International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 5 Issue 04, April-2016

Developing a Pseudo-Linear Regression Framework for Predicting Volume of Work of Construction Companies in Qatar

Karim Abdel Warith,
PhD, Post-Doctoral Associate
School of Civil Engineering
Qatar University
Doha, PO Box: 2713

Sherif Attallah,
PhD, Assistant Professor
Department of Technology
Ball State University
Muncie, IN, 47306

Mohamed Mansour,
PhD Student
School of Civil Engineering
Purdue University
West Lafayette, IN, 47907

Amr Kandil,
Associate Professor
Division Construction Engineering and Management
School of Civil Engineering
Purdue University
West Lafayette, IN, 47907

Abstract- **Understanding the current trends in emerging construction market is of great importance to various stakeholders of this critical industry including investors, developers, designers, contractors, suppliers and researchers. Many variables affect these trends and the optimum approach to forecast work volume trends will be through appropriate modeling of the most significant variables. However, modeling of continuous variables, where standard least squares linear regression is inadequate, has been a challenge to many researchers in modeling similar trends. This paper proposes a methodological framework that can be used to model continuous variables affecting the work volume in emerging construction market similar to Qatar. The framework is composed of a hierarchy of models namely; linear regression, regression trees, regression scales and quantile regression. Construction company data from Qatar was collected through interviews with key decision makers in the companies and used to develop and demonstrate the framework's logic. The framework integrates several statistical modeling techniques that are applicable and relevant to the research objective in order to optimize the validity of the variables' significance and insure a solid-based forecast of future trends. The framework was used to predict the volume of work of companies given some characteristics of the company. The framework identified a number of significant factors including; attitude towards risk, aggressive price competition and targeting publicity. This research represents a new utilization of statistical analysis for emerging construction market through determining the significant factors affecting the market share and volume of work for construction companies. The developed framework represents a new tool for predicting trends in this critical industry and can be utilized by interested researchers in different areas of the world or even different industries.**

Keywords-Volume Of Work; Emerging Markets; Regression Analysis; Linear Regression; Non-Linear Regression; Statistical Analysis; Qatar.

I. INTRODUCTION

Modeling continuous variables is very common in the field of construction management. Many researchers have relied on least square regression to model continuous variables [1][2][3] [4][5][6][7][8][9]. However least square regression has many restrictive conditions, including normality, multicolinearity and linearity [10][11]. Other researchers turned to transformations, rejecting collinear variables and removing outliers [12]. In cases where non-linearity was expected, some researchers have estimated non-linearity using Regression trees [13][14][15][16], Regression scales [17] and Quantile regression [18][19].

However, no publication has been targeted to systematically determine how to proceed from one model attempt to the next. In addition, in the authors' knowledge, Regression scales and Quantile regression have not been applied in the area of construction management. This paper presents a framework that provides a systematic approach to test least squares regression and propose solution in case least squares regression does not provide the desired solution or the data set violates some of the required assumptions. The framework was used to analyze and predict the future volume of work to be undertaken by construction companies in emerging markets, with Qatar as an example. The framework determined the most significant variables, which influence the future volume of work.

A. Problem Statement

A general problem faced by many analysts is that of modeling data when the nature of relationships is unknown. In many instances engineers are faced by the task of modeling data when little or no information about the nature of model is available [20][21][22]. Many researchers attempt linear regression as a first step [6][7][8][9]. The work herein describes a statistical framework that can be followed in case linear regression does not provide acceptable results.

Another problem being addressed by this paper is that of strategy and characteristics that companies need to expand in emerging markets. Other researchers have explored models for contractors' financial failure [23], financial performance of construction companies [24] and contractor's markup behavior [25]. Kim et al. (2013) [26] have studied the use of real options in situations where contractors which to enter a new market. The approach presented herein has a similar application but focuses on the company characteristics rather than the tools to be used, such as real options. Companies wish to understand their chances in emerging markets more accurately. This will allow them to make more informed decisions on where and when to pursue certain markets. Or what elements they need to develop before expanding beyond their current markets.

B. Objective

The objective of this paper is to develop a statistical framework capable of handling continuous data in the absence of predefined relationships between the predicted variable and predictor variables. A secondary objective is to apply this framework to determine which characteristics affect the volume of work that a given company can attain by pursuing work in an emerging market.

II. MODELING METHODOLOGY

The methodology introduced by the authors is meant to be used in cases where the dependent variable is continuous and little information is known about the factors affecting it. FIGURE 1 illustrates the proposed methodology.

The framework utilizes 4 main models; Ordinary Least Squared Linear Regression, Regression Scales, Regression Trees and Quantile Regression; through a series of steps.

A. Step 1 Apply OLS

The first step in every linear modeling exercise is to attempt an Ordinary Least Squares (OLS) linear regression model [25][6][8][9].

OLS is a technique used to estimate coefficients in in a linear regression model. The method relies on minimizing the sum of squared vertical distances between the observed responses the predicted responses [27].

The OLS is based on a number of assumptions [27]:

1. Relation is linear in nature (covered in Step1)
2. All variables should be exogenous i.e are independent of all response values.
3. No or minor multicollinearity
4. No heteroskedasticty

a. Step 1a Check Linearity

Linear regression techniques operate under the assumption that the underlying function is inherently linear. In other words, the nature of the relation between dependent and independent variables should be linear. Unfortunately there is no procedure to determine whether the relationship is inherently linear. If there is any doubt, from prior knowledge on the subject, that the relation is linear, the simplest solution is to perform a transformation on the data.

b. Step 1b Apply Transformation if needed

There is no strict technique to be followed to transform data points. Typical transformations include ln, log, exponential, and raising data to any power. OLS (Step 2) is then applied to the resulting data set. If the results are unsatisfactory, Regression Scales (Step 4) could be attempted. It is important to note that while applying transformations is presented as part of the first step in this framework, it is possible to utilize it at any point in the framework. Any of the models presented herein can be applied on transformed data.

c. Step 1c Apply OLS

It is important to note that even with the relatively stringent assumptions associated with OLS, OLS is still very useful. OLS could help determine which factors are more significant, even in the absence of a true linear relationship. Also, OLS can help in identifying the general trends between the dependent and independent variables.

B. Step 2 Model Diagnostics

a. Step 2a Remove Non- Significant Variables

The t-statistic is a ratio of the difference between the result of a coefficient and its standard error [28]. A Student t Test is performed on every variable in the equation. Variables that have a t statistic with low confidence (i.e. less than 90%) should be removed.

b. Step 2b Remove Endogenous Variables

If a variable is suspected to be endogenous (not exogenous) it should be removed from the model. There is no strict test to determine whether a variable is exogenous or not. This determination should in inferred by logic.

c. Step 2c Check for Mutlicolinearity

Multicollinearity (also known as collinearity) is a situation where two or more predictor variables are highly correlated, i.e. one can be linearly predicted from the other [29]. This causes errors in the prediction of model coefficients. The Variable Inflation Factor test (VIF) was used to determine which variables are collinear [29].

