Digital Development Initiative of World Bank

# TABLE OF CONTENTS

**List of Appendices**

# List of Figures

# List of Tables

## List of Appendices

*Appendix 1.*  Effective Governance

*Appendix 2.* GNI per capita

*Appendix 3.* Literature review of ARIMA, ACF and PACF

*Appendix 4.* Auto Arima

## 1. Introduction

Technological developments have accelerated faster than any other invention in humanity, impacting nearly half of the world's population in just 2 decades and shaping communities. Technology can help to level the playing field by improving communication, financial inclusion, trade access, and access to public services (United Nation, (n.d.)). Modern technology will aid in overcoming some of the most difficult growth obstacles, such as supplying knowledge control, tackling remoteness, isolation, and providing economic mobility (World Bank, 2016).

The Internet is transforming our work ethics, interact, develop, and exchange knowledge, and coordinate the global migration of information, thoughts, and objects. The size of this transition, however, is still underestimated. Over the last five years, the Internet has accounted for 21% of GDP growth in developed markets. Single customers and small, ambitious young entrepreneurs have earned the greatest benefits from the Internet's empowering impact, whereas large corporations and national economies have profited the greatest benefits from this technology advancement (Manyika and Roxburgh, 2011).

Centered on these views, this paper presents the use of different statistical tools and different analytics models such as (*Clustering (descriptive method)*, *Forecasting using Time series ARIMA (predictive method)*), etc. on the data provided by World Bank to select the best 10 countries which are the best candidates for Digital Development initiative of World Bank which is mainly focused to expand fast and affordable access to Internet and developing different online platforms to increase good governance, social accountability and Rule of laws.

## 2. Literature Review

Literature review done was primarily focused on the Good governance and Corruption control impacts on the technological development, economy, and the management of donor-funds by different organizations such as World bank.

Improvements in governance have seemed to proceed together with the advancement. The World Bank created a data set of good governance focused on perceptions of speech and transparency, democratic stability and the absence of conflict, government efficiency, institutional quality, the legal system, and corruption levels, using more than 100 indicators. The Bank boosted optimism that it had discovered a clear link between its governance metrics and economic performance by claiming to have discovered the secret to economic growth (Sundaram, 2015). It can further be seen in Appendix(1-2) regarding the relation of Good Governance with Economies technological development.

More detailed allocating funds of expenses to a specific collection of tasks for which clear goals, objectives, and the resources needed to achieve them can be specified is possible with project aid. The donor-funded project will also be included in the State's spending, subject to legislative criteria specific to the project and the area in which it is located, and the funds can be disbursed and accounted for by government processes (Kareiva, et al, 2008). World Bank ventures are often conducted in this manner, while normal government financial practices will need to be supplemented with additional expenditure statements and detailed project accounts. For the technological developments in the countries having structured governance and Control of corruption makes use of foreign funding's in the most benefitted way (SOAS, (n.d.)).

Aid techniques are being re-evaluated from the ground up. In recent times, improving good governance in emerging regions has become both a goal and a requirement for receiving development aid (Santiso, 2001). The emerging economies that are primarily focused on technological advancements are working on their good governance to get funds from World bank to develop different sorts of projects.

Based on the literature review, the selection criteria were based on the following indicators such as Corruption Control, Government effectiveness, Rule of law, Political stability in parallel to Internet users. Because the countries having good governance had a better foreign aid management

system which can be a major indicator for cutting down the countries for the World bank aid and can use that aid for the technological development of their country.

## 3. Methodology and Data

The research methodology done in this paper was quantitative based and the data mainly gathered was from the secondary source "World Bank Indicators website". After the completion of literature review regarding the selection criteria of countries, different combination of variables was shortlisted (shown in table 1).

