

# Church Population Growth Prediction Using Predictive Analytic (Linear Regression Model) Technique

\*Ayeni J.A. Makinde O.E. Ipeyeda F.W.  
Department of Computer Sciences,  
Ajayi Crowther University, Oyo, Nigeria

## Abstract

The world today has become highly technological and Information and Communication Technology (ICT) has been identified as one of the basic pillars on which the modern society stands. In Nigeria today, competency in basic ICT skills is now regarded as one of the basic requirements for our day-day activities in religion, economic, financial and political. The use of available ICT tools equally enhances the validity of results and assist management in planning and budgeting for growth. The church cannot be an exception in this regard as “spreading the gospel” is the goal of the churches by growing the number of new members using available tools. The issue of church growth transcends many areas of church activities and the application of the ICT infrastructures could be gainfully put into use in some of these areas for greater impact on church development. The difficulties confronting church organisations in respect of church growth are becoming unsurmountable for those that have refused to embrace and deploy Information and Communication Technologies to enhance growth and its resultant complexities. In this paper, a system capable of predicting church growth with the application of Linear Regression Model (LRM) using the historic data of a growing church is presented. The results demonstrated the importance of predicting the population of organisations such as religious centers that often have a growing teeming population of members. Such results of prediction using ICT tools would be an essential guide in the budgeting requirements of such organisations.

**Keywords:** Prediction, growth, tools, ICT, normalize, linear, budgeting

**DOI:** 10.7176/CEIS/12-1-05

**Publication date:** January 31<sup>st</sup> 2021

## 1. General Overview

The world today has become highly technological and Information and Communication Technology (ICT) has been identified as one of the basic pillars on which the modern society stands, (Ayeni, Akinyemi and Lawal, 2020). In Nigeria today, competency in basic ICT skills is now regarded “as part of core of education alongside, reading, writing and numeracy” (UNESCO, 2015). There has been an evolution of ICT since mid-20<sup>th</sup> century and it was a proven phenomenon by 2004 and has also been embraced as the engine room of modern development of the 21<sup>st</sup> century. Some of the areas of human activities greatly impacted by ICT include social, political, educational and economical communication. These areas of ICT impact operate mainly on information, and churches are not excluded from this community as they are involved in communicating spiritual information and education on weekly basis. Aluko (2014), observed the extensiveness of the revolution of ICT and posited that the probability of it spreading and impacting all areas of the earth was very high. The issue of church growth transcends many areas of church activities and the application of the ICT infrastructures could be gainfully put into use in some of these areas for greater impact on church development. Aluko (2014) opines that the “challenge of Church growth is rapidly becoming a difficult matter to church organizations who are not ready to deploy Information and Communication Technologies in order to enhance their growth and expand their circle of influence to cover the teeming wired generation who are consistently living, and can hardly stay one hour away from the use of ICT gadgets”, (Aluko, 2014).

The use of ICT tools and the application of the Technology to improving the church growth and development have been the major focus of most church organizations the world over and Nigeria could not be left out. In view of the phenomenal of growth in some Pentecostal Churches in terms of membership, there is an urgent need to employ ICT in order to be able to cope with the corresponding needs of the Church in terms of budgeting and provisions of essential ICT tools for its growth.

## 2. Scope of the Work

The scope of this work is limited to the application of ICT tools for predicting The Light of Christ Church – LCC (Members population growth) Ishaga, Lagos, Nigeria, as a case study. The paper addresses the issues related to the application of Machine Learning - predictive analytics using Linear Regression Model, to predicting church population growth.

## 3. Research Type

In his theses titled “An analysis of dissertations on church growth”, Virginia (2009), classified Research types

differently according to viewpoints and cited examples of divisions of research types as follows: “Quantitative and Qualitative Researches,” “Basic and Applied Researches,” “Experimental and Non-Experimental Researches,” or “Philosophical, Historical, Descriptive, and Experimental Researches” (Virginia, 2009). Morris (2018), defined research as “the systematic investigation into and study of materials and sources in order to establish facts and reach new conclusions” and “concluded that nearly all of us have engaged in research at some time” (Morris, 2018). Therefore, it will not be out of context to conclude that ICT and research can be gainfully applied to church growth related issues.

