ASSIGNMENT 2 FRONT SHEET

Qualification	**BTEC Level 5 HND Diploma in Computing**		
Unit number and title	**Unit 8: Data Analytics**		
Submission date	08/04/2024	**Date Received 1st submission**	08/04/2024
Re-submission Date		**Date Received 2nd submission**	
Group number:	**Student names & codes**	**Final scores**	**Signatures**
	1. Nguyễn Hoàng Thanh Nam BS00499		Nam
	2. Hoàng Anh Quý BS00311		Quý
	3. Phạm Anh Tuấn BS00304		Tuấn
Class	**DA06201**	**Assessor name**	Ms. Thai Thi Thanh Thao

Student declaration

I certify that the assignment submission is entirely my own work and I fully understand the consequences of plagiarism. I understand that making a false declaration is a form of malpractice.

Grading grid

P5	P6	M3	D2	P7	P8	M4	D3

OBSERVATION RECORD

Student 1	**Nguyễn Hoàng Thanh Nam**
Description of activity undertaken	
P5 Identify predictive analytic techniques and describe them with examples. **M3** Compare a range of predictive analytical techniques for forecasting purposes. **D2** Evaluate how predictive analytic techniques can be used for forecasting purposes.	
Assessment & grading criteria	
How the activity meets the requirements of the criteria	

Student signature:	**Nam**	**Date:**	
Assessor signature:		**Date:**	
Assessor name:			
Student 2	**Hoàng Anh Quý**		
Description of activity undertaken			

P6 Apply an appropriate tool or programming language to demonstrate these predictive analytic techniques.

P7 Analyse prescriptive analytic methods with appropriate examples.

P8 Demonstrate these methods using an appropriate programming language or tool.

Assessment & grading criteria

How the activity meets the requirements of the criteria

Student signature:	**Quý**	**Date:**	
Assessor signature:		**Date:**	
Assessor name:			

Student 3	**Phạm Anh Tuấn**
Description of activity undertaken	
P6 Apply an appropriate tool or programming language to demonstrate these predictive analytic techniques. **M4** Describe how these prescriptive analytic methods are used to find the best course of action in a situation. **D3** Apply an appropriate programming language or tool to demonstrate how these prescriptive analytic methods are used to find the best course of action in a situation.	
Assessment & grading criteria	

How the activity meets the requirements of the criteria			
Student signature:	Tuấn	**Date:**	
Assessor signature:		**Date:**	
Assessor name:			

Summative Feedback:	☐ Resubmission Feedback:	
Grade:	Assessor Signature:	Date:

Internal Verifier's Comments:
Signature & Date:

Contents

Introduction

In this assignment, we will explore the basic steps to apply predictive analytics and prescriptive analytics methods. Firstly, we will explore predictive analytics techniques such as linear regression, decision trees, and time series analysis. Then, we will compare and evaluate these methods to determine the most suitable applications. Lastly, we will grasp and apply prescriptive analytics methods such as optimization modeling and simulation to find optimal actions in real-world situations.

Main Body

P5 Identify predictive analytic techniques and describe them with examples.

In the construction of predictive analytical models, the analyst needs to have a clear understanding of the mathematical properties and descriptive characteristics of various functional relationships. Charts used in the analysis process assist the analyst in understanding these relationships, enabling them to select the relationships that are most suitable for the situation or requirements to be incorporated into the analytical model. Two types of regression models are commonly used in the business environment:

Regression models of cross-sectional data: Cross-sectional data typically represents data at a single point in time for multiple entities. Scatter charts are often used to visualize cross-sectional data.

Regression models of time-series data: The focus of this model is future prediction, where time-series data is data collected at specific time intervals. Line charts are typically used to visualize time-series data.

Examples cross-sectional data and time-series and their chart:

Respondent	Age	Education	Income
1	23	Some undergraduate courses	$10,000 to < $20,000
3	24	Bachelor Degree	$60,000 to < $70,000
4	23	Bachelor Degree	$30,000 to < $40,000
5	23	Bachelor Degree	$30,000 to < $40,000
6	24	Bachelor Degree	$10,000 to < $20,000
7	30	Doctorate Degree	$60,000 to < $70,000
8	28	Bachelor Degree	$50,000 to < $60,000
9	20	Some undergraduate courses	$0 to < $10,000
11	24	Associate Degree	$20,000 to < $30,000
12	24	Bachelor Degree	$30,000 to < $40,000
13	27	Bachelor Degree	$70,000 to < $80,000
14	24	Master Degree	$0 to < $10,000
15	21	Bachelor Degree	$0 to < $10,000
16	23	Master Degree	$70,000 to < $80,000
17	23	Master Degree	$50,000 to < $60,000
18	20	Some undergraduate courses	$0 to < $10,000
19	46	J.D.	$10,000 to < $20,000
20	23	Bachelor Degree	$40,000 to < $50,000
21	23	Some undergraduate courses	$10,000 to < $20,000
23	21	Some undergraduate courses	$0 to < $10,000
26	59	Associate Degree	$150,000 or more
27	65	Doctorate Degree	$70,000 to < $80,000
29	50	Doctorate Degree	$40,000 to < $50,000
31	31	Some undergraduate courses	$20,000 to < $30,000
34	55	J.D.	$110,000 to < $130,000

Figure 1: Cross-sectional data

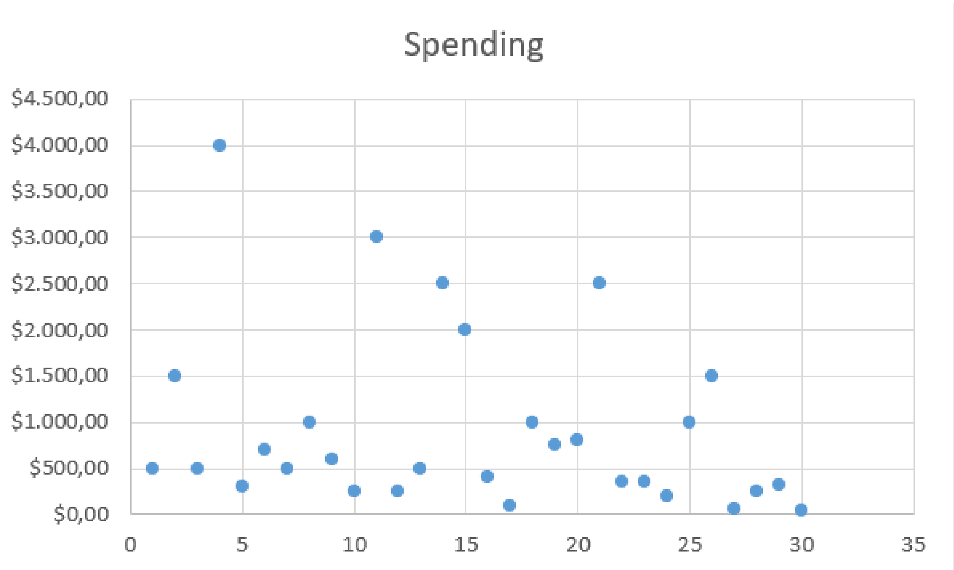


Figure 2: Cross-sectional Chart

Time to Pay Suppliers	
Month	**Working Days**
01/01/2010	8,32
01/02/2010	8,28
01/03/2010	8,29
01/04/2010	8,32
01/05/2010	8,36
01/06/2010	8,35
01/07/2010	8,34
01/08/2010	8,32
01/09/2010	8,36
01/10/2010	8,33
01/11/2010	8,32
01/12/2010	8,29

Figure 4: Time-series data

Figure 3: Time-series chart

Simple linear regression

Simple linear regression involves finding a linear relationship between one independent variable, X, and one dependent variable, Y. The relationship between two variables can assume many forms, as illustrated in Figure 5. The relationship may be linear or nonlinear, or there may be no relationship. Because we are focusing our discussion on linear regression models, the first thing to do is to verify that the relationship is linear, as in (a). We would not expect the data to line up perfectly along a straight line; we simply want to verify that the general relationship is linear. If the relationship is nonlinear, as in (b), then alternative approaches must be used, and if no relationship is evident, as in (c), then it is pointless to consider developing a linear regression model.

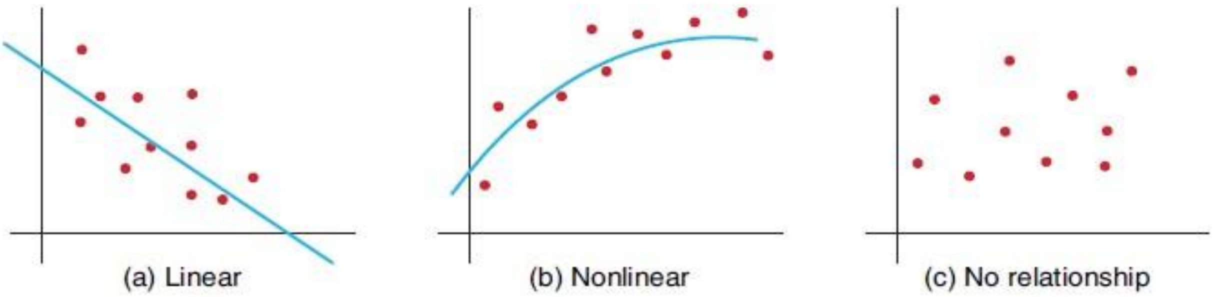


Figure 5: Simple Linear regression

To determine if a linear relationship exists between the variables, we recommend that you create a scatter chart that can show the relationship between variables visually.

Multiple linear regression

Multiple linear regression is a statistical method used to model the relationship between a dependent variable and two or more independent variables. This method helps predict the values of dependent variables based on the values of independent variables, by finding the most optimal straight line reflecting the relationship between them. This is an important tool in predictive analysis and is widely applied in many business fields.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k + \varepsilon$$

Figure 6: Multiple linear regression

In multiple linear regression, more than one independent variable is used to predict the dependent variable. The mathematical formula of the model has the form:

In which:

1. Y is a dependent variable that should be expected.
2. X1, X2..., Xn are independent variables.
3. b0, b1, ..., bn are regression coefficients, in which b0 is the intercept coefficient and b1, ..., bn are the slope coefficients.
4. e is bad randomly.

Logistic Regression

Describe the logistic regression model formulation and its estimation from data.

Logistic regression is a statistical method used to model the probability of a binary outcome based on one or more independent variables. It is typically used in situations where the dependent variable is a binary variable with two classes, such as "true" or "false", "success" or "failure". Logistic regression uses the logistic function to transform the result of a linear function into a probability value ranging from 0 to 1. This allows us to predict the probability of an event occurring or not occurring based on the independent variables.

Formulation:

A logistic regression model is built based on a logit function, defined by:

a.k.a. **Log Odds**

or **Logit**

Intercept

$$\boxed{\log\left(\frac{P}{1-P}\right)} = \beta_0 + \beta_1 X$$

Figure 7: Formulation (1)

Where *P* is the probability of the dependent variable.

The logit function is a linear function of the independent variables:

$$\text{logit P} = b_0 + b_1 x_1 + b_2 x_2 + \cdots + b_k x_k$$

Figure 8: Formulation (2)

Where β0, β1…, βn are the model coefficients, and X1, X2…, Xn are the values of the independent variables.

Estimation from data:

The logistic regression model is estimated from data using an optimization method like Gradient Descent or second-order optimization methods like Newton-Raphson.

The objective is to adjust the coefficients β so that the loss function (typically cross-entropy loss function) is minimized.

Understand two phases in logistic regression

First Phase:

- Estimating the probabilities of each class or dependent variable.
- In the binary case, we estimate the probability P(Y=1) (class 1), from which we can estimate the probability of the other class (P(Y=0)).
- This is typically done by using the sigmoid function to convert logit values to probabilities.

Second Phase:

- Classifying each case into one of the classes.

- This is usually done by setting a threshold on the estimated probabilities. For example, with a threshold of 0.5, cases with a probability P(Y=1) greater than or equal to 0.5 are classified into class 1, and the rest into class 0.

Solve a multi-class classification problem.

Logistic Regression can be extended to solve multi-class classification problems.

Logistic regression models for each class will predict the probabilities of each class, and thus, we only need to estimate $m-1$ probabilities (where m is the number of classes) as the sum of probabilities of all classes must equal 1.

Finally, to predict the final class, we can use the voting method, where the class predicted with the highest probability is chosen as the final output.

Apply regression analysis

Through studying the data set and what I have learned, I decided to apply linear regression to this essay. Here are some reasons why I might choose simple linear regression:

- Simplicity and Ease of Interpretation: Simple linear regression is straightforward and easy to understand, especially when there's only one independent variable. Explaining the results of the model to non-statisticians becomes much easier.
- Simple Linear Relationship: When there's a simple linear relationship between the dependent and independent variables, simple linear regression is the most appropriate choice. This is the case when there's no complex interaction between variables or when there's no need to consider additional independent variables.
- Less Data Requirement: Simple linear regression often doesn't require a large amount of data to estimate the parameters. This can be useful when your dataset is small or when you want to perform quick and efficient analysis.
- Specific Analysis Objective: If your objective is simply to predict or explain the dependent variable based on a specific independent variable, simple linear regression is a reasonable choice.
- Good Predictive Performance: In some cases, simple linear regression can provide good predictive performance for your specific dataset. This is especially true when the relationship between variables is simple and not too complex.

In summary, the choice of simple linear regression can be explained by the simplicity of the method, its suitability for the relationship between variables, the minimal data requirement, and the specific analysis objectives

P6 Apply an appropriate tool or programming language to demonstrate these predictive analytic techniques.

From the received document, it is clear that Ms. Elizabeth Burke is concerned about the increasing defects in the quality of materials supplied by the suppliers. An initiative has been proposed to mitigate this issue: to manage deliveries more strictly and improve material quality by restructuring the production policies of the suppliers. The following will be the analysis results from the data to visualize how the supplier initiative will affect PLE and what would have happened if the supplier initiative had not been implemented.

Evaluate the quality of the Mower through the "Mower Test" in the data, and from there visualize the ratio of defective or that do not meet performance of Mowers. The Mower Test is a dataset of 3000 Mowers divided into 30 Samples, each Sample consists of 100 Mowers evaluated based on two values: Pass or Fail.

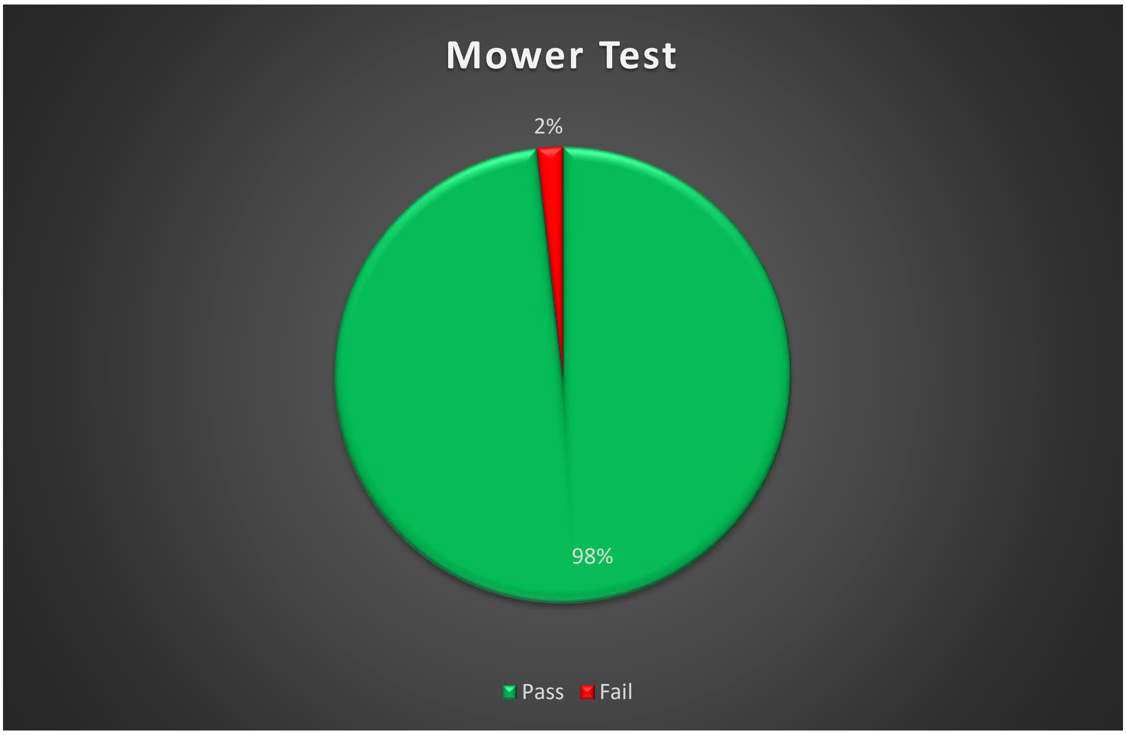


Figure 9: Mower test

Out of a total of 3000 data points for Mower, the majority of results show that the products meet the quality standards with specific figures of 2946/3000 items, accounting for 98%, and a small portion of products do not meet the quality standards with specific figures of 54/3000 items, accounting for 2%.

The document provided on pages 168-170 states that the standard weight of a Blade is 5.00 ± 0.2, indicating that a Blade with a weight ranging from 4.80 to 5.20 is considered acceptable. “Blade Weight” is a dataset consisting of 350 data points of mower blade weights used to estimate the error rate of the manufacturing process.

