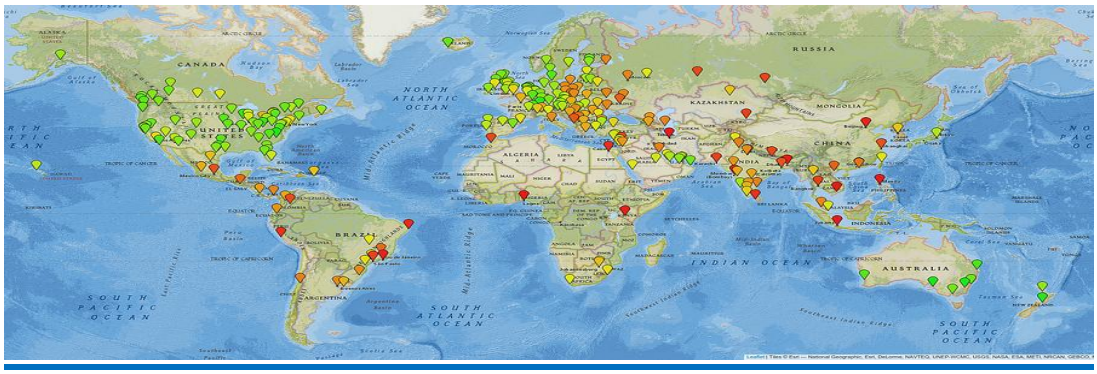
RAM LAL ANAND COLLEGE

UNIVERSITY OF DELHI

DEPARTMENT OF STATISTICS

LINEAR MODELS PROJECT: -

Comparison of Quality of life index 2019, 2021 and 2022



Dr. Mukta Datta Mazumdar
Department of Statistics
Ram Lal Anand College

Prabhjyot Kaur (5019)
Adishi Agrawal (5039)

ACKNOWLEDGEMENT

We would like to express our gratitude towards our honorable principal Dr. Rakesh Kumar Gupta of Ram Lal Anand College for their support in the accomplishment of our project on the Quality-of-Life Index by Country 2022. Our sincere thanks go to Dr. Mukta Datta Mazumdar, as without her support and guidance the completion of this project would not have been possible. We would like to extend our deep appreciation to our group members without their support, cooperation, and coordination we would not have been able to complete this project.

INDEX

AIM AND OBJECTIVES

To fit the multiple regression and simple linear regression to the data which here in our study is the quality of life index 2022 which ranks various countries based on certain factors and parameters and we aim to check whether the rank and quality of life index allotted to various countries in different continents are related to parameters on which it is measured. Also, we are checking whether the quality of life index counted does depend on the given parameters in the population.

INTRODUCTION AND ABOUT DATA

The quality of life is a concept made diverse definitions for several disciplines. In general, it can be considered as the sum of quantitative and qualitative value that an individual's lifestyle, health, relation with the community. QOL concept should not be confused with the income-based concept of standard of living. It refers to the physical, psychological, and sociological state of being of people. It is very arduous, to make a single definition of the concept of life quality that includes large scope. In the literature, there are various definitions of the concept of life quality.

Quality of life is a broader concept than happiness because it entails factors such as enjoyment and achievement. Quality of life is also broader than satisfaction because it concerns variables such as aspiration and recollection. It is also broader than wellbeing because the quality of life is neutral. It is broader than health because it involves being in the context of one or several. It has long been accepted that material wellbeing, as measured by GDP per person, cannot alone explain the broader quality of life in a country. One strand of the literature has tried to adjust GDP by quantifying facets that are omitted by the GDP measure—various nonmarket activities and social ills such as environmental pollution. But the approach has faced insurmountable difficulties in assigning monetary values to the various factors and intangibles that comprise a wider measure of socio-economic well-being. There have been numerous attempts to construct alternative, non-monetary indices of social and economic wellbeing by combining in a single statistic a variety of different factors that are thought to influence the quality of life. To summarize life quality combines both quantitative and qualitative indicators.

This report is written based on secondary data collected from Source:

Quality of Life Index by Country 2022 by Numbeo.

In this study 87 countries have been taken into consideration and ranked accordingly. Moreover, statistical methodologies were used to analyze the secondary data, and inferences were drawn accordingly based on the study.

The data was further analyzed by drawing ANOVA tables, plots, etc., and suitable statistical tests with help of Microsoft Excel and R programming.

We have tried to explain the columns of the data which are considered as parameters to fit multiple regression model and simple linear regression model to the data taken which are listed as follows.

- Rank- A country's overall score reflects the weighted sum of its sub-ranking scores. The sub ranking and overall scores were rescaled so that the top country in each sub-category received a value of 100, and others were calculated as a proportion of that top score. Scores were ranked in descending order.
- Quality of Life Index- *Quality of Life Index* (higher is better) is an estimation of overall quality of life by using an empirical formula which takes into account purchasing power index (higher is better), pollution index (lower is better), house price to income ratio (lower is better), cost of living index (lower is better), safety index (higher is better), health care index (higher is better), traffic commute time index (lower is better) and climate index (higher is better).
- Purchasing Power Index- It measures the weighted average of prices of consumer goods and services, in particular, transportation, food, and medical care.
- Safety Index-THE SAFETY INDEX REPRESENTS THE PERCENTAGE OF THE PROJECT'S CONSTRUCTION PLUS RIGHT-OF-WAY COSTS THAT IS RETURNED TO THE MOTORISTS AS SAVINGS IN THE COST OF PREVENTED ACCIDENTS. IN REALITY, IT IS A SAFETY BENEFIT-COST RATIO. THE DECIMAL POINT HAS BEEN MOVED TO THE RIGHT 2 PLACES, AND THE RATIO IS CALLED AN INDEX.
- Health Care Index-Health Care Index is an estimation of the overall quality of the health care system, health care professionals, equipment, staff, doctors, cost, etc.
- Cost of Living Index-A cost-of-living index is a theoretical price index that measures the relative cost of living over time or regions. It is an index that measures differences in the price of goods and services and allows for substitutions with other items as prices vary.
- Property Price to Income Ratio-Price-to-income ratio is the ratio between the price of a median home to that of the median annual household income in a particular area. The concept of price-to-income ratio is used to measure the affordability of homes in a certain area.

- Traffic Commute Time Index- a composite index of time consumed in traffic due to job commute, estimation of time consumption dissatisfaction, CO2 consumption estimation in traffic, and overall inefficiencies in the traffic system.
- Pollution Index-An air pollution index (API) is a quantitative measure that describes ambient air quality. The index is obtained by combining figures for various air pollutants into a single measurement.
- Climate Index- A climate index is a simple diagnostic quantity that is used to characterize an aspect of a geophysical system such as a circulation pattern.

Also, for the sake of simplicity and to find out the valuable results we had divided the data into 4 regions per the countries provided in the data:-

- AMERICA
- ASIA
- EUROPE
- OCEANIA AND AFRICA

We had combined the countries in Oceania and Africa as the number of observations are less than the parameters taken which would lead to inappropriate and insignificant results.

DATA ANALYSIS

- Assuming the data to be normal, we had carried out our analysis.
- We are comparing the three dataset models regarding the quality-of-life index releases for three different years i.e., 2019, 2021 and 2022. We are analyzing how well different parameters are able to predict the quality of life index for various countries.
- In every region of the world, we had divided the countries according to the continent in which it lies.
- First, we check how well rank is related to all the independent regressors in the model using ANOVA.
- Second, we check how well quality of life is related to all the independent regressors in the model using ANOVA.
- Third, we carried out partial f test to check the significance of some of the unimportant regressors in the model.
- Fourth, we had tried to represent the correlation between several parameters numerically.

