Analysis of Multiple Regression Model on COVID-19

ADEOYE A.O, SANUSI A.O., ALAO A.N

Abstract: **This research analysis of multiple regressions on covid-19 in Nigeria was aimed to access the pattern of daily increase in covid-19 cases in Nigeria to determine a model for the covid-19 cases in Nigeria and to examine the significant reliability of the fitted model. Data was obtained from the record of the world health organization, time plot and regression analysis was used to analysis the data with the use of EXCEL. The result of the analysis shows that the cases of covid-19 are increasing on daily bases on the available data. Also the fitted model is $Y=70293.51+18.17X_1+16.73X_2-1772.898X_3$ and the reliability shows that the regression equation is reliable that means, this equation can be used for future perdition of covid-19.**

Keywords: Model, Pattern, Reliability, Regression, covid19

I. INTRODUCTION

The novel human corona virus disease COVID-19 has become the fifth documented pandemic since the 1918 flu pandemic. COVID-19 was first reported in Wuhan, China, and subsequently spread worldwide(Chen et al., 2019).. The corona virus was officially named severe acute respiratory syndrome corona virus 2 (SARS-CoV-2) by the International Committee on Taxonomy of Viruses based on phylogenetic analysis(Burke et al., 2020).. SARS-CoV-2 is believed to be a spillover of an animal corona virus and later adapted the ability of human-to-human transmission. Because the virus is highly contagious, it rapidly spreads and continuously evolves in the human population (Gorbalenya et al., 2020).. Currently, people all over the world have been affected by corona virus disease 2019 (COVID-19), which is the fifth pandemic after the 1918 flu pandemic. As of now, we can trace the first report and subsequent outbreak from a cluster of novel human pneumonia cases in Wuhan City, China, since late December 2019(Guan et al., 2019). The earliest date of symptom onset was 1 December 2019. The symptomatology of these patients, including fever, malaise, dry cough, and dyspnea, was diagnosed as viral pneumonia(Kamp et al., 2020). Initially, the disease was called Wuhan pneumonia by the press because of the area and pneumonia symptoms. Whole-genome sequencing results showed that the causative agent is a novel coronavirus. Therefore, this virus is the seventh member of the corona virus family to infect humans (Rothe et al., 2019). The World Health Organization (WHO) temporarily termed the new virus 2019 novel corona virus (2019-nCoV) on 12 January 2020 and then officially named this infectious disease corona virus disease 2019 (COVID-19) on 12 February 2020(Wu et al., 2020).. Later, the International Committee on Taxonomy of Viruses (ICTV) officially designated the virus as SARS-CoV-2 based on phylogeny, taxonomy and established practice (Shen et al., 2020). Subsequently, human-to-human transmission of COVID-19 occurring within Hong Kong has been shown in clinical data (Wang et al., 2019). Since COVID-19 initially emerged in China, the virus has evolved for four months and rapidly spread to other countries worldwide as a global threat. On 11 March 2020, the WHO finally made the assessment that COVID-19 can be characterized as a pandemic, following 1918 Spanish flu (H1N1), 1957 Asian flu (H2N2), 1968 Hong Kong flu (H3N2), and 2009 Pandemic flu (H1N1), which caused an estimated 50 million, 1.5 million, 1 million, and 300,000 human deaths, respectively.

II. METHODOLOGY

MULTIPLE LINEAR REGRESSION

Multitude regression, is a statistical technique that uses several explanatory variables to predict the outcome of a response variable. The goal of multiple linear regression is to model the linear relationship between the explanatory (independent) variables. In essence, multiple regressions is the extension of ordinary least-squares (OLS) regression because it involves more than one explanatory variable.

SOURCE OF DATA

Data used in this research work is a secondary data. The data that was sourced online from (http// nairametric. Com)

III. METHOD OF DATA ANALYSIS

In other to carry out a research successfully, appropriate statistical tools are to be used for the analysis of data collected for the study. In this research, time plot was used to explore the data collected and multiple linear regression model was fitted on the data to access best fit for the data.

REGRESSION MODELS

Multiple Regression Model

Multiple linear regressions refer to a statistical technique that is used to predict the outcome of a variable based on the value of two or more variables. The variable that we want to predict is known as the dependent variable, while the variables we use to predict the value of the dependent variable are known as independent or explanatory variables.