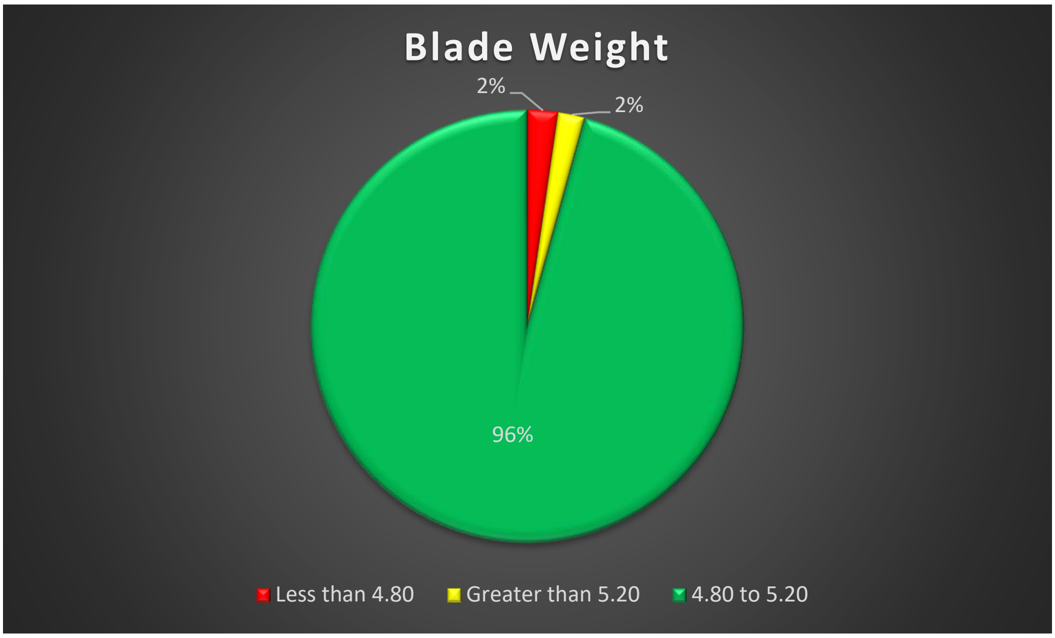


Figure 10: Blade Weight

Out of a total of 350 data points for Blade Weight, the majority of blade weights meet the standard (4.80 to 5.20) with 335 out of 350, accounting for approximately 96% of the total. In addition, the number of defective blades is 15 out of 350, representing about 4% of the total.

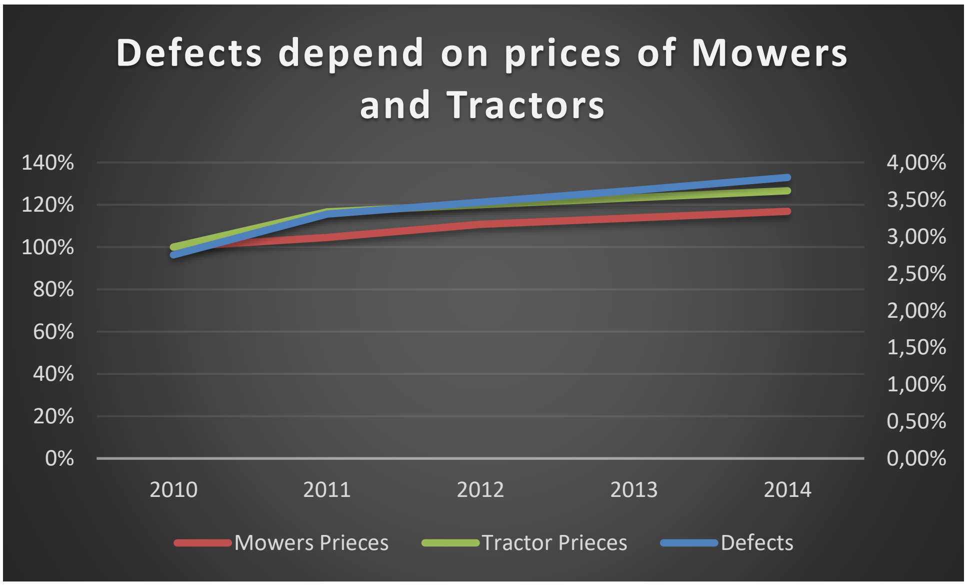


Figure 11: Defects depend on prices of Mowers and Tractors

	Defects	Mowers Prieces	Tractor Prieces
2010	2,75%	100%	100%
2011	3,30%	104,62%	116,67%
2012	3,46%	110,77%	120,00%
2013	3,62%	113,85%	123,33%
2014	3,80%	116,92%	126,67%

Figure 12: Data of Defects depend on prices of Mowers and Tractors

The chart above illustrates the relationship between Defects (represented by the blue line) and two variables: Mower and Tractor prices. The Defects value is calculated as the percentage of complaints per year, determined by the total sales volume of the two types of machines from the "World" column of the "Mower Unit Sales" and "Tractor Unit Sales" sheets, compared to the total number of Complaints from the "World" column of the "Complaints" sheet. The data for Mower and Tractor prices are calculated based on the increase or decrease each year, with 2010 as the reference point. The chart's results show that the price data for the two types of machines gradually increased from 2010 to 2014, and the Defects rate also steadily increased from 2010 to 2014. The provided document does not mention any improvements in machine performance, yet the prices increase each year. This clearly shows the correlation between these three variables.

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0,998982219
R Square	0,997965474
Adjusted R Square	0,995930949
Standard Error	0,000255229
Observations	5

ANOVA

	df	*SS*	*MS*	*F*	*ignificance F*
Regression	2	6,39E-05	3,2E-05	490,515	0,002035
Residual	2	1,3E-07	6,51E-08		
Total	4	6,4E-05			

	Coefficients	*andard Err*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*ower 95,0%*	*pper 95,0%*
Intercept	-0,016971742	0,002383	-7,12123	0,019154	-0,02723	-0,00672	-0,02723	-0,00672
Mower	0,016114295	0,005167	3,118758	0,089259	-0,00612	0,038346	-0,00612	0,038346
Tractor	0,028342493	0,003424	8,27697	0,014285	0,013609	0,043076	0,013609	0,043076

DEFECTS = -0,017+0,016 * MOWER + 0,028 * TRACTOR

Figure 13: Formula of Defects depends on prices of Mowers and Tractors

Based on the compiled data, our team has used Excel's Data Analysis Regression to derive a predictive formula for Defects. From this formula, it can be predicted that if the prices of Mowers and Tractors decrease, the Defects rate will also decrease.

From the PLE company data file, I see a correlation between the Dealer Satisfaction worksheet and the Complaints worksheet. So, I compiled the two tables into one table below to show the dealer's satisfaction level and total complaints from 2010 to 2014. In which, the level of satisfaction is measured on a scale of 0–5 (0 = very poor, 1 = poor, 2 = below average, 3 = average, 4 = above average, and 5 = excellent).

Year	level 0	level 1	level 2	level 3	level 4	level 5	Sample size	% poor	% average	% positive	Complaints compared to the previous year (%)	Complaints
2010	1	0	4	21	37	17	80	6%	26%	68%	0%	2439
2011	0	0	4	19	37	20	80	5%	24%	71%	16%	2841
2012	1	1	4	16	63	37	122	5%	13%	82%	6%	3009
2013	1	3	9	23	74	87	197	7%	12%	82%	8%	3251
2014	3	4	10	30	98	128	273	6%	11%	83%	4%	3395

Figure 14: Dealer satisfaction and complaints.

In that table, I calculated the percentage of satisfaction with three levels as follows: poor (level 0 - level 2), average (level 3), and positive (level 4 - level 5). Besides, I also calculated the percentage change in the number of complaints this year compared to the previous year.

From the above data, we found that the poverty level percentage is an insignificant variable, so I removed it from the Multiple Linear Regression model.

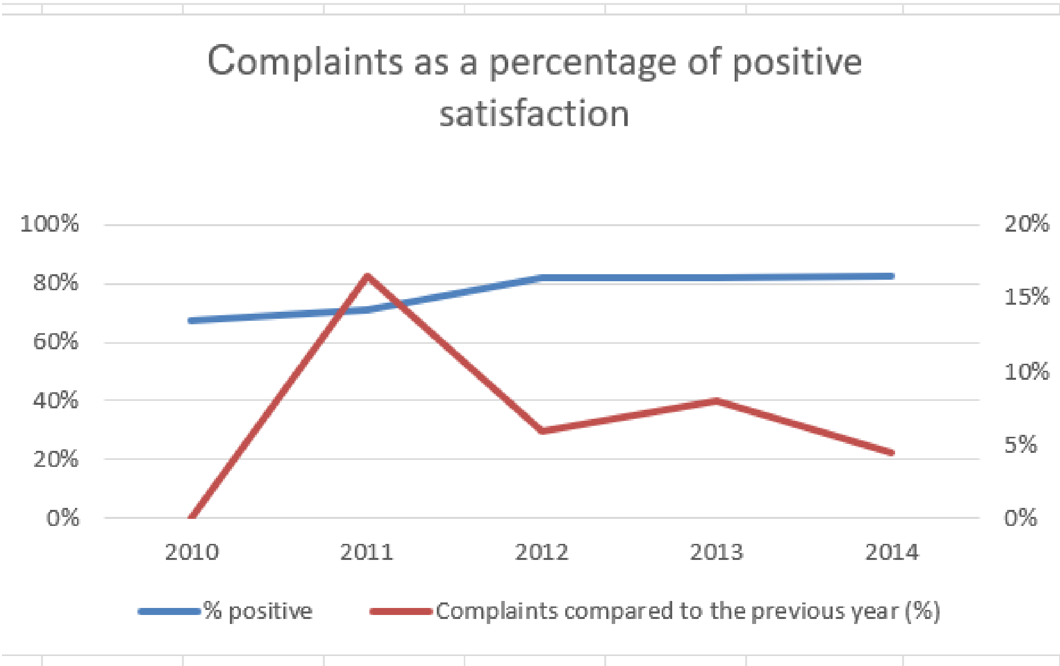


Figure 15 Complaints as a percentage of positive satisfaction.

From the graph above, we can see that the complaint rate from 2011 to 2014 tends to decrease as the level of satisfaction increases.

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.922231
R Square	0.850511
Adjusted R	0.701021
Standard E	204.2725
Observatio	5

ANOVA

	df	*SS*	*MS*	*F*	*ignificance F*
Regressio	2	474809.5	237404.8	5.689442	0.149489
Residual	2	83454.5	41727.25		
Total	4	558264			

	Coefficients	*andard Err*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*ower 95.0%*	*pper 95.0%*
Intercept	4446.48	12533.55	0.354766	0.756682	-49481	58373.98	-49481	58373.98
% everage	-5420.12	13137.63	-0.41256	0.719947	-61946.8	51106.56	-61946.8	51106.56
% positive	-687.407	13355.76	-0.05147	0.96363	-58152.6	56777.79	-58152.6	56777.79

Complaint = 4446.48 -5420.12***Average(%)** -687.407* **Positive(%).**

Figure 16 Correlation between dealer satisfaction and complaints.

Based on aggregated data, our team used Excel's analytical regression tool to express the formula. From the above formula, we realize that if we want to reduce complaints, we need to increase the satisfaction rate of the Dealer.

Based on the results of the data analysis, in order to reduce negative feedback in the future, it is necessary to improve the manufacturing process to optimize the quality of the Blade and reduce the error rate during production. The Blade directly affects the performance of the Mower. For instance, a non-standard weight can cause uncomfortable vibrations during use, or insufficient sharpness can reduce the machine’s efficiency and increase operating time, leading to reduced lifespan and increased maintenance effort. Therefore, optimizing the quality of the Blade enhances the performance of the Mower and improves the selling price of the Mower and Tractor, improves customer satisfaction, and can thus reduce the Defects rate received in the coming years.

P7 Analyse prescriptive analytic methods with appropriate examples.

Describe the purpose of prescriptive analytics and the fundamental methods in prescriptive analytics.

According to the book Business Analytics written by James Evans, Prescriptive Analytics is the next stage after descriptive and predictive analytics. It focuses on providing recommendations on how to optimize business practices to accommodate a variety of predicted outcomes. (Evans, 2017)

The main purposes of prescriptive analysis are:

- Decision Optimization: Prescriptive analytics helps optimize decisions based on certain constraints and limitations, aiming to maximize performance or benefits.
- Identify strategies: Helps identify optimal strategies and action plans in a business or organizational environment.
- Predict the results of decisions: Prescriptive analysis allows predicting the results of different decisions, helping decision makers have an overall view before implementation.

The fundamental methods in prescriptive analysis may include:

- Linear Regression: Although linear regression is commonly used in predictive analytics, it can also be applied in prescriptive analytics to optimize decisions based on linear relationships between variables. Use to model relationships and trends in data. Predict a dependent variable based on one or more independent variables.
    - For example: Suppose you want to predict the selling price of a house based on the area. Using linear regression, you can build a model to predict the selling price based on this independent variable.
- Multiple Regression: Like linear regression, but multiple regression allows evaluating the influence of multiple independent variables on a dependent variable, helping in making more complex decisions.
    - For example: In financial analysis, you can use multivariate regression to predict a company's profits based on many factors such as revenue, costs, and tax rates.
- Logistic Regression: In prescriptive analysis, logistic regression can be used to predict the probability of a binary dependent variable (yes or no) based on the independent variables of categorical events, thereby assisting in making strategic decisions.
    - For example, in the medical field, logistic regression can be used to predict the probability that a patient will develop cancer based on factors such as age, gender, and family history.

(Funix, 2021) (Hạnh, 2024)

State that the prescriptive analytic methods are appropriate to find the total cost of meeting demand over the next year.

Linear regression and multiple regression methods are suitable in prescriptive analysis to predict total demand response costs in the coming year because these models can describe the relationship between independent variables (such as the number quantity of lawn mowers and tractors sold, production costs, Operating and Interest Expenses) and dependent variable (total cost) by using excel tools such as pivot table, pivot chart, trendline to calculate, total Aggregate and visualize data to see relationships, dependencies, and trends. Besides, I use Excel's data analysis tool to come up with a specific formula for total costs, thereby predicting the total costs needed for the next year and analyzing the effects of different factors on costs. costs, thereby optimizing resources. Prescriptive analytics using regression helps make decisions based on data and predictions, supporting effective financial management.

P8 Demonstrate these methods using an appropriate programming language or tool.

Demonstrate these methods using an appropriate programming language or tool and produce a linear optimization model for minimizing the total cost of meeting demand over the next year.

From the dataset provided by the company, I compiled and calculated the number of lawn mowers and tractors sold in the world, annual revenue and usage costs from 2010 to 2014.

Year	Sum of World lawn mower unit.
2010	110520
2011	108490
2012	109270
2013	110560
2014	110043
Grand Total	**548883**

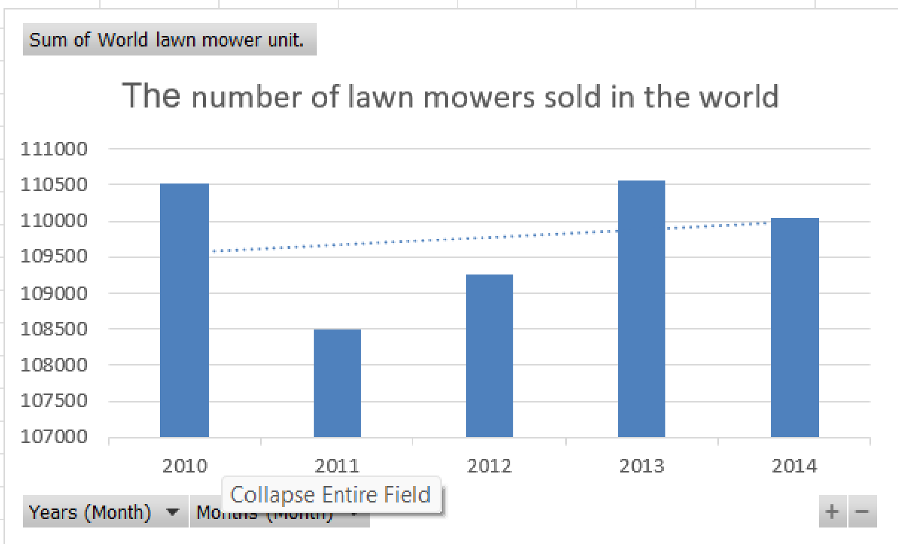


Figure 17 The number of lawnmowers sold in the world 1.

From the chart above, we see that the number of lawn mowers sold in each year is not equal, there is a sharp decrease from 2011 to 2012 and a slight increase in the following years.

