VIDYASAGAR UNIVERSITY

Constructing an Index to Compare Rural Health System Condition Between Different States of India

Registration No: - 1160457 of 2020-2021

Roll No: - 1126116-200134

Session: - 2022-2023

Acknowledgement:

I would like to take this opportunity to express my heartfelt gratitude to all those who have contributed towards the successful completion of my project work.

First and foremost, I would like to express my sincere appreciation to my teachers, Dr. Shyamsundar Sahoo, Mr. Tanmay Maity, Mr. Sibsankar Karan, and Mr. Bijitesh Halder, for their unmatched support, guidance, and constant motivation throughout the project. Their expertise, valuable insights, and constructive feedback have been instrumental in shaping my ideas and developing my skills.

I would also like to extend my sincere thanks to my friends, particularly Nilanjan Samanta, for their unflinching support and encouragement during the project. Their feedback and suggestions have been immensely helpful in overcoming challenges and staying focused on my goals.

I would also like to express my appreciation to the university for including project work in the curriculum. This project has provided me with a unique opportunity to apply the knowledge and skills I have acquired during my course of study in a practical setting, and has been an enriching experience.

Finally, I would like to thank everyone who has contributed to this project in one way or another. Your support and encouragement have been invaluable, and I am grateful for your contribution to this achievement.

Contents:

1. Introduction..

2. Objective..

3. Data Collection…

4. Data Description..

5. Methodology..

6. Result And Analysis..

7. Conclusion…

8. References..

9. Appendix…

Introduction:

India's rural health system is facing significant challenges that are affecting the health and well-being of millions of people living in rural areas. Despite some progress in recent years, there are still significant disparities in healthcare access and quality between urban and rural areas. This is due to a lack of adequate infrastructure, shortage of healthcare professionals, and limited access to essential medicines.

Real-life examples illustrate the challenges faced by the rural health system in India. For instance, many people living in rural areas do not have access to basic healthcare services, and maternal and infant mortality rates are significantly higher than in urban areas. The prevalence of chronic diseases such as diabetes and hypertension is also rising in rural areas, but access to specialist care and medication remains limited.

To better understand the condition of the rural health system in India, we have constructed an index that measures key indicators such as healthcare infrastructure, availability of essential stuffs, and quality of healthcare services in a particular state. This index is unique because there is no such well-known index previously constructed, and we hope that it will be helpful in visualizing the condition of rural health system in India. By identifying areas of improvement, policymakers can develop targeted interventions to improve rural health outcomes and ensure that all people have access to quality healthcare services, regardless of where they live

Objectives:

1. To assess the current state of rural health systems in India using our constructed index.

2. To identify areas of improvement in rural health systems and develop recommendations for policymakers.

3. To raise awareness about the challenges faced by the rural health system in India and the need for targeted interventions to improve health outcomes.

4. To contribute to the development of evidence-based policies and strategies to improve rural health outcomes in India.

5. To compare rural health system condition between different state and UT's of India

6. To compare it with a existing reputed index

Data Collection:

The entire data is collected from Annual Rural Health Statistics report (2020-21) published by Ministry of Health and Family Welfare.

Data Description:

In India, the health care system is organized into a hierarchical structure, with different levels of health facilities providing care to the population.

1.   **Sub-Centre:** The Sub-Centre is the first point of contact between the primary health care system and the community. It is typically staffed by an Auxiliary Nurse Midwife (ANM) and a Male Health Worker (MHW) and serves a population of around 5,000 in rural areas and 3,000 in hilly, tribal and difficult areas. The Sub-Centre provides basic health care services such as maternal and child health care, immunization, family planning, and health education.

2.   **Primary Health Centre:** The Primary Health Centre (PHC) is a higher level of health care facility that provides comprehensive health care services to a population of around 30,000 in rural areas and 20,000 in hilly, tribal and difficult areas. The PHC is staffed by a Medical Officer, a Pharmacist, a Staff Nurse, and other paramedical staff. The PHC provides curative, preventive, and promotive health care services, including outpatient care, maternal and child health care, family planning, immunization, and laboratory services.

3.   **Community Health Centre:** The Community Health Centre (CHC) is a higher level of health care facility that provides specialist care to a population of around 80,000 in rural areas and 60,000 in hilly, tribal and difficult areas. The CHC is staffed by a Medical Officer, specialists such as surgeons, obstetricians, and pediatricians, and other paramedical staff. The CHC provides inpatient and outpatient care, emergency obstetric care, and specialist care in surgery, obstetrics, and pediatrics.

4.   **ANM:** The Auxiliary Nurse Midwife (ANM) is a female health worker who is the key link between the community and the health care system. ANMs are usually posted at the Sub-Centre and are responsible for providing basic health care services such as maternal and child health care, family planning, and immunization. ANMs are also responsible for health education and creating awareness about health issues in the community.

