

1
2

3 **Data Mining and Statistical Analysis for Available Budget Allocation**
4 **Preprocurement of Manufacturing Equipment**

5 **ABSTRACT**

6 In a situation where a decision maker faces problems of allotting the available budget on the strategic
7 decisions in a manufacturing industry, data information plays an important role to maintain long run
8 profit in the industry. Statistical analysis was incorporated to determine the correlational strength
9 between the number of years and each of the strategic decisions, their confidence level, and the
10 predicted values. This study identified the strategic areas of addressing the issues which are machine
11 (A_1), accessory (B_1), spare part (C_1) and miscellaneous (D_1), exploring the hidden data of the selected
12 strategic decisions from International Brewery Plc, Ilesha and statistical analysis between the number
13 of years and each of the selected strategic decisions. The model used in this work is simple linear
14 regression while Statistical Analysis Software "SAS" was used for its applications. After exploring the
15 hidden data from a case study, the suggested cost of procurement for machines, accessories, spare-
16 parts and miscellaneous are: ₦119,975,000.00; ₦127,968,000.00; ₦134,965,000.00 and
17 ₦33,491,500.00 respectively. From appendix, the probability of each of the strategic decision is less
18 than 0.05 which implies that the Null-Hypothesis is rejected. The number of years has significant
19 effect on Machines, Accessories, Spare-parts and Miscellaneous. As the number of years increases,
20 the cost of procurement of the strategic decisions increases due to high rate of demand and
21 consumption of their products. However, the cost of procurement may fall depending on the level of
22 demand and maintenance culture. Besides, management of the company may ask decision maker to
23 maintain the cost before procurement. This result may be used for further research on optimization of
24 the available budget for equipment procurement.

25 Keywords: Data mining; statistical analysis; preprocurement; budget allocation; manufacturing
26 equipment.

27 **1. INTRODUCTION**

28 Data Mining (DM) and Statistics are the two disciplines which are commonly used in data analysis
29 and knowledge extraction. Though Statistics is a traditional branch that has evolved from applied
30 Mathematics while Data Mining is a multidisciplinary branch that has evolved from computer science,
31 but both are used for the same purpose [1]. The growth of data mining has been massive in past
32 decade. Its application has increased with the increase of data generation as more and more data
33 being captured through various means of Information Technology like internet. There is a growing
34 research in the area of databases with the help of data mining. Since data mining can be used in
35 advance data research analysis and is capable of extracting valuable knowledge from large data sets
36 [2].

37 It has emerged as a new scientific and engineering discipline to meet such requirements. Data Mining
38 is commonly quoted as “solving problems by analysing data that already exists in databases”. In
39 addition to the mining of structured and numeric data stored in data warehouses, more and more
40 interest is now being experienced in the mining of unstructured and non-numeric data such as text
41 and web in recent times [3].

42 DM is a combination of computational and statistical techniques to perform exploratory data analysis
43 (EDA) on rather large and mostly not very well cleaned data sets (or data bases). In recent times, the
44 issue of capturing data is not considered to be a major issue but since a huge amount of data does
45 not convey any information, screening of useful and non-useful data has become a major challenge.
46 Most modern problems can electronically deal with the cumulative data from many years ago [4]. This
47 leads to a requirement for training the data miners in statistics or statistics graduates in data mining
48 [5].

49 **1.1 Major goals of data mining**

50 [6] distinguished the major goals of data mining into two types:

- 51 a. Verification of user’s hypothesis
- 52 b. Discovery of new patterns that can be used for *prediction* and *description*

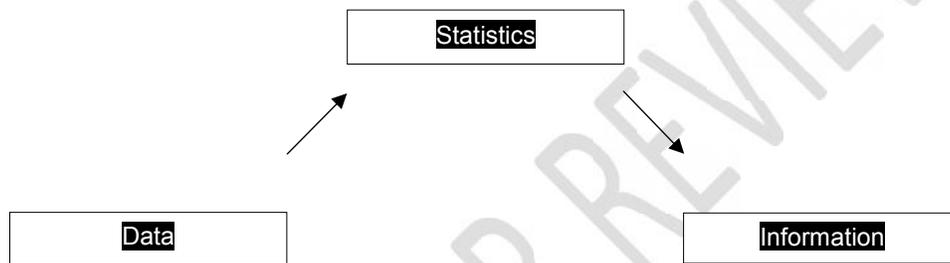
53 Data mining methods seek to discover unexpected and interesting regularities, called patterns, in
54 presented data sets. Statistical significance testing also called as Hypothesis testing can be applied in
55 these scenarios to select the surprising patterns that do not appear as clearly in random data. As
56 each pattern is tested for significance, a set of statistical hypotheses are considered simultaneously.

57 The multiple comparisons of several hypotheses simultaneously are often used in Data Mining [7].

58 [8] added that prediction involves using some variables or fields in the database to forecast unknown
 59 or future values of other variables of interest. Description focuses on finding human-interpretable
 60 patterns describing the data.

61 Various complexities in the stored data (data interrelations) have limited the use of Verification–Driven
 62 Data Mining in decision-making. It must be complemented with the discovery-driven data mining.
 63 Furthermore, in the context of Data Mining, description tends to be more important than prediction.
 64 This is contrast to pattern recognition and machine learning applications where prediction is often the
 65 primary goal [9]; [10].

