A REVIEW ON OBSERVING POPULATION PROJECTION METHOD TO THE POPULATION OF GREECE , DENMARK AND AUSTRIA

A Project Submitted In Partial Fulfilment of the requirements for the Degree of Bachelor of Science in Statistics

By

TOMOJIT UKIL

Semester-6

Paper: DSE-B2

Registration No: 544-1111-0661-20

Roll No: 203544-21-0121

VIVEKANANDA COLLEGE

AFFILIATED TO THE UNIVERSITY OF CALCUTTA

CERTIFICATE

This is certified that the project paper entitled "**A REVIEW ON POPULATION PROJECTION METHOD TO THE POPULATION OF GREECE, DENMARK & AUSTRIA** " is submitted by **TOMOJIT UKIL** in partial fulfilment of the requirement for the Bachelor degree of Statistics (Honours) is based upon the result of benefited research work carried out by the investigator under guidance and supervision of the professors of Department of Statistics, Vivekananda College , Thakurpukur .

The results of the investigator reported in this project paper have not so far been submitted for any degree or diploma.

DECLARATION

I **TOMOJIT UKIL** , a student of B.Sc. Semester-6, Statistics Honours, of University of Calcutta, Registration no: 544-1111-0661-20 ,Roll no:203544-21-0121, hereby declare that I have done this piece of project work entitled as "A Review Work On Fitting a Suitable Curve To The Population Of Greece , Denmark And Austria " under the supervision of Professors of Department of Statistics, Vivekananda College) as a part of B.Sc. Sem-6 examination according to the syllabus paper DSE-B2.

I further declare that the piece of project work has not been published elsewhere for any degree or diploma or taken from any published project.

Tomojit Ukil

(Signature Of The Student)

(NAME OF THE STUDENT)

ACKNOWLEDGEMENT

I am indebted to number of person for helping me in the preparation of this project.
Firstly, Dr. Tapan Kumar Poddar Principal, Vivekananda College, University of Calcutta. Without whose help I couldn’t have been a part of this prestigious college.
I owe a deep debt of gratitude to my supervisors Mr. Nilkanta Mukherjee , Head of the Department of Statistics , Vivekananda College , Prof. Sutapa Biswas , Prof Ridhhi Das Mazumder (faculty Members) for necessary guidance, for this presentation of this dissertation , valuable comments and suggestions. I am extremely grateful to them for the the necessary stimulus, support and valuable time., often took pains and stood by me in adverse circumstances. Without their encouragement and inspiration it was not possible for me to complete this project.

Finally my earnest thanks go to my friends who were always beside me when I needed them without any excuses and made these three years worthwhile.

This project is not only a mere project. It is the memories spend with the whole department which has created a mutual understanding among us. There are many emotions related to this piece of work, especially respect and duty towards teachers and vice versa; educational attachment with my friends; social attachment with my college.

Tomojit Ukil

Student of Department of Statistics

Vivekananda College

Signature of the Examiner

Table of Contents

OBJECTIVE OF THE PROJECT

The objective of population projection is to estimate and forecast the future size, structure, and characteristics of a population over a specific period. This process involves using demographic data, historical trends, and assumptions about future demographic behaviour to make informed predictions about how a population will change over time.

The key objectives of population projection are as follows:

1. Forecasting population size: Population projections aim to estimate the total number of people who will be living in a given region or country in the future. This helps in understanding the scale of demographic changes that may occur and assists in planning for various services and resources accordingly.

2. Analyzing age and sex distribution: Population projections provide insights into how the age and sex composition of a population will change over time. This is essential for planning social services, healthcare facilities, pension systems, and other programs tailored to specific age groups.

3. Assessing demographic trends: By analyzing past demographic trends, population projections attempt to identify patterns and factors that have influenced population changes historically. This helps in making assumptions and developing models for future projections.

    It's important to note that population projections are based on assumptions about future demographic behaviour, such as birth rates, death rates, migration, and other factors. These assumptions can vary, leading to different projection scenarios and outcomes. Therefore, population projections are meant to be used as a guide, and they may change over time as new data becomes available or as the underlying assumptions are updated.

Here our main focus will be on Forecasting Population size varying over different factors over a period of time for the countries viz:- GREECE , DENMARK, AUSTRIA .

ABSTRACT FOR THE PROJECT :-A REVIEW ON POPULATION PROJECTION METHOD OF THE POPULATION OF DENMARK , GREECE &AUSTRIA through Mathematical Curve Fitting(Logistic Curve)

RELEVANCE :- A series of observations recorded in accordance with the time of occurrence are called Time Series Data. One of the Main Objective of these data to forecast the future behaviour and thereby facilitate in proper planning for future operation , which having a large application among the top most corporates all over the world . **Stock Market Predictions , Customer Feedbacks , Forecasting Climatic Changes , Population Projections of the countries etc** . are the part of these analysis and an integral part of the world economics. **Mathematical Curve fitting is one of the most important components of Time series , in predicting future trend values , forecasting large data etc, and widely used for predicting the population growth of the countries**

DATA SOURCE:- For forecasting purpose , the data ;consists of three countries ; viz:- **DENMARK , GREECE, AUSTRIA** ,is collected from the website named:_ **STATISTA** and the source is given below :-

For **DENMARK** :- https://www.statista.com/statistics/1008531/total-population-denmark-1769-2020/

For **GREECE** :- https://www.statista.com/statistics/1014317/total-population-greece-1821-2020/

For **AUSTRIA** :- https://www.statista.com/statistics/1008043/population-austria-1910-2020/

BRIEF DESCRIPTION OF THE ANALYSIS :- As mentioned earlier (i.e. in RELEVANCE) time series data are integral part of our daily life . The various features in the actual population gives us an insight of how the population behaves. However as the study ranges over a long period of time (approximately 70 years) , great care should betaken in choosing the Mathematical model we assume to fit the given population. In this situation Logistic curve into foreplay. Logistic Curve fitting comes under the section of Population Projection in Statistical studies. Here I have obtained estimates of the population figures of these three countries by a certain mathematical method. The Population Figures are considered to be of some particular function of time t, say Pt, as a function of t. Population Projection is done using the Logistic Curve and the fitting is observed. Generally a data ranging over such a long range of time shows many different features. The study intends to study how Logistic Curve is used to estimate those population figures. Here we have used the MAPE Test method to draw inference on the desired fitting. How much is it appropriate to model the given data using our mathematical assumptions.

PRELIMINARY CONCLUSION: As we all know all the above mentioned countries are a great sufferer of the WORLD WAR – II , thus over a period of time from (1950-2020) , how the population of the three early mentioned countries grows or behave , after facing Rapid Industrialisation or Urbanization happened in the European Continent over a period of time . Moreover this also gives us estimate how LOGISTIC CURVE is more appropriate in EUROPIAN countries rather than Asian Countries.

COUNTRY :- GREECE

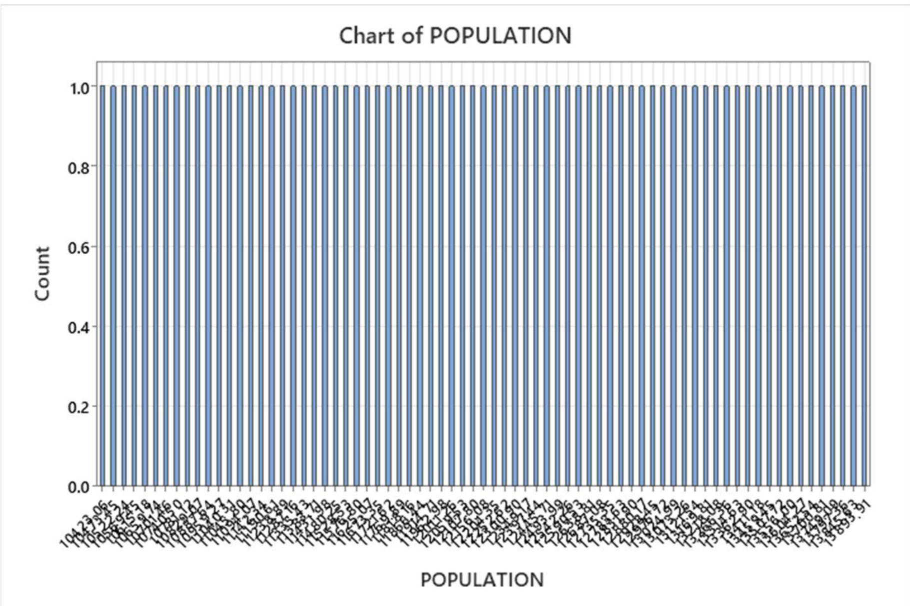
For the above mentioned country we have population data over 70 odd years i.e. from 1948-2020 , post World War-II period . We intend to study the population growth , due to rapid industrialisation , urbanization in that post war period . We thought of fitting Pearl-reed Logistic Curve for that population data varying over 72 years.

DATA VISUALISATION

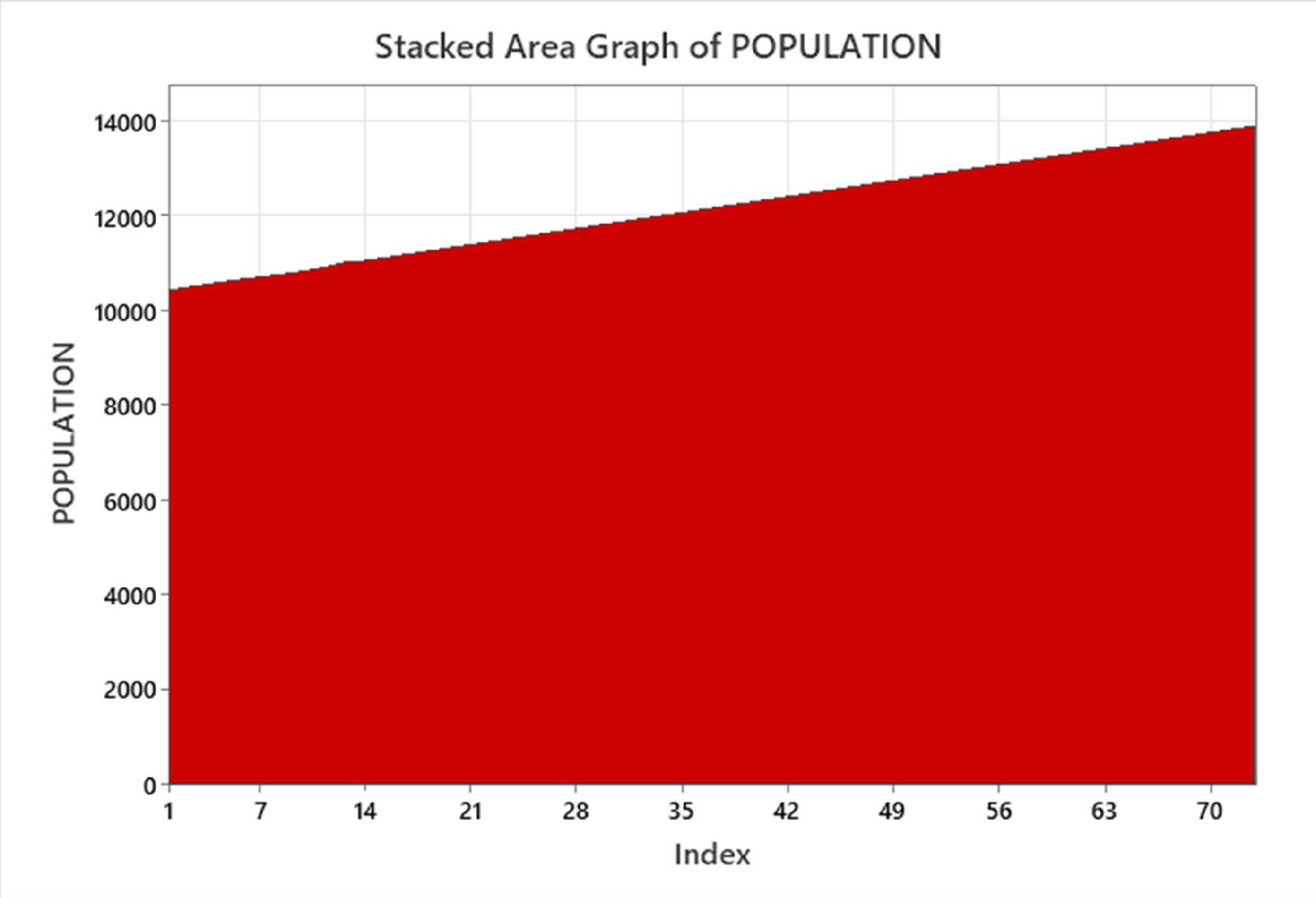
Data visualization is an effective way to explore and understand population data over the years. There are various types of visualizations that can be used to represent population trends, such as line charts, bar charts, area charts, and heatmaps. Let's explore a few examples**:-**

Line Chart: A line chart is useful for showing the trend of population change over time. You can plot the population on the y-axis and the years on the x-axis. Each point on the line represents the population count for a specific year.