VERIFYING THE PARAMETERS ABLE TO PREDICT THE INDEX CALCULATED

- **FOR YEAR 2019**

The coefficients obtained for the model when Quality of life is a response variable are interpreted in the predicted model equation as follows: -

$$\hat{Y} = 99.9785 + 0.40005\,X_1 + 0.50008X_2 + 0.40007\,X_3 - 0.1\,X_4 - 1\,X_5 - 0.4998\,X_6 - 0.6666\,X_7 + 0.33334\,X_8$$

Where,

X_1 = Purchasing power index
X_2 = Safety index
X_3 = Health care index
X_4 = Cost of living index
X_5 = Property price to income ratio
X_6 = Traffic commute time index
X_7 = Pollution index
X_8 = Climate index.

TABLE:

	Coefficients
Intercept	99.9784689
X Variable 1	0.40004982
X Variable 2	0.50007929
X Variable 3	0.40007291
X Variable 4	-0.1000159
X Variable 5	-0.9999726
X Variable 6	-0.4998206
X Variable 7	-0.6666151
X Variable 8	0.33334495

From the dataset taken, we can take the three countries having best, worst index and taking third country for verifying the quality of index with that which is given in the dataset :-

1ST COUNTRTY

Best score is maintained by Denmark. Using the predicted model equation written above, we can calculate $\hat{Y}$ which is quality of life index Using the corresponding values of Xi (i = 1,2,…,8).

Hence $\hat{Y}$ obtained for Denmark is 198.5807329 which is very close to Y which is 198.57 given as in the dataset.

2nd COUNTRTY

Worst score is maintained by Egypt. Using the predicted model equation written above, we can calculate $\hat{Y}$ which is quality of life index
Using the corresponding values of Xi (i = 1,2,…,8)

The table for intercept and slopes is same as Table *

Hence $\hat{Y}$ obtained for Egypt is 83.9777494 which is very close to Y which is given as 83.98 in the dataset.

3rd COUNTRY

We take third country as India. Using the predicted model equation written above, we can calculate $\hat{Y}$ which is quality of life index
Using the corresponding values of Xi (i = 1,2,…,8)

Hence $\hat{Y}$ obtained for India is 117.5074592 which is very close to Y which is given as 117.51 in the dataset.

Hence, the predicted model is same as the original model and we can say that the data is significant in nature.

- **FOR YEAR 2021**

The coefficients obtained for the model when Quality of life is a response variable are interpreted in the predicted model equation as follows: -

$$\hat{Y} = 100.0055 + 0.39997 X_1 + 0.49990 X_2 + 0.40000 X_3 - 0.0999 X_4 - 0.9999 X_5 - 0.5000 X_6 - 0.6666 X_7 + 0.33329 X_8$$

Where,

X_1 = Purchasing power index
X_2 = Safety index
X_3 = Health care index
X_4 = Cost of living index
X_5 = Property price to income ratio
X_6 = Traffic commute time index
X_7 = Pollution index
X_8 = Climate index.

From the dataset taken, we can take the three countries having best, worst index and taking third country for verifying the quality of index with that which is given in the dataset :-

Table **

	Coefficients
Intercept	100.0055409
X Variable	0.399974388
X Variable	0.499905235
X Variable	0.400003841
X Variable	-0.09991795
X Variable	-0.999953004
X Variable	-0.500085744
X Variable	-0.666623688
X Variable	0.333295881

1ST COUNTRTY

Best score is maintained by Switzerland. Using the predicted model equation written above, we can calculate $\hat{Y}$ which is quality of life index Using the corresponding values of Xi (i = 1,2,…,8)

Hence $\hat{Y}$ obtained for Switzerland is 190.8291395 which is very close to Y which is given as 190.82 in the dataset.

2nd COUNTRTY

Worst score is maintained by Nigeria. Using the predicted model equation written above, we can calculate $\hat{Y}$ which is quality of life index Using the corresponding values of Xi (i = 1,2,…,8)

The table for intercept and slopes is same as Table *

Hence $\hat{Y}$ obtained for Nigeria is 52.00218 which is very close to Y which is given as 52 in the dataset.

3rd COUNTRY

We take third country as India. Using the predicted model equation written above, we can calculate $\hat{Y}$ which is quality of life index Using the corresponding values of Xi (i = 1,2,…,8)

Hence $\hat{Y}$ obtained for India is 104.5215878 which is very close to Y which is given as 104.52 in the dataset.

Hence, the predicted model is same as the original model and we can say that the data is significant in nature.

- **FOR YEAR 2022**

The coefficients obtained for the model when Quality of life is a response variable are interpreted in the predicted model equation as follows: -

$$\hat{Y} = 100.0075 + 0.399964\,X_1 + 0.500010\,X_2 + 0.399998\,X_3 - 0.10002\,X_4 - 0.99993\,X_5 - 0.49999\,X_6 - 0.66669\,X_7 + 0.333268\,X_8$$

Where,

X_1 = Purchasing power index
X_2 = Safety index
X_3 = Health care index
X_4 = Cost of living index
X_5 = Property price to income ratio
X_6 = Traffic commute time index
X_7 = Pollution index
X_8 = Climate index.

From the dataset taken, we can take the three countries having best, worst index and taking third country for verifying the quality of index with that which is given in the dataset :-

1ST COUNTRTY

Best score is maintained by Switzerland. Using the predicted model equation written above, we can calculate $\hat{Y}$ which is quality of life index Using the corresponding values of Xi (i = 1,2,…,8)

Table *

	Coefficients
Intercept	100.0075772
X Variable	0.399964854
X Variable	0.500010422
X Variable	0.399998311
X Variable	-0.1000232
X Variable	-0.999937522
X Variable	-0.499990435
X Variable	-0.666694454
X Variable	0.333268822

Hence $\hat{Y}$ obtained for Switzerland is 195.273981 which is very close to Y which is given as 195.27 in the dataset.

2nd COUNTRTY

Worst score is maintained by Nigeria. Using the predicted model equation written above, we can calculate $\hat{Y}$ which is quality of life index Using the corresponding values of Xi (i = 1,2,…,8)

The table for intercept and slopes is same as Table *

Hence $\hat{Y}$ obtained for Nigeria is 52.44006257 which is very close to Y which is given as 52.44 in the dataset.

3rd COUNTRY

We take third country as India. Using the predicted model equation written above, we can calculate $\hat{Y}$ which is quality of life index
Using the corresponding values of Xi (i = 1,2,…,8)

Hence $\hat{Y}$ obtained for India is 110.9845774 which is very close to Y which is given as 110.99 in the dataset.

Hence, the predicted model is same as the original model and we can say that the data is significant in nature.

DIVIDING THE WORLD INTO 4 REGIONS viz EUROPE, AMERICA, ASIA, OCEANIA AND AFRICA AND ANALYSING THE MODEL REGARDING THE QUALITY OF LIFE INDEX RELEASES IN 2022

1. EUROPE

1. Multiple Linear Regression when Rank is Response variable: -

Null Hypothesis: -
H_0: $\beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = \beta_7 = \beta_8 = \beta_9 = 0$
(There is no relation between X & Y)

Alternative Hypothesis: -
H_1: $\beta_j \neq 0$ for all at least one j = 1, 2,….,9
(There exists a relation between X and Y)

Table 1.1: - ANOVA Table for testing significance of hypothesis

Source of variation	*d.f*	*SS*	*MS*	*F*	*Significance F*
Regression	9	13596.44626	1510.716	673.6896283	1.82832E-28
Residual	26	58.30373645	2.242451		
Total	35	13654.75			

From the table given the calculated F statistic (F_o) is 673.6896283 for 9 and 26 degrees of freedom for regression and residuals. We are checking the significance of the model at a 5% level of significance.
The tabulated value of F at 0.05 with corresponding degrees of freedom is 2.2655. On comparing the two values we found that
$F_o > F_{0.05,9,26}$. Hence, we rejected our null hypothesis.

