



# ANALYSIS OF MATERNAL MORTALITY RATE AND IMPACT OF OTHER RELATED ATTRIBUTES

Varshini Madaka<sup>1</sup>, A. M. Sowjanya<sup>2</sup>

Department of Computer Science & Systems Engineering, College of Engineering (A), Andhra University Visakhapatnam,  
India<sup>1,2</sup>

**ABSTRACT:** Maternal Mortality Rate(MMR) refers to deaths due to complications from pregnancy and/or childbirth. Complications of childbirth and pregnancy are predominant reasons of death among women of reproductive age across the world. Setting up of goals for mortality reduction is vital, but it we must also acknowledge the fact that accurate measurement of maternal mortality remains still challenging and many deaths still go uncounted or unnoticed. As per the data published by WHO, the WHO African region has got the highest Maternal Mortality Rate when compared it with its other five regions. In this paper we predict how Maternal Mortality Rate(MMR) is related to the following attributes - GDP growth, Maternal Leave Policy, Literacy rate and Health care expenditure among the different regions of the world . As part of the predictions, Child Mortality Rate(CMR) is shown in relation with Maternal Mortality Rate(MMR); CMR among different age groups; Maternal mortality due to direct and indirect health causes like cerebrovascular arrest, cardio vascular arrest, infection; women age impact on Maternal mortality; Impact of Delivery methods on maternal mortality. Linear Regression algorithm and several Data visualization techniques are being used to visualize the impact of maternal mortality rate due to the above said parameters using normal Python and Microsoft Power BI tool for better visualization. From the predictions depicted from this paper, it is very clearly evident that GDP, health care expenditure have positive correlation with decline of MMR over the past few years; Child Mortality Rate is more due to Maternal Mortality Rate; C-Section Delivery method rate increasing over the past few years; Cause of maternal mortality due to other cardio vascular diseases tends to decline over years. The WHO recommends that oxytocin should be administered within a minute of birth to lower the risk of postpartum hemorrhage. Our model concludes that attributes like Maternal leave policy, fertility rate, literacy, health care expenditure have a major impact on Maternal Mortality Rate(MMR) compared to other attributes.

**KEYWORDS:** Maternal Mortality Rate, GDP, Health care expenditure, Child Mortality Rate, Fertility rate, Linear Regression, Ordinary Least Square Algorithm, Exploratory Data Analysis, Power BI .

## I.INTRODUCTION

Maternal mortality represents an uncomfortable medical state of complication across the world. A pregnancy-related death is defined as the death of a woman during pregnancy or within one year of the end of pregnancy from a pregnancy complication or, a chain of events related to pregnancy. The World Health Organization's (WHO's) 10th revision of the International Statistical Classification of Diseases and Related Health Problems (ICD-10) defines maternal mortality as "the death of a woman while pregnant or within 42 days of termination of pregnancy irrespective of the duration and the site of the pregnancy, from any cause related to or aggravated by the pregnancy or its management but not from accidental or incidental causes"[1]. From 2000 to 2017, the global maternal mortality ratio declined by 38 per cent from 342 deaths to 211 deaths per 100,000 live births[2]. According to UN inter-agency estimates which is less than half the 6.4 percent annual rate needed to achieve the Sustainable Development global goal of 70 maternal deaths per 100,000 live births by 2030. Thus, the data suggests that there is a lot of work left behind to reduce the MMR and bring it to the actual estimates. As there is no realistic assessment of the reasons for MMR, Maternal Mortality Review Committees were constituted to identify, review, and characterize maternal deaths; and identify prevention measures by using all the available opportunities. Statistics reveal that about 295,000 women died during and following pregnancy and childbirth in 2019 across the globe marking the MMR is unacceptably High. Large proposition of these deaths (approx. 94%) have occurred only in undeveloped countries due to lack of resources, illiteracy, lower GDP. The primary possibility for maternal mortality include but no limited to, poverty and poor nutritional values, malnutrition prior to pregnancy and during pregnancy, first pregnancy or higher than fourth pregnancy (excessive fertility), maternal age younger than twenty or older than thirty-four years, infections, female genital mutilation, unintended pregnancies and unsafe abortions, smoking, and inadequate health care during pregnancy and delivery with either direct or indirect impact. From the time FDA issued guidance, pregnancy safety studies conducted by FDA have expanded data from pregnancy exposure registries (pregnancy registries)328 to also include other types of epidemiologic studies. The development of pregnancy 31 safety studies requires specialized knowledge in a variety of areas, including expertise in the 32 fields of epidemiology, clinical teratology, obstetrics, pediatrics, clinical genetics, and statistics when designing a study[3].

