| **2001** | 222.25 | 203.1621 | 153.946288 | 191330.6149 | 4.89455E+15 |
| **2002** | 240.7231 | 217.1476 | 168.1075179 | 191344.5864 | 4.9011E+15 |
| **2003** | 241.28 | 231.1332 | 183.5714128 | 191358.551 | 4.90766E+15 |
| **2004** | 261.6 | 245.1187 | 200.4578023 | 191372.5085 | 4.91422E+15 |
| **2005** | 281.6 | 259.1042 | 218.8975391 | 191386.4591 | 4.9208E+15 |
| **2006** | 281.6 | 273.0897 | 239.0335126 | 191400.4028 | 4.92737E+15 |
| **2007** | 281.6 | 287.0752 | 261.0217565 | 191414.3395 | 4.93396E+15 |
| **2008** | 324.7333 | 301.0607 | 285.0326577 | 191428.2693 | 4.94054E+15 |
| **2009** | 280.7 | 315.0463 | 311.2522766 | 191442.1921 | 4.94714E+15 |
| **2010** | 312.3 | 329.0318 | 339.883789 | 191456.108 | 4.95374E+15 |

Here we observe that prices given by the Polynomial and Linear are not realistic hence we do not consider it.

**SUM OF SQUARES OF RESIDUALS (SSR):**

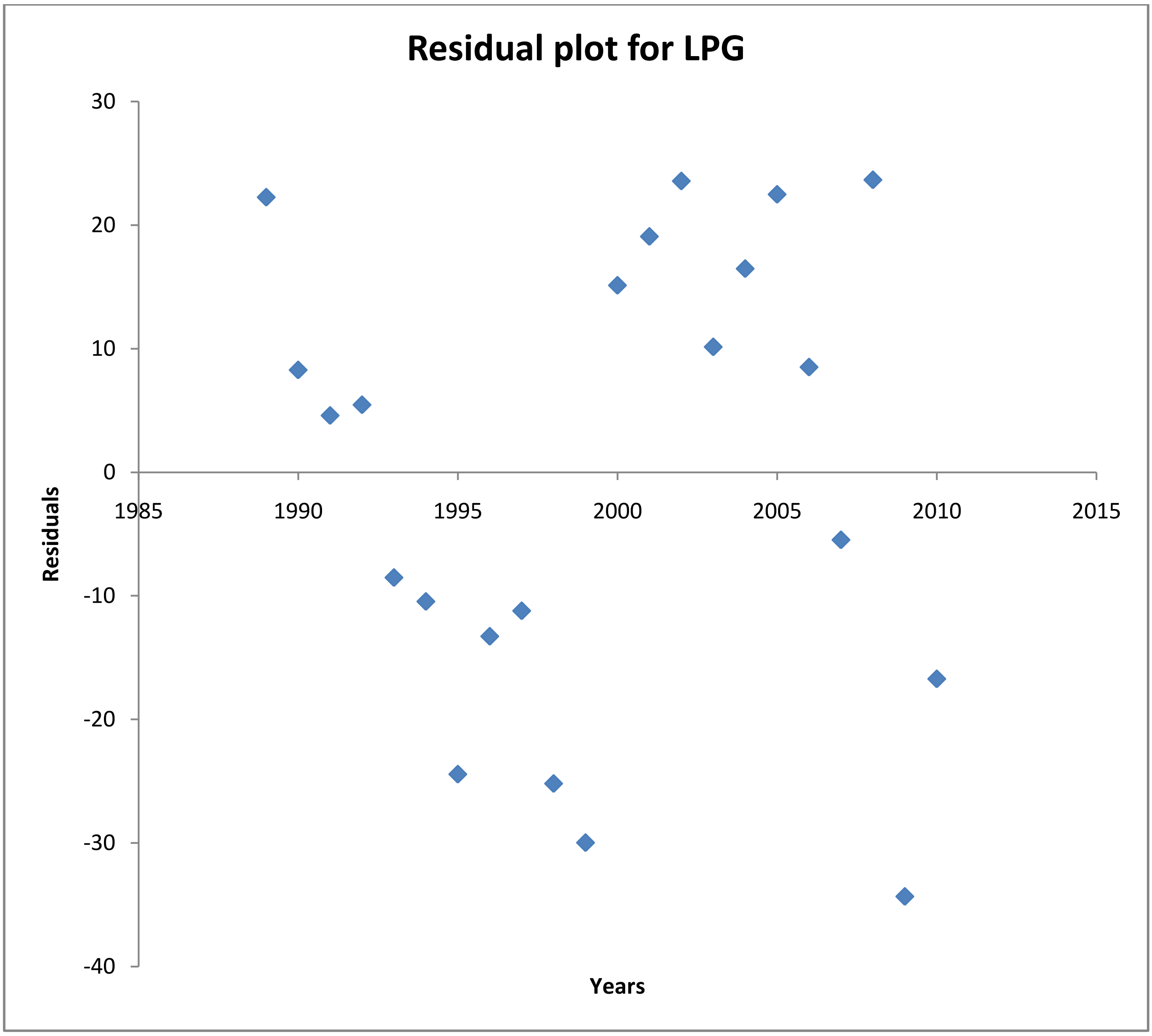
For the other fits we calculate the residuals and their SSR. The one with least SSR is considered as the best fit.

| Year | Price of LPG | Linear | | Exponential | |
|---|---|---|---|---|---|
| | | Residual | SSR | Residual | SSR |
| **1989** | 57.6 | 22.26408 | 495.689 | 4.050644 | 16.40772 |
| **1990** | 57.6 | 8.278558 | 68.53453 | -0.87526 | 0.76608 |
| **1991** | 67.9 | 4.593041 | 21.09603 | 4.04571 | 16.36777 |
| **1992** | 82.75 | 5.457525 | 29.78457 | 13.02187 | 169.5692 |
| **1993** | 82.75 | -8.52799 | 72.72665 | 6.607715 | 43.6619 |
| **1994** | 94.8 | -10.4635 | 109.485 | 11.65353 | 135.8047 |
| **1995** | 94.8 | -24.449 | 597.7549 | 4.005041 | 16.04036 |
| **1996** | 119.95 | -13.2845 | 176.4791 | 20.80298 | 432.7641 |
| **1997** | 136 | -11.2201 | 125.8897 | 27.73264 | 769.0991 |
| **1998** | 136 | -25.2056 | 635.3211 | 17.77332 | 315.8911 |
| **1999** | 145.2 | -29.9911 | 899.4657 | 16.09787 | 259.1416 |
| **2000** | 204.3083333 | 15.13172 | 228.969 | 63.33035 | 4010.733 |
| **2001** | 222.25 | 19.08787 | 364.3469 | 68.30371 | 4665.397 |
| **2002** | 240.7230769 | 23.57543 | 555.801 | 72.61556 | 5273.019 |
| **2003** | 241.28 | 10.14684 | 102.9583 | 57.70859 | 3330.281 |
| **2004** | 261.6 | 16.48132 | 271.634 | 61.1422 | 3738.368 |
| **2005** | 281.6 | 22.49581 | 506.0613 | 62.70246 | 3931.599 |
| **2006** | 281.6 | 8.510288 | 72.42501 | 42.56649 | 1811.906 |
| **2007** | 281.6 | -5.47523 | 29.97813 | 20.57824 | 423.4641 |
| **2008** | 324.7333333 | 23.67259 | 560.3914 | 39.70068 | 1576.144 |
| **2009** | 280.7 | -34.3463 | 1179.666 | -30.5523 | 933.4416 |
| **2010** | 312.3 | -16.7318 | 279.9524 | -27.5838 | 760.8654 |
| | **Total-** | | **7384.41** | | 32630.73 |

The SSR of Linear is the minimum. Hence Linear fit is the best fit for the LPG prices.

## RESIDUAL PLOT:

After fitting the appropriate Trend Line to the oil prices, the prices were estimated for the collected data using the equations. Then the residuals were obtained as the difference of the observed prices and the estimated prices. Then a graph of the year and the corresponding residual were plotted to obtain a Residual Plot to study the goodness of the fit.