66 [11] defined Statistics in different ways but the most suitable for this work is as illustrated below:



72 Source: [11]

73 Data: Facts, especially numerical facts, collected together for reference or information [11]; [12].

74 Statistics: Knowledge communicated concerning some particular facts. Statistics is a way to get
 75 information from data. It is a tool for creating an understanding from a set of numbers [11]; [13].

76 1.2 Major Approaches in Statistics

77 [14] introduced the major approaches in solving statistical problems:

S/No.	Statistics Technique	Description
1	Descriptive Statistics	Central Tendency Dispersion Shape (Graphical Display)
2	Regression -Linear -Logistic -Non linear	-Prediction
3	Correlation Analysis	-Modelling

	<ul style="list-style-type: none"> -Pearson correlation -Spearman correlation 	-Association
4	Probability Theory <ul style="list-style-type: none"> -Marginal -Union -Joint -Conditional 	Prediction of the behaviour of the system defined
5	Probability Distribution <ul style="list-style-type: none"> -Discrete Probability Distribution -Continuous Probability Distribution 	
6	Bayesian Classification	Bayes' Theorem and Naïve Bayesian classification
7	Estimation Theory	<ul style="list-style-type: none"> -Model Selection -Estimating Confidence interval and significance level -ROC Curves
8	Analysis of Variance (ANOVA)	Test equality of more than two groups mean
9	Factor Analysis (FA)	Reduction of large no. of variables into some general ones, also known as Data reduction Technique
10	Discriminate Analysis	Predict a categorical response variable
11	Time series analysis <ul style="list-style-type: none"> -Moving Average Method -Exponential smoothing 	Forecasting trends and seasonality

	-Auto regression method	
12	Quality Control Charts -Attributes Charts -Variable charts	Display the spread of individual observation with reference to mean
13	Principal Component Analysis	Data Reduction
14	Canonical Correlation Analysis	
15	Cluster Analysis -Hierarchal -Non Hierarchal	
16	Sampling -Random Sampling -Non Random Sampling	

78 Source: [14].

79 A decision maker needs to be aware of the limited available resources. However, in order to minimize
80 shortages, the past procurement records must be critically analysed to prevent unforeseeable
81 occurrences. Hence, the development of the model on machines cost, accessories cost, spare parts
82 cost and miscellaneous cost. The study would help to determine the cost of purchase of any selected
83 strategic decisions beforehand and create a room for adjustment due to flexibility of the developed
84 model and software. The study proposed to use Statistical Analysis Software (SAS) to analyse the
85 extracted data of the key strategic decisions used in International Brewery Plc, Ilesha, Nigeria, and
86 determine the level of confidence, error terms and to predict the cost of parameters (i.e. machines
87 cost. accessories cost, spare parts cost and miscellaneous cost) with the available budget allocation
88 before procurement.

89

90

91 2. METHODOLOGY

92 In order to analyse the extracted data for preprocurement of manufacturing equipment, the
93 International Brewery Plc, Ilesha was visited to explore past procurement records. These are the
94 following steps taken:

- 95 i. Identification of the equipment procurement such as machines cost, accessories cost,
96 spare parts cost and miscellaneous cost.
- 97 ii. Historical data from a case study International Brewery Plc, Ilesha, Nigeria to determine
98 the correlational strength between the number of years and each of the strategic
99 decisions, their confidence interval, and to predict the cost for each parameter.
- 100 iii. Modified adopted models for prediction of the cost of purchase of each strategic decision.
- 101 iv. Determination of the hypothesis of (iii) above.

102

103 **2.1 Strategic Decisions for Model Development**

104 In this study for proper analysis, four strategic decisions were identified for preprocurement of
105 manufacturing equipment. They are:

- 106 a) **Machine (A):** A machine is a tool that consists of one or more parts, and uses energy to meet
107 a particular goal e.g. labeller, washer, filler, pasteurizer etc.
- 108 b) **Accessories (B):** An accessory aids the performance of a machine e.g. beer spoon, beer
109 paddle, beer siphon etc.
- 110 c) **Spare parts (C):** Spare part is an interchangeable part that is kept in an inventory and used
111 for the repair or replacement of failed parts e.g. hose tail, cask racking spear, female equal
112 tee etc.
- 113 d) **Miscellaneous (D):** Other costs not planned for but can still occur.

114 **2.2 Statistical Analysis of the Data**

115 **2.2.1 Simple Linear Regression Analysis from Data Set**

116 The models below explain the simple linear regression of the relationship between the number of
117 years of procurement and each of the strategic decisions (i.e. machine, accessory, spare-part and
118 miscellaneous).

119 [15] expressed the general form of a simple linear regression analysis as:

$$120 \hat{y} = \beta_0 + \beta_1 x + e_t \quad \dots 2.1$$

121 Where:

122 \hat{y} is the predicted value for machine, accessory, spare-part and miscellaneous.