## II. RELATED WORK

Several methods do exist which could help us to predict Maternal Mortality Rate due to various factors. Ordinary Least Square algorithm best fits the model that proves all the assumptions of Linear Regression. Development of the MMR estimates, consultations with countries were carried out during July–August 2010 by WHO. Miscoding by ICD-9 or ICD-10, most often seen in cases of deaths caused by cerebrovascular diseases and cardiovascular diseases[4]. At a Period from 1990 to 2015, many countries have made notable progress in collecting data through civil registration systems, surveys, censuses and specialized studies over the past decade. The methodology employed by the Maternal Mortality Estimation Inter-Agency Group (MMEIG) in this round generated using a Bayesian approach[5]. From 2015 consultations with countries were carried out during May and June 2019. This process generated additional data for inclusion in the maternal mortality estimation model, demonstrating widespread expansion of in-country efforts to monitor maternal mortality[6]. Objective was to describe maternal mortality rates in a large, multi-country dataset and to determine maternal, pregnancy-related, delivery and postpartum characteristics that are associated with maternal mortality. collected data describing all pregnancies from 2010 to 2013 among women enrolled in the multi-national Global Network for Women's and Children's Health Research Maternal and Neonatal Health Registry (MNHR). We reported the proportion of mothers who died per pregnancy and the maternal mortality ratio (MMR). Generalized linear models were used to evaluate the relationship of potential medical and social factors and maternal mortality and to develop point and interval estimates of relative risk associated with these factors[7]. WHO African Region. Maternal mortality on GDP was estimated using a double log econometric model. The analysis is based on cross-sectional data for 45 of the 46 Member States in the WHO African Region. All the explanatory variables included in the double log model were found to have statistically significant effect on per capita gross domestic product (GDP) at 5% level in a t-distribution test. The study used Production Function (PF) conceptual framework to estimate the GDP loss attributable to maternal mortality in the Region. Maternal mortality has an adverse effect on future human capital creation process, on the quality of future labor force, and hence, future levels of GDP[8]. The model combines the rate of change implied by a multilevel regression model with a time-series model to capture data-driven changes in country-specific MMRs, and includes a data model to adjust for systematic and random errors associated with different data sources.

## III. RESEARCH METHODOLOGY

A description of methodology is represented about the framework that is undertaken. It consists of various milestones that need to be achieved to fulfill the objective. We have used Linear Regression Algorithm[9]. Our target was predicting mortality rate with the rest of the factors being our independent variables.

### 1. DATA COLLECTION

CSV files related to maternal mortality are collected from the “WHO”[10] used as datasets for the project. Data sets used are Mortality, Fertility Rate, Population mean, health care expenditure, fertile women population, maternity leave, GDP. Worked on few Data sets such as Child Mortality Rate, Probability of dying 1000 births per country of Neonatal, under age 1, age 5 babies, Delivery by type, Maternal and child Mortality Rate, Maternal by age, cause of pregnancy due to other related diseases to show the relationship with Maternal Mortality Rate using different data visualization techniques.

Since Maternal Mortality Rate is a dependent variable, rest of the variables like Fertility rate, GDP, healthcare expenditure are treated as Independent variables. Table 1 list outs the parameters used for modeling.

Parameters	Description	Datatype
Country Name	Name of the country	String
Year	Year	Numerical
Population	Population of the country	Numerical
w_lit_pop	Women population who are literate	Numerical
fertile_women_pop	Population of Fertile women	Numerical
fertility_rate	Fertility of women in a country	Numerical
maternity_leaves	Leave policy associated in a country	Boolean
Gdp	GDP of a country	Numerical
Healthcare	Health Care expenditure associated to the country	Boolean
hc_expend	Health care expenditure rate allocated to country	Numerical
mortality_rate	Maternal Mortality Rate	Numerical
CMR	Child Mortality Rate	Numerical
CMR_nonMMR	Child Mortality Rate due other reasons rather than Maternal Mortality Rate	Numerical
GHO_CODE	Code of Infant mortality rate (probability of dying between birth at different age groups)	String
SEX	Gender	String
Display_Value	Mortality Rate of Baby	Numerical

Parameters	Description	Datatype
DTC	Code of Delivery By Type	String
Value	Rate of Delivery type performed in different countries over years	Numerical
Tot_Per	Percentage of Maternal women died at different age groups	Numerical
age_15_to_17	Maternal Women died at age 15 to 17	Numerical
age_18_to_19	Maternal Women died at age 18 to 19	Numerical
age_less_than_20	Maternal women died at age less than 20	Numerical
age_more_than_20	Maternal women died at age more than 20	Numerical
age_20_to_34	Maternal women died at age 20 to 34	Numerical
age_35_to_49	Maternal women died at age 35 to 49	Numerical
Hemorrhage	The rate of Hemorrhage	Numerical
Infection	The rate of Infection	Numerical
Amniotic fluid embolism	The rate of Amniotic fluid embolism	Numerical
Thrombotic pulmonary embolism	The rate of Thrombotic pulmonary embolism	Numerical
Hypertensive disorders of pregnancy	The rate of Hypertensive disorders of pregnancy	Numerical
Anesthesia complications	The rate of Anesthesia complications	Numerical
Cerebrovascular accident	The rate of Cerebrovascular accident	Numerical
Cardiomyopathy	The rate of Cardiomyopathy	Numerical
Other cardiovascular conditions	The rate of Other cardiovascular conditions	Numerical
Other non cardiovascular conditions	The rate of Other non cardiovascular conditions.	Numerical