Year	Sum of World lawn mower unit.
⊟2010	110520
Jan	7020
Feb	9280
Mar	9780
Apr	11100
May	11930
Jun	12240
Jul	10740
Aug	10080
Sep	8430
Oct	7650
Nov	6650
Dec	5620
⊟2011	108490
Jan	7020
Feb	9030
Mar	10050
Apr	10890

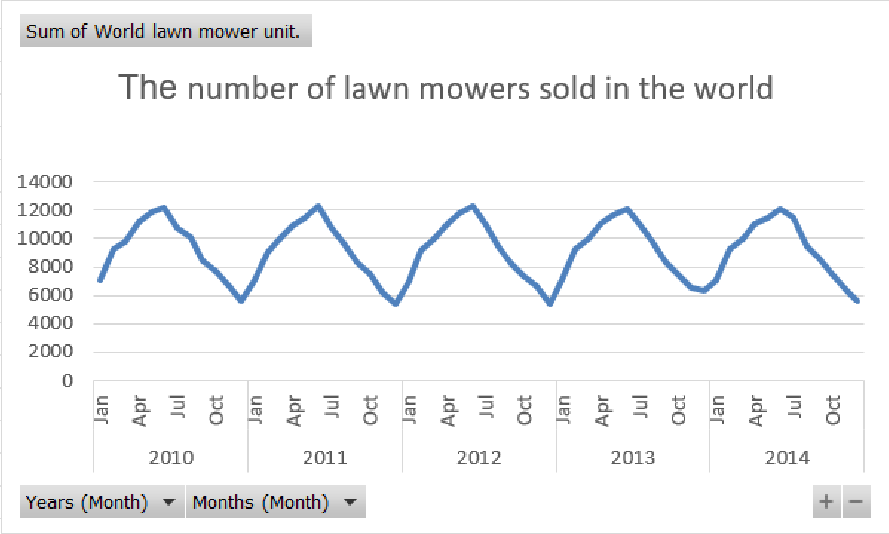


Figure 18 The number of lawnmowers sold in the world 2.

Besides, from the chart above we see that the number of lawn mowers sold changes by month and has a certain cycle. Specifically, customers will buy lawn mowers in large quantities from April 4 to the peak in July every year. This is when the weather is usually warm and grass will thrive. Based on this chart, the company should increase the quantity of lawn mowers produced from March to July to meet and ensure availability for sale.

Year	The number of tractors sold in the world
⊞2010	20909
⊞2011	24970
⊞2012	29593
⊞2013	36745
⊞2014	46194
Grand Total	**158411**

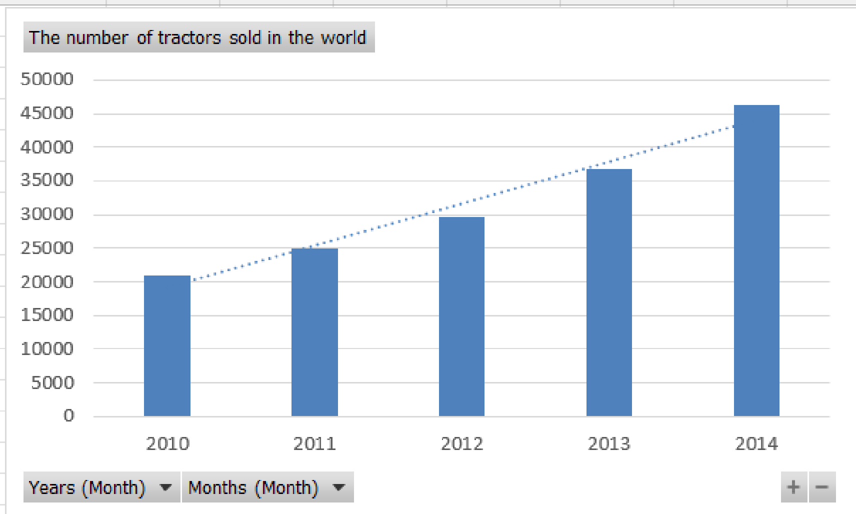


Figure 19 The number of tractors sold in the world 1.

Unlike lawn mowers, from the chart we can see that the number of tractors sold tends to increase steadily over the years.

Year	The number of tractors sold in the world
⊟2010	20909
Jan	1592
Feb	1711
Mar	1810
Apr	1867
May	1779
Jun	1740
Jul	1826
Aug	1695
Sep	1681
Oct	1663
Nov	1825
Dec	1720
⊟2011	24970
Jan	1761
Feb	2035
Mar	2142
Apr	2340
May	2280
Jun	2271
Jul	2154

Figure 20 The number of tractors sold in the world 2.

And it also follows a certain cycle, increasing from January to July and then gradually decreasing to January of the following year.

Month	Unit Production Cost of Tractor	The number of tractors sold	Revenue from tractors	Production cost of tractor	Unit Production Cost of Mower	The number of lawn mowers sold	Revenue from lawn mowers	Production cost of mower	Total revenue	Total production cost	Administrative expenses	Depreciation expenses	Interest expenses	Tractor Transmission Cost	Shipping cost	Total cost
Jan-10	$1,750	1592	$ 5,174,000	$ 2,786,000	$50	7020	$ 1,053,000	$ 351,000	$ 6,227,000	$ 3,137,000	$633,073	$140,467	$7,244	$ 4,616.80	$ 13,548.28	$ 3,935,949
Feb-10	$1,755	1711	$ 5,560,750	$ 3,002,805	$50	9280	$ 1,392,000	$ 464,000	$ 6,952,750	$ 3,466,805	$607,904	$165,636	$7,679	$ 4,961.90	$ 17,146.54	$ 4,270,132
Mar-10	$1,763	1810	$ 5,882,500	$ 3,191,030	$51	9780	$ 1,467,000	$ 498,780	$ 7,349,500	$ 3,689,810	$630,687	$142,853	$6,887	$ 5,249.00	$ 18,083.60	$ 4,493,570
Apr-10	$1,770	1867	$ 6,067,750	$ 3,304,590	$51	11100	$ 1,665,000	$ 566,100	$ 7,732,750	$ 3,870,690	$613,401	$160,139	$6,917	$ 5,414.30	$ 20,160.98	$ 4,676,722
May-10	$1,778	1779	$ 5,781,750	$ 3,163,062	$51	11930	$ 1,789,500	$ 608,430	$ 7,571,250	$ 3,771,492	$607,664	$165,876	$8,316	$ 5,159.10	$ 21,226.96	$ 4,579,734
Jun-10	$1,785	1740	$ 5,655,000	$ 3,105,900	$51	12240	$ 1,836,000	$ 624,240	$ 7,491,000	$ 3,730,140	$632,967	$140,573	$7,428	$ 5,046.00	$ 21,613.20	$ 4,537,767
Jul-10	$1,792	1826	$ 5,934,500	$ 3,272,192	$51	10740	$ 1,611,000	$ 547,740	$ 7,545,500	$ 3,819,932	$609,604	$163,936	$8,737	$ 5,295.40	$ 19,545.04	$ 4,627,049
Aug-10	$1,795	1695	$ 5,508,750	$ 3,042,525	$51	10080	$ 1,512,000	$ 514,080	$ 7,020,750	$ 3,556,605	$607,749	$165,791	$7,054	$ 4,915.50	$ 18,307.50	$ 4,360,422
Sep-10	$1,801	1681	$ 5,463,250	$ 3,027,481	$52	8430	$ 1,264,500	$ 438,360	$ 6,727,750	$ 3,465,841	$603,367	$170,173	$8,862	$ 4,874.90	$ 15,821.84	$ 4,268,940
Oct-10	$1,804	1663	$ 5,404,750	$ 3,000,052	$52	7650	$ 1,147,500	$ 397,800	$ 6,552,250	$ 3,397,852	$629,083	$144,457	$8,488	$ 4,822.70	$ 14,624.72	$ 4,199,327
Nov-10	$1,810	1825	$ 5,931,250	$ 3,303,250	$52	6650	$ 997,500	$ 345,800	$ 6,928,750	$ 3,649,050	$611,995	$161,545	$7,049	$ 5,292.50	$ 13,449.00	$ 4,448,381
Dec-10	$1,813	1720	$ 5,590,000	$ 3,118,360	$52	5620	$ 843,000	$ 292,240	$ 6,433,000	$ 3,410,600	$625,712	$147,828	$8,807	$ 4,988.00	$ 11,710.60	$ 4,209,646
Jan-11	$1,835	1761	$ 5,987,400	$ 3,231,435	$55	7020	$ 1,228,500	$ 386,100	$ 7,215,900	$ 3,617,535	$656,123	$175,447	$7,430	$ 5,106.90	$ 13,876.14	$ 4,475,518
Feb-11	$1,841	2035	$ 6,919,000	$ 3,746,435	$55	9030	$ 1,580,250	$ 496,650	$ 8,499,250	$ 4,243,085	$652,679	$178,891	$6,791	$ 5,901.50	$ 17,402.60	$ 5,104,750
Mar-11	$1,848	2142	$ 7,282,800	$ 3,958,416	$55	10050	$ 1,758,750	$ 552,750	$ 9,041,550	$ 4,511,166	$655,521	$176,049	$8,013	$ 6,211.80	$ 19,129.98	$ 5,376,091
Apr-11	$1,854	2340	$ 7,956,000	$ 4,338,360	$55	10890	$ 1,905,750	$ 598,950	$ 9,861,750	$ 4,937,310	$676,581	$154,989	$8,979	$ 6,786.00	$ 20,765.70	$ 5,805,411
May-11	$1,860	2280	$ 7,752,000	$ 4,240,800	$56	11420	$ 1,998,500	$ 639,520	$ 9,750,500	$ 4,880,320	$676,581	$154,989	$7,484	$ 6,612.00	$ 21,439.00	$ 5,747,425
Jun-11	$1,866	2271	$ 7,721,400	$ 4,237,686	$56	12270	$ 2,147,250	$ 687,120	$ 9,868,650	$ 4,924,806	$656,440	$175,130	$7,858	$ 6,585.90	$ 22,688.04	$ 5,793,508
Jul-11	$1,872	2154	$ 7,323,600	$ 4,032,288	$56	10720	$ 1,876,000	$ 600,320	$ 9,199,600	$ 4,632,608	$661,969	$169,601	$7,424	$ 6,246.60	$ 20,151.56	$ 5,498,000
Aug-11	$1,878	2146	$ 7,296,400	$ 4,030,188	$56	9650	$ 1,688,750	$ 540,400	$ 8,985,150	$ 4,570,588	$677,212	$154,358	$6,848	$ 6,223.40	$ 18,541.74	$ 5,433,771
Sep-11	$1,885	2085	$ 7,089,000	$ 3,930,225	$56	8310	$ 1,454,250	$ 465,360	$ 8,543,250	$ 4,395,585	$653,545	$178,025	$6,751	$ 6,046.50	$ 16,426.80	$ 5,256,379
Oct-11	$1,892	1970	$ 6,698,000	$ 3,727,240	$57	7510	$ 1,314,250	$ 428,070	$ 8,012,250	$ 4,155,310	$657,388	$174,182	$8,160	$ 5,713.00	$ 15,011.70	$ 5,015,765
Nov-11	$1,897	1936	$ 6,582,400	$ 3,672,592	$57	6250	$ 1,093,750	$ 356,250	$ 7,676,150	$ 4,028,842	$672,475	$159,095	$7,898	$ 5,614.40	$ 13,068.34	$ 4,886,993
Dec-11	$1,903	1850	$ 6,290,000	$ 3,520,550	$57	5370	$ 939,750	$ 306,090	$ 7,229,750	$ 3,826,640	$656,325	$175,245	$8,953	$ 5,365.00	$ 11,590.30	$ 4,684,118
Jan-12	$1,925	2000	$ 7,200,000	$ 3,850,000	$59	6970	$ 1,254,600	$ 411,230	$ 8,454,600	$ 4,261,230	$723,594	$226,526	$9,443	$ 5,800.00	$ 14,265.30	$ 5,240,858

Figure 21 Necessary revenue and expenses each month.

Next, we created a worksheet to summarize and calculate total revenue, Total production cost, Administrative expenses, Depreciation expenses, Interest expenses, Tractor Transmission Cost, Shipping cost, Total cost.

Year	Total revenue in world	Total production cost in world	Administrative expenses in world	Depreciation expenses in world	Interest expenses in world	Tractor Transmission Cost in world	Shipping cost in world	Total cost in world
⊞2010	$ 84,532,250.00	$ 42,965,817.00	$ 7,413,206.00	$ 1,869,274.00	$ 93,468.00	$ 60,636.10	$ 205,238.26	$ 52,607,639.36
⊞2011	$ 103,883,750.00	$ 52,723,795.00	$ 7,952,839.00	$ 2,026,001.00	$ 92,589.00	$ 72,413.00	$ 210,091.90	$ 63,077,728.90
⊞2012	$ 126,203,400.00	$ 64,516,024.00	$ 8,917,696.00	$ 2,483,744.00	$ 110,277.00	$ 85,819.70	$ 220,222.72	$ 76,333,783.42
⊞2013	$ 156,410,100.00	$ 79,197,693.00	$ 7,964,649.00	$ 2,251,791.00	$ 121,465.00	$ 106,560.50	$ 236,019.70	$ 89,878,178.20
⊞2014	$ 196,445,370.00	$ 104,128,027.00	$ 7,738,943.00	$ 2,486,977.00	$ 117,855.00	$ 133,962.60	$ 253,580.43	$ 114,859,345.03
Grand Total	**$ 667,474,870.00**	**$ 343,531,356.00**	**$ 39,987,333.00**	**$ 11,117,787.00**	**$ 535,654.00**	**$ 459,391.90**	**$ 1,125,153.01**	**$ 396,756,674.91**

Figure 22 Necessary revenue and expenses each year.

SUMMARY OUTPUT

Regression Statistics	
Multiple R	1
R Square	1
Adjusted R Square	65535
Standard Error	0
Observations	5

ANOVA

	df	*SS*	*MS*	*F*	*ignificance F*
Regression	6	2.36E+15	3.93E+14	#NUM!	#NUM!
Residual	0	0	65535		
Total	6	2.36E+15			

	Coefficients	*andard Err*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*ower 95.0%*	*Upper 95.0%*
Intercept	174055.1603	0	65535	#NUM!	#VALUE!	#VALUE!	174055.2	174055.1603
Total production cost in world	1.0020	0	65535	#NUM!	1.00197	1.00197	1.00197	1.001970368
Administrative expenses in world	0.9994	0	65535	#NUM!	0.999431	0.999431	0.999431	0.999430808
Depreciation expenses in world	0.9875	0	65535	#NUM!	0.987481	0.987481	0.987481	0.987481275
Interest expenses in world	1.3721	0	65535	#NUM!	1.372119	1.372119	1.372119	1.372118665
Tractor Transmission Cost in world	0	0	65535	#NUM!	0	0	0	0
Shipping cost in world	0	0	65535	#NUM!	0	0	0	0

total cost = 174055.1603 + 1.002* Total production cost + 0.9994*Administrative expenses +0.9875*Depreciation expenses + 1.3721*Interest expenses.

Figure 23 Multivariate regression analysis of related factors and prediction of necessary costs.

Next, we used the data analysis tool to conduct multivariate regression analysis between the dependent variable Total Cost, and independent variables including production cost, administrative fees, depreciation fees, interest rates, and transmission fees to produce tractor gearboxes and shipping costs. From there we came up with the formula shown above.

To predict the costs needed to meet demand for the next year (2015), we forecast the number of lawn mowers and tractors sold, production prices, and related costs using the method of linear regression analysis to come up with the formula.

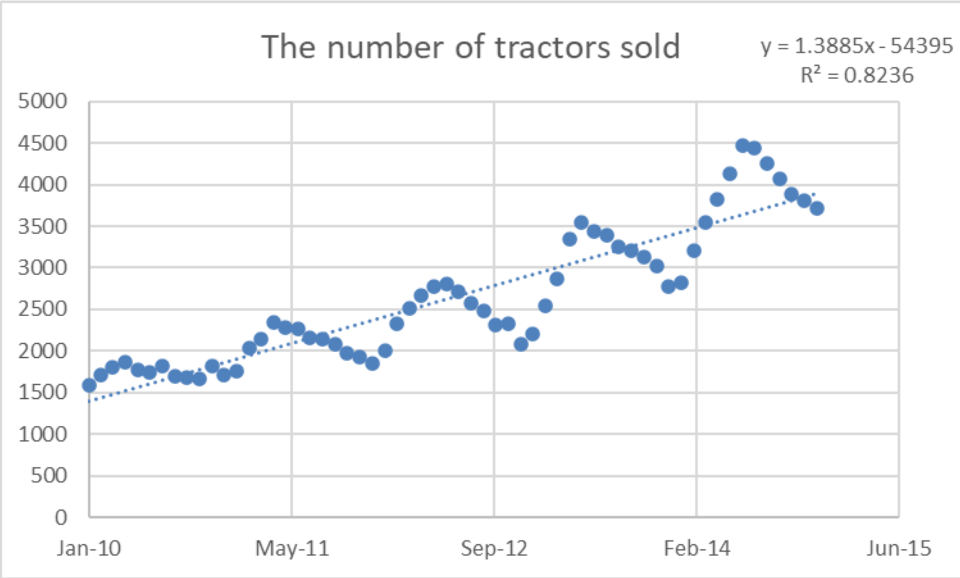


Figure 24 Linear regression analysis of tractor sales.

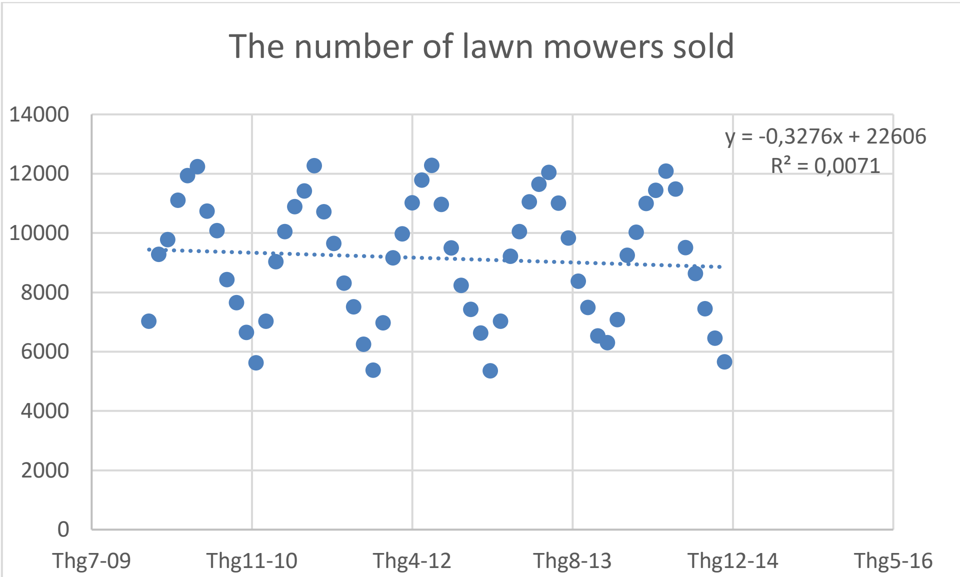


Figure 25 Linear regression analysis of lawn mower sales.