Table. 1 Selected variables and purpose of their usage

| Variables | Purpose for Using |
|---|---|
| Internet users by % of population | Used to check the number of internet users in specific country to analyze the availability of internet providence telling the status of technology development of a country. |
| Rule of Law | Used to analyze the check and balance of law to understand how well the country and the government is handling their laws which indirectly promotes the development of a country. |
| Political Stability | It is used because economic growth and political stability are closely tied. On the other side, the instability that comes with an unpredictable political situation could discourage investment and delay economic growth (Hussain, 2014). So, this helped in filtering the countries which either were less than our min requirement or were more than our requirement. |
| Government Effectiveness | It is used because good governance plays major role in helping the country in getting the aid through IMF and World bank and with this it was easy to filter out countries in a much better way. |
| Regulatory Quality | The evaluation of the government's capacity to develop and enforce sound laws and laws that enable and encourage private sector growth is captured by regulatory quality. It was used to understand the countries private sector development and how well government is managing it. |
| Corruption Control | Corruption is a multifaceted problem. It has deep roots in administrative and legislative institutions, and its impact on growth differs depending on the circumstances of the region. However, while costs can vary and structural corruption may coexist with good economic success, history shows that corruption is detrimental to growth (World Bank, (n.d.).). It was used to analyze which country can be the best candidate in shortlisting for receiving aid while checking the corruption control part of it. |
| Social Contribution | Real or apportioned payments to social insurance programs are made to provide for the payment of social |

| | insurance benefits. It was used to analyze the relation between the public and government regarding taxation and insurance stuff which helps in the development of different programs in a country. |
|---|---|

The variable *"Internet users"* was used to filter the countries which were developing economies and had the growing number internet users with the % population of a specific country, which includes the developing world countries and are not the ones and are fully developed like USA,UK or China etc.

The variables *Rule of Law, Political Stability, Government Effectiveness, Regulatory Quality, Corruption Control* and *Social Contribution* are all together used to understand the Good governance factor and how well the countries are dealing with their laws while handling different foreign investments which includes funds and aids from organizations like *IMF* or *World Bank*.

By using the tool R, first *214 countries* were filtered by using the above indicators and the selection criteria was to select the countries that are either on the track on developing or are developing with a good governance and are having *40-50%* of internet users by population in order to get the countries which are on their path to develop technological advancements. The countries were filtered by using the R code that uses the values *between $1^{st}$ and $3^{rd}$ Quartiles* removing all the other countries and the data collected for these variables of countries was from 2010-2017.

Figure 1. Summary of Data which helped in filtering

```
> summary(df)
   iso2c              country              year        IT.NET.USER.ZS      CC.EST             GE.EST
Length:2184       Length:2184       Min.    :2010   Min.   : 0.00    Min.   :-1.8264   Min.   :-2.4837
Class :character  Class :character  1st Qu.:2012   1st Qu.:16.95    1st Qu.:-0.7407   1st Qu.:-0.7259
Mode  :character  Mode  :character  Median :2014   Median :40.88    Median :-0.2469   Median :-0.1062
                                    Mean    :2014   Mean   :43.01    Mean   : 0.0000   Mean   : 0.0000
                                    3rd Qu.:2015   3rd Qu.:67.59    3rd Qu.: 0.7465   3rd Qu.: 0.8332
                                    Max.    :2017   Max.   :99.55    Max.   : 2.4049   Max.   : 2.2414
                                                    NA's   :190     NA's   :501       NA's   :502

    PV.EST             RQ.EST             RL.EST         GC.REV.SOCL.ZS
Min.   :-3.1310   Min.   :-2.5296   Min.   :-2.4234   Min.   : 0.000
1st Qu.:-0.6348   1st Qu.:-0.7442   1st Qu.:-0.7758   1st Qu.: 0.000
Median : 0.0928   Median :-0.1053   Median :-0.1569   Median : 3.797
Mean   : 0.0000   Mean   : 0.0000   Mean   : 0.0000   Mean   :12.318
3rd Qu.: 0.8694   3rd Qu.: 0.7504   3rd Qu.: 0.7847   3rd Qu.:26.511
Max.   : 1.9651   Max.   : 2.2605   Max.   : 2.1003   Max.   :55.544
NA's   :492       NA's   :502       NA's   :494       NA's   :857
```

Figure 2. 38 Countries shown highlighted by red label after filtering (*using MapChart*)



Analytic Techniques performed were.

- Clustering (Descriptive Analytic Technique)

Clustering using K-means algorithm was applied on the countries which were already filtered by using all the 7 variables by using the code in R. The clustering was done between the Government effectiveness and Political Stability, in order to get the clusters which can help in providing us with countries which are Politically stable, and they are having Good governing body of countries using their data of year 2010-2017 (shown in Appendix). The optimal number of clusters were identified with Elbow-method (shown in Appendix).