**Table 1. List of Parameters used for Analysis**

## 2. DATA PREPROCESSING

To apply machine learning algorithms[11][12], we examine the available data and make it appropriate for applying the models on them to get the best possible accuracies during modeling.

### 2.1 DATA CLEANING

It is a process of transforming the raw, complex data into systematic understandable and meaningful data.

### 2.2 MISSING OR NULL DATA POINTS

Finding out missing and redundant data that are in the dataset. Entire dataset is checked for NaN and whichever observation consists of NaN will be deleted. Thus, this brings uniformity in the dataset. It means that every record was constituted its corresponding feature values.

### 2.3 OUTLIERS

The data point which lies far away from other observations which can mislead the final predictions which reduces the accuracy of the results. The Outliers occur due to some kind of experimental errors and they need to be excluded from the dataset. So, we have to remove the outliers of dataset.

In our data we don't have much outliers. Instead of removing the outliers, we will just manage to make some of our models robust on them. Fig.1. clearly shows there are less outliers in our dataset.

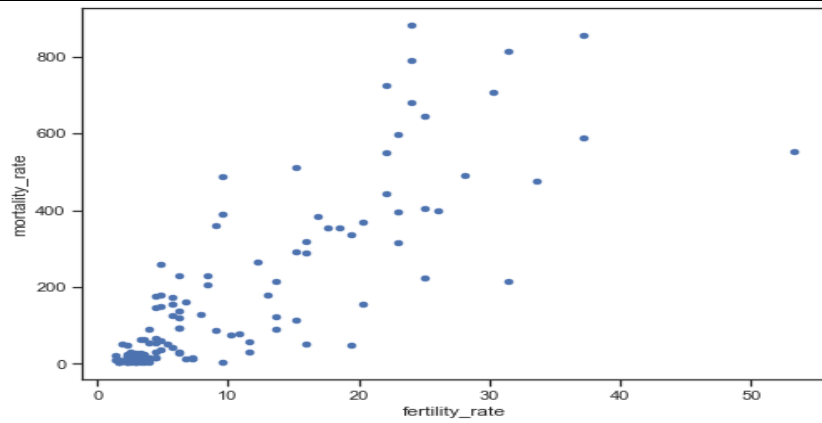


Fig.1. Scatter plot between Mortality Rate and Fertility Rate

### 3 DATA ANALYSIS

We have to determine the characteristics of the dataset before applying any model to the selected dataset to find out any model. Thus, the data which is collected should be analyzed followed by cleaning and processing. During this process, we need to study the different parameters and their relationships. In this paper we have used EDA[13] for Data Analysis.

#### Exploratory Data Analysis (EDA)

Exploratory Data Analysis (EDA) is an approach/philosophy for data analysis that employs a variety of techniques to

1. Maximize insight into a data set.
2. Extract important variables.
3. Detect outliers and anomalies.
4. Test underlying assumptions.

In this paper we used EDA technique by maximizing insight into the data set and extracting important variables in the datasets. We also scaled and adjust to create a fit model.

### 4 .EVALUATION METHOD

Train/Test Split: The model would be over fit if we run same data for both training and testing the model. So, we can use some of the data to be feed to the algorithm and the remaining to the test data to evaluate our model. Using python Machine Learning (ML) the dataset is being split into train set and test set around 30%-70%.

### 5. DATA CORRELATION USING A HEATMAP

Correlation map to see how features are correlated with each other. A correlation number shows the degree of combination between two variables and between +1 to -1 exists the correlation number. A positive number gives positive number between two variables and vice versa and if the correlation number is 0, then it results no correlation between two variables and thus, there is no similarity. We use heatmap for visualize the correlations. Here the correlation factor is set to 0.7.