According to Menard (1995) [30] a VIF value more than 10 is a clear indication of multicolinearity, a value from 4 to 10 is cause for concern. To be conservative, the authors chose a value of 4 as a cutoff.

If multicolinearity exists, the variables causing colinearity should be removed from the model. VIF will give an indication to which variable is causing the co-linearity. However, once the variable is removed the model should be tested again to insure colinearity no longer exists or if other variables need to be removed.

d. Step 2d Check for Heteroskedasticity

Heteroskedasticity is a statistical phenomenon that occurs when the errors of within a model are inconsistent [31]. In other words the magnitude of the error varies depending in the value of the independent variables. The authors tested for heteroskedasticity using the Studentized Breusch-Pagan test (BP) [32] and the Non-constant Variance Score Test (NCV) [31]. In case of heteroskedasticty, transformations can be employed to solve the issue.

e. Step 2e Check Goodness of Fit

There are a number of techniques to check the goodness of fit of the model. The authors recommend using the coefficient of determination, R squared, as it is well suited to standard regression techniques. R squared is a value that indicates how well data fit a statistical model, i.e. how far off the model is from the actual data. An Rsq generally ranges from 0 to 1. Rsq of 1 generally means the model is perfectly fit with the data.

C. Step 3 Apply Regression Scales

The second procedure that can be applied to solve the non-linearity issue, is applying a regression analysis that is based on the samples distribution. The technique was developed by Mizeria et al. (2002) [17]. Regression scales simply scales the data to various distributions and then applies standard regression. Distributions include normal distribution, exponential, gamma, beta and Huber.

D. S tep 4 Apply Regression Trees

Another solution for nonlinear estimation is regression trees [13][14][15][16]. The procedure is very simple. The data set is divided into a subset using one of independent variables. These subsets are then divided using another independent variable. In some case it could the same independent variable, but the cut off limit could be different. The procedure is similar to clustering. The code used searches through all possibilities of independent variables and their values to find the most suitable independent variable and the corresponding cut off value.

E. Step 5 Apply Quantile Regression

If Regression trees fail, quantile regression can be used to approximate nonlinear relationships [33][19]. Unlike OLS, Quantile regression targets the median not the mean of the data. It could also be used to target different quantiles. Quantile regression estimates different values for coefficients of the independent variables of each quantile. If there is a trend in these values, that trend could be estimated and then a secondary equation can be developed that captures both the change in the independent variables and that of the coefficient. While there is no clear minimum criterion for the number of data points at the terminal nodes of each branch, it is obvious that 1 point is unacceptable. Since this does not allow for data averaging in this particular node. Some researchers suggest 10% of the data [34]. It is generally not recommended to use regression trees in very small datasets [35]. Nevertheless, it is helpful in the sense of identifying the more prominent factors affecting the data

F. Step 6 Iterations

Quantile regression is considered to be the last resort in this analysis. The quantiles can be changed, however a model with an acceptable number of points in each quantile and a satisfactory goodness of fit, is not guaranteed. For this reason a counter to the number of iterations is introduced in the framework. The user may choose any maximum number of iterations; the authors do not recommend any value greater than 100. A much smaller value can be used if the exercise is not programmed.

III. FRAMEWORK APPLICATION

A. Qatar as an Emerging Market

Many construction companies are interested in expanding their activities beyond their home countries. In some cases this expansion is rewarded, in other times it is not. The outcome of the expansion relies on many factors [36][37][38][39][40]. According to Johnson and Tellis (2008) [41], the factors should be differentiated into region specific and firm specific. Furthermore, Johnson and Tellis (2008) [41] divide the firm specific factors into firm resources and firm strategy. The focus of this paper is on both types of firm specific factors. A number of investigations where performed in different countries and regions [42][43]. Other researchers have focused on consulting firms [44].

Other entities have suggested more specific factors. According to business consultant Accenture (2009) [45], there are three factors that help in the success of firms in emerging markets: being authentically local, networking the organization, and creating geographic options.

Abidalli and Harris (1995) [46] have suggested a number of indicators that are related to failure of construction companies. The most important of which are; a weak financial director, autocratic chief executive, lack of engineering skills, poor response to market changes and making losses in projects [46].

In coming years the construction industry is expected to grow rapidly in emerging economies such as the Middle East and Africa [47] in these emerging markets is expected to double to become a $6.7 trillion business by 2020 [48]. One of these emerging markets is Qatar [49]. After being granted the right to organize the FIFA world cup 2022, the volume of the construction industry in Qatar has quadrupled [50]. Spending is forecasted to be in the $100Bn range [51]. The Qatar Statistics Authority (2013) [52] reports a 14% year-on-year nominal growth in the economy. About $100bn of infrastructure projects are currently planned or announced in Qatar [53][60]. Qatar plans to invest $25bn in the development of rail lines. Other major transport infrastructure projects include: the construction of the New Doha International Airport (USD 9bn), a new port ($9bn) and the construction of new road networks ($20bn) [54]. Furthermore, Qatar is continuing to expand its power sector through the construction of more power plants, which will help to combat the power shortages in the country and is also tackling the issue of

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 5 Issue 04, April-2016

water demand through developing independent water and power plants [54]. All this has made Qatar one of the most promising new emerging markets. The volume of construction work in Qatar has attracted a large number of companies to the country and completion is rapidly increasing [53]. However, these opportunities did not result in profits for all construction companies in Qatar. In fact many of them lost considerable amounts of money and withdrew from the market completely [55]. This makes Qatar a very suitable case study for the development of the proposed models.

There is no simple answer to why some companies succeeded while others failed. Researchers are yet to introduce a unified theory of the drivers of success in emerging markets [14][56][57]. Therefore, this paper investigates some of the company characteristics that lead to successful expansions in emerging markets.

IV. DATA COLLECTION AND ANALYSIS

A. Data Collection

The total number of construction companies in Qatar is 373 [58]. The number of tier 1 companies is estimated to be about 70 companies [55]. The research team interviewed 25 construction companies. The companies included 3 developers, 12 contractors and 10 consultants, all of which are considered tier 1 companies, such as AECOM, Atkins and Bechtel, with turnovers over 500 million USD.

TABLE 1 lists the questions that are relevant to the discussion herein. The variables are referenced according to their position in the original data.

The analysis used the entire data set. However, Log/Ln models excluded 1 zero value from the calculation to avoid mathematical errors.

B. Data Preliminary Analysis

The data was collected to represent the population of tier 1 contractors/developers and consultants in Doha.

C. Apply OLS

In order to present all the aspects of the framework the authors assumed that the underlying relationship is nonlinear. The authors applied a number of transformations and are presenting the natural log transformation herein.

The log linear model shown in Table 2 produced an Rsq of 0.898 and an adjusted Rsq 0.787.