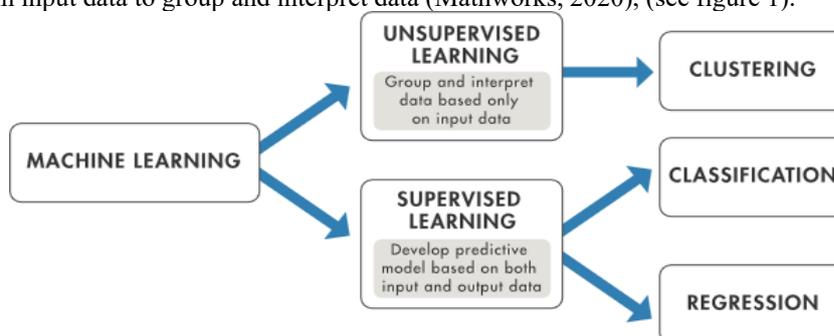
#### 4. Review of Related Works

Most of the work on Church Growth has been qualitative and very few on quantitative research because of the complexity of religion itself. However, “while qualitative research could add to knowledge about these hypotheses, such as provide a richer understanding of the mechanisms of growth of a religious movement, they cannot by themselves test whether the growth is taking place” (Storm, 2012). The Venerable Bob Jackson was the first researcher to study the application of statistical data in the area of church growth (Norris, 2012). According to Norris (2012), findings Bob’s work such as “that growth in attendance seems lower in team ministries and that there can be significant decline during vacancies are interesting (Norris, 2012). In his book titled “Hope for the Church”, Bob (2002), took a challenging look at recent church attendance figures and sets out to interpret their message for today’s Church and highlighted areas where the Church was growing, analyzed causes of decline and outlines strategic responses to the problem of declining numbers (Bob, 2002).

David de Wetter et al. (2010) in their final report on the research conducted by Towers Watson for the United Methodist Church (UMC) in the USA in 2010 - a particularly interesting and thorough study, carried out a large number of interviews, focus groups and surveys across the denomination in order to develop hypotheses on what drives church vitality, calculating a vitality index for over 32,000 churches and considered three sub-factors of vitality namely attendance, growth and engagement (David de Wetter et al., 2010). In the final report, David de Wetter et al. (2010) found that: all kinds of UMC churches showed high vitality – small, large, across different geographies, and church settings. The four drivers of vitality are fairly consistent across different churches namely (David de Wetter et al., 2010): effective lay leadership, the existence and number of small groups and programs (including for Children and Youth), a mix of traditional and contemporary styles of worship, the Pastor exhibiting excellence in key attributes and the length of his or her appointment (David de Wetter et al., 2010).

#### 5. Literature Review

According to Mathworks (2020), “Machine learning is a data analytics technique that teaches computers to do what comes naturally to humans and animals: learn from experience” (Mathworks, 2020; Weng *et al.*, 2017), and further concluded that “Machine learning algorithms use computational methods to ‘learn’ information directly from data without relying on a predetermined equation as a model” (Mathworks, 2020). These learning algorithms are often designed to identify patterns in data that are used to make critical decisions and predictions; areas of application include but not limited to stock trading, medical diagnosis, customer behavior, energy etc. in most of the literature, emphasis is being laid on two types of Machine Learning algorithms; the supervised learning and unsupervised learning. In the former, the model or prototype is trained on known input and output data to predict future outputs (Weng *et al.*, 2017). In the latter technique, the algorithm attempts to discover hidden patterns or intrinsic structures in input data to group and interpret data (Mathworks, 2020), (see figure 1).