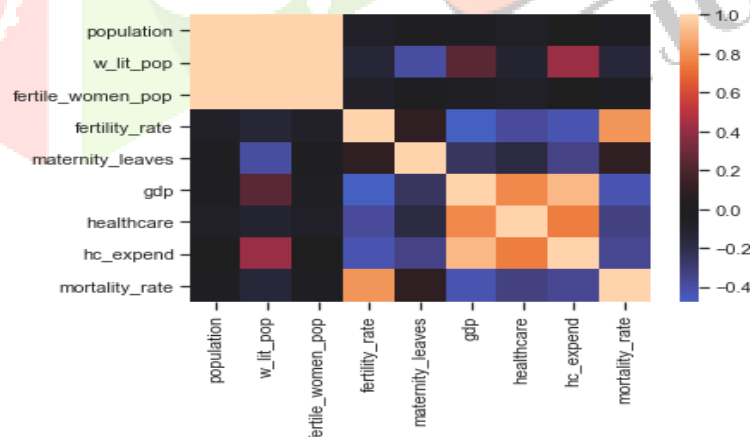


Fig.2. Data correlation of dataset

### 6. DATA VISUALIZATION TECHNIQUES

Data visualization[14] refers to the techniques used to communicate data or provide information by encoding it as visual objects. In this paper, different visualization techniques were used to plot using datasets to visualize relation between Maternal Mortality Rate and Child Mortality Rate, Probability of babies dying among different ages with respective gender rates, Maternal Mortality among different ages, Maternal mortality due to direct and indirect causes. In this paper visualization was shown in Experimental results section using Python Libraries and Microsoft Power BI[20].

7. APPLYING ALGORITHMS

We gain more insights about the dataset once the dataset is clean and then we can apply an effective Machine Learning model suites to our dataset. We are using Linear Regression in our dataset to predict the dependent variable. We are training them to predict the continuous values. Thus, the algorithms are implemented using python’s SciKit-learn , Scipi, numpy, pandas, seaborn, matplotlib, statsmodels libraries[15][16].

LINEAR REGRESSION

Linear Regression[17]is a supervised machine learning algorithm where the predicted output is continuous and has a constant slope. It is used to predict values within a continuous range instead of trying to classify them into classes. Based on the dataset the data points are plotted using the linear equation. The linear regression equation form is stated as  $Y = a + bX$ , where Y is the dependent variable which is plotted on Y axis, X is the independent variable which is plotted on the X axis and b is the slope of the line and is the y-intercept.

Ordinary Least Square (OLS) estimates

There are six assumptions[18] associated with the linear regression model. Ordinary Least Squares (OLS) is the most common estimation method for linear models. Based on the model assumptions, we are able to derive estimates on the intercept and slope that minimize the **sum of squared residuals (SSR)**. Minimizing the SSR is a desired result, since we want the error between the regression function and sample data to be as small as possible. The coefficient estimates that minimize the SSR are called the **Ordinary Least Squared (OLS) estimates**. Ordinary Least Squares (OLS) is used for estimating the regression coefficients.

MODELING

In linear regression, the relationships are modeled using linear predictor functions whose unknown model parameters are estimated from the data. Such models are called linear models. Before starting the model we have created a list of countries as an index. The test data size 0.2 and training data are been set up ad we fitted the model using considered attributes by adjusting r-squared value by 0.80.

