INSTITUTO
UNIVERSITÁRIO
DE LISBOA

Understanding Price variations of electronic components in supply chain

Rafaela Filipa de Jesus Besugo

Master's degree in Integrated Business Intelligence Systems

Supervisor:
Doctor João Carlos Amaro Ferreira, Assistant Professor with Habilitation
ISCTE-IUL - University Institute of Lisbon

Co-Supervisor:
Fabian Tinkl-Hennighausen
Siemens SA

September 2023

TECNOLOGIAS
E ARQUITETURA

Department of Information Science and Technology

Understanding Price variations of electronic components in supply chain

Rafaela Filipa de Jesus Besugo

Master's degree in Integrated Business Intelligence Systems

Supervisor:
Doctor João Carlos Amaro Ferreira, Assistant Professor with Habilitation
ISCTE-IUL - University Institute of Lisbon

Co-Supervisor:
Fabian Tinkl-Hennighausen
Siemens SA

September 2023

Acknowledgements

If some years ago people would tell me I was finishing a master's degree in business intelligence, I would tell those people were crazy.

But now here we are, and life is being full of challenges such as finishing this dissertation on an area that I was not expertise.

With hard work and will, is with happiness and feeling of achievement that I can say "I AM A MASTER!!"

I want to thank my supervisor João Carlos Ferreira for the guidance.

To my co-supervisor Fabian Tinkl-Hennighausen, thank you so much for all the help and patient with me during these months. Really glad I met you and could work with you.

Also thank you to Siemens SA to be able to do this dissertation with this amazing company that give us these opportunities of knowing other work areas and knowing another people inside the company.

Thank you to my friend that I could take from this master's degree that also supported me during this period: Bernardo Santos.

Finally, a big thank you to my family, friends and boyfriend Sandro that always believes in me and supports all my craziness.

Rafaela Filipa de Jesus Besugo

Resumo

A indústria dos componentes eletrónicos caracteriza-se por rápidos avanços tecnológicos e por uma cadeia de abastecimento global complexa. Um dos desafios enfrentados pelos fabricantes, distribuidores e utilizadores finais é a variação dos preços.

Compreender os fatores que contribuem para as variações de preços na cadeia de abastecimento é essencial para otimizar as estratégias de aquisição, gerir os custos e garantir a sustentabilidade do negócio.

Esta tese de mestrado tem como objetivo investigar e analisar as causas às variações de preços dos componentes eletrónicos, com destaque para fatores-chave como a procura do mercado, a dinâmica da cadeia de abastecimento e as influências externas.

Ao obter uma compreensão abrangente destes fatores, as partes interessadas podem tomar decisões informadas para atenuar o impacto das variações de preços e aumentar a sua vantagem competitiva.

A fixação de preços é essencial para as empresas no mercado da eletrónica. De acordo com a análise das estatísticas dos componentes eletrónicos, os *chips* são os mais caros, em média, seguidos dos relés. Geograficamente, os componentes são abundantes na Europa e na América do Norte, incluindo a Suíça, a Alemanha, Nova Iorque e *Boston*, sendo os preços médios mais elevados na Suécia e em França. Os componentes mais recentes tendem a custar mais quando são introduzidos pela primeira vez devido ao facto de serem novidade no mercado e não serem tão reconhecidos, por isso existe a necessidade de elevar o preço. A variável SE Grade tem impacto no preço, o que mostra que as empresas devem prestar atenção aos componentes contrafeitos e introduzir este fator na escolha do componente a comprar. É aconselhada a implementação de medidas adicionais de garantia de qualidade para componentes com elevada probabilidade de contrafação.

Palavras-chave: Inteligência no Negócio, Análise de Preços, Variações, Cadeia de Abastecimento, Materiais Eletrónicos, Ciência de Dados.

Abstract

The electronic components industry is characterized by rapid technological advancements and a complex global supply chain. One of the critical challenges faced by manufacturers, distributors, and end-users is the fluctuating prices of electronic components.

Understanding the factors that contribute to price variations within the supply chain is essential for optimizing procurement strategies, managing costs, and ensuring business sustainability.

This master thesis aims to investigate and analyze the underlying causes of price variations in electronic components, with a focus on key factors such as market demand, supply chain dynamics, and external influences. By gaining a comprehensive understanding of these factors, stakeholders can make informed decisions to mitigate the impact of price variations and enhance their competitive advantage.

Pricing is essential for businesses in the cutthroat electronics market. According to an analysis of electronic component statistics, Chip components are the most expensive on average, followed by Relays. Geographically, components are abundant in middle Europe and western North America, including Switzerland, Germany, New York, and Boston, with the highest median prices being found in Sweden and France. Newer components tend to cost more when they are first introduced due to the fact that they are new to the market and not as recognized, so there is a need to raise the price. The SE Grade has an impact in price showing that companies must pay attention to the counterfeit components and introduce this factor when choosing what component to buy. Also implement additional quality assurance measures for components with a high probability of counterfeiting.

Keywords: Business Intelligence, Price Analysis, Variations, Supply Chain, Electronic Materials, Data Science.

Index

Figures Index

Table Index

Map Index

List of abbreviations

AI – Artificial Intelligence

ANOVA – Analysis of Variance

CRISP-DM - Cross Industry Standard Process for Data Mining

EC – Electronic Components

ICT - Information and Communication Technologies

PLM – Product Lifecycle Management

PRISMA - Preferred Reporting Items for Systematic Reviews and Meta Analyses

R&D - Research and Development

SC – Supply Chain

SE – Silicon Expert

USA – United States of America

VIF – Variance Inflation Factor

CHAPTER 1

Introduction

The electronics sector has experienced unprecedented growth during the last few decades. This expansion has been made possible by the increasing need for electronic components in a number of sectors, including technology, automotive, healthcare, and telecommunications. Despite this exponential growth, the supply chain price fluctuation of electronic components continues to be a significant issue for the business. Thorough research and comprehension are required since the volatility of component pricing has the potential to significantly affect product cost, profitability, and competitiveness.

1.1 Problem Statement

Price changes of electronic components across the supply chain can be attributed to a number of reasons, including the cost of raw materials, manufacturing expenses, demand-supply dynamics, geopolitical concerns, market competition, and unforeseen disruptions like natural disasters or pandemics. Despite their significance, these variations, their underlying causes, and their larger implications for the global economy and the electronics industry are not fully understood. This information gap hinders effective strategic planning, budgeting, and risk mitigation in the electronics industry.

1.2 Purpose of the Study

The investigation and comprehension of the mechanisms influencing price fluctuations of electronic components throughout the supply chain is the primary goal of this study. This thesis intends to give a thorough examination of these aspects, their interdependencies, and their impact on the pricing of the electronic components by combining quantitative and qualitative research methodologies.

1.3 Research questions

The primary goal guiding this study is: "What can impact the price variations in electronic components in the supply chain?"

To answer this overarching goal, the study will address several sub-goals such as: How do geopolitical factors and market competition influence these price variations? Are risk factors also impacting the price?

1.4 Significance of the Study

Researchers, politicians, suppliers, and manufacturers in the electronics industry might find value for the study's findings. By having a greater grasp of the factors influencing price variations in electronic components, industry stakeholders may be able to anticipate future price changes and create effective strategies for managing these variations. The study may also establish the framework for more research in this area and contribute to the body of information already available on supply chain management in the electronics industry.

1.5 Methodologic approach

In Supply Chain, where various stakeholders are engaged in a collaborative and cooperative process, a powerful model to address the Big Data Analytics issues is CRISP-DM [3].

CRISP-DM is the standard and an industry-independent process model for applying data mining projects [1]. It consists of six phases from business understanding to deployment.

Figure 1.1 shows the phases that must be followed when using CRISP-DM method [2]:

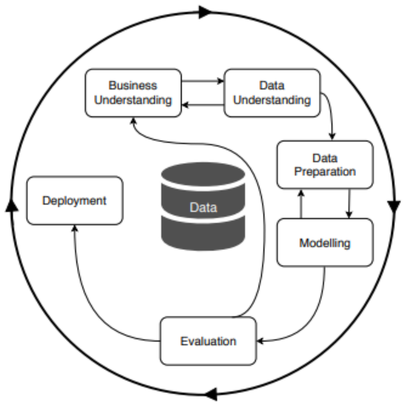


Figure 1.1 - The CRISP-DM process model of data mining

- **Business Understanding →** To gain a sense of the resources that are needed and accessible, the business environment should be evaluated. One of the most crucial components of this phase is choosing the data mining objective. First, the data mining kind and data mining success criteria should be described. The creation of a project plan is required.
- **Data understanding →** In this phase, it is crucial to gather data from data sources, explore and describe it, and assess the quality of the data.
- **Data preparation →** By establishing inclusion and exclusion criteria, data selection should be carried out. Cleaning data can be used to deal with poor data quality. It is necessary to generate derived characteristics based on the model that is being used.
- **Modeling →** The development of the test case, the model, and the modeling approach are all part of the data modelling step. Specific parameters must be set in order to build the model. It is appropriate to compare the model to the evaluation criteria and choose the best ones for assessment.
- **Evaluation →** Results are compared to the established company objectives during the evaluation phase. As a result, it is necessary to analyze the findings and decide the next course of action. Another argument is that a broad examination of the process is necessary.
- **Deployment →** It may be software or a final report. Planning the deployment, monitoring it and maintenance must be done.