Conclusion: - We rejected our Ho and we can conclude that there exists some relationship between the explanatory variables and response variable for the multiple linear regression model.
We can say that the Rank allotted to different countries in the Europe region of the world does on different explanatory variables taken which are **Quality of life index, Purchasing power index, Safety index, Health care index, Cost of living index, Property price to income ratio, Traffic commute time index, Pollution index, Climate index.**

Hence, the predicted model for rank as a response variable obtained as :-

$$\hat{y} = 6605.1058 - 65.4296\,X_1 + 25.8707\,X_2 + 32.3807\,X_3 + 25.8445\,X_4 - 6.431\,X_5 - 64.4968\,X_6 - 32.2063\,X_7 - 43.1272\,X_8 + 21.5268\,X_9$$

Where, the independent variables or regressors in the model are as follows:-

X_1 = Quality of life index
X_2 = Purchasing power index
X_3 = Safety index
X_4 = Health care index
X_5 = Cost of living index
X_6 = Property price to income ratio
X_7 = Traffic commute time index
X_8 = Pollution index
X_9 = Climate index.

2. Multiple Linear Regression when Quality of Life is Response variable: -

Null Hypothesis: -
H_0: $\beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = \beta_7 = \beta_8 = 0$
(There is no relation between X & Y)

Alternative Hypothesis: -
H_1: $\beta_j \neq 0$ for all at least one j = 1, 2,….,8
(There exists a relation between X and Y)

Table 1.2: - ANOVA Table for testing significance of hypothesis

Soure of variation	*df*	*SS*	*MS*	*F*	*Significance F*
Regression	8	24412.59	3051.574	119706256.1	7.3816E-100
Residual	27	0.000688	2.55E-05		
Total	35	24412.59			

From the table given the calculated F statistic (F_o) is 119706256.1 for 8 and 27 degrees of freedom for regression and residuals. We are checking the significance of the model at 5% level of significance.
The tabulated value of F at 0.05 with corresponding degrees of freedom is 2.3053. On comparing the two values we found that
$F_o > F_{0.05,8,27}$. Hence, we rejected our null hypothesis.

Conclusion: - We rejected our Ho and we can conclude that there is some relationship between the explanatory variables and response variable for the multiple linear regression model.

We can say that the Quality-of-life index allotted to different countries in the Europe region of the world does on different explanatory variables taken which are **Purchasing power index, Safety index, Health care index, Cost of living index, Property price to income ratio, Traffic commute time index, Pollution index, Climate index.**

Hence, the predicted model for quality of life index as a response variable obtained as :-

$$\hat{y} = 100.0336 + 0.3999\,X_1 + 0.4999\,X_2 + 0.3998\,X_3 - 0.1001\,X_4 - 1.0005\,X_5 - 0.5002\,X_6 - 0.6668\,X_7 + 0.33336\,X_8$$

Where, the independent variables or regressors in the model are as follows:-

X_1 = Purchasing power index
X_2 = Safety index
X_3 = Health care index
X_4 = Cost of living index
X_5 = Property price to income ratio
X_6 = Traffic commute time index
X_7 = Pollution index
X_8 = Climate index.

3. ***SIMPLE LINEAR REGRESSION BETWEEN QUALITY-OF-LIFE INDEX AND RANK ALLOTTED TO THE COUNTRIES***

This shows the linear negative relationship between rank and quality of life index as quality-of-life index increases rank decreases.

Null Hypothesis: -
H_0: $\beta_1 = 0$ (There is no relation between X & Y)

Alternative Hypothesis: -
H_1: $\beta_1 ! = 0$ (There exists a relation between X and Y)

Table 1.3: - ANOVA Table for testing significance of hypothesis

Soure of variation	*df*	*SS*	*MS*	*F*	*Significance F*
Regression	1	13559.38	13559.38	4834.241997	3.04972E-38
Residual	34	95.36533	2.804863		
Total	35	13654.75			

From the table given the calculated F statistic (F_o) is 4834.241997 for 1 and 34 degrees of freedom for regression and residuals. We are checking the significance of the model at 5% level of significance.
The tabulated value of F at 0.05 with corresponding degrees of freedom is 4.11. On comparing the two values we found that
$F_o > F_{0.05,1,34}$. Hence, we rejected our null hypothesis.

Conclusion: - We rejected our Ho and we can conclude that there exists some relationship between the explanatory variables and response variable for the multiple linear regression model.
We can say that the Rank allotted to different countries in the Europe region of the world does on **Quality-of-life index.**

Hence, the predicted model for quality of life index as a response variable obtained as :-

$\hat{y} = 142.436201 - 0.745269013\ X$ **Where, X = Rank**

4. PARTIAL F TEST TO CHECK SIGNIFICANCE OF REDUCED MODEL WHEN RANK IS A RESPONSE VARIABLE

Null Hypothesis :- Th full model and reduced model do not differ significantly

Where Full model and reduced model are

```
model<-lm(Rank~Quality.of.life.index+Purchasing.power.index+Safety.index +
Health.care.index+Cost.of.living.index+Property.price.to.income.ratio
+Traffic.commute.time.index + Pollution.index + Climate.index)
summary(model)
```

```
reduced_model<-
lm(Rank~Quality.of.life.index+Purchasing.power.index+Safety.index +
Health.care.index)
summary(reduced_model)
```

Alternative Hypothesis :- It is necessary to include cost other f living index, property price to income ratio, traffic commute time index, pollution index, and climate index in the model, and the Full model is significant.

Conclusion:- By carrying out the partial test for a set of regressors to check their significance, we obtained an F value as 2.1258. and the p value obtained is 0.09414 at 5% level of significance. On comparing the p-value of the test conducted and $\alpha=0.05$, the p-value is greater than α and hence we reject the null hypothesis being considered and the two models do not differ significantly. On removing the cost of living index, property price to income ratio, traffic commute time index, pollution index, and climate index, the full model and the model omitting these are not different. Hence, we can omit the said regressors which do not contribute significantly to the model.

5. MULTIPLE LINEAR REGRESSION OF REDUCED MODEL WHEN RANK IS A RESPONSE VARIABLE

Null Hypothesis: -
H_0: $\beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$
(There is no relation between X & Y)

Alternative Hypothesis: -
H_1: $\beta_j \neq 0$ for all at least one j = 1,2,3,4
(There exists a relation between X and Y)

TABLE 1.5: - ANOVA Table for testing significance of hypothesis

	df	SS	MS	F	Significance F	
Regressio	4	13572.61	3393.153	1280.618	6.22E-34	
Residual	31	82.13829	2.649622			
Total	35	13654.75				

From the table given the calculated F statistic (F_o) is 1280.618 for 4 and 31 degrees of freedom for regression and residuals. We are checking the significance of the model at 5% level of significance.
The tabulated value of F at 0.05 with corresponding degrees of freedom is 2.6896. On comparing the two values we found that
$F_o > F_{0.05,4,31}$. Hence, we rejected our null hypothesis.

Conclusion: - We rejected our Ho and we can conclude that there exists some relation between the leftover explanatory variables in the reduced model and response variable for the multiple linear regression model.
We can say that the Rank allotted to different countries in the Europe region of the world does on leftover explanatory variables taken which are **Quality of life index, Purchasing power index, Safety index, Health care index.**

R^2 value for this model is 0.99398 means 99.398% of the variation explained by the regressors x.