Bar Chart: A bar chart can be used to compare population data for different years. You can represent each year as a separate bar and the height of each bar corresponds to the population count. This allows you to easily compare population sizes across multiple years.



Area Chart: An area chart is similar to a line chart, but the area under the line is filled with colour, making it easier to visualize the cumulative population growth. It is particularly useful for showing the composition of population changes over time, such as the growth of different age groups or demographic categories.



Heatmap: A heatmap can be used to visualize population density or distribution across different regions. Each region is represented by a cell, and the color intensity of the cell indicates the population density. This type of visualization is helpful for identifying patterns and variations in population distribution.

DESCRIPTIVE MEASURES

Descriptive measures are statistical calculations that provide summary information about a population over the years. Here are some common descriptive measures used in population analysis.

Mean: The mean, also known as the average, represents the sum of all population values divided by the total number of years. It provides a measure of central tendency and helps to understand the average population size over the years.

Median: The median is the middle value in a sorted list of population values. It is a robust measure of central tendency that is less influenced by extreme values. The median can give an idea of the population size that separates the higher and lower halves.

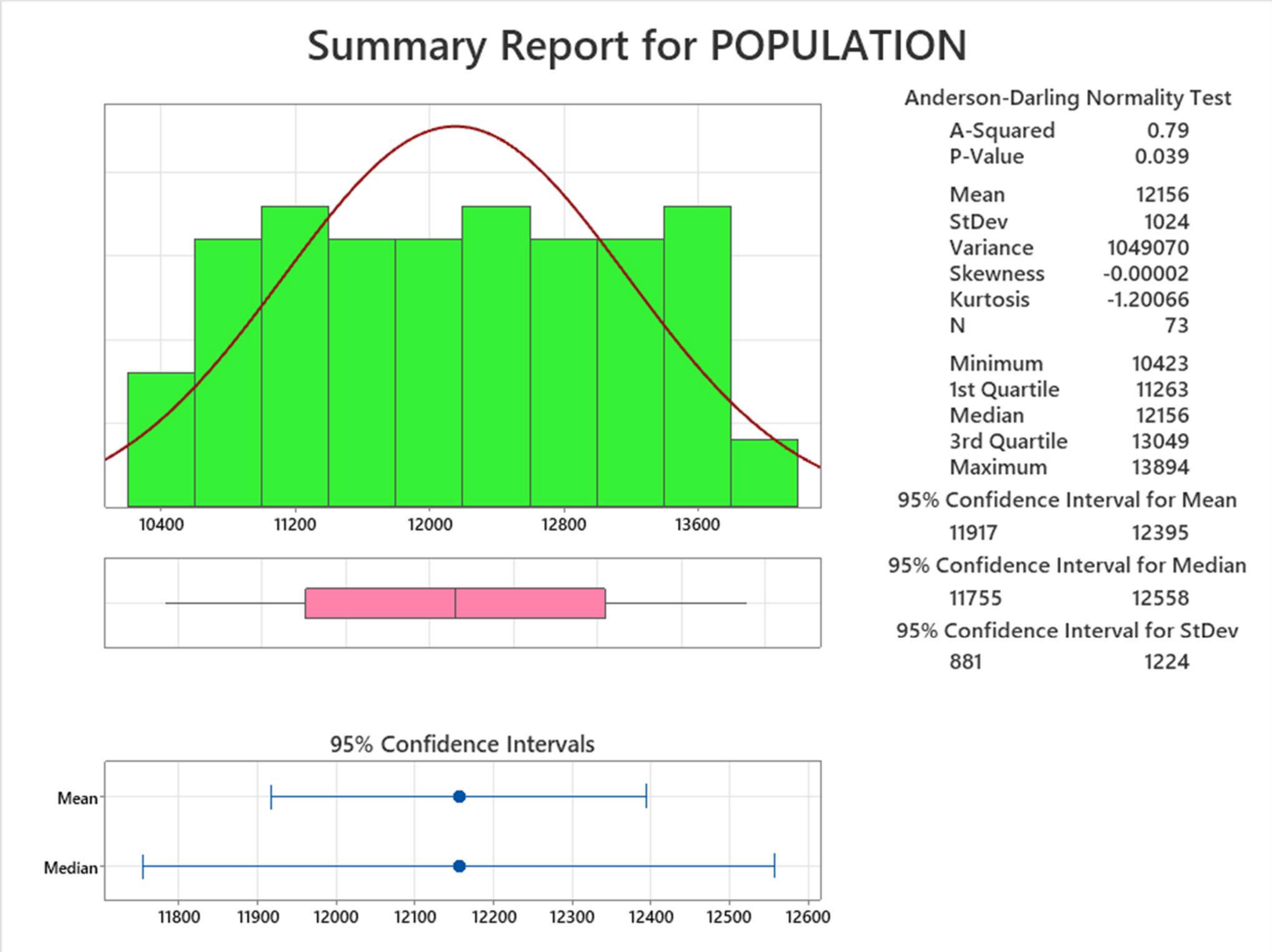
Mode: The mode represents the most frequently occurring population value over the years. It helps to identify the population count that appears with the highest frequency, which can be useful for understanding dominant population sizes.

Range: The range is the difference between the maximum and minimum population values over the years. It provides a simple measure of dispersion, indicating the spread of population sizes across different years.

Standard Deviation: The standard deviation measures the average amount of variation or dispersion in the population data over the years. It quantifies how much the population sizes deviate from the mean. A higher standard deviation indicates greater variability in population counts.

Percentiles: Percentiles represent the values below which a certain percentage of the population falls. For example, the 25th percentile indicates the population value below which 25% of the years' data falls. Percentiles help to understand the distribution of population sizes and identify specific thresholds.

Growth Rate: The growth rate is the percentage change in population size over consecutive years. It measures the rate at which the population is increasing or decreasing. Positive growth rates indicate population growth, while negative growth rates indicate population decline.



These descriptive measures provide valuable insights into the characteristics and changes in population over time.

The above table with the respective graphs provides us the idea of the descriptive measures present on the data of the country Greece , over the period of time 1948-2020.

Now ;

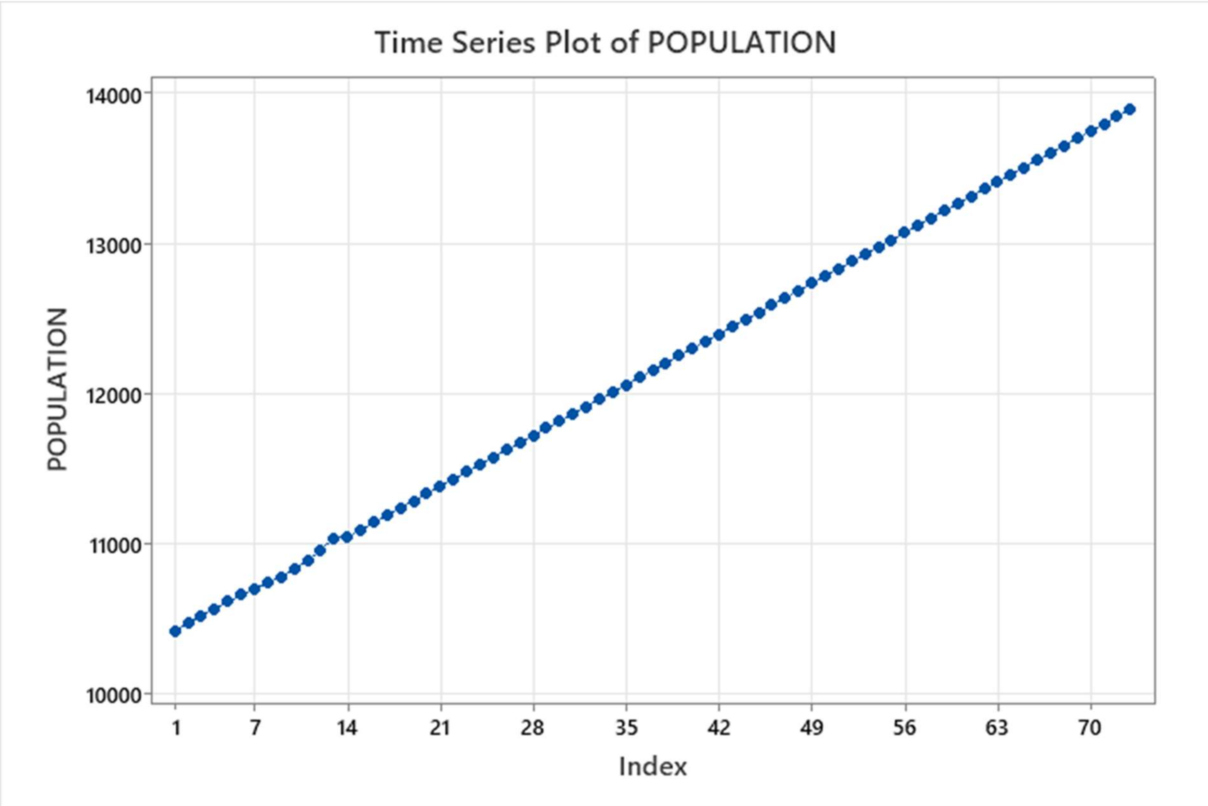
The Pearl-Reed Logistic Curve Model is given by :-

$$\boxed{Y_t = (10^a) / (\beta_0 + \beta_1 * \beta_2^t).}$$

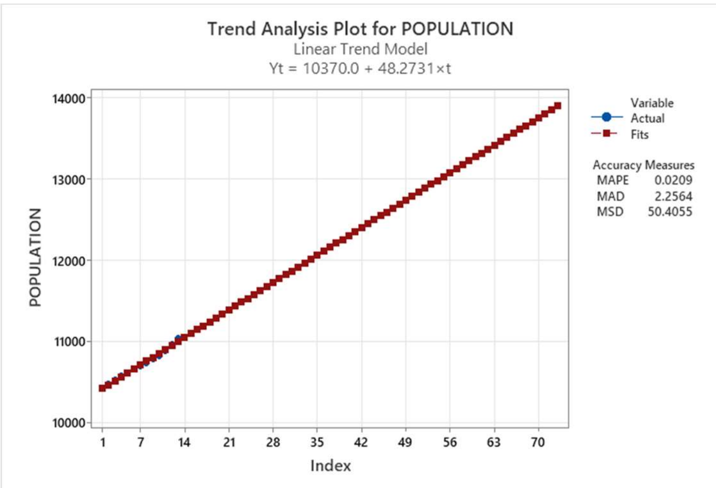
The reasons for choosing Logistic Curve as a measure of fit is given below with appropriate graphs .

Measure of Goodness of Fit :- The mean absolute percentage error (MAPE), also known as mean absolute percentage deviation (MAPD), is a measure of prediction accuracy of a forecasting method in Statistics. It usually expresses the accuracy as a ratio.

Graph of actual values for Greece from 2020-1948



Fitted values for the country Greece (LINEAR TREND)



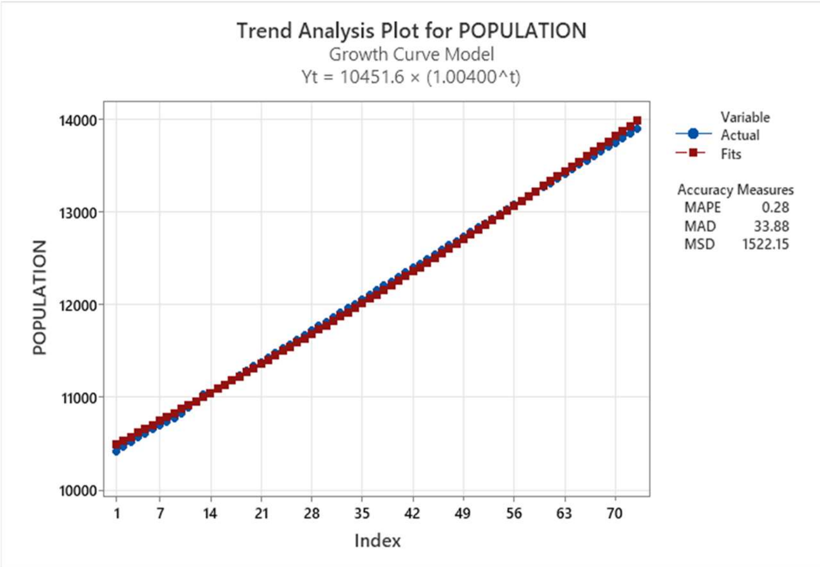
Method

Model type	Linear Trend Model
Data	POPULATION
Length	73
NMissing	0

Fitted values for the country Greece (EXPONENTIAL CURVE)

Method

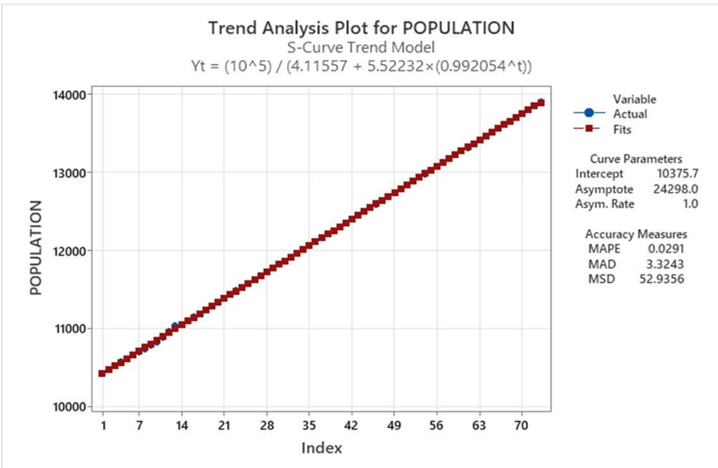
Model type	Growth Curve Model
Data	POPULATION
Length	73
NMissing	0



Fitted values for the country Greece (LOGISTIC CURVE)

Method

Model type	S-Curve Trend Model
Data	POPULATION
Length	73
NMissing	0



Fitted Trend Equation

Yt = (10^5) / (4.11557 + 5.52232×(0.992054^t))

<u>CONCLUSION:-</u> Among all the fitted curves the logistic curve produces the least **MAPE** value i.e. .0291< 2% , which implies **this is a best fit for our data** and make it quite obvious to choose the Logistic Curve .