The equation that describes how y is related to x is known as the regression model. The multiple regression model is $\hat{y}$= $a+b_1x_1+b_2x_2+b_3x_3$-------- b_kx_k

Where

a= intercept, value of y when all predictor variable are 0

b_1= Regression coefficient for first predictor variable, x_1

b_2= Regression coefficient for second predictor variable, x_2

b_3= Regression coefficient for third predictor variable, x_3

y= dependent variable

IV. PRESENTATION OF RESULTS

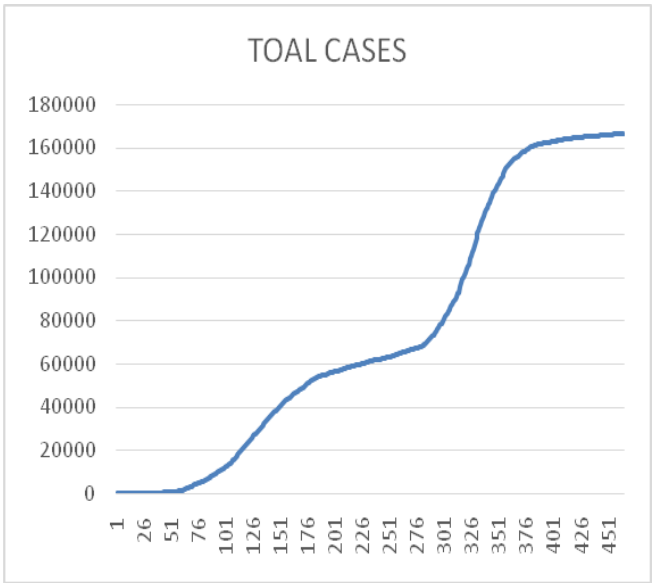

Fig 1: THE PLOT OF TOTAL CASES OF COVID 19 IN NIGERIA

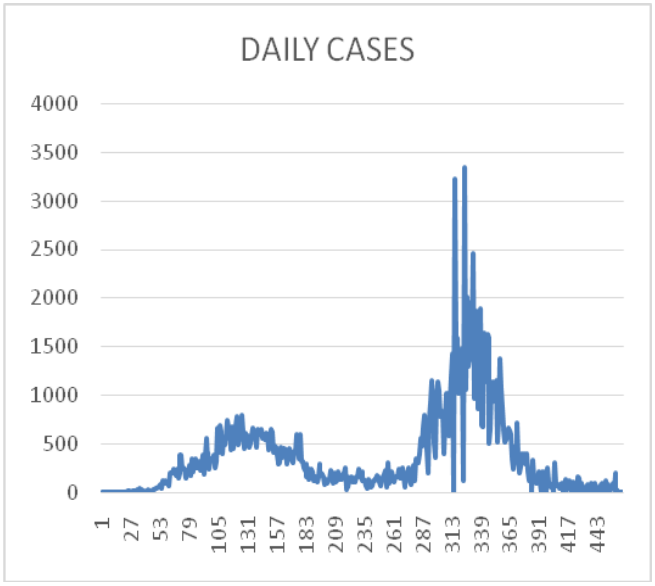

Fig 2: TIME PLOT FOR DAILY CASES

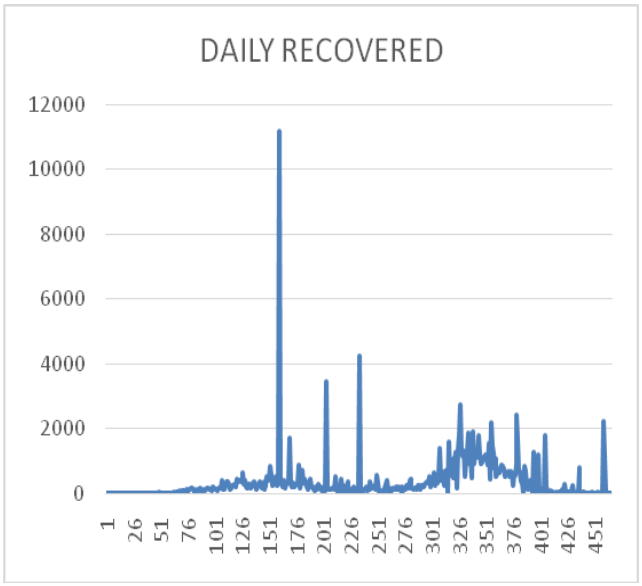

Fig 3: TIME PLOT FOR DAILY CASES

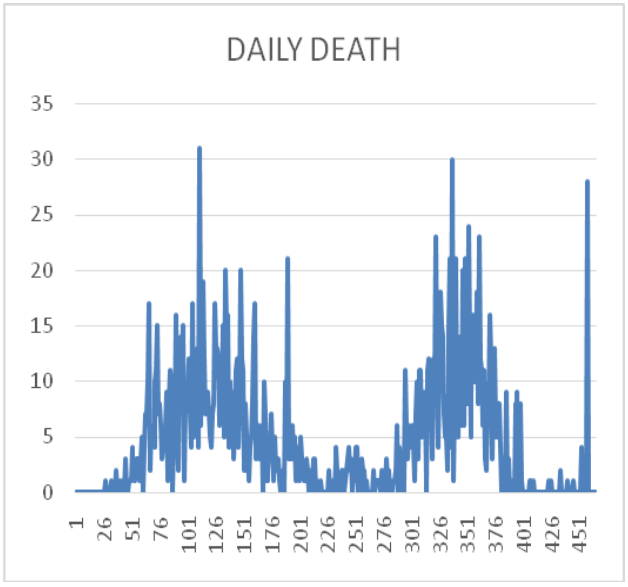

Fig 4: TIME PLOT FOR DALY DEATH CASES

V. ANALYSIS OF THE MODEL

$$\hat{Y} = a_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$$

Where

a = intercept, value of y when all predictor variable are 0

b_1= Regression coefficient for first predictor variable, x_1

b_2= Regression coefficient for second predictor variable, x_2

b_3= Regression coefficient for third predictor variable, x_3

y = dependent variable

DATA ANALYSIS

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.241652329
R Square	0.058395848
Adjusted R Square	0.052281536
Standard Error	58031.78193
Observations	466

ANOVA

	Df	*SS*	*MS*	*F*	*Significance F*
Regression	3	96491130432	32163710144	9.550680726	3.9457E-06
Residual	462	1.55587E+12	3367687714		
Total	465	1.65236E+12			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*
Intercept	70293.50699	3692.656966	19.0360241	4.51576E-60	63037.0224	77549.99158	63037.0224
DAILY CASES	18.17321145	7.378873606	2.462870679	0.014146679	3.67289823	32.67352467	3.67289823