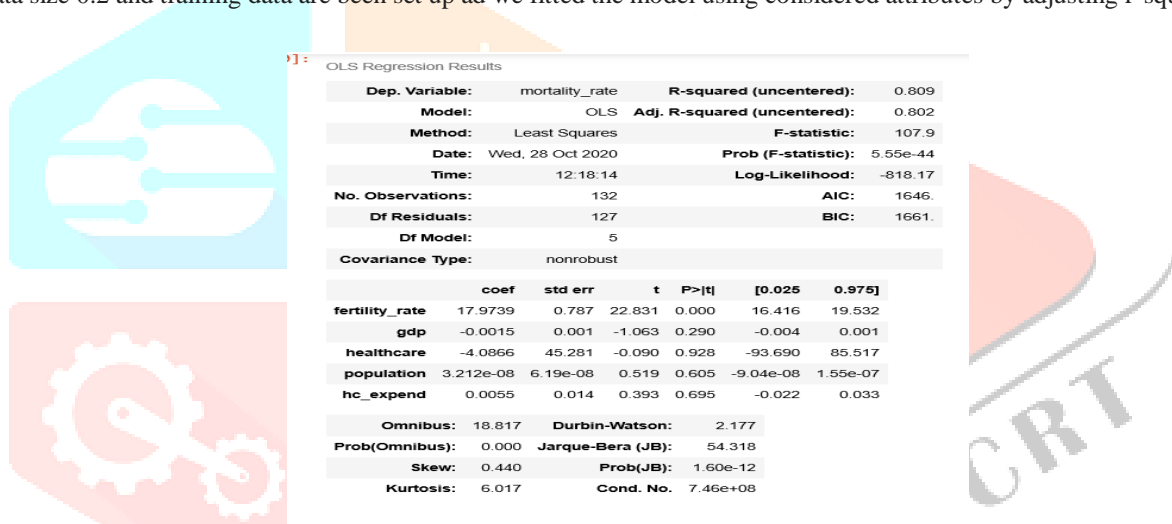


Fig.3 Results of fitness of the Model

IV. EXPERIMENTAL RESULTS

The following section shows the results of the algorithm we have applied. Implementing the Linear Regression on the datasets by training the data fitting the model among the independent attributes like Fertility Rate, GDP and Health care with dependent variable Mortality Rate the results are shown below. Our first test shows the effect of fertility rate on mortality rate. We hypothesized that there would be a strong positive correlation. The results of Regression residual plots obtained between Fertility Rate and Mortality Rate determine there is a pattern and residual line which is close to zero have all its residual plots closer to it. Fig. 4 proves there is a linear relation between Fertility and Maternal Mortality Rate(MMR).

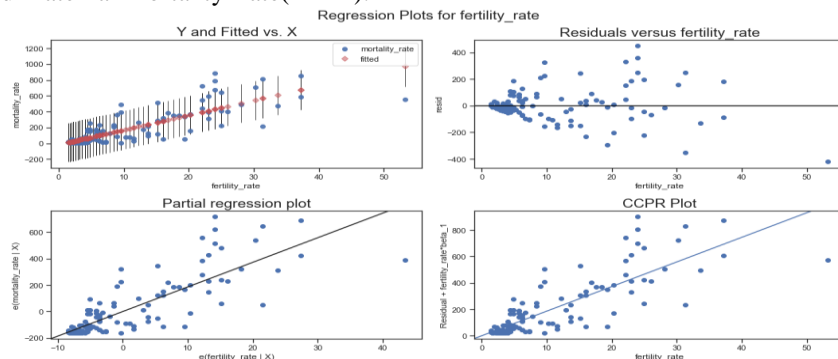


Fig.4 Regression Residual Plots between Fertility rate and MMR

In our second test, we compared a country's GDP in relation to it’s expenditure on healthcare and it's affect on maternal mortality rate. From the Fig.5 we found an inverse relationship (negative correlation) and Fig.6 and Fig.7 represents positive correlation. This result proves a lower Maternal Mortality Rate can be obtained by the countries who provides better healthcare for the women.

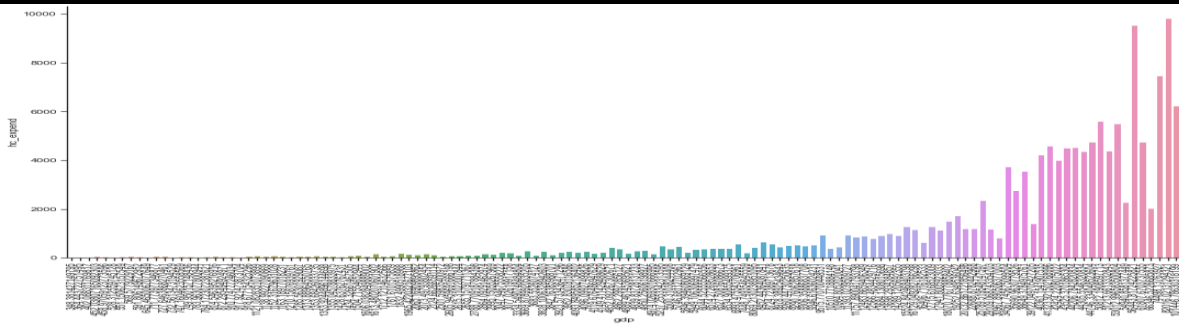


Fig.5 Visualization between GDP and Health care expenditure.

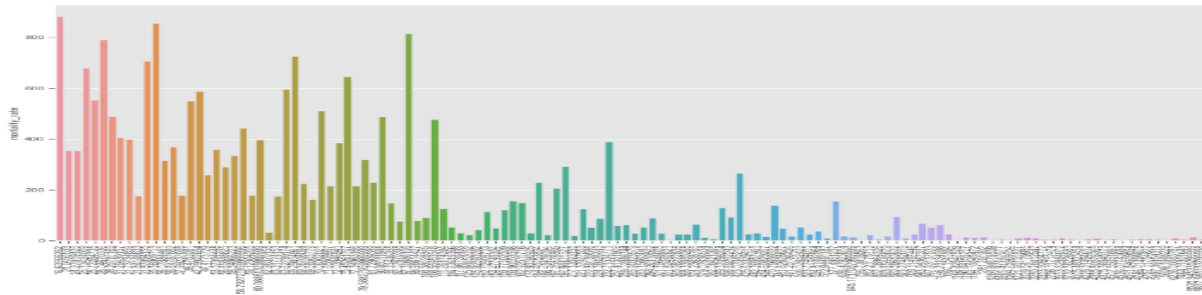


Fig.6 Visualization between Mortality Rate and Health care expenditure.