- ARIMA Forecasting (Predictive Analytic Technique)

After the Clustering, ARIMA was applied on the final 10 countries which were outcome of the shortlisted countries with the best Governance helping our selection criteria to be specific and bringing up the best countries that are having control of corruption, showing that effective behavior will be seen  towards aids and foreign investments. The forecast was done on the data of "Internet users" and the data was taken from year 1998-2017.

The Internet users % by population of a country forecasted for next 5 years was then taken into a mean value and then combined with the having the  mean of data 2010-2017 to calculate the final value output by *weighted mean* of all these values and then the countries were listed in ascending order with the maximum value at the top. The ranking made was primarily based on 40% part of

mean of 5 years forecast value of  Internet users % by population for a country because for the digital development there must be awareness of online platforms and internet usage to people that's why it took major part in weighted mean, while the remaining 60% weightage was divided between the (government effectiveness, control of corruption and political stability) each having 20% weightage with the mean of their data from year 2017 for a country. These variables played 60% part because this was important to know how effectively the financial aid can be managed by a government of a country.

## 4. Analysis and Results

In this part of the paper the complete analysis and results are shown. The analysis was done in following steps.

Figure 3. Steps taken to Analyze the data for ranking 10 countries
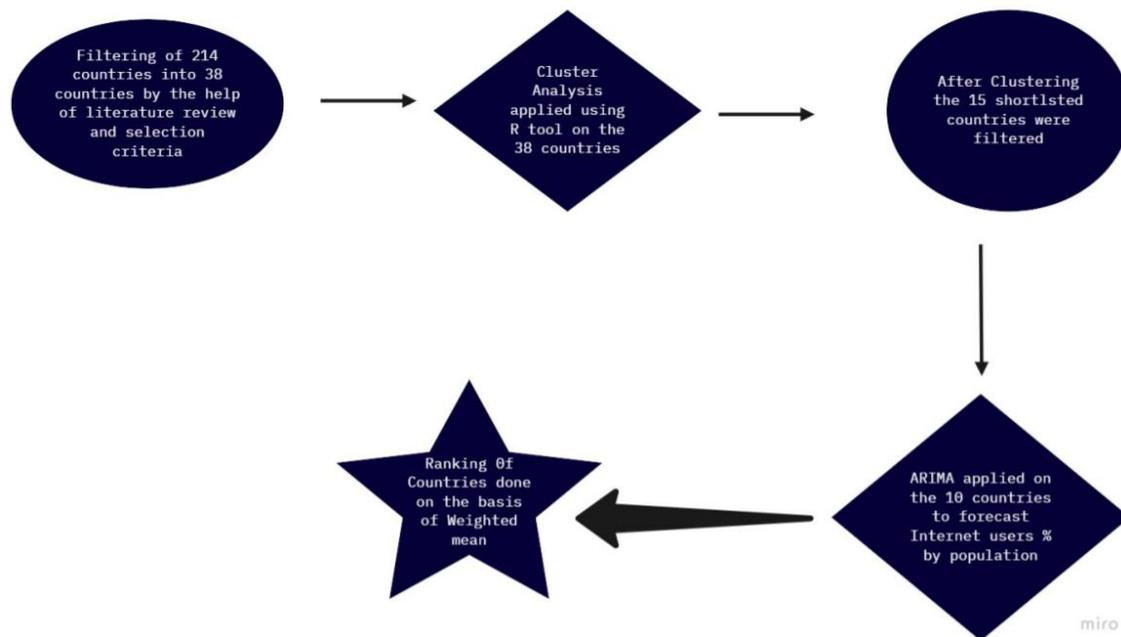


## 4.1. Filtering of Countries from 214 to 38

In the first part, the 214 countries selected were, shortlisted by using variable *"Internet users"*, *Rule of Law, Political Stability, Government Effectiveness, Regulatory Quality, Corruption*

*Control* and *Social Contribution.* Internet users was used as a variable to understand the awareness of technology and its usage in a country by understanding percentage of Internet users to its population. R tool was used in this case to filter out countries. The variables regarding the good governance were used by reviewing the literature by reading different articles and understanding the impact of it on the economic and technological development of a country. gg-plots of variables in relation to 38 countries have been shown in Appendix (3-4)