Factors to work with:

1. Average Population Covered by a Sub Centre

2. Average Population Covered by a Primary Health Centre

3. Average Population Covered by a Community Health Centre

4. Average Rural Area Covered by a Sub Centre

5. Average Rural Area Covered by a Primary Health Centre

6. Average Rural Area Covered by a Community Health Centre

7. Average Radial Distance Covered by a Sub Centre

8. Average Radial Distance Covered by a Primary Health Centre

9. Average Radial Distance Covered by a Community Health Centre

10. Average Number of Villages Covered by a Sub Centre

11. Average Number of Villages Covered by a Primary Health Centre

12. Average Number of Villages Covered by a Community Health Centre

13. Average Rural Population Covered by a Health Worker

14. Health Workers (Female) / ANM at Rural Areas (In Position/ Required Ratio)

15. Doctors at Primary Health Centres in Rural Areas (In Position/Required Ratio)

16. Total Specialists at CHCs in Rural Area (In Position/Required Ratio)

17. Radiographers at CHCs in Rural Area (In Position/Required Ratio)

18. Pharmacists at PHCs and CHCs in Rural Area (In Position/Required Ratio)

19. Laboratory Technicians at PHCs and CHCs in Rural Area (In Position/Required Ratio)

20. Nursing Staff at PHCs and CHCs in Rural Areas (In Position/Required Ratio)

21. Statewise Rural Infant Mortality Ratio

Methodology:

First of all, we are dividing the states in two categories, 1. Larger States and 2. Smaller States according to their size and population, we couldn't consider all the states due to lack of information and eliminated the UT's completely

Larger States:

Andhra Pradesh, Assam, Bihar, Chattisgarh, Gujarat, Haryana, Jharkhand, Karnataka, Kerala, Madhya Pradesh, Maharashtra, Odisha, Punjab, Rajasthan, Tamil Nadu, Telangana, Uttar Pradesh, Uttarakhand and West Bengal

Smaller States:

Goa, Manipur, Meghalaya, Mizoram, Nagaland

We will perform the calculations for these two categories sepeartely.

Now among those factors we are considering, which ones are actually significant for the index, we have to eliminate those which are not significant. For that we have performed Backward Step Regression using R taking Infant Mortality Rate as the response variable and eliminate the insignificant factors.

Backward step regression is a feature selection technique used in linear regression analysis to identify the most significant independent variables that have a strong correlation with the dependent variable and exclude the less significant variables. The process starts with all independent variables included in the model, and the variable with the highest p-value is removed. The model is then refitted without the removed variable, and the process is repeated until all remaining variables have a p-value below a certain threshold or until a specified number of variables remain in the model. The advantage of backward step regression is that it can help simplify the model by removing unnecessary variables that do not contribute significantly to the model. However, this technique assumes that all variables are independent of each other, and it may not work well if there are strong interactions between variables.

We will use the t-value of the factors we got from backward step regression as their respective weights to form the integer.

The t-value for each variable in backward step regression is calculated as the ratio of the estimated coefficient to its standard error. The formula for the t-value of a variable in the model is:

t-value = (estimated coefficient / standard error)

Then we have to normalize the significant factors and bring them within same scale

There are two types of factors:

Positive Factors: Higher the value, better the result

Ex: Doctors at Primary Health Centres in Rural Areas (In Position/Required Ratio)

Negative Factors: Lower the value, better the result

Ex: Average number of villages covered by a sub-centre

$$\text{Scaled value } (S_i) \text{ for positive indicator} = \frac{(X_i - \text{Minimum value}) \times 100}{(\text{Maximum value} - \text{Minimum value})}$$

$$\text{Scaled value } (S_i) \text{ for negative indicator} = \frac{(\text{Maximum value} - X_i) \times 100}{(\text{Maximum value} - \text{Minimum value})}$$

Now we just have to calculate the composite indicator using the scaled values of factor and our index is ready

$$\text{Composite Index} = \frac{\Sigma\, W_i * S_i}{\Sigma\, W_i}$$

Result & Analysis:

By performing backward step regression taking Infant Mortality Rate as response variable, we are eliminating the following factors:

1. Average Population Covered by a Primary Health Centre
2. Average Radial Distance Covered by a Sub Centre
3. Average Number of Villages Covered by a Primary Health Centre
4. Health Workers (Female) / ANM at Rural Areas (In Position/ Required Ratio)
5. Laboratory Technicians at PHCs and CHCs in Rural Area (In Position/Required Ratio

After that, following the process mentioned in methodology, we calculate the index values

For the larger states, the index value table is given below

State	index value
Kerala	85.32731
Tamil nadu	73.73068
Assam	65.20467
Telangana	64.49315
Gujrat	61.83552
Andhra Pradesh	59.97826
karnataka	58.63135
Punjab	57.87452
Rajasthan	55.71368
Haryana	54.49678
West Bengal	54.47929
Chattisgarh	51.61902
Uttar Pradesh	49.40358
Odisha	46.55075
Maharashtra	45.94177
Bihar	41.17796
Madhya Pradesh	39.58766
Jharkhand	38.85999
Uttarakhand	37.47354