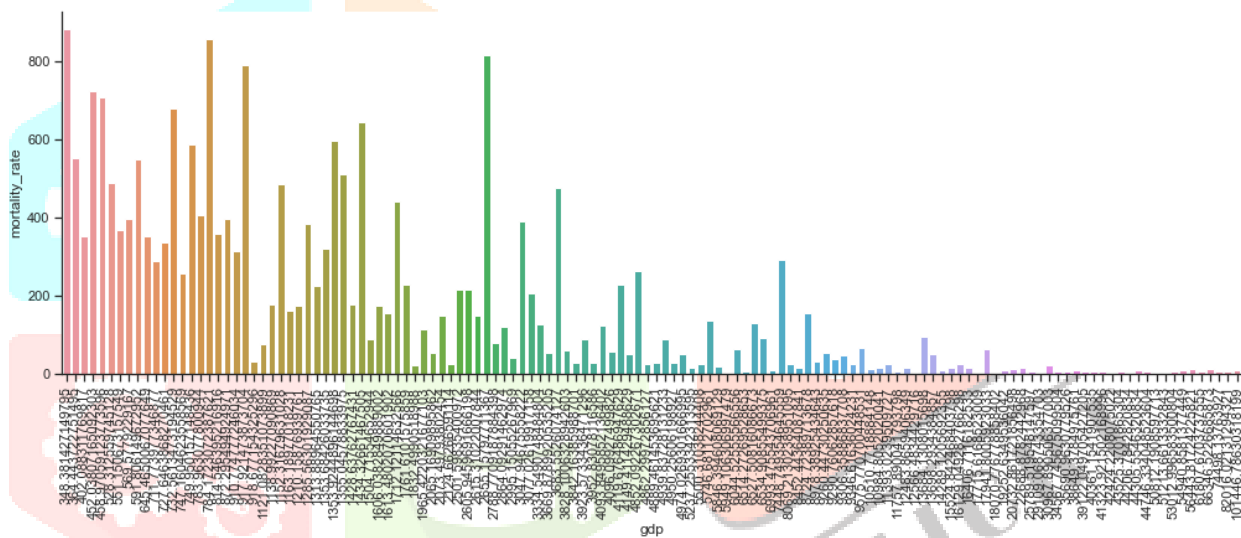


Fig.7 Visualization between Mortality Rate and GDP.

The visualization of impact of Maternal Mortality Rate(MMR) progress due to other related factors are shown in below figures.



Fig.8 Progress of Mortality across the world over years.

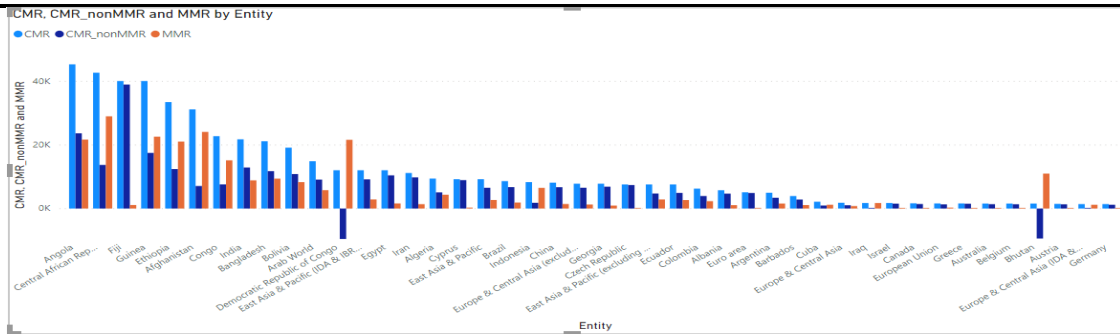


Fig.9 Impact of Child Mortality due to Maternal Mortality

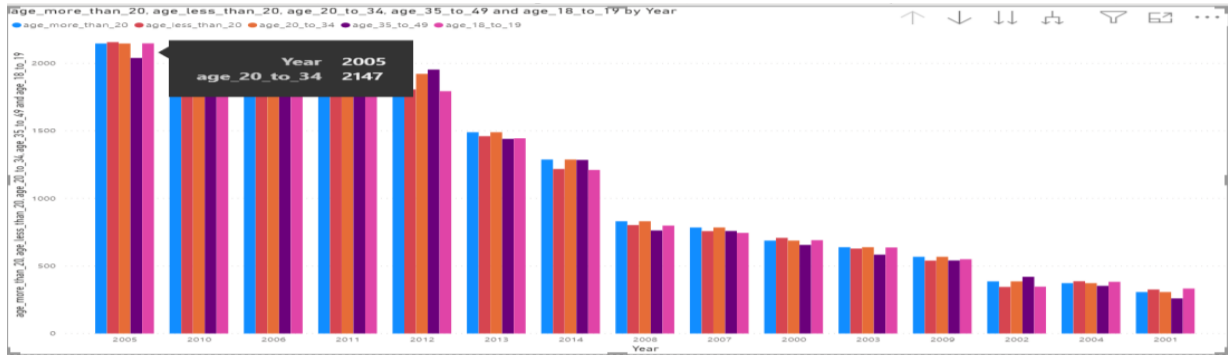


Fig.10 Maternal Mortality Rate among Different age groups throughout the world over years.