Figure 4. gg-plot of Internet users% by Population to 38 Countries
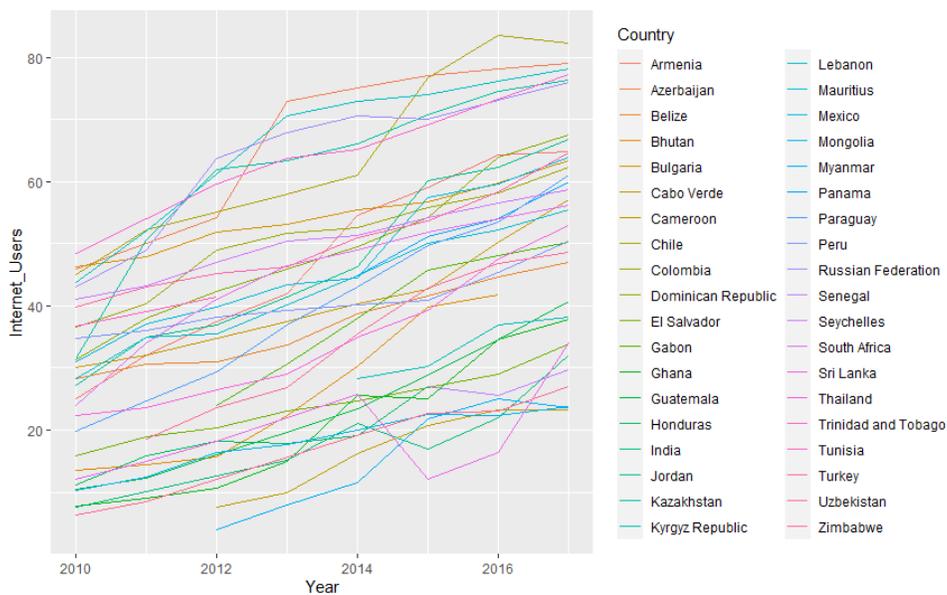


## 4.2. Clustering Performed

In the second part, clustering was performed over the 38 countries. Before the clustering was performed, the variables to be selected for clustering their correlation was observed. The variables for clustering used were *Government effectiveness* and *Political stability.* These variables were used to observe such countries which already had a 40-50% of Internet users by population and are on their track to development and how these countries managed different funds and aid from foreign organizations over the years to select the best ones which are politically stable and have an effective government to handle aid in a best way. Clustering performed were of 2,3,4 and 5 clusters but the number of clusters to be selected were identified by Elbow method which is (shown in Appendix 7-8). On that basis, k=2 clusters were selected. Out of those two clusters one was

having high government effectiveness and political stability which constituted of 15 countries. Clustering performed has been shown below with k=2 and the remaining in Appendix (5-6).

Figure 5. Cluster plot shown of 38 countries while using k=2
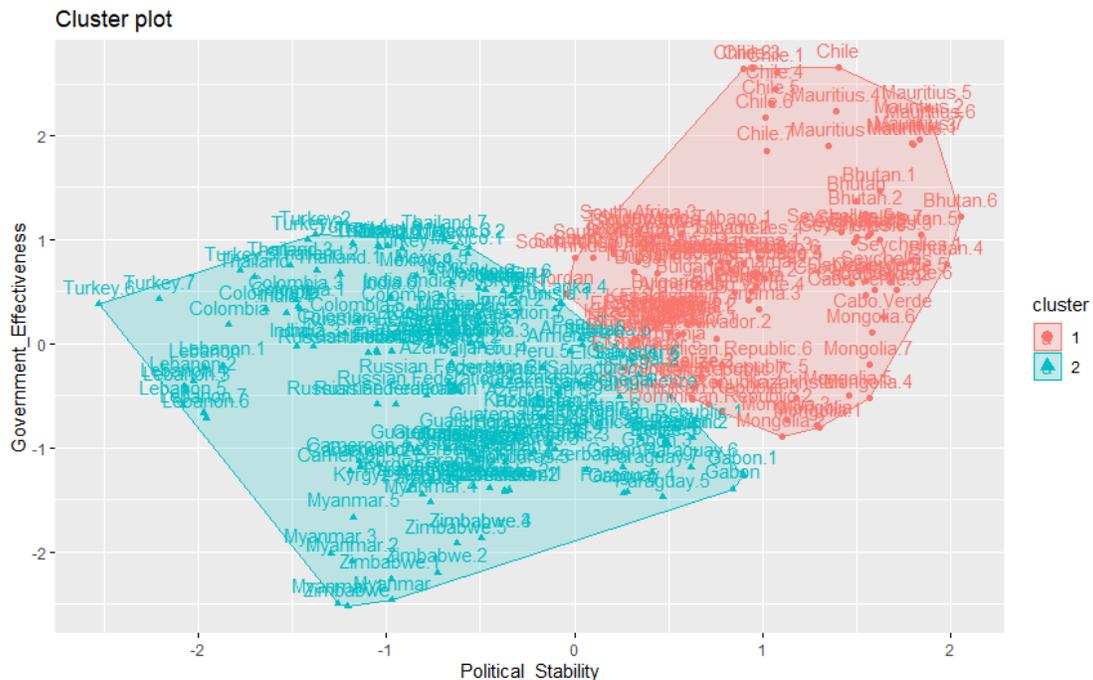