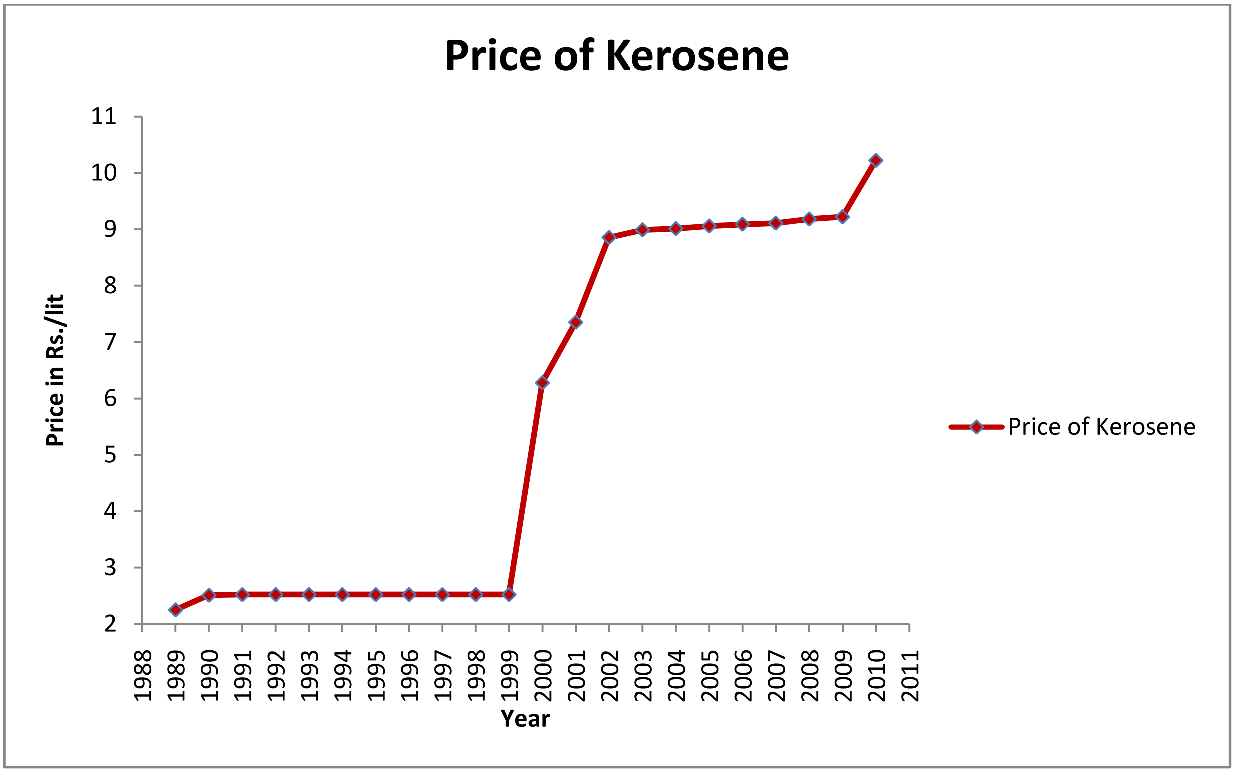

## CONCLUSION:

From the graphs we can conclude that the Linear fit for LPG is good as the Residual Plot does not show any pattern.

# KEROSENE:

**DATA:**

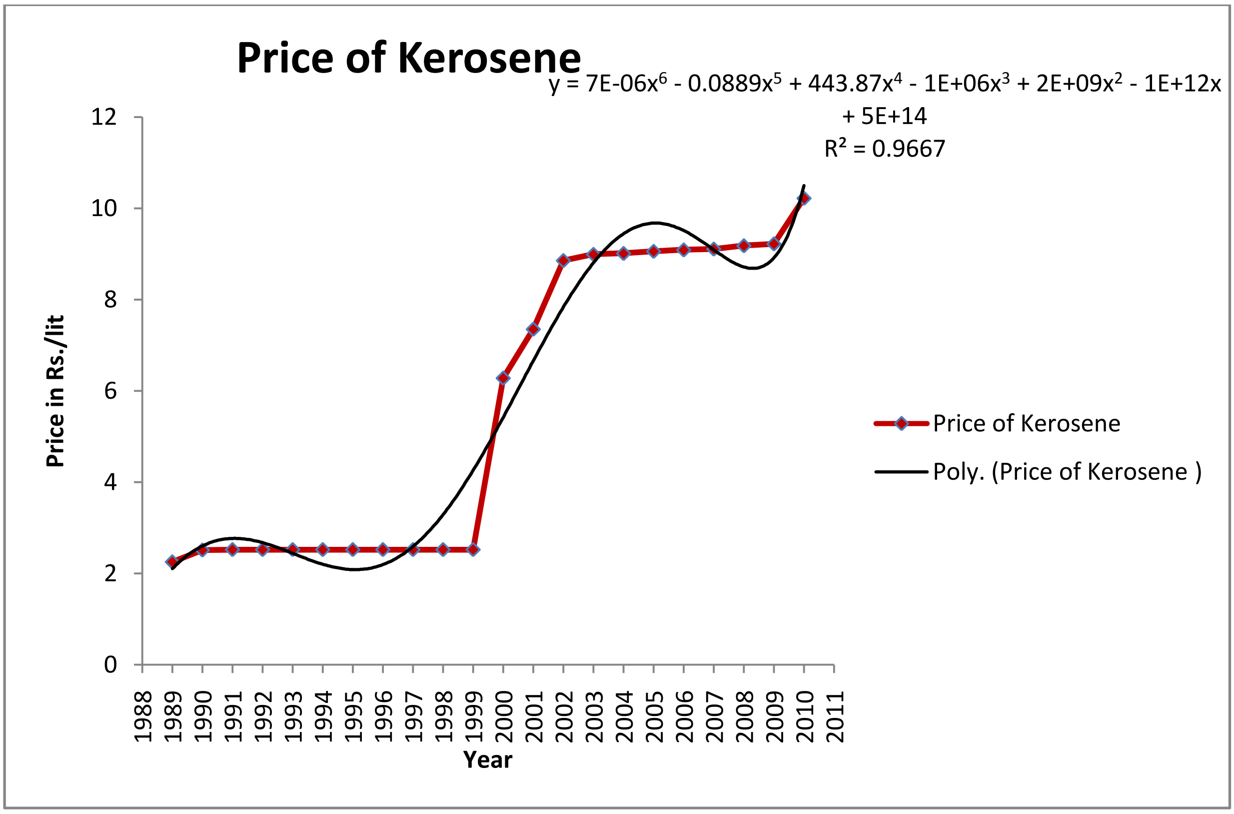
| Year | Price of Kerosene |
|---|---|
| **1989** | 2.25 |
| **1990** | 2.51 |
| **1991** | 2.52 |
| **1992** | 2.52 |
| **1993** | 2.52 |
| **1994** | 2.52 |
| **1995** | 2.52 |
| **1996** | 2.52 |
| **1997** | 2.52 |
| **1998** | 2.52 |
| **1999** | 2.52 |
| **2000** | 6.278333 |
| **2001** | 7.35 |
| **2002** | 8.854615 |
| **2003** | 8.99 |
| **2004** | 9.01 |
| **2005** | 9.056667 |
| **2006** | 9.0875 |
| **2007** | 9.1075 |
| **2008** | 9.18 |
| **2009** | 9.22 |
| **2010** | 10.22 |



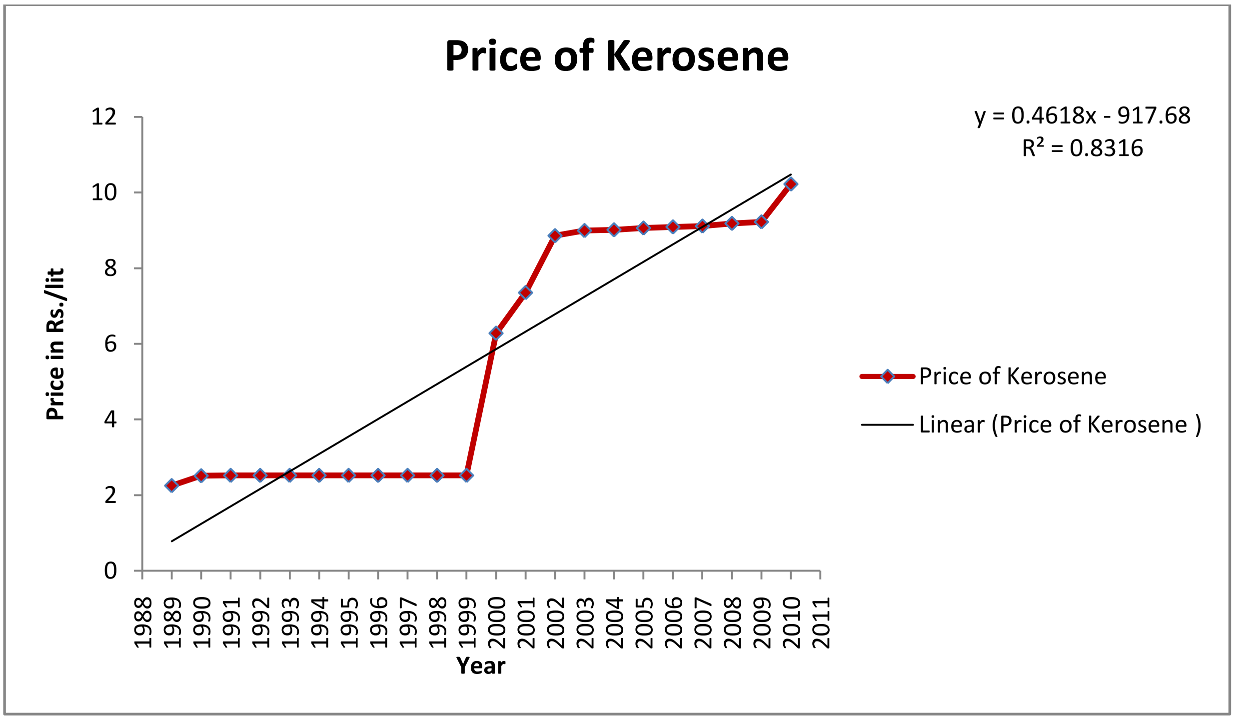

**TREND LINE:**

We fit the following Trend Lines to the above data:

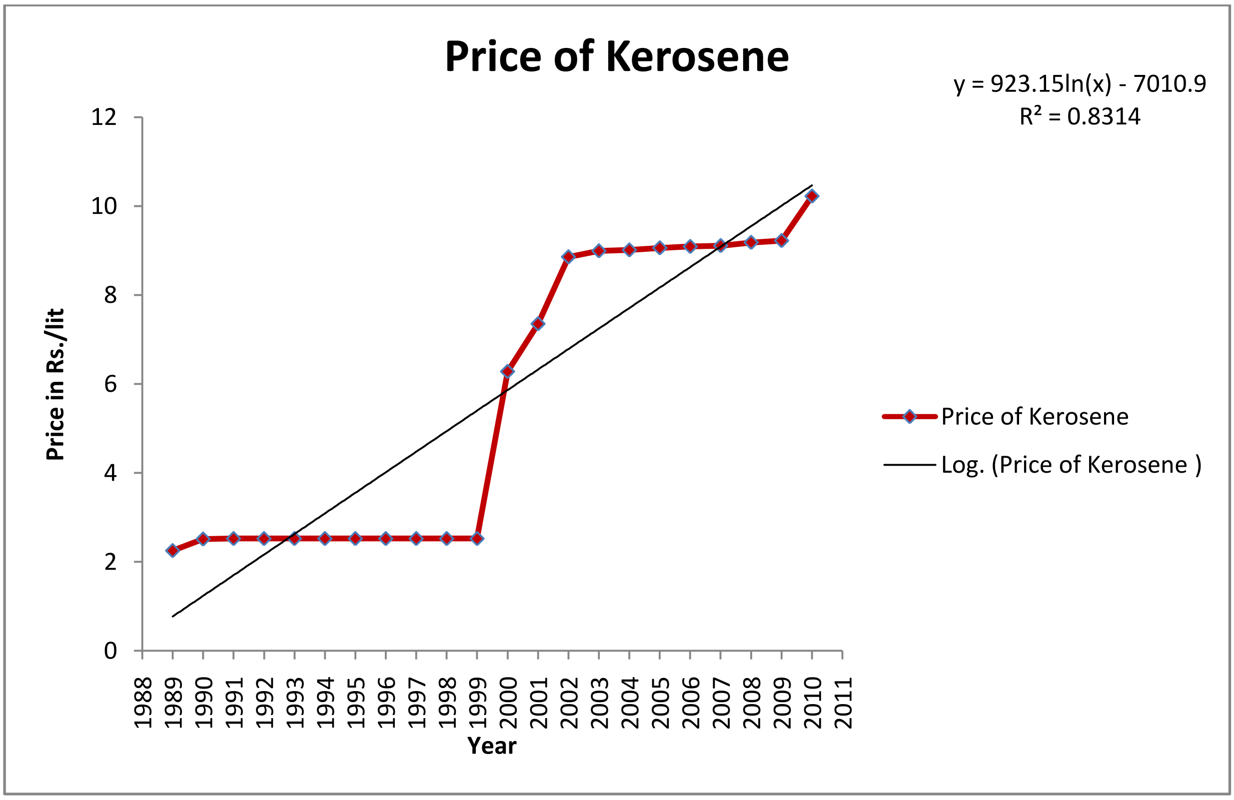
1. Polynomial of degree six-



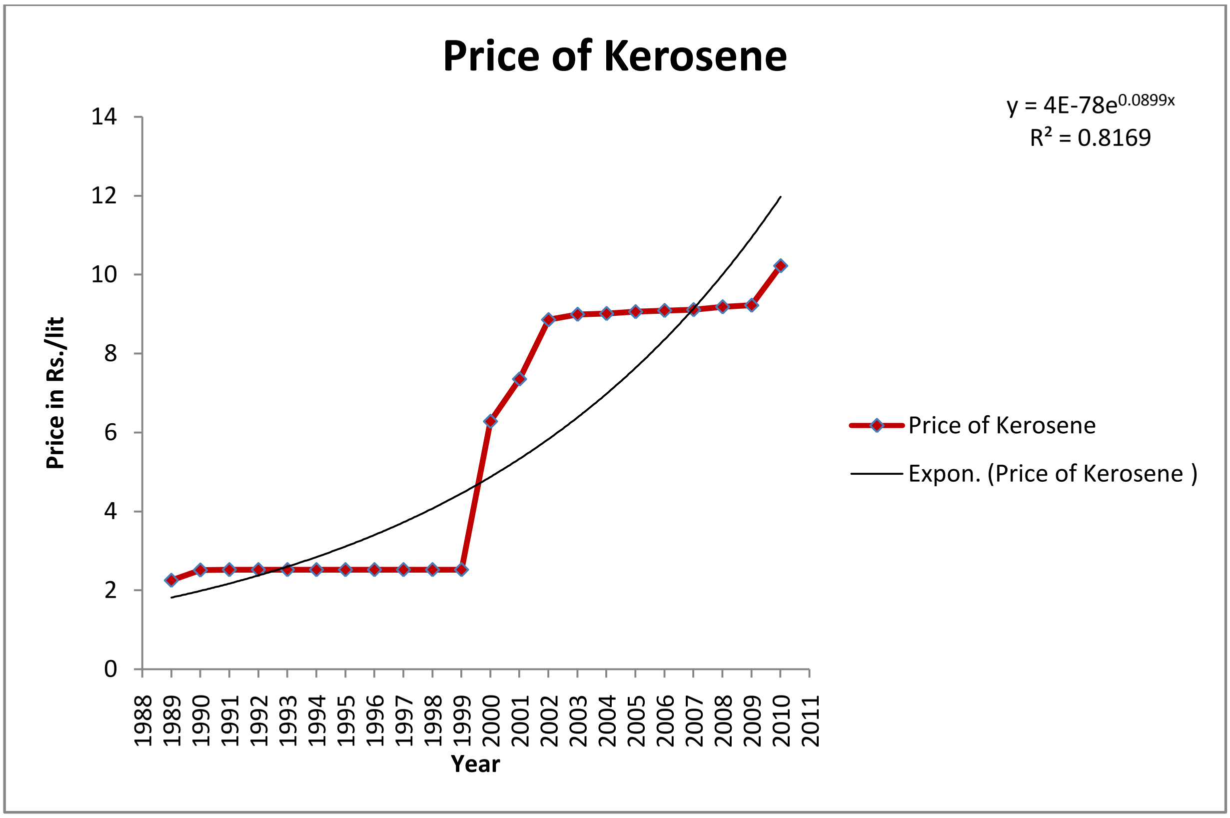
## 2. Linear-




## 3. Logarithmic-



4. Exponential-




## EQUATIONS:

Hence we obtain the equations of the fit from the graphs as follows:

| Trend Line | Equation |
|---|---|
| Exponential | $y = 4E\text{-}78e^{0.089x}$ |
| Linear | $y = 0.461x - 917.6$ |
| Logarithmic | $y = 923.1\ln(x) - 7010$ |
| Polynomial | $y = 7E\text{-}06x^6 - 0.088x^5 + 443.8x^4 - 1E{+}06x^3 + 2E{+}09x^2 - 1E{+}12x + 5E{+}14$ |

**ESTIMATES:**

Using these equations the Kerosene prices were estimated as follows:

| Year | Price of Kerosene | Estimates | | | |
|---|---|---|---|---|---|
| | | Linear | Logarithmic | Exponential | Polynomial |
| **1989** | 2.25 | 1.30199711 | 0.778442536 | 0.302866863 | 3.1944E+15 |
| **1990** | 2.51 | 1.765983046 | 1.240212577 | 0.33105789 | 3.1979E+15 |
| **1991** | 2.52 | 2.229735882 | 1.701982619 | 0.361872954 | 3.2014E+15 |
| **1992** | 2.52 | 2.693255851 | 2.16375266 | 0.395556302 | 3.2048E+15 |
| **1993** | 2.52 | 3.156543188 | 2.625522702 | 0.432374916 | 3.2083E+15 |
| **1994** | 2.52 | 3.619598127 | 3.087292743 | 0.472620629 | 3.2118E+15 |
| **1995** | 2.52 | 4.082420899 | 3.549062785 | 0.516612436 | 3.2153E+15 |
| **1996** | 2.52 | 4.545011738 | 4.010832827 | 0.564699027 | 3.2188E+15 |
| **1997** | 2.52 | 5.007370876 | 4.472602868 | 0.617261546 | 3.2223E+15 |
| **1998** | 2.52 | 5.469498545 | 4.93437291 | 0.674716616 | 3.2258E+15 |
| **1999** | 2.52 | 5.931394977 | 5.396142951 | 0.737519637 | 3.2293E+15 |
| **2000** | 6.278333 | 6.393060403 | 5.857912993 | 0.806168401 | 3.2328E+15 |
| **2001** | 7.35 | 6.854495054 | 6.319683034 | 0.881207032 | 3.2363E+15 |
| **2002** | 8.854615 | 7.315699161 | 6.781453076 | 0.963230303 | 3.2398E+15 |
| **200** | 8.99 | 7.7766729 | 7.2432231 | 1.0528883 | 3.2433E+ |

| | | | | | |
|---|---|---|---|---|---|
| **3** | | 53 | 17 | 49 | 15 |
| **2004** | 9.01 | 8.237416661 | 7.704993159 | 1.150891818 | 3.2468E+15 |
| **2005** | 9.056667 | 8.697930515 | 8.166763201 | 1.258017509 | 3.2504E+15 |
| **2006** | 9.0875 | 9.158214743 | 8.628533242 | 1.375114522 | 3.2539E+15 |
| **2007** | 9.1075 | 9.618269574 | 9.090303284 | 1.503110994 | 3.2574E+15 |
| **2008** | 9.18 | 10.07809524 | 9.552073325 | 1.643021453 | 3.2609E+15 |
| **2009** | 9.22 | 10.53769196 | 10.01384337 | 1.795954861 | 3.2645E+15 |
| **2010** | 10.22 | 10.99705997 | 10.47561341 | 1.963123401 | 3.268E+15 |

Here we observe that prices given by the Polynomial is not realistic hence we do not consider it.