**Figure 1. Machine learning techniques include both unsupervised and supervised learning**

(Source: mathworks, 2020)

The Supervised Machine Learning develops the algorithm that builds a model to make predictions on evidence input in the presence of uncertainty (Mathworks, 2020; Weng *et al.*, 2017). Again, as elicited in mathworks (mathworks, 2020), a “supervised learning algorithm takes a known set of input data and known responses to the data (output) and trains a model to generate reasonable predictions for the response to new data” (Mathworks, 2020; Weng *et al.*, 2017). The models developed by supervised machine learning are based on one

of these two techniques; classification or regression techniques. The former predicts discreet responses and classifies input data into categories while regression techniques predict continuous responses such as processing data range, temperature, power load and fluctuations in power demand (Mathworks, 2020). The unsupervised learning technique attempts to discover hidden patterns or intrinsic structures in input data to group and interpret such data (Weng *et al.*, 2020). Datasets consisting of input data without labeled responses are used to draw inferences, (Mathworks, 2020; Weng *et al.*, 2017). The most common unsupervised learning technique is Clustering often used for exploratory data analysis to find hidden patterns or groupings in data, (Mathworks, 2020; Weng *et al.*, 2017).

### 5.1 Predictive Analytics

Predictive analytics is the process of using data analytics to make predictions based on data and the process uses data along with analysis, statistics, and machine learning techniques to create a predictive model for forecasting future events (Mathworks and Researchswan, 2020). In Mathworks (2020) and Researchswan, (2020) the term predictive analytics is defined “as the application of a statistical or machine learning technique to create a quantitative prediction about the future” (Mathworks, 2020; Weng *et al.*, 2017). Frequently, supervised machine learning techniques are used to predict a future value or estimate a probability (Mathworks, 2020).

### 5.2 Predictive Analysis Using Linear Regression

Today’s ICT applications have proven the importance of Artificial Intelligence (AI) as the engine room of development. Various professionals have successfully applied AI to make their tasks a lot much easier. The Application of AI has gained considerable influence in the social science, Finance, administration, scientific and engineering fields. It is now being used to predict future market prices, stock exchange, medical fields to determine ailment and even carry out surgical operations, meteorologists for weather forecasting, Human Resources to determine if an applicant meets the basic requirements for a job and the self-driven car. Ahmed (2020) and Rohith (2017) concluded that the driving force behind the ubiquitous use of AI is machine learning algorithms. Seif (2018) defined “Regression is a technique used to model and analyze the relationships between variables and often times how they contribute and are related to producing a particular outcome together” (Seif, 2008). Consequently, Linear regression could be described as a regression model that is composed of linear variables. A regression technique is determined by the number of independent variables and the type of relationship existing between the independent and dependent variables (Ahmed, 2020; Rohith, 2017, Bhis and Saini, 2019).

Therefore, linear regression is a type of regression analysis where the number of independent variables is one and there is a linear relationship between the independent (x) and dependent (y) variable (Attia, 2020; Rohith, 2017, Bhis and Saini, 2019).

According to Ognjanovski (2018), “linear simply means a straight line and given multiple (x,y) coordinates like on the graph (Figure 3), then the straight line would represent the linear regression equation” (Ognjanovski, 2018). Actually it is the straight line that best fits (make the best prediction) for the given data.

This straight line can be mathematically expressed by the following (equation 1):

$$h_{\theta}(x) = \theta_0 + \theta_1 x \quad (\text{eqn.1})$$

Where  $h$  is called *hypothesis* and  $\theta$  is called *theta*. This equation is same as:

$$y = a + bx \quad (\text{eqn. 2})$$

Ognjanovski (2018) listed the steps involved in the development process of a Linear Regression Model (LRM) as follows:

- i. Data Collection and Preparation – Use the applicable data collection technique for data acquisition – in Excel CSV format.
- ii. Visualize Data (Understand the Data) – Load the raw data into memory using applicable software method (Ognjanovski, 2018).
- iii. Feature(s) Normalization - This step is only required if there is the need to use the Gradient Descent algorithm. Before one is able to compute the theta values there is the need to normalize the features so it will optimize the calculations. That is done by subtracting the average value and divide by standard deviation for all the features. Normalization is often employed to standardize the range (values) of independent variables or features of data (Ognjanovski, 2018). (see Feature Normalization formula – eqn. 3)

$$x' = \frac{x - \bar{x}}{\sigma} \quad (\text{eqn. 3})$$

- a. Create Linear Regression Model - The computation of the theta values for the equation that best fits the data visualized in the previous step using one of the stated methods below (Ognjanovski, 2018):
- b. Computing Normal Equation (one-step algorithm).

c. Using Gradient Descent algorithm

**6. Research Approach**

For the purpose of this research, the Light of Christ Church – LCC (local Church) will form the basis of Case Study for the project because of the ease of access to historical data and other useful information considered appropriate for the research. The schematic representation of the research approach used in this work is presented in (Figure 2) and consists basically of the following processes:

- i. Harvesting historical data in respect of the Church growth – Light of Christ Church (Blessed Sanctuary 2014 - 2019)
- ii. Collation of data for the purpose of analysis and mining
- iii. Develop Predictive Model of the membership Growth Using Linear Regression (Machine Learning)
- iv. Apply the Predictive Model to predict the membership population till the year 2027.
- v. Present the outcomes to assist decision makers for the purpose of branch, members, staff and infrastructure development.

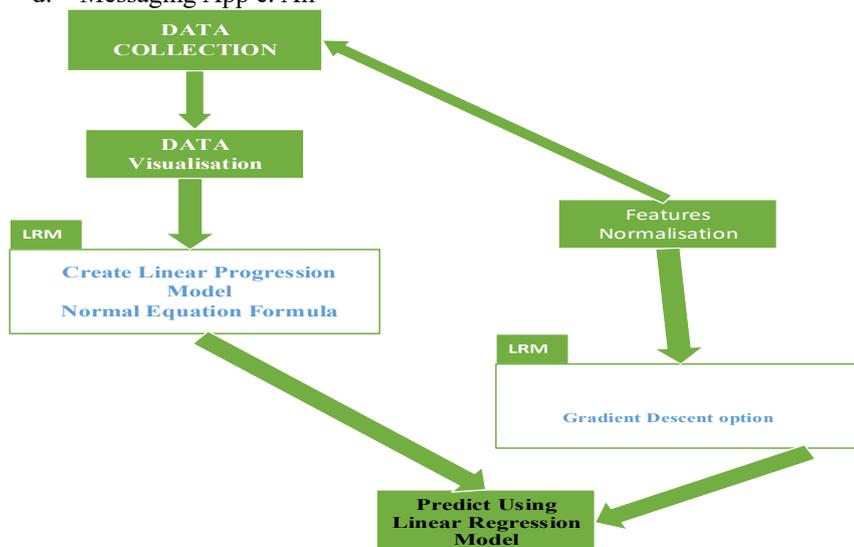
*6.1 Research Assumptions*

For the purpose of this research, it is assumed that the application of modern ICT tools such as sermon and church activities deliveries through the use of projector screen, SMS base text messaging, WhatsApp, evangelism and feedbacks from members could have been responsible and may not have affected the outcomes of the predictive modelling.

*6.2 Data Collection – Questionnaire*

As the starting point in the development process, data form was developed and used for the acquisition of the following data:

- i. Data Form 001 - Light of Christ Church
- ii. Name of Church: (LCC – Blessed Sanctuary – (Year)
- iii. Church Population/Membership (Year)
- iv. Size of Auditorium(Year)
- v. Number of Pastors/Staff (Year)
- vi. Size of Car Parks (Year)
- vii. Residential Qtrs. - Officials/Members (Year)
- viii. Have ICT tools been introduced for the purpose of evangelism in the year of study (yes/No)?
- ix. Which of the following ICT Tools has been introduced in (viii) above?
  - a. email
  - b. WhatsApp group
  - c. Tele-Evangelism
  - d. Messaging App e. All



