Informatics Institute of Technology

in collaboration with

the University of Westminster, UK

Department of Computing

(B.Eng.) in Software Engineering

Module: 6DATA005C.1 Advanced Analytics Course Work (Individual Report)

Student Name: Wehan Withana

UOW Number: w1761296

IIT Student ID: 2019349

Date of Submission: 13/01/2023

Contents

List of Figures

1. Introduction

Companies are always trying to increase their profit and expand their business. Because of that they need a proper way to deliver their products and services. Sales forecasting according to time periods is very important for every business specially textile companies because most of companies waste their money in buying unnecessary things. This document includes the scope of the problem, the domain of the problem, mathematical model, dataset, and the solution.

2. Project Scope

The scope of the project is finding a solution for a textile company to improve their sales using sales forecasting because most of the textile companies have lost their money with the buying of unnecessary items.

3. Literature Review

The statistics in this study (Dikov, 2020). cover the years 2013 through 2019. The goal is to develop a model that will enable us to predict GoPro's sales for the 2020 fiscal year. This may now display both the new projections and the actual historical numbers.

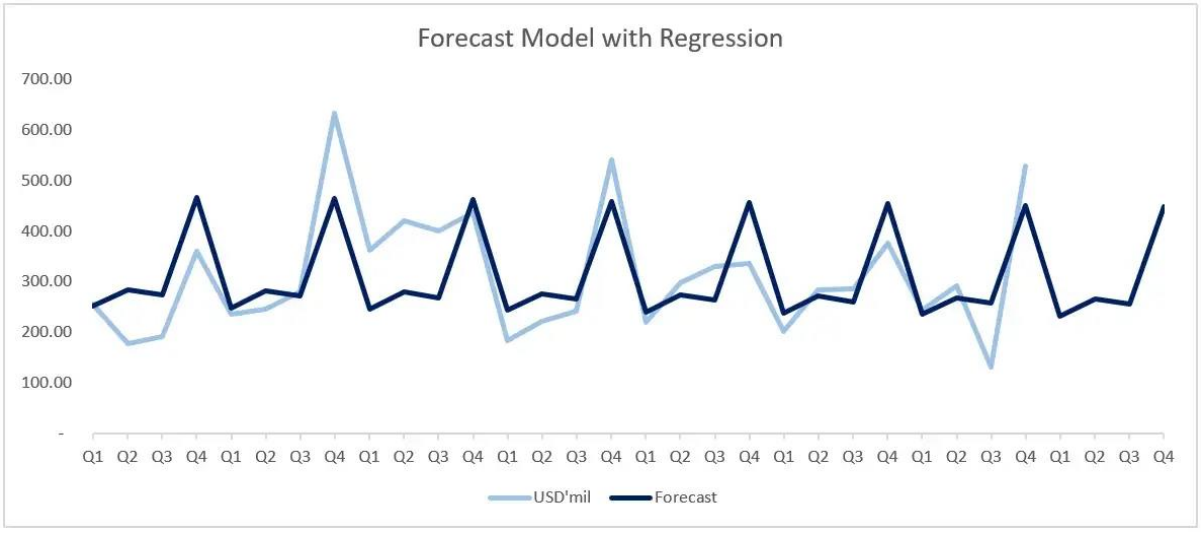


Figure 1: Forecast Model with Regression

4. Problem Domain

The challenges Sri Lankan Fashion industry is having with the seasonal sales are included in the problem domain. To forecast the sales of each and every product in the future month, they require a system. When the company manager does his duties well, the company will naturally make a profit. Targets must be accomplished by the manager. To increase sales, a system must forecast the most appropriate, profitable, and popular product. Due to Sri Lanka's recent problems and the skyrocketing pricing of services, they are unable to estimate sales using their previous experience. Therefore, it's critical to keep sales in line with demand.

5. Evaluate the Problem

Below is the detailed description of the problem that was found, the mathematical model and the data collection.

5.1. Company Details

The ASB fashion brand has 20 branches dispersed around the island. Leading retailer in Sri Lanka, ASB Fashion sells not only clothing but also accessories and household goods. The products are available for purchase by people of all ages and genders. They are aiming for the public to improve sales during the next Christmas season by making adjustments to their sales habits. As of right now, we're focusing on the three branches in the western province: Kalutara, Panadura, and Mathugama.

Appendix A – The company details can be found related to the proposed problem.