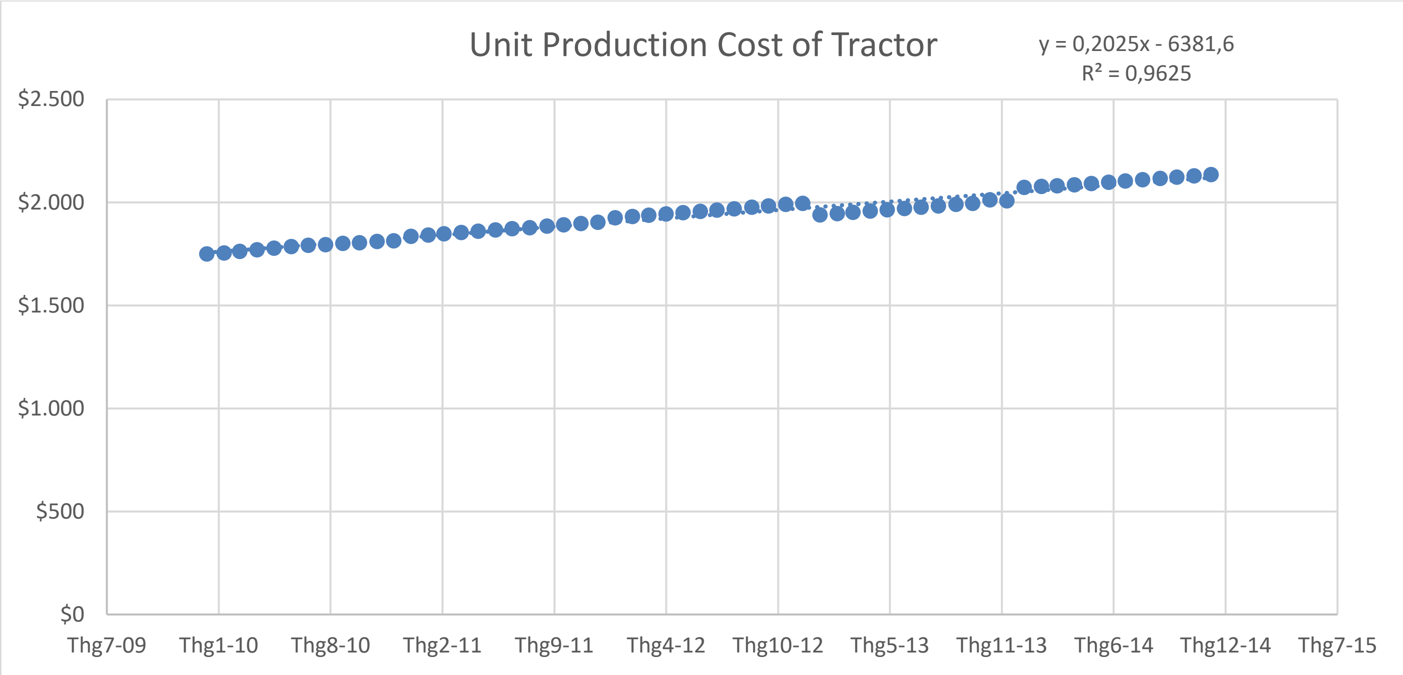


Figure 26 Linear regression analysis of unit production cost of tractor.

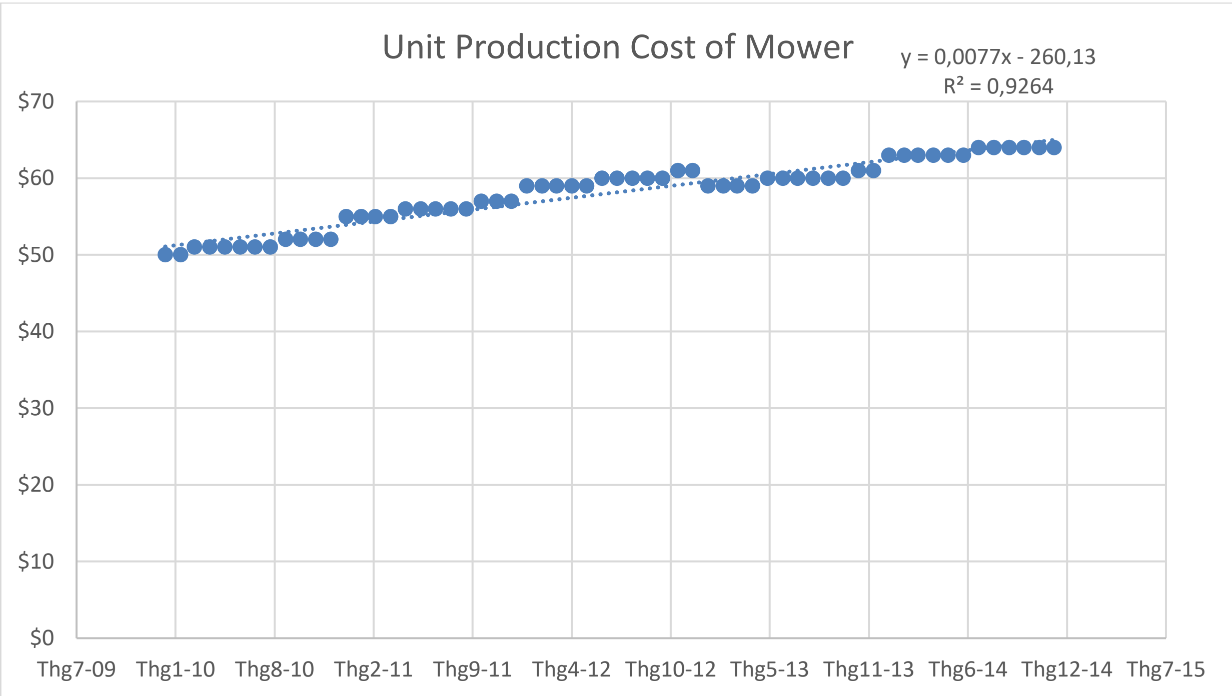


Figure 27 Linear regression analysis of unit production cost of mower.

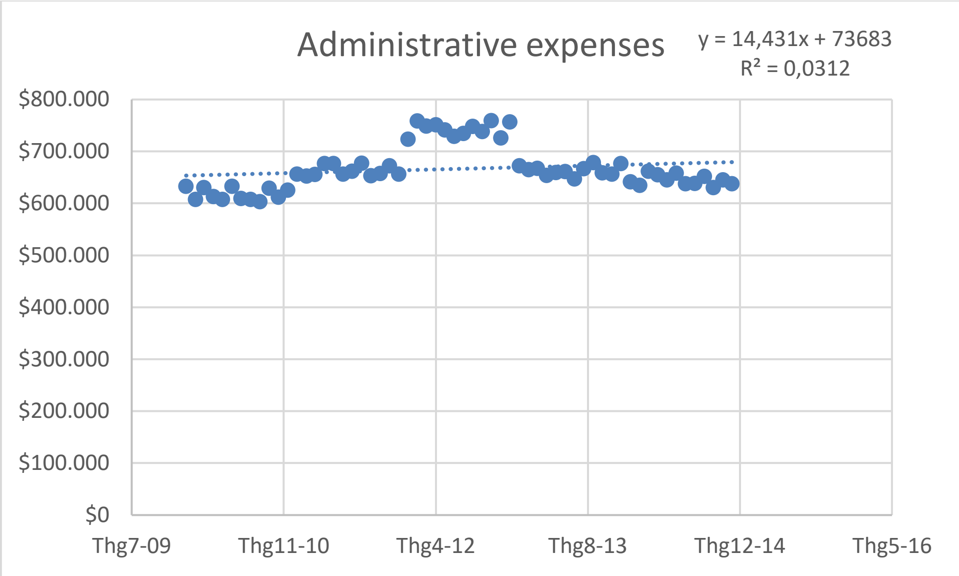


Figure 28 Linear regression analysis of Administrative expenses.

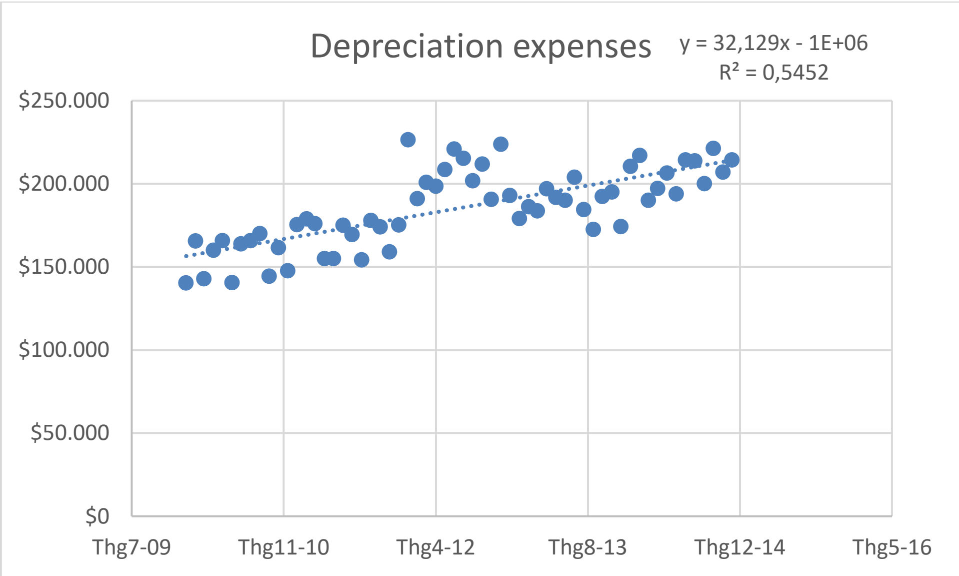


Figure 29 Linear regression analysis of Depreciation expenses.

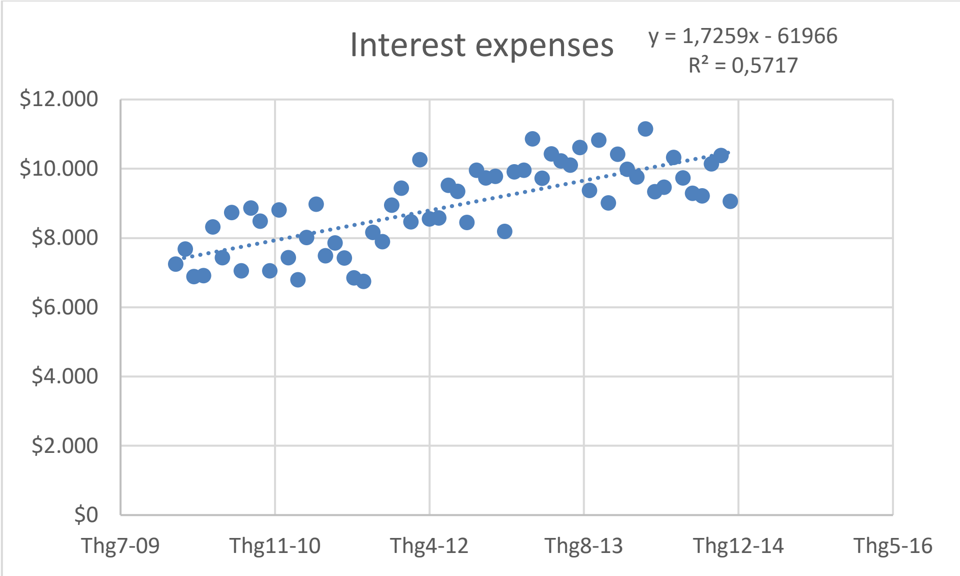


Figure 30 Linear regression analysis of interest expenses.

Applying the formula we have predicted the costs shown in the table below.

Month	Unit Production Cost of Tractor	The number of tractors sold	Revenue from tractors	Production cost of tractor	Unit Production Cost of Mower	The number of lawn mowers sold	Revenue from lawn mowers	Production cost of mower	Total revenue	Total production cost	Administrative expenses	Depreciation expenses	Interest expenses
Jan-15	$2,124	3929		$ 8,346,695	$63	8845		$ 559,974		$ 8,906,668	$679,857	$349,579	$10,530
Feb-15	$2,131	3972		$ 8,463,071	$64	8835		$ 561,440		$ 9,024,511	$680,305	$350,575	$10,584
Mar-15	$2,136	4011		$ 8,568,649	$64	8826		$ 562,760		$ 9,131,409	$680,709	$351,474	$10,632
Apr-15	$2,143	4054		$ 8,686,054	$64	8816		$ 564,217		$ 9,250,271	$681,156	$352,470	$10,686
May-15	$2,149	4096		$ 8,800,186	$64	8806		$ 565,622		$ 9,365,808	$681,589	$353,434	$10,738
Jun-15	$2,155	4139		$ 8,918,655	$64	8796		$ 567,069		$ 9,485,724	$682,036	$354,430	$10,791
Jul-15	$2,161	4180		$ 9,033,816	$65	8786		$ 568,465		$ 9,602,281	$682,469	$355,394	$10,843
Aug-15	$2,167	4223		$ 9,153,347	$65	8776		$ 569,903		$ 9,723,250	$682,917	$356,390	$10,896
Sep-15	$2,174	4266		$ 9,273,419	$65	8766		$ 571,335		$ 9,844,755	$683,364	$357,386	$10,950
Oct-15	$2,180	4308		$ 9,390,133	$65	8756		$ 572,717		$ 9,962,850	$683,797	$358,350	$11,002
Nov-15	$2,186	4351		$ 9,511,268	$66	8746		$ 574,141		$ 10,085,409	$684,244	$359,346	$11,055
Dec-15	$2,192	4393		$ 9,629,010	$66	8736		$ 575,513		$ 10,204,524	$684,677	$360,310	$11,107

Figure 31 Necessary expenses per month in 2015.

Row Labels	Sum of Total production cost	Sum of Administrative expenses	Sum of Depreciation expenses	Sum of Interest expenses
2010	$42,965,817	$7,413,206	$1,869,274	$93,468
2011	$52,723,795	$7,952,839	$2,026,001	$92,589
2012	$64,516,024	$8,917,696	$2,483,744	$110,277
2013	$79,197,693	$7,964,649	$2,251,791	$121,465
2014	$104,128,027	$7,738,943	$2,486,977	$117,855
2015	$114,587,459	$8,187,119	$4,259,137	$129,814
Grand Total	**$458,118,815**	**$48,174,452**	**$15,376,924**	**$665,468**
total cost = 174055.1603 + 1.002* Total production cost + 0.9994*Administrative expenses +0.9875*Depreciation expenses + 1.3721*Interest expenses.				
the total cost of meeting demand in 2015	$127,556,911.17			

After regression analysis, we have predicted the total cost of meeting demand in 2015 to be $127,556,911.17.

Also from the above formula, we see that **if we want to minimize total costs to meet demand in 2015, we need to minimize production costs, administrative expenses, depreciation expenses, and interest expenses.**

Year	Total revenue in world	Total cost in world	
2010	$ 84,532,250.00	$ 52,607,639.36	
2011	$ 103,883,750.00	$ 63,077,728.90	
2012	$ 126,203,400.00	$ 76,333,783.42	
2013	$ 156,410,100.00	$ 89,878,178.20	
2014	$ 196,445,370.00	$ 114,859,345.03	
Grand Total	**$ 667,474,870.00**	**$ 396,756,674.91**	
Profits earned in the 5 years from 2010 to 2014			**$270,718,195.09**

Figure 32 Profits earned in the 5 years from 2010 to 2014

The table above shows that the company's total profit is $270,718,195.09. From this source of money the company can:

- Pay off part of the bank loan to thereby lower interest expenses.
- Invest in expanding existing factories or building new facilities, improving and enhancing production efficiency thereby reducing production costs.

Fixed Costs of Capacity Increase or New Construction		
Current Plants	**Additional Capacity**	**Cost**
Kansas City	10000	$605,000.00
Kansas City	20000	$985,000.00
Santiago	5000	$381,000.00
Santiago	10000	$680,000.00
Proposed Locations	**Maximum capacity**	**Cost**
Auckland	15,000	$917,000.00
Auckland	20,000	$1,136,000.00
Birmingham	15,000	$962,000.00
Birmingham	20,000	$1,180,000.00
Frankfurt	15,000	$874,000.00
Frankfurt	20,000	$1,093,000.00
Mumbai	15,000	$750,000.00
Mumbai	25,000	$959,000.00
Singapore	15,000	$839,000.00
Singapore	20,000	$1,058,000.00
total to capacity increase	$2,651,000	
total to new construction	$9,768,000	
total	$12,419,000	

Figure 33 Fixed Costs of Capacity Increase or New Construction.

- Replace the current process used to produce tractor gearboxes with proposed process A to reduce Transmission Costs.