The actual fitted Trend Equation thus given by :-

$$\boxed{Yt= (10^5)/(4.1157+5.52232*(0.992054^t))} \quad \text{............ (1)}$$

Where the values of the estimated parameters thus given by :-

$\hat{a} = 5$

$\hat{\beta}_0 = 4.1157$

$\hat{\beta}_1 = 5.52232$

$\hat{\beta}_2 = .992054$

Now by putting the values of "t" (calculating through usual method) we will get the fitted values of the Population From the year 2020 to 1948.

TABLE SHOWING DATA FOR GREECE (FITTED POPULATION) BASED ON THE TREND EQUATION Yt= (10^5)/(4.1157+5.52232*(0.992054^t))

YEARS	Actual Population Data	t (1984-years)	Fitted Population Data
2020	10,423.06	36	12107.5444
2019	10,473.45	35	12059.0852
2018	10,522.24	34	12010.6288
2017	10,569.45	33	11962.1769
2016	10,615.18	32	11913.7308
2015	10,659.74	31	11865.2923
2014	10,701.46	30	11816.8628
2013	10,740.50	29	11768.4438
2012	10,781.12	28	11720.037
2011	10,829.07	27	11671.6437
2010	10,887.64	26	11623.2657
2009	10,959.27	25	11574.9043
2008	11,040.31	24	11526.5612
2007	11,045.80	23	11478.2378
2006	11,094.07	22	11429.9357
2005	11,142.34	21	11381.6564
2004	11,190.61	20	11333.4014
2003	11,238.89	19	11285.1723
2002	11,287.16	18	11236.9705
2001	11,335.43	17	11188.7976
2000	11,383.71	16	11140.655
1999	11,431.98	15	11092.5443
1998	11,480.25	14	11044.4669
1997	11,528.53	13	10996.4245
1996	11,576.80	12	10948.4183
1995	11,625.07	11	10900.45
1994	11,673.35	10	10852.521
1993	11,721.62	9	10804.6328
1992	11,769.89	8	10756.7869
1991	11,818.16	7	10708.9846
1990	11,866.44	6	10661.2276
1989	11,914.71	5	10613.5172
1988	11,962.98	4	10565.8548
1987	12,011.26	3	10518.242
1986	12,059.53	2	10470.6801
1985	12,107.80	1	10423.1707
1984	12,156.08	0	10375.715
1983	12,204.35	-1	10328.3146

1982	12,252.62	-2	10280.9708
1981	12,300.90	-3	10233.685
1980	12,349.17	-4	10186.4587
1979	12,397.44	-5	10139.2932
1978	12,445.71	-6	10092.1899
1977	12,493.99	-7	10045.1502
1976	12,542.26	-8	9998.17538
1975	12,590.53	-9	9951.26689
1974	12,638.81	-10	9904.42605
1973	12,687.08	-11	9857.6542
1972	12,735.35	-12	9810.95267
1971	12,783.63	-13	9764.32278
1970	12,831.90	-14	9717.76585
1969	12,880.17	-15	9671.28319
1968	12,928.45	-16	9624.8761
1967	12,976.72	-17	9578.54586
1966	13,024.99	-18	9532.29375
1965	13,073.26	-19	9486.12105
1964	13,121.54	-20	9440.02902
1963	13,169.81	-21	9394.0189
1962	13,218.08	-22	9348.09195
1961	13,266.36	-23	9302.24938
1960	13,314.63	-24	9256.49243
1959	13,362.90	-25	9210.82231
1958	13,411.18	-26	9165.24022
1957	13,459.45	-27	9119.74735
1956	13,507.72	-28	9074.34488
1955	13,556.00	-29	9029.03398
1954	13,604.27	-30	8983.81583
1953	13,652.54	-31	8938.69155
1952	13,700.81	-32	8893.6623
1951	13,749.09	-33	8848.7292
1950	13,797.36	-34	8803.89338
1949	13,845.63	-35	8759.15593
1948	13,893.91	-36	8714.51795

LINE DIAGRAM OF LOGISTIC CURVE PLOTTED AGAINST FITTED POPULATION AGAINST GREECE

Method

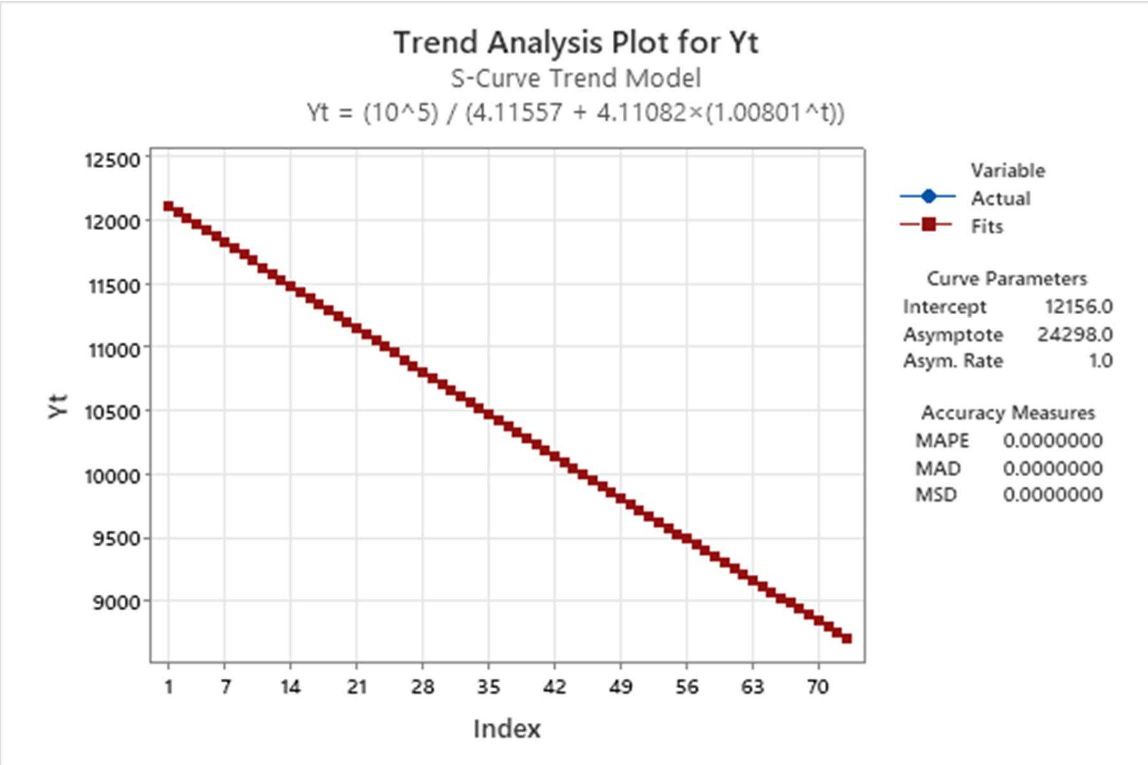
Model type	S-Curve Trend Model
Data	Yt
Length	73
NMissing	0

Accuracy Measures

MAPE	0.0000000
MAD	0.0000000
MSD	0.0000000

Fitted Trend Equation

Yt = (10^5) / (4.11557 + 4.11082×(1.00801^t))



Thus from the above figure we can see that Fitted Population gives a Straight Line , when plotted against a Pearl-Reed (Logistic Curve) with MAPE value almost 0 , so thus on the basis of our Analysis we can conclude that our curve fitting is near to perfection.

COUNTRY :- DENMARK

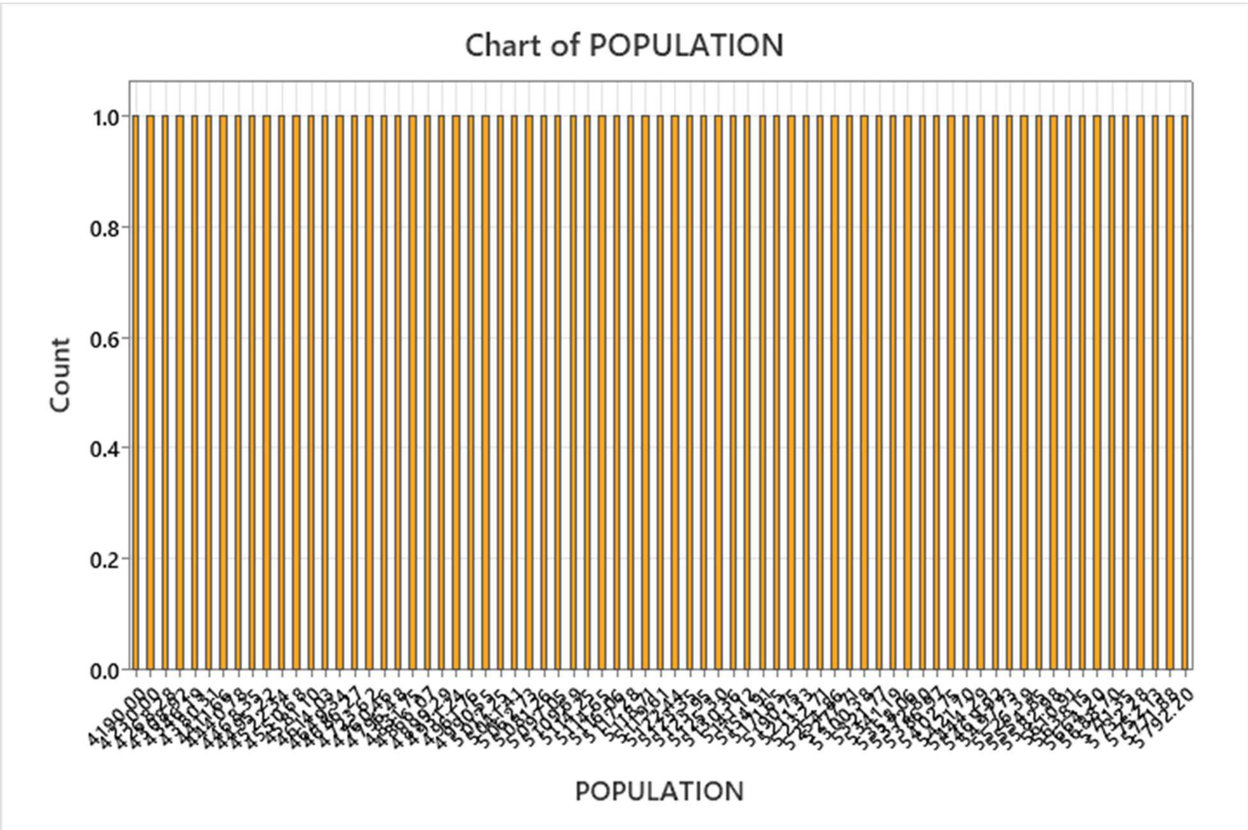
For the above mentioned country we have population data over 70 odd years i.e. from 1948-2020 , post World War-II period . We intend to study the population growth , due to rapid industrialisation , urbanization in that post war period . We thought of fitting Pearl-reed Logistic Curve for that population data varying over 72 years.