6. CORRELATION BETWEEN VARIABLES IN REDUCED MODEL

	RANK	*QOLI*	*PURCHASING POWER INDEX*	*SAFETY INDEX*	*HEALTH CARE INDEX*
RANK	1				
QOLI	-0.9965	1			
PURCHASING POWER INDEX	-0.83361	0.848179	1		
SAFETY INDEX	-0.56349	0.570637	0.243396415	1	
HEALTH CARE INDEX	-0.82717	0.829936	0.733473956	0.378899349	1

Through above the table we can see that there is negative correlation between rank and other variables while there is positive correlation between other variables(excluding rank).

2. *AMERICA*


United States
Mexico
Cuba
Haiti
Dominican Republic
Puerto Rico
Belize
Cayman Islands 1 species
Jamaica
Martinique - 1 species
Trinidad & Tobago - 2 species
Guyana - 4 species
Suriname
French Guiana 4 species
Guatemala
El Salvador
Honduras
Nicaragua
Costa Rica
Panama
Venezuela 5
Colombia 6
Ecuador 5 species
Peru 5
Brazil 6
Bolivia 5
Paraguay
Uruguay
Argentina
CROCODILIANS PER COUNTRY
= 0
= 1
= 2
= 3
= 4 - 6

1. Multiple Linear Regression when Rank is Response variable: -

Null Hypothesis: -
H_0: $\beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = \beta_7 = \beta_8 = \beta_9 = 0$
(There is no relation between X & Y)

Alternative Hypothesis: -
H_1: $\beta_j \neq 0$ for all at least one j = 1, 2,….,9
(There exists a relation between X and Y)

Table 2.1: - ANOVA Table for testing significance of hypothesis

Source of variation	*df*	*SS*	*MS*	*F*	*Significance F*
Regression	9	5182.858	575.8731	79.82560961	0.000374503
Residual	4	28.85656	7.214139		
Total	13	5211.714			

From the table given the calculated F statistic (F_o) is 79.82560961 for 9 and 4 degrees of freedom for regression and residuals. We are checking the significance of the model at 5% level of significance.

The tabulated value of F at 0.05 with corresponding degrees of freedom is 5.9988. On comparing the two values we found that
$F_o > F_{0.05,9,4.}$ Hence, we rejected our null hypothesis.

Conclusion: - We rejected our Ho and we are able to conclude that there exists some relation between the explanatory variables and response variable for multiple linear regression model.
We can say that Rank allotted to different countries in America region of the world does depends on different explanatory variables taken which are **Quality of life index, Purchasing power index, Safety index, Health care index, Cost of living index, Property price to income ratio, Traffic commute time index, Pollution index, Climate index.**

Hence, the predicted model for rank as a response variable obtained as: -

$$\hat{y} = -15569.04296 + 156.5147\ X_1 - 62.7484\ X_2 - 78.3760\ X_3 - 63.2493\ X_4 + 15.4638 X_5 + 157.6748\ X_6 + 78.58570\ X_7 + 104.8441\ X_8 - 52.44339\ X_9$$

Where, the independent variables or regressors in the model are as follows:-

X_1 = Quality of life index
X_2 = Purchasing power index
X_3 = Safety index
X_4 = Health care index
X_5 = Cost of living index
X_6 = Property price to income ratio
X_7 = Traffic commute time index
X_8 = Pollution index
X_9 = Climate index.

2. Multiple Linear Regression when Quality of Life is Response variable: -

Null Hypothesis: -
H_0: $\beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = \beta_7 = \beta_8 = 0$
(There is no relation between X & Y)

Alternative Hypothesis: -
H_1: β_j! = 0 for all at least one j = 1, 2,….,8
(There exists a relation between X and Y)

Table 2.2: - ANOVA Table for testing significance of hypothesis

Source of variation	*df*	*SS*	*MS*	*F*	*Significance F*
Regression	8	8776.76	1097.095038	29570087.85	9.37693E-19
Residual	5	0.000186	3.71015E-05		
Total	13	8776.76			

From the table given the calculated F statistic (F_o) is 29570087.85 for 8 and 5 degrees of freedom for regression and residuals. We are checking the significance of the model at 5% level of significance.
The tabulated value of F at 0.05 with corresponding degrees of freedom is 4.8183. On comparing the two values we found that
$F_o > F_{0.05,8,5}$. Hence, we rejected our null hypothesis.

Conclusion: - We rejected our Ho and we are able to conclude that there exists some relation between the explanatory variables and response variable for multiple linear regression model.
We can say that Quality of life index allotted to different countries in America region of the world does depends on different explanatory variables taken which are **Purchasing power index, Safety index, Health care index, Cost of living index, Property price to income ratio, Traffic commute time index, Pollution index, Climate index.**

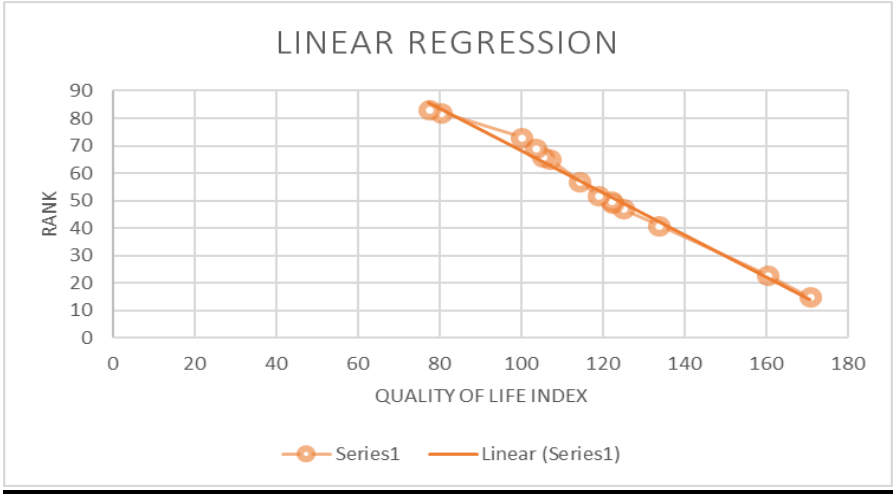
Hence, the predicted model for quality of life index as a response variable obtained as: -

$$\hat{y} = 99.9810 + 0.3998\ X_1 + 0.4996\ X_2 + 0.4005\ X_3 - 0.0998\ X_4 - 1.0005\ X_5 - 0.5001 X_6 - 0.6665\ X_7 + 0.33333\ X_8$$

Where, the independent variables or regressors in the model are as follows:-

X_1 = Purchasing power index
X_2 = Safety index
X_3 = Health care index
X_4 = Cost of living index
X_5 = Property price to income ratio
X_6 = Traffic commute time index
X_7 = Pollution index
X_8 = Climate index.

3. ***SIMPLE LINEAR REGRESSION BETWEEN QUALITY-OF-LIFE INDEX AND RANK ALLOTED TO THE COUNTRIES***



This shows linear negative relationship between rank and quality of life index as quality of life index increases rank decreases.

Null Hypothesis: -
H_0: $\beta_1 = 0$ (There is no relation between X & Y)

Alternative Hypothesis: -
H_1: $\beta_1 ! = 0$ (There exists a relation between X and Y)

Table 2.3: - ANOVA Table for testing significance of hypothesis

Source of variation	*df*	*SS*	*MS*	*F*	*Significance F*
Regression	1	5141.451	5141.451	878.0919367	1.36243E-12
Residual	12	70.26305	5.855254		
Total	13	5211.714			

From the table given the calculated F statistic (F_o) is 878.0919367 for 1 and 12 degrees of freedom for regression and residuals. We are checking the significance of the model at 5% level of significance.
The tabulated value of F at 0.05 with corresponding degrees of freedom is 4.7472. On comparing the two values we found that
$F_o > F_{0.05,1,12}$. Hence, we rejected our null hypothesis.