Rural Health Index (Larger States)
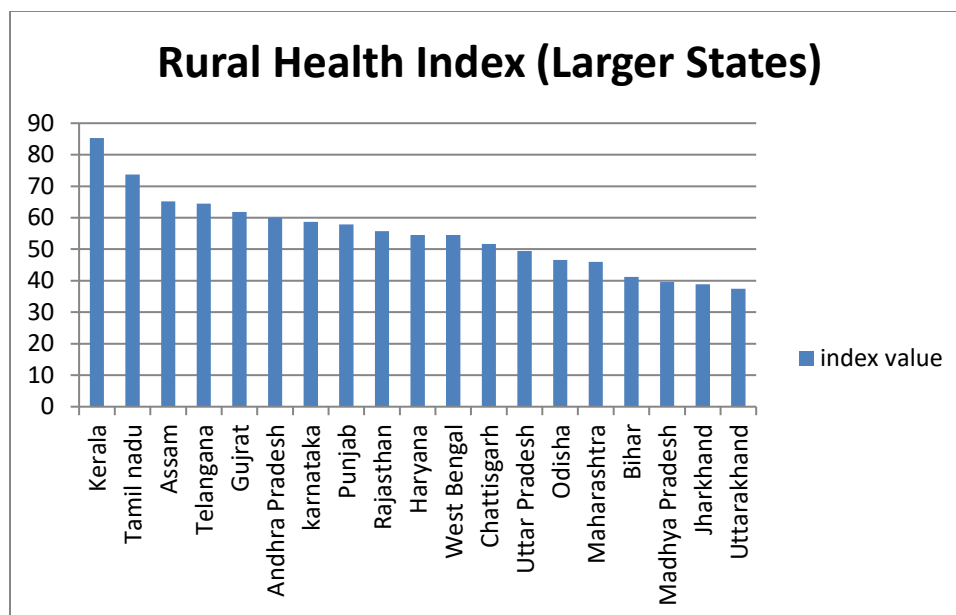
For the smaller states, the index value table is given below

State	index value
Goa	69.41978
Manipur	27.29415
Meghalaya	34.87745
Nagaland	68.25072
Sikkim	42.51475

Rural Health Index(Smaller States)
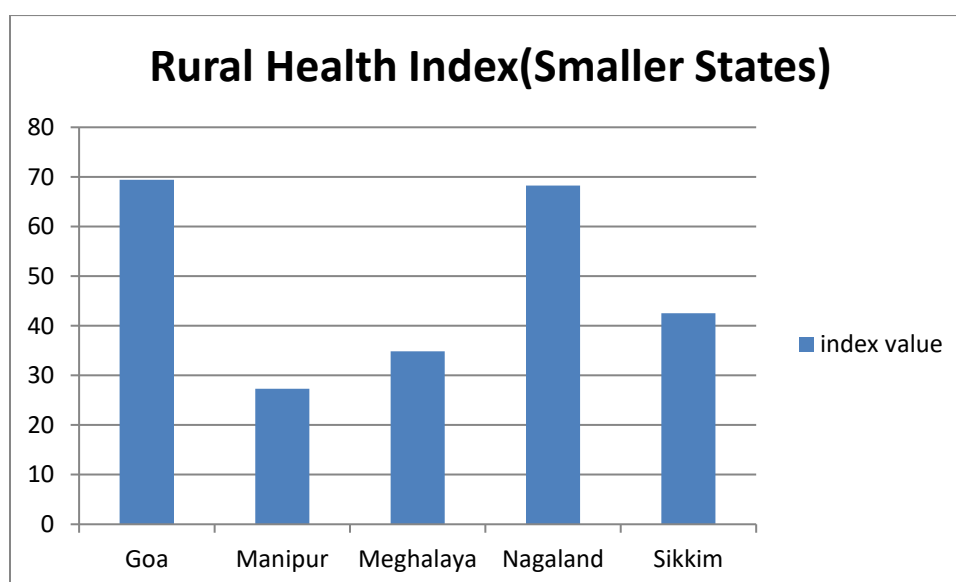
Comparison with NITI AAYOG Health Index (2019-20)

Similarities and Differences in Methodology:

1. Like NITI AAYOG Health Index, we have divided the states in two categories (larger states and smaller states) here also

2. The positive factors and negative factors are normalized in similar way

3. In NITI AAYOG Health Index, the factors are categorized as key factors, intermediate factors, minor factors etc. and a particular weight is assigned to each key factors. Similarly, a particular weight is assigned to each intermediate and minor factors . Here we ran a backward regression taking infant mortality rate as the response variable and eliminated few factors. For the remaining factors, we have taken their t-value as their respective weight.

Creating an index using t-values as weights after performing backward step regression can have several potential advantages over using fixed weights for factors. Here are some reasons why our index might be considered better:

1. **Statistical significance:** The t-values obtained from backward step regression indicate the statistical significance of each factor included in the index. By assigning weights proportional to these t-values, we are giving more importance to factors that are statistically more significant in explaining the variation in the data. This can result in a more robust and reliable index.

2. **Variable importance:** Backward step regression helps identify the most important factors for predicting the outcome variable. By assigning higher weights to factors with larger t-values, we are effectively prioritizing the most influential variables in the index construction. This can lead to a more accurate representation of the underlying factors driving the phenomenon we are measuring.

3. **Adaptability:** Unlike fixed weights, which remain constant regardless of the data, using t-values as weights allows our index to adapt to changes in the significance and importance of each factor. If the relative importance of certain factors changes over time

or due to new data, our index will automatically reflect these changes by adjusting the weights accordingly. This adaptability can enhance the relevance and responsiveness of the index.