1.6 Work Structure

With the goals and methodology established, there will be six chapters including:

Chapter 2 → PRISMA approach is explained to outline a systematic literature evaluation of the literature review in pricing analysis of electronic components in supply chains.

Chapter 3 → CRISP-DM Methodology is explained and applied following four phases.

Chapter 4 → Results are shown from the analysis done.

Chapter 5 → Conclusions and future work is presented giving it some challenges.

Chapter 6 → Articles and websites used to support literature review and analysis done.

CHAPTER 2

Literature Review

2.1 Review Methodology

The systematic literature review was conducted using the Preferred Reporting Items for Systematic Reviews and Meta Analyses (PRISMA) approach. By making the authors' findings and research easier for readers to understand, PRISMA aims to help authors improve and more thoroughly assess the systematic review and meta-analysis reporting process. In order to find and choose research publications linked to price and data analysis on supply chain disruption, the search on the subject was done in electronic database: Scopus.

The research phase was mainly done by searching for the main topics in title, abstract and keywords of the articles. Some of the topics were "Price Analysis", "Supply Chain", "Electronic Components".

The 32 selected articles were considered based on their abstract and full text. If the article was interesting for the study, it was saved in the Mendeley tool to be referenced later.

Query: "data analysis" OR "pric* analysis" AND "component*" OR "electronic component*" AND "supply chain*" OR "supply chain* disruption*".

Figure 2.1 represents the process of selecting the documents.

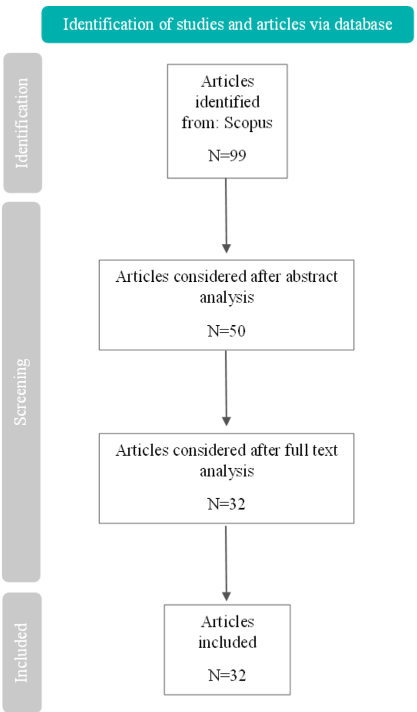


Figure 2.1 – PRISMA Methodology applied in this study

2.2 Results

Analyzing the results of the articles considered, Figure 2.2, shows that there is a range of 24 years because it was concluded that articles on electronic components in supply chains were not abundant in more recent years, so the search was extended. Most of the articles are from the years 2007/2009, 2013/2014 and 2021/2022.

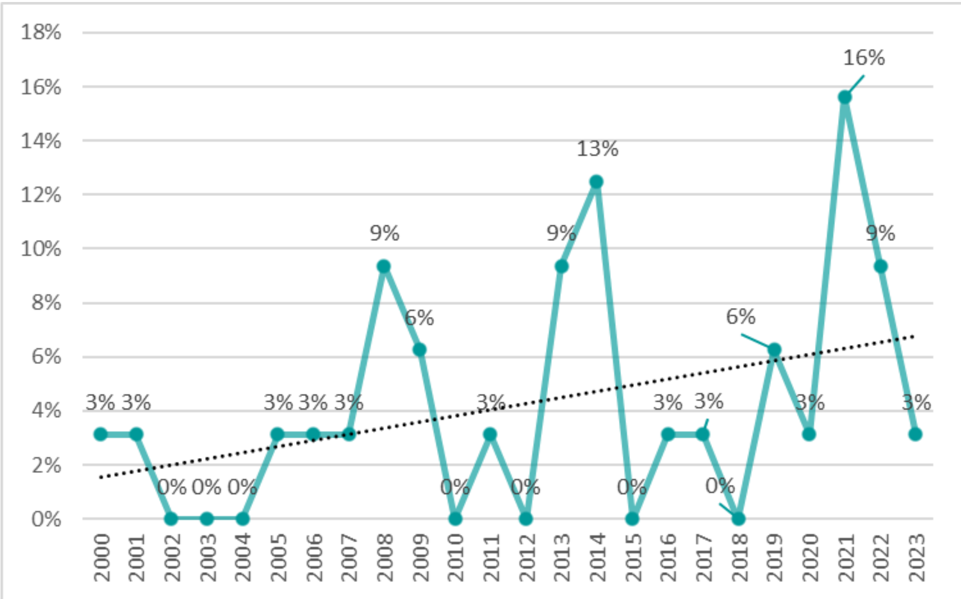


Figure 2.2 - Articles years

By having a wider range of years, the topics "component" and "supply chain" were the most abundant with 29% and 37%, respectively, possible to check in Figure 2.3.

The topic "Industry 4.0" was the least mentioned topic with 4%.

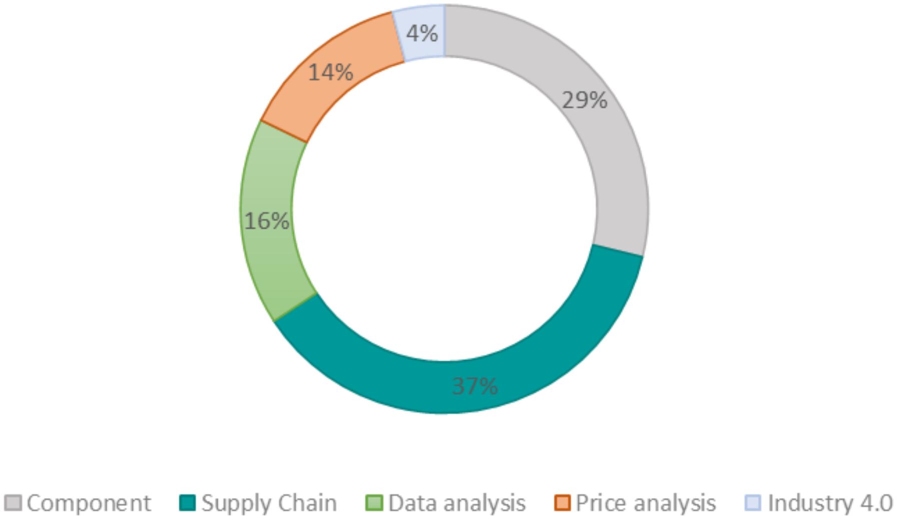


Figure 2.3 - Article topics percentage

Figure 2.4 represents de correlations between the several topics most present in the studies and articles considered. The map was done in VOSviewer. The words that have a higher occurrence are the ones with the biggest cluster – Supply Chain, Product and Analysis.

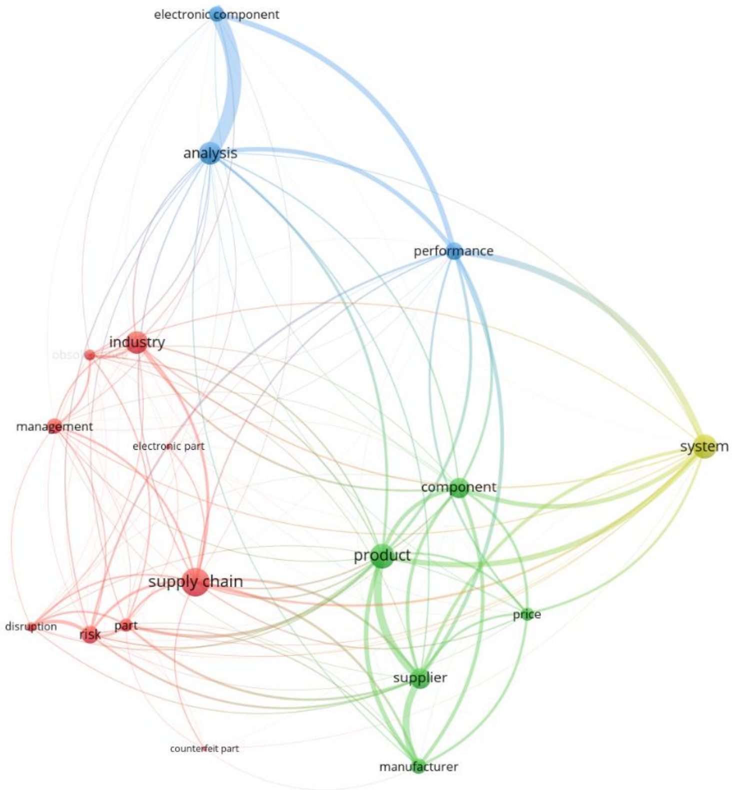


Figure 2.4 - Correlation map of article topics

2.2.1 Electronic components

Electronic components (ECs) are frequently used in crucial systems in the industrial, transportation, aviation, power production, and cybersecurity industries [4] and even more, manufacturers are compelled to reuse and recycle electrical devices because of resource depletion and environmental concerns.