## SUM OF SQUARES OF RESIDUALS (SSR):

For the other fits we calculate the residuals and their SSR.The one with least SSR is considered as the best fit.

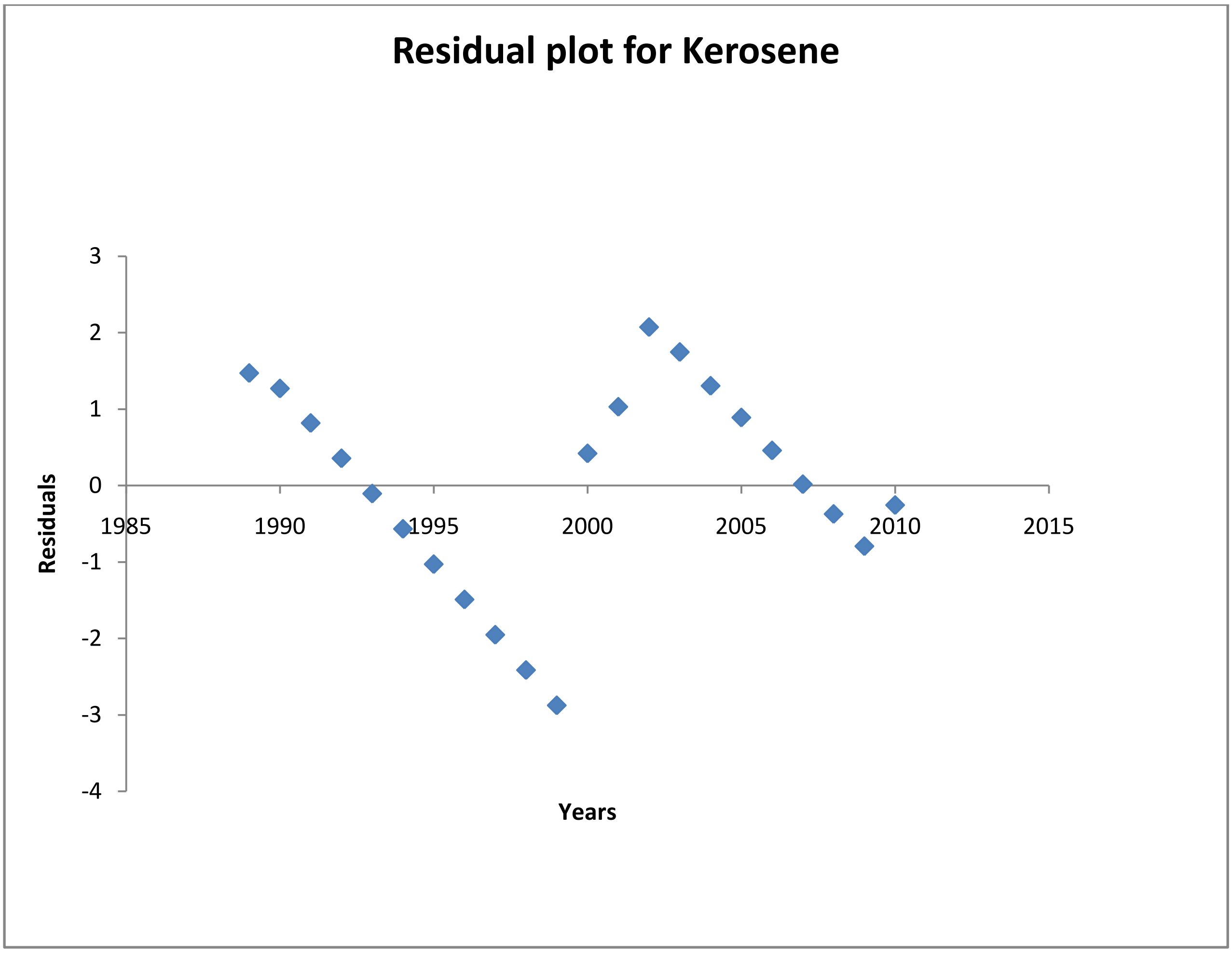
| Year | Price of Kerosene | Linear | | Logarithmic | | Exponential | |
|---|---|---|---|---|---|---|---|
| | | Residual | SSR | Residual | SSR | Residual | SSR |
| **1989** | 2.25 | 0.948003 | 0.898709479 | 1.471557 | 2.165481 | 0.302867 | 1.947133137 |
| **1990** | 2.51 | 0.744017 | 0.553561228 | 1.269787 | 1.61236 | 0.331058 | 2.17894211 |
| **1991** | 2.52 | 0.290264 | 0.084253258 | 0.818017 | 0.669152 | 0.361873 | 2.158127046 |
| **1992** | 2.52 | -0.17326 | 0.03001759 | 0.356247 | 0.126912 | 0.395556 | 2.124443698 |
| **1993** | 2.52 | -0.63654 | 0.405187231 | -0.10552 | 0.011135 | 0.432375 | 2.087625084 |
| **1994** | 2.52 | -1.0996 | 1.20911604 | -0.56729 | 0.321821 | 0.472621 | 2.047379371 |
| **1995** | 2.52 | -1.56242 | 2.441159065 | -1.02906 | 1.05897 | 0.516612 | 2.003387564 |
| **1996** | 2.52 | -2.02501 | 4.100672539 | -1.49083 | 2.222583 | 0.564699 | 1.955300973 |
| **1997** | 2.52 | -2.48737 | 6.187013875 | -1.9526 | 3.812658 | 0.617262 | 1.902738454 |
| **1998** | 2.52 | -2.9495 | 8.699541669 | -2.41437 | 5.829197 | 0.674717 | 1.845283384 |
| **19** | 2.52 | - | 11.6376 | - | 8.2721 | 0.737 | 1.78248 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **99** | | 3.41139 | 1569 | 2.87614 | 98 | 52 | 0363 |
| **2000** | 6.278333 | -0.11473 | 0.013162301 | 0.42042 | 0.176753 | 0.806168 | 5.472164933 |
| **2001** | 7.35 | 0.495505 | 0.245525151 | 1.030317 | 1.061553 | 0.881207 | 6.468792968 |
| **2002** | 8.854615 | 1.538916 | 2.368263144 | 2.073162 | 4.298002 | 0.96323 | 7.891385082 |
| **2003** | 8.99 | 1.213327 | 1.472162523 | 1.746777 | 3.051229 | 1.052888 | 7.937111651 |
| **2004** | 9.01 | 0.772583 | 0.596885015 | 1.305007 | 1.703043 | 1.150892 | 7.859108182 |
| **2005** | 9.056667 | 0.358736 | 0.128691627 | 0.889903 | 0.791928 | 1.258018 | 7.798649158 |
| **2006** | 9.0875 | -0.07071 | 0.005000575 | 0.458967 | 0.21065 | 1.375115 | 7.712385478 |
| **2007** | 9.1075 | -0.51077 | 0.260885558 | 0.017197 | 0.000296 | 1.503111 | 7.604389006 |
| **2008** | 9.18 | -0.8981 | 0.806575055 | -0.37207 | 0.138439 | 1.643021 | 7.536978547 |
| **2009** | 9.22 | -1.31769 | 1.736312103 | -0.79384 | 0.630187 | 1.795955 | 7.424045139 |
| **2010** | 10.22 | -0.77706 | 0.6038222 | -0.25561 | 0.065338 | 1.963123 | 8.256876599 |
| | **Total** | | 44.48413292 | | **38.22989** | | 103.9947279 |

The SSR of Logarithmic is the minimum. Hence <u>Logarithmic fit is the best fit</u> for the Kerosene prices.

## RESIDUAL PLOT:

After fitting the appropriate Trend Line to the oil prices, the prices were estimated for the collected data using the equations. Then the residuals were obtained as the difference of the observed prices and the estimated prices. Then a graph of the year and the corresponding residual were plotted to obtain a Residual Plot to study the goodness of the fit.




## CONCLUSION:

The Residual Plot of Diesel shows a certain pattern, which means that the fit is not very good. But due to the limitations of Microsoft Excel, other fits, which could be better, cannot be used for the estimation. Hence we proceed with Logarithmic Fit as the best fit for Kerosene Prices.