DATA VISUALISATION

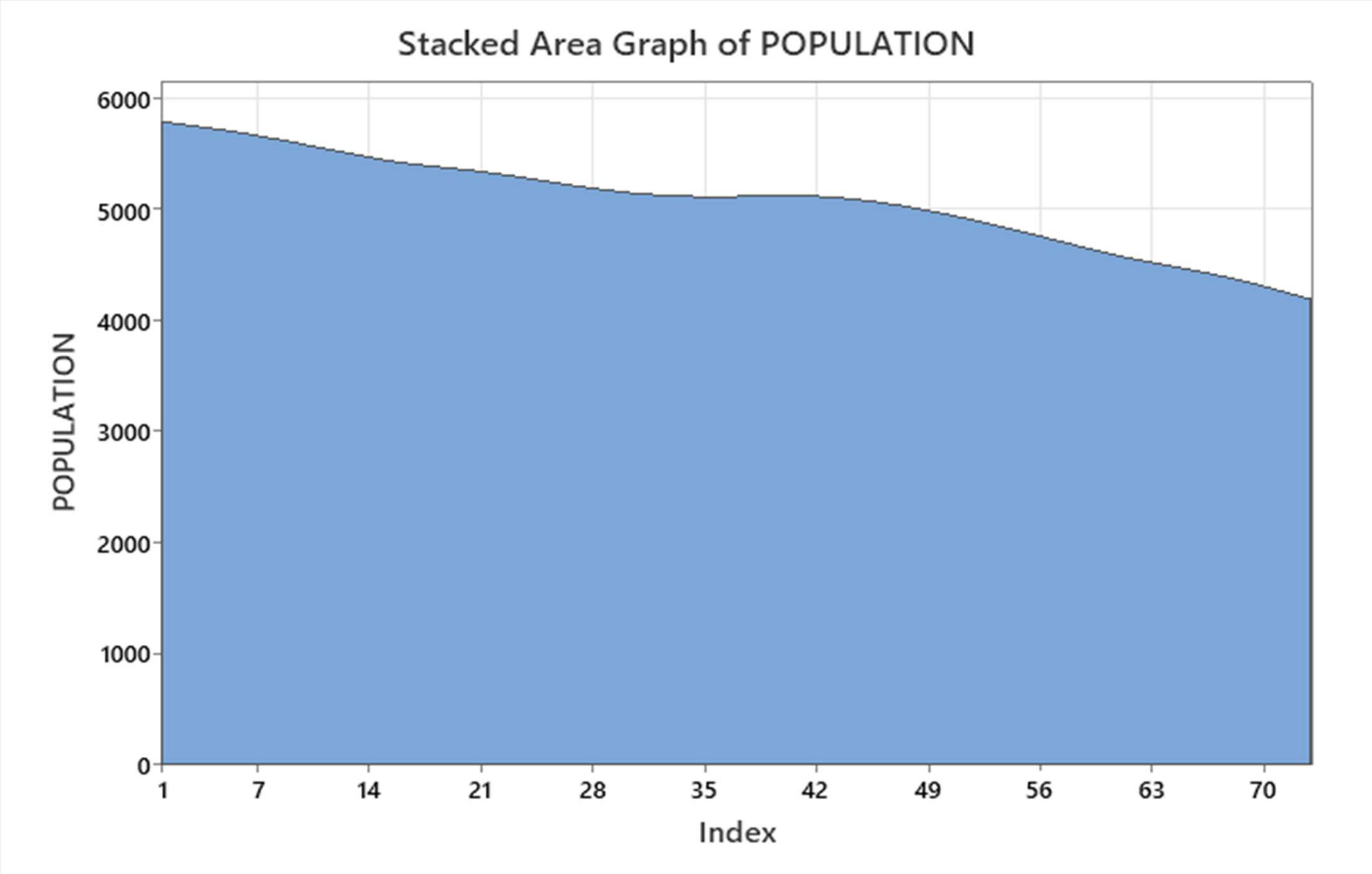
Data visualization is an effective way to explore and understand population data over the years. There are various types of visualizations that can be used to represent population trends, such as line charts, bar charts, area charts, and heatmaps. Let's explore a few examples**:-**

Line Chart: A line chart is useful for showing the trend of population change over time. You can plot the population on the y-axis and the years on the x-axis. Each point on the line represents the population count for a specific year.

Bar Chart: A bar chart can be used to compare population data for different years. You can represent each year as a separate bar and the height of each bar corresponds to the population count. This allows you to easily compare population sizes across multiple years.



Area Chart: An area chart is similar to a line chart, but the area under the line is filled with colour, making it easier to visualize the cumulative population growth. It is particularly useful for showing the composition of population changes over time, such as the growth of different age groups or demographic categories.



Heatmap: A heatmap can be used to visualize population density or distribution across different regions. Each region is represented by a cell, and the color intensity of the cell indicates the population density. This type of visualization is helpful for identifying patterns and variations in population distribution.

DESCRIPTIVE MEASURES

Descriptive measures are statistical calculations that provide summary information about a population over the years. Here are some common descriptive measures used in population analysis.

Mean: The mean, also known as the average, represents the sum of all population values divided by the total number of years. It provides a measure of central tendency and helps to understand the average population size over the years.

Median: The median is the middle value in a sorted list of population values. It is a robust measure of central tendency that is less influenced by extreme values. The median can give an idea of the population size that separates the higher and lower halves.

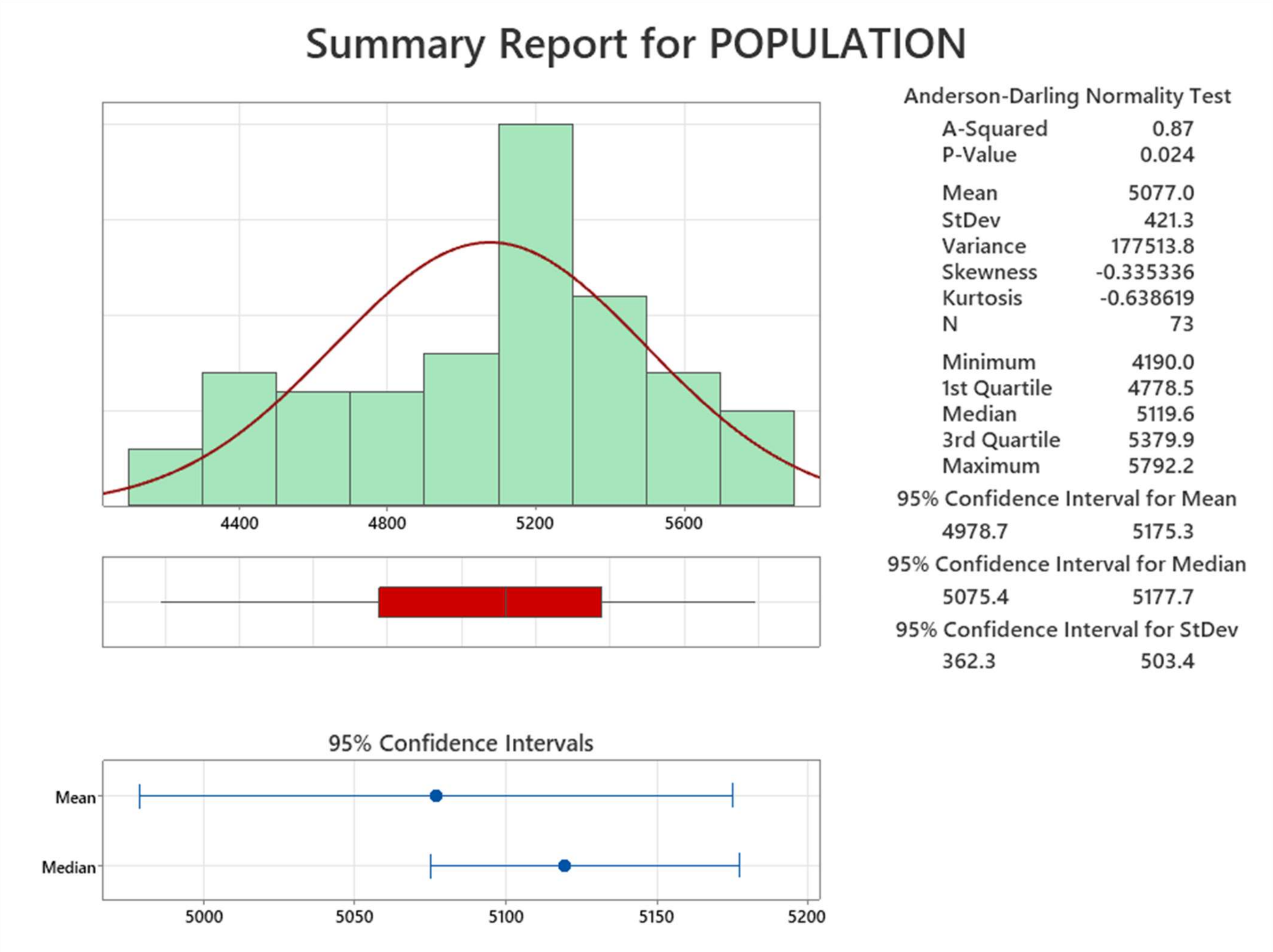
Mode: The mode represents the most frequently occurring population value over the years. It helps to identify the population count that appears with the highest frequency, which can be useful for understanding dominant population sizes.

Range: The range is the difference between the maximum and minimum population values over the years. It provides a simple measure of dispersion, indicating the spread of population sizes across different years.

Standard Deviation: The standard deviation measures the average amount of variation or dispersion in the population data over the years. It quantifies how much the population sizes deviate from the mean. A higher standard deviation indicates greater variability in population counts.

Percentiles: Percentiles represent the values below which a certain percentage of the population falls. For example, the 25th percentile indicates the population value below which 25% of the years' data falls. Percentiles help to understand the distribution of population sizes and identify specific thresholds.

Growth Rate: The growth rate is the percentage change in population size over consecutive years. It measures the rate at which the population is increasing or decreasing. Positive growth rates indicate population growth, while negative growth rates indicate population decline.



These descriptive measures provide valuable insights into the characteristics and changes in population over time.

The above table with the respective graphs provides us the idea of the descriptive measures present on the data of the country Denmark , over the period of time 1948-2020.

Now ;

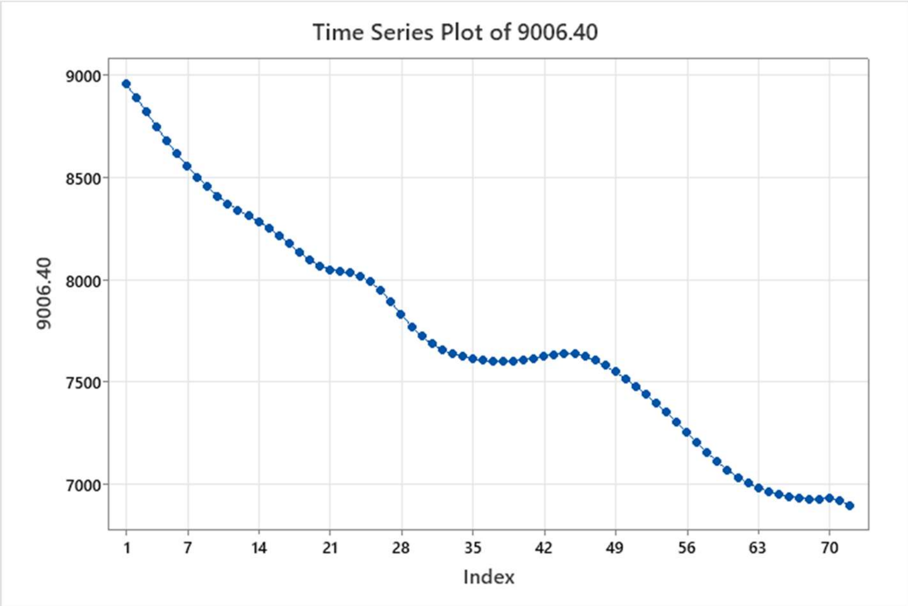
The Pearl-Reed Logistic Curve Model is given by :-

$$\boxed{Y_t = (10^a) / (\beta_0 + \beta_1 * \beta_2^t).}$$

The reasons for choosing Logistic Curve as a measure of fit is given below with appropriate graphs .

Measure of Goodness of Fit :- The mean absolute percentage error (MAPE), also known as mean absolute percentage deviation (MAPD), is a measure of prediction accuracy of a forecasting method in Statistics. It usually expresses the accuracy as a ratio.