Refurbishing and reusing these items are one green manufacturing technique. Decision-makers and manufacturing strategies are impressed by secondary market consumer input. Reusing and recycling electronic parts is a good way to cut waste, increase productivity, and lower costs. Value has been extracted from e-waste through sustainable development methods, such as closed-loop supply chain management [5], [6].

The 21st century's increased globalization has led to a major increase in the complexity of the electronics supply chain. Today, electronic components are produced, distributed, and bought all over the world. [7].

The selection of the most efficient and affordable components affects both the design and functionality of an electrical device choosing appropriate components for a specific application is influenced by technical requirements as well as availability, standardization, and pricing [8].

2.2.2 Price

Price receives a lot of attention because manufacturers want suppliers to be able to deliver components with process capacity that satisfy the anticipated quality level and are also reasonable. Low cost is one strategic advantage that businesses can employ in a competitive market. Simply said, customers like to get high-quality goods at a discount. As a result, the two main factors are cost and quality.

Differentiated items in diverse consumer marketplaces frequently exhibit comparable price patterns, which can be partially accounted for by trends and volatility in common components. For businesses looking to increase the accuracy of their pricing estimates as well as learn more about how customers value certain product characteristics, the observed connections between the prices of different items may be a valuable source of data [9].

A high-quality and expensive supplier is undesired in the supplier selection process due to financial considerations. Customers like vendors who supply products of acceptable quality at competitive prices [10] when purchasing a product, they consider both the price and the product's quality level, which is influenced by the quality of each component. In accordance with this demand structure, the manufacturer determines the final product's market price, and each supplier makes quality investments [11] willing to upgrade or switch from their preferred quality and style product to a more expensive one if the price is right for them [12].

Price and quality are still crucial in today's competitive market, especially when supply chain partners' cooperation on quality investments and pricing might affect supply chain competitiveness.

Manufacturers provide a variety of identical items with varying features or options in many different industries. In this situation, a product's demand is influenced by both its own price and the costs of competing goods. Sales of linked products are positively (negatively) impacted by a product's price change [13]. In real life, this substitution effect occurs frequently.

For instance, when the manufacturer finishes a component in his inventory, he is faced with two choices: to buy again the same component or to move the production over another product which does not include that component. The price of the parts contained in the product consider a multi-agent supply chain environment because the importance of quick forecast models based on an agent framework is much more important in these days [14].

The operations network links supply-side, manufacturing, and demand-side businesses and operations in the supply chain to run the most efficient business that can react fast to a variety of customer demands. Decision-making at the strategic, tactical, and operational levels is required to design the most cost-effective supply chain network [14].

2.2.3 Supply chain disruption

The fact that industries with a high import content are the main sources of pricing pressure shows that interruptions in global supply networks may be of major importance. It concludes that supply considerations are the primary cause of the current rise in producer prices in industry and in each of the manufacturing sectors experiencing the greatest price pressures, accounting for at least 80% of the increase [15].

A supply chain disruption is a mismatch between supply and demand that would result in backordered parts if there were no mitigating factors such as buffered parts or second sources. While the primary effect of a disruption is the same, the source/cause of each event varies. There are, at least, four disruption categories: part-specific, supplier-specific, customer-specific, and external [16].

A supply chain disruption is the occurrence of an unplanned, unusual triggering event somewhere in the supply chain or its environment and a resulting circumstance that materially undermines the supply chain companies' ability to do their regular activity. In comparison to normal commercial operations, it is a unique circumstance for the affected organizations.

Also, an organization needs a comprehensive knowledge extraction from all its business segments. The management level of each supply chain needs to be able to continuously analyze the processes within their units under their control from various angles, the outcome of this analysis depends on the degree of their as-is process awareness, for instance, a distribution part of a SC manages a complex process whose output directly depends on the performance of various parts [17], [18].

The disruption is distinguished by its severity, as well as by its direct and indirect repercussions, and it has a certain probability of occurring. Supply chain disruptions involve time pressure because the resulting harm typically depends on time, meaning that decisions on mitigation must be made quickly. Depending on the severity, various phrases like glitch, disturbance, accident, tragedy, or crisis may be used. Disruptions in a supply chain can occur both inside and outside of the network. As a result, they could be very different. The supply chain is impacted differently by crises with very distinct characteristics, such as a supplier's financial breakdown and an earthquake that damages industrial capacity.

Many supply chain decision-makers were caught off-guard by the intensity of some disasters which highlighted the lack of preparedness in many supply chains. For this reason, many firms have started to take supply chain disruptions more seriously and to rethink their supply chain strategy and design [19].

A supply chain's continuity and regular functioning could be severely hampered by a failure at any one point in the network. The size or severity, frequency, likelihood or probability of occurrence, and length are the factors that describe a disruption's profile [20].

Many firms have used backup sourcing as the most straightforward and yet highly effective procurement technique to overcome uncertain supply disruptions and to lessen the effects of supply shortages.

Counterfeit, data loss, human error, operational risks, transactional hazards, risks of foodborne illness and knowledge asymmetries are just a few of the demand risks that can cause supply chain disruptions [21].

2.2.4 Counterfeit

Since counterfeits have been a hazard to the electronic supply chain for decades, stakeholders for safety-critical industrial systems place a high price on the authenticity and quality of electronic components [4] becoming an issue for the electronics business as well as for everyone who uses electronics daily, whether it be for home care or transportation.

The issue of counterfeit electronic parts and the risks involved with them is one of the main areas of concern in the supply chain for electronic parts [22] becoming a $600 billion industry and a menace that accounts for more than 8% of the world's merchandise trade. There are currently few non-intrusive, cost-effective methods for identifying counterfeit electronic parts [5].

Particularly, the supply chain for electronic parts is now greatly concerned about counterfeit electronic parts. These can be rescued from scrap or new or surplus parts that have undergone some sort of modification.

Because of the reliance on defensive and offensive systems, fake electronics can seriously compromise security and safety. Because dishonest individuals, including brokers, are able to enter the supply chain, it is getting more difficult to detect counterfeit parts or parts with a dubious history [23].

Counterfeit parts are produced for a variety of reasons, including the quick obsolescence of electronic components and the lengthy wait times of components from reliable sources. The price of inspection/testing procedures makes it more difficult for part users to identify counterfeit electronic components, while the lack of high-quality verification tools in the electronic part supply chain and the availability of inexpensive tools and parts to create counterfeits make the counterfeiting of electronic parts a relatively low risk operation for counterfeiters [24]. There have been 1,800 of suspect counterfeit electronic parts being, and the total number of parts involved is around 1 million.

2.2.5 Obsolescence issue

When a part can no longer be purchased from the original manufacturer, it is deemed obsolete [25] and the performance of a system can be impacted, which presents a difficult issue for spare parts management [26], being also subject to high frequency involuntary procurement obsolescence.

Many electronic parts are only procurable from their original manufacturer for a few years, then they are discontinued in favor of newer, higher performing parts.

When parts become obsolete, considerable resources must be expended to resolve the problem. A potential drawback of reusing the same component in multiple products that has been articulated by electronics system manufactures is described by the following scenario:

All products in use (depend on) a common part. An unexpected problem develops with the part. Instead of fighting a fire for one product, exists simultaneously an issue on every product, i.e., effectively created a "single point of failure" scenario by reusing the part.

Short term problems that usually only affect a limited number of products that share the part for a short period of time, e.g., you receive a bad batch or lot of parts. Lead-time issues and inventory management are temporary supply chain disruptions that impact a manufacturing or product-specific organization and are therefore not the focus of this paper.

Long-term problems such as a fundamental supply chain or wear out problem that affects all products that share the part and for which a permanent solution (often a replacement part) must be found. For long-term problems, under some conditions, the financial penalties associated with the delays may offset savings due to part commonality [27].

Long life cycle goods require the availability of compatible parts for production and service, which over time exposes the parts to supply chain disruptions such vendors leaving the market, allocation challenges, hazards associated with counterfeit parts, and part obsolescence [28].

Obsolescence has been a long-standing problem for businesses who manufacture items. The supporting systems' hardware and software components are particularly impacted. The pressure and lack of interoperability in this industry are exacerbated by the absence of standardized, open, and modular technologies for specific timing limitations. Systems based on Industry 4.0 provide this [29].

2.2.6 Industry 4.0

Companies and sectors are facing a number of changes and problems as a result of Industry 4.0. Software has a particularly significant impact. In fact, the use of software has substantially increased demand, particularly for electric and electronic parts. In this situation, it is obvious that there is a propensity to speed up development, add complexity, and go through numerous iterations (sprints) to ensure that the final product meets market and consumer expectations [30].

Information technology-driven changes to production systems are referred to as industry 4.0. Information and communication technologies (ICT) and manufacturing processes are integrated in this concept, which is why ICT and manufacturing technologies are seen as two pillars of Industry 4.0 [31]. It is now possible to collect and analyze data from much equipment, enabling more rapid, adaptable, and efficient production of higher-quality goods at lower costs.

The manufacturing revolution will boost productivity, alter economics, promote industrial development, and alter workforce demographics, ultimately altering the competitiveness of businesses and areas [32].