D. Model Diagnostics

a. Heteroskedasticity

The model was tested for heteroskedasticity using the Studentized Breusch-Pagan test (BP) [32] and the Non-constant Variance Score Test (NCV) [31] p value of NCV 0.896 and a BP value of 8.8.

The model attained a BP value of 8.8 (p value = 0.66) and a NCV chi-square p value =0.896. By observing the p-values of both tests, neither test suggests that there is significant heteroskedasticity. Thus no correction is needed for heteroskedasticity.

b. Multicolinearity

The Variable Inflation Factor test (VIF) was used to determine which variables are collinear. As seen in Table 2, none of the values for VIF exceeded 4. Thus multicolinearity is unlikely to exist.

E. Applying Regression Scales

The distribution of the population was unknown. The general assumption was that the population is normally distributed, since there was no information to indicate otherwise. A Q-Q plot was used to test if the population is in fact normally distributed. FIGURE 2 shows the quantile distribution with a confidence interval of 90%.

It is clear that there is a significant deviation from the normal distribution. The points at the high end of the quantiles are beyond the extreme limits of the distribution. Also, the profile of the points is curved, instead of being intertwined with the middle line as would be expected from normally distributed data. A number of distributions were applied using regression scales, including normal distribution, exponential, gamma, beta and Huber. The Huber Distribution gave the best results. Table 3 illustrates the results of the analysis. The Regression Scale model produced an Rsq of 0.946, Adjusted Rsq of 0.891 and Log Likelihood of 463.5.

F. Model Diagnostics

The significance of the independent variables is very high, which is a good indication of the robustness of the model. The Rsq is also high indicating a good fit. The only drawback is that there are too many independent variables, which may be a problem in terms of model interpretation. This will be discussed in the model interpretation section.

G. Applying Regression Trees

Incase both options to resolve the normality had failed; the next step in the framework is to apply regression trees. Other researchers have applied a similar technique to construction industry data to determine job performance and satisfaction [14].

Two types of data were used; data without transformations and data with transformations. Figure 4 shows the use of data with no transformation, while Figure 5 shows the use of log transformed data. The three trees in Figure 4 show three attempts each with a higher number of points in the terminal branches. Figure 4 a, b & c show a minimum of 1,2 & 3 points in the terminal node.

H. Model Diagnostics

a. Terminal nodes

As mentioned earlier there is no clear minimum criterion for the number of data points at the terminal nodes of each branch. However, 1 point at the terminal node is unacceptable. Thus the model with 1 observation in the lightest branch is for illustration only. The models with 2 and more points at their terminal nodes are acceptable, if the 10% minimum rule is followed.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 5 Issue 04, April-2016

b. Goodness of Fit

Table 4 shows the Rsq and Adjusted Rsq for all the models previously listed.

The model with terminal nodes having 1 observation (suffix a) attained the best fit, which is expected. However, as explained earlier these models should not be used since they will not be robust in other cases. Within the rest of the models; the natural log models attained good results. The best model, in this subset, in the author's opinion is the RegTree1b, since it received no data transformation, yet the Rsq is within an acceptable range.

I. Quantile Regression

The common trait between all of the previous models is that the models assume there is a linear relationship between the dependent and independent variable. This assumption is most likely untrue, but given the infinite possibilities of nonlinear relationships, it is impossible to guess the true relationship with out more information. The linearity assumption is a good starting point.

In this analysis quantile regression was used to estimate the regression coefficients. Different quantiles were attempted. Table 5 shows the results of the set of quantiles that produced the best fit.

J. Model Diagnostics

Observing the profile of each coefficient as it changes with each quantile does not show any distinct trend in the coefficients. This reinforces the assumption that the relationship between the dependent variable and independent variables is nonlinear.

The significance, indicated by the t value, is very high for all coefficients. The Rsq is 0.846, which indicates a good fit. However, Rsq may be misleading in this case. The sample size is also too small to be dealt with in such dissection. Another issue is that there are too many independent variables, which may be a problem in terms of model interpretation. This will be discussed in the model interpretation section.

K. Model Interpretation

While model correctness and goodness of fit are important from the statistical perspective, it is also important to be able to understand the model. The previous analysis focused on the data modeling from an empirical perspective without tackling the logical merit of the model. This section serves as an example of how the results of the framework can be interpreted. The section assumes all models have been attempted.

Interpretation of any given model relies on explaining the relationship between the dependent and independent variables. Table 6 identifies the typical trend between the dependent and independent variables in the Qatari market data. The relationships are noted as P (positive or directly proportional) or N (negative or inversely proportional).

The Value of work is inversely proportional with respect to risk attitude (V6). When the risk attitude is aggressive the volume of work handled by the company is less. At first this may seem counter intuitive, but when this piece of information is combined with the knowledge that the Value of work increases when the company has unique features (V8), the bigger picture becomes clearer. It seems that in this emerging market, like most emerging markets, there is a shortage of expertise, thus companies that retain certain expertise are more likely to get more work. Also, Value of works handled by consultants seems to be higher than that of contractors (V25) and more work is handled by foreign companies (V26) this supplements the lack of expertise argument. Aggressive risk taking may be a result of a company's lack of unique features, in which case it is the result of the lower value of works gained rather than the cause.

The volume of work is inversely proportional to the number of years spent in an emerging market. This relation is counter intuitive and highly suspect of interactions with other variables. In case the statistical relation is in fact mimicking the true relation, a reasonable explanation would suggest that the number of years is endogenous to another variable, or coincidences that companies who have spent lesser time in Qatar gained lesser volumes of work. It could be that companies do not move into the construction market until they have secured a project. Then they remain in the market hoping to get more projects. Thus companies who have stayed longer in Qatar may have completed their original project and have lower volume compared to companies who have just entered the market with a secure project. Another explanation for this counter intuitive relation is unidentified colinearity with another variable. Upon closer examination it was observed that attempting to use new practices (V14) is always present in equations where Years in Qatar (V27) is present. There may be some interaction between these two variables and the Volume of work dependent. The Volume of work is proportional to attempting new practices. Thus may be the fact that new practices are appealing to this market and that new companies entering the region may try to market themselves as having newer practices, is the reason for the inverse relation between years in Qatar and volume of work.

Lowering prices (V7) seems to be consistent with gaining more work, as would be expected in any market. Furthermore, pursuing higher profit (V11) or attempting to lead the market (V12) seems to result in less overall work, while pursuing publicity (V13) and targeting emerging markets (V15) seem to have a positive effect on the overall volume of work gained. This seems to be logical, in an emerging market it is not yet known who is better in certain aspects, with publicity and proper advertising companies may gain certain types of work much quicker than others, even if in reality they are not better than the competition. The actual performance of the company could be much worse than it seems. This is clear when considering that; providing high quality services (V9), participating in training programs (V10), communicating corporate strategy (V36), monitoring the competition (V37), being aware of weaknesses (V38) and having a strong quality management system (V39), which can be considered signs of being a 'good' company, are also inversely proportional to the volume of work received by the company. This can be attributed to the fact that most of the work if offered by the government. In the case of emerging markets, governments may be inexperienced in determining which companies are better.