Conclusion: - We rejected our Ho and we are able to conclude that there exists some relation between the explanatory variables and response variable for multiple linear regression model.
We can say that Rank allotted to different countries in America region of the world does depends on **Quality-of-life index.**

Hence, the predicted model for quality of life index as a response variable obtained as :-

$\hat{y}$ = 144.8702578 - 0.765377582 X **Where X = Rank**

4. PARTIAL F TEST TO CHECK SIGNIFICANCE OF REDUCED MODEL WHEN RANK IS A RESPONSE VARIABLE

Null Hypothesis: - The full model and reduced model do not differ significantly

Where Full model and reduced model are

```
model<-lm(Rank~Quality.of.life.index+Purchasing.power.index+Safety.index +
Health.care.index+Cost.of.living.index+Property.price.to.income.ratio
+Traffic.commute.time.index + Pollution.index + Climate.index)
summary(model)
```

```
reduced_model<-
lm(Rank~Quality.of.life.index+Purchasing.power.index+Safety.index +
Health.care.index)
summary(reduced_model)
```

Alternative Hypothesis: - It is necessary to include cost of living index, property price to income ratio, traffic commute time index, pollution index and climate index in the model and Full model is significant.

Conclusion: - By carrying out the partial test for set of regressors to check their significance, we obtained F value as 0.6168. and p value obtained is 0.6984 at 5% level of significance. On comparing the p value of the test conducted and α=0.05, p value is greater than α and hence we reject the null hypothesis being considered and the two models do not differ significantly.

On removing cost of living index, property price to income ratio, traffic commute time index, pollution index and climate index, the full model and the model omitting these are not different. Hence, we can omit out the said regressors which do not contribute significantly to the model.

5. MULTIPLE LINEAR REGRESSION OF REDUCED MODEL WHEN RANK IS A RESPONSE VARIABLE

Null Hypothesis: -
H_0: $\beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$
(There is no relation between X & Y)

Alternative Hypothesis: -
H_1: $\beta_j \neq 0$ for all at least one j = 1,2,3,4
(There exists a relation between X and Y)

TABLE 1.5: - ANOVA Table for testing significance of hypothesis

	df	SS	MS	F	Significance F	
Regressio	4	5160.61	1290.153	227.2106	4.99E-09	
Residual	9	51.10402	5.678225			
Total	13	5211.714				

From the table given the calculated F statistic (F_o) is 227.2106 for 4 and 9 degrees of freedom for regression and residuals. We are checking the significance of the model at 5% level of significance.
The tabulated value of F at 0.05 with corresponding degrees of freedom is 3.6331. On comparing the two values we found that
$F_o > F_{0.05,4,9}$. Hence, we rejected our null hypothesis.

Conclusion: - We rejected our Ho and we are able to conclude that there exists some relation between the leftover explanatory variables in the reduced model and response variable for multiple linear regression model.
We can say that Rank allotted to different countries in America region of the world does depends on leftover explanatory variables taken which are **Quality of life index, Purchasing power index, Safety index, Health care index.**

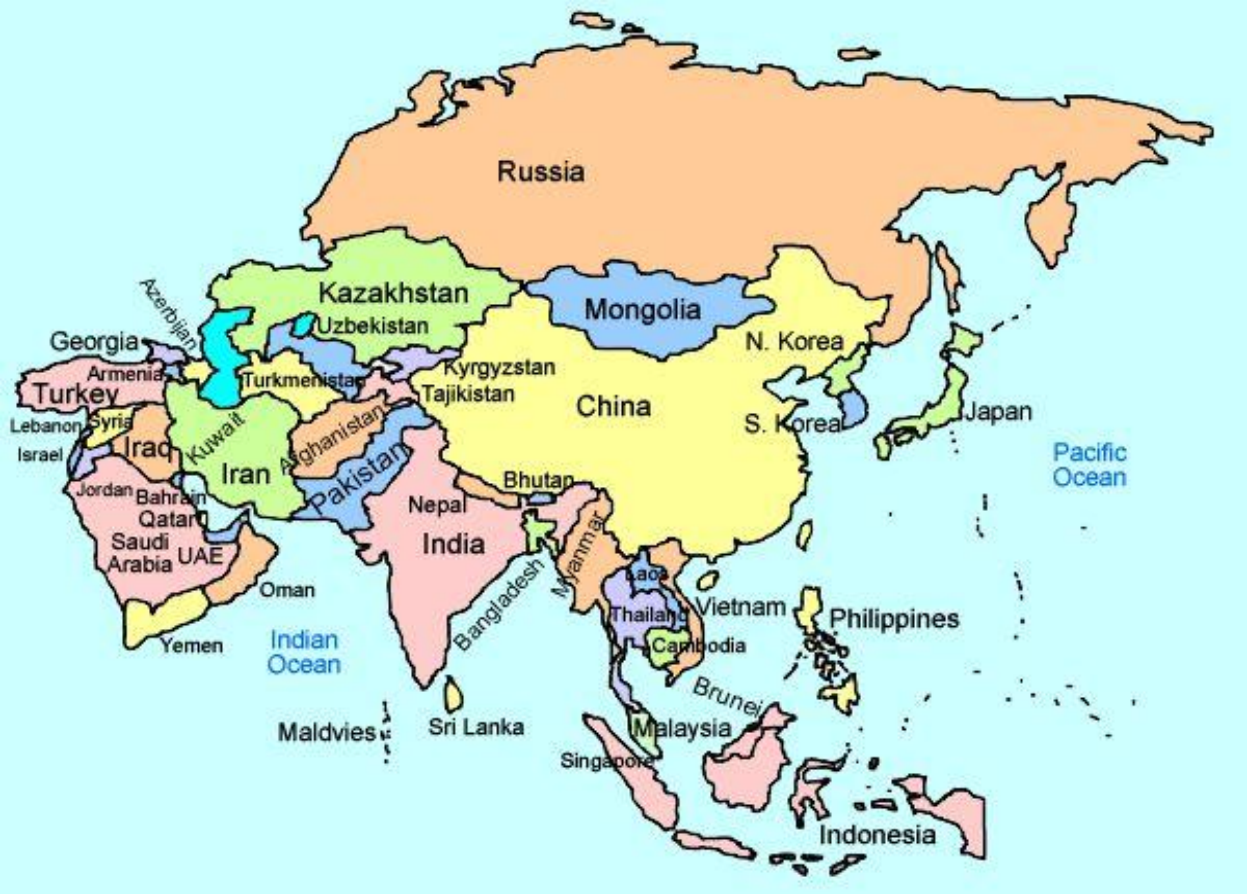
R^2 value for this model is 0.99019 means 99.019% of the variation explained by the regressors x.

6. CORRELATION BETWEEN VARIABLES IN REDUCED MODEL

	RANK	*QOLI*	*PURCHASING POWER INDEX*	*SAFETY INDEX*	*HEALTH CARE INDEX*
RANK	1				
QOLI	-0.99324	1			
PURCHASING POWER INDEX	-0.88796	0.902651	1		
SAFETY INDEX	-0.69245	0.718484	0.52418289	1	
HEALTH CARE INDEX	-0.58023	0.612856	0.375437422	0.774137365	1

Through above table we can see that there is negative correlation between rank and other variables while there is positive correlation between other variables.