## 4.3. Filtering on the Clustered Data for selection of 10 Countries

After the selection of the *Cluster 1,* we received 15 countries which then were further filtered using the "Internet Users" variable. Range of Internet users' values was used in R-studio to remove the values (of Countries which constituted values less than 12.4999 and values greater than 67.57125). the reason setting for this range was, in case of range higher than 67.57125 showing the countries which were developed economies and are already having a lot development programs and they were not considered eligible while lower range showed there had been no awareness among people of technology, internet and online platforms. The gg-plots of the shortlisted 10 countries has been (shown in Appendix 21).

## 4.4. ARIMA Forecasting of Internet Users

ARIMA forecasting was applied on Internet Users % by population of the shortlisted 10 countries which proved to have an effective government system which can handle foreign aids and

investments in the best way for the technological development of their country. This forecasting was done to observe the future population increase or decrease of Internet users in these countries to use that data to rank the countries. The data used for forecasting was from 1998-2017 because maximum the samples are better and accurate the future predicted values are. The forecasted values of countries have been (shown in Appendix 9-20)

Figure 6. Final 10 countries highlighted on world map (*using MapChart*)
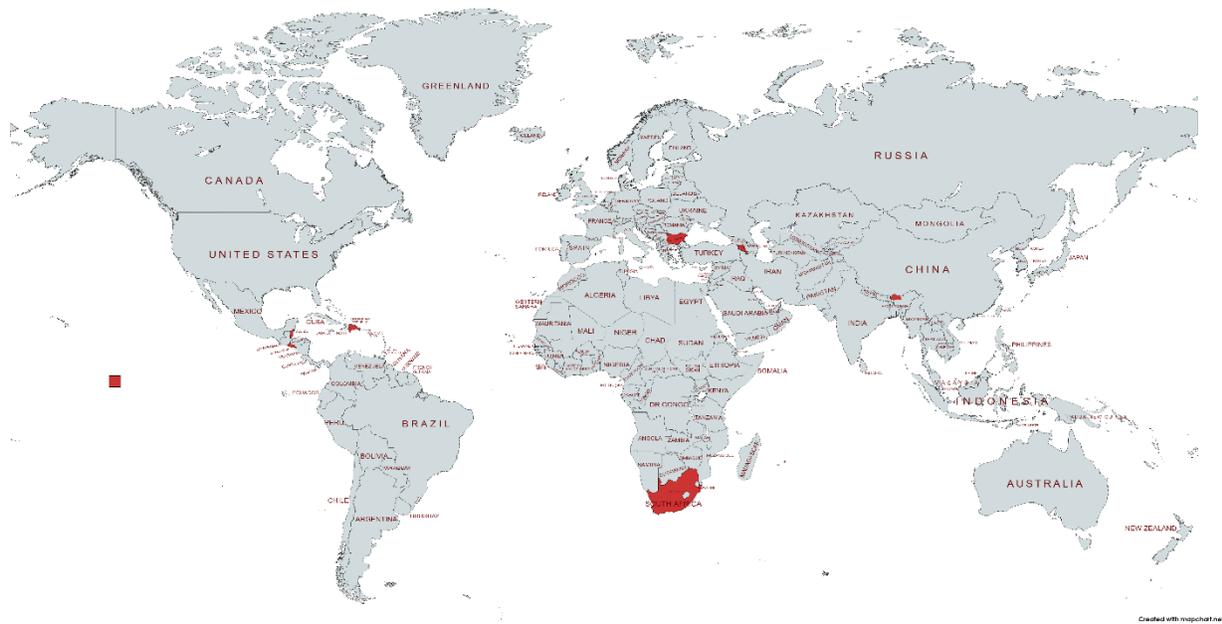


