DATA MODELLING AND ANALYTICAL TECHNIQUES EMPLOYED FOR SALES FORECASTING

Author
ANDREA GONZALEZ DUENAS

Co-Author
Dr RACHEL JOHN ROBINSON

Email: info@rachel-johnrobinson.de
Affiliation: Arden University, Berlin, Germany
ORCID: 0000-0002-1079-1358

Data Modelling and Analytical Techniques Employed for Sales Forecasting

Author: Andrea Gonzalez Duenas

Published by
Bohr Publishers, Chennai, India
Old Door No. 4, New Door No. 3, Rishilaya Building, 2nd Street,
Golden Avenue, Devi Karumariamman Nagar, Velachery,
Chennai – 600 042, India

©2023 Bohr Publishers
All rights reserved. No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording or storage in any information or retrieval system, without the prior permission of the copyright owner.

ISBN: 978-93-92892-67-7

Price: Rs.

Typeset and Cover designed by
Chennai Publishing Services Pvt. Ltd., Chennai, India
Printed by Ridha Prints Pvt. Ltd., Madurai, India

ABSTRACT

Sometimes companies forget that sales is more than just selling and that analysing and forecasting is a key factor for a company´s success. The company under study does not have any type of sales analysis or forecast method at the moment, which has brought inventory and demand fulfilment problems. The researcher has decided to do a sales analysis per family products and focus on the top 10 products with highest sales for the forecasting analysis. There are different methods to create a forecast analysis, most of them and the more accurate ones are based on quantitative methods rather than qualitative methods. The company under study does not has a normal distribution of the sales and has a seasonality of 12 months; having 3 years and 7 months of sales information, the researcher has decided that the best approach for the forecast analysis was a quantitative method based on historical data. The analysis shows that the top 10 products represent more than 50% of the total pieces sold, that only three products families have products in the top 10 category, and that 2 out of 10 products forecasts cannot be trusted and need to be reviewed on a monthly basis due to quality rejections from customers.

Keywords: Sales analysis, forecast, predictive analysis, data analysis, historical data

Statement of Declaration: The author declared that there is no conflict of interest.

TABLE OF CONTENTS

LIST OF TABLES **VII**

LIST OF FIGURES **IX**

1. INTRODUCTION **1**
- 1.1. STUDY PURPOSE 1
- 1.2. BACKDROP 1
- 1.3. STUDY RATIONALE 2
- 1.4. CASE STUDY BACKGROUND AND PROBLEM STATEMENT 3
- 1.5. RESEARCH AIM & OBJECTIVES 4
- 1.6. DISSERTATION STRUCTURE 4

2. LITERATURE REVIEW **6**
- 2.1. INTRODUCTION 6
- 2.2. SALES ANALYSIS 7
- 2.3. SIGNIFICANCE TEST 8
- 2.4. SALES FORECASTING 9

3. METHODOLOGY AND METHOD **12**
- 3.1. INTRODUCTION 12
- 3.2. PHILOSOPHICAL ASSUMPTIONS 12
- 3.3. RESEARCH QUESTIONS 13
- 3.4. VALIDITY AND RELIABILITY 14
- 3.5. DATA SELECTION AND COLLECTION 16
- 3.6. ETHICS AND BIAS 17
- 3.7. LIMITATIONS 17

4. RESULTS AND DISCUSSION **19**
- 4.1. INTRODUCTION 19
- 4.2. RESULTS 19
  - 4.2.1 Gather Data 20
  - 4.2.2 Discover and Assess Data 20
  - 4.2.3 Clean and Validate Data 23
  - 4.2.4 Transform and Enrich Data 25

4.2.5 Data Analysis 25
4.2.6 Products with the Highest Sales in the Past 3.5 Years 26
4.2.7 Testing Significance of Top 10 Products 28
4.2.8 Demand Fluctuation Analysis 32
4.2.9 Forecasting 34
4.3. SUMMARY 42

5. CONCLUSIONS 44
5.1. INTRODUCTION 44
5.2. GENERAL CONCLUSIONS 44
5.3. RESEARCH QUESTIONS CONCLUSIONS 45
5.3.1. Objective 1 45
5.3.2. Objective 2 46
5.3.3. Objective 3 46
5.3.4. Objective 4 47
5.4. RECOMMENDATIONS 47
5.5. ERROR AND LIMITATIONS 48
5.6. RECOMMENDATIONS FOR FURTHER STUDY 49

REFERENCES 50

LIST OF TABLES

Table 1.	Forecasting techniques	11
Table 2.	Variable description	21
Table 3.	Descriptive analysis statistics	22
Table 4.	Normality test	22
Table 5.	Duplicate products	24
Table 6.	Products per family	25
Table 7.	Top 10 products	26

LIST OF FIGURES

Figure 1.	Flowchart data analysis process	19
Figure 2.	Data sample	20
Figure 3.	Total pieces sold distribution	22
Figure 4.	Normality box-plot	23
Figure 5.	Data filter by family	24
Figure 6.	Unite price "0"	24
Figure 7.	Data with format and top 10 variables	25
Figure 8.	Total sales per family per year	26
Figure 9.	Top sales from the top 10 products	27
Figure 10.	Top 10 sales per year	27
Figure 11.	Final data in SPSS	28
Figure 12.	SPSS variables view	29
Figure 13.	Performing test, part 1	29
Figure 14.	Performing test, part 2	30
Figure 15.	Test results	31
Figure 16.	Total pieces sold according to the top 10 category per year	31
Figure 17.	Top 10 total sales per trimester	32
Figure 18.	Bottles sales per quarter	33
Figure 19.	Lid sales per quarter	33
Figure 20.	Label sales per quarter	34
Figure 21.	Sales data per SKU worksheets	35
Figure 22.	Forecast sheet, steps 1–3	35
Figure 23.	Forecast sheet, steps 4–8	36
Figure 24.	Forecast results (table and chart)	37
Figure 25.	T45020 forecast for 2022 and 2023	37
Figure 26.	EMS042 forecast for 2022 and 2023	38
Figure 27.	BZE060 forecast for 2022 and 2023	38
Figure 28.	BTWS150 forecast for 2022 and 2023	38
Figure 29.	BZE040 forecast for 2022 and 2023	39
Figure 30.	ETGE01000 forecast for 2022 and 2023	39
Figure 31.	T45180 forecast for 2022 and 2023	40
Figure 32.	EMS041 forecast for 2022 and 2023	40
Figure 33.	EMS043 forecast for 2022 and 2023	41
Figure 34.	BTWS040 forecast for 2022 and 2023	41
Figure 35.	Top 10 products sales and forecast summary	43