The size and resources of the firm also impact the value of works handled. The larger the size (V28) and the greater the resources (V35), the greater the value of works handled by the firm. This is expected, as larger firms are more capable in handling larger contracts.

Another factor the boosts gaining work in emerging markets seems to be attempting to use new practices (V14, V16). This may be specific to the Qatari market, as the country boasts in leading the market in a number of fields. New innovations and practices are encouraged in Qatar in all fields [59][52].

V. CONCLUSION

It is important to note that the objective of the investigation was to develop a framework capable of modeling continuous variables. This has been performed and tested on a data set from construction companies in Qatar. The framework proved useful and applicable.

For this specific data set the resulting model is OLS after natural log transformation. The transformation was introduced to account for the normality of the data. The resulting model produced an Rsq of 0.898. The remaining steps of the framework were demonstrated on the data for illustration purposes.

Although, the models themselves are unlikely to be transferable, the model form and the significant factor are expected to be transferrable to other emerging markets. Additionally, the modeling methodology developed can be easily applied to other markets and other industry sectors and helps simplify similar problems. However, it needs further testing on other types of problems to insure its generalizability.

The overall limitation of the analysis can be summarized in the sample size. While the sample size with respect to the entire population may be acceptable, the actual number of points is too small for some of the analysis techniques. Therefore, future work by the authors includes getting more responses and testing the modeling framework on other problems.

A secondary objective was to identify the relevant factors that impact the volume of work to be received by a construction company in an emerging market. The model interpretation section summarized all these factors. One of the most important factors is publicity. Proper publicity seems to be the key in attaining more work by a given firm. Other important factors are the size of the firm, the number of years present and offering unique services.

ACKNOWLEDGEMENTS

The authors would like to thank Dr. Josh Mills for his valuable support. The authors would like to express their gratitude to the R Project open source community for their continued contribution to the R software. This study is supported by Qatar Foundation NPRP Award No.:09-717-2-274. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the writers and do not necessarily reflect the views of Qatar Foundation.

REFERENCES

[1] Park, P. Y., Miranda-Moreno, L. F., & Saccomanno, F. F. (2010). Estimation of speed differentials on rural highways using hierarchical linear regression models . Canadian Journal Of Civil Engineering, 37(4), 624-637.

[2] Sonmez, R. (2004). Conceptual cost estimation of building projects with regression analysis and neural networks. *Canadian Journal Of Civil Engineering, 31*(4), 677-683.

[3] Flood, I., & Issa, R. (2010). Empirical Modeling Methodologies for Construction. *Journal of Consturction Engineering and Management, 136*, 36-48.

[4] Lovegrove, G. R., & Sayed, T. (2006). Macro-level collision prediction models for evaluating neighbourhood traffic safety. *Canadian Journal Of Civil Engineering, 33*(5), 609-621.

[5] Duffy, G., Oberlender, G., & Seok Jeong, D. (2011). Linear Scheduling Model with Varying Production Rates . *Journal of Contruction Engineering and Management, 137*(8), 574–582.

[6] Serrano, M. A., Calero, C., Sahraoui, H. A., & Piattini, M. (2007). Empirical Studies to Assess the Understandability of Data Warehouse Schemas Using Structural Metrics. *Software Qual J, 16*, 79-106.

[7] Pasupathy, R. K., Ivan, J. N., & Ossenbruggen, P. J. (2000). *Single and Multi-Vehicle Crash Prediction Models for Two-Lane Roadways.* USDOT.

[8] Miller, J. C., Eddy, D. R., & Fischer, J. (2004). *he Sensitivity and Specificity of Oculometrics Under Fatigue Stress Compared to Performance and Subjective Measures.* US Air Force Research Laboratory.

[9] March, S. T., & Smith, G. F. (1995). Design and natural science research on information technology. *Decision Support Systems, 15*, 251-266.

[10] J.B.Ramsey. (1969). Tests for Specification Errors in Clasical Linear Least-Squares Regression Analysis. *Journal of the Royal Statistical Society. Series B (Methodological), 31*(2), 350-371.

[11] Anselin, L., & Rey, S. (2010). Properties of Tests for Spatial Dependence in Linear Regression Models. *Geographical Analysis, 23*(2).

[12] Intriligator, M. D., Bodkin, R. G., & Hsiao, C. (1978). *Econometric Models, Techniques and Application.* Prentice-Hall.

[13] Jong-Suk, J., Owusu-Antwi, E. B., & Ji-Hwan, A. (2006). Analytical procedures for evaluating factors that affect joint faulting for jointed plain concrete pavements using the Long-Term Pavement Performance database. *Canadian Journal Of Civil Engineering, 33*(10), 1279-1286.

[14] Leung, M., Chen, D., & Yu, J. (2008). Demystifying Moderate Variables of the Interrelationships among Affective Commitment, Job Performance, and Job Satisfaction of Construction Professionals . *Journal of Construction Engineering and Management, 134*(12), 963–971.

[15] Segal, M. R. (1988). Regression Trees for Censored Data. *Biometric, 44*(1), 35-47.

[16] Loh, W.-Y. (2002). Regression Trees With Unbiased Variable Selection And Interaction Detection. *Statistica Sinica, 12*, 361-386.

[17] Mizera, I., & Muller, C. (2002). Breakdown points of Cauchy Regression-Scale Estimators. *Statistics & probability letters, 57*(1), 79-89.

[18] Koenker, R. (2005). *Quantile Regression* (Vol. 38). Cambridge University Press.

[19] Xiao, Z. (2009). Quantile Cointegrating Regression. *Journal of Econometrics , 150*(2), 248-260.

[20] Smith, R. P., & Morrow, J. A. (1999). Product Development Process Modeling. *Design Studies, 20*(3), 237-261.

[21] Gainsburg, J. (2006). The Mathematical Modeling of Structural Engineers. *Mathematical Thinking and Learning, 8*(1), 3-36.

[22] Wielinga, B., Schreiber, A., & Breuker, J. (1992). KADS: A Modelling Approach to Knowledge Engineering. *Knowledge Acquisition, 4*(1), 5-53.

[23] Russell, J., & Zhai, H. (1996). Predicting Contractor Failure Using Stochastic Dynamics of Economic and Financial Variables. *122*(2), 183–191.

[24] Chen, H. (2009). Model for Predicting Financial Performance of Development and Construction Corporations. *Journal of Construction Engineering and Management, 135*(11), 1190–1200.

[25] Oo, B., Drew, D., & Lo, H. (2010). Modeling the Heterogeneity in Contractors' Mark-Up Behavior. *Journal of Construction Engineering and Management, 136*(7), 720–729.