5.2. Identification of the Problem

The ASB Company is a leading Fashion company in Sri Lanka having 20 branches island wide. They mainly focused on distributing the latest clothing styles during the festive season. The main problem in the fashion industry is spending money on unnecessary items in

unnecessary times. The ASB Fashion needs to find the sales forecasting system to choose items accordingly to suit the time to minimize the wastage and improve their sales profit.

5.3. Benefits of the Sales Forecasting System

The sales forecasting system will minimize the wastage and gave crucial benefits to the company as below

- Helps in overall business planning, budgeting and risk management.
- Allows companies to efficiently allocate resources for future growth and manage cash flow.
- Helps to achieve their goals by identifying early warning signals in their sales.
- Helps to estimate their costs.

5.4. Objective of the Research

In this section the author aims to describe the objective of this research. Here are some objectives.

- Identify the highly demanding and low demanding sales items in a season.

From this, managers can avoid unnecessary invests on low demanding items and this will save money. Therefore managers can invest more on highly demanding items.

- Manage stocks and storage more efficiently.

This is another benefit of our research. After a season every shop has a redundancy. They need to clear these stocks to buy new items. After the season, shop owners sell those items at a very cheap price. It'll reduce their profit. Therefore using this system they will able to minimize the redundancy and it will help to keep the stocks according to the demand of the product

5.5. Data Collection of the Problem

To overcome this problem the team requested data related to the fashion industry from ASB Fashion. We explained the purpose of this research and after the explanation they realized the usefulness of this research and agreed to provide the data as requested.

We used this data to forecast the price of given items to maximize their selling's and profits.

5.6. Mathematical Model of the Problem

Identifying a mathematical model is the first and the most important step of solving this type of problem. In this problem we needed to forecast the sales. We wanted to identify the most selling items. Therefore, we choose multiple linear regression models. Give below the multiple linear regression model formulation for this problem.

- Independent Variables (X) - **Number of sales in each item (C0756 (x_1), A0112(x_2), R3699(x_3), A0136(x_4))**
- Dependent Variable (Y) - **Monthly Profit**
- Mathematical Model

$y = mx_1 + mx_2 + mx_3 + mx_4 + C$

Y(Profit) = { (Coefficients of C0756)* C0756 + (Coefficients of A0112)* A0112+(Coefficients of R3699)* R3699 + (Coefficients of A0136)* A0136 + Coefficients of Intercept }

6. Way of Solving Problem

6.1 Steps of solving

We used excel analyse tool for solve this problem.

Appendix B showing the steps of solving the problem.

6.2 Solution

After the steps in 6.1 we can get the 2023 predicted more accurate profit for chosen items.

Appendix C : Previous profit and predicted profit chart.

27		A0136	536,789	597994.4062	26	0	0	0	1
28	2022	C0756	1,675,787	1543945.156	27	1	0	0	0
29		A0112	889,568	676527.5366	28	0	1	0	0
30		R3699	1,579,086	1192220.912	29	0	0	1	0
31		A0136	835,786	608662.5366	30	0	0	0	1
32	2023 - Predicted Profit in items	C0756		1554613.286	31	1	0	0	0
33		A0112		687195.667	32	0	1	0	0
34		R3699		1202889.042	33	0	0	1	0
35		A0136		619330.667	34	0	0	0	1
36									
37									
38									
39									
40									

Figure 2: Predicted Profit

7. Assumption of the Problem

We needed to establish some assumptions before we could take on this challenge. We must determine the environment of the company and the products it sells. This business doesn't keep accurate data to analyse the seasonal sales of its products.There are numerous ways to increase sales and profits for fashion businesses, but ASB Fashion only looks at sales data from the previous three months.

8. Usefulness of the Proposed Model

As the company did not have any proper plan to invest more efficient items on seasonal periods, they had to face a lot of difficulties. Due to awareness of proper management company owners spend lots of money on buying unnecessary items. Then they faced profit loss more often and there were times they had stopped buy new items due to lack of money to invest and storage problems. As a result of this unable to buy new items they need to clear their stocks any and sell redundancy at cheap price. Proposed model is the most suitable solution for this problem. The model developed by the team for this problem is very useful to the ASB Company as the proposed solution allows them to properly manage their stocks and increase the profit. .All these benefits help the company to get to the top of the Clothing industry.

Thorough this crisis situation most of the fashion industry companies are going down. The owners can’t buy more items due to this corruption. Therefore they need to invest that limited money on more efficiently. Hence this model is most useful for this company to minimize their redundancy and avoid investing on unnecessary items.