3. ASIA


Russia
Kazakhstan
Mongolia
Uzbekistan
Azerbijan
Georgia
Armenia
Turkey
Turkmenistan
Kyrgyzstan
Tajikistan
China
N. Korea
S. Korea
Japan
Pacific Ocean
Lebanon
Syria
Israel
Iraq
Kuwait
Iran
Afghanistan
Pakistan
Bhutan
Nepal
India
Jordan
Bahrain
Qatar
Saudi Arabia
UAE
Oman
Yemen
Indian Ocean
Bangladesh
Myanmar
Laos
Thailand
Vietnam
Philippines
Cambodia
Brunei
Malaysia
Singapore
Maldvies
Sri Lanka
Indonesia

1. Multiple Linear Regression when Rank is Response variable: -

Null Hypothesis: -
H_0: $\beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = \beta_7 = \beta_8 = \beta_9 = 0$
(There is no relation between X & Y)

Alternative Hypothesis: -
H_1: $\beta_j \neq 0$ for all at least one j = 1, 2,….,9
(There exists a relation between X and Y)

Table 3.1: - ANOVA Table for testing significance of hypothesis

Source of variation	*df*	*SS*	*MS*	*F*	*Significance F*
Regression	9	13779.57	1531.064	121.9023152	7.0083E-15
Residual	19	238.6354	12.55976		
Total	28	14018.21			

From the table given the calculated F statistic (F_o) is 121.9023152 for 9 and 19 degrees of freedom for regression and residuals. We are checking the significance of the model at 5% level of significance.

The tabulated value of F at 0.05 with corresponding degrees of freedom is 2.4227. On comparing the two values we found that
$F_o > F_{0.05,9,19.}$ Hence, we rejected our null hypothesis.

Conclusion: - We rejected our Ho and we are able to conclude that there exists some relation between the explanatory variables and response variable for multiple linear regression model.
We can say that Rank allotted to different countries in Asia region of the world does depends on different explanatory variables taken which are **Quality of life index, Purchasing power index, Safety index, Health care index, Cost of living index, Property price to income ratio, Traffic commute time index, Pollution index, Climate index.**

Hence, the predicted model for rank as a response variable obtained as: -

$$\hat{y} = 4708.9634 - 46.4390\,X_1 + 18.2805\,X_2 + 22.8499\,X_3 + 18.3130\,X_4 - 4.5452\,X_5 - 45.688\,X_6 - 23.084\,X_7 - 30.359\,X_8 + 15.2563\,X_9$$

Where, the independent variables or regressors in the model are as follows:-

X_1 = Quality of life index
X_2 = Purchasing power index
X_3 = Safety index
X_4 = Health care index
X_5 = Cost of living index
X_6 = Property price to income ratio
X_7 = Traffic commute time index
X_8 = Pollution index
X_9 = Climate index.

2. ***Multiple Linear Regression when Quality of Life is Response variable: -***

Null Hypothesis: -
H_0: $\beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = \beta_7 = \beta_8 = 0$
(There is no relation between X & Y)

Alternative Hypothesis: -
H_1: $\beta_j \neq 0$ for all at least one j = 1, 2,….,8
(There exists a relation between X and Y)

Table 3.2 : - ANOVA Table for testing significance of hypothesis

Source of variation	*df*	*SS*	*MS*	*F*	*Significance F*
Regression	8	26064.42	3258.053	129784528.3	2.0116E-75
Residual	20	0.000502	2.51E-05		
Total	28	26064.42			

From the table given the calculated F statistic (F_o) is 129784528.3 for 8 and 20 degrees of freedom for regression and residuals. We are checking the significance of the model at 5% level of significance.

The tabulated value of F at 0.05 with corresponding degrees of freedom is 2.4471. On comparing the two values we found that
$F_o > F_{0.05,8,20}$. Hence, we rejected our null hypothesis.

Conclusion: - We rejected our Ho and we are able to conclude that there exists some relation between the explanatory variables and response variable for multiple linear regression model.
We can say that Quality of life index allotted to different countries in Asia region of the world does depends on different explanatory variables taken which are **Purchasing power index, Safety index, Health care index, Cost of living index, Property price to income ratio, Traffic commute time index, Pollution index, Climate index.**

Hence, the predicted model for quality-of-life index as a response variable obtained as: -

$$\hat{y} = 99.9925 + 0.40009\,X_1 + 0.49990\,X_2 + 0.40003\,X_3 - 0.1000\,X_4 - 0.9998\,X_5 - 0.4998\,X_6 - 0.6666\,X_7 + 0.33331\,X_8$$

Where, the independent variables or regressors in the model are as follows:-

X_1 = Purchasing power index
X_2 = Safety index
X_3 = Health care index
X_4 = Cost of living index
X_5 = Property price to income ratio
X_6 = Traffic commute time index
X_7 = Pollution index
X_8 = Climate index.

3. ***SIMPLE LINEAR REGRESSION BETWEEN QUALITY-OF-LIFE INDEX AND RANK ALLOTED TO THE COUNTRIES***

This shows linear negative relationship between rank and quality of life index as quality of life index increases rank decreases.
Null Hypothesis: -
H_0: $\beta_1 = 0$ (There is no relation between X & Y)

Alternative Hypothesis: -
H_1: $\beta_1 ! = 0$ (There exists a relation between X and Y)

Table 3.3 : - ANOVA Table for testing significance of hypothesis

Source of variation	*df*	*SS*	*MS*	*F*	*Significance F*
Regression	1	13661.92	13661.92	1035.308344	4.5338E-23
Residual	27	356.2916	13.19599		
Total	28	14018.21			

From the table given the calculated F statistic (F_o) is 1035.308344 for 1 and 27 degrees of freedom for regression and residuals. We are checking the significance of the model at 5% level of significance.
The tabulated value of F at 0.05 with corresponding degrees of freedom is 4.2100. On comparing the two values we found that
$F_o > F_{0.05,1,27}$. Hence, we rejected our null hypothesis.

Conclusion: - We rejected our Ho and we are able to conclude that there exists some relation between the explanatory variables and response variable for multiple linear regression model.
We can say that Rank allotted to different countries in Asia region of the world does depends on **Quality-of-life index.**

Hence, the predicted model for quality of life index as a response variable obtained as :-

$\hat{y}$ = 140.3621221 - 0.723988625 X **Where X = Rank**

4. PARTIAL F TEST TO CHECK SIGNIFICANCE OF REDUCED MODEL WHEN RANK IS A RESPONSE VARIABLE

Null Hypothesis :- The full model and reduced model do not differ significantly

Where Full model and reduced model are

```
model<-lm(Rank~Quality.of.life.index+Purchasing.power.index+Safety.index +
Health.care.index+Cost.of.living.index+Property.price.to.income.ratio
+Traffic.commute.time.index + Pollution.index + Climate.index)
summary(model)
```

```
reduced_model<-
lm(Rank~Quality.of.life.index+Purchasing.power.index+Safety.index +
Health.care.index)
summary(reduced_model)
```

Alternative Hypothesis: - It is necessary to include cost of living index, property price to income ratio, traffic commute time index, pollution index and climate index in the model and Full model is significant.

Conclusion: - By carrying out the partial test for set of regressors to check their significance, we obtained F value as 1.7638 and p value obtained is 0.1687 at 5% level of significance. On comparing the p value of the test conducted and α=0.05, p value is greater than α and hence we reject the null hypothesis being considered and the two models do not differ significantly. On removing cost of living index, property price to income ratio, traffic commute time index, pollution index and climate index, the full model and the model omitting these are not different. Hence, we can omit out the said regressors which do not contribute significantly to the model.