4. **Overfitting avoidance:** Backward step regression helps mitigate the risk of overfitting, which occurs when a model is excessively complex and tailored to the specific characteristics of the training data, leading to poor generalization to new data. By selecting factors based on their statistical significance and excluding less relevant variables, we reduce the likelihood of overfitting. Consequently, our index is more likely to perform well on unseen data and provide more accurate predictions.

5. **Transparency and reproducibility**: By using t-values as weights, we make the index construction process more transparent and reproducible. The t-values are derived from statistical tests, which are generally well-documented and widely understood. This allows others to replicate our index or understand the rationale behind its construction, promoting transparency and facilitating comparisons with other indices or models.

NITI AAYOG Index values for larger states:

State	Index Value
Kerala	82.2
Tamil nadu	72.42
Telangana	69.96
Andhra Pradesh	69.95
Maharashtra	69.14
Gujarat	63.59
Punjab	58.08
karnataka	57.93
Chattisgarh	50.7
Haryana	49.26
Assam	47.74
Jharkhand	47.55
Odisha	44.31
Uttarakhand	44.21
Rajasthan	41.33
Madhya Pradesh	36.72
Bihar	31
Uttar Pradesh	30.57

Comparison with our index:

1. Kerala in top position and Tamil Nadu in 2nd position according to both NITI AAYOG & our Rural Health Index

2. There are 3 states common in top 5, 8 states common in top 10, 3 states common in below 5

3. Assam holds a good position according to our index but not according the other one

4. Maharasthra does not hold a good position in our index despite holding a good position in that NITI AAYOG index

NITI AAYOG Index Value for Smaller States:

State	Index Value
Goa	55.53
Sikkim	53.38
Meghalaya	43.05
Manipur	34.26
Nagaland	27

Comparison with our index:

1. Goa tops in both cases

2. There are 2 common states in top 3 and 2 common in below 3

3. Nagaland holding good position in our index but that's not the case for NITI AAYOG Index

Conclusion:

In this study, a rural health index was constructed to evaluate the health infrastructure and outcomes in Indian states. Despite acknowledging the scope for improvements in our index, it is notable that the results obtained align closely with the existing NITI Aayog health index, which primarily focuses on overall health performance without specifically emphasizing rural areas. This similarity suggests that the constructed rural health index is moving in the right direction.

The fact that our rural health index and the NITI Aayog health index demonstrate similar findings indicates that the existing index effectively captures the overall health performance of states, encompassing both rural and urban areas. However, the development of a dedicated rural health index is crucial to shed light on the specific challenges and disparities faced by rural populations in accessing healthcare.

While there are opportunities for improvement in our index, such as refining indicators, data sources, and methodology, the alignment with the NITI Aayog health index signifies that our efforts have been in the correct direction. By focusing on rural areas, our index provides a more nuanced understanding of the state of rural healthcare in India, emphasizing the need for targeted interventions and resource allocation.

The findings of this study hold significant implications for policymakers, healthcare organizations, and other stakeholders. The constructed rural health index can serve as a valuable tool in identifying areas for improvement and guiding the development of tailored interventions to address the unique healthcare needs of rural communities. By leveraging the insights provided by the rural health index, policymakers can work towards bridging gaps in healthcare infrastructure, enhancing primary healthcare services, and reducing health disparities in rural India.

Future research should aim to further refine the index by incorporating additional relevant indicators, improving data quality and availability, and engaging in ongoing collaboration with experts and stakeholders. Longitudinal studies can also be conducted to monitor the progress of states over time and assess the impact of interventions aimed at improving rural healthcare.

In conclusion, the construction of a rural health index represents a vital step towards understanding and addressing the healthcare challenges faced by rural populations in India. While improvements are possible, the similarity between our index and the NITI Aayog health index validates our approach and underscores the importance of prioritizing rural healthcare to build a more inclusive and equitable healthcare system for all.

References:

1. Rural Health Statistics (2020-21): https://main.mohfw.gov.in/newshighlights-90

2. NITI AAYOG Health Index (2019-20) : http://social.niti.gov.in/hlt-ranking/?round=4

3. NITI AAYOG Health Index Methodology: https://social.niti.gov.in/health-index

4. Fundamental of Statistics,Vol-1, Vol-2,A.M Gun , M.K Gupta , B. Dasgupta

Appendix:

```r
library(ggplot2)
library(moments)
library(cowplot)
library(LambertW)
library(tidyverse)
library(ggpubr)
library(lawstat)
library(devtools)

df <- read.csv("C:\\Users\\hp\\Desktop\\Project\\pratyay\\final.csv")

# str(df)
# df<- df[,-1]

f_model <- lm(F21~., data = df)
#summary(model)
nu_model <- lm(F21~0,data=df)

stepAIC(f_model,direction = "backward",scope =list(upper=f_model,lower =
nu_model))

## Start:  AIC=104.74
## F21 ~ F1 + F2 + F3 + F4 + F5 + F6 + F7 + F8 + F9 + F10 + F11 +
##      F12 + F13 + F14 + F15 + F16 + F17 + F18 + F19 + F20
##
##          Df Sum of Sq     RSS     AIC
## - F11    1       0.75  308.31  102.81
## - F14    1       1.72  309.28  102.88
## - F3     1       2.94  310.50  102.98
## - F9     1       4.22  311.78  103.09
## - F4     1       9.70  317.27  103.52
## - F7     1      10.78  318.34  103.61
## - F5     1      18.15  325.71  104.18
## - F19    1      18.35  325.91  104.19
## - F6     1      24.70  332.26  104.68
## <none>              307.56  104.75
## - F12    1      50.13  357.69  106.52
## - F10    1     106.69  414.25  110.19
## - F1     1     129.14  436.70  111.51
## - F18    1     131.48  439.04  111.64
## - F16    1     147.06  454.62  112.52
## - F17    1     163.51  471.07  113.40
## - F15    1     198.77  506.33  115.21
## - F13    1     252.81  560.37  117.74
## - F20    1     285.31  592.87  119.15
## - F8     1     313.08  620.64  120.30
## - F2     1     367.03  674.59  122.38
##
## Step:  AIC=102.81
```

```
## F21 ~ F1 + F2 + F3 + F4 + F5 + F6 + F7 + F8 + F9 + F10 + F12 +
##      F13 + F14 + F15 + F16 + F17 + F18 + F19 + F20
##
##          Df Sum of Sq    RSS    AIC
## - F14    1      1.02 309.33 100.89
## - F3     1      3.06 311.37 101.05
## - F9     1      6.39 314.70 101.32
## - F7     1     10.05 318.37 101.61
## - F4     1     16.66 324.98 102.12
## - F19    1     17.61 325.92 102.19
## <none>             308.31 102.81
## - F5     1     32.13 340.44 103.28
## - F6     1     34.12 342.43 103.43
## - F12    1     76.88 385.19 106.37
## - F1     1    135.41 443.72 109.91
## - F18    1    164.04 472.35 111.47
## - F16    1    175.22 483.53 112.06
## - F17    1    200.31 508.62 113.32
## - F15    1    224.77 533.08 114.50
## - F10    1    226.96 535.27 114.60
## - F13    1    253.50 561.81 115.81
## - F8     1    318.83 627.15 118.56
## - F20    1    326.19 634.50 118.85
## - F2     1    543.62 851.93 126.22
##
## Step:  AIC=100.89
## F21 ~ F1 + F2 + F3 + F4 + F5 + F6 + F7 + F8 + F9 + F10 + F12 +
##      F13 + F15 + F16 + F17 + F18 + F19 + F20
##
##          Df Sum of Sq    RSS     AIC
## - F3     1      2.34  311.66  99.076
## - F7     1     10.85  320.18  99.750
## - F19    1     16.61  325.93 100.195
## - F9     1     16.70  326.03 100.203
## <none>              309.33 100.888
## - F4     1     74.50  383.83 104.283
## - F12    1     77.96  387.29 104.507
## - F6     1    107.69  417.02 106.356
## - F5     1    129.30  438.63 107.620
## - F1     1    134.43  443.76 107.910
## - F18    1    166.66  475.99 109.663
## - F10    1    242.31  551.63 113.350
## - F13    1    283.32  592.64 115.143
## - F16    1    321.98  631.30 116.723
## - F17    1    366.45  675.78 118.425
## - F8     1    391.19  700.52 119.324
## - F15    1    658.97  968.30 127.417
## - F2     1    725.20 1034.53 129.071
## - F20    1    769.81 1079.14 130.126
##
```

```
## Step:  AIC=99.08
## F21 ~ F1 + F2 + F4 + F5 + F6 + F7 + F8 + F9 + F10 + F12 + F13 +
##     F15 + F16 + F17 + F18 + F19 + F20
##
##          Df Sum of Sq      RSS      AIC
## - F19    1      14.29   325.95   98.197
## - F7     1      14.45   326.11   98.209
## <none>                  311.66   99.076
## - F12    1      77.24   388.90  102.611
## - F9     1     116.06   427.73  104.990
## - F5     1     126.97   438.63  105.620
## - F18    1     165.14   476.80  107.706
## - F4     1     174.17   485.83  108.175
## - F10    1     252.06   563.72  111.892
## - F6     1     257.16   568.83  112.118
## - F13    1     283.73   595.40  113.259
## - F16    1     324.58   636.24  114.917
## - F17    1     364.36   676.02  116.434
## - F8     1     470.11   781.78  120.067
## - F1     1     573.71   885.37  123.178
## - F15    