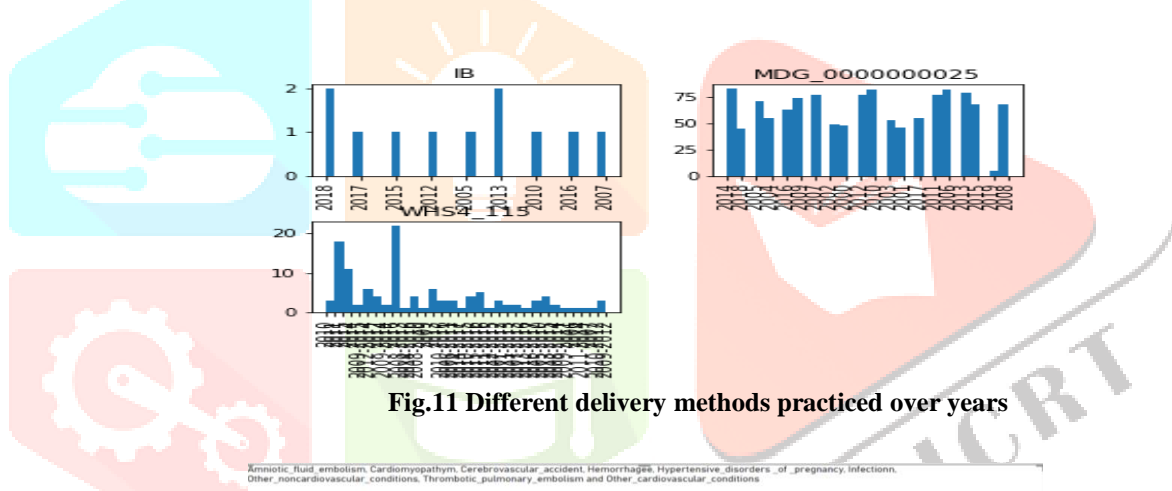


Fig.11 Different delivery methods practiced over years

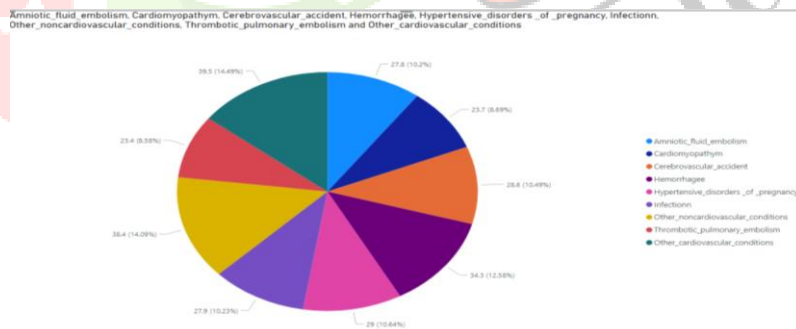


Fig.12 Maternal Mortality due to direct and indirect causes over different years in US

Finally, to show the accuracy[19] among the Observed data(i.e., trained data) and Predicted data(i.e., test data) using our model the accuracy table and Residual plot is provided below.

Accuracy	Value
MSE Test data	2.294684e-23
MSE Train data	3.444602e-22
MAE Train data	6.761019e-12
MAE Test data	4.012898e-12
RMSE Train data	1.855964e-11
RMSE Test data	4.790286e-12

Table.2 Accuracy of the Model

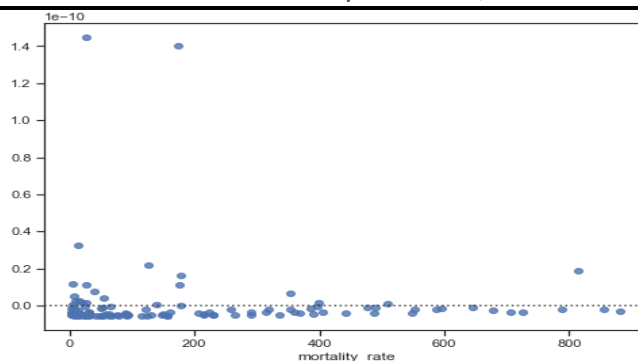


Fig.13 Residual plot of Model(Predicted Vs Expected)

## V. CONCLUSIONS

We can conclude that our considered model Linear Regression best fits for Maternal Mortality Rate Analysis. From above obtained plots and tables, it is proven from our model that there is no multi collinearity and there is no normal distribution. Adjusting the r-squared value to 80% with a data size of 0.2, the p-values obtained for our model are less than 1. In our model, it is demonstrated that accuracy values for trained data and test data are relatively similar and small and also Linear score for our considered model is 1. Thus, it indicates that it is a best fit line of our model beside leaving few outliers which is close to zero. By examining all the above-mentioned factors satisfied all the assumptions of OLS[10]. Hence it can be concluded that our model works well and is able to predict maternal mortality rate given the independent variables.