5. MULTIPLE LINEAR REGRESSION OF REDUCED MODEL WHEN RANK IS A RESPONSE VARIABLE

Null Hypothesis: -
H_0: $\beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$
(There is no relation between X & Y)

Alternative Hypothesis: -
H_1: β_j! = 0 for all at least one j = 1,2,3,4
(There exists a relation between X and Y)

TABLE 1.5: - ANOVA Table for testing significance of hypothesis

	df	SS	MS	F	Significance F	
Regressio	4	13668.81	3417.202	234.7255	7.30E-19	
Residual	24	349.399	14.55829			
Total	28	14018.21				

From the table given the calculated F statistic (F_o) is 234.7255 for 4 and 24 degrees of freedom for regression and residuals. We are checking the significance of the model at 5% level of significance.
The tabulated value of F at 0.05 with corresponding degrees of freedom is 2.7763. On comparing the two values we found that
$F_o > F_{0.05,4,24}$. Hence, we rejected our null hypothesis.

Conclusion: - We rejected our Ho and we are able to conclude that there exists some relation between the leftover explanatory variables in the reduced model and response variable for multiple linear regression model.
We can say that Rank allotted to different countries in Asia region of the world does depends on leftover explanatory variables taken which are **Quality of life index, Purchasing power index, Safety index, Health care index.**

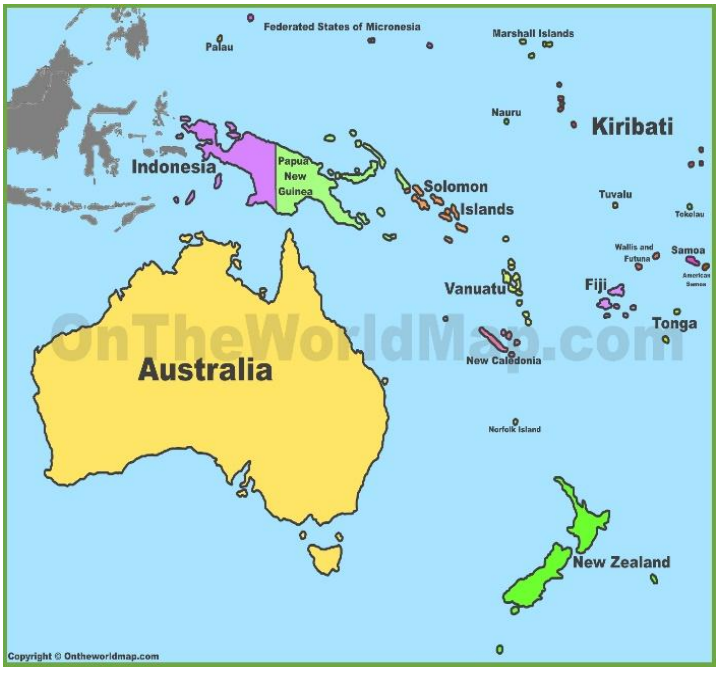
R^2 value for this model is 0.97508 means 97.508% of the variation explained by the regressors x.

6. CORRELATION BETWEEN VARIABLES IN REDUCED MODEL

	RANK	*QOLI*	*PURCHASING POWER INDEX*	*SAFETY INDEX*	*HEALTH CARE INDEX*
RANK	1				
QOLI	-0.98721	1			
PURCHASING POWER INDEX	-0.84402	0.846063	1		
SAFETY INDEX	-0.73568	0.751444	0.723104126	1	
HEALTH CARE INDEX	-0.37688	0.376778	0.401745336	0.420913256	1

Through above table we can see that there is negative correlation between rank and other variables while there is positive correlation between other variables.

4. OCEANIA AND AFRICA



Here the number of countries in Africa region were just 6 in number and number of countries in Oceania region were just 2 which were less than the number of parameters engulfed into multiple linear regression model. So, we had combined the countries for these two regions. After combining the countries, again the number of countries which are observations in our model are less than the number of parameters taken to check the relation between quality-of-life index and other such indices presented in the data. Consequently, we built up two multiple linear regression model and a simple linear regression respectively each for: -

1. When response variable was rank allotted to the countries.
2. When response variable was quality of life index of different countries.
3. When response variable was rank and explanatory variable is quality of life index.

1.1. <u>Multiple Linear Regression when Rank is Response variable: -</u>

Here the number of regressors are 3 which are Purchasing power index, Safety index and Health care index.

Null Hypothesis: -
H_0: $\beta_1 = \beta_2 = \beta_3 = 0$
(There is no relation between X & Y)

Alternative Hypothesis: -
H_1: β_j! = 0 for all at least one j = 1,2,3
(There exists a relation between X and Y)

Table 4.1.1: - ANOVA Table for testing significance of hypothesis

Source of variation	*df*	*SS*	*MS*	*F*	*Significance F*
Regression	3	6647.094	2215.698	25.4381	0.004572935
Residual	4	348.4061	87.10153		
Total	7	6995.5			

From the table given the calculated F statistic (F_o) is 25.43810487 for 3 and 4 degrees of freedom for regression and residuals respectively for the model. We are checking the significance of the model at 5% level of significance.

The tabulated value of F at 0.05 with corresponding degrees of freedom is 6.5914. On comparing the two values we found that
$F_o > F_{0.05,3,4.}$ Hence, we rejected our null hypothesis.

Conclusion: - We rejected our Ho and we are able to conclude that there exists some relation between the explanatory variables and response variable for multiple linear regression model.
We can say that Rank allotted to different countries in Africa and Oceania region of the world does depends on different explanatory variables taken which are **Purchasing power index, Safety index, Health care index.**

1.2. Multiple Linear Regression when Rank is Response variable: -

Here the number of regressors are 5 which are Cost of living index, Property price to income ratio, Traffic commute time index, Pollution index, Climate index.

Null Hypothesis: -
H_0: $\beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = 0$
(There is no relation between X & Y)

Alternative Hypothesis: -
H_1: $\beta_j ! = 0$ for all at least one j = 1, 2 ,…..,5
(There exists a relation between X and Y)

Table 4.1.2: - ANOVA Table for testing significance of hypothesis

Source of variation	*df*	*SS*	*MS*	*F*	*Significance F*
Regression	5	6955.679	1391.136	69.86868	0.01417039
Residual	2	39.82144	19.91072		
Total	7	6995.5			

From the table given the calculated F statistic (F_o) is 69.86868486 for 5 and 2 degrees of freedom for regression and residuals. We are checking the significance of the model at 5% level of significance.

The tabulated value of F at 0.05 with corresponding degrees of freedom is 19.2964. On comparing the two values we found that
$F_o > F_{0.05,5,2}$. Hence, we rejected our null hypothesis.

Conclusion: - We rejected our Ho and we are able to conclude that there exists some relation between the explanatory variables and response variable for multiple linear regression model.
We can say that Quality of life index allotted to different countries in Africa and Oceania region of the world does depends on different explanatory variables taken which are **Cost of living index, Property price to income ratio, Traffic commute time index, Pollution index, Climate index.**

2.1 ***Multiple Linear Regression when Quality of Life is Response variable: -***

Here the number of regressors are 3 which are Purchasing power index, Safety index and Health care index.

Null Hypothesis: -
H_0: $\beta_1 = \beta_2 = \beta_3 = 0$
(There is no relation between X & Y)

Alternative Hypothesis: -
H_1: $\beta_j \neq 0$ for all at least one j = 1,2,3
(There exists a relation between X and Y)

Table 4.2.1 : - ANOVA Table for testing significance of hypothesis

Source of variation	*df*	*SS*	*MS*	*F*	*Significance F*
Regression	3	13541.32	4513.773	42.9529	0.001682425
Residual	4	420.3463	105.0866		
Total	7	13961.66			

From the table given the calculated F statistic (F_o) is 42.95289535 for 3 and 4 degrees of freedom for regression and residuals respectively for the model. We are checking the significance of the model at 5% level of significance.