These were the final 10 countries selected for ranking.

## 4.5. Ranking of 10 countries by Weighted Mean

In the last step, countries were ranked based on the weighted mean of 5 years forecasting values of Internet users % by population and mean of the values of government effectiveness, control of corruption and political stability of data 2017. This combination was set to understand and analyze the countries which were trying their best to innovate and advance towards technological development with the help of Government and its policies which can act as an effective tool for managing financial aid.

The table below shows the Internet users mean of 5 years forecasting values and the values of government effectiveness, control of corruption and political stability of data 2017 which are converted into ==range of 0-100 to apply weighted mean==.

Table 2. Final 10 countries data for Ranking

| Country | Internet Users % by Population (Mean of 5 YEARS forecasted values) | Corruption Control (2017) % | Government Effectiveness (2017) % | Political Stability (2017) % |
|---|---|---|---|---|
| Armenia | 73.48 | 29.9 | 47.3 | 49.36 |
| Bulgaria | 73.08 | 39.3 | 55.5 | 70 |
| Bhutan | 61.82 | 69.7 | 60.9 | 88.8 |
| Belize | 53.88 | 36.7 | 35.4 | 64.2 |
| Cabo Verde | 77.64 | 62.8 | 53.4 | 80.95 |
| Dominican Republic | 83.4 | 25.2 | 41.4 | 67.0 |
| El Salvador | 44.9 | 31.1 | 41.4 | 62.7 |
| Seychelles | 67.64 | 59.2 | 59 | 81.3 |
| Mauritius | 63.7 | 47.9 | 69.8 | 85.4 |
| South Africa | 63.58 | 42.6 | 56.1 | 57.1 |

Table 3. Countries ranked according to weighted mean of the variables selected

| Country by Rank | Weighted mean (= Internet users *0.4 + CC*0.2 + GE*0.2 + PS*0.2) | Rank |
|---|---|---|
| Cabo Verde | 70.486 | 1 |
| Bhutan | 68.608 | 2 |
| Seychelles | 66.956 | 3 |
| Mauritius | 66.1 | 4 |
| Bulgaria | 62.192 | 5 |
| Dominican Republic | 60.08 | 6 |
| South Africa | 56.592 | 7 |
| Armenia | 54.704 | 8 |
| Belize | 48.812 | 9 |
| El Salvador | 45 | 10 |

## 5. Conclusion and Limitations

After the following steps of analytical techniques and tools, 10 countries have been concluded and finalized as the most suitable candidates for Digital Development Initiative Aid of World bank because of their eagerness towards Technological development observed by the Internet Users indicator which helped in analyzing the awareness of people among online platforms and gadgets

and the effectiveness of Governance shown by these countries while handling all external and internal affairs and leading their country towards Innovative development by help of Foreign funds and investments by organizations like IMF, World Bank etc.

Clustering helped us shortlisting to 15 countries on the basis of Good governance which indirectly leads to better economy and made it possible to filter it while ARIMA forecasting made our analysis more accurate and specific by cutting down the countries which were involved in both these areas(Technological development or awareness and Effective Government and its policies) in an effective manner.

In parallel to analysis, number of limitations occurred in the path which included missing data of most the countries in terms of years and the variables which have been used. During the filtration of data according to our selection criteria some data lost resulted in the removal of countries which could go ahead in this analysis. For forecasting the minimum samples used for it are 50-100 but due to limitation of years data, just had to pick 20 samples meaning 20 years data from 1998-2017.

# 6. References

Data Flair. (n.d.). *Clustering in R.* Available at: https://data-flair.training/blogs/clustering-in-r-tutorial/. . [Accessed 27th March 2021].