2.2.7 Data Analysis

Automatic forecasting has developed into a critical component of business analytics tools for many businesses. Traditional manual data analysis is impractical in such settings due to the abundance of data, the short analytical life cycle, and the rapid pace of company processes [33].

2.2.8 Summary

In Table 2.1 is a summary of the research articles mentioned above.

Table 2.1 – Research Articles summary

Topic	References	# Doc	% Doc
Electronic Components	[5] - [9]	5	17%
Price	[10] - [16]	7	23%
Supply Chain disruption	[17] - [22]	6	20%
Counterfeit	[5], [23] - [25]	4	13%
Obsolescence	[26] - [30]	4	13%
Industry 4.0	[31] - [33]	3	10%
Data Analysis	[34]	1	3%

CHAPTER 3

Price analysis on electronic components in supply chain

As mentioned in the previous chapter, the methodology that this dissertation will be following is the CRISP-DM. Chapter 3 will cover three phases – Business Understanding, Data Understanding and Data Preparation**.**

3.1 Business Understanding

The first focus is comprehending the project's requirements and goals from a business standpoint. Based on this understanding, a data mining problem definition and a rough project schedule are then created [34].

For this analysis, the stakeholders have the following points as requirements:

- Reducing Manual effort;
- Increase quality;
- Increase transparency;
- Time saving;
- Cost optimization (money saved because of the solution);
- Avoiding errors;
- Sharing data knowledge;
- Cheaper prices of procured material;
    - Identification of compliance risk for materials;
    - Saving resources and time;
    - Increasing customer satisfaction with high quality products.

-

A lot of manual effort and time-consuming clarification rounds are needed by Supply Chain Management, Procurement and Product Lifecycle Management (PLM) departments when it comes to supplier strategy, supply chain efficiency.

Insufficient data quality and availability of information in PLM systems hinders development in research and development (R&D), cost value engineering and component engineering like obsolescence information of parts, better alternatives or finding similar parts already available within SIEMENS.

The solution would be combining various sources from PLM, external data sources and Artificial Intelligence (AI) Services. Then, is possible to enhance:

- Today's PLM process since PLM systems are often fragmented, with no immediate access to business information;
- Digital transformation within the procurement organization;
- Coordination of product development with the supply chain.

3.2 Data Understanding

The data understanding phase begins with initial data collecting and continues with actions to familiarize with the data, find data quality issues, gain preliminary comprehension, or select intriguing subsets to generate hidden information hypotheses.

Data understanding and business understanding are closely related. It is necessary to have at least a basic grasp of the available data in order to formulate the data mining challenge and create the project plan [34].

The data used in this analysis is from SiliconExpert (SE), one of the data suppliers used by Siemens. With SE is possible to extract data and analyze it so the company can avoid redesigns and mitigate obsolescence as well managing risk of more than a billion parts [35].

To start to use the data is necessary to go to Snowflake system and then the SE data is available to download as csv. file the several part details output that will be used in the analysis [36]:

- Summary data;
- Pricing data;
- Risk data.

3.2.1 Summary Data

Summary data gives the general information about a specific part (electronic component).

In Table 3.1 is possible to verify which are the variables in Summary data and its description:

Table 3.1 - Metadata Summary Data [36]

Summary Data	
Variable	**Description**
DataProviderID	The SE component id
PLName	Product line name
PartNumber	Part number that matches the input part number in SE database
Manufacturer	Manufacturer of the part number
ManufacturerID	The SE manufacturer id
PartDescription	Description for product name, it consists of values of some important parameters to give brief information about each product name
FamilyName	Root Part Number of the Product Name
CountriesOfOrigin	A list of countries at which the last significant manufacturing process is carried out

3.2.2 Pricing Data

In the pricing dataset, is possible to analyze the price information from each part.

In Table 3.2 is possible to verify the variables and description in Price dataset:

Table 3.2 - Metadata Price Data [36]

Pricing Data	
Variable	**Description**
MinimumPrice	Minimum Price according to SiliconExpert Resources
AveragePrice	Average Price according to SiliconExpert Resources
MinLeadtime	Minimum Lead Time according to SiliconExpert Resource
Maxleadtime	Maximum Lead Time according to SiliconExpert Resource
LastUpdatedate	The last updated date of component/EEE price

The pricing dataset is a CSV file composed of 907,872 entries and 10 columns.

In Table 3.3, some information regarding the variables from the dataset can be verified, concluding that all columns except REQUESTEDCOMID have null values, with a higher number of nulls in columns 3,4, 5, 6, and 7. The most abundant data type is the object type with five columns, followed by float64 with four columns and only one as int64.

Table 3.3 - Pricing data Information

#	Column	Non-null count	Dtype
0	REQUESTEDCOMID	907872	int64
1	PARTNUMBER	907173	object
2	MANUFACTURER	907173	object
3	MINIMUMPRICE	297885	float64
4	AVERAGEPRICE	297885	float64
5	MINLEADTIME	218274	float64

6	MAXLEADTIME	218274	float64
7	LASTUPDATEDATE	297885	object
8	MANUFACTURER_CLEANED	901476	object
9	PARTNUM_MAN	901476	object

3.2.3 Risk Data

In risk dataset is available the information about lifecycle or obsolescence for each part number.

In Table 3.4 is possible to verify the variables and description in Risk dataset.

Table 3.4 - Risk data Metadata

Risk Data	
Variable	**Description**
RohsRisk	Restriction of Hazardous Substances in Electrical and Electronic Equipment - Risk based on ROHS availability (Low, Medium, High)
MultiSourcingRisk	Risk based on Crosses availability (Low, Medium, High)
InventoryRisk	Risk based on the availability of the part in the authorized distributors inventories
LifecycleRisk	Estimated Lifecycle Risk (Low, Medium, High)
PredictedObsolescenceYear	
LifecycleStage	Current part lifecycle stage. One of the following: Preliminary, Introduction, Growth, Mature, Decline, Phase Out, Obsolete or Discontinued
YearsEOL	Estimated Number of years till Obsolescence
NumberOfDistributors	Count of authorized distributors that have the part in their inventories

NumberOfOtherSources	Count of available crosses for that Part
CrossesPartCategory	Indicates whether or not this product category typically contains crosses. Please note that some parts may not carry crosses even though the product category may be crossable
SEGrade	Using historical data & detailed analysis of past counterfeit events, SiliconExpert has found patterns for parts most likely to be targeted by counterfeiters
LastUpdateDate	The last time SE checked and imported the lifecycle status for the part
YeolComment	Comment on estimated number of years till obsolescence

3.3 Data Preparation

The process of creating the final dataset from the initial raw data falls under the data preparation step. Tasks for data preparation may be completed more than once and not necessarily in that order. Table, record, and attribute selection, data cleaning, the creation of new attributes, and data processing for modeling tools are among the tasks [34].

From SE to Snowflake, already some changes are done in the data:

1. PARTNUMBER and MANUFACTURER columns are included in the datasets;
2. Each PARTNUMBER receives an ID named REQUESTEDCOMID;
3. A new column is added named MANUFACTURER_CLEANED that consists in the full name of the MANUFACTURER;
4. A new column is added named PARTNUM_MAN that's joins the PARTNUMBER and MANUFACTURER_CLEANED.

3.3.1 Data selection and joining data

For the data preparation was important to analyze and select the important attributes from each dataset so the analysis could be as complete and correct as possible.

From dataset 1 was important to have the Manufacturer so is possible to know who is producing the materials, minimum and average price being the main focus of the analysis, latitude and longitude so is possible to have a geographical analysis, product line and description for each component, minimum and maximum lead time could be interesting to check for the same material type, how much time it takes to be delivered.

For dataset 2 the information of price and risk data was joined. This is in a separate dataset because the risk data has no summary data information available and then a lot of null data was being generated.

The inclusion of specific columns, such as ROHS Risk, MultiSourcing Risk, Inventory Risk, Lifecycle Risk, Predicted Obsolescence Year, Lifecycle Stage, Number of Distributors, Number of Other Sources, Crosses Part Category, SE Grade, is of utmost importance in the field of price analysis for electronic components and equipment. These variables provide insights into the dynamic market for electrical components, assisting enterprises and stakeholders in decision-making and risk-mitigation.

Availability of parts compatible with the Restriction of Hazardous Substances directive is evaluated in this column's ROHS Risk section. Low ROHS risk suggests simpler sourcing of environmentally friendly components, which could result in cost savings and a competitive advantage.

The danger of supply chain interruptions is reduced by the availability of many sources for a component. In pricing negotiations, a high MultiSourcing Risk warns purchasers of probable difficulties in sourcing alternatives.

Inventory Risk: Prices may be greatly impacted by the availability of components in approved distributor inventories. While low risk implies cost stability, large inventory risk may signal shortage, which could result in higher pricing.

Determining a component's lifetime risk enables procurement tactics and price negotiations to be influenced by the possibility of supply shortages or discontinuation in the future.

Due to expanded supply alternatives, pricing may become more competitive the more distributors that carry a component and pricing discussions are impacted by the availability of cross references or substitutes. Better pricing is frequently the result of more options.

Crosses Part Category: Identifying whether a product category typically has crosses helps gauge supply and demand dynamics, affecting pricing strategies.