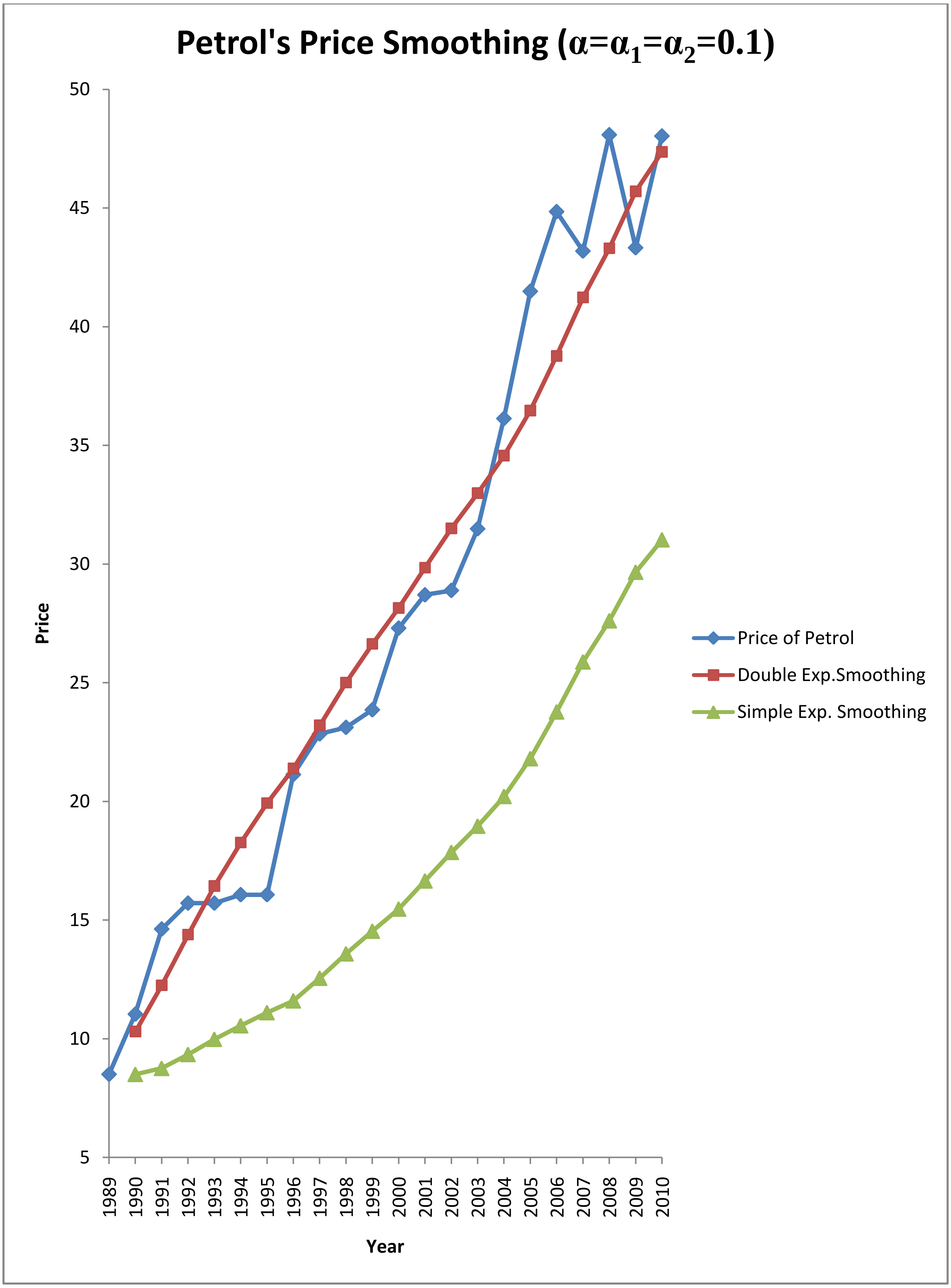
# DOUBLE EXPONENTIAL SMOOTHING

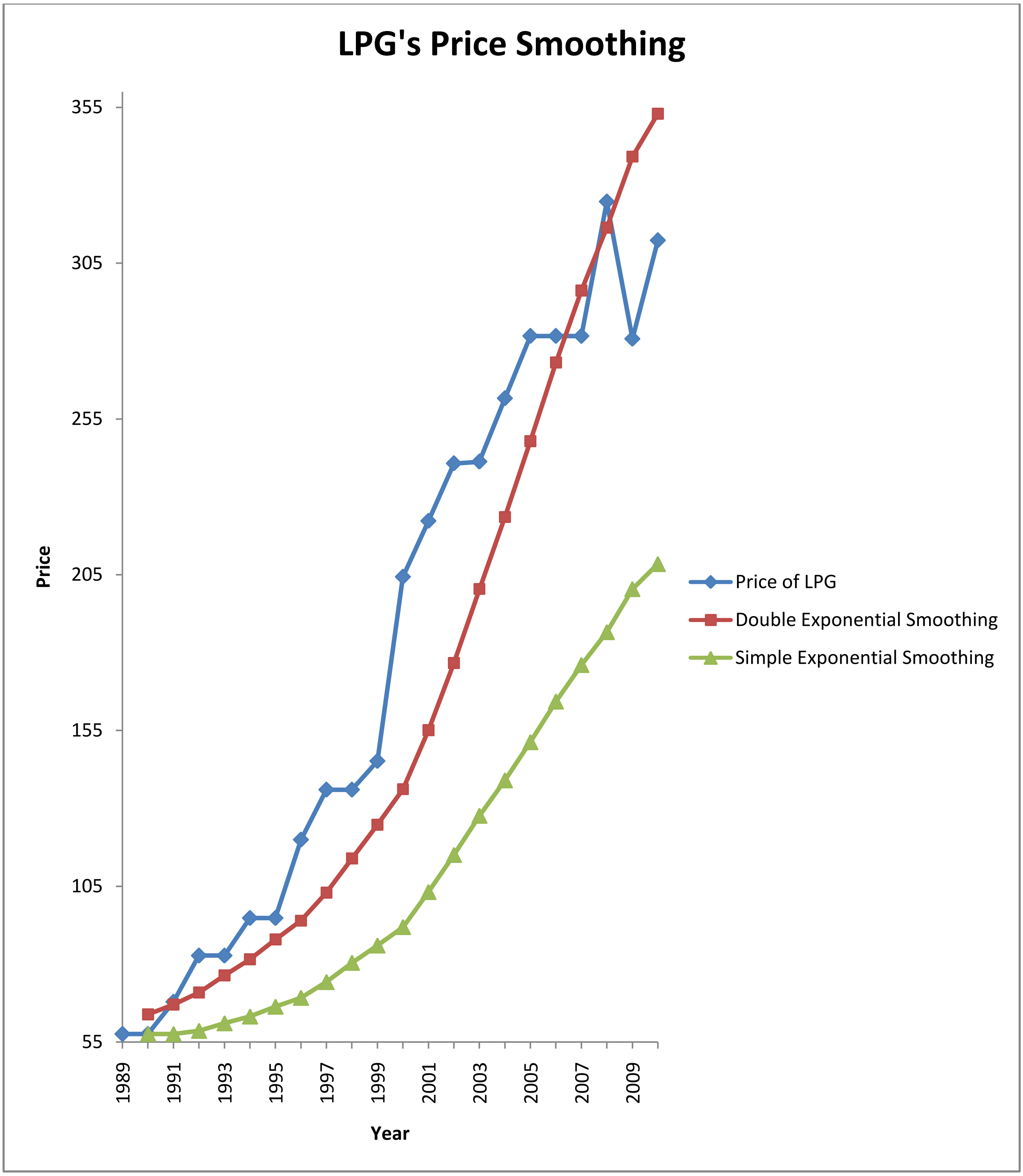
The Prices of Petrol & LPG have a Linear increasing trend & no seasonal variation, hence Double Exponential Smoothing Procedure can be applied to Petrol and LPG's Prices to produce smoothed values of Prices. Simple Exponential Smoothing (SES) requires time series to be stagnant; therefore Double Exponential Smoothing (DES) is applied below.

For both the time series mentioned above the value of smoothing constant is 0.1 for Simple Exponential Smoothing. For Double Exponential Smoothing $\alpha_1=\alpha_2=0.1$ for Petrol's Price Smoothing & $\alpha_1=0.1$, $\alpha_2=0.4$ for LPG's Price Smoothing.

| Year | Price of Petrol | DES | SES | Price of LPG | DES | SES |
|---|---|---|---|---|---|---|
| **1989** | 8.5 | __ | __ | 57.6 | __ | __ |
| **1990** | 11.035 | 10.3025 | 8.5 | 57.6 | 63.8875 | 57.6 |
| **1991** | 14.62 | 12.2515 | 8.7535 | 67.9 | 67.03125 | 57.6 |
| **1992** | 15.71 | 14.387785 | 9.34015 | 82.75 | 70.925375 | 58.63 |
| **1993** | 15.71 | 16.43266365 | 9.977135 | 82.75 | 76.3880725 | 61.042 |
| **1994** | 16.0666667 | 18.2658278 | 10.5504215 | 94.8 | 81.5589774 | 63.2128 |
| **1995** | 16.0666667 | 19.92935059 | 11.10204602 | 94.8 | 87.9474326 | 66.37152 |
| **1996** | 21.13 | 21.38789426 | 11.59850808 | 119.95 | 93.9711451 | 69.214368 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **1997** | 22.84 | 23.20433795 | 12.55165727 | 136 | 102.94664 | 74.2879312 |
| **1998** | 23.115 | 25.0064939 | 13.58049155 | 136 | 113.951721 | 80.45913808 |
| **1999** | 23.856 | 26.63701931 | 14.53394239 | 145.2 | 124.738224 | 86.01322427 |
| **2000** | 27.2983333 | 28.15078199 | 15.46614815 | 204.3083 | 136.184548 | 91.93190184 |
| **2001** | 28.7 | 29.84887724 | 16.64936667 | 222.25 | 155.122025 | 103.169545 |
| **2002** | 28.8838462 | 31.50584087 | 17.85443 | 240.7231 | 176.645039 | 115.0775905 |
| **2003** | 31.4826667 | 32.9892728 | 18.95737162 | 241.28 | 200.426181 | 127.6421391 |
| **2004** | 36.1283333 | 34.56917753 | 20.20990112 | 261.6 | 223.519054 | 139.0059252 |
| **2005** | 41.49 | 36.47125001 | 21.80174434 | 281.6 | 247.857878 | 151.2653327 |
| **2006** | 44.845 | 38.76946941 | 23.77056991 | 281.6 | 273.112504 | 164.2987994 |
| **2007** | 43.185 | 41.23412217 | 25.87801292 | 281.6 | 296.181167 | 176.0289195 |
| **2008** | 48.0833333 | 43.30581844 | 27.60871163 | 324.7333 | 316.359718 | 186.5860275 |
| **2009** | 43.3233333 | 45.70795356 | 29.6561738 | 280.7 | 339.168691 | 200.4007581 |
| **2010** | 48.03 | 47.37002897 | 31.02288975 | 312.3 | 352.954686 | 208.4306823 |


Petrol's Price Smoothing (α=α1=α2=0.1)
50
45
40
35
30
25
20
15
10
5
Price
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
Year
Price of Petrol
Double Exp.Smoothing
Simple Exp. Smoothing




## CONCLUSION:

It can be seen that Single Exponential Smoothing does not follow the data well when there is a trend in the data hence this drawback of Simple Exponential Smoothing is overcome by Double Exponential Smoothing.