Average of Current	Average of Process A	Average of Process B
$289.60	$285.50	$298.43

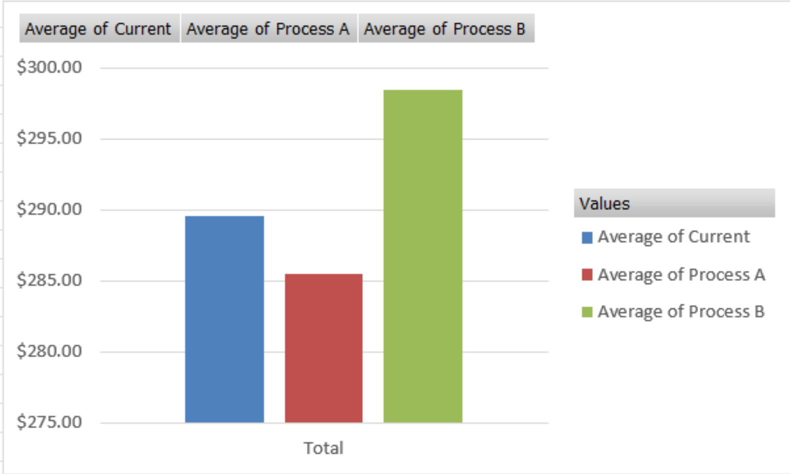


Figure 34 Average cost of the processes used to produce tractor gearboxes.

➔ Leading to minimizing the total cost of meeting demand in 2015.

Compare a range of predictive analytical techniques for forecasting purposes. (M3)

State the range of predictive analytical techniques

What is Forecasting?

Process of predicting a future event based on historical data.

A statement about the future value of a variable of interest such as demand.

Forecasting is used to make informed decisions.

The underlying basis of all business decisions

- Production
- Inventory
- Personnel
- Facilities

Types of Forecasts

Business analysts may choose from a wide range of forecasting techniques to support decision making.

Selecting the appropriate method depends on the characteristics of the forecasting problem, such as the time horizon of the variable being forecast, as well as available information on which the forecast will be based.

Three major categories of forecasting approaches are

Qualitative and judgmental techniques:

Qualitative and judgmental techniques rely on experience and intuition. They are necessary when historical data are not available or when the decision maker needs to forecast far into the future. Another use of judgmental methods is to incorporate nonquantitative information, such as the impact of government regulations or competitor behavior, in a quantitative forecast.

For example: Predicting the Price of Oil

- Early 1988 oil price was about $22 a barrel.
- Mid 1988 oil price dropped to $11 a barrel because of oversupply, high production in non-OPEC regions, and lower than normal demand.
- In the past, OPEC would raise the price of oil.

- Historical analogy would forecast a higher price.
- However, the price continued to drop even though OPEC agreed to cut production.
- Historical analogies cannot always account for current realities

Figure 35: Predicting the price of oil

Short-term and long-term forecasting and decision-makers

Qualitative and judgmental forecasting models are often utilized for both short-range and long-range forecasts in different situations, depending on the nature of the forecast and the required level of confidence. and decisions regarding the use of qualitative and judgmental forecasting models are typically made by senior managers and relevant experts according to specific situations. Here's how they are typically applied

a) Short-term forecasting

- **How it is used:** In situations where factors are rapidly changing and difficult to predict, but short-term forecasts are needed to manage resources and production planning, qualitative and judgmental forecasting models are often preferred. For example, businesses may use them to predict market supply and demand dynamics in the next few months or to estimate the amount of products needed for production in the short term.
- **Who decides:** In short-term forecasting, managers and teams directly work with short-term data and are responsible for production management, resource planning, and supply chain management. Therefore, decisions regarding the use of qualitative and judgmental forecasting models in short-term forecasting are often made by senior managers in production, planning, or finance departments, based on business needs and information from experts and data analysis teams.

b) Long-term forecasting:

- **How it is used:** In situations where forecasts are needed for long-term strategic decisions such as business development plans, major investments, or product strategies, qualitative and judgmental forecasting models can be employed. For instance, when planning to expand into new markets or develop new product lines, managers may rely on expert opinions and subjective assessments to forecast long-term demand and trends, which may extend from one year to several years or even decades

- **Who decides:** In long-term forecasting, decisions are usually made at a higher strategic level and often relate to the long-term strategic decisions of the company, such as product development plans, market expansion, or major investments. In this case, decisions regarding the use of qualitative and judgmental forecasting models are typically made by senior managers such as chief executives or other top-level leaders, as they have a broader perspective and the ability to evaluate the long-term impact of strategic decisions.

Statistical time-series models

Statistical time-series models find greater applicability for short-range forecasting problems. A time series is a stream of historical data, such as weekly sales. The values of a time series over T periods as At, t = 1, 2, …, T. Time-series models assume that whatever forces have influenced sales in the recent past will continue into the near future; thus, forecasts are developed by extrapolating these data into the future. Time series generally have one or more of the following components:

- random behavior
- trends
- seasonal effects
- cyclical effects

For example: Identifying Trends in a Time Series

- Total Energy Production & Consumption.
- The general upward trend with some short downward trends.
- The time series is composed of several different short trends.

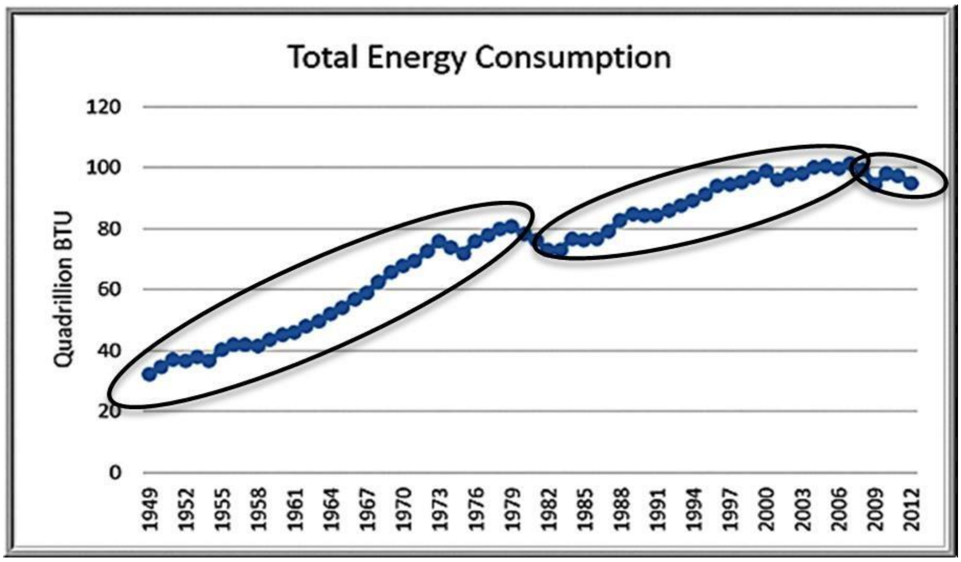


Figure 36: Total Energy Consumption Time Series

Short-term and long-term forecasting and decision-makers

Statistical time-series models can be used for both short-range and long-range forecasts, depending on the nature of the data and the specific forecasting objectives. In both short-term and long-term forecasting processes, decisions regarding the use of statistical time-series models are typically made by appropriate managers and experts based on each specific situation.

a) Short-term:

- **How it is used:** Production management is used to forecast the demand for raw materials, resources, and labor in the short term, aiding production managers in efficient planning for daily or weekly production processes. Supply and demand forecasting is employed to predict short-term trends in supply and demand in the market, assisting in inventory management and price adjustments or marketing strategies. Transportation and logistics planning is utilized to forecast anticipated shipment volumes and transportation schedules in the near term, optimizing transportation and delivery operations.
- **Who decides:** Decision-making regarding the use of statistical time-series models is typically made by production managers, planners, and supply chain managers. They need to forecast specific variables in the near term, such as raw material demand, production schedules, or transportation and logistics factors. These decisions are often based on specific business needs, input from experts and data analysis teams, as well as the reliability and accuracy of the forecasting models.

b) Long-term:

- **How it is used:** Business development planning is used to provide forecasts of market demand over the long term, supporting decisions regarding business expansion, new product development, or market expansion. Financial strategy is employed to forecast future revenue, profitability, and cash flow for long-term financial planning, including investment decisions, financing, and management of financial resources. Human resource management is utilized to forecast workforce needs and develop long-term personnel plans, assisting in building effective and sustainable human resource strategies.
- **Who decides:** Decisions are usually made at a higher strategic level and relate to the long-term strategic decisions of the company. Decisions regarding the use of statistical time-series models in long-range forecasting are often made by senior leaders such as CEOs or other high-level managers. They have a broader perspective and the ability to assess the long-term impact of strategic decisions on the business. These decisions are typically based on the strategic goals of the business, such as business development plans, financial management, or human resource management, and also rely on detailed data analysis and forecasting.

Explanatory/causal methods

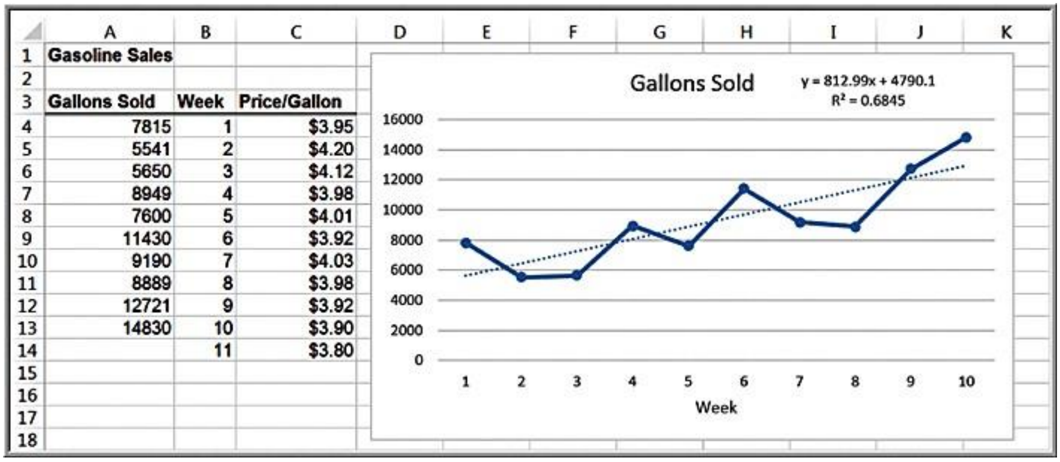
Explanatory/causal models, often called econometric models, seek to identify factors that explain statistically the patterns observed in the variable being forecast, usually with regression analysis. In many forecasting applications, other independent variables besides time, e.g., economic indexes or demographic factors, may influence the time series.

For example: Forecasting Gasoline Sales Using Simple Linear Regression

Excel file Gasoline Sales

Simple trendline using week as the independent variable. We would predict sales for week 11:

Sale = 4,790.1 + 812.99(11) = 13,733 Gallons



Gallons Sold	Week	Price/Gallon
7815	1	$3.95
5541	2	$4.20
5650	3	$4.12
8949	4	$3.98
7600	5	$4.01
11430	6	$3.92
9190	7	$4.03
8889	8	$3.98
12721	9	$3.92
14830	10	$3.90
	11	$3.80

Figure 37: Gallons Sold

Short-term and long-term forecasting and decision-makers

Explanatory/Causal Methods can be used for both short-term and long-term forecasting. The decision-makers for both short-term and long-term forecasting typically consist of management executives or management teams within an organization. Specifically:

a) Short-term

- **How it is used:** For short-term forecasting, Explanatory/Causal Methods can be used to identify and forecast short-term fluctuations in data, helping to manage minor fluctuations and respond quickly to changes.
- **Who decides:** Short-term forecasting decisions are often made by senior managers or specific departmental teams, such as production, sales, or finance management. These decisions may involve adjusting production, ordering, or short-term cash flow management to respond to short-term market fluctuations or customer demand.

b) Long-term

- **How it is used:** For long-term forecasting, these methods can be used to forecast long-term trends and the impact of long-term economic, social, or political factors on dependent variables, assisting in strategic decision-making and long-term planning.
- **Who decides:** Long-term forecasting decisions are typically made by higher-level executives or strategic planning departments. These decisions often relate to long-term business strategies, financial investments, product development, market expansion, or significant resource and personnel decisions.

Explain how the chosen of the appropriate forecasting model to forecast the sale number of mowers and tractors to plan manufacturing capacity.

Explain the choice of forecasting model

In forecasting and data analysis, selecting appropriate mathematical functions is extremely important for building accurate and effective forecasting models. Below is the importance of mathematical functions in this process:

- **Modeling Complex Relationships:** Mathematical functions provide the foundation for modeling complex relationships between variables. By selecting the appropriate mathematical function, we can accurately reflect the relationship between variables in the data and create accurate forecasting models.
- **Model Testing and Evaluation:** Mathematical functions help test and evaluate the performance of forecasting models. With evaluation metrics such as R-squared, mean squared error, we can assess the accuracy of the model and adjust it if necessary.
- **Forecasting and Risk Management:** Mathematical functions allow forecasting future values based on current data. This aids in risk management and making informed business decisions based on accurate forecasts.
- **Model Diversification and Testing:** By using various types of mathematical functions, we can diversify and test the performance of models. This helps us find the most suitable model for the data and specific objectives.
- **Enhanced Understanding of Data:** Mathematical functions help us gain a deeper understanding of the relationships between variables in the data. This provides valuable information to support business decisions and strategy development.

To explain the selection of an appropriate forecasting model for predicting the sales volume of mowers and tractors to plan production capacity, I will apply the method of Modeling Relationships and Trends in Data. Additionally, I will utilize various common mathematical functions for forecast analysis, as the data for Industry Mower Total Sales and Industry Tractor Total Sales are time-series data. Therefore, I will employ the trendline tool to utilize line charts. This facilitates accurate calculations and quicker application of mathematical formulas. Consequently, it provides us with a comprehensive overview and aids in devising the most optimal and suitable production capacity plan for each region.

Explain how to apply the model

To discuss the application of modeling relationships and trends in data, I will employ mathematical formulas to forecast the sales volume of the Industry Mower and Tractor Total Sales tables for PLE company, as well as forecast future increases in production costs. Upon applying the formulas to the data, the trendline tool will provide us with constants, independent variables, and dependent variables for each formula. Consequently, we can forecast the number of mowers and tractors sold in the upcoming year more accurately.

To select the appropriate mathematical function for each data type, I will simultaneously apply three different types of mathematical functions to a dataset to determine the most suitable model for that data. However, as per the requirements of the task, we must select the appropriate forecasting model for each region. Therefore, I will choose five regions where mowers and tractors are sold within the data for analysis. These regions are North America (NA), South America (SA), Europe (Eur), the Pacific Rim (Pac), and China.

NA region

Firstly, based on the Industry Mower Total Sales and Industry Tractor Total Sales datasets, I have created a table containing information on the total sales of mowers and tractors over five years. From there, I will model with three appropriate mathematical functions for the year 2010 to assess the accuracy of each model and evaluate which model is the most suitable.

Year	NA of Mower	NA of Tractor
2010	880392	81516
2011	853795	75474
2012	863023	78501
2013	867283	97342
2014	890359	130743

Figure 38: NA regression(1)

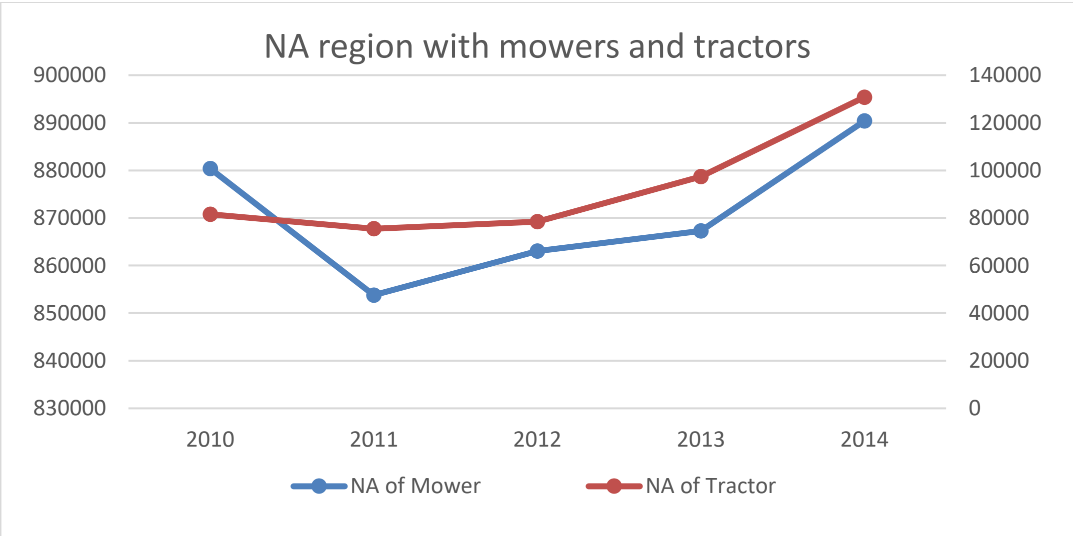


Figure 39: NA region (2)

Based on the data and chart, I will utilize the trendline tool to apply three mathematical methods of modeling relationships and trends to find the appropriate computational approach for the data. These methods are:

1. **Linear function:**
   Linear of tractor: $y = 12032x + 56618$

   Linear of mower: $y = 3342,1x + 860944$
2. **Polynomial function:**
   Polynomial of tractor: $y = 6764,2x2 - 28553x + 103968$

   Polynomial of mower: $y = 6741,3x2 - 37106x + 908134$
3. **Logarithmic function:**
   Logarithmic of tractor: $y = 24792ln(x) + 68977$

   Logarithmic of mower: $y = 3377,5ln(x) + 867737$

	Linear	Polynomial	logarithmic
Mower	864.286	877.769	867.737
Tractor	68.650	82.179	68.977

Sample		
Year	NA of Mower	NA of Tractor
2010	880.392	81.516

Figure 40: Sample NA region

➔ Based on the computational results of the NA region data with mowers and tractors, we conclude that the most suitable computational method for the data is the Polynomial method for both mower and tractor data. This conclusion is drawn because the Polynomial method yields forecast results that closely match the sample data. Therefore, we will select the Polynomial method for the given data.