Historical data on counterfeit events provides insights into the authenticity of components. Higher SE grades can indicate lower counterfeit risks and potentially lower prices.

The dataset has 12 columns and 297885 entries.

For data cleaning techniques in some of the columns the nulls were dropped or replaced with helpful information, also was checked duplicate entries. Upfront these are explained in more detail.

3.3.2 Data cleaning

In first step duplicate entries were verified and eliminated. 4793 entries were deleted, and the dataset is now composed with 903079 entries and 12 columns.

Table 3.5 represents the change on non-null values after duplicate entries check.

Table 3.5 - Dataset Information

#	Column	Non-null count	Dtype
0	REQUESTEDCOMID	903079	int64
1	PARTNUMBER	902380	object
2	MANUFACTURER	902380	object
3	ADDRESS	293294	object
4	PHONENUMBER	293294	object
5	EMAIL	293294	object
6	LATITUDE	293294	float64
7	LONGITUDE	293294	float64
8	MINIMUMPRICE	293795	float64
9	AVERAGEPRICE	293795	float64

10	MINLEADTIME	217970	object
11	MAXLEADTIME	217970	object
12	LASTUPDATEDATE	293795	object
13	MANUFACTURER_CLEANED	896683	object
14	PARTNUM_MAN	896683	object
15	SUMMARYDATA_PARTDESCRIPTION	903079	object
16	SUMMARYDATA_PLNAME	903079	object

Regarding null values, Table 3.6 shows the amount of these in each column of the dataset.

Table 3.6 - Dataset null values count

#	Column	Null count
0	REQUESTEDCOMID	903079
1	PARTNUMBER	902380
2	MANUFACTURER	902380
3	ADDRESS	0
4	PHONENUMBER	0
5	EMAIL	0
6	LATITUDE	0
7	LONGITUDE	0
8	MINIMUMPRICE	609284
9	AVERAGEPRICE	609284
10	MINLEADTIME	685109
11	MAXLEADTIME	685109

12	LASTUPDATEDATE	609284
13	MANUFACTURER_CLEANED	6396
14	PARTNUM_MAN	6396
15	SUMMARYDATA_PARTDESCRIPTION	0
16	SUMMARYDATA_PLNAME	0

Before proceeding with the analysis was important to address the quality data issue that is the null values. If not possible to fill them with information, they must be deleted.

For columns 8 and 9 the null values were deleted besides that the 609284 entries represented around 67% of all dataset, but after clarifying with the stakeholders it was discussed that was fine continuing the analysis with the around 300000 entries left.

Table 3.7 represents the null count after deleting null values in MINIMUMPRICE and AVERAGEPRICE columns.

Table 3.7 - Data set null count after cleaning

#	Column	Null count
0	REQUESTEDCOMID	0
1	PARTNUMBER	0
2	MANUFACTURER	0
3	ADDRESS	0
4	PHONENUMBER	0
5	EMAIL	0
6	LATITUDE	0
7	LONGITUDE	0
8	MINIMUMPRICE	0
9	AVERAGEPRICE	0

10	MINLEADTIME	75825
11	MAXLEADTIME	75825
12	LASTUPDATEDATE	0
13	MANUFACTURER_CLEANED	1698
14	PARTNUM_MAN	1698
15	SUMMARYDATA_PARTDESCRIPTION	0
16	SUMMARYDATA_PLNAME	0

After analyzing the content from columns MINLEADTIME and MAXLEADTIME the null value corresponds to the figure "N/A", so its fine to keep these entries as 0, representing that certain PARTNUMBER is always available. Doesn't make sense to delete it, because then, around 76000 entries with price data would be lost.

To solve the remaining null entries, for MANUFACTURER_CLEANED the 1698 entries were filled with the content from column MANUFACTURER and for PARTNUM_MAN the content from PARTNUMBER and MANUFACTURER were joined. This way, it was possible to keep the 1698 entries in the dataset.

For the risk dataset, Table 3.8 represents the non-null count of the several columns.

Table 3.8 - Null count of dataset 2

#	Column	Null count	Dtype
0	REQUESTEDCOMID	0	int64
1	PRICINGDATA_MINIMUMPRICE	0	float64
2	RISKDATA_ROHSRISK	0	object
3	RISKDATA_MULTISOURCINGRISK	21	object
4	RISKDATA_INVENTORYRISK	21	object
5	RISKDATA_LIFECYCLERISK	21	object
6	RISKDATA_PREDICTEDOBSOLESCENCEYEAR	13912	float64

7	RISKDATA_LIFECYCLESTAGE	13912	object
8	RISKDATA_NUMBEROFDISTRIBUTORS	21	float64
9	RISKDATA_NUMBEROFOTHERSOURCES	21	float64
10	RISKDATA_CROSSPARTYCATEGORY	21	object
11	RISKDATA_SEGRADE	13912	float64

The columns with 21 null registers were deleted for representing a lower number from the entire dataset, not harming the analysis.

The columns RISKDATA_LIFECYCLESTAGE and RISKDATA_SEGRADE can be found in the dataset. These columns likely contain details regarding the SE Grade and lifecycle stage of the electronic components.

In these columns, null entries often denote missing or ambiguous data. It implies that the dataset will miss information on the lifecycle stage or SE Grade for some records.

To handle the nulls for RISKDATA_LIFECYCLESTAGE and RISKDATA_SEGRADE, was decided to populate these entries with the value "Unknown" as opposed to leave them as null (missing data).

By filling up nulls with "Unknown" these records are covered for analysis and are not being excluded. This is crucial because discarding records that have missing data could result in the loss of potentially critical data, even though still be able to draw conclusions or notice trends from these records.

The expected year of obsolescence for specific goods or products is represented by the risk data field RISKDATA_PREDICTEDOBSOLESCENCEYEAR.

Like the null entries in the previous columns, these entries show that some records years of obsolescence are unknown or cannot be forecast.

Instead of leaving these entries blank, it was decided to give them the value "0000." This option effectively indicates that there is no estimated obsolescence year available for these items.

Null entries are filled with the number "0000" to indicate that there is no expected obsolescence data. It enables the preservation of these records for study and guarantees that the absence of this prediction is clearly stated.

3.3.3 Feature Engineering

The column SUMMARYDATA_PLNAME gives the information of the product line that the electronic component is inserted.

However, was possible to join several of these components in more general categories and then are easier to find and analyze.

An example would be the LEDs category. The process to create the LEDs category was searching for all material components containing “LEDs” or “LED” in the product line description. This way would be possible to analyze the LEDs in an all group.

Not every material could be added in a category for being a unique material for the product line description.

In total, 49 general categories were created by grouping components with the same product line.

The columns COUNTRY and CONTINENT were created to understand from LATITUDE and LONGITUDE variables which countries and continents can impact more the price in the sense that if, for example, the major of electronic parts are coming from a certain country or continent, maybe the price will be high.

The next steps after Data Preparation are the Modelling and Evaluation that will be done in the next chapter as Results.

CHAPTER 4

Results

4.1 Visualization Analysis on Summary and Price

Starting the analysis on the dataset 1 that contains the summary and price data, Figure 4.1 shows that in Category variable is possible to conclude that the most existing electronic part are the Connectors, followed by Resistors and Relays. These three categories have around 25000 components each.

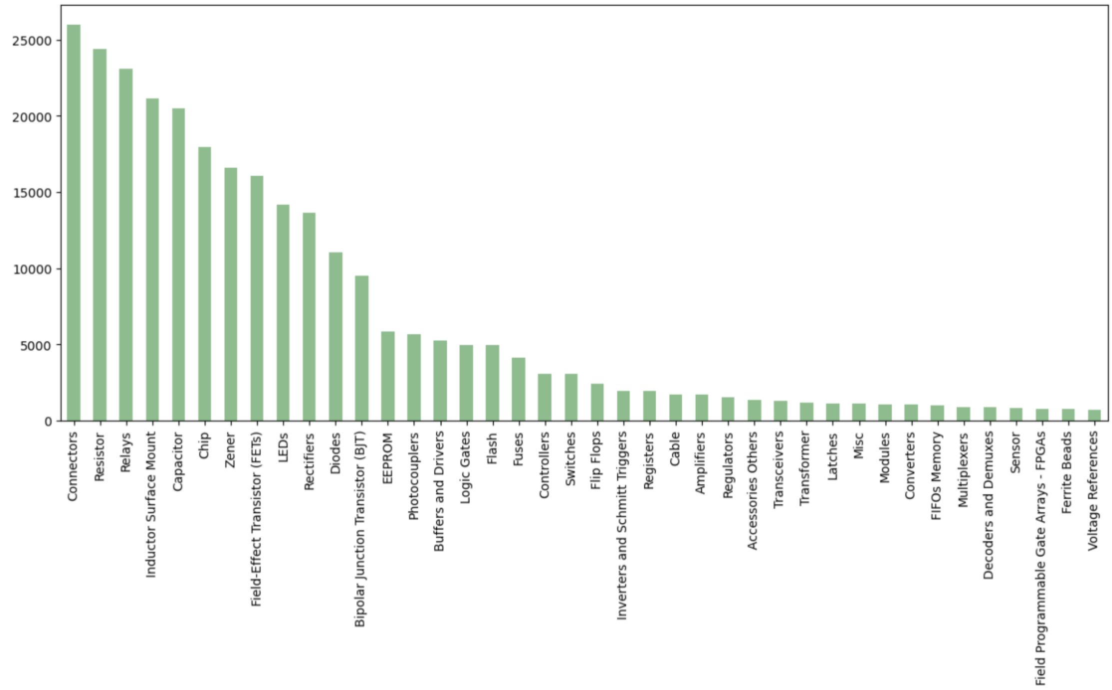


Figure 4.1 – EC count by Category

Upon examination of the pricing data in Figure 4.2, it is evident that the Chip category possesses the highest median price, approximately 33, succeeded by the Relays category with a median price of around 15. Although the Connectors category demonstrates a higher availability, its median price remains relatively low, at approximately 3.