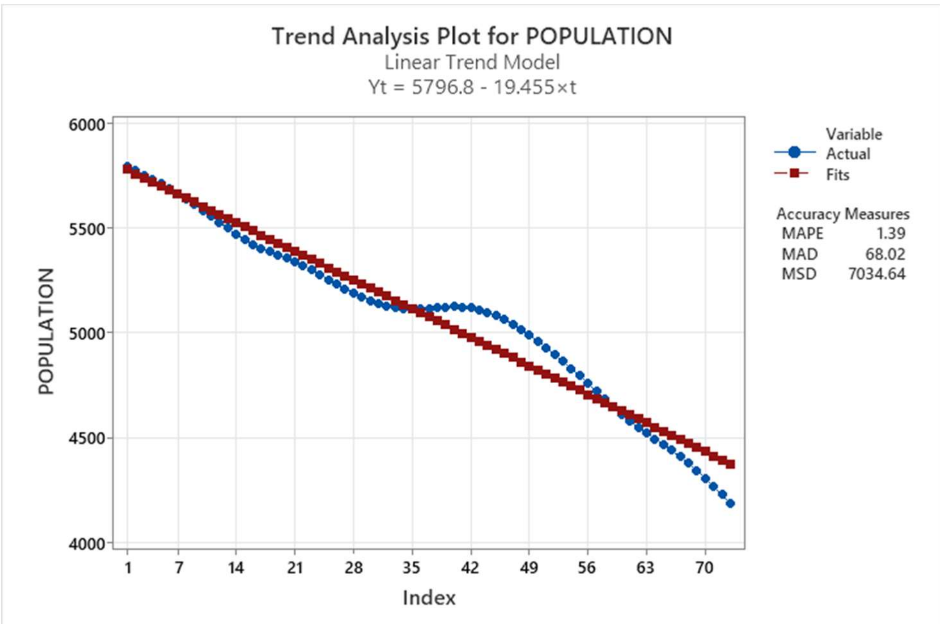
Graph of Actual values for the Country Denmark



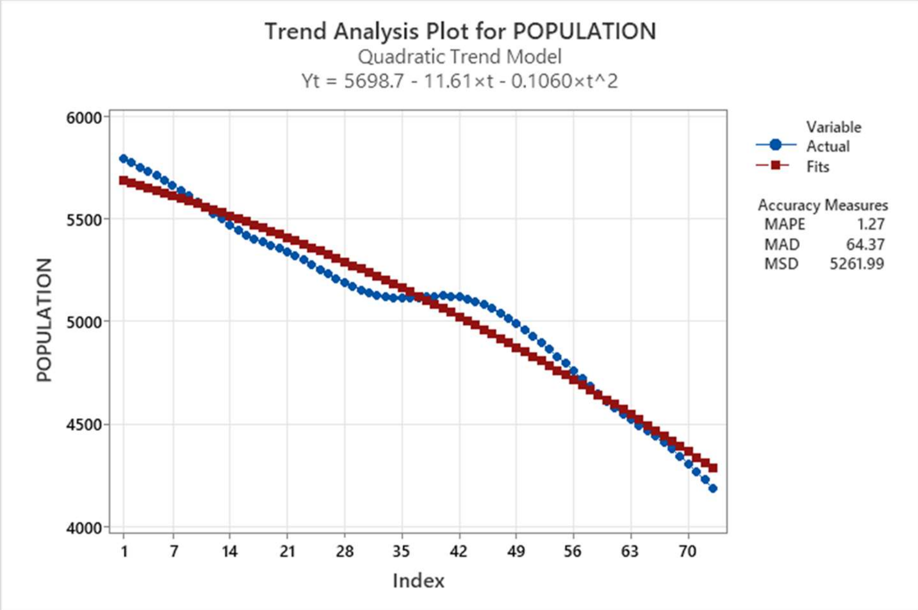
Graph of Linear trend fitting for Denmark

Method

Model type	Linear Trend Model
Data	POPULATION
Length	73
NMissing	0



.Graph for Quadratic curve fitting for Denmark



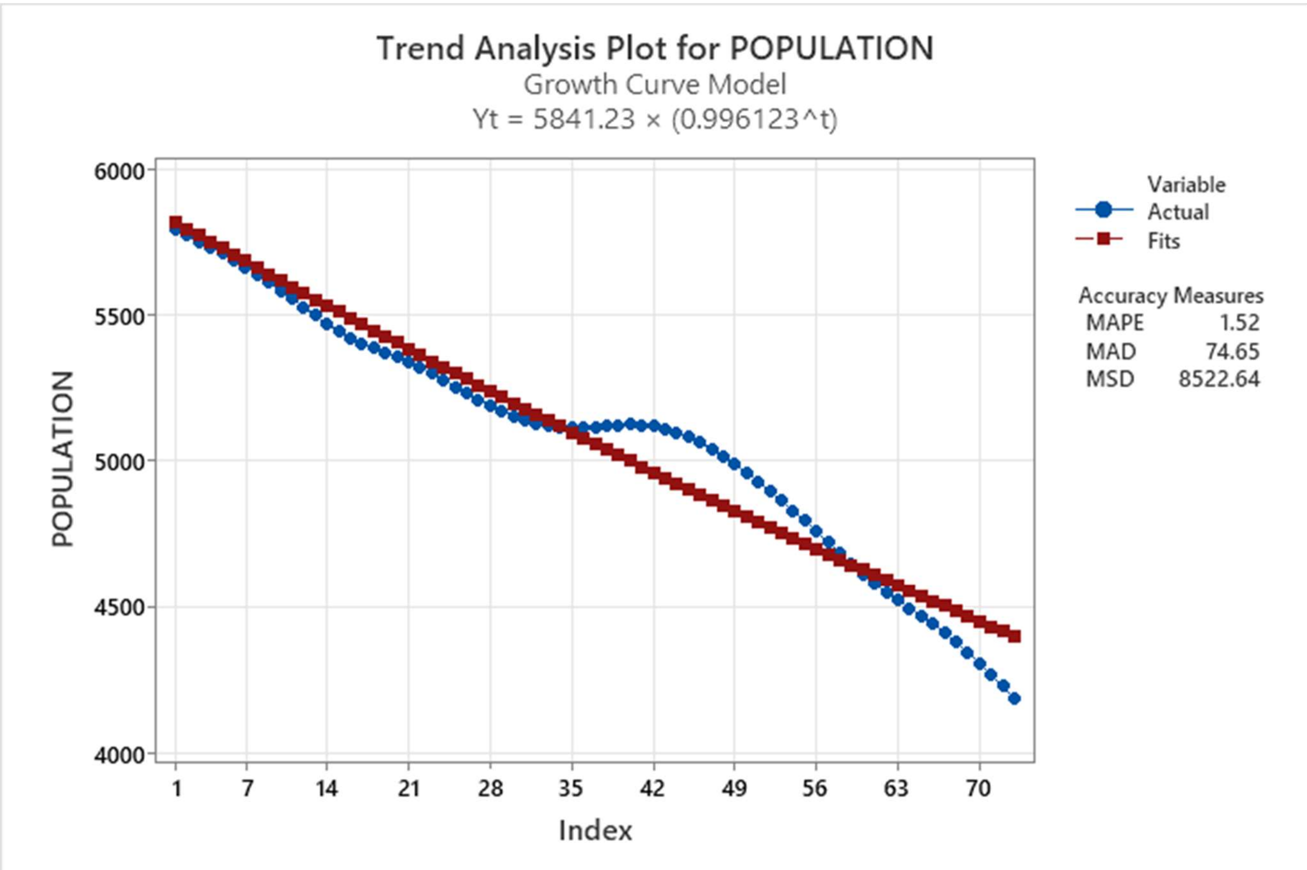
Method

Model type	Quadratic Trend Model
Data	POPULATION
Length	73
NMissing	0

Graph for exponential curve fitting for Denmark

Method

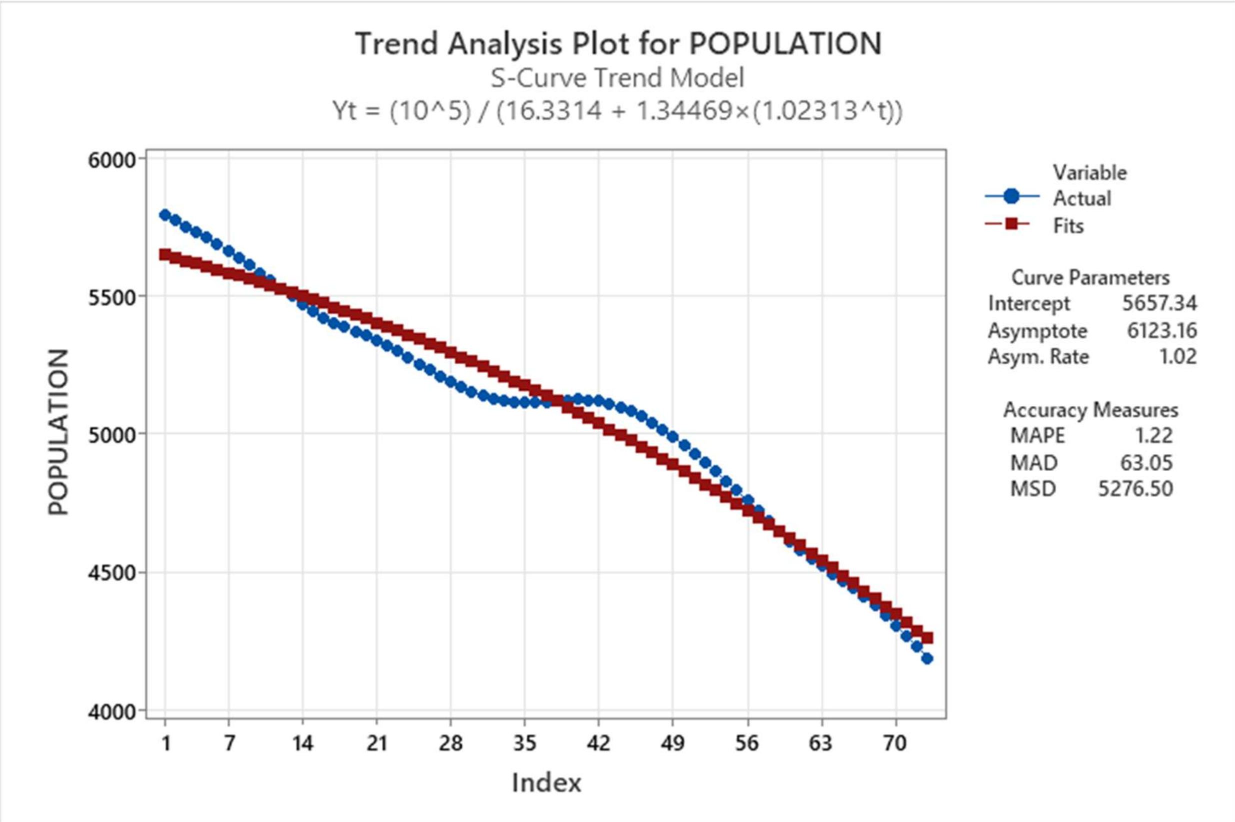
Model type	Growth Curve Model
Data	POPULATION
Length	73
NMissing	0



Graph for Logistic Curve fitting for Denmark

Method

Model type	S-Curve Trend Model
Data	POPULATION
Length	73
NMissing	0



Fitted Trend Equation

Yt = (10^5) / (16.3314 + 1.34469×(1.02313^t))

Conclusion:- From all of the above graphs we can see that **MAPE** value for the **logistic curve** is the least , 1.22<2% , implies **best fit** so we choose the model for fitting the data.

The actual fitted Trend Equation thus given by :-

$$Yt= (10^5)/(16.3314+1.34469*(1.02313^t)) \quad \text{............ (2)}$$

Where the values of the estimated parameters thus given by :-

$\hat{a} = 5$

$\hat{\beta}_0 = 16.3314$

$\hat{\beta}_1 = 1.34469$

$\hat{\beta}_2 = 1.02313$

Now by putting the values of "t" (calculating through usual method) we will get the fitted values of the Population From the year 2020 to 1948.

TABLE SHOWING DATA FOR DENMARK (FITTED POPULATION) BASED ON THE TREND EQUATION Yt= (10^5)/(16.3314+1.34469*(1.02313^t))

YEARS	ACTUAL POPULATION DATA	t (1984-years)	FITTED POPULATION (YT)
2020	5,792.20	36	5156.156
2019	5,771.88	35	5174.631
2018	5,752.13	34	5192.817
2017	5,732.28	33	5210.715
2016	5,711.35	32	5228.329
2015	5,688.70	31	5245.659
2014	5,664.20	30	5262.71
2013	5,638.15	29	5279.482
2012	5,610.91	28	5295.979
2011	5,582.98	27	5312.202
2010	5,554.85	26	5328.156
2009	5,526.39	25	5343.841
2008	5,497.73	24	5359.262
2007	5,469.92	23	5374.42
2006	5,444.29	22	5389.318
2005	5,421.70	21	5403.96
2004	5,402.75	20	5418.348
2003	5,386.97	19	5432.484
2002	5,372.80	18	5446.373
2001	5,358.06	17	5460.016
2000	5,341.19	16	5473.418
1999	5,321.77	15	5486.579
1998	5,300.38	14	5499.505
1997	5,277.71	13	5512.198
1996	5,254.86	12	5524.66
1995	5,232.71	11	5536.895
1994	5,211.33	10	5548.906
1993	5,190.75	9	5560.696
1992	5,171.65	8	5572.268
1991	5,154.91	7	5583.625
1990	5,141.12	6	5594.77
1989	5,130.36	5	5605.706
1988	5,122.44	4	5616.436
1987	5,117.28	3	5626.964
1986	5,114.65	2	5637.292
1985	5,114.25	1	5647.423