9. Self-Reflection about the Project

Finding a research topic and challenge was the main goal because the coursework said that students needed to identify a real-world problem and its solution. The team first proposed problem is find minimum cost for Uber rides. But we couldn’t find a dataset for Sri Lankan rides. That’s the main reason to avoid that question. The problem was proposed without any clear idea about models or problems. Hence the team decided to study thoroughly multiple linear regression model and excel analysing in order to get more knowledge and find a suitable problem. After learning about various problem-solving techniques, we concentrated on one of the most prevalent issues in the modern world. In that situation, we discovered that the majority of businesses experience profit losses as a result of their lack of management knowledge, and some researchers have further demonstrated this through their research. At that point, upon choosing the problem, we had a solid understanding of the models and real-world problems.

Then we found this sales forecasting problem. It’s a current major problem for Sri Lankan fashion companies in this crisis situation. After doing this, it will be a big help to these fashion industry. That why we choose this problem. After confirming the research area, the team decided to learn more about the company's sales background. The hardest part in this research was find a business with this issue. Then we requested the dataset that we need from the ASB Company for our research. I've discovered a lot along this procedure. This includes identifying a project-related problem, discovering how to collaborate with others, and communicating with the business. More significantly, I've improved my ability to communicate well with others because getting my point through was crucial. Also I could get a huge technical knowledge about regression models, excel solver and analysing tools.

References

FCCA, D.D. (2020). Multiple Linear Regression Analysis in Excel. [online] Medium. Available at: https://medium.com/magnimetrics/multiple-linear-regression-analysis-in-excel-6f1436acd896.

john-galt. (n.d.). Multiple Regression. [online] Available at: https://johngalt.com/forecasting-methods/simple-regression [Accessed 12 Jan. 2023].

Needle, F. (2020). How to Use Regression Analysis to Forecast Sales: A Step-by-Step Guide. [online] blog.hubspot.com. Available at: https://blog.hubspot.com/sales/regression-analysis-to-forecast-sales.

Appendix A

Company Name	ASB Fashion (pvt)Ltd.
Head Office	478 Athur V. Dias Mawatha. Panadura
Telephone Number	038 2241284
Web Site	http://asbfashion.com/

Appendix B

Step 1 - This is the chosen items sales date set in 8 Years.

	A	B	C	D	E	F	G	H	I	J
1	**Year**	**Item Code**	**'rofit in the mont**	**Forecast**	**t**	C0756	A0112	R3699	A0136	
2	**2015**	C0756	1,353,451	1474602.308	1	1	0	0	0	
3		A0112	615,750	607184.689	2	0	1	0	0	
4		R3699	1,221,300	1122878.064	3	0	0	1	0	
5		A0136	534,800	539319.689	4	0	0	0	1	
6	**2016**	C0756	1,390,760	1485270.438	5	1	0	0	0	
7		A0112	640,450	617852.8194	6	0	1	0	0	
8		R3699	1,225,600	1133546.194	7	0	0	1	0	
9		A0136	549,890	549987.8194	8	0	0	0	1	
10	**2017**	C0756	1,475,634	1495938.569	9	1	0	0	0	
11		A0112	680,230	628520.9498	10	0	1	0	0	
12		R3699	1,250,760	1144214.325	11	0	0	1	0	
13		A0136	580,450	560655.9498	12	0	0	0	1	
14	**2018**	C0756	1,540,328	1506606.699	13	1	0	0	0	
15		A0112	728,657	639189.0802	14	0	1	0	0	
16		R3699	1,310,655	1154882.455	15	0	0	1	0	
17		A0136	610,655	571324.0802	16	0	0	0	1	
18	**2019**	C0756	1,587,678	1517274.83	17	1	0	0	0	
19		A0112	726,867	649857.2106	18	0	1	0	0	
20		R3699	1,330,780	1165550.586	19	0	0	1	0	
21		A0136	625,670	581992.2106	20	0	0	0	1	
22	**2020**	A0112	320,450	657858.3084	21	0	1	0	0	
23		R3699	560,560	1173551.683	22	0	0	1	0	
24		A0136	325,890	589993.3084	23	0	0	0	1	
25	**2021**	A0112	540,878	665859.4062	24	0	1	0	0	
26		R3699	789,656	1181552.781	25	0	0	1	0	
27		A0136	536,789	597994.4062	26	0	0	0	1	
28	**2022**	C0756	1,675,787	1543945.156	27	1	0	0	0	
29		A0112	889,568	676527.5366	28	0	1	0	0	
30		R3699	1,579,086	1192220.912	29	0	0	1	0	
31		A0136	835,786	608662.5366	30	0	0	0	1	