[26] Kim, D., Ashuri, B., & Han, S. (2013). Financial Valuation of Investments in International Construction Markets: Real-Options Approach for Market-Entry Decisions. *Journal of Managment Engineering, 29*(4), 355–368.

[27] Pedroni, P. (2001). Fully modified OLS for heterogeneous cointegrated panels. *Advances in econometrics, 15*, 93-130.

[28] Bailey, B. J. (1977). Tables of the Bonferroni t statistic. *Journal of the American Statistical Association, 72*(358), 469-478.

[29] Mansfield, E. R., & Helms, B. P. (1982). Detecting Multicollinearity. *The American Statistician, 36*(3), 158-160.

[30] Menard, S. (1995). Applied Logistic Regression Analysis. *Sage University Series on Applications in the Social Sciences*.

[31] Koenker, R. (1981). A Note on Studentizing a Test for Heteroscedasticity. *Journal of Econometrics, 17*, 107-112.

[32] Breusch, T., & Pagan, A. (1979). A Simple Test for Heteroscedasticity and Random Coefficient Variation. *Econometrica, 47*, 1287–1294.

[33] Koenker, R., & Hallock, K. F. (2001). Quantile Regression. *Journal of Economic Perspectives, 15*(4), 143-156.

[34] Wilkinson, L. (2004). Classification and regression trees. In *Systat 11* (pp. 35-56).

[35] Torgo, L. F. (1999). *Inductive Learning of Tree-Based Regression Models.* Departamento de Ciência de Computadores.

[36] Hawk, D. (2006). Conditions of success: a platform for international construction development. *Construction Management and Economics, 24*(7), 735-742.

[37] Betts, M., & Ofori, G. (1992). Strategic planning for competitive advantage in construction. *Construction Management and Economics, 10*, 511-532.

[38] Ng, S. T., Tang, Z., & Palaneeswaran, E. (2009). Factors contributing to the success of equipment-intensive subcontractors in construction. *International Journal of Project Management, 27*(7), 736-744.

[39] Ghapanchi, A. H., Wohlin, C., & Aurum, A. (2014). Resources contributing to gaining competitive advantage for open source software projects: An application of resource-based theory. *International Journal of Project Management, 32*(1), 139-152.

[40] Wang, S. Q. (2004). Risk management framework for construction projects in developing countries. *Construction Management and Economics, 22*(3), 237-252.

[41] Johnson, J., & Tellis, G. J. (2008). Drivers of Success for Market Entry into China and India. *Journal of Marketing, 72*, 1-13.

[42] Raftery, J., Pasadilla, B., Chiang, Y., & Hui, E. C. (1998). Globalization and construction industry development: implications of recent developments in the construction sector in Asia. *Construction Management & Economics, 16*(6), 729-737.

[43] Öz, Ö. (2001). Sources of competitive advantage of Turkish construction companies in international markets . *Construction Management & Economics, 19*(2), 135-144.

[44] Tang, L. C., Atkinson, B., & Zou, R. R. (2012). An entropy-based SWOT evaluation process of critical success factors for international market entry: a case study of a medium-sized consulting company. *Construction Management and Economics , 30*(10), 821-834.

[45] Accenture. (2009). *Expansion into Africa:Challenges and Success Factors Revealed.* Accenture.

[46] Abidali, A. F., & Harris, F. (1995). A methodology for predicting company failure in the construction industry. *Construction management and economics, 13*(3), 189-196.

[47] Accenture. (2012). *Achieving High Performance in The Construction Industry.* Accenture.

[48] Global Construction Perspectives and Oxford Economics. (2009). *Global Construction 2020: A global forecast for the construction industry over the next decade to 2020*.

[49] KDI. (2007). Minister Lee Visited Qatar, an emerging construction market.

[50] MBT. (2009). *Rise with Qatar*. Ministry of Business and Trade.

[51] Benham, J., & Smith, M. (2010). *Factbox: Qatar's construction plan for the World Cup*. (Reuters) Retrieved June 2013, from http://www.reuters.com/article/2010/12/06/us-qatar-construction-idUSTRE6B533D20101206

[52] Qatar Foundation. (2013). *Building the Future of Education*. Retrieved from World Innovation Summit for Education: www.wise-qatar.org

[53] Thompson, R. (2011). Qatar projects companies sector face intense competition.

[54] Deloitte. (2012). GCC powers of construction 2012: Construction Sector Overview. Retrieved from http://www.deloitte.com/assets/Dcom-MiddleEast/Local%20Assets/Documents/Industries/Real%20Estate/Construction/me_real-estate_gcc_construction_ppt_12.pdf

[55] Eissa, K. (2013, June 12). Country Manager, Alghanim International Qatar. (K. A. Warith, Interviewer) Doha, Qatar.

[56] Gale, A., & Luo, J. (2004). Factors affecting construction joint ventures in China. *International Journal of Project Management, 22*(1), 33-42.

[57] Cserháti, G., & Szabó, L. (2013). The relationship between success criteria and success factors in organisational event projects. *International Journal of Project Management*.

[58] Qatcom. (2013). *Construction Companies in Doha Qatar*. Retrieved July 2013, from Qatcom.com: http://www.qatcom.com/listings/qatar-search/construction-companies

[59] Sharaa, A. (2013). *Doha Hosts the World Innovation Summit for Health.* Retrieved 2013, from Kuwait Times: http://news.kuwaittimes.net/doha-hosts-world-innovation-summit-health-2013/

[60] QSA. (2013). *Statistics Sector*. (Ministry of Development Planning and Statistics) Retrieved from http://www.qsa.gov.qa/eng/index.htm

[61] Ben-Gal, I. (2005). *Outlier detection, In: Maimon O. and Rockach L. (Eds.) Data Mining and Knowledge Discovery Handbook: A Complete Guide for Practitioners and Researchers.* Kluwer Academic Publishers, .

[62] Cullen, A. C., & Frey., H. C. (1999). *Probabilistic techniques in exposure assessment: a handbook for dealing with variability and uncertainty in models and inputs.* Springer.

[63] MEED. (2013). Qatar's construction boom gathers pace.

[64] Michael, J. R. (1983). The stabilized probability plot. *Biometrika, 70*(1), 11-17.