1984	5,116.06	0	5657.36
1983	5,119.61	-1	5667.106
1982	5,123.35	-2	5676.665
1981	5,125.30	-3	5686.038
1980	5,123.95	-4	5695.23
1979	5,118.71	-5	5704.243
1978	5,109.69	-6	5713.079
1977	5,097.05	-7	5721.742
1976	5,081.26	-8	5730.235
1975	5,062.73	-9	5738.56
1974	5,041.41	-10	5746.721
1973	5,017.25	-11	5754.719
1972	4,990.55	-12	5762.558
1971	4,961.76	-13	5770.241
1970	4,931.24	-14	5777.77
1969	4,899.29	-15	5785.147
1968	4,866.07	-16	5792.376
1967	4,831.75	-17	5799.459
1966	4,796.48	-18	5806.399
1965	4,760.46	-19	5813.198
1964	4,723.62	-20	5819.858
1963	4,686.27	-21	5826.383
1962	4,649.34	-22	5832.774
1961	4,614.03	-23	5839.035
1960	4,581.10	-24	5845.167
1959	4,550.68	-25	5851.172
1958	4,522.34	-26	5857.054
1957	4,495.22	-27	5862.815
1956	4,468.35	-28	5868.456
1955	4,440.78	-29	5873.98
1954	4,411.66	-30	5879.389
1953	4,380.31	-31	5884.686
1952	4,346.19	-32	5889.872
1951	4,308.92	-33	5894.949
1950	4,268.28	-34	5899.921
1949	4,230	-35	5904.788
1948	4,190	-36	5909.553

LINE DIAGRAM OF LOGISTIC CURVE PLOTTED AGAINST FITTED POPULATION AGAINST DENMARK

Method

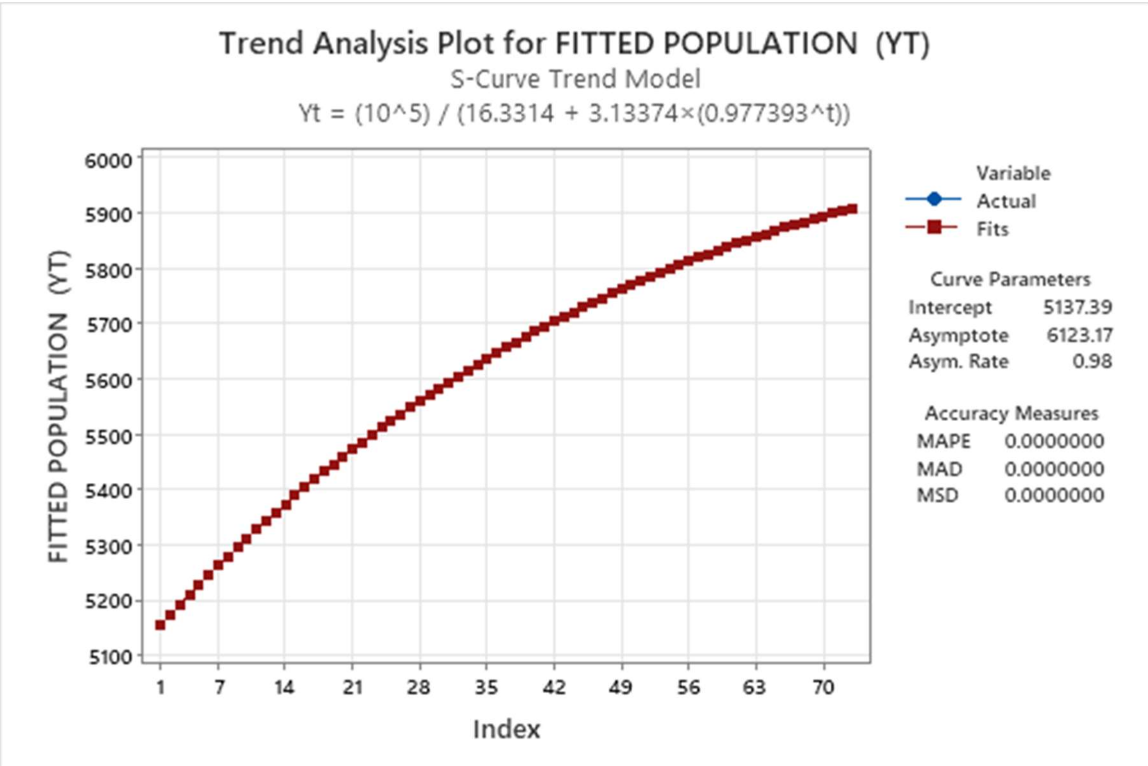
Model type	S-Curve Trend Model
Data	FITTED POPULATION (YT)
Length	73
NMissing	0

Accuracy Measures

MAPE	0.0000000
MAD	0.0000000
MSD	0.0000000

Fitted Trend Equation

Yt = (10^5) / (16.3314 + 3.13374×(0.977393^t))



Thus from the above figure we can see that Fitted Population gives a Straight Line , when plotted against a Pearl-Reed (Logistic Curve) with MAPE value almost 0 , so thus on the basis of our Analysis we can conclude that our curve fitting is good .

COUNTRY :- AUSTRIA

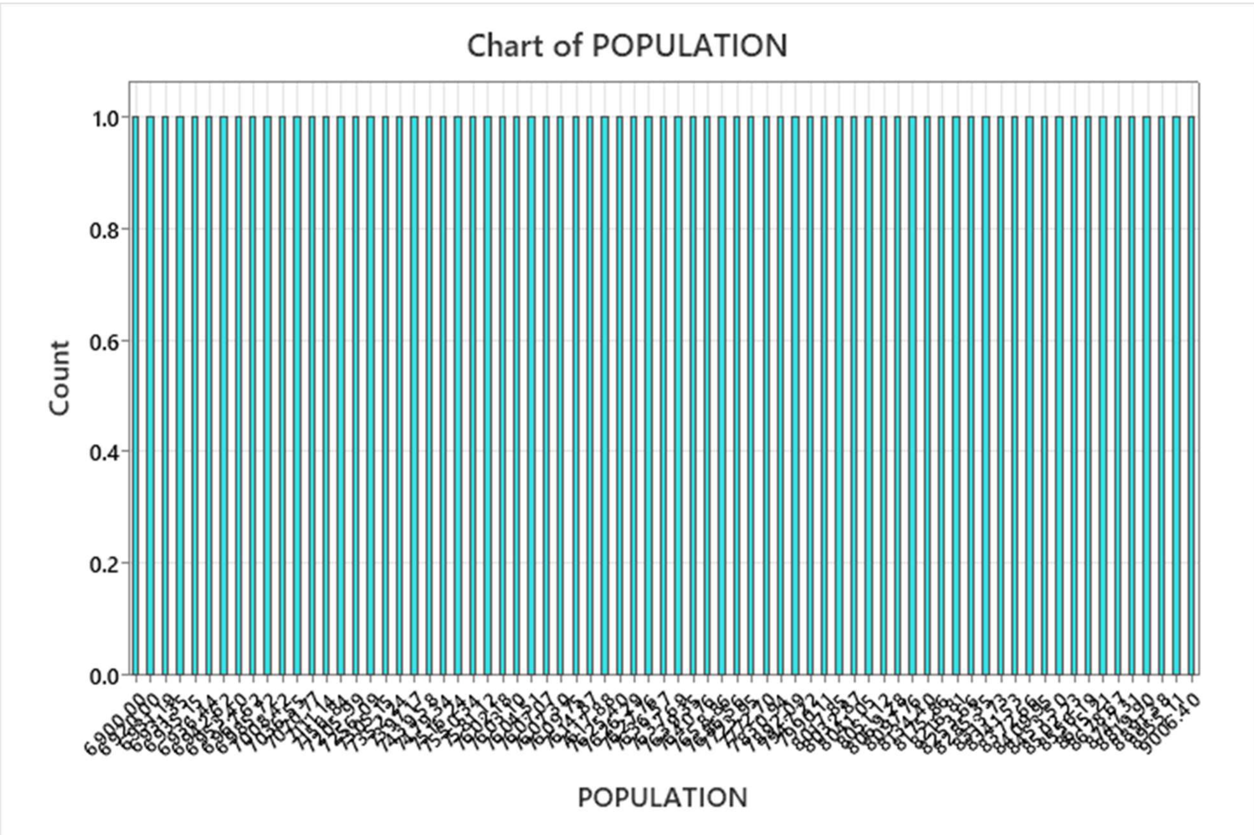
For the above mentioned country we have population data over 70 odd years i.e. from 1948-2020 , post World War-II period . We intend to study the population growth , due to rapid industrialisation , urbanization in that post war period . We thought of fitting Pearl-reed Logistic Curve for that population data varying over 72 years.

DATA VISUALISATION

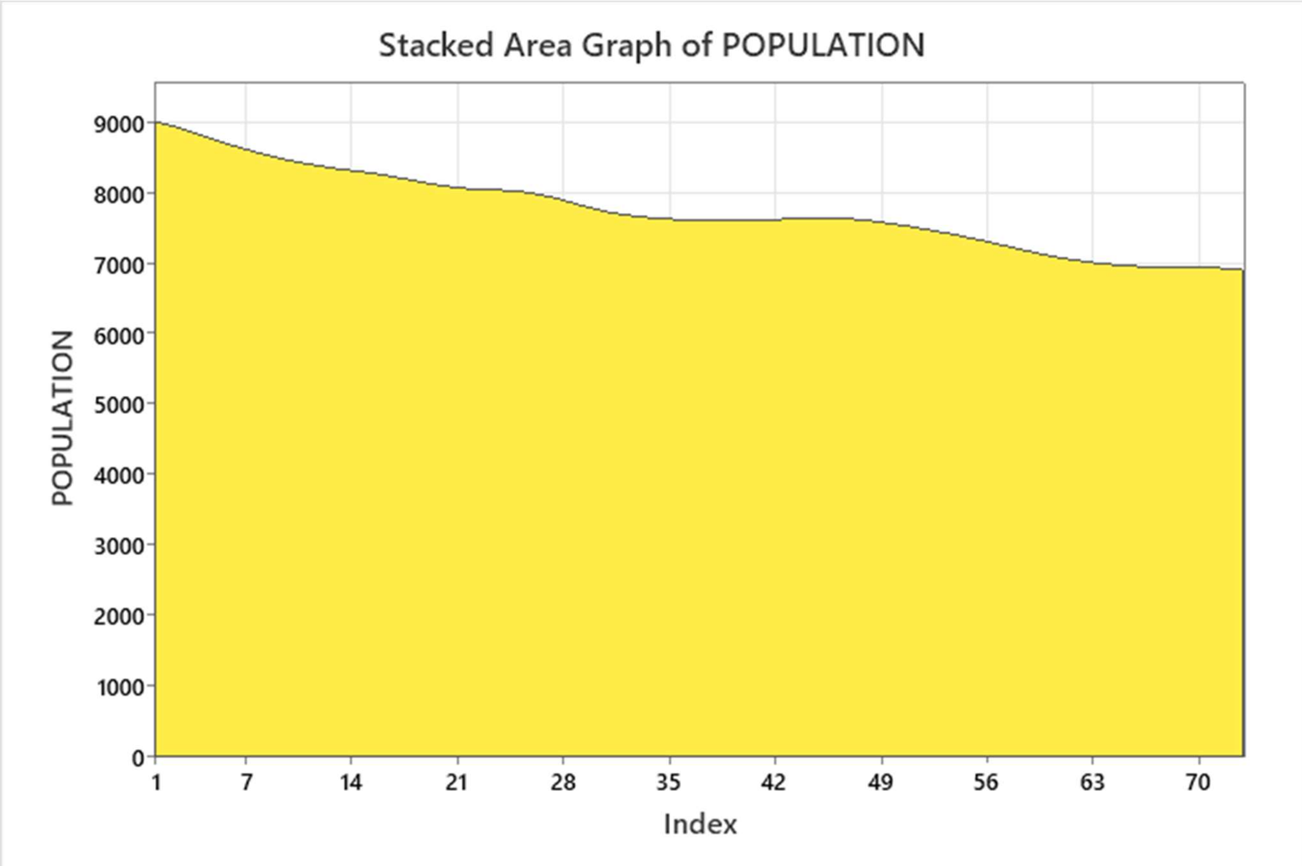
Data visualization is an effective way to explore and understand population data over the years. There are various types of visualizations that can be used to represent population trends, such as line charts, bar charts, area charts, and heatmaps. Let's explore a few examples**:-**

Line Chart: A line chart is useful for showing the trend of population change over time. You can plot the population on the y-axis and the years on the x-axis. Each point on the line represents the population count for a specific year.

Bar Chart: A bar chart can be used to compare population data for different years. You can represent each year as a separate bar and the height of each bar corresponds to the population count. This allows you to easily compare population sizes across multiple years.



Area Chart: An area chart is similar to a line chart, but the area under the line is filled with colour, making it easier to visualize the cumulative population growth. It is particularly useful for showing the composition of population changes over time, such as the growth of different age groups or demographic categories.



Heatmap: A heatmap can be used to visualize population density or distribution across different regions. Each region is represented by a cell, and the color intensity of the cell indicates the population density. This type of visualization is helpful for identifying patterns and variations in population distribution.

DESCRIPTIVE MEASURES

Descriptive measures are statistical calculations that provide summary information about a population over the years. Here are some common descriptive measures used in population analysis.

Mean: The mean, also known as the average, represents the sum of all population values divided by the total number of years. It provides a measure of central tendency and helps to understand the average population size over the years.