Step 2 – Next generate the regression using analyse tool

	A	B	C	D	E	F
1	SUMMARY OUTPUT					
2						
3	*Regression Statistics*					
4	Multiple R	0.745928946				
5	R Square	0.556409992				
6	Adjusted R Square	0.453655917				
7	Standard Error	299395.6524	Standard error *2 =	598791.3047	Standard error *3 =	898186.9571
8	Observations	32				
9						
10	ANOVA					
11		*df*	*SS*	*MS*	*F*	*Significance F*
12	Regression	5	3.03576E+12	6.07153E+11	8.466753932	7.37235E-05
13	Residual	27	2.42022E+12	89637756655		
14	Total	32	5.45598E+12			
15						
16		*Coefficients*	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*
17	Intercept	1321225.004	136776.7525	9.659719072	2.96836E-10	1040582.29
18	t	-3345.691964	5774.719015	-0.579368789	0.567143408	-15194.43666
19	C0756	0	0	65535	#NUM!	0
20	A0112	-624837.683	149809.1671	-4.170890842	#NUM!	-932220.7037
21	R3699	-105798.6161	150142.6944	-0.704653773	0.487060013	-413865.9783
22	A0136	-686011.2991	150696.9329	-4.552257873	0.000101427	-995215.8647
23						
24						
25						
26	RESIDUAL OUTPUT					
27						
28	*Observation*	*edicted Profit in the mor*	*Residuals*			
29	1	1317879.313	35571.6875			
30	2	689695.9375	-73945.9375			
31	3	1205389.313	15910.6875			
32	4	621830.9375	-87030.9375			
33	5	1304496.545	86263.45536			

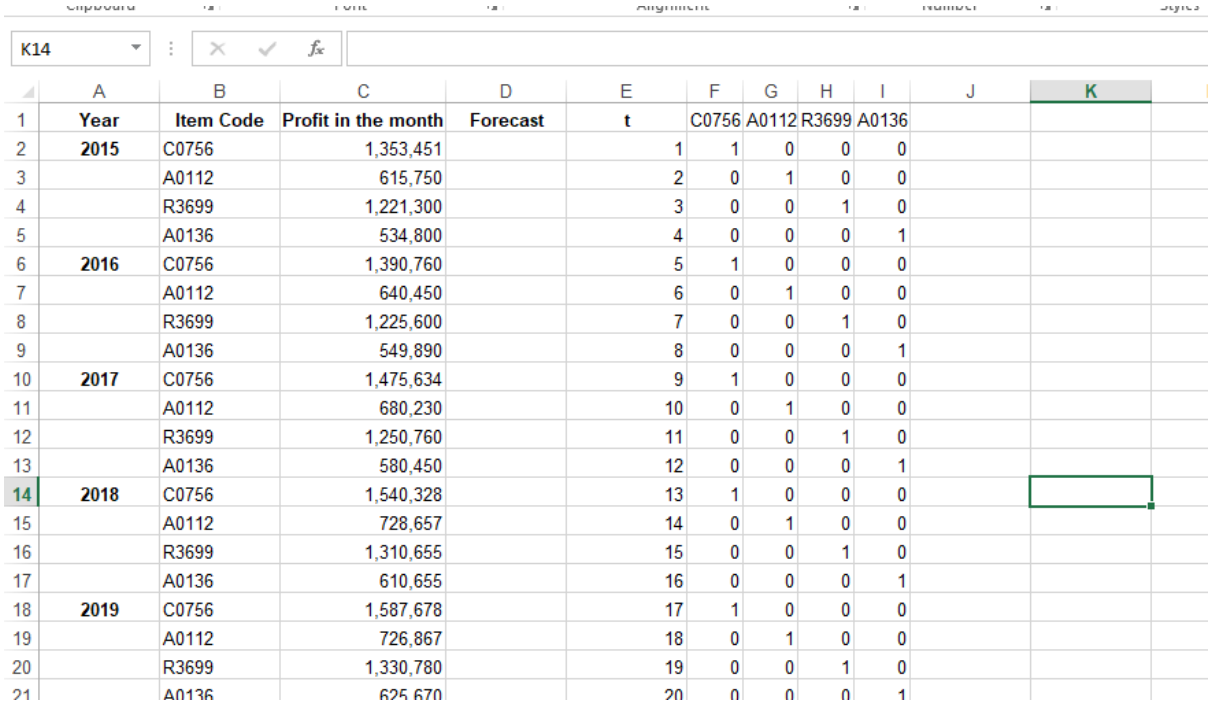
Step 3 : Check if Therese a above value to Standard error*3 in residuals , if not check Standard error*2 in residuals. If there is above value, remove that index value (profit) from the dataset. Again repeat that and get regression 2.