[65] Ng, S. T., & Tang, Z. (2010). Labour-intensive construction sub-contractors: Their critical success factors. *International Journal of Project Management, 28*(7), 732-740.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 5 Issue 04, April-2016

TABLE 1 INTERVIEW RESULTS

Variable	Question	Average Response	Std. Dev
Dependent	Value of works in the last 3 years (Million QAR)	14660.1	34923.9
V6	Attitude towards risk (0 moderate, 3 aggressive)	1.3	0.677
V7	Your firm is successful through offering lower prices than your competitors (1 = Strongly Disagree ; 2 = Disagree; 3 = neither agree nor disagree; 4 = agree; 5 = strongly agree)	2.72	0.9798
V8	Your firm is successful because its services/products have more unique features (1 = Strongly Disagree…….5 = strongly agree)	4.04	1.0198
V9	Your firm is successful because it has a reputation of high-quality services/products (1 = Strongly Disagree…….5 = strongly agree)	4.4	0.70711
V10	Your firm regularly arranges for its employees to attend and participate in training programs (1 = Strongly Disagree…….5 = strongly agree)	4.2	0.70711
V11	Your firm selects the projects it participates in based on profit/economic considerations (1 = Strongly Disagree…….5 = strongly agree)	4.08	0.95394
V12	Your firm selects the projects it participates in with the objective of leading the market or increasing market share regardless of profit (1 = Strongly Disagree…….5 = strongly agree)	3.2	0.8165
V13	Your firm selects the projects it participates in taking into consideration the publicity value (1 = Strongly Disagree…….5 = strongly agree)	2.96	1.05987
V14	Your firm is reluctant to try new practices until they are proved successful in the market (1 = Strongly Disagree…….5 = strongly agree)	2.64	1.11355
V15	Your firm targets to explore new markets (1 = Strongly Disagree…….5 = strongly agree)	4.4	0.6455
V16	Your firm successfully introduced new practices over the past 5 years (1 = Strongly Disagree……5 = strongly agree)	4.08	0.75939
Error! Not a valid result for table. Cont.			
Variable	Question	Average Response	Std. Dev
V25	Firm Type (1 if Contractor, 0 otherwise)	0.36	0.4899
V26	Home of Origin/Ownership (1 if foreign, 0 otherwise)	0.48	0.5099
V27	Years in Qatar	8.88	6.1327
V28	Number of Employees on the payroll	464.2	418.299
V35	Your firm has no shortages in financial and technological resources (1 = Strongly Disagree…….5 = strongly agree)	4.16	0.89815
V36	The strategy of the firm is well communicated to the employees (1 = Strongly Disagree…….5 = strongly agree)	4.12	0.88129
V37	Your firm monitors the competitors and is aware of their activities and practices (1 = Strongly Disagree…….5 = strongly agree)	3.88	0.66583
V38	Your firm is generally aware of its own weaknesses (1 = Strongly Disagree…….5 = strongly agree)	4.04	0.67577
V39	Your firm has a well enforced quality management system (1 = Strongly Disagree…….5 = strongly agree)	4.28	0.73711

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 5 Issue 04, April-2016

TABLE 2 LOG LINEAR MODEL

Coefficient	Estimate	Std. Error	t value	Pr(>|t|)	VIF
(Intercept)	-0.63229	2.578444	-0.245	0.809844	
V9	-1.73051	0.542592	-3.189	0.006558	2.446781
V13	0.96073	0.302426	3.177	0.006724	1.971886
V14	-0.64301	0.262412	-2.45	0.028022	1.599541
V15	1.924294	0.405517	4.745	0.000313	1.320935
V16	1.518279	0.397053	3.824	0.001861	1.820058
V25	-1.47526	0.74171	-1.989	0.066609	2.448665
V26	1.88439	0.519916	3.624	0.002761	1.34417
V27	-0.11349	0.051403	-2.208	0.044441	1.881693
V28	0.002508	0.000871	2.88	0.012116	2.521389
				Rsq	0.898
				Adj Rsq	0.787

TABLE 3 RS1 MODEL RESULTS

Coefficient	Estimate	Std. Error	t stat	Pr(>|z|)
(Intercept)	-8525.45	23013	-1.85	
V7	4364.921	1948	11.2	0
V8	13833.81	2057	33.625	0
V11	-18869.4	1844	-51.165	0
V13	12588.84	1845	34.12	0
V14	-8611.14	1594	-27.005	0
V15	46705.57	3109	75.105	0
V25	-46686	5572	-41.89	0
V26	24126.98	3142	38.39	0
V27	-6388.86	357	-89.42	0
V28	84.11	6.23	67.555	0
V35	14064.59	2138	32.89	0
V36	-14514.5	1749	-41.485	0
V37	-32888.6	2940	-55.93	0
V38	-9029.44	2970	-15.2	0
			Rsq	0.946
			Adj Rsq	0.891

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 5 Issue 04, April-2016

TABLE 4 GOODNESS OF FIT FOR REGRESSION TREE MODELS

Model	Rsq	Adj. Rsq
RegTree 1a	0.984261	0.980119
RegTree 1b	0.658193	0.609363
RegTree 1c	0.210959	0.139228
LnRegTree1	0.66626	0.618583

TABLE 5 QUANTILE REGRESSION RESULTS

Coefficient	Target									
	5% (0.05)		25% (0.25)		50% (0.5)		75% (0.75)		95% (0.95)	
	Value	t value	Value	tvalue	Value	t value	Value	t value	Value	t value
(Intercept)	-1.05E+05	-7.03E+14	-104946	-3.220	-19227	-18.290	-17606	-0.713	-27334	-3.447
V6	2.64E+04	1.48E+15	25051	6.415	13244	105.157	8877	3.000	8399	8.841
V7	1.53E+04	1.29E+15	14197	5.475	15964	190.912	10846	5.521	11722	18.585
V8	2.54E+04	2.06E+15	26213	9.752	19646	226.628	23801	11.688	24011	36.722
V11	-2.40E+04	-1.90E+15	-25484	-9.24	-27820	-312.78	-29251	-13.999	-28000	-41.736
V14	-1.58E+04	-1.69E+15	-15204	-7.468	-18647	-284.003	-13524	-8.768	-12590	-25.422
V15	5.27E+04	2.69E+15	53336	12.470	51711	374.893	47178	14.559	46925	45.102
V25	-4.04E+04	-1.39E+15	-40405	-6.361	-6575	-32.097	-16691	-3.468	-19668	-12.729
V26	3.71E+04	1.72E+15	36753	7.783	27633	181.448	24511	6.851	26007	22.64
V27	-8.20E+03	-3.85E+15	-8284	-17.790	-5220	-347.644	-5480	-15.535	-5498	-48.545
V28	1.10E+02	2.86E+15	108	12.784	74	273.253	53	8.300	56	27.127
V36	-1.40E+04	-1.29E+15	-13191	-5.551	-14852	-193.784	-4010	-2.227	-5229	-9.045
V37	-2.20E+04	-1.46E+15	-21195	-6.455	-27326	-258.046	-29299	-11.778	-27963	-35.010
									Rsq	0.846