Median: The median is the middle value in a sorted list of population values. It is a robust measure of central tendency that is less influenced by extreme values. The median can give an idea of the population size that separates the higher and lower halves.

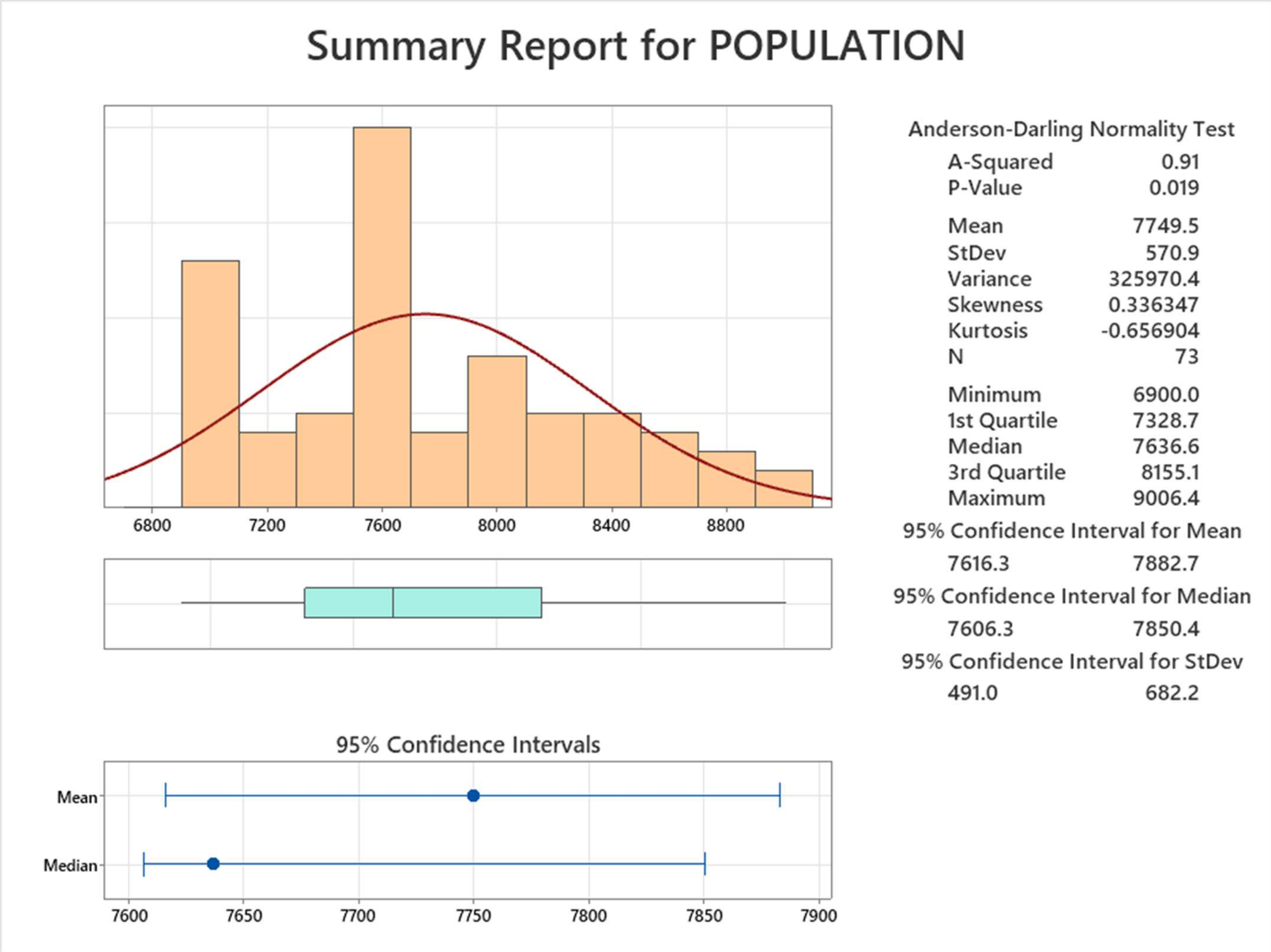
Mode: The mode represents the most frequently occurring population value over the years. It helps to identify the population count that appears with the highest frequency, which can be useful for understanding dominant population sizes.

Range: The range is the difference between the maximum and minimum population values over the years. It provides a simple measure of dispersion, indicating the spread of population sizes across different years.

Standard Deviation: The standard deviation measures the average amount of variation or dispersion in the population data over the years. It quantifies how much the population sizes deviate from the mean. A higher standard deviation indicates greater variability in population counts.

Percentiles: Percentiles represent the values below which a certain percentage of the population falls. For example, the 25th percentile indicates the population value below which 25% of the years' data falls. Percentiles help to understand the distribution of population sizes and identify specific thresholds.

Growth Rate: The growth rate is the percentage change in population size over consecutive years. It measures the rate at which the population is increasing or decreasing. Positive growth rates indicate population growth, while negative growth rates indicate population decline.



These descriptive measures provide valuable insights into the characteristics and changes in population over time.

The above table with the respective graphs provides us the idea of the descriptive measures present on the data of the country Austria , over the period of time 1948-2020.

Now ;

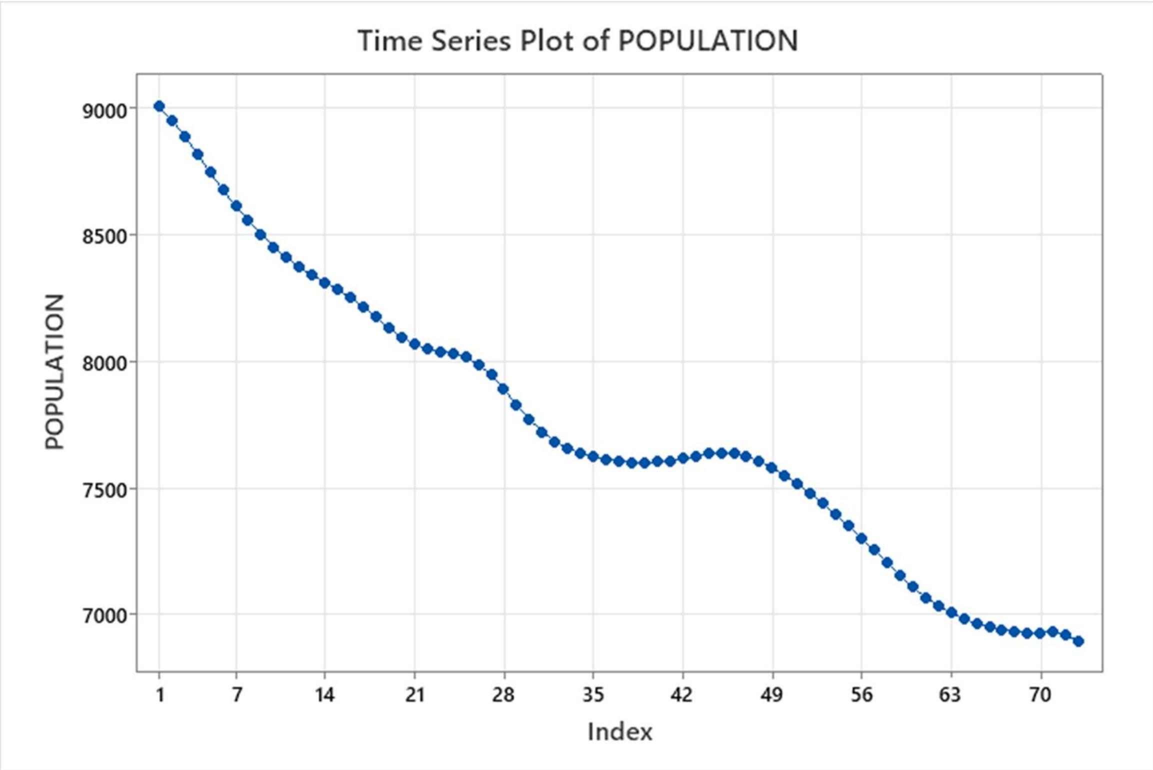
The Pearl-Reed Logistic Curve Model is given by :-

$$\boxed{\mathbf{Y_t = (10^a) / (\beta_0 + \beta_1 * \beta_2^t).}}$$

The reasons for choosing Logistic Curve as a measure of fit is given below with appropriate graphs .

Measure of Goodness of Fit :- The mean absolute percentage error (MAPE), also known as mean absolute percentage deviation (MAPD), is a measure of prediction accuracy of a forecasting method in Statistics. It usually expresses the accuracy as a ratio.

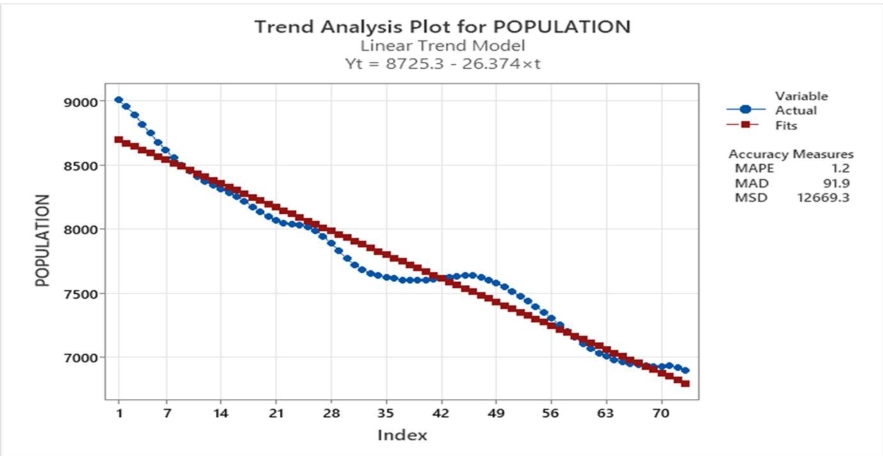
Actual Graph for the Country Austria



Linear fitted graph for Austria

Method

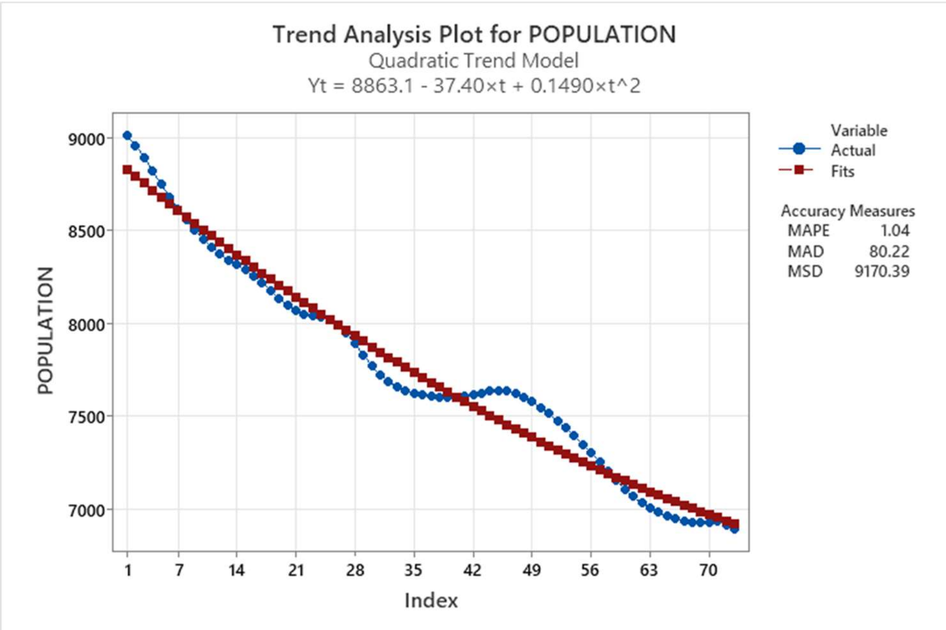
Model type	Linear Trend Model
Data	POPULATION
Length	73
NMissing	0



Graph for Quadratic curve for Austria

Method

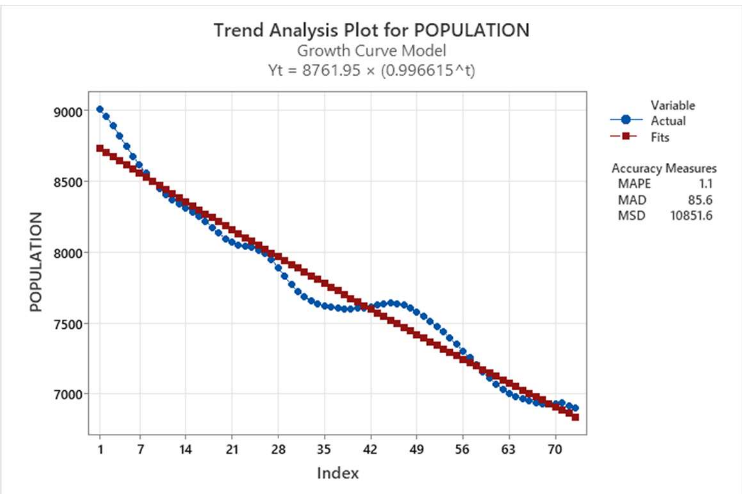
Model type	Quadratic Trend Model
Data	POPULATION
Length	73
NMissing	0