E53

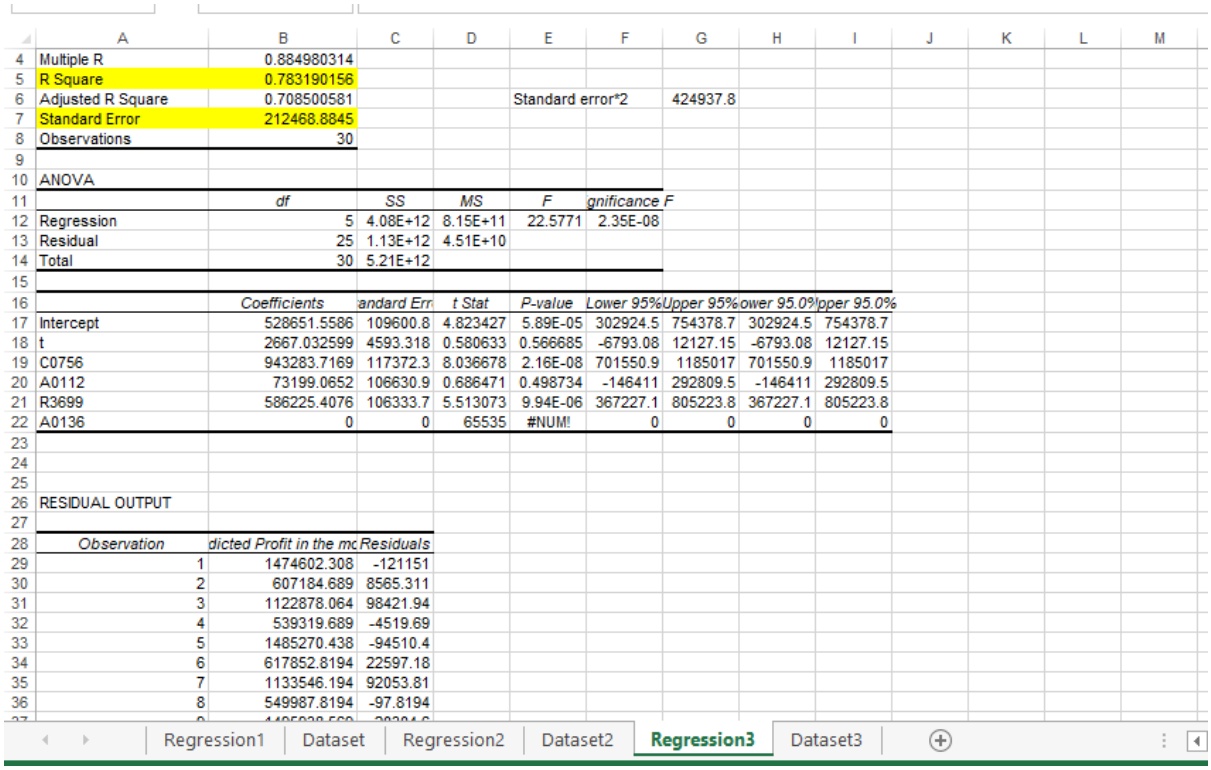
	A	B	C	D	E	F	G	H	I
1	SUMMARY OUTPUT								
2									
3	*Regression Statistics*								
4	Multiple R	0.811567439							
5	R Square	0.658641708							
6	Adjusted R Square	0.56766351							
7	Standard Error	263803.9924		standard error*2 =	527607.9848		standard error*3 =	791412	
8	Observations	31							
9									
10	ANOVA								
11		*df*	*SS*	*MS*	*F*	*Significance F*			
12	Regression	5	3.4912E+12	6.9824E+11	12.54157644	3.71882E-06			
13	Residual	26	1.80941E+12	69592546403					
14	Total	31	5.30061E+12						
15									
16		*Coefficients*	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*ower 95.0%*	*pper 95.0%*
17	Intercept	594347.2829	132751.9695	4.477125914	0.000133623	321471.7015	867222.8644	321471.7	867222.9
18	t	-1098.214634	5359.830697	-0.204897262	0.839248028	-12115.50442	9919.07515	-12115.5	9919.075
19	C0756	798646.1199	138017.0343	5.786576447	4.27673E-06	514948.0429	1082344.197	514948	1082344
20	A0112	65668.57073	132336.8722	0.496222781	0.623906421	-206353.7659	337690.9073	-206354	337690.9
21	R3699	582460.1604	132010.8495	4.412214319	0.000158558	311107.973	853812.3477	311108	853812.3
22	A0136	0	0	65535	#NUM!	0	0	0	0
23									
24									
25									
26	RESIDUAL OUTPUT								
27									
28	*Observation*	*redicted Profit in the mont*	*Residuals*						
29	1	1391895.188	-38444.18815						
30	2	657819.4244	-42069.42438						
31	3	1173512.799	47787.20062						
32	4	589954.4244	-55154.42438						
33	5	1387502.33	3257.670387						