The tabulated value of F at 0.05 with corresponding degrees of freedom is 6.5914. On comparing the two values we found that
$F_o > F_{0.05,3,4}$. Hence, we rejected our null hypothesis.

Conclusion: - We rejected our Ho and we are able to conclude that there exists some relation between the explanatory variables and response variable for multiple linear regression model.
We can say that Quality of life index calculated for different countries in Africa and Oceania region of the world does depends on different explanatory variables taken which are **Purchasing power index, Safety index, Health care index.**

2.2 Multiple Linear Regression when Quality of Life is Response variable:-

Here the number of regressors are 5 which are Cost of living index, Property price to income ratio, Traffic commute time index, Pollution index, Climate index .

Null Hypothesis: -
H_0: $\beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = 0$
(There is no relation between X & Y)

Alternative Hypothesis: -
$H_{1:}$ $\beta_j ! = 0$ for all at least one j = 1,2,..,5
(There exists a relation between X and Y)

Table 4.2.2 : - ANOVA Table for testing significance of hypothesis

Source of variation	*df*	*SS*	*MS*	*F*	*Significance F*
Regression	5	13893.6	2778.719	81.64648	0.012143684
Residual	2	68.06709	34.03355		
Total	7	13961.66			

From the table given the calculated F statistic (F_o) is 81.64648027 for 5 and 2 degrees of freedom for regression and residuals. We are checking the significance of the model at 5% level of significance.

The tabulated value of F at 0.05 with corresponding degrees of freedom is 19.2964. On comparing the two values we found that
$F_o > F_{0.05,5,2}$. Hence, we rejected our null hypothesis.

Conclusion: - We rejected our Ho and we are able to conclude that there exists some relation between the explanatory variables and response variable for multiple linear regression model.
We can say that Quality of life index allotted to different countries in Africa and Oceania region of the world does depends on different explanatory variables taken which are **Cost of living index, Property price to income ratio, Traffic commute time index, Pollution index, Climate index.**

3. ***SIMPLE LINEAR REGRESSION BETWEEN QUALITY-OF-LIFE INDEX AND RANK ALLOTED TO THE COUNTRIES***

This shows linear negative relationship between rank and quality of life index as quality of life index increases rank decreases.

Null Hypothesis: -
H_0: $\beta_1 = 0$ (There is no relation between X & Y)

Alternative Hypothesis: -
H_1: $\beta_1 ! = 0$ (There exists a relation between X and Y)

Table 4.3: - ANOVA Table for testing significance of hypothesis

Source of variation	*df*	*SS*	*MS*	*F*	*Significance F*
Regression	1	6737.164	6737.164	156.4744	1.59609E-05
Residual	6	258.3361	43.05602		
Total	7	6995.5			

From the table given the calculated F statistic (F_o) is 156.4743734 for 1 and 6 degrees of freedom for regression and residuals. We are checking the significance of the model at 5% level of significance.
The tabulated value of F at 0.05 with corresponding degrees of freedom is 5.9874. On comparing the two values we found that
$F_o > F_{0.05,1,6}$. Hence, we rejected our null hypothesis.

Conclusion: - We rejected our Ho and we are able to conclude that there exists some relation between the explanatory variables and response variable for multiple linear regression model.
We can say that Rank allotted to different countries in Africa and Oceania region of the world does depends on **Quality-of-life index.**

RESULTS AND CONCLUSION

- After analyzing data for 2019,2021 and 2022 and obtaining the predicted model for India, the country with the highest rank and the country with the lowest rank for all three years we can say that the predicted model is the same as the original model in all the cases hence data is significant for all three years.
- In this study, we tried to find out what is the relationship between rank allotted to different countries while calculating the quality-of-life index. In this we have constructed simple linear regression models and multiple linear regression models using rank and quality of life index maintained by the countries in year 2022.

- We, after carrying out the analysis of the study conducted are able to conclude that rank and quality of life index depends on various indices like safety index, health index, cost of living index etc. Using MS Excel and R programming, we did the analysis part by constructing Analysis of variance table (ANOVA) for various models and we found the f statistic and then compared it with the tabulated values of f-distribution at 5% level of significance. We tested the significance of all the regressors present in the total model and then tested the significance of few regressors w.r.t to response variable which was RANK and QUALITY OF LIFE in our case and found that if some of the unimportant regressors are omitted out, then also reduced model and total model have similar significance to the response variable taken. By this, we are able to find evidence that quality of life

index measured does not depend on cost-of-living index, property price to income ratio, traffic commute time index, pollution index and climate index much and while seeing the relation among rank and other independent variables, these can be omitted out. If not omitted, these unimportant regressors would significantly increase the error part or residuals part present in the model which is not at all desirable.

- Also, we had tried to present the analysis in a graphical form including scatter plots showing regression lines between quality-of-life index and rank which tells us the negative relation between the two. Concluding it, we can say that as quality-of-life index improves, the countries had a lower rank at the global level which is a good socio-economic indicator. Also, we had shown that data follows normal distribution to a little extent by representing QQ plots. Adding to this, we have calculated correlation between rank and quality of life index and between some of the other regressors which again validates the point of negative relation between rank and quality of life index. Furthermore, we had shown the R^2 for the models being analyzed which is coefficient of determination which is the ratio of measure of model fit and tells us how the regressors and regressed variable are fit well and what is their extent of relationship.

- At last, we would like to comment that the countries located in Asia, America, Oceania, Africa and Europe participated and gave their valuable data to the higher functionaries while calculating the index. Some the indices which are independent regressors in this model contributed

significantly to the model while some other do not and we can neglect them. In each region of the world, we were able to find that rank and quality of life index relates and depends on various parameters or indices. Also, every region shows a negative relationship between rank and quality of life index which validates that as quality-of-life increases, rank decreases drastically. Using various tools of data analysis, we carried out this study and we hope that this study finds its importance to the world leaders.

- We haven't applied the partial F test in case of multiple regression when Y is the Quality of Life Index as all the X variables were having P value less than 0.05 hence they are contributing significantly to the model.

FUTURE SCOPE

- Quality of Life index is a growing field of research. Many disciplines have already embedded QOL within their research domain which can lead many countries and citizens to imbibe those disciplines and maintain standards of living.

- Quality of life also depends on political regime of the country in the sense that are citizens able to practice their rights in a fundamental way or not. As many countries are under

autocratic regime, citizens living in those areas might be compromising their personal choices and needs which is disastrous for countries to propel and to score better in quality-of-life index.

- Quality of Life index studies should be promoted through additional outlets of research. Besides, international scholarly publications, efforts should be made to present study results in local media or on the Web to familiarize people about basic needs and choices.

- Efforts should be made to ensure that local government officials are made aware of this index findings, so as to inform the multitude of planning and policy decisions that need to be made in their respective jurisdictions to perform better next time and compete with top 10 countries of the world.

- Many developing countries who performed worst in the index released and are under the shackles of poverty and hunger can be brought into the spotlight where well off or developed nations can lend a helping hand to them as a token of gratitude to eliminate the social evils at a global level.

- This index can give the social scientists a better perspective to comment on the countries who performed moderately or worst and can suggest the leaders of that particular country on better governance and participation of the people in the political institutions.

- Quality of life can be also understood as combining certain parameters of 17 sustainable development goals founded by united nations and it can help achieve the agenda 2030 by calculating the position of different countries on this scale globally.

REFERENCES

- https://datatofish.com/multiple-linear-regression-in-r/
- https://www.youtube.com/watch?v=G_obrpV70QQ
- https://www.numbeo.com/quality-of-life/rankings_by_country.jsp
- https://pdf.sciencedirectassets.com/
- https://iovs.arvojournals.org/article.aspx?articleid=2128582