# ESTIMATION OF FUEL PRICES

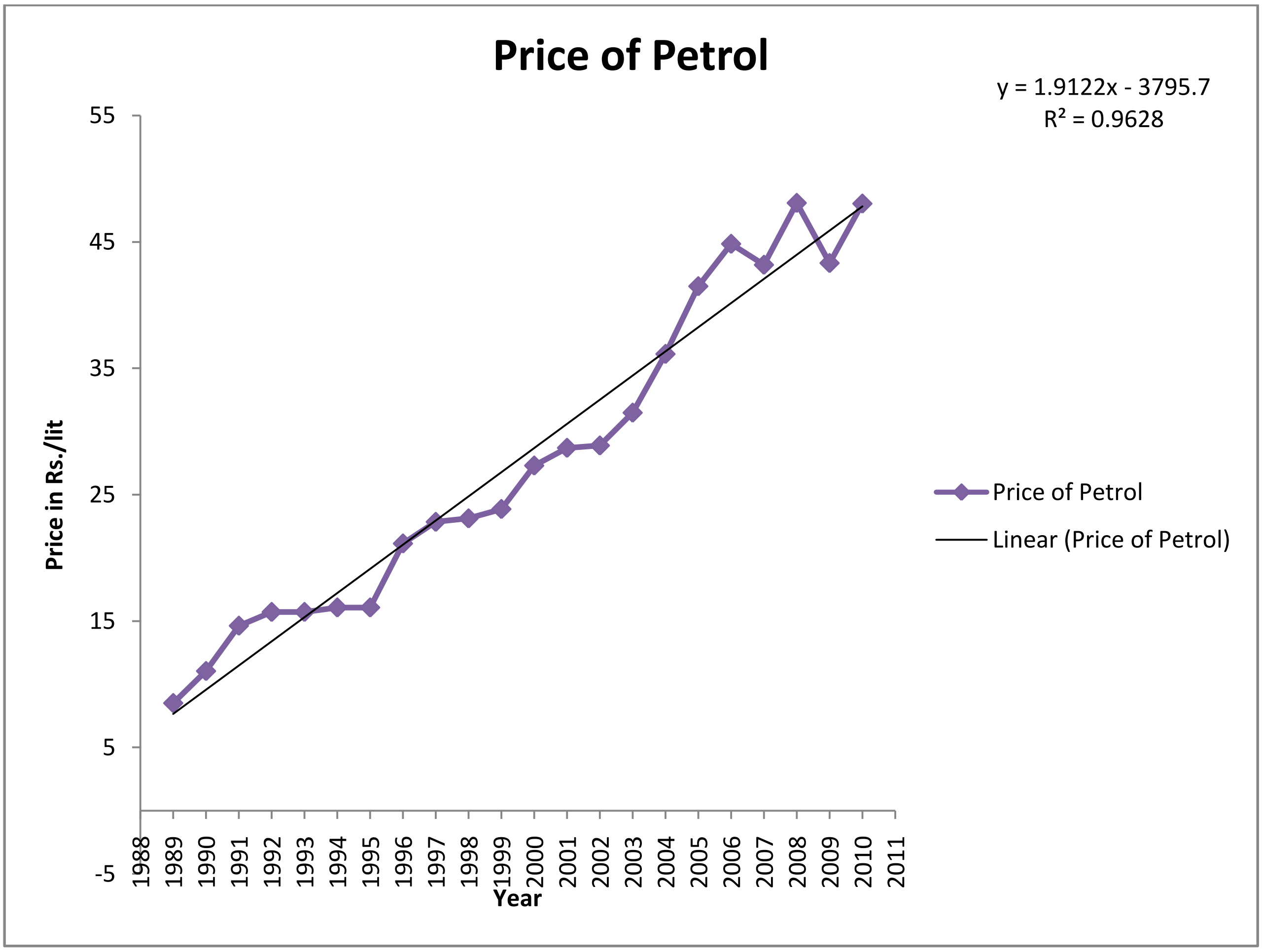
For estimation purpose we select an equation which gives a realistic value of price of the fuel. Selecting an equation which has a good R^2 value does not always give realistic value of the price. Hence we choose the following equations for the following fuels:

| *FUEL* | *Equation Used* |
|---|---|
| Petrol | Linear |
| Diesel | Logarithmic |
| LPG | Linear |
| Kerosene | Logarithmic |

# PETROL

For estimation of Petrol prices in the following years after 2011 we fit a Linear function to the data of Petrol prices.

We choose this fit because it gave realistic values of prices and its Sum of Squares of Residual is Minimum for this fit. Hence we get the following graph:




The equation of this Exponential function is

$$y = 1.912x - 3795$$

Till 2011 the observed Petrol price and the estimated prices are as follows

| Year | Observed Prices | Estimated Price |
|---|---|---|
| **1989** | 8.5 | 7.6658 |
| **1990** | 11.035 | 9.578 |
| **1991** | 14.62 | 11.4902 |
| **1992** | 15.71 | 13.4024 |
| **1993** | 15.71 | 15.3146 |
| **1994** | 16.06666667 | 17.2268 |
| **1995** | 16.06666667 | 19.139 |
| **1996** | 21.13 | 21.0512 |
| **1997** | 22.84 | 22.9634 |
| **1998** | 23.115 | 24.8756 |
| **1999** | 23.856 | 26.7878 |
| **2000** | 27.29833333 | 28.7 |
| **2001** | 28.7 | 30.6122 |
| **2002** | 28.88384615 | 32.5244 |
| **2003** | 31.48266667 | 34.4366 |
| **2004** | 36.12833333 | 36.3488 |
| **2005** | 41.49 | 38.261 |
| **2006** | 44.845 | 40.1732 |

| | | |
|---|---|---|
| **2007** | 43.185 | 42.0854 |
| **2008** | 48.08333333 | 43.9976 |
| **2009** | 43.32333333 | 45.9098 |
| **2010** | 48.03 | 47.822 |

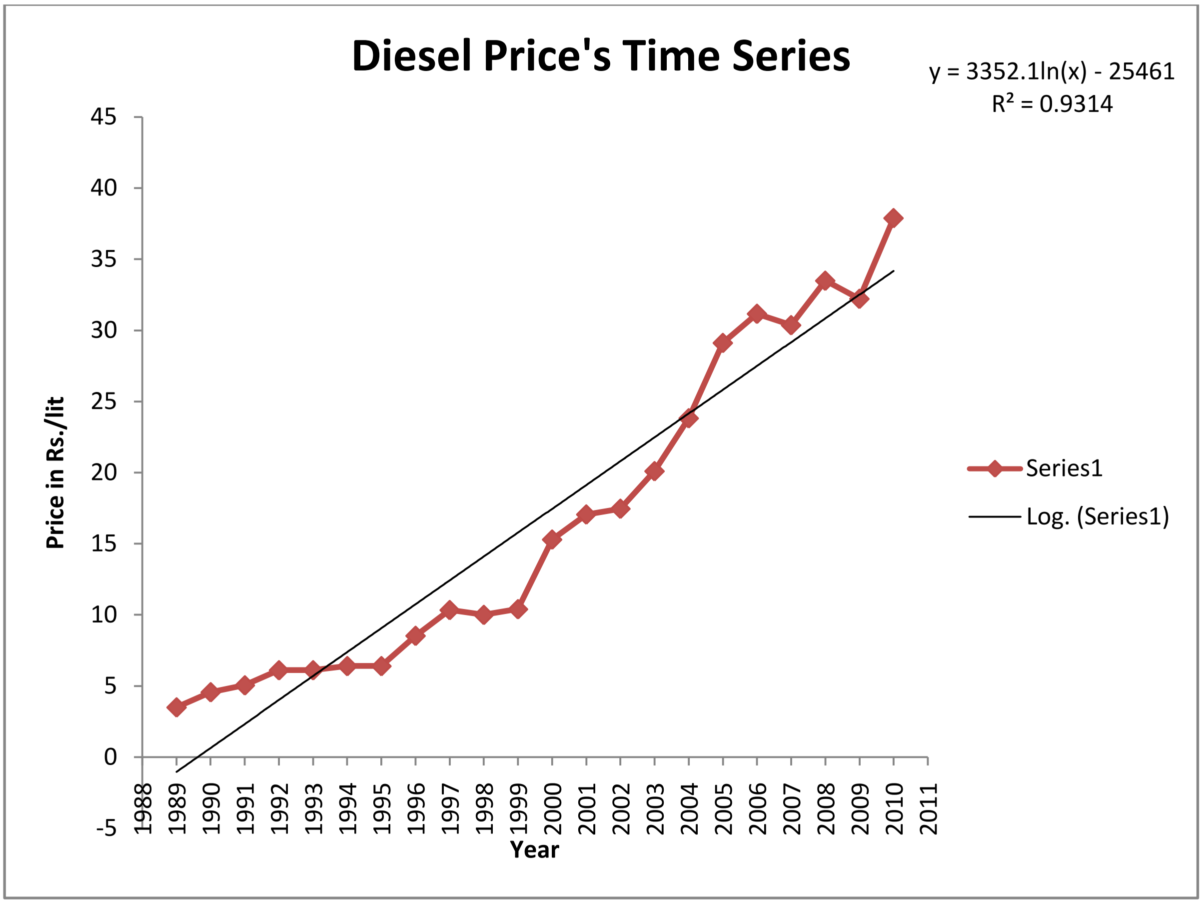
For the successive years the Petrol prices are predicted to be as follows:

| *Year* | *Estimated price (in Rs.)* |
|---|---|
| ***2011*** | ***49.7342*** |
| ***2012*** | ***51.6464*** |
| ***2013*** | ***53.5586*** |
| ***2014*** | ***55.4708*** |
| ***2015*** | ***57.383*** |
| ***2016*** | ***59.2952*** |
| ***2017*** | ***61.2074*** |
| ***2018*** | ***63.1196*** |
| ***2019*** | ***65.0318*** |
| ***2020*** | ***66.944*** |

# <u>DIESEL:</u>

For estimation of Diesel prices in the following years after 2011 we fit a Logarithmic function to the data of Diesel prices.

We choose this fit because it gave realistic values of prices and its Sum of Squares of Residual is Minimum for this fit. Hence we get the following graph:




The equation of this Linear function is

$$y = 3352.\ln(x) - 25461$$

Till 2011 the observed Diesel price and the estimated prices are as follows:

| Year | Observed prices | Estimated prices (in Rs.) |
|---|---|---|
| **1989** | 3.5 | -1.26184 |
| **1990** | 4.565 | 0.42300 |
| **1991** | 5.05 | 2.107003 |
| **1992** | 6.11 | 3.790157 |
| **1993** | 6.11 | 5.472465 |
| **1994** | 6.4 | 7.15393 |
| **1995** | 6.4 | 8.834552 |
| **1996** | 8.53 | 10.51433 |
| **1997** | 10.34 | 12.19327 |
| **1998** | 10 | 13.87105 |
| **1999** | 10.4 | 15.54863 |
| **2000** | 15.295 | 17.22504 |
| **2001** | 17.06 | 18.90063 |
| **2002** | 17.46384615 | 20.57537 |
| **2003** | 20.10266667 | 22.24928 |
| **2004** | 23.82166667 | 23.92235 |
| **2005** | 29.11666667 | 25.59459 |

| | | |
|---|---|---|
| **2006** | 31.165 | 27.26599 |
| **2007** | 30.365 | 28.93656 |
| **2008** | 33.49 | 30.6063 |
| **2009** | 32.21666667 | 32.27521 |
| **2010** | 37.89 | 33.94328 |

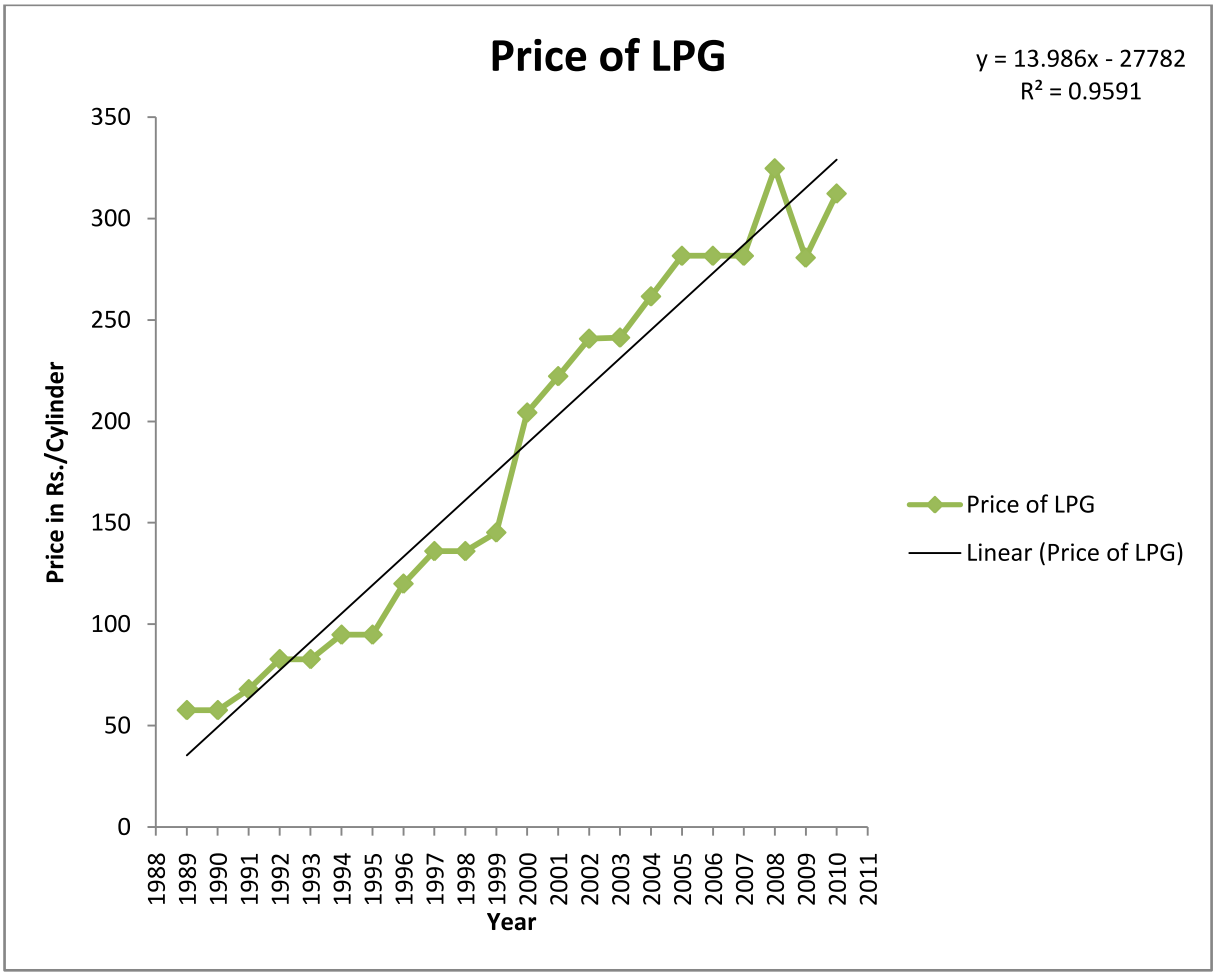
For the successive years the Diesel prices are predicted to be as follows:

| ***Year*** | ***Estimated price (in Rs.)*** |
|---|---|
| ***2011*** | ***35.61053052*** |
| ***2012*** | ***37.27694865*** |
| ***2013*** | ***38.94253874*** |
| ***2014*** | ***40.60730163*** |
| ***2015*** | ***42.27123812*** |
| ***2016*** | ***43.93434905*** |
| ***2017*** | ***45.59663522*** |
| ***2018*** | ***47.25809746*** |
| ***2019*** | ***48.91873658*** |
| ***2020*** | ***50.5785534*** |

# LPG:

For estimation of LPG prices in the following years after 2011 we fit a Linear function to the data of LPG prices.

We choose this fit because it gave realistic values of prices and its Sum of Squares of Residual is Minimum for this fit. Hence we get the following graph:




The equation of this Linear function is

$$y = 13.98x - 27782$$

Till 2011 the observed LPG price and the estimated prices are as follows:

| Year | Observed Prices | Estimate Prices |
|---|---|---|
| **1989** | 57.6 | 36.154 |
| **1990** | 57.6 | 50.14 |
| **1991** | 67.9 | 64.126 |
| **1992** | 82.75 | 78.112 |
| **1993** | 82.75 | 92.098 |
| **1994** | 94.8 | 106.084 |
| **1995** | 94.8 | 120.07 |
| **1996** | 119.95 | 134.056 |
| **1997** | 136 | 148.042 |
| **1998** | 136 | 162.028 |
| **1999** | 145.2 | 176.014 |
| **2000** | 204.3083333 | 190 |
| **2001** | 222.25 | 203.986 |
| **2002** | 240.7230769 | 217.972 |
| **2003** | 241.28 | 231.958 |
| **2004** | 261.6 | 245.944 |
| **2005** | 281.6 | 259.93 |
| **2006** | 281.6 | 273.916 |

| | | |
|---|---|---|
| **2007** | 281.6 | 287.902 |
| **2008** | 324.7333333 | 301.888 |
| **2009** | 280.7 | 315.874 |
| **2010** | 312.3 | 329.86 |