**Figure 2: Architectural representation of the LRM**

### 6.3 System Tools Used

The process was realized on a Laptop System with running Windows 8.1, Intel Celeron Processor CPU 560 @ 2.13 GHz. 4GB Ram, 500GB HDD. 32 Bit Operating System. Software Tools are MS Excel 2015, GNU Octave (compatible with Matlab)

### 6.4 Data Collection – Outcome

**Table 1: Data Collation and Categorization**

#	Year	Members Population	Male	Female	Children
1	2014	106	45	52	9
2	2015	140	60	61	19
3	2016	160	68	75	17
4	2017	185	70	75	40
5	2018	370	120	150	100
6	2019	600	190	210	200
7	2020				
8	2021				
9	2022				

## 7. Data Aggregation

For the purpose of our work only the first two columns will be used in the prediction analysis process (year, membership strength) and stored in a csv file of Excel format (Table 2).

**Table 2: Data Collected and aggregated in csv format (Chc Pop\_2014-2019.csv)**

Year	Population
2014	105
2015	140
2016	160
2017	185
2018	370
2019	600

### 7.1 Data Visualization

The data visualization process involves loading the .csv file using the function for loading the csv data is *getCsvData.m* (Ognjanovski, 2018) file using the filename (LCCdata.csv) as the only parameter. The first is the year and the second column is the membership for that particular year.

### 7.2 Create Linear Regression Model – Using Normal Equation Formula

The analytical solution for Linear Regression computation was used which include multiplication of matrices and computation of the value of theta.

$$\Theta = (X^T X)^{-1} X^T y \quad (\text{eqn.4})$$

### 7.3 Predict using the Linear Regression Model (LRM)

Having computed the theta values for the normal equation, the membership prediction for some of the next years at the LCC blessed sanctuary was computed till 2020. Using the theta values that was derived using the **Normal Equation**, the expected membership strength was computed by multiplying with the theta values as illustrated using this code:

$$pvalue = pyear\_val * (\text{theta}) \quad (\text{eqn. 5})$$

## 8. System Implementation

The membership data from 2014 to 2019 was loaded into file LCCdat.csv and the following steps were taken:

- i. Data Loading into memory
- ii. Data Visualization (see Figure 2)
- iii. Computation of the normal LRM equation (see Figure 3)
- iv. the output was plotted

### 8.1 Predicting Church Population for year 2020 using LRM – Normal equation

Predict\_year = 2020

PyVal = Pred\_year\_normalized \* theta

$$\begin{aligned}
 &= [1 \ 2020] \cdot [-183241.34717] \\
 &= (91 \cdot 2020) - 183241.34717 \\
 &= 578.65283 = 579
 \end{aligned}$$

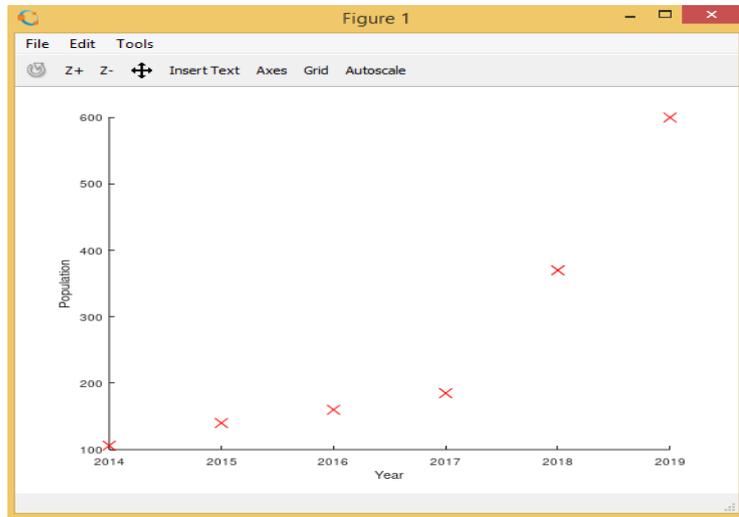


Figure 3: Visualization of the captured data using Octave tools

Table 3: Output from Octave workspace

Name	Class	Dimension	Value
X	double	7x2	[1, 2014; 1, 2015...
m	double	1x1	6
pred_value	double	1x1	578.67
pred_year	double	1x1	2020
pred_year_val	double	1x2	[1, 2020]
statistics_data	cell	6x2	...
theta	double	2x1	[-183241.34717; ...
y	double	6x1	[106; 140; 160; 1...