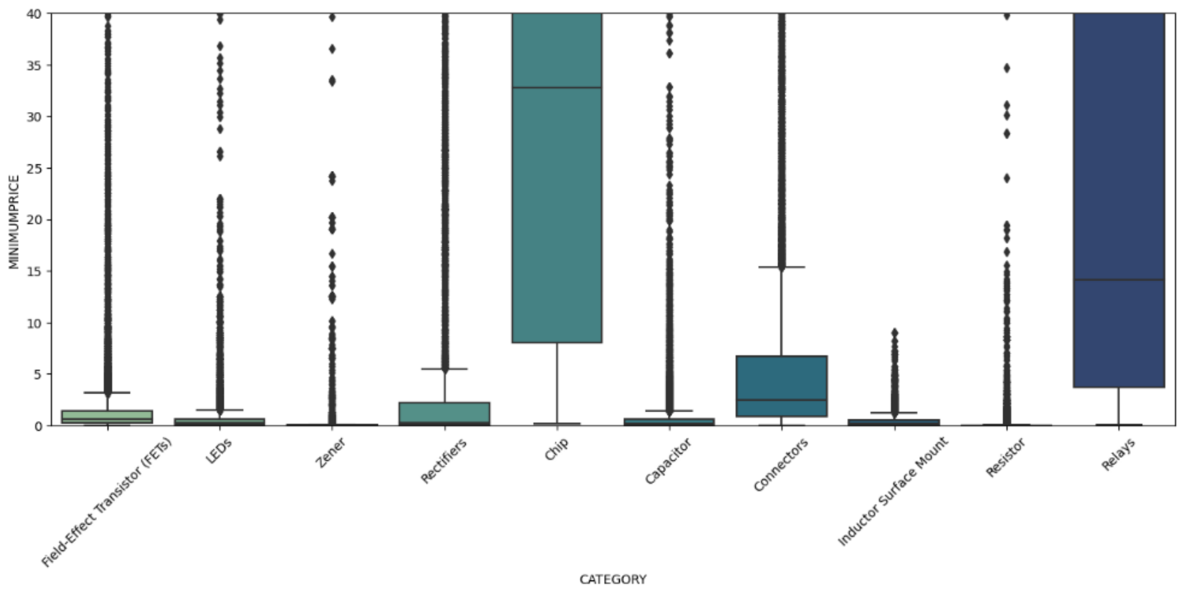


Figure 4.2 - Boxplot showing the distribution of Price across Category (Top 10)

For the Manufacturer, in Figure 4.3, the company Vishay is the one that sells more electronic components, around 20000, followed by Microchip Technology and Texas Instruments that sells between 15000 and 17500 parts. SIEMENS comes in 25th selling around 5000 electronic parts.

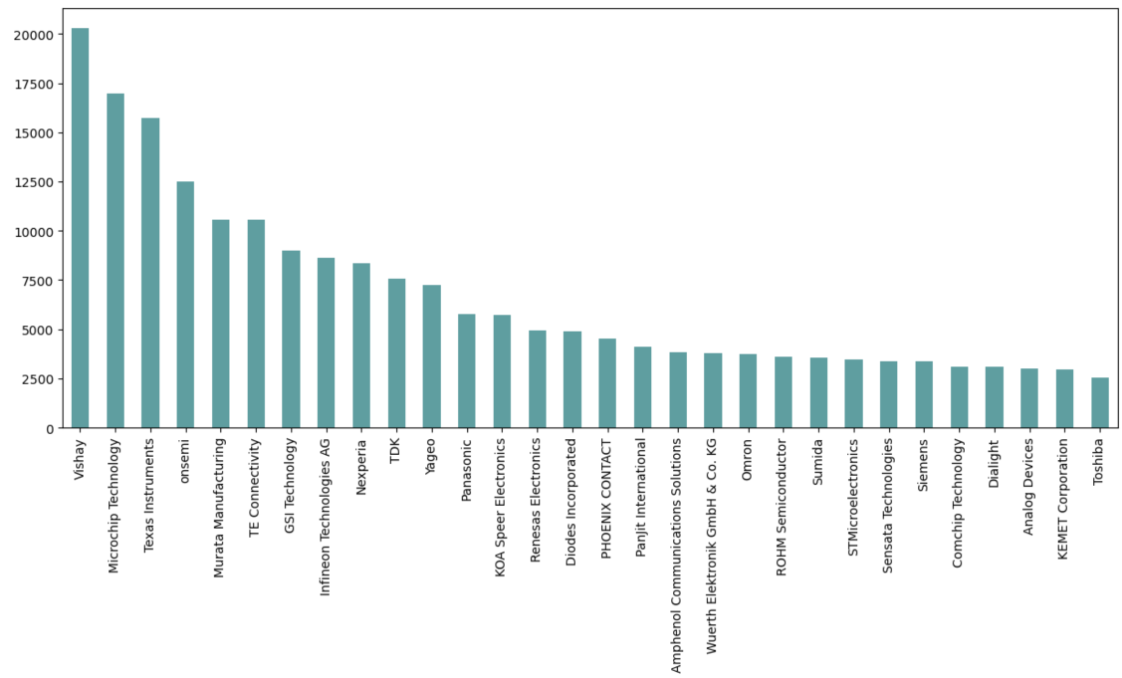


Figure 4.3 - EC count by Manufacturer

Upon cross-referencing with pricing data, it has been determined that GSI Technology implements the highest median price of 70, followed by TE Connectivity with a median price of 7, as depicted in Figure 4.4.

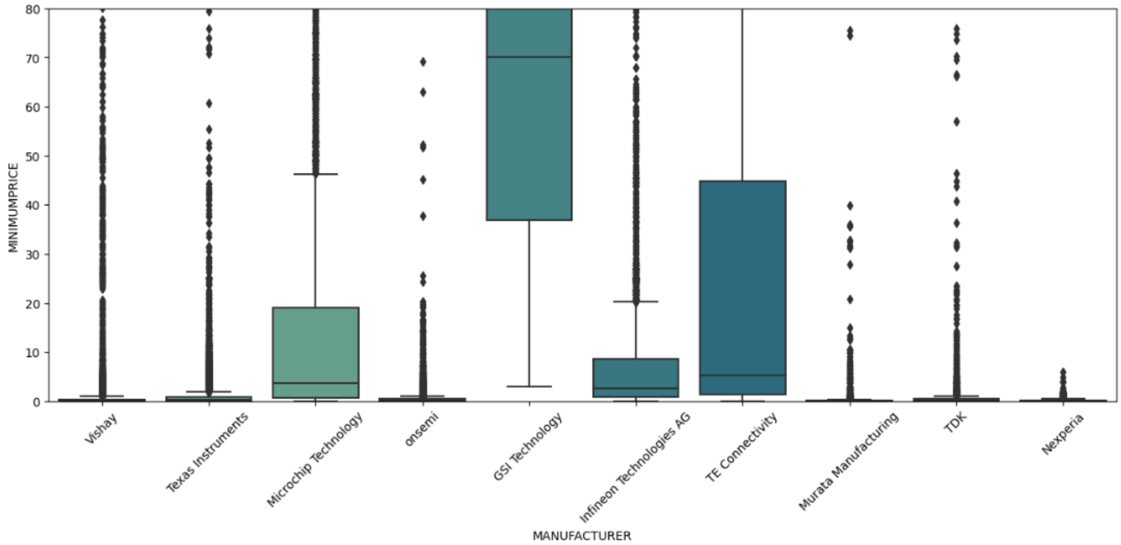


Figure 4.4 - Boxplot showing the distribution of Price across Manufacturer (Top 10)