1     672.16   983.83  125.814
## - F2     1     723.61  1035.28  127.089
## - F20    1     788.06  1099.72  128.598
##
## Step:  AIC=98.2
## F21 ~ F1 + F2 + F4 + F5 + F6 + F7 + F8 + F9 + F10 + F12 + F13 +
##     F15 + F16 + F17 + F18 + F20
##
##          Df Sum of Sq      RSS      AIC
## - F7     1      14.41   340.36   97.278
## <none>                  325.95   98.197
## - F12    1      64.25   390.20  100.695
## - F9     1     157.39   483.34  106.046
## - F4     1     170.36   496.31  106.708
## - F5     1     181.61   507.56  107.268
## - F10    1     239.08   565.03  109.950
## - F13    1     285.69   611.65  111.932
## - F18    1     301.29   627.24  112.561
## - F6     1     330.10   656.05  113.684
## - F16    1     333.89   659.84  113.828
## - F8     1     455.84   781.79  118.068
## - F17    1     527.59   853.54  120.263
## - F1     1     565.14   891.09  121.339
## - F2     1     731.65  1057.60  125.622
## - F20    1     791.84  1117.79  127.006
## - F15    1     850.30  1176.25  128.280
##
## Step:  AIC=97.28
## F21 ~ F1 + F2 + F4 + F5 + F6 + F8 + F9 + F10 + F12 + F13 + F15 +
##     F16 + F17 + F18 + F20
```

```
##
##          Df Sum of Sq       RSS       AIC
## <none>                    340.36   97.278
## - F12   1      49.84    390.20   98.695
## - F5    1     167.62    507.99 105.289
## - F9    1     262.76    603.12 109.581
## - F10   1     269.44    609.80 109.856
## - F18   1     290.65    631.01 110.711
## - F13   1     313.38    653.74 111.596
## - F16   1     386.37    726.73 114.242
## - F6    1     400.23    740.59 114.714
## - F8    1     450.06    790.42 116.342
## - F17   1     530.39    870.75 118.762
## - F1    1     551.10    891.46 119.350
## - F4    1     553.85    894.22 119.427
## - F2    1     769.69   1110.05 124.832
## - F20   1     777.52   1117.88 125.008
## - F15   1     861.77   1202.14 126.825


##
## Call:
## lm(formula = F21 ~ F1 + F2 + F4 + F5 + F6 + F8 + F9 + F10 + F12 +
##      F13 + F15 + F16 + F17 + F18 + F20, data = df)
##
## Coefficients:
## (Intercept)            F1              F2              F4              F5
F6
##   -2.211e-15     1.083e-02    -1.618e-03    -2.313e+00     1.299e-01      9.745e-
02
##           F8            F9             F10             F12             F13
F15
##    1.241e+01    -1.002e+01     4.520e+00    -8.236e-02     4.742e-03        -
1.243e+02
##          F16            F17             F18             F20
##    5.408e+01    -3.103e+01    -3.405e+01     1.246e+02

# str(df1)
# d <- cor(df1)

summary(stepAIC(nu_model,direction = "forward",scope
=list(upper=f_model,lower = nu_model)),steps = 10000)

## Start:  AIC=127.01
## F21 ~ 0
##
##          Df Sum of Sq    RSS     AIC
## + F13   1    1698.14  2322.4  115.29
## + F15   1    1418.30  2602.3  118.13
## + F1    1    1377.60  2643.0  118.52
## + F2    1     707.83  3312.7  124.17
## + F20   1     619.90  3400.7  124.82
```

```
## + F3    1     525.99 3494.6 125.50
## + F6    1     436.69 3583.9 126.13
## + F11   1     391.80 3628.8 126.44
## + F19   1     391.29 3629.3 126.45
## + F10   1     347.31 3673.3 126.75
## <none>              4020.6 127.01
## + F9    1     277.91 3742.7 127.22
## + F14   1     251.51 3769.1 127.39
## + F4    1     250.37 3770.2 127.40
## + F12   1     183.31 3837.3 127.84
## + F18   1     171.98 3848.6 127.92
## + F17   1     170.42 3850.1 127.92
## + F7    1     127.43 3893.1 128.20
## + F5    1     119.37 3901.2 128.25
## + F8    1      31.84 3988.7 128.81
## + F16   1      19.37 4001.2 128.89
##
## Step:  AIC=115.29
## F21 ~ F13 - 1
##
##         Df Sum of Sq    RSS    AIC
## + F15   1    268.186 2054.2 114.22
## <none>              2322.4 115.29
## + F10   1    153.948 2168.5 115.57
## + F14   1    124.995 2197.4 115.90
## + F12   1     78.831 2243.6 116.42
## + F19   1     67.058 2255.4 116.56
## + F20   1     58.127 2264.3 116.65
## + F1    1     46.442 2276.0 116.78
## + F7    1     43.582 2278.8 116.81
## + F8    1     24.856 2297.6 117.02
## + F2    1     21.409 2301.0 117.06
## + F11   1     13.315 2309.1 117.14
## + F3    1     12.057 2310.4 117.16
## + F4    1     11.182 2311.2 117.17
## + F17   1      8.551 2313.9 117.19
## + F9    1      6.252 2316.2 117.22
## + F16   1      6.248 2316.2 117.22
## + F18   1      5.702 2316.7 117.23
## + F6    1      5.309 2317.1 117.23
## + F5    1      0.001 2322.4 117.29
##
## Step:  AIC=114.22
## F21 ~ F13 + F15 - 1
##
##         Df Sum of Sq    RSS    AIC
## + F20   1    247.673 1806.6 113.01
## + F14   1    237.377 1816.9 113.15
## + F10   1    172.105 1882.1 114.03
## <none>              2054.2 114.22
```