## VI. REFERENCES

- <https://www.ncbi.nlm.nih.gov/books/NBK361917/#:~:text=The%20International%20Classification%20of%20Diseases,but%20not%20from%20accidental%20or>
- <https://data.unicef.org/topic/maternal-health/maternal-mortality/>
- [www.fda.gov.in](http://www.fda.gov.in)
- Trends in Maternal Mortality: 1990 to 2008 “John Wilmoth, Sarah Zureich, Nobuko Mizoguchi, Mie Inoue, Mikkel Oestergaard”, 14 September 2010
- Trends in Maternal Mortality: 1990 to 2015 - Alkema L, Chou D, Hogan D, Zhang S, Moller AB, Gemmill A, Fat DM, Boerma T, Temmerman M, Mathers C, Say L; United Nations Maternal Mortality Estimation Inter-Agency Group collaborators and technical advisory group. Global, regional, and national levels and trends in maternal mortality between 1990 and 2015, with scenario-based projections to 2030: a systematic analysis by the UN Maternal Mortality Estimation Inter-Agency Group. Lancet. 2016 Jan 30;387(10017):462-74. doi: 10.1016/S0140-6736(15)00838-7. Epub 2015 Nov 13. PMID: 26584737; PMCID: PMC5515236.
- Trends in Maternal Mortality-2000 to 2017. - WHO, UNICEF, UNFPA, World Bank Group and the United Nations Population Division ,September 2019. Available at <https://apps.who.int/iris/handle/10665/327596>.
- Risk factors for maternal death and trends in maternal mortality in low- and middle-income countries: a prospective longitudinal cohort analysis , Citation :Melissa Bauserman, Adrien Lokangaka, Vanessa Thorsten, Antoinette Tshetu, Shivaprasad S Goudar, Fabian Esamai, Ana Garces, Sarah Saleem, Omrana Pasha, Archana Patel, Albert Manasyan, Mabel Berrueta, Bhala Kodkany, Elwyn Chomba, Edward A Liechty, K Hambidge, Nancy F Krebs, Richard J Derman, Patricia L Hibberd, Fernando Althabe, Waldemar A Carlo, Marion Koso-Thomas, Robert L Goldenberg, Dennis D Wallace, Elizabeth M McClure, Carl L Bose PMID: 26062992 PMCID: PMC4464034 DOI: 10.1186/1742-4755-12-S2-S5 <https://pubmed.ncbi.nlm.nih.gov/26062992/>
- Effects of Maternal Mortality on GDP <https://pubmed.ncbi.nlm.nih.gov/17348747/> Kirigia JM, Oluwole D, Mwabu GM, Gatwiri D, Kainyu LH. Effects of maternal mortality on gross domestic product (GDP) in the WHO African region. Afr J Health Sci. 2006 Jan-Jun;13(1-2):86-95. doi: 10.4314/ajhs.v13i1.30821. PMID: 17348747.
- [https://www.tutorialspoint.com/machine\\_learning\\_with\\_python/machine\\_learning\\_with\\_python\\_ecosystem.htm](https://www.tutorialspoint.com/machine_learning_with_python/machine_learning_with_python_ecosystem.htm)
- <https://apps.who.int/gho/data/node.imr>
- Data Mining: Concepts and Techniques , 3rd ed. The Morgan Kaufmann Series in Data Management Systems
- <https://medium.com/analytics-vidhya/data-cleaning-and-preprocessing-a4b751f4066f>
- <https://towardsdatascience.com/exploratory-data-analysis-8fc1cb20fd15>
- <https://www.geeksforgeeks.org/seaborn-categorical-plots/>
- <https://www.dataquest.io/blog/sci-kit-learn-tutorial/>
- <https://www.edureka.co/blog/python-libraries/>
- <https://towardsdatascience.com/linear-regression-using-python-b136c91bf0a2>
- <https://www.albert.io/blog/key-assumptions-of-ols-econometrics-review/>
- <https://www.datatechnotes.com/2019/10/accuracy-check-in-python-mae-mse-rmse-r.html>
- <https://www.udemy.com/topic/microsoft-power-bi/>