123 x is the independent variable (i.e. number of years)

124 β_0 is the intercept of the regression

125 β_1 is the slope

126 e_t is error term or residual

127 Where:

128
$$\beta_0 = \frac{-y - m(\sum x)}{n} \quad \dots 2.2$$

129
$$\beta_1 = \frac{n(\sum xy) - (\sum x)(\sum y)}{n(\sum x^2) - (\sum x)^2} \quad \dots \quad 2.3$$

130 **2.2.2 Correlational Strength between the Number of Years of Procurement and the**
131 **Strategic Decisions**

132 [16] made known that, in calculating a correlation coefficient, three different sums of squares (SS) are
133 needed. The sum of squares for variable X, the sum of square for variable Y and the sum of the
134 cross-product of XY.

135 The sum of squares for variable X is:

136
$$SS_{XX} = \sum(x_i - \bar{x})^2 \quad \dots 2.4$$

137 Where:

138 SS_{XX} is the sum of squares for variables X

139 \bar{x} is the average value of X

140 x_i denotes data point

141 The sum of squares for variable Y is:

142
$$SS_{YY} = \sum(y_i - \bar{y})^2 \quad \dots 2.5$$

143 The sum of the cross-products (SS_{XY}) is:

144
$$SS_{XY} = \sum(x_i - \bar{x})(y_i - \bar{y}) \quad \dots 2.6$$

145 Therefore, Pearson's correlation coefficient is given by:

146
$$r = \frac{SS_{XY}}{\sqrt{(SS_{XX})(SS_{YY})}} \quad \dots 2.7$$

147 [15] added that coefficient of determination established a relationship between two variables which
148 determines their best of fits:

149
$$R^2 = \frac{S^2_{XY}}{S_{XX}S_{YY}} \quad \dots \quad 2.8$$

150

151 **2.2.3 Rule of Thumb**

152 To test the hypothesis, let the null hypothesis represents $\mu_1 \approx \mu_0$ which means that there is no
153 significant difference and alternate hypothesis represents $\mu_1 \neq \mu_0$ shows that there is significant

154 difference between the number of years and each of the strategic decisions. If probability $P_r \leq 0.05$,
 155 “reject null hypothesis”.

156 The hypothesis can be tested with a t-statistic:

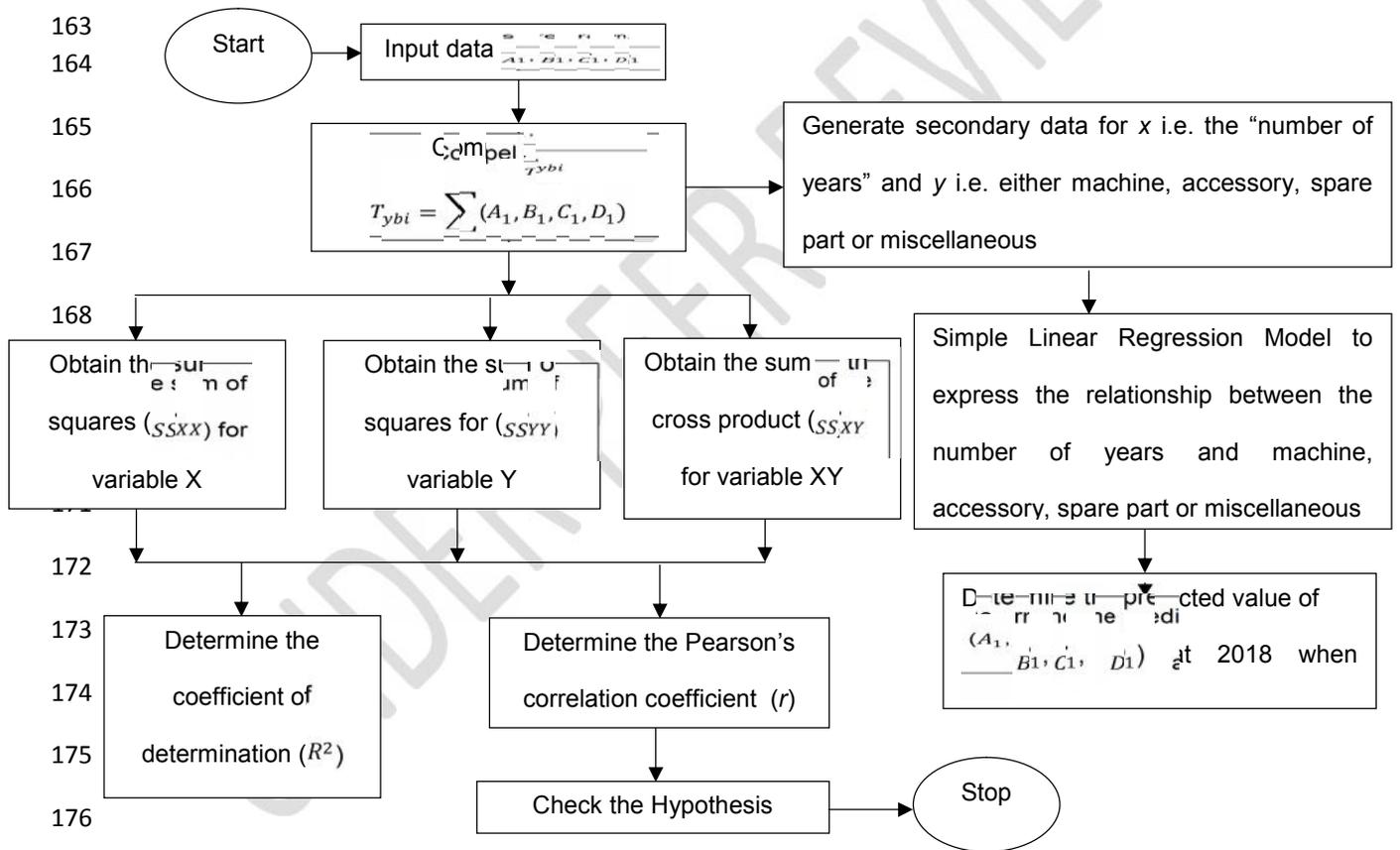
157
$$t_{stat} = \frac{r}{se_r} \quad \dots 2.9$$