Heckman, E. (2016). *Fitting an ARIMA model.* Available at: https://blog.minitab.com/en/starting-out-with-statistical-software/fitting-an-arima-model. [Accessed 25th March 2021].

Hussain, Z. (2014). *Can political stability hurt economic growth?* Available at: https://blogs.worldbank.org/endpovertyinsouthasia/can-political-stability-hurt-economic-growth. [Accessed 24th March 2021].

Kareiva, P., Chang, A. and Marvier, M. (2008). Development and conservation goals in World Bank projects. *Science*, *321*(5896), pp.1638-1639.

Manyika, J. and Roxburgh, C. (2011). *The great transformer: The impact of the Internet on economic growth and prosperity.* Available at: https://www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/the-great-transformer#. [Accessed 22 March 2021]

Santiso, C. (2001). Good governance and aid effectiveness: The World Bank and conditionality. *The Georgetown public policy review*, *7*(1), pp.1-22.

SOAS. (n.d.). *The role of projects in the new aid framework*. Available at: https://www.soas.ac.uk/cedep-demos/000_P534_PPM_K3637-Demo/unit1/page_15.htm. [Accessed 21st March 2021].

Sundaram, J. (2015). *Does good governance always boost development?* Available at: https://www.weforum.org/agenda/2015/06/does-good-governance-always-boost-development/. [Accessed 25 March 2021]

United Nations (n.d.). *The Impact of Digital Technologies.* Available at: https://www.un.org/en/un75/impact-digital-technologies. [Accessed 20 March 2021]

University of Cincinnati (n.d.). K-Means Cluster Analysis [Computer Program]. Available at: http://uc-r.github.io/kmeans_clustering. [Accessed 19th March 2021].

University of Essex (2021) *BE 277 Business analytics for managers and entrepreneurs.* [Computer program]. Available at: https://moodle.essex.ac.uk/pluginfile.php/1589710/mod_resource/content/0/wk5example.R (Downloaded: 24th March 2021).

World Bank (2016). *Digital Development Partnership (DDP).* Available at: https://www.worldbank.org/en/programs/digital-development-partnership. [Accessed 22 March 2021]

World Bank IDA. (n.d.). *Borrowing Countries.* Available at: http://ida.worldbank.org/about/borrowing-countries. [Accessed 22nd March 2021].

World Bank. (n.d.). *Indicators.* Available at: https://data.worldbank.org/indicator

World Bank. (n.d.). *Combating Corruption.* Available at: https://www.worldbank.org/en/topic/governance/brief/anti-corruption. [Accessed 26th March 2021].

**List of Appendices**

1. Scholars, politicians, public rights organizations, and aid supporters are increasingly agreeing that governance is necessary for the development and, as a result, economic development. This opinion has become more evident because of several longitudinal studies conducted in the last decade that have shown the powerful beneficial benefits of effective governance (World Bank, 2008).

2. A country's relative poverty, defined as GNI per capita below a set threshold and regularly updated ($1,185 in fiscal year 2021), determines its eligibility for IDA assistance. IDA also assists countries that are over the operating threshold but lack the credit ratings needed to lend from the Foreign Bank for Restructuring and Advancement such as some small island economies (IBRD) (World Bank IDA, (n.d.)).

3. Literature review of ARIMA, ACF and PACF

 ARIMA stands for autoregressive integrated moving average, and it's a predictive analysis method that incorporates time series data to better explain the data set or forecast future patterns.

The letters ACF and PACF stand for Autocorrelation Function and Partial Autocorrelation Function, respectively. The association between results of a time series split by k time units is known as autocorrelation. Partial autocorrelations, on the other hand, calculate the magnitude of the relationship when other terms are factored into the equation, in this case the intermediate lags in the model (Heckman, 2016).

4. Auto Arima

Even though ARIMA is a successful technique for predicting time series data, the data processing and variable tuning processes take a long time. Until you can use ARIMA, you must first make the series stationary and use the plots we mentioned earlier to calculate the values of p and q. Since it removes steps, Auto ARIMA makes this job very easy for us. Auto ARIMA considers the AIC and BIC values provided when identifying the appropriate variable configuration. The estimators AIC and BIC are used to evaluate models. The model is stronger if these values are lower (Heckman, 2016).