DAILY RECOVERED	16.72659372	4.284722805	3.903774988	0.00010881	8.306633429	25.146554	8.306633429
DAILY DEATH	-1772.898101	570.8685135	-3.105615496	0.002015632	-2894.718679	-651.0775235	-2894.718679

$$\hat{Y} = 70293.50669 + 18.17321145x_1 + 16.72659372x_2 - 1772$$

WHERE
X_1 = Daily cases
X_2 = Daily recovered
X_3 = Daily death
Y = Total cases
FROM THE ANALYSIS
X_1 is significant since p-value < α= 0.5
X_2 is significant since p-value < α= 0.5
X_3 is significant since p-value < α= 0.5

TEST FOR RELIABILITY
H_0 : The regression is not reliable
H_1 : The regression model is reliable
Level of significant
α= 0.05
Test statistic :- ANOVA TABLE
Decision rule : Reject H_0 if p-value < α = 0.05 otherwise do not reject H_0
COMPUTATIONS

Anova: Single Factor

TO FIT THE PREDICTOR INTO THE MODEL

SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
Total cases	466	34737743	74544.51288	3553468504
DAILY CASES	466	166560	357.4248927	191877.5395
DAILY RECOVERED	466	159946	343.2317597	485024.1139
DAILY DEATH	466	2099	4.504291845	31.11933638

ANOVA

Source of Variation	*SS*	*Df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	1.92994E+12	3	6.43313E+11	724.0145507	.000	2.609686593

Within Groups	1.65268E+12	1860	888536359.2
Total	3.58262E+12	1863	

Decision : From our ANOVA table p-value < $\alpha = 0.05$ we reject our H_0
Conclusion: The regression model is reliable

VI. SUMMARY AND CONCLUSION

This study access the pattern of daily increase in covid-19 cases in Nigeria and also examine the significant (reliability) of the fitted model. The analysis was used to check the significant of our predictors. Based on our study from the time plot we can see that the daily cases of covid-19 is increasing daily but later drop .The analysis in this research shows that the model fitted was reliable and total cases shows an increase from the time plot.