SA region

Secondly, based on the Industry Mower Total Sales and Industry Tractor Total Sales datasets, I have created a table containing information on the total sales of mowers and tractors over five years for the SA region. From there, I will model with three appropriate mathematical functions for the year 2010 to assess the accuracy of each model and evaluate which model is the most suitable.

	SA of mower	SA of tractor
2010	8236	12425
2011	8069	18866
2012	8181	26068
2013	7900	31497
2014	8165	36719

Figure 41: SA regression(1)

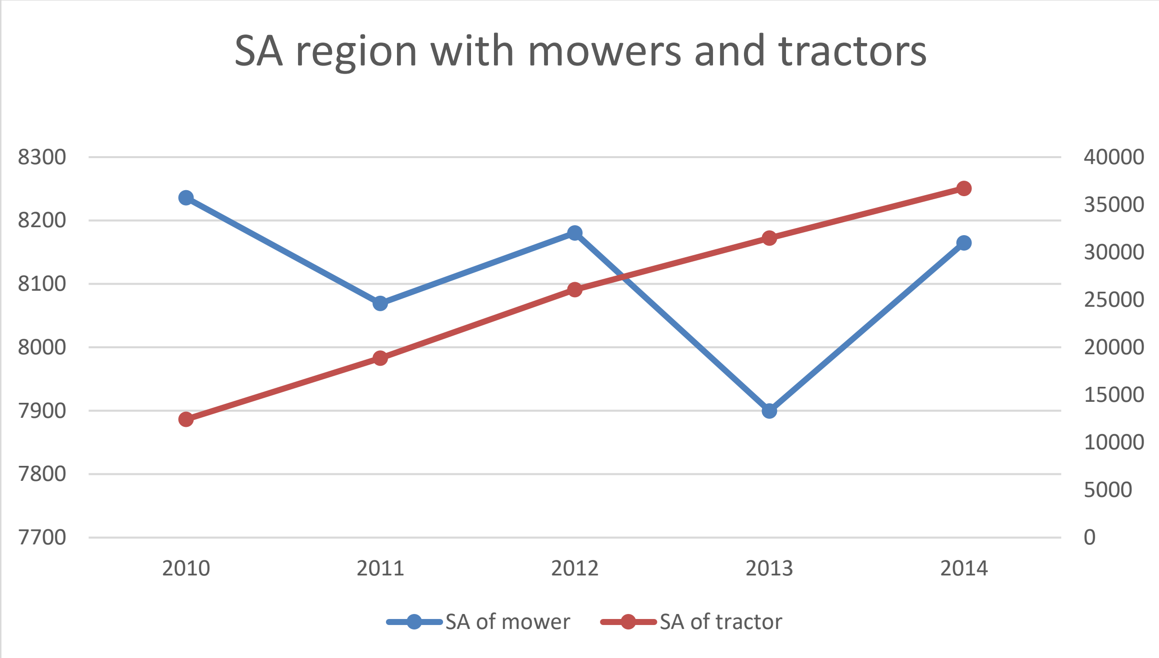


Figure 42: SA Regression(2)

Based on the data and chart, I will utilize the trendline tool to apply three mathematical methods of modeling relationships and trends to find the appropriate computational approach for the data. These methods are:

1. **Exponential function:**
   Exponential of tractor: $y = 10488e0{,}268x$
   Exponential of mower: $y = 8203{,}6e - 0{,}004x$

2. **Polynomial function:**
   Polynomial of tractor: $y = -300{,}73x2 + 7926{,}4x + 4643{,}6$
   Polynomial of mower: $y = 33{,}68x2 - 233{,}24x + 8439{,}3$

3. **Power function:**
   Power of tractor: $y = 8203{,}6e - 0{,}004x$
   Power of mower: $y = 8201{,}4x - 0{,}012$

	Polynomial	Exponential	Power
Mower	8.240	8.171	8.201
Tractor	12.269	13.711	12.213

Sample		
Year	SA of Mower	SA of Tractor
2010	8.236	12.425

Figure 43: Sample SA regression

➔ Based on the computational results of the SA region data with mowers and tractors, we conclude that the most suitable computational method for the data is the Polynomial method for both mower and tractor data. This conclusion is drawn because the Polynomial method yields forecast results that closely match the sample data. Therefore, we will select the Polynomial method for the given data.

EUR region

Thirdly, based on the Industry Mower Total Sales and Industry Tractor Total Sales datasets, I have created a table containing information on the total sales of mowers and tractors over five years for the EUR region. From there, I will model with three appropriate mathematical functions for the year 2010 to assess the accuracy of each model and evaluate which model is the most suitable.

	EUR of mower	EUR of tractor
2010	281053	74732
2011	279370	79868
2012	241357	83862
2013	226743	77199
2014	238679	70508

Figure 44: EUR regression(1)

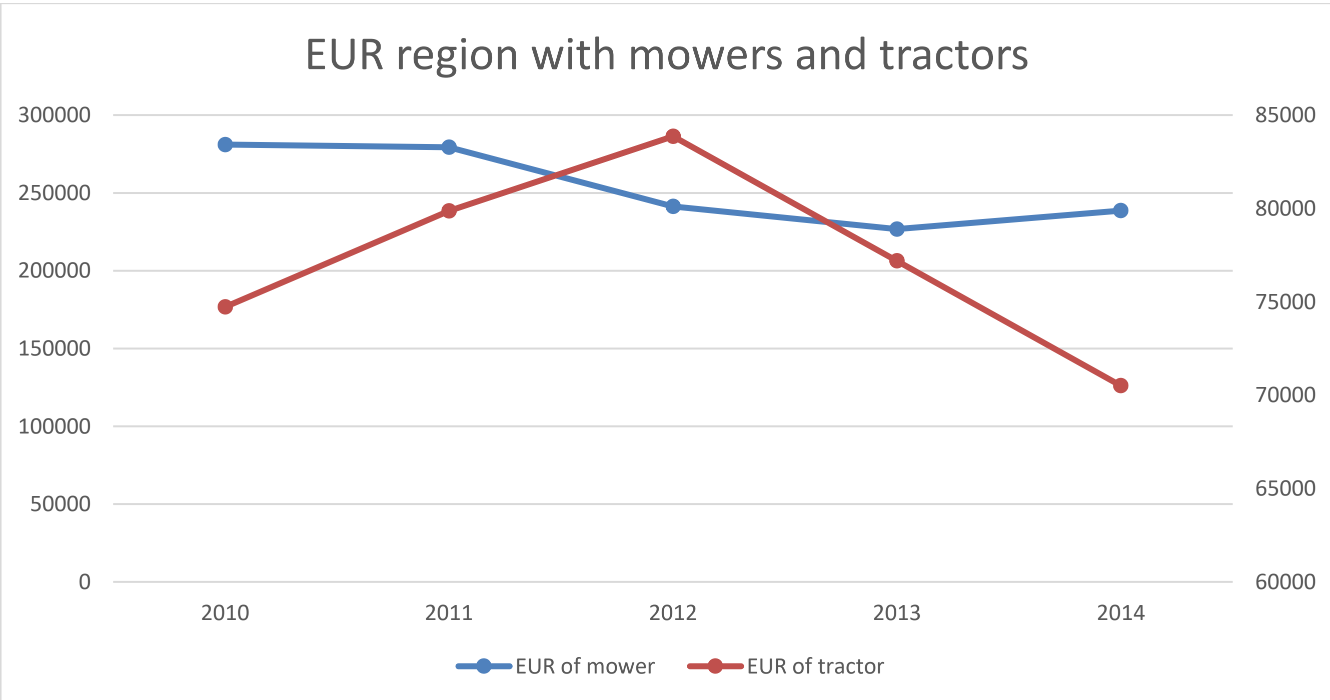


Figure 45: EUR regression(2)

Based on the data and chart, I will utilize the trendline tool to apply three mathematical methods of modeling relationships and trends to find the appropriate computational approach for the data. These methods are:

1. **Exponential function:**

Exponential of tractor: $y = 10488e0{,}268x$

Exponential of mower: $y = 8203{,}6e - 0{,}004x$

2. **Polynomial function:**

Polynomial of tractor: $y = -2450{,}7x2 + 13592x + 63414$

Polynomial of mower: $y = 3616{,}9x2 - 35439x + 319972$

3. **Linear function:**

Linear of tractor: $y = -1111{,}7x + 80569$

Linear of mower: $y = -13737x + 294653$

	Polynomial	Linear	Exponential
Mower	288.150	280.916	280.880
Tractor	74.555	79.457	79.457

Sample		
Year	EUR of Mower	EUR of Tractor
2010	281.053	74.732

Figure 46: Sample EUR regression

➔ Based on the computational results of the EUR region data with mowers and tractors, we conclude that the most suitable computational methods for the data are not just one but two methods. Specifically, the Polynomial method is suitable for tractor data, while the Linear method is suitable for mower data. This conclusion suggests that not every computational method will be appropriate for the data, and different methods may be needed for different subsets of the data.

Pac region

Next, based on the Industry Mower Total Sales and Industry Tractor Total Sales datasets, I have created a table containing information on the total sales of mowers and tractors over five years for the Pac region. From there, I will model with three appropriate mathematical functions for the year 2010 to assess the accuracy of each model and evaluate which model is the most suitable

	Pac of mower	Pac of tractor
2010	14565	14142
2011	16643	15914
2012	18929	16758
2013	23687	18700
2014	23870	13850

Figure 47: Pac Regression(1)

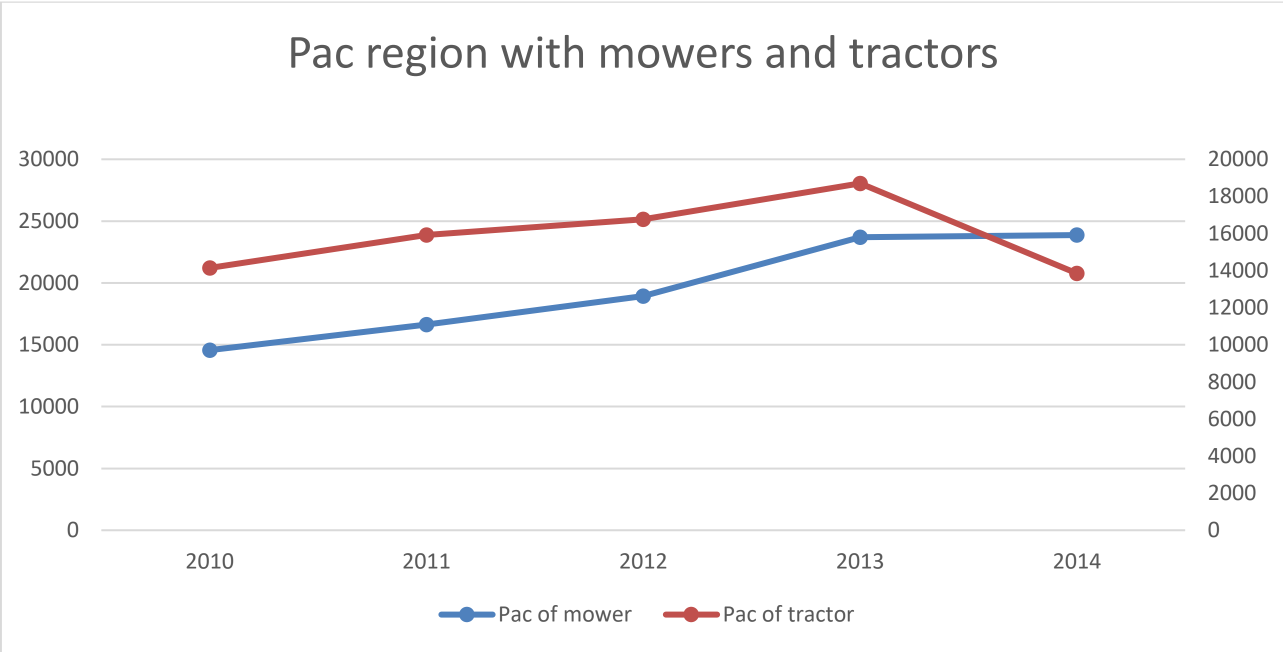


Figure 48: Pac regression(2)

Based on the data and chart, I will utilize the trendline tool to apply three mathematical methods of modeling relationships and trends to find the appropriate computational approach for the data. These methods are:

1. **Logarithmic function:**

   Logarithmic of tractor: $y = 1030{,}9ln(x) + 14886$

   Logarithmic of mower: $y = 6217{,}2ln(x) + 13586$

2. **Linear function:**

   Linear of tractor: $y = 220{,}36x + 15212$

   Linear of mower: $y = 2565{,}3x + 11843$

3. **Power function:**

   Power of tractor: $y = 14876x0{,}0613$

   Power of mower: $y = 13989x0{,}3295$

	Logarithmic	Linear	Power
Mower	13.586	14.408	13.989
Tractor	14.886	15.432	14.876

Sample		
Year	Pac of Mower	Pac of Tractor
2010	14.565	14.142

Figure 49: Sample Pac regression(2)

➔ Based on the computational results of the Pac region data with mowers and tractors, we conclude that the most suitable computational methods for the data are not just one but two methods. Specifically, the Logarithmic method is suitable for tractor data, while the Linear method is suitable for mower data. This conclusion further supports the notion that not every computational method will be appropriate for the data, and different methods may be needed for different subsets of the data.

China region

Finally, I will analyze the data from the Industry Tractor Total Sales table since there is no data available for mower sales in the China region. Therefore, I have created a dataset containing information on the total sales of tractors over five years for the China region. From there, I will model with three appropriate mathematical functions for the year 2010 to assess the accuracy of each model and evaluate which model is the most suitable.

Year	China of Tractor
2010	3473
2011	3746
2012	7081
2013	20034
2014	29859

Figure 50: China Regression (1)

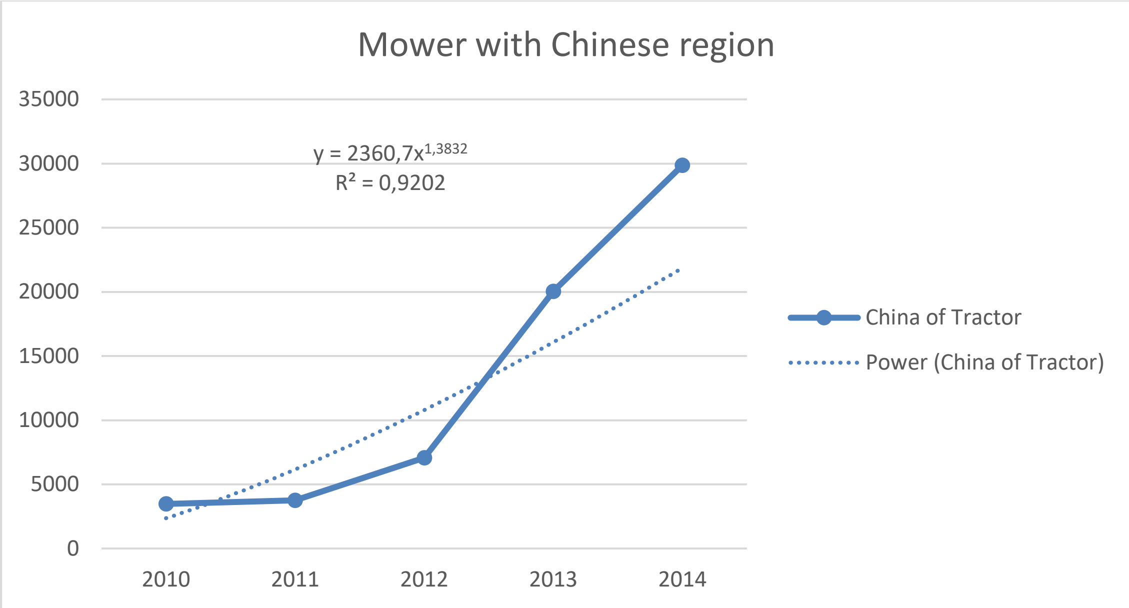


Figure 51: Pac regression (2)

Based on the data and chart, I will utilize the trendline tool to apply three mathematical methods of modeling relationships and trends to find the appropriate computational approach for the data. These methods are

1. **Polynomial function:**

   Exponential of tractor: $y = 1030{,}9ln(x) + 14886$

   Exponential of mower: $y = 6217{,}2ln(x) + 13586$

2. **Exponential function:**

   Polynomial of tractor: $y = 220{,}36x + 15212$

   Polynomial of mower: $y = 2565{,}3x + 11843$

3. **Power function:**

   Logarithmic of tractor: $y = 14876x0{,}0613$

   Logarithmic of mower: y = $13989x0{,}3295$

	Polynominal	Exponential	Power
Tractor	3.130	2.684	2.361

Sample	
Year	Tractor
2010	3.473

Figure 52: Sample Pac regression

➔ Based on the computational results of the Mower with Chinese region data, we conclude that the most suitable computational method for the data is the Polynomial method. This conclusion is drawn because the forecast results of this method closely match the actual results of the year 2010. Therefore, we will select the Polynomial method for the Mower with Chinese region data.