```
## + F19    1    153.432 1900.8 114.28
## + F7     1     89.910 1964.3 115.10
## + F12    1     51.235 2003.0 115.59
## + F8     1     40.395 2013.8 115.72
## + F1     1     36.009 2018.2 115.78
## + F4     1     35.437 2018.8 115.78
## + F5     1     33.820 2020.4 115.81
## + F16    1     27.351 2026.9 115.88
## + F11    1     26.613 2027.6 115.89
## + F2     1     15.779 2038.5 116.03
## + F9     1     10.018 2044.2 116.10
## + F6     1      3.203 2051.0 116.18
## + F3     1      2.022 2052.2 116.19
## + F17    1      1.393 2052.8 116.20
## + F18    1      0.885 2053.3 116.21
##
## Step:  AIC=113.01
## F21 ~ F13 + F15 + F20 - 1
##
##          Df Sum of Sq    RSS    AIC
## <none>                1806.6 113.01
## + F10    1    109.387 1697.2 113.45
## + F1     1    109.001 1697.6 113.45
## + F16    1     79.179 1727.4 113.89
## + F14    1     60.578 1746.0 114.16
## + F17    1     58.635 1747.9 114.18
## + F18    1     50.918 1755.7 114.29
## + F3     1     39.008 1767.6 114.46
## + F12    1     37.024 1769.5 114.49
## + F7     1     26.760 1779.8 114.64
## + F2     1     20.770 1785.8 114.72
## + F6     1     10.358 1796.2 114.86
## + F11    1      5.991 1800.6 114.92
## + F4     1      4.240 1802.3 114.95
## + F8     1      2.706 1803.9 114.97
## + F19    1      1.184 1805.4 114.99
## + F5     1      0.891 1805.7 115.00
## + F9     1      0.782 1805.8 115.00

##
## Call:
## lm(formula = F21 ~ F13 + F15 + F20 - 1, data = df)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -17.942  -6.430  -2.714   7.728  14.889
##
## Coefficients:
##       Estimate Std. Error t value Pr(>|t|)
## F13   0.005093   0.001636   3.113  0.00507 **
```

```
## F15 -39.322562   16.655273   -2.361   0.02750 *
## F20  26.994798   15.543767    1.737   0.09642 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 9.062 on 22 degrees of freedom
## Multiple R-squared:  0.5507, Adjusted R-squared:  0.4894
## F-statistic: 8.987 on 3 and 22 DF,  p-value: 0.000447
```

```
Start:   AIC=104.74
F21 ~ F1 + F2 + F3 + F4 + F5 + F6 + F7 + F8 + F9 +
F10 + F11 +
    F12 + F13 + F14 + F15 + F16 + F17 + F18 + F19 +
F20


        Df Sum of Sq      RSS      AIC
- F11    1       0.75 308.31 102.81
- F14    1       1.72 309.28 102.88
- F3     1       2.94 310.50 102.98
- F9     1       4.22 311.78 103.09
- F4     1       9.70 317.27 103.52
- F7     1      10.78 318.34 103.61
- F5     1      18.15 325.71 104.18
- F19    1      18.35 325.91 104.19
```

```
- F6      1      24.70 332.26 104.68
<none>                 307.56 104.75
- F12     1      50.13 357.69 106.52
- F10     1     106.69 414.25 110.19
- F1      1     129.14 436.70 111.51
- F18     1     131.48 439.04 111.64
- F16     1     147.06 454.62 112.52
- F17     1     163.51 471.07 113.40
- F15     1     198.77 506.33 115.21
- F13     1     252.81 560.37 117.74
- F20     1     285.31 592.87 119.15
- F8      1     313.08 620.64 120.30
- F2      1     367.03 674.59 122.38


Step:   AIC=102.81
F21 ~ F1 + F2 + F3 + F4 + F5 + F6 + F7 + F8 + F9 +
F10 + F12 +
    F13 + F14 + F15 + F16 + F17 + F18 + F19 + F20


        Df Sum of Sq      RSS      AIC
- F14     1       1.02 309.33 100.89
- F3      1       3.06 311.37 101.05
- F9      1       6.39 314.70 101.32
- F7      1      10.05 318.37 101.61
- F4      1      16.66 324.98 102.12
- F19     1      17.61 325.92 102.19
<none>                 308.31 102.81
- F5      1      32.13 340.44 103.28
- F6      1      34.12 342.43 103.43
- F12     1      76.88 385.19 106.37
```

```
- F1      1      135.41 443.72 109.91
- F18     1      164.04 472.35 111.47
- F16     1      175.22 483.53 112.06
- F17     1      200.31 508.62 113.32
- F15     1      224.77 533.08 114.50
- F10     1      226.96 535.27 114.60
- F13     1      253.50 561.81 115.81
- F8      1      318.83 627.15 118.56
- F20     1      326.19 634.50 118.85
- F2      1      543.62 851.93 126.22


Step:   AIC=100.89
F21 ~ F1 + F2 + F3 + F4 + F5 + F6 + F7 + F8 + F9 +
F10 + F12 +
    F13 + F15 + F16 + F17 + F18 + F19 + F20


        Df Sum of Sq      RSS      AIC
- F3      1         2.34  311.66   99.076
- F7      1        10.85  320.18   99.750
- F19     1        16.61  325.93  100.195
- F9      1        16.70  326.03  100.203
<none>                    309.33  100.888
- F4      1        74.50  383.83  104.283
- F12     1        77.96  387.29  104.507
- F6      1       107.69  417.02  106.356
- F5      1       129.30  438.63  107.620
- F1      1       134.43  443.76  107.910