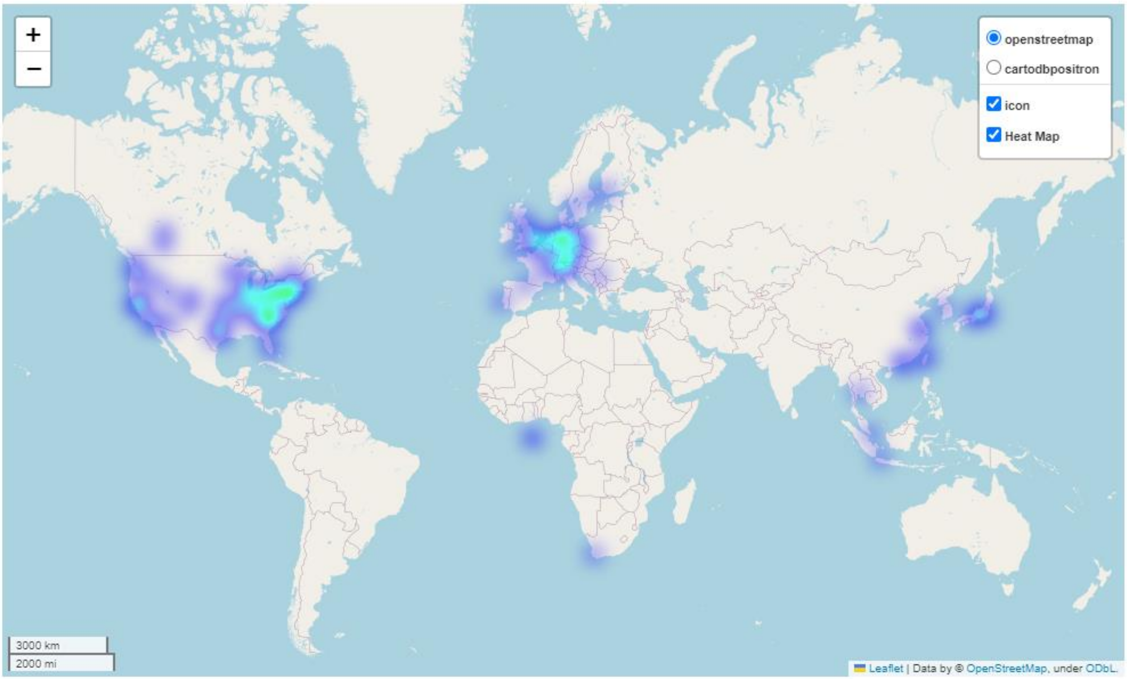
It is also interesting to check if GSI Technology had similar components to those of other companies and to compare the prices each company charged. This company has electronic components similar to those of Infineon Technologies AG, which is also represented in Table 4.1 as one of the 10 most expensive companies. It is possible to conclude that the variable "Manufacturer" certainly has an impact on the price.

Table 4.1 - Comparison of the same ECs between companies GSI Technology and Infineon Technologies AG

Part Description	Price GSI Technology	Price Infineon Technologies AG
SRAM Chip Sync Dual 1.8V 36M-bit 1M x 36 0.45ns 165-Pin FBGA Tray	**48.33**	83.84
SRAM Chip Sync Dual 1.8V 144M-bit 16M x 9-bit 0.45ns 165-Pin FBGA Tray	206.83	**25.00**

SRAM Chip Sync Single 1.8V 144M-bit 8M x 18 0.45ns 165-Pin FBGA Tray	227.99	**25.00**
SRAM Chip Sync Single 1.8V 36M-bit 1M x 36 0.45ns 165-Pin FBGA Tray	**49.56**	90.03
SRAM Chip Sync Single 1.8V 36M-bit 2M x 18 0.45ns 165-Pin FBGA Tray	**48.30**	94.30
SRAM Chip Sync Dual 1.8V 36M-bit 2M x 18 0.45ns 165-Pin FBGA Tray	44.45	**19.66**
SRAM Chip Sync Single 1.8V 72M-bit 4M x 18 0.45ns 165-Pin FBGA Tray	138.35	**25.00**
SRAM Chip Sync Dual 1.8V 144M-bit 8M x 18 0.45ns 165-Pin FBGA Tray	229.71	**25.00**

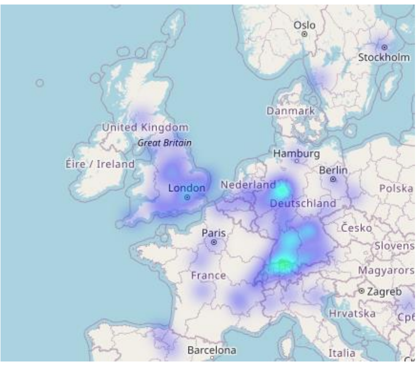
Regarding geolocation analysis, in Map 4.1, is possible to verify that the countries with more affluence of electronic parts are produced in middle Europe and west North America.



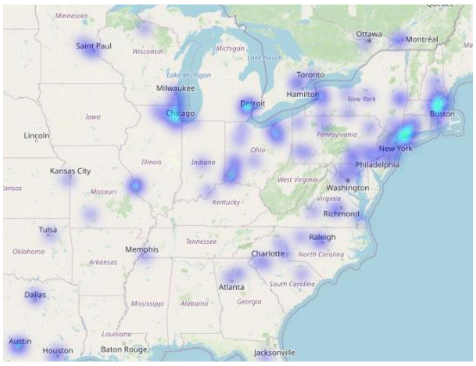
Map 4.1 - Affluence of EC worldwide to be bought

In Maps 4.2 and 4.3 is possible to verify in more detail that, in Europe, Switzerland and Germany are highlighted on getting more parts produced.

In North America, New York and Boston are also the cities with most electronic components produced.



Map 4.3 - Region of Europe that has more EC produced



Map 4.2 - Region of USA with most EC produced

Cross checking with price, in Figure 4.5, is possible to check that parts produced in Sweden have a higher medium price of around 60, followed by parts produced in France with a medium price of 40. Parts produced in Serbia have also one of the higher medium prices of around 20, what indicates that doesn’t sell too many parts, but the medium price can be high.

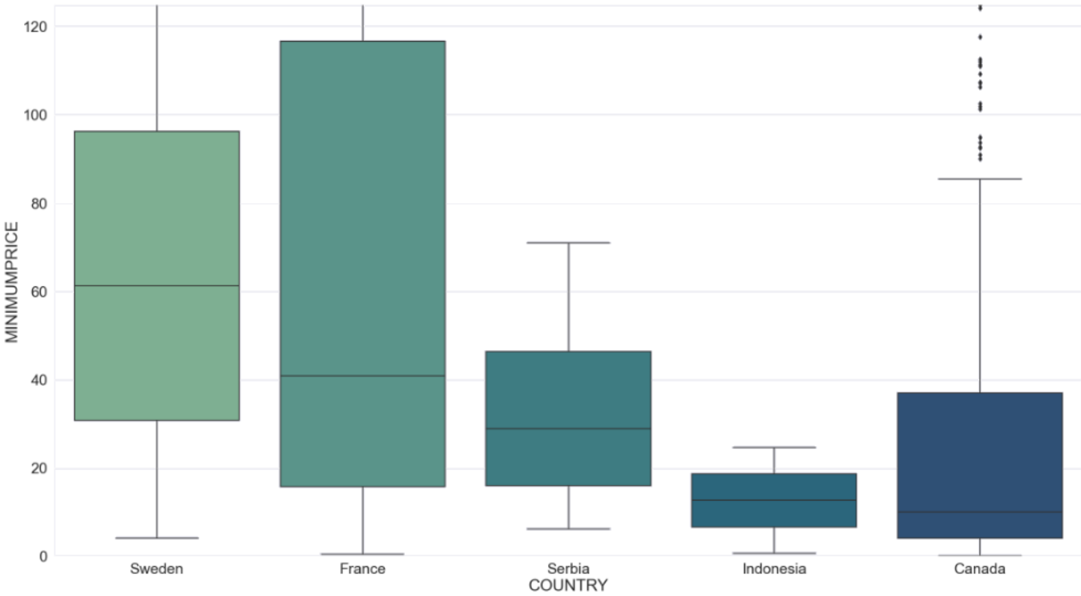


Figure 4.5 - Boxplot showing the distribution of Price across Country - Top 5

Also was compared if the parts that are sold from Sweden are also sold in other countries. In Table 4.2 is possible to check that for example, in Japan the same parts can be bought with a lower price than Sweden.

Table 4.2 - Price comparison of the same EC between Countries

Part Description	Price Sweden	Price Germany	Price France	Price United States	Price Canada	Price Switzerland	Price Japan
Power Transformers	4.09	4.53	8.81	26.76	94.88	34.89	2.58
Bearings	20.05	-	-	1278.77	-	27.27	9.36
Telecom Transformers	6.98	13.2	-	1.73	-	-	0.60

Per continent, Europe has the higher medium price of around 4 compared with the other continents. In Figure 4.6, is possible to conclude that Europe is the main source of electronic components in terms of amount available but also practice a higher value.