Show the result of forecasting future increases in production costs.

The "Unit Production Costs" dataset provides monthly accounting estimates of the variable cost per unit for manufacturing tractors and mowers over the past five years. Therefore, I have generated a spreadsheet calculating the average quantity of each unit for tractors and mowers from the Unit Production Costs data. Based on this data, I will apply modeling of relationships and trends within the dataset to analyze and forecast how much the average quantity per unit of tractors and mowers will increase production costs over the next two years

Unit Production Costs		
Year	**Tractor**	**Mower**
2010	$1.785	$51
2011	$1.869	$56
2012	$1.960	$60
2013	$1.975	$60
2014	$2.102	$64

Figure 53: Unit Production Costs (1)

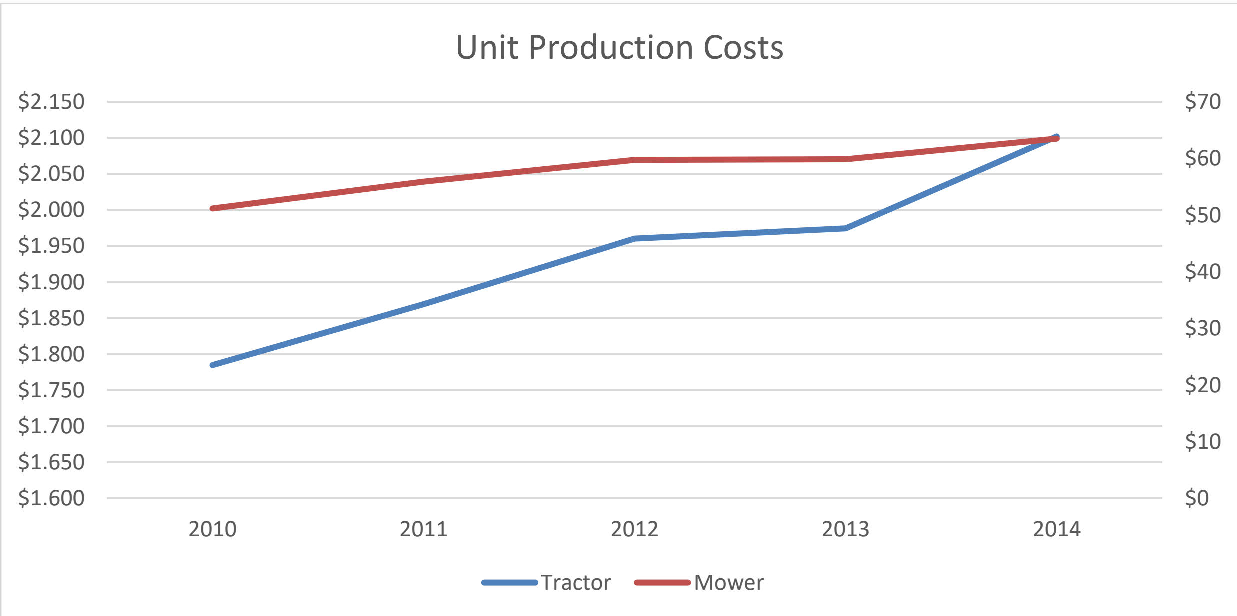


Figure 54: Unit Production Costs (2)

Based on the data and chart, I will utilize the trendline tool to apply three mathematical methods of modeling relationships and trends to forecast the pattern for the year 2012. These methods are:

1. **Linear function:**

   Linear of tractor: $y = 73{,}983x + 1716{,}2$

   Linear of mower: $y = 2{,}8583x + 49{,}458$

2. **Exponential function:**

   Exponential of tractor: $y = 1725{,}6e0{,}0382x$

   Exponential of mower: $y = 49{,}822e0{,}05x$

3. **Logarithmic function:**

   Logarithmic of tractor: $y = 179{,}92ln(x) + 1765{,}8$

   Logarithmic of mower: $y = 7{,}2501ln(x) + 51{,}091$

	Linear	Exponential	Logarithmic
Tractor	$1.938	$1.863	$1.963
Mower	$58	$58	$59

Sample		
2012	Tractor	Mower
	$1.960	$60

Figure 55: Sample Production Cost

➔ Based on the computational results of the Unit Production Costs dataset, we conclude that the most suitable computational method for the data is the Logarithmic method. This conclusion is drawn because the forecast results of this method closely match the actual results of the year 2012. Therefore, we will select the Logarithmic method for the Unit Production Costs data.

To forecast future increases in production costs specifically for the years 2015 and 2016, I will apply modeling of relationships and trends using the Logarithmic method. This method has been demonstrated to be the most suitable computational approach for the data. By utilizing the Logarithmic model, we can predict the trends in production costs for the selected time frame with greater accuracy, considering the historical patterns and relationships captured by the Logarithmic function.

- **Logarithmic function:**

  Logarithmic of tractor: $y = 179{,}92ln(x) + 1765{,}8$

  Logarithmic of mower: $y = 7{,}2501ln(x) + 51{,}091$

	2015	2016
Mower	$64	$65
Tractor	$2.088	$2.116

Figure 56: Forecast Unit Production Costs

➔ Based on the modeling of relationships and trends, I applied the logarithmic computational method to the Unit Production Costs data to derive the results above. As you may understand, I utilized the logarithmic mathematical function to generate forecasts for both mowers and tractors. Specifically, I forecasted that in 2015, the average cost per unit for mowers would be $64 and for tractors would be $2,088. Similarly, in 2016, I forecasted that the average cost per unit for mowers would be $65 and for tractors would be $2,116. I hope that with the demonstrated data, it will be beneficial for Elizabeth Burke.

Evaluate how the predictive analytic techniques can be used for forecasting purposes of Ms. Burke, including forecast the sale number of mowers and tractors and future increases in production costs. D2

Evaluation of forecasting analysis techniques for predicting the sales volume of mowers and tractors.

To evaluate the appropriate forecasting model for predicting the sales volume of mowers and tractors to plan manufacturing capacity, we will rely on the forecast results from each region: North America (NA), South America (SA), Europe (Eur), the Pacific Rim (Pac), and China. This approach will allow us to select the most suitable general method, thereby assessing the most accurate and reliable data.

NA region

	Linear	Polynomial	logarithmic
Mower	864.286	877.769	867.737
Tractor	68.650	82.179	68.977

Sample		
Year	**NA of Mower**	**NA of Tractor**
2010	880.392	81.516

Figure 57: NA region (D2)

SA region

	Polynomial	Exponential	Power
Mower	8.240	8.171	8.201
Tractor	12.269	13.711	12.213

Sample		
Year	**SA of Mower**	**SA of Tractor**
2010	8.236	12.425

Figure 58: SA region (D2)

EUR region

	Polynomial	Linear	Exponential
Mower	288.150	280.916	280.880
Tractor	74.555	79.457	79.457

Sample		
Year	**EUR of Mower**	**EUR of Tractor**
2010	281.053	74.732

Figure 59: EUR region (D2)

Pac Region

	Logarithmic	Linear	Power
Mower	13.586	14.408	13.989
Tractor	14.886	15.432	14.876

Sample		
Year	Pac of Mower	Pac of Tractor
2010	14.565	14.142

Figure 60: Pac region (D2)

China Region

	Polynominal	Exponential	Power
Tractor	3.130	2.684	2.361

Sample	
Year	Tractor
2010	3.473

Figure 61: China region (D2)

➔ From the sample forecast results of tractor sales volume in 2010, we observe that we are utilizing modeling relationships and trends in data using three different mathematical methods for analysis. With these three distinct mathematical methods, I have been able to identify the most appropriate mathematical functions for my data, namely the Polynomial and Linear functions. These two mathematical functions appear most frequently in this dataset and exhibit the highest accuracy compared to the actual results of the Sample in 2010.

To provide a detailed and accurate explanation of why our data aligns best with the Polynomial and Linear functions, we will further illustrate with examples to better understand the calculation formulas and the selection of the most appropriate data for accuracy.

For example

1. Understanding both the mathematics and the descriptive properties of different functional relationships is crucial in constructing predictive analytical models. We often initiate the process by plotting the data on a chart to comprehend their patterns and select the suitable type of functional relationship to integrate into an analytical model. With regards to my dataset, I have observed that it is a time series data. Therefore, we utilize a line chart for visualization purposes.

→ Time-series

Year	NA of Mower	NA of Tractor
2010	880392	81516
2011	853795	75474
2012	863023	78501
2013	867283	97342
2014	890359	130743

Figure 62: Example NA region (1)

2. To explain why I use two mathematical functions, Polynomial and Linear functions, we will rely on a line chart, as I mentioned earlier, because it is a type of time-series data visualization. From there, we will utilize trendline tools to observe how the chart represents a straight line and the R-squared value. Specifically, I will use the example of the EUR region, and from the dataset below, I will use the Mower column for a linear function and the Tractor column for a Polynomial function to display them on the chart in the most intuitive way.

	Polynomial	Linear	Exponential
Mower	288.150	280.916	280.880
Tractor	74.555	79.457	79.457

Sample		
Year	**EUR of Mower**	**EUR of Tractor**
2010	281.053	74.732

Figure 63: Example NA region (2)

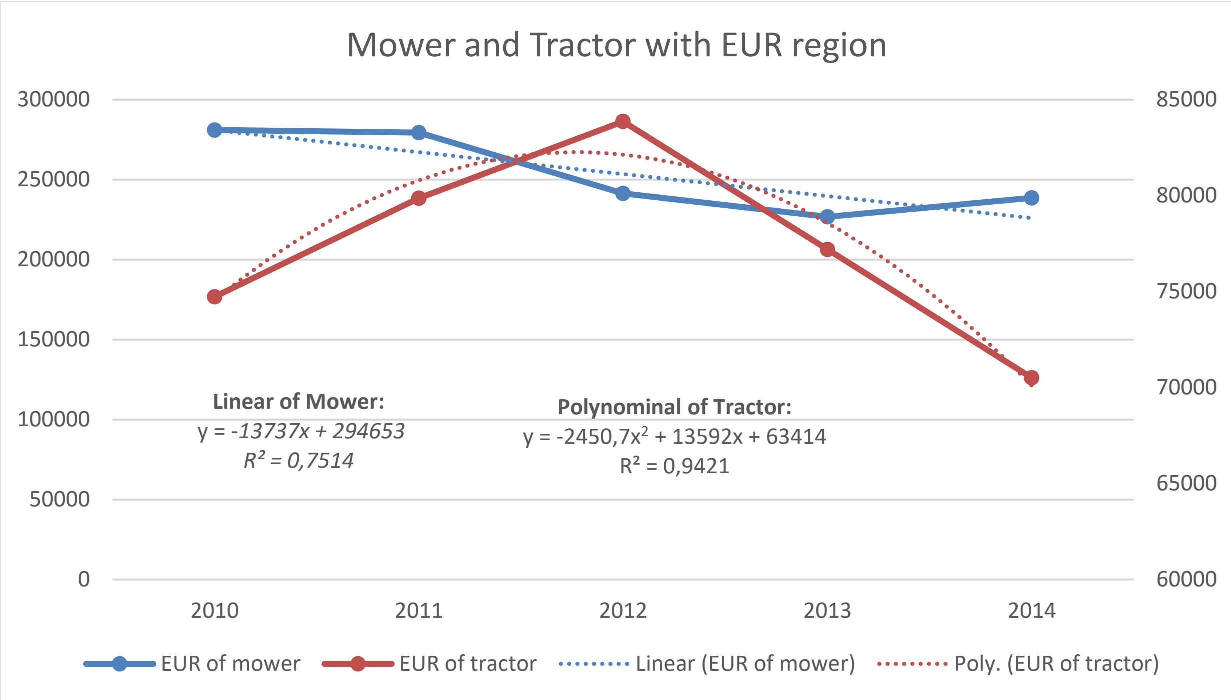


Figure 64: Example NA region (3)

➔ Because the Mower column shows a downward trend, forming a stable decreasing line, and it is easy to understand and can approximate behavior rather well over small ranges of values. Additionally, we have used trendline tools to display the formula, constant values, and along with it, the R-squared value, which is a measure of the "fit" of the line to the data. The value of R-squared will be between 0 and 1. The larger the value of R-squared, the better the fit, and in the Mower line chart, the R-squared value is 0.75. This number indicates high accuracy and suitability with the data, so I use the linear function to represent the Mower column. For the Tractor column, we observe a parabolic shape with only one peak, accompanied by an R-squared value of 0.94, indicating very high accuracy and suitability with the data. Therefore, I choose the Polynomial function to use for this column of data. The combination of data type, suitable mathematical function, and high R-squared values demonstrate high suitability, accuracy, and reliability, and I believe that this information will provide Elizabeth Burke with valuable insights.

Evaluate predictive analytical techniques for future increases in production costs

In this dataset, I have applied modeling techniques to analyze relationships and trends, utilizing three mathematical functions to assess and forecast the average number per unit of tractor and cutter machines, which will increase production costs over the next two years. Subsequently, I will evaluate the predictive analysis techniques regarding the increase in production costs in the future for accuracy and suitability with the given data. I will employ the sample results from 2012 to confirm the accuracy of the data through these three mathematical functions and choose the most suitable one based on the evaluation.

	Linear	Exponential	Logarithmic
Tractor	$1.938	$1.863	$1.963
Mower	$58	$58	$59

Sample		
2012	**Tractor**	**Mower**
	$1.960	$60

Figure 65: Example Unit Production Costs

➔ Based on the sample results of the Unit Production Cost, I utilized modeling techniques to analyze relationships and trends using three mathematical functions to forecast the production cost of mowers and tractors in the year 2012. From the three results of these different functions, it was evident that the computational function that provided the closest approximation to the actual values was the Logarithmic function. Both columns of data for mowers and tractors from the Unit Product Cost table best fit the Logarithmic mathematical function.

To be able to explain in detail and accurately why our data best fits the Logarithmic mathematical function, we will provide an additional example so that we can understand more about the calculation formula and how to choose the most appropriate data for accuracy.

For Example

1. To create predictive analytics models, it is important to understand the mathematical and descriptive properties of various functional relationships. We usually start by creating a diagram of the data to understand it and select the appropriate types of functional relationships that we want to incorporate into the analytical model. As for this data set, I realize that this is time series data, we use a line chart.

Unit Production Costs

Year	Tractor	Mower
Time-series 2010	$1.785	$51
2011	$1.869	$56
2012	$1.960	$60
2013	$1.975	$60
2014	$2.102	$64

Figure 66: Example Unit Production Costs (1)

2. To explain why I use the Logarithmic mathematical function, we will rely on a line chart, as I mentioned earlier, because it is a type of time-series data visualization. From there, we will utilize trendline tools to observe how the chart represents the line and the R-squared(R2) value. Specifically, I will provide an example using the Unit Production Cost table, and from the dataset below, both the Mower and Tractor columns will use the Logarithmic function to visualize the data, helping us to gain the most objective insight.

	Linear	Exponential	Logarithmic
Tractor	$1.938	$1.863	$1.963
Mower	$58	$58	$59

Sample		
2012	**Tractor**	**Mower**
	$1.960	$60

Figure 67: Example Unit Production Costs (2)

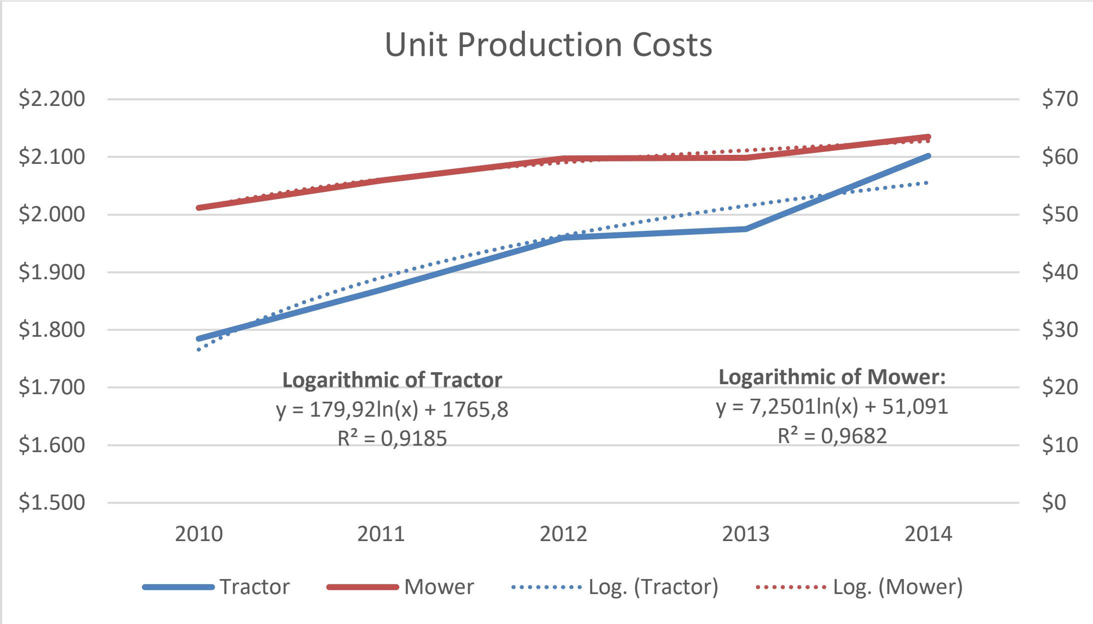


Figure 68: Example Unit Production Costs (3)

➔ Because both the Mower and Tractor columns exhibit a rapid increase followed by a stabilization, we can observe a changing variable that initially rises quickly and then stabilizes. Moreover, we used trendline tools to display the formula, constant values, and along with it, the R-squared value, which is a measure of the "fit" of the line to the data. The value of R-squared will be between 0 and 1. In the line chart of the Tractor column, the astonishingly high R-squared value of 0.9185 indicates high accuracy and a high degree of logic and suitability with the data. Following that, the line chart of the Mower column boasts an impressive R-squared value of 0.9682, signifying the highest accuracy, logic, and suitability with the data. The combination of data type, appropriate mathematical functions, and high R-squared values demonstrates a high level of suitability, accuracy, and reliability of the Unit Production Cost table. I believe this information will provide Elizabeth Burke with valuable insights.