158 Where:

159 se_r represents the standard error of the correlation coefficient.

160
$$se_r = \sqrt{\frac{1-r^2}{n-2}} \quad \dots 2.10$$

161 [16] stated that under null hypothesis, t-statistics has $n - 2$ degrees of freedom but test results are
 162 converted to P_r before conclusions are drawn.



177 Fig. 1: Flowchart for Statistical Models Developed

178

179 Table 1: Available Data from International Brewery, Ilesha Nigeria

Date	Number of	Machine (A_1)	Accessory (B_1)	Spare parts (C_1)	Miscellaneous (D_1)	TOTAL

	Years					
1971	1	2,000,000	1,000,000	1,700,000	500,000	5,200,000
1978	8	2,500,000	1,500,000	2,000,000	600,000	6,600,000
1980	10	2,600,000	1,600,000	2,100,000	650,000	6,950,000
1981	11	2,600,000	1,600,000	2,100,000	650,000	6,950,000
1982	12	2,600,000	1,600,000	2,100,000	650,000	6,950,000
1983	13	2,600,000	1,600,000	2,100,000	650,000	6,950,000
1985	15	2,650,000	1,700,000	2,300,000	1,000,000	7,650,000
1986	16	2,650,000	1,700,000	2,300,000	1,000,000	7,650,000
1988	18	3,000,000	5,500,000	8,000,000	2,000,000	18,500,000
1989	19	3,000,000	5,500,000	8,000,000	2,000,000	18,500,000
1991	21	5,500,000	6,000,000	10,000,000	2,500,000	24,000,000
1992	22	5,500,000	6,000,000	10,000,000	2,500,000	24,000,000
1993	23	7,300,000	6,500,000	10,500,000	3,000,000	27,300,000
1994	24	32,200,000	40,000,000	45,000,000	10,000,000	127,200,000
1995	25	32,200,000	40,000,000	45,000,000	10,000,000	127,200,000
1996	26	32,200,000	40,000,000	45,000,000	10,000,000	127,200,000
1997	27	32,200,000	40,000,000	45,000,000	10,000,000	127,200,000
1998	28	32,200,000	40,000,000	45,000,000	10,000,000	127,200,000
2001	31	42,000,000	45,500,000	50,550,000	10,500,000	148,550,000

2002	32	42,000,000	45,500,000	50,550,000	10,500,000	148,550,000
2007	37	82,000,000	85,000,000	95,000,000	20,000,000	282,000,000
2008	38	82,000,000	85,000,000	95,000,000	20,000,000	282,000,000
2009	39	82,000,000	85,000,000	95,000,000	20,000,000	282,000,000
2013	43	95,000,000	96,000,000	100,000,000	25,000,000	316,000,000
2014	44	95,000,000	96,000,000	100,000,000	25,000,000	316,000,000
2017	47	120,000,000	128,000,000	135,000,000	33,500,000	416,500,000
TOTAL		845,500,000	907,800,000	1,009,300,000	232,200,000	2,994,800,000

180 Source: International Brewery Plc, Ilesha, 2017

181 3.0 Results and Discussion

182 3.1 Application of Simple Linear Regression Model between the Number of Years 183 and the Strategic Decisions

184 In order to predict or forecast the costs of procurement of machines, accessories, spare parts and
185 miscellaneous for the year 2018, the method below suggests the amount to be spent for each of them
186 before procuring them:

187 3.1.1 Predicted Value for Machines

188 Machine = $\beta_0 + \beta_1$ (Number of years)

189 Machine = 119,975,000.00

190 **Standard Error:**

$$191 se_r = \sqrt{\frac{1-r^2}{n-2}}$$

192 $se_r = 11,171,425$

193 95% Confidence Limits

194 95% C.I. = predicted \pm S.E (2.064)

195 Upper bound = 141,870,591.60

196 Lower bound = 98,079,408.40

197 To determine how well the model fits the data: variables machine and number of years:

198 $R^2 = \frac{S^2_{XY}}{S_{XX}S_{YY}} = 0.8589$

199 **3.1.2 Predicted Value for Accessories**

200 Accessories = $\beta_0 + \beta_1$ (Number of years)

201 Accessories = 127,968,000.00

202 **Standard Error:**

203 $se_r = \sqrt{\frac{1-r^2}{n-2}}$

204 $se_r = 12,446,682$

205 **95% Confidence Limits**

206 95% C.I. = predicted \pm S.E (2.064)