VII. RECOMMENDATION

Based on the findings in this research we thereby recommend that Nigerian Center for Disease Control (NCDC should take a proper precaution so as to fight against covid-19.

REFERENCES

1. Burke RM, Midgley CM, Dratch A, Fenstersheib M, Haupt T, Holshue M, .(2020) Active Monitoring of Persons Exposed to Patients with Confirmed COVID-19 - United States, January-February MMWR
2. Chen N, Zhou M, Dong X, Qu J, Gong F, Han Y, (2019). Epidemiological and clinical characteristics of 99 cases of 2019 novel corona virus pneumonia in Wuhan, China: a descriptive study. Lancet 2020;395(10223): 507–13.
3. Gorbalenya AE, Baker SC, Baric RS, de Groot RJ, Drosten C, Gulyaeva AA,(2020). Severe acute Respiratory syndrome-related corona virus: The species and its viruses – a statement of the Corona virus .
4. Guan WJ, Ni ZY, Hu Y, Liang WH, Ou CQ, He JX, (2020). Clinica characteristics of corona virus disease China. N Engl J Med. 2020 Feb 28. doi: 10.1056/NEJMoa2002032. .
5. Kamp f G, Todt D, Pfaender S, Steinmann E. (2020) Persistence of corona viruses on inanimate surfaces and their inactivation with biocidal agents. J Hosp Infect 2020; 104(3): 246–51.
6. National Health Commission of the People's Republic of China. New corona virus cases rise to 571 I Chinese mainland. Available from: URL: http://en.nhc.gov.cn/2020-01/23/c_76004.htm. Accessed, a. January 23, 2020.
7. Rothe C, Schunk M, Sothmann P, Bretzel G, Froeschl G, Wallrauch C, (2020).Transmission of 2019-nCoV Infection from an Asymptomatic Contact in Germany. N Engl J Med 2020; 382(10): 970–1.
8. Report of the WHO-China Joint Mission on Corona virus Disease 2019 (COVID2019). February 16-24, 2020. http://www.who.int/docs/default-source/coronaviruse/who-china-joint-mission-on-covid-19-finalreport.pdf. Accessed, March 4, 2020 Morb Mortal Wkly Rep 2020; 69(9): 245–6. [CrossRef]
9. Report of the WHO-China Joint Mission on Corona virus Disease 2019 (COVID-2019). February 16-24, 2020. Available from: URL: http://www.who.int/docs/default-source/coronaviruse/who-china-join -mission-on-covid-19- final-report.pdf. Accessed, March 04, 2020.
10. Shen K, Yang Y, Wang T, Zhao D, Jiang Y, Jin R, (2020). Diagnosis, treatment, and prevention of 2019 novel corona virus infection in children: experts' consensus statement. World J Pediatr. 2020 Feb 7. doi: 10.1007/s12519-020-00343-7.
11. Wang D, Hu B, Hu C, Zhu F, Liu X, Zhang J, (2020). Clinical characteristics of 138 hospitalized patients with 2019 novel corona virus-infected pneumonia in Wuhan, China. JAMA 2020; 323(11): 1061–9.
12. World Health Organization. Novel corona virus situation report -2. January 22, 2020. Available from: URL: https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200122-sitrep-2-2019-ncov.pdf. Accessed, January 23, 2020.
13. World Health Organization. Director-General's remarks at the media briefing on 2019-nCoV on 11 February 2020. Available from: URL: https://www.who.int/dg/speeches/detail/who-director-general-s-remarks-at-the-media-briefing-on-2019-ncov-on-11-february-2020. Accessed, February 12, 2020.
14. World Health Organization. WHO Director-General's opening remarks at the media briefing on COVID-19 24 February 2020. Available from: URL: https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-themedia-briefing-on-covid-19---24-februar y-2020.Accessed, February 26, 2020.
15. Wu C, Chen X, Cai Y, Xia J, Zhou X, Xu S, et al. Risk factors associated withacute respiratory distress syndrome and death in patients with corona virus disease 2019 pneumonia in Wuhan, China. JAMA Intern Med. 2020 Mar 13. doi:10.1001/jamainternm

[1]**ADEOYE A.O, [2]SANUSI,A.O. AND [3]ALAO A.N**
[1]Department of Statistics Federal Polytechnic Offa
[2]Department of Statistics Federal Polytechnic Offa
[3]Department of Statistics Kwara State Polytechnic Ilorin