$$\begin{aligned}
 &= \text{Predyearval} = (x)\text{Pred\_year\_normalize} * (\Theta)\text{theta} \\
 &y = x * \Theta
 \end{aligned}$$

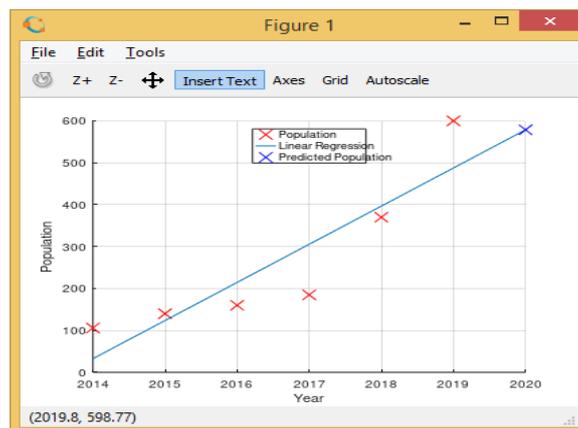
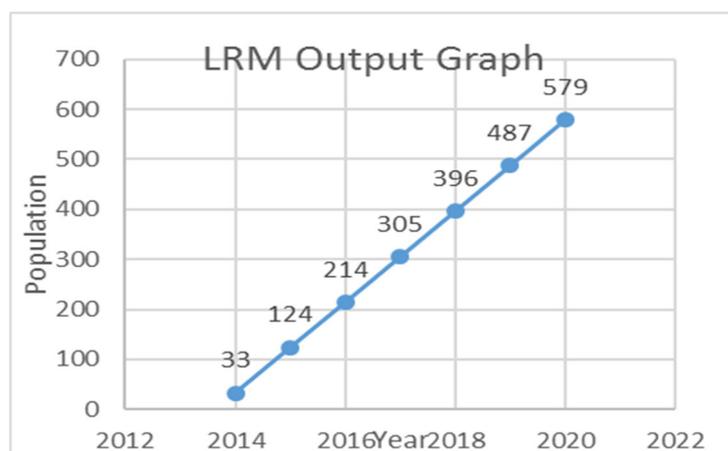


Figure 4: LRM computation output using Octave tools

**Table 4: Showing prediction for year 2020**

Year	$y = x * \Theta$
2014	33
2015	124
2016	215
2017	306
2018	397
2019	488
2020	579



**Figure 5: Representation of the LRM using Normal Equation (Plotted in Excel)**

## 9. Conclusion

The system for the predictive analysis of the membership strength as experimented on the LCC Blessed Sanctuary membership historic data using Linear Regression Model (Normal Equation for LRM) demonstrated an upward trend in its futuristic data. The results as shown in figure 4 indicate the accuracy of the linear regression algorithm. Although the population understudy is relatively small, it is important to state here that the size has no effect on the results of the prediction. The prediction for subsequent years could also be made with the same approach and the model could be applied to other areas and even highly populated churches for planning purposes and budgeting.

## 10. Acknowledgement

We freely acknowledge the contributions of the postgraduate group of the department of Computer Sciences for their contributions and inputs towards the realization of this paper. Dr. Esolomo Oluwadare, Dr. Binuyo Kayode of the Departments of Physics and Mathematics respectively are equally recognized for their invaluable contributions on topical issues of the project.