207 Upper bound = 152,363,049.30

208 Lower bound = 103,572,950.70

209 To determine how well the model fits the data: variables accessory and number of years:

210 $R^2 = \frac{S^2_{XY}}{S_{XX}S_{YY}} = 0.8716$

211 **3.1.3 Predicted Value for Spare-parts**

212 Spare-parts = $\beta_0 + \beta_1$ (Number of years)

213 Spare-parts = 134,965,000.00

214 **Standard Error:**

215 $se_r = \sqrt{\frac{1-r^2}{n-2}}$

216 $se_r = 13,392,197$

217 **95% Confidence Limits**

218 95% C.I. = predicted \pm S.E (2.064)

219 Upper bound = 161,213,224.30

220 Lower bound = 108,716,775.70

221 To determine how well the model fits the data: variables spare-part and number of years:

222 $R^2 = \frac{S^2_{XY}}{S_{XX}S_{YY}} = 0.8808$

223 **3.1.4 Predicted Value for Miscellaneous**

224 Miscellaneous = $\beta_0 + \beta_1$ (Number of years)

225 Miscellaneous = 33,491,500.00

226 **Standard Error:**

$$227 \quad se_r = \sqrt{\frac{1-r^2}{n-2}}$$

$$228 \quad se_r = 3,087,991.40$$

229 **95% Confidence Limits**

230 95% C.I. = predicted \pm S.E (2.064)

231 Upper bound = 39,543,851.93

232 Lower bound = 27,439,148.06

233 To determine how well the model fits the data: variables miscellaneous and number of years:

$$234 \quad R^2 = \frac{S^2_{XY}}{S_{XX}S_{YY}} = 0.8727$$

235 After exploring the hidden data from a case study, the suggested cost of procurement for machines,
236 accessories, spare-parts and miscellaneous are: ₦119,975,000.00; ₦127,968,000.00;
237 ₦134,965,000.00 and ₦33,491,500.00 respectively. From the appendix table A2, B2, C2 and D2, the
238 probability of each of the strategic decision is less than 0.05 which means that the Null-Hypothesis
239 has to be rejected. The coefficient of determination between the number of years and each of the
240 strategic decisions has strong correlations and 95% C.I. (confidence interval) means that the amount
241 proposed for budgeting is within the range of upper bound and lower bound which implies that the
242 amount sets cannot exceed the upper bound but falling under the limit is good while the amount sets
243 for lower bound cannot fall under but exceeding the limit is fine. The amount predicted is within the
244 range of the upper and lower bound.

245

246 **4. CONCLUSION and RECOMMENDATION**

247 **4.1 Conclusion**

248 The number of years has significant effect on Machines, Accessories, Spare-parts and Miscellaneous.
249 As the number of years increases, the cost of procurement of those strategic decisions increases due
250 to high rate of demand and consumption of their products. However, the cost of procurement may fall
251 depending on the level of demand and maintenance culture. Besides, management of the company
252 may ask decision maker to maintain the cost before procurement. The model used in this work is
253 simple linear regression while Statistical Analysis Software "SAS" was used for its applications.

254 **4.2 Recommendation**

255 As it is stated, data mining is the extraction of hidden information in the company. This study made
256 use of old records for preprocurement of the manufacturing equipment (such as machine, accessory,
257 spare-part and miscellaneous) for available budget allocation which will subsequently be used for
258 budgeting with the limited available budget. Therefore, this work is recommended that the procedures
259 developed with the software "SAS" be used for budgeting, to determine the cost of procurement
260 beforehand with the use of past procurement data and the limited available budget. This would further
261 assist decision maker to forecast the amount to be spent on them using another tools.

262

263 **References**

- 264 1. **Jaya, Srivastava. and Abhay, Kumar Srivastava. (2015):** Understanding Linkage between
265 Data Mining and Statistics. International Journal of Engineering Technology, Management
266 and Applied Sciences (IJETMAS). www.ijetmas.com. Vol. 3, Issue 10, ISSN 2349-4476.
- 267 2. **Ganesh, S. (2002):** Data mining: Should it be Included in The Statistics Curriculum? The 6th
268 international conference on teaching statistics (ICOTS 6), Cape Town, South Africa.
- 269 3. **Goodman, A., Kamath, C. and Kumar, V. (2008):** Data Analysis in the Twenty-First Century,
270 Vol. 1, No. 1, Journal Volume: 1; Journal Issue: 1, Lawrence Livermore National Laboratory
271 (LLNL), Livermore, CA, 1-3.
- 272 4. **Fayyad, U., Chaudhuri, S. and Bradley, P. (2000):** Data mining and its rule in database
273 systems, Proceeding of 26th VLDB Conference. Cairo, Egypt, Morgan Kaufmanu, 63 – 124.
- 274 5. **Stephen, M. Stigler. (2002):** Statistics on the table: the history of statistical concepts and
275 methods, Cambridge, Mass: Harvard University Press.
- 276 6. **Kurt, Thearling. (2010):** An Introduction to Data Mining,
277 <http://www.thearling.com/text/dmwhite/dmwhite.htm>.
- 278 7. **Neal, Leavitt. (2011):** Data Mining Corroborate Masses.
279 http://www.leavcom.com/ieee_may02.htm [30] Robert Nisbet Et al, The Handbook of
280 Statistical Analysis and Data Mining Applicants, Academic Press, ISBN: 0123747651, (2009),
281 www.elsevierdirect.com/datamining.
- 282 8. **Jure, Leskovec. (2010):** Data Mining: Introduction,
283 <http://www.stanford.edu/class/cs345a/slides/01-intro.pdf>.