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 5 Issue 04, April-2016

TABLE 6 TYPICAL RELATION BETWEEN VALUE OF WORK AND INDEPENDENT VARIABLES

Variable ID	Description	Typically
V6	Attitude towards risk (0 moderate, 3 aggressive)	N*
V7	Your firm is successful through offering lower prices than your competitors (1 = Strongly Disagree ; 2 = Disagree; 3 = neither agree nor disagree; 4 = agree; 5 = strongly agree)	P
V8	Your firm is successful because its services/products have more unique features (1 = Strongly Disagree…….5 = strongly agree)	P
V9	Your firm is successful because it has a reputation of high-quality services/products (1 = Strongly Disagree…….5 = strongly agree)	N
V10	Your firm regularly arranges for its employees to attend and participate in training programs (1 = Strongly Disagree…….5 = strongly agree)	N
V11	Your firm selects the projects it participates in based on profit/economic considerations (1 = Strongly Disagree…….5 = strongly agree)	N
V13	Your firm selects the projects it participates in taking into consideration the publicity value (1 = Strongly Disagree……5 = strongly agree)	P
V14	Your firm is reluctant to try new practices until they are proved successful in the market (1 = Strongly Disagree……5 = strongly agree)	N
V15	Your firm targets to explore new markets (1 = Strongly Disagree…….5 = strongly agree)	P
V16	Your firm successfully introduced new practices over the past 5 years (1 = Strongly Disagree……5 = strongly agree)	P
V25	Firm Type (1 if Contractor, 0 otherwise)	N
V26	Home of Origin/Ownership (1 if foreign, 0 otherwise)	P
V27	Years in Qatar	N
V28	Number of Employees on the payroll	P
V35	Your firm has no shortages in financial and technological resources (1 = Strongly Disagree…….5 = strongly agree)	P
V36	The strategy of the firm is well communicated to the employees (1 = Strongly Disagree…….5 = strongly agree)	N
V37	Your firm monitors the competitors and is aware of their activities and practices (1 = Strongly Disagree…….5 = strongly agree)	N
V38	Your firm is generally aware of its own weaknesses (1 = Strongly Disagree…….5 = strongly agree)	N
V39	Your firm has a well enforced quality management system (1 = Strongly Disagree…….5 = strongly agree)	N