## References

- UNESCO (2015). ICT in Education UNESCO Regional Office for Eastern Africa <http://www.unesco.org/new/en/nairobi/communication-and-information/ict-in-education/#topPage>.
- Ayo, Aluko (2014). Deployment of ICT for Church Growth, URL:<https://ayoolaluko.blogspot.com/2014/07/deployment-of-ict-for-churchgrowth.html>
- Weng SF, Reys J, Kai J, Garibaldi JM, Qureshi N (2017) Can machine-learning improve cardiovascular risk prediction using routine clinical data? PLOS ONE 12(4): e0174944. <https://doi.org/10.1371/journal.pone.0174944> Predicting cardiovascular
- Virginia, L. (2009). An analysis of dissertations on church growth, Liberty Baptist Seminary, <https://digitalcommons.liberty.edu/cgi/viewcontent.cgi?article=1175&context=doctoral> 2009.
- Morris, M. (2018). Principles of church growth: research, <https://ministry.acst.com/principles-church-growth-research/2018>.
- Ingrid, Storm (2012). Researching religion using quantitative method, Institute for Social Change, University of Manchester, <https://www.kent.ac.uk/religionmethods/documents/Researching%20religion%20using%20quantitative%20data.pdf>
- Ayeni, J. A., Akinyemi, O. and Lawal O.O. (2020). Application of ICT to Enhancing Basic Education System in Nigeria. International Journal of Research and Innovation in Social Science (IJRISS) |Volume IV, Issue II, February 2020|ISSN 2454-6186

- Christian A. Schwarz* (1966): Natural Church Development: A Guide to Eight Essential Qualities of Healthy Churches  
Sep 1, 1996.  
[https://www.amazon.com/Christian...Schwarz//B001K83BVQ%3Fref=db\\_s\\_a\\_mng\\_r...](https://www.amazon.com/Christian...Schwarz//B001K83BVQ%3Fref=db_s_a_mng_r...)
- Kevin Norris. (2012). Church Growth Research Programme, Note by Kevin Norris, Senior Strategy Officer (Analyst) (Resource) Strategy and Development Unit, Resourcing Mission Bulletin, 2012
- Bob, Jackson (2002). Hope for the Church. **Church** House Publ. ISBN: 9780715155516. Published:  
<https://www.chpublishing.co.uk/books/9780715155516/hope-for-the-church>
- David de Wetter, Ilene G., Rich L. and Rick S. (2010). UMC Call to Action: Vital Congregations Research Project Findings Report for Steering Team June 28, 2010.<http://s3.amazonaws.com/Website Properties /news-media/docu- ments/umc-call-to-action-vital-congregations-research-project .PDF>
- Borislav Savkovic. (2017). Predictive Analytics.<https://www.mathworks.com/discovery/predictive-analytics.html#how-it-works>.
- Mathworks (2020). “What Is Machine Learning? 3 things you need to know”.  
<https://www.mathworks.com/discovery/machine-learning.html>
- Mathworks and Researchswan (2020). Predictive Analytics with MATLAB.  
<https://researchswan.com/whitepaper/predictive-analytics-with-matlab>
- Rohith G. (2017). Introduction to Machine Learning Algorithms: Linear Regression.  
<https://towardsdatascience.com /introduction-to-machine-learning-algorithms-linear-regression-c4e325882a 2017>.
- Attia, A.(2020) Machine Supervised Learning: Regress in Python 3 and Math.  
<https://www.udemy.com/course/machine-supervised-learning-regression/>
- Ognjanovski G. (2018). Predict Population Growth Using Linear Regression — Machine Learning Easy and Fun.  
<https://medium.com/analyti-vidhya/predict-population-growth-using-linear-regression-machine-learning d555b1ff8f38>
- Bhis, S. and Saini, P. (2019). Linear Regression. <https://www.ques10.com/p/40524/linear-regression-1/>
- Seif, G. (2018). 5 Types of Regression and their properties. URL:<https://towardsdatascience.com /5-types-of-regression-and-their-properti es-c5e1fa12d55e>