284 9. **Gorunescu, F. (2011):** Data Mining Concepts, Models and Techniques, Vol. 12, intelligent
 285 systems reference library, Springer.

286 10. **Berry, M. J. A. and Linoff, G. S. (2000):** Mastering Data Mining -The Art and Science of
 287 Customer Relationship Management, New York, Wiley.

288 11. **Moore, David. S. (2003):** "The Basic Practice of Statistics." Third edition. W.H. Freeman and
 289 Company. New York.

290 12. **Wiley Inter Science, (2007):** Data Analysis in the 21st Century, www.interscience.wiley.com/.

291 13. **Stephen, M. Stigler. (1986):** The History of Statistics: The Measurement of Uncertainty
 292 before 1900, Cambridge, MA: Belknap Press of Harvard University Press.

293 14. **Lomax, R. G. (2007):** An Introduction to Statistical Concepts for Education and Behavioural
 294 Sciences (2nd ed.). New York: Routledge.

295 15. **Shen, K. S., Crossley, J. N. and Lun, W. C. (1999):** *The Nine Chapters on the Mathematical*
 296 *Art. Companion and Commentary.* Oxford: Oxford University Press.

297 16. **Witten, I. H. and Frank, E. (2005):** "Data Mining: Practical machine learning tools and
 298 techniques." Morgan Kaufmann, San Francisco, 2nd edition, 2005.

299

300 **Appendices**

301 **Table A**

Pearson Correlation Coefficients, N = 26					
Prob > r under H0: Rho=0					
	years	Machine	Accessory	Spare_Part	Miscellaneous
years	1.00000	0.92678 <.0001	0.93359 <.0001	0.93849 <.0001	0.93421 <.0001
Machine	0.92678 <.0001	1.00000	0.99682 <.0001	0.99450 <.0001	0.99454 <.0001
Accessory	0.93359 <.0001	0.99682 <.0001	1.00000	0.99894 <.0001	0.99681 <.0001
Spare_Part	0.93849 <.0001	0.99450 <.0001	0.99894 <.0001	1.00000	0.99366 <.0001
Miscellaneous	0.93421 <.0001	0.99454 <.0001	0.99681 <.0001	0.99366 <.0001	1.00000

302

The SAS System

303

The REG Procedure

304

Model: MODEL1

305

Dependent Variable: Machine

306

Table A1

Number of Observations Read	26
Number of Observations Used	26

307

308

309

Table A2

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	2.922218E16	2.922218E16	146.11	<.0001
Error	24	4.800132E15	2.000055E14		
Corrected Total	25	3.402232E16			

310

311

312

313

314

315

316

Table A3

Root MSE	14142330	R-Square	0.8589
Dependent Mean	32519231	Adj R-Sq	0.8530
Coeff Var	43.48913		

320

Table A4

Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > t
Intercept	1	-35979715	6309256	-5.70	<.0001
years	1	2826941	233874	12.09	<.0001

325

Table A5

Forecasts for Variable Machine					
Obs	Time	Forecasts	Standard Error	95% Confidence Limits	
48	2018	119975000.0	11171425	98079408.4	141870591.6
49	2019	119975000.0	15790884	89025436.0	150924563.9
50	2020	119975000.0	19336579	82076001.5	157873998.5