Graph for Exponential Curve for Austria

Method

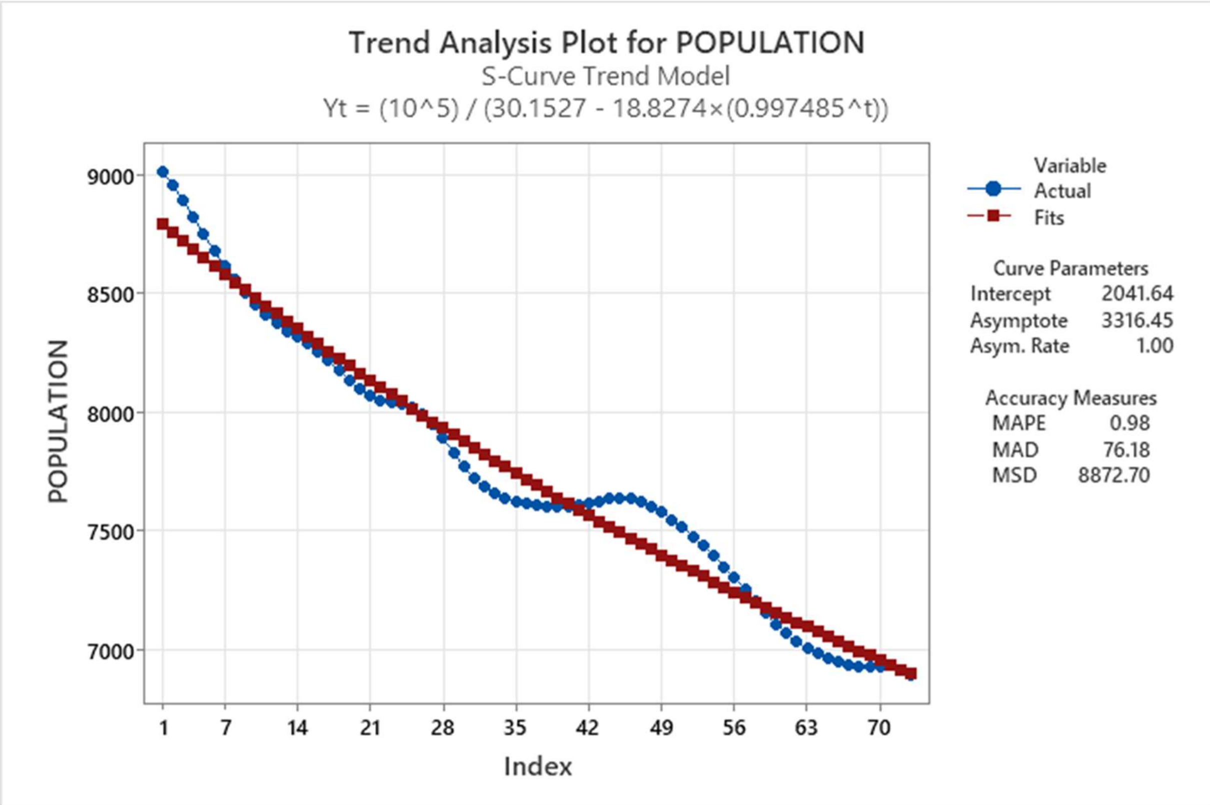
Model type	Growth Curve Model
Data	POPULATION
Length	73
NMissing	0



Graph for Logistic Curve for Austria

Method

Model type	S-Curve Trend Model
Data	POPULATION
Length	73
NMissing	0



Fitted Trend Equation

$$Y_t = (10^5) / (30.1527 - 18.8274 \times (0.997485^t))$$

Conclusion:- From all of the above curves Logistic Curve produces the least MAPE value which is 0.98<2% , which **implies best fit** , for our data .

The actual fitted Trend Equation thus given by :-

$$\boxed{Yt= (10\text{^}5)/(30.1527-18.8274*(0.997485\text{^}t))} \quad \ldots\ldots\ldots\ldots \textbf{(3)}$$

Where the values of the estimated parameters thus given by :-

^a = 5

^ β_0 = 30.1527

^β_1 = 18.8274

^ β_2 =0.997485

Now by putting the values of "t" (calculating through usual method) we will get the fitted values of the Population From the year 2020 to 1948.

TABLE SHOWING DATA FOR AUSTRIA (FITTED POPULATION) BASED ON THE TREND EQUATION Yt= (10^5)/(30.1527-18.8274*(0.997485^t))

YEARS	ACTUAL POPULATION DATA	t (1984-years)	FITTED POPULATION DATA
2020	9,006.40	36	2112.004
2019	8,955.11	35	2110.072
2018	8,891.38	34	2108.138
2017	8,819.90	33	2106.203
2016	8,747.31	32	2104.267
2015	8,678.67	31	2102.33
2014	8,615.21	30	2100.391
2013	8,556.19	29	2098.451
2012	8,502.23	28	2096.51
2011	8,453.50	27	2094.567
2010	8,409.95	26	2092.623
2009	8,372.66	25	2090.678
2008	8,341.53	24	2088.731
2007	8,313.73	23	2086.784
2006	8,285.35	22	2084.835
2005	8,253.66	21	2082.884
2004	8,216.81	20	2080.933
2003	8,175.86	19	2078.98
2002	8,134.40	18	2077.026
2001	8,097.76	17	2075.071
2000	8,069.28	16	2073.114
1999	8,051.12	15	2071.156
1998	8,041.05	14	2069.197
1997	8,032.87	13	2067.237
1996	8,017.85	12	2065.276
1995	7,990.11	11	2063.313
1994	7,947.22	10	2061.349
1993	7,892.09	9	2059.384
1992	7,830.94	8	2057.418
1991	7,772.70	7	2055.451
1990	7,723.95	6	2053.482
1989	7,686.56	5	2051.512
1988	7,658.86	4	2049.541
1987	7,639.35	3	2047.569
1986	7,625.40	2	2045.596
1985	7,614.87	1	2043.621
1984	7,607.30	0	2041.645

1983	7,603.10	-1	2039.669
1982	7,602.18	-2	2037.691
1981	7,604.51	-3	2035.712
1980	7,609.75	-4	2033.731
1979	7,617.88	-5	2031.75
1978	7,627.76	-6	2029.768
1977	7,636.57	-7	2027.784
1976	7,640.76	-8	2025.799
1975	7,637.69	-9	2023.814
1974	7,626.29	-10	2021.827
1973	7,607.07	-11	2019.839
1972	7,581.12	-12	2017.85
1971	7,550.34	-13	2015.859
1970	7,516.24	-14	2013.868
1969	7,479.34	-15	2011.876
1968	7,439.58	-16	2009.882
1967	7,397.17	-17	2007.888
1966	7,352.24	-18	2005.892
1965	7,305.15	-19	2003.896
1964	7,256.09	-20	2001.898
1963	7,205.99	-21	1999.899
1962	7,156.84	-22	1997.9
1961	7,111.14	-23	1995.899
1960	7,070.77	-24	1993.897
1959	7,036.55	-25	1991.894
1958	7,008.42	-26	1989.89
1957	6,985.72	-27	1987.886
1956	6,967.62	-28	1985.88
1955	6,953.40	-29	1983.873
1954	6,942.62	-30	1981.865
1953	6,935.15	-31	1979.856
1952	6,931.19	-32	1977.847
1951	6,931.35	-33	1975.836
1950	6,936.44	-34	1973.824
1949	6,920	-35	1971.811
1948	6900	-36	1969.798

LINE DIAGRAM OF LOGISTIC CURVE PLOTTED AGAINST FITTED POPULATION AGAINST AUSTRIA

Method

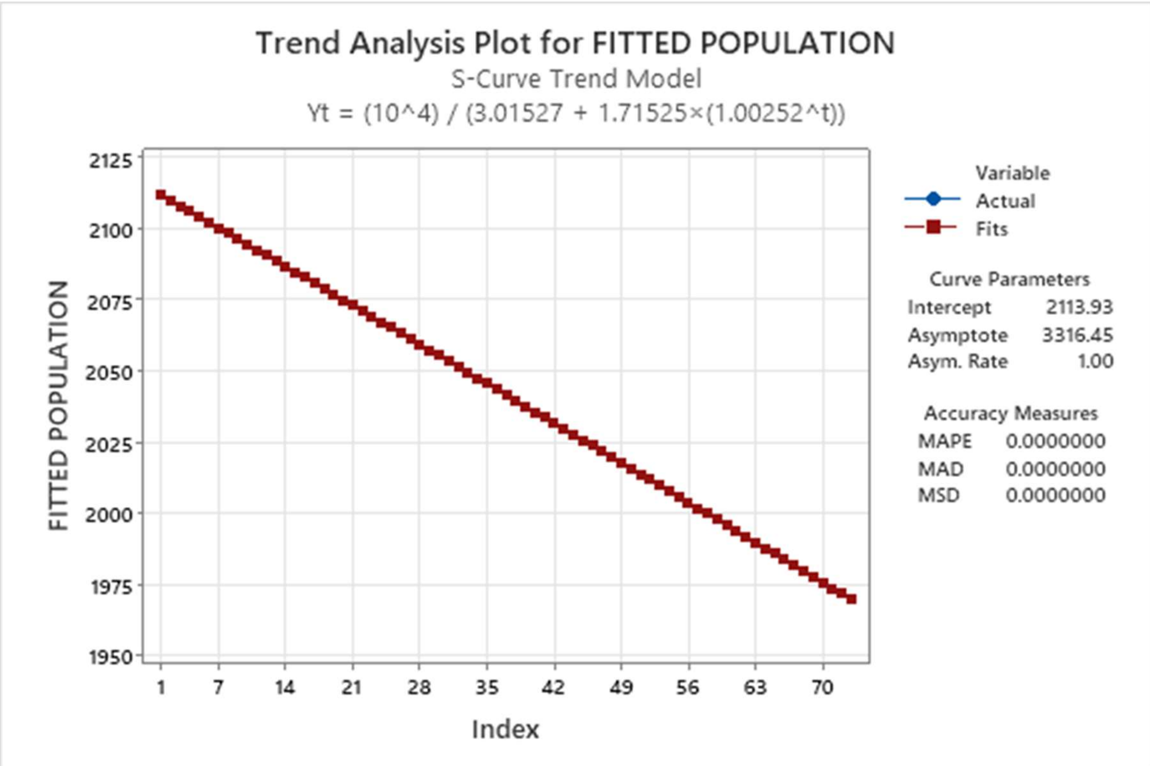
Model type	S-Curve Trend Model
Data	FITTED POPULATION (YT)
Length	73
NMissing	0

Accuracy Measures

MAPE	0.0000000
MAD	0.0000000
MSD	0.0000000

Fitted Trend Equation

$$Yt = (10^4) / (3.01527 + 1.71525\times(1.00252^t))$$



Thus from the above figure we can see that Fitted Population gives a Straight Line , when plotted against a Pearl-Reed (Logistic Curve) with MAPE value almost 0 , so thus on the basis of our Analysis we can conclude that our curve fitting is near to perfection.

CONCLUSION

From the above done study it is readily observed that the Population Projection method done using the Logistic Curve Equation is more than satisfactory if we consider European countries. For the above three countries we have

Austria and Greece STRAIGHT LINE ie. the fit is very good and for **Denmark almost a line ie. the fit is good**. Hence the results obtained justifies the fact that **European Countries has experienced the Three phases of Logistic Curve and thus can be modelled successfully using the Logistic Equation**. Hence in our case the Population Projection Method used is a success.

The study helps us to understand different scenarios which affected the population of the countries through drawing of Line Diagram. As such we gain an idea about the past history of these countries.

Although there may be different mathematical models which can be used to model a Population Data but for European Countries The Logistic Curve yields satisfactory results.

REFERENCE & SOURCES

1. For **DENMARK** :- Sources: UN DESA; Macmillan Publishers(Google Books)
   LINK https://www.statista.com/statistics/1008531/total-population-denmark-1769-2020/

2. **GREECE** :- Sources: UN DESA; Macmillan Publishers(Google Books)

   LINK:- https://www.statista.com/statistics/1014317/total-population-greece-1821-2020/

3. For **AUSTRIA** :- Sources: UN DESA; Macmillan Publishers(Google Books)
   https://www.statista.com/statistics/1008043/population-austria-1910-2020/

4. Fundamentals of Statistics, Volume II ,A.M. GUN, M.K. DASGUPTA,

   B.DASGUPTA

5. Fundamentals Of Applied Statistics, S.C. GUPTA, V.K. KAPOOR

6. The growth of population and the factors which controls it, H.G. Yule, 1925

7. The fitting of growth curves, K.R. Nair, 1954.

8. A special thanks to Minitab for making the project much much easier.