* One exception

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 5 Issue 04, April-2016

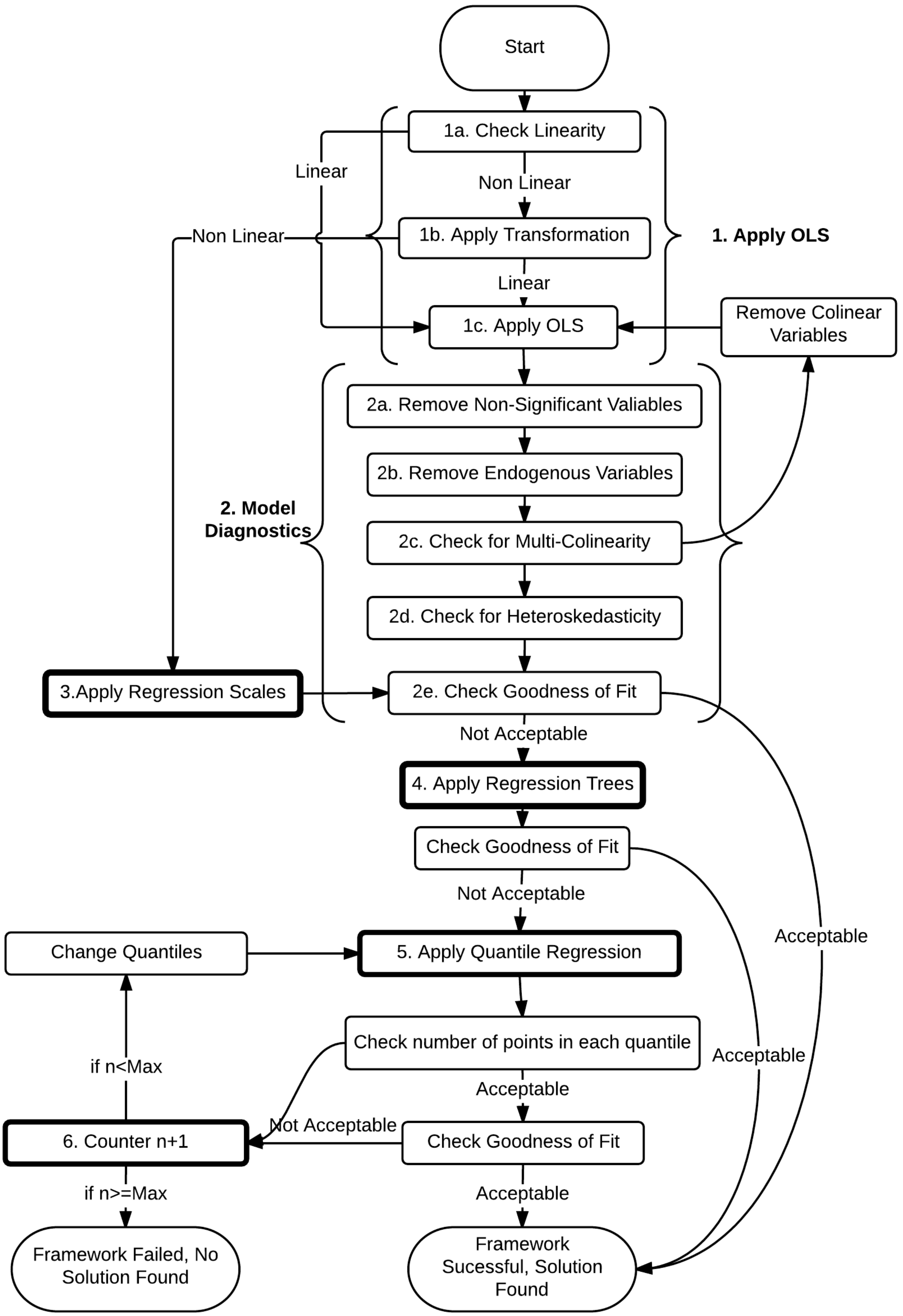


FIGURE 1 MODELING METHODOLOGY

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 5 Issue 04, April-2016

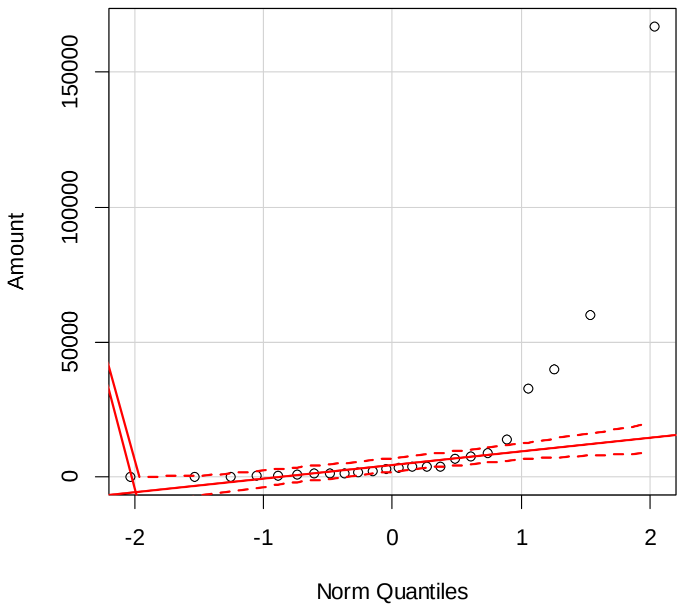
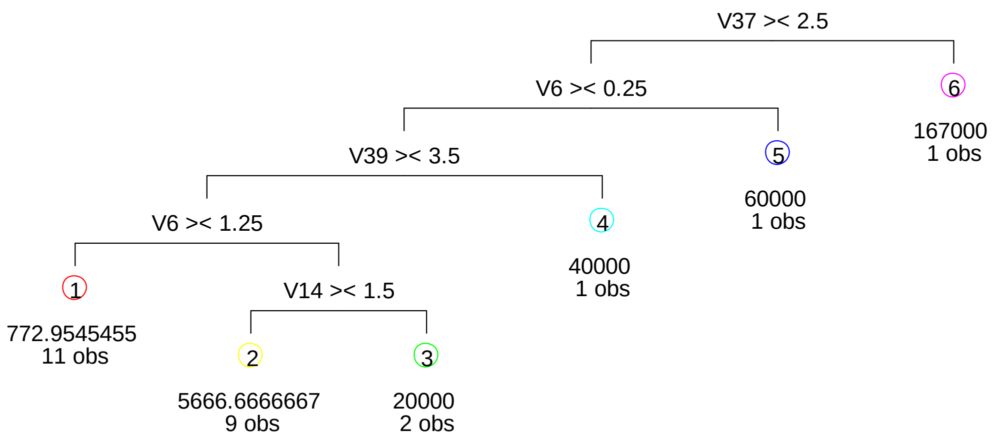
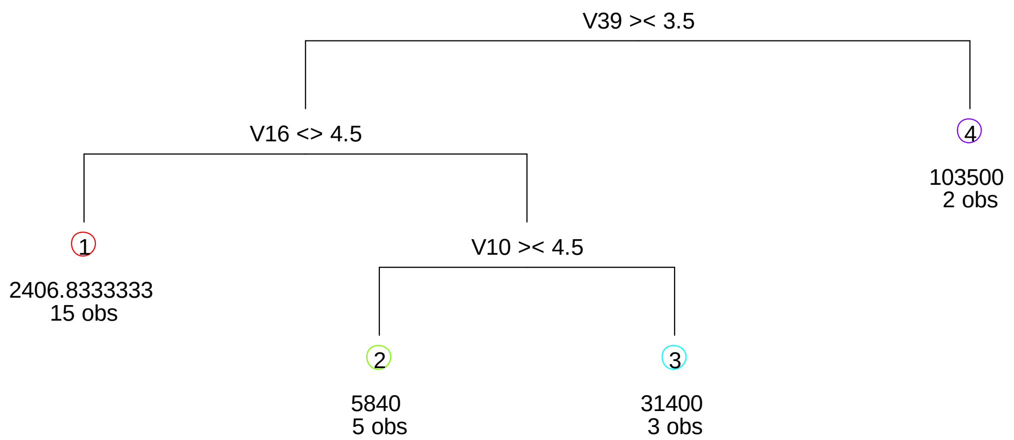


FIGURE 2 Q-Q PLOT

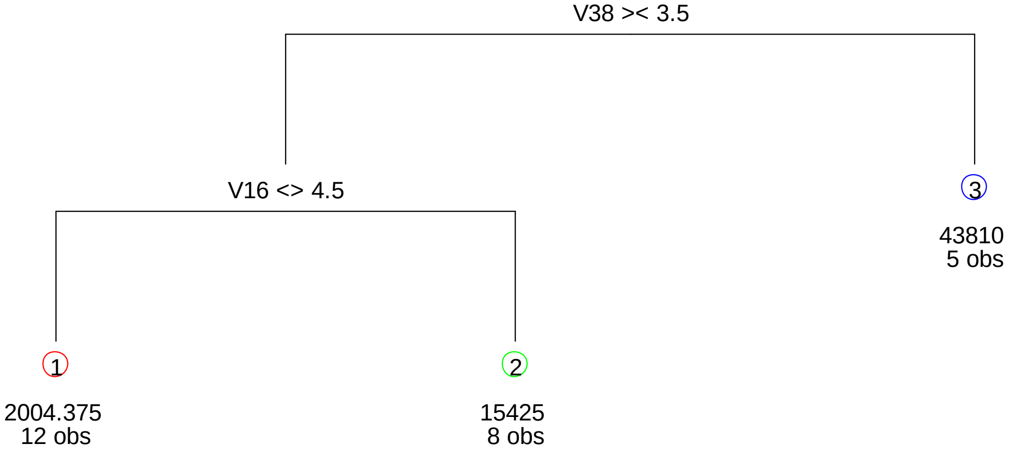


(A)



International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 5 Issue 04, April-2016

(B)



(C)

FIGURE 3 (A) REG TREE 1A (B) REG TREE 1B (C) REG TREE 1C

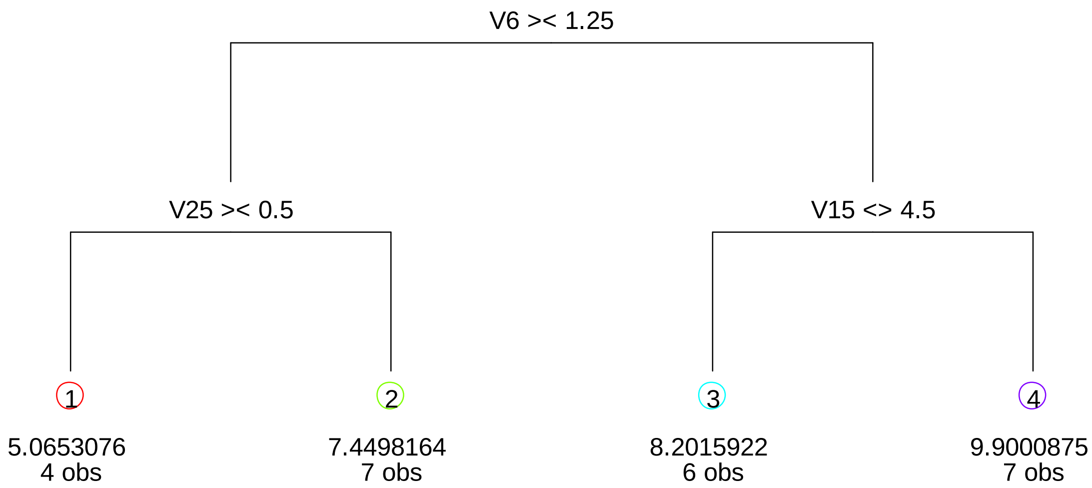


FIGURE 4 LN REG TREE 1