326

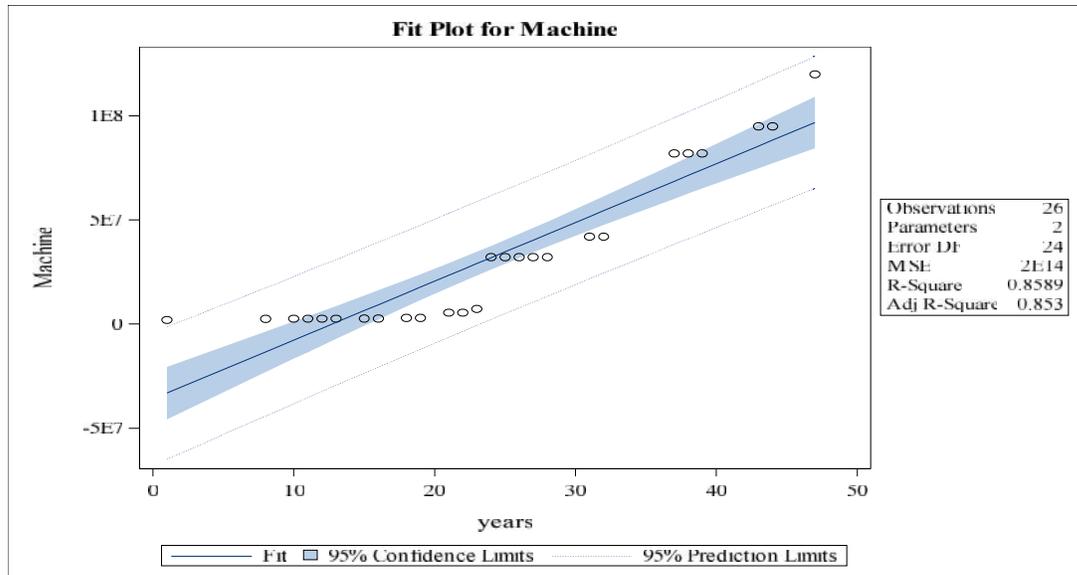
327

328

329

330

331



332

333

The SAS System

334

The REG Procedure

335

Model: MODEL1

336

Dependent Variable: Accessory

337

Table B1

Number of Observations Read	26
Number of Observations Used	26

338

339

Table B2

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	3.236141E16	3.236141E16	162.90	<.0001
Error	24	4.767925E15	1.986635E14		
Corrected Total	25	3.712933E16			

344

Table B3

Root MSE	14094805	R-Square	0.8716
Dependent Mean	34915385	Adj R-Sq	0.8662
Coeff Var	40.36847		

347

348

349

350

351

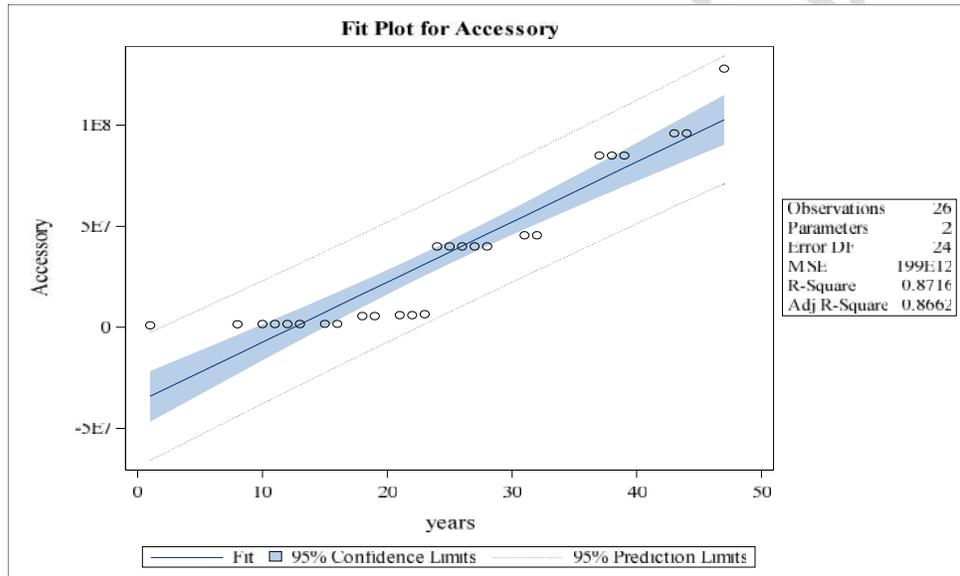
352

353 **Table B4**

Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > t
Intercept	1	-37169013	6288054	-5.91	<.0001
years	1	2974912	233088	12.76	<.0001

358 **Table B5**

Forecasts for Variable Accessory					
Obs	Time	Forecasts	Standard Error	95% Confidence Limits	
48	2018	127968000.0	12446682	103572950.7	152363049.3
49	2019	127968000.0	17593468	93485436.0	162450564.0
50	2020	127968000.0	21543917	85742699.4	170193300.6



363

364

365

366

367

368

The SAS System
The REG Procedure
Model: MODEL1
Dependent Variable: Spare-part

Table C1

Number of Observations Read	26
Number of Observations Used	26

371

372

373

374 **Table C2**

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	3.684116E16	3.684116E16	177.28	<.0001
Error	24	4.987551E15	2.078146E14		
Corrected Total	25	4.182872E16			