Regression1 | Dataset | **Regression2** | Dataset2 | Regression3 | Dataset3

Likewise get 3 regression analysis and get the 3rd calculated dataset.



	Year	Item Code	Profit in the month	Forecast	t	C0756	A0112	R3699	A0136
2	2015	C0756	1,353,451		1	1	0	0	0
3		A0112	615,750		2	0	1	0	0
4		R3699	1,221,300		3	0	0	1	0
5		A0136	534,800		4	0	0	0	1
6	2016	C0756	1,390,760		5	1	0	0	0
7		A0112	640,450		6	0	1	0	0
8		R3699	1,225,600		7	0	0	1	0
9		A0136	549,890		8	0	0	0	1
10	2017	C0756	1,475,634		9	1	0	0	0
11		A0112	680,230		10	0	1	0	0
12		R3699	1,250,760		11	0	0	1	0
13		A0136	580,450		12	0	0	0	1
14	2018	C0756	1,540,328		13	1	0	0	0
15		A0112	728,657		14	0	1	0	0
16		R3699	1,310,655		15	0	0	1	0
17		A0136	610,655		16	0	0	0	1
18	2019	C0756	1,587,678		17	1	0	0	0
19		A0112	726,867		18	0	1	0	0
20		R3699	1,330,780		19	0	0	1	0
21		A0136	625,670		20	0	0	0	1



4	Multiple R	0.884980314										
5	R Square	0.783190156										
6	Adjusted R Square	0.708500581			Standard error*2		424937.8					
7	Standard Error	212468.8845										
8	Observations	30										
9												
10	ANOVA											
11		*df*	*SS*	*MS*	*F*	*gnificance F*						
12	Regression	5	4.08E+12	8.15E+11	22.5771	2.35E-08						
13	Residual	25	1.13E+12	4.51E+10								
14	Total	30	5.21E+12									
15												
16		*Coefficients*	*andard Err*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*ower 95.0%*	*pper 95.0%*			
17	Intercept	528651.5586	109600.8	4.823427	5.89E-05	302924.5	754378.7	302924.5	754378.7			
18	t	2667.032599	4593.318	0.580633	0.566685	-6793.08	12127.15	-6793.08	12127.15			
19	C0756	943283.7169	117372.3	8.036678	2.16E-08	701550.9	1185017	701550.9	1185017			
20	A0112	73199.0652	106630.9	0.686471	0.498734	-146411	292809.5	-146411	292809.5			
21	R3699	586225.4076	106333.7	5.513073	9.94E-06	367227.1	805223.8	367227.1	805223.8			
22	A0136	0	0	65535	#NUM!	0	0	0	0			
23												
24												
25												
26	RESIDUAL OUTPUT											
27												
28	*Observation*	*dicted Profit in the mc*	*Residuals*									
29	1	1474602.308	-121151									
30	2	607184.689	8565.311									
31	3	1122878.064	98421.94									
32	4	539319.689	-4519.69									
33	5	1485270.438	-94510.4									
34	6	617852.8194	22597.18									
35	7	1133546.194	92053.81									
36	8	549987.8194	-97.8194									