M4 and D3 Describe how these prescriptive analytic methods are used to find the best course of action in a situation and Apply an appropriate programming language or tool to demonstrate how these prescriptive analytic methods are used to find the best course of action in a situation.

Summarize the strategies

Based on the information received, our team will note and summarize the strategies that PLE has proposed. Then, we will create a table showing the parameters of each strategy.

PLE is confronted with a binary decision-making scenario: either to launch their product on a global scale, which would entail an expenditure of $850,000, or to conduct a market test in North America, costing $200,000.

In the case of a global product launch, the probability distribution for the product's response is bifurcated into two potential outcomes: a high response with a probability of 0.6 and a low response with a probability of 0.4. The expected gross revenue, contingent on a high response, is projected to be $2,000,000, whereas in the event of a low response, the anticipated revenue is $450,000.

Alternatively, if the decision is made to commence with a North American market test, the probability of a high product response is 0.7, with a 0.3 probability for a low response. The expected gross revenue under a high response scenario is $1,200,000, while a low response would yield an estimated $200,000.

However, PLE acknowledges the uncertainty inherent in these projections, as they may or may not accurately reflect the global market potential. Consequently, following the market research, PLE must then decide whether sales to North America, expand to the global market, or discontinue the product.

In the event of a high response in the North American market, PLE would allocate an additional $200,000 for global marketing. The probability of a high product response under these circumstances is 0.9, with a 0.1 probability for a low response. The expected gross revenue in the case of a high response is $2,000,000, while a low response would result in an estimated $450,000.

Conversely, if the North American market response is low, PLE would invest an additional $600,000 in global marketing. The probability of a high product response in this scenario is significantly lower at 0.05, with a high probability of 0.95 for a low response. The expected gross revenue for a high response remains at $2,000,000, while a low response would yield an estimated $450,000.

Prescriptive analytic methods to find optimal selling strategy.

Situation	Cost	Revenues (HP)	HP	Revenues(LP	LP
Global	$850.000	$2.000.000	0,6	$450.000	0,4
North America	$200.000	$1.200.000	0,7	$200.000	0,3
Expand globally (**NAHP**)	$200.000	$2.000.000	0,9	$450.000	0,1
Expand globally (**NALP**)	$600.000	$2.000.000	0,05	$450.000	0,95

Figure 69: Strategy for snow blower

Notes
NAHP: North America with hight probabilities
NALP: North America with Low probabilities
HP: Hight probabilities
LP: Low probabilities

Figure 70: Notes Strategy for snow blower

With this compiled information, our team presents a Decision Tree to determine the best selling strategy and evaluate risk.

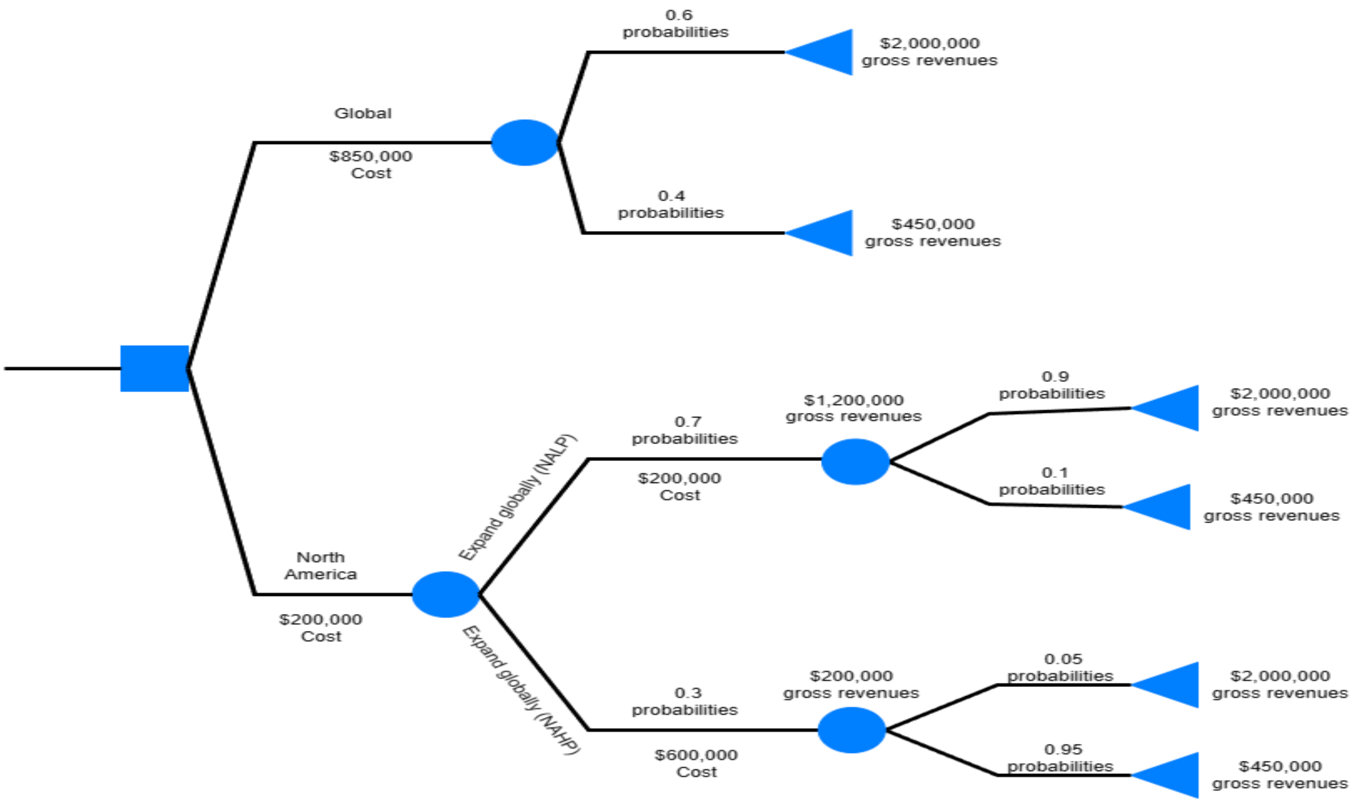


Figure 71: Snow blower Decision Tree

The formula to calculate the expected revenue for each strategy is based on the probability and corresponding revenue of each event. Here's how to calculate for each strategy:

Launch product on a global scale: net revenues are calculated as follows: (high response probability * high response revenue) + (low response probability * low response revenue) - Launch product on a global cost.

$$(\mathbf{0.6} * \$\mathbf{2,000,000}) + (\mathbf{0.4} * \$\mathbf{450,000}) - \$\mathbf{850,000} = \$\mathbf{530,000}$$

Situation	Cost	Revenues (HP)	HP	Revenues(LP)	LP
Global	$850.000	$2.000.000	0,6	$450.000	0,4
North America	$200.000	$1.200.000	0,7	$200.000	0,3
Expand globally (**NAHP**)	$200.000	$2.000.000	0,9	$450.000	0,1
Expand globally (**NALP**)	$600.000	$2.000.000	0,05	$450.000	0,95
Situation					
Global	=$C$4*$D$4 + $E$4*$F$4 - $B$4			$530.000	

Figure 72: Launch product on a global scale

North American market test: net revenues are calculated as follows: Net revenues = (high response probability in North America * high response revenue in North America) + (low response probability in North America * low response revenue in North America) - market test cost

$$(\mathbf{0.7} * \$\mathbf{1,200,000}) + (\mathbf{0.3} * \$\mathbf{200,000}) - \$\mathbf{200,000} = \$\mathbf{700,000}$$

Situation	Cost	Revenues (HP)	HP	Revenues(LP)	LP
Global	$850.000	$2.000.000	0,6	$450.000	0,4
North America	$200.000	$1.200.000	0,7	$200.000	0,3
Expand globally (**NAHP**)	$200.000	$2.000.000	0,9	$450.000	0,1
Expand globally (**NALP**)	$600.000	$2.000.000	0,05	$450.000	0,95
Situation					
North America	=$C$5*$D$5+$E$5*$F$5-$B$5			$700.000	

Figure 73: North American market test

Expand globally: To calculate net revenues, we need to consider both high response in North America and low response in North America.

If the response in North America is high (probability is 0.7), the high global response probability is 0.9 and the low global response probability is 0.1. Gross revenue in this case is $2,000,000 (high global response probability) and $450,000 (low global response probability) which already includes the Gross revenue of North

America so the expected revenue in this case is: revenue = high response probability in North America * (high global response probability * high global response revenue) + (low global response probability * low global response revenue - global expansion cost)

$$\mathbf{0,7 * (0,9 * \$2,000,000 + \ 0,1 * \$450,000 \ - \$200,000) \ = \ \$1,151,500}$$

If the response in North America is low (probability is 0.3), the high global response probability is 0.05 and the low global response probability is 0.95. Gross revenue in this case is $2,000,000 (high global response probability) and $450,000 (low global response probability) which already includes the Gross revenue of North America so the expected revenue in this case is: revenue = low response probability in North America * (high global response probability * high global response revenue) + (low global response probability * low global response revenue - global expansion cost)

$$\mathbf{0.3 * (0,05 * \$2,000,000 + \ 0,95 * \$450,000 \ - \$600,000) \ = \ -\$21,750}$$

With 2 results $1,151,500 and -$21,750 for 2 cases. Next, we subtract the North American market trial cost ($200,000). Therefore, the net revenues in this case are: net revenues = revenue If the response in North America is high + revenue If the response in North America is low - trial cost

$$\mathbf{\$1,151,500 \ + \ (-\$21,750) \ - \ \$200,000 \ = \ \$929,750}$$

Situation	Cost	Revenues (HP)	HP	Revenues(LP)	LP
Global	$850.000	$2.000.000	0,6	$450.000	0,4
North America	$200.000	$1.200.000	0,7	$200.000	0,3
Expand globally (**NAHP**)	$200.000	$2.000.000	0,9	$450.000	0,1
Expand globally (**NALP**)	$600.000	$2.000.000	0,05	$450.000	0,95

Situation	
Expand globally	= $D$5*($D$6*$C$6+$F$6*$E$6-$B$6) + $F$5*($D$7*$C$7+$E$7*$F$7-$B$7)-$B$5 $929.750

Figure 74: Expand globally

The summary of the results of the 3 specific cases is as follows:

Situation	Cost	Revenues (HP)	HP	Revenues(LP)
Global	$850.000	$2.000.000	0,6	$450.000
North America	$200.000	$1.200.000	0,7	$200.000
Expand globally (**NAHP**)	$200.000	$2.000.000	0,9	$450.000
Expand globally (**NALP**)	$600.000	$2.000.000	0,05	$450.000

Situation	Net revenues
Global	$530.000
North America	$700.000
Expand globally	$929.750

Figure 75: Situation

- Launch product on a global: $530,000
- North American market test: $700,000
- Expand globally: $929,750

Risk profile

Summarizing the results we have just obtained; our team present a profit statistics table to assess the risk:

	Global		North America		Expand globally (NAHP)		Expand globally (NALP)	
Gross revenue	$2.000.000	$450.000	$1.200.000	$200.000	$2.000.000	$450.000	$2.000.000	$450.000
Cost	$850.000	$850.000	$200.000	$200.000	$400.000	$400.000	$800.000	$800.000
Net revenue	$1.200.000	$180.000	$840.000	$60.000	$1.260.000	$31.500	$30.000	$128.250
Probability	60%	40%	70%	30%	63%	7%	2%	29%
Profit	$350.000	-$670.000	$640.000	-$140.000	$860.000	-$368.500	-$770.000	-$671.750

Figure 76: Risks profile

The probability of Expand globally (NAHP) and Expand globally (NALP) is calculated by multiplying the probabilities on the event branches along the path to the terminal outcome (Decision tree).

Expand globally (NAHP) with high probability: 0,7 $* \ 0.9 \ = \ 0.63$

Expand globally (NAHP) with low probability: $0.7 \ * \ 0.1 \ = \ 0.07$

Expand globally (NALP) with high probability: $0.3 \ * \ 0.05 \ = \ 0.02$

Expand globally (NALP): with low probability: $0.3 \ * \ 0.95 \ = \ 0.285$

First, in the case of Global with high probability, the profit brings in $350,000; otherwise, the loss goes up to $670,000, nearly twice the profit of $350,000. The likelihood of incurring this loss is 40%. Although this scenario yields relatively low profit, the associated risk is very high, and the cost for this option is the highest among all remaining alternatives at $850,000. Overall, this is a high-cost, high-risk, low-profit option.

In the North American market test case, the profit amounts to $640,000, and conversely, the loss is $140,000 with low probability. The damage rate for this option is 30%, lower than that of the Global case. Additionally, the cost for this option is significantly lower than Global by $200,000, which is only a quarter of Global's cost. The profit gained is much higher compared to the cost incurred. Therefore, this is a low-cost, low-risk, high-profit option.

Next, after testing the North American market, PLE will expand the snow blower globally based on two scenarios: high probability and low probability.

In the case of North America with high probability, the total profit is $860,000, the highest among all scenarios. Conversely, if it fails, the loss is still relatively high at $368,500. However, the damage rate is only 7%, considering the cost of $400,000 for both expanding the global market and testing the North American market. Overall, if the North American market test is successful, this will be the next option with low cost, low risk, and high profit.

Finally, in the case of North America with low probability, both high and low probability scenarios result in substantial losses ranging from $671,000 to $770,000, without any corresponding profit. This can be considered the worst option compared to the remaining alternatives.

Summarize the risks in each case. The worst-case scenario could be a low response rate in North America affecting the Global low response rate, causing damage up to $770,000. In any case with a low response rate, it will cause financial damage to PLE, with the damage ranging from $140,000 to $770,000. With the plan to test the market in North America and then expand globally with a low response rate, PLE could decide to stop product development to minimize the damage with $140,000.

Result

Based on these proposed scenarios and risk profiles, **expand globally** seems to yield the most impressive results. If positive feedback is received from the North American market, PLE could confidently deploy the product internationally. This is evidenced by the breakthrough growth in the global reach ratio, from 0.6 to 0.9, reflecting the warm acceptance from consumers after successful testing in North America. Moreover, the difference in net revenue between the two scenarios **launching the product globally** and testing **the North American market before expanding** is quite significant, with $929,750 compared to $530,000. This demonstrates the ability to minimize risk through product testing in a specific market before proceeding with expansion. This method allows PLE to more accurately assess product reception in a controlled environment while reducing initial costs. In summary, this is the option with the lowest cost and lowest risk occurrence rate, and brings the highest benefit to PLE.

Conclusion

The paper on predictive and prescriptive analytics in this research emphasizes the importance of data-driven decision-making in a business environment. By using techniques such as linear regression, logistic regression, and multivariate analysis, we have demonstrated the ability to effectively forecast and optimize business outcomes. The application of these methods using appropriate tools like Excel has facilitated the generation of sales forecast results, the most effective business strategies, and the costs for operating a business. The results obtained from this analysis serve as important information for decision-making in shaping the business strategies and operations of an organization.

Evaluation

In the field of predictive analytics, we used the Sales Performance Lawn Equipment Database which was analysed using methods such as time series analysis, linear regression, multiple linear regression, and modelling relationships and trends in data. These methods provide reliable forecasts and valuable insights to support business decision-making for Elizabeth Burke and Performance Lawn Equipment (PLE). We ensure a seamless integration of predictive techniques, meticulously applying each step logically and effectively. Forecast calculations based on probability and expected revenue for each scenario help generate realistic and feasible simulation results. This allows managers to gain a deeper understanding of potential situations and corresponding risks. Specifically, we apply predictive methods and tests on each sample set, utilizing evaluation metrics and decision trees to assess accuracy, suitability, and risks. Through this process, we achieve precise and reliable selections, providing Elizabeth Burke with critical insights for objective decision-making. In summary, the application of analytical methods and techniques in this analysis results in accurate forecasts and detailed risk assessments, facilitating intelligent and effective business decisions. This approach enables businesses to seize opportunities and mitigate risks in uncertain business environments.

References

Evans, J., 2017. *Business Analytics.* 3rd ed. s.l.:Pearson.

Funix, 2021. *What is prescriptive analytics?.* [Online]

Available at: https://funix.edu.vn/chia-se-kien-thuc/phan-tich-de-xuat-prescriptive-analytics-la-gi/

[Accessed 23 3 2024].

Hạnh, L. H., 2024. *Phân tích dữ liệu là gì? Quy trình, phương pháp và công cụ hỗ trợ.* [Online]

Available at: https://base.vn/blog/phan-tich-du-lieu/

[Accessed 23 3 2024].