379 **Table C3**

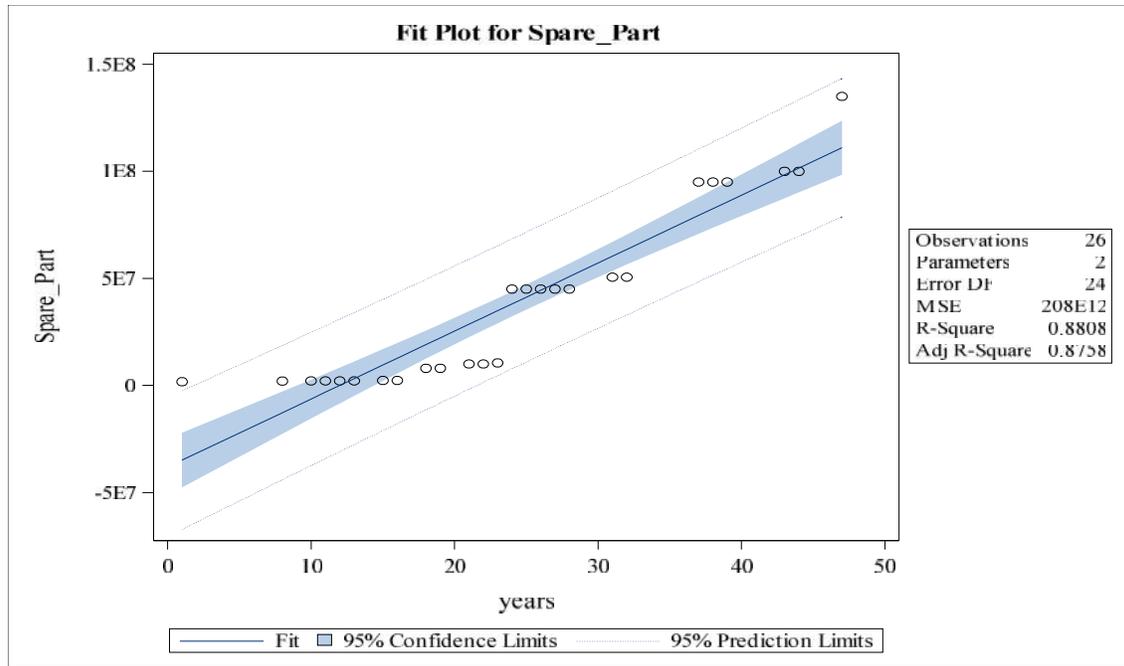
Root MSE	14415777	R-Square	0.8808
Dependent Mean	38819231	Adj R-Sq	0.8758
Coeff Var	37.13566		

382 **Table C4**

Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > t
Intercept	1	-38092792	6431248	-5.92	<.0001
years	1	3174147	238396	13.31	<.0001

388 **Table C5**

Forecasts for Variable Spare_Part					
Obs	Time	Forecasts	Standard Error	95% Confidence Limits	
48	2018	134965000.0	13392197	108716775.7	161213224.3
49	2019	134965000.0	18929960	97862960.8	172067039.2
50	2020	134965000.0	23180505	89532045.7	180397954.3



393

394

The SAS System

395

The REG Procedure

396

Model: MODEL1

397

Dependent Variable: Miscellaneous

398

Table D1

Number of Observations Read	26
Number of Observations Used	26

399

400

Table D2

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	1.966149E15	1.966149E15	164.60	<.0001
Error	24	2.86676E14	1.194483E13		
Corrected Total	25	2.252825E15			

401

Table D3

Root MSE	3456130	R-Square	0.8727
Dependent Mean	8930769	Adj R-Sq	0.8674
Coeff Var	38.69913		

402

403

404

405

406

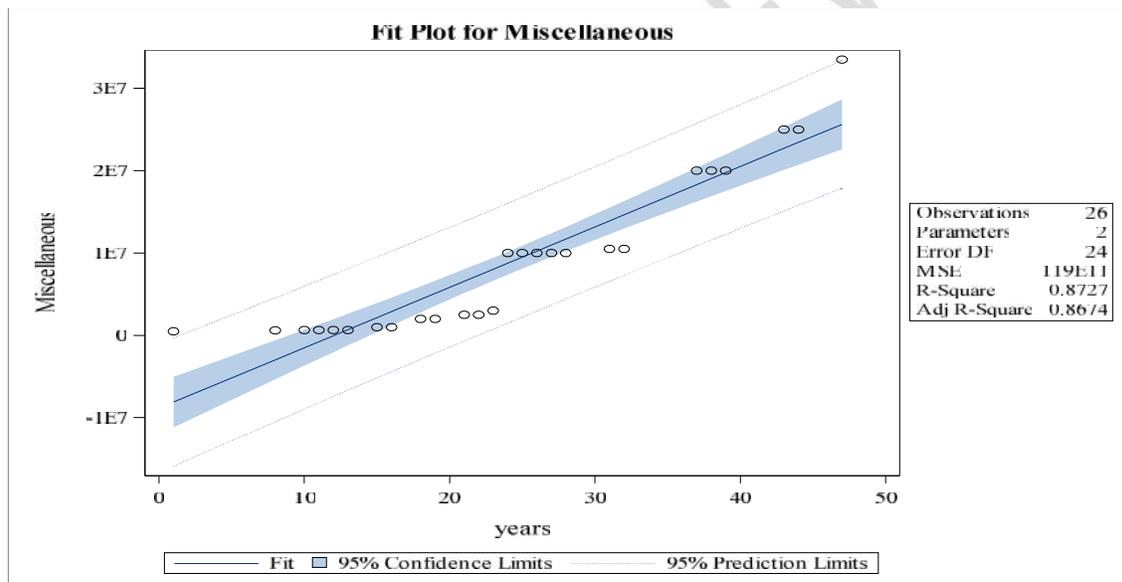
407 **Table D4**

Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > t
Intercept	1	-8837119	1541868	-5.73	<.0001
years	1	733278	57154	12.83	<.0001

408

409 **Table D5**

Forecasts for Variable Miscellaneous					
Obs	Time	Forecasts	Standard Error	95% Confidence Limits	
48	2018	33491500.00	3087991.4	27439148.06	39543851.93
49	2019	33491500.00	4364896.3	24936460.39	42046539.60
50	2020	33491500.00	5344992.9	23015506.42	43967493.57



414