Regression1 | Dataset | Regression2 | Dataset2 | **Regression3** | Dataset3

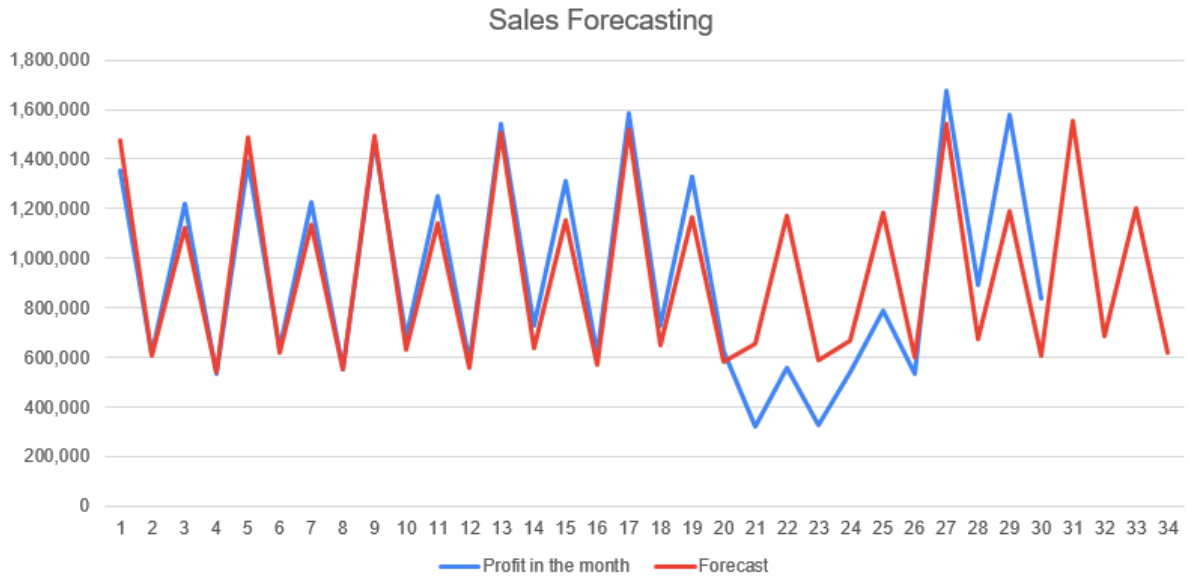
Step 3: 3rd dataset after calculating

	A	B	C	D	E	F	G	H	I	J
1	**Year**	**Item Code**	**Profit in the month**	**Forecast**	**t**	C0756	A0112	R3699	A0136	
2	**2015**	C0756	1,353,451	1474602.308	1	1	0	0	0	
3		A0112	615,750	607184.689	2	0	1	0	0	
4		R3699	1,221,300	1122878.064	3	0	0	1	0	
5		A0136	534,800	539319.689	4	0	0	0	1	
6	**2016**	C0756	1,390,760	1485270.438	5	1	0	0	0	
7		A0112	640,450	617852.8194	6	0	1	0	0	
8		R3699	1,225,600	1133546.194	7	0	0	1	0	
9		A0136	549,890	549987.8194	8	0	0	0	1	
10	**2017**	C0756	1,475,634	1495938.569	9	1	0	0	0	
11		A0112	680,230	628520.9498	10	0	1	0	0	
12		R3699	1,250,760	1144214.325	11	0	0	1	0	
13		A0136	580,450	560655.9498	12	0	0	0	1	
14	**2018**	C0756	1,540,328	1506606.699	13	1	0	0	0	
15		A0112	728,657	639189.0802	14	0	1	0	0	
16		R3699	1,310,655	1154882.455	15	0	0	1	0	
17		A0136	610,655	571324.0802	16	0	0	0	1	
18	**2019**	C0756	1,587,678	1517274.83	17	1	0	0	0	
19		A0112	726,867	649857.2106	18	0	1	0	0	
20		R3699	1,330,780	1165550.586	19	0	0	1	0	
21		A0136	625,670	581992.2106	20	0	0	0	1	
22	**2020**	A0112	320,450	657858.3084	21	0	1	0	0	

Regression1 | Dataset | Regression2 | Dataset2 | Regression3 | **Dataset3**

Then we get a more accurate forecasting.

Appendix C


Sales Forecasting
1,800,000
1,600,000
1,400,000
1,200,000
1,000,000
800,000
600,000
400,000
200,000
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34
Profit in the month
Forecast