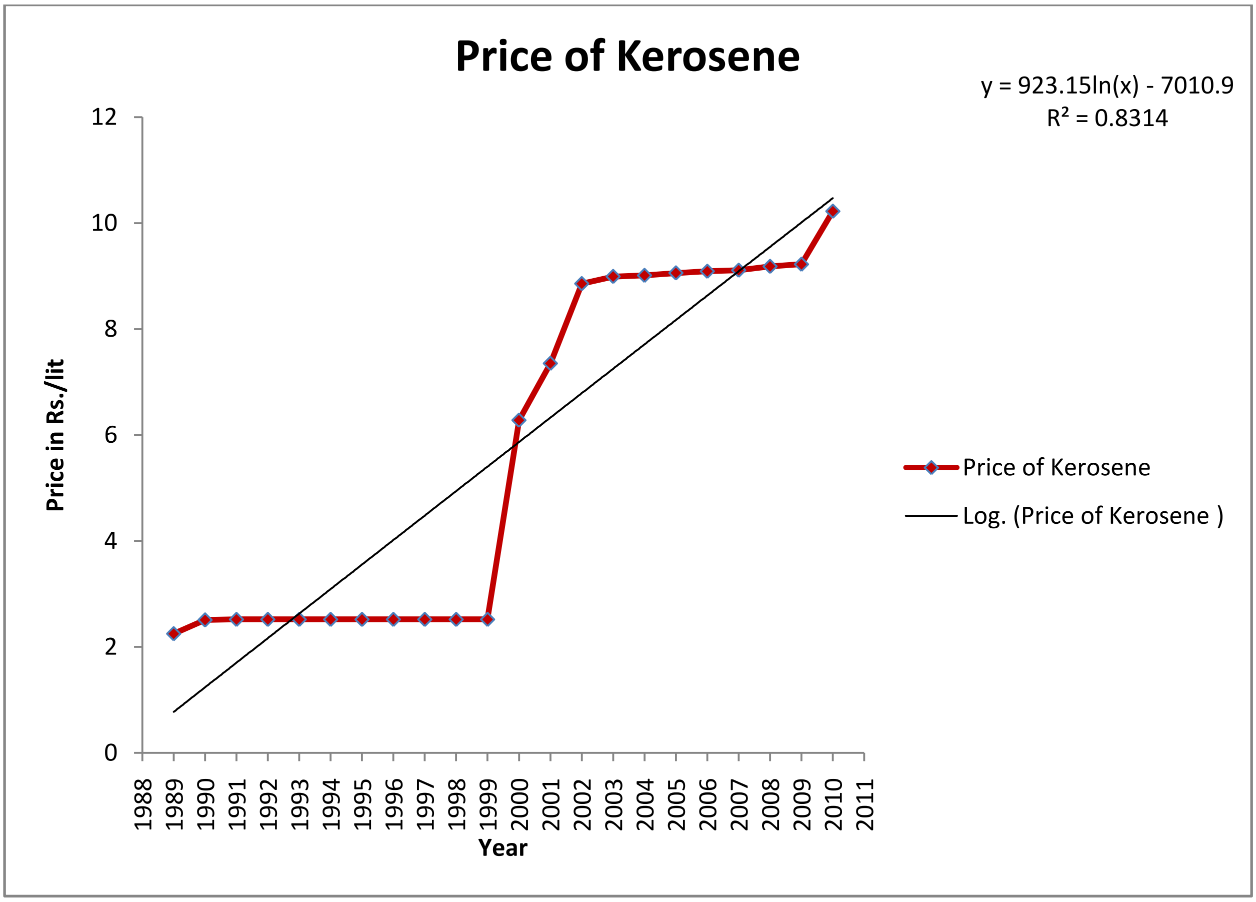
For the successive years the LPG prices are predicted to be as follows:

| *Year* | *Estimated price (in Rs.)* |
|---|---|
| **2011** | **343.846** |
| **2012** | **357.832** |
| **2013** | **371.818** |
| **2014** | **385.804** |
| **2015** | **399.79** |
| **2016** | **413.776** |
| **2017** | **427.762** |
| **2018** | **441.748** |
| **2019** | **455.734** |
| **2020** | **469.72** |

# <u>KEROSENE:</u>

For estimation of Kerosene prices in the following years after 2011 we fit <u>a Logarithmic function</u> to the data of Kerosene prices.

We choose this fit because it gave realistic values of prices and its <u>Sum of Squares of Residual is Minimum</u> for this fit. Hence we get the following graph:




The equation of this Logarithmic function is

$$y = 923.1\ln(x) - 7010$$

Till 2011 the observed Kerosene price and the estimated prices are as follows:

| Year | Observed Prices | Estimate Prices |
|---|---|---|
| **1989** | 2.25 | 0.781766474 |
| **1990** | 2.51 | 1.245777542 |
| **1991** | 2.52 | 1.709555497 |
| **1992** | 2.52 | 2.173100573 |
| **1993** | 2.52 | 2.636413004 |
| **1994** | 2.52 | 3.099493024 |
| **1995** | 2.52 | 3.562340865 |
| **1996** | 2.52 | 4.024956761 |
| **1997** | 2.52 | 4.487340943 |
| **1998** | 2.52 | 4.949493643 |
| **1999** | 2.52 | 5.411415094 |
| **2000** | 6.278333333 | 5.873105526 |
| **2001** | 7.35 | 6.334565171 |
| **2002** | 8.854615385 | 6.795794259 |
| **2003** | 8.99 | 7.25679302 |
| **2004** | 9.01 | 7.717561684 |
| **2005** | 9.056666667 | 8.178100482 |
| **2006** | 9.0875 | 8.638409641 |

| | | |
|---|---|---|
| **2007** | 9.1075 | 9.098489391 |
| **2008** | 9.18 | 9.558339961 |
| **2009** | 9.22 | 10.01796158 |
| **2010** | 10.22 | 10.47735447 |

For the successive years the Kerosene prices are predicted to be as follows:

| ***Year*** | ***Estimated price (in Rs.)*** |
|---|---|
| **2011** | **10.93651887** |
| **2012** | **11.395455** |
| **2013** | **11.85416308** |
| **2014** | **12.31264335** |
| **2015** | **12.77089602** |
| **2016** | **13.22892134** |
| **2017** | **13.68671951** |
| **2018** | **14.14429077** |
| **2019** | **14.60163535** |
| **2020** | **15.05875345** |

# CONCLUSION

# CONCLUSION

All the Oil prices show an increasing trend.

Based on the increasing pattern, we observe that the Petrol and LPG prices show a Linear trend while Diesel and Kerosene show a Logarithmic trend.

Also as prices of Petrol and LPG have a Linear increase and no seasonal variation, Double Exponential Smoothing can be applied to them to obtain smoothed values of prices.

Using the selected fits we estimate the oil prices for the successive ten years that is from 2011 to 2020. Hence, by studying the trends in the changes in Oil Prices from 1989-2010, the oil prices for the years 2014 and 2015 are estimated to be as follows:

| Fuel | Estimated Prices for 2014 (Rs.) | Estimated Prices for 2015 (Rs.) |
|---|---|---|
| **Petrol** | 55.47 | 57.38 |
| **Diesel** | 40.60 | 42.27 |
| **LPG** | 385.804 | 399.79 |
| **Kerosene** | 12.31 | 12.77 |

The estimates are approximate, due to the limitations of the data set. Hence the actual prices may or may not differ from estimated values.

# SCOPE

# AND

# LIMITATIONS

# SCOPE

The statistical analysis of Oil Prices can be made on a larger extent by using larger data for a greater number of years. Having studied the previous patterns in changes of Oil Prices, the study enables the approximate prediction of Average Annual Oil Prices for years 2011 – 2020.

If the data set is large enough and consists of monthly, weekly or daily Oil Prices, the use of similar statistical techniques could be extended to predicting the Oil Prices for forthcoming years, on an Annual, Monthly, Weekly and / or Daily basis with Sufficient Accuracy.

# LIMITATIONS

- The Residual Plots of Diesel and Kerosene show a certain pattern, which means that their respective fits are not very good. But due to the limitations of Microsoft Excel, other fits, which could be better, cannot be used for estimation.
- The trend line equations obtained in Microsoft Excel seem to be incorrect. They do not provide realistic values of Oil Prices for a given year, even for the trend line having the best R-Squared Value.
- The data represents the Oil Prices in Delhi, Only. They are not the average Oil Prices in the Nation.
- The data of Oil Prices has been obtained for the period 1989-2010 only. There were major changes in Oil Prices after 2010, which have not been studied and taken into consideration, while estimating the forecast values for Oil Prices. This may result discrepancies between the Forecasted and Observed Values.
- The Oil Prices in the data are Annual Averages. Oil Prices on a Monthly, Weekly and Daily basis could not be obtained. Hence, Prediction is limited to Annual Prices only. Oil Prices cannot be predicted on a Monthly, Weekly and Daily basis.

# REFERENCES

# REFERENCES


## ❖ BOOKS

- Discrete Probability for S.Y.B.Sc. - (Nirali Publications)

## ❖ WEBSITES

- http://www.statslab.cam.ac.uk/~rrw1/timeseries/t.pdf
- http://en.wikipedia.org/wiki/Time_series
- http://www.public.iastate.edu/~alicia/stat328/Time%20Series.pdf
- http://www.lancaster.ac.uk/media/lancaster-university/content-assets/documents/lums/forecasting/Chap4.pdf