- F18     1       166.66  475.99  109.663
- F10     1       242.31  551.63  113.350
- F13     1       283.32  592.64  115.143
```

```
- F16    1     321.98   631.30 116.723
- F17    1     366.45   675.78 118.425
- F8     1     391.19   700.52 119.324
- F15    1     658.97   968.30 127.417
- F2     1     725.20  1034.53 129.071
- F20    1     769.81  1079.14 130.126


Step:   AIC=99.08
F21 ~ F1 + F2 + F4 + F5 + F6 + F7 + F8 + F9 + F10 +
F12 + F13 +
    F15 + F16 + F17 + F18 + F19 + F20


         Df Sum of Sq       RSS       AIC
- F19    1      14.29   325.95   98.197
- F7     1      14.45   326.11   98.209
<none>                  311.66   99.076
- F12    1      77.24   388.90 102.611
- F9     1     116.06   427.73 104.990
- F5     1     126.97   438.63 105.620
- F18    1     165.14   476.80 107.706
- F4     1     174.17   485.83 108.175
- F10    1     252.06   563.72 111.892
- F6     1     257.16   568.83 112.118
- F13    1     283.73   595.40 113.259
- F16    1     324.58   636.24 114.917
- F17    1     364.36   676.02 116.434
- F8     1     470.11   781.78 120.067
- F1     1     573.71   885.37 123.178
- F15    1     672.16   983.83 125.814
- F2     1     723.61  1035.28 127.089
```

```
- F20    1       788.06 1099.72 128.598


Step:   AIC=98.2
F21 ~ F1 + F2 + F4 + F5 + F6 + F7 + F8 + F9 + F10 +
F12 + F13 +
     F15 + F16 + F17 + F18 + F20


         Df Sum of Sq       RSS      AIC
- F7      1      14.41  340.36   97.278
<none>                  325.95   98.197
- F12     1      64.25  390.20  100.695
- F9      1     157.39  483.34  106.046
- F4      1     170.36  496.31  106.708
- F5      1     181.61  507.56  107.268
- F10     1     239.08  565.03  109.950
- F13     1     285.69  611.65  111.932
- F18     1     301.29  627.24  112.561
- F6      1     330.10  656.05  113.684
- F16     1     333.89  659.84  113.828
- F8      1     455.84  781.79  118.068
- F17     1     527.59  853.54  120.263
- F1      1     565.14  891.09  121.339
- F2      1     731.65 1057.60  125.622
- F20     1     791.84 1117.79  127.006
- F15     1     850.30 1176.25  128.280


Step:   AIC=97.28
F21 ~ F1 + F2 + F4 + F5 + F6 + F8 + F9 + F10 + F12 +
F13 + F15 +
     F16 + F17 + F18 + F20
```

```
        Df Sum of Sq      RSS      AIC
<none>                   340.36   97.278
- F12    1      49.84   390.20   98.695
- F5     1     167.62   507.99  105.289
- F9     1     262.76   603.12  109.581
- F10    1     269.44   609.80  109.856
- F18    1     290.65   631.01  110.711
- F13    1     313.38   653.74  111.596
- F16    1     386.37   726.73  114.242
- F6     1     400.23   740.59  114.714
- F8     1     450.06   790.42  116.342
- F17    1     530.39   870.75  118.762
- F1     1     551.10   891.46  119.350
- F4     1     553.85   894.22  119.427
- F2     1     769.69  1110.05  124.832
- F20    1     777.52  1117.88  125.008
- F15    1     861.77  1202.14  126.825


Call:
lm(formula = F21 ~ F1 + F2 + F4 + F5 + F6 + F8 + F9 +
F10 + F12 +
    F13 + F15 + F16 + F17 + F18 + F20, data = df)


Residuals:
    Min      1Q  Median      3Q     Max
-6.8043 -1.8328 -0.0138  1.2456  8.2085


Coefficients:
          Estimate Std. Error t value
```

```
(Intercept)  -2.211e-15  1.230e+00    0.000
F1            1.083e-02  2.836e-03    3.817
F2           -1.618e-03  3.587e-04   -4.511
F4           -2.313e+00  6.043e-01   -3.827
F5            1.299e-01  6.172e-02    2.105
F6            9.745e-02  2.996e-02    3.253
F8            1.241e+01  3.598e+00    3.450
F9           -1.002e+01  3.800e+00   -2.636
F10           4.520e+00  1.693e+00    2.669
F12          -8.236e-02  7.174e-02   -1.148
F13           4.742e-03  1.647e-03    2.879
F15          -1.243e+02  2.603e+01   -4.774
F16           5.408e+01  1.692e+01    3.196
F17          -3.103e+01  8.285e+00   -3.745
F18          -3.405e+01  1.228e+01   -2.772
F20           1.246e+02  2.748e+01    4.534
             Pr(>|t|)
(Intercept)  1.00000
F1           0.00411 **
F2           0.00146 **
F4           0.00405 **
F5           0.06456 .
F6           0.00995 **
F8           0.00728 **
F9           0.02709 *
F10          0.02566 *
F12          0.28057
F13          0.01822 *
F15          0.00101 **
F16          0.01090 *
```

```
F17          0.00459 **
F18          0.02167 *
F20          0.00142 **
---
Signif. codes:
0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


Residual standard error: 6.15 on 9 degrees of freedom
Multiple R-squared:  0.9153,    Adjusted R-squared:
0.7743
F-statistic: 6.488 on 15 and 9 DF,  p-value: 0.003833
```