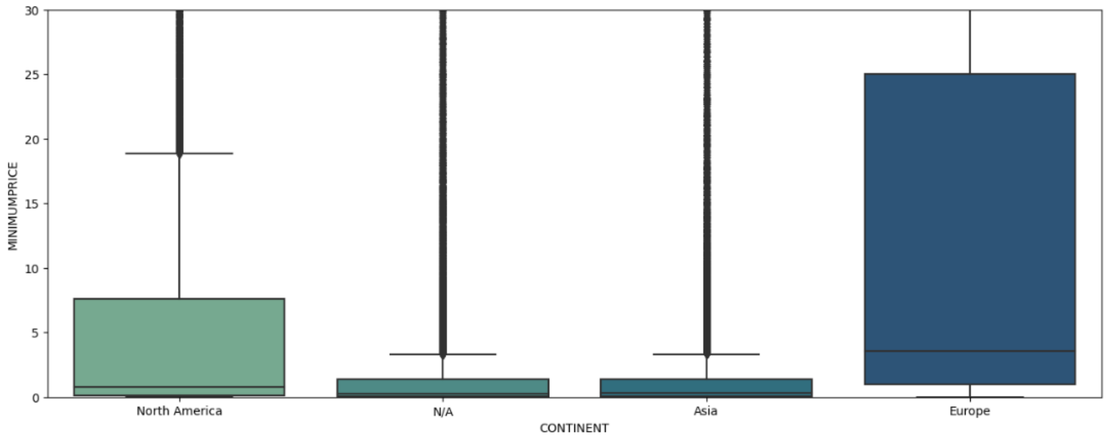


Figure 4.6 - Boxplot showing the distribution of Price across Continent

In resume, companies can focus on categories with high demand and sales, such as Connectors, Resistors, and Relays, to ensure they have a robust product portfolio in these categories.

For categories with high median prices like Chips, companies can explore opportunities to offer more premium or value-added products to capture higher margins.

Analyzing manufacturer insights studying the methods of industry giants like as Vishay, Microchip Technology, and Texas Instruments, manufacturers might potentially imitate their successful techniques. Businesses like GSI Technology, who have high median prices, could wish to highlight the superiority and distinctiveness of their goods in order to defend their premium pricing.

Also think about growing their operations or distribution networks in areas like middle Europe, West North America, Switzerland, Germany, New York, and Boston where there is a greater demand for electronic components.

Such information can be used to modify pricing methods. Parts made in France and Sweden, for example, might command a higher price, but parts made in Serbia might need to provide a lower price to increase sales.

Given that the median price of electronic components is higher in Europe, it is possible that European businesses will be able to keep or raise their prices, which will represent the market worth of their goods.

For cross-country comparison, to make sure they stay competitive in various areas, businesses can undertake pricing benchmarking and competitive analysis. Price modifications might be required to maintain competitiveness if the identical parts are marketed in several nations.

4.2 Visualization Analysis on Risk and Price

It has been observed that 85% of the components (around 250000) are in the Mature stage, characterized by market saturation, intensified competition, and price wars. On the other hand, components in the Introduction stage have relatively higher average prices, which can be attributed to minimal sales and efforts to create initial demand and awareness among customers. [37].

In Figure 4.7 is possible to check the number of electronics components by Lifecycle Stage.

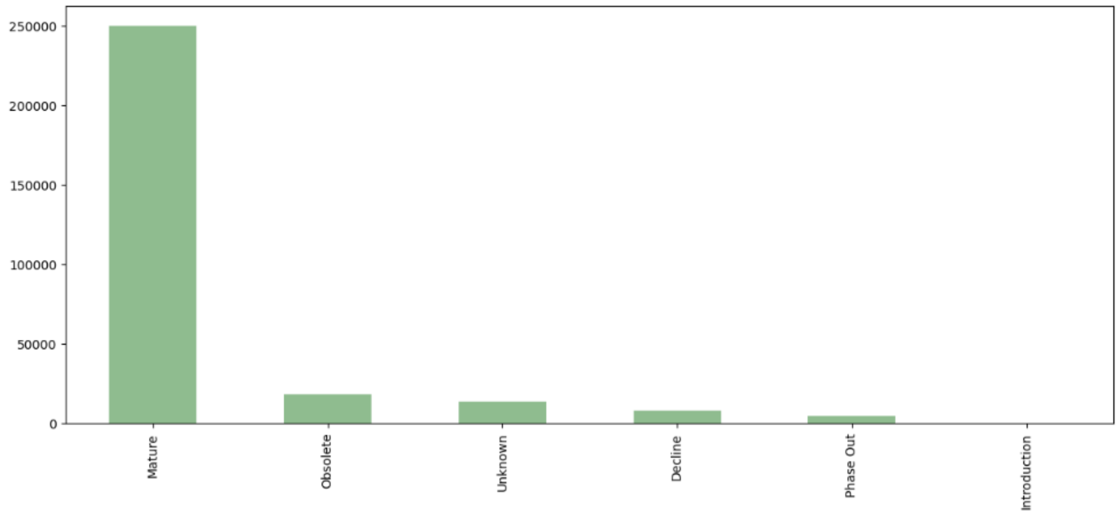


Figure 4.7 - EC count by Lifecycle Stage

However, in Figure 4.8, the introduction stage has the higher medium price of around 3 comparing with the mature stage. This is because customers are introduced to the product and, before consumers are made aware of the product and its advantages, sales are minimal [37].

During the introduction stage, the production volume is typically lower, and economies of scale have not yet been achieved. As a result, the per-unit production cost is higher, which could lead to a higher average price for the product compared to the mature stage when production costs have decreased due to larger production volumes.

Companies may initially set a higher price for their products during the introduction stage to capitalize on early adopters who are willing to pay a premium for innovative products. This strategy allows companies to maximize their profit margins before the market becomes more competitive during the mature stage, at which point they may need to lower prices to remain competitive and attract a broader customer base.

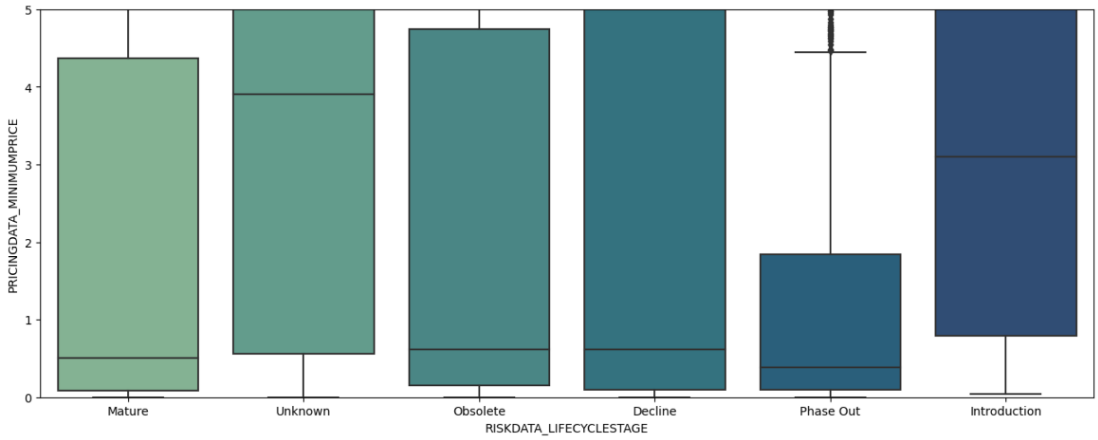


Figure 4.8 - Boxplot showing the distribution of Price across Lifecycle Stage

In Figure 4.9, regarding RoHS risk, it is observed that over 200000 parts are classified as low risk, indicating minimal potential harm to human health and the environment due to electrical waste. A possible explanation for the trend between RoHS risk and pricing could be that components with medium risk have higher average prices, potentially reflecting higher manufacturing costs or regulatory compliance requirements. However, it is important to note that high-risk components exhibit the lowest average prices. This observation may be due to the presence of other factors influencing the pricing of these components, such as variations in quality and technology [38].

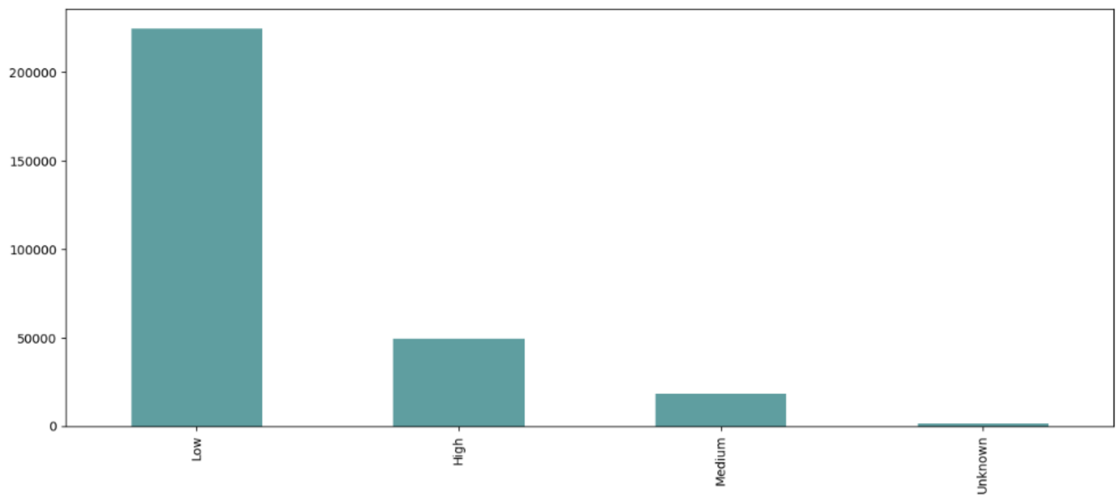


Figure 4.9 - EC count by RoHS Risk

Checking with price, the parts considered as medium risk have the higher medium price around 20. Could be that more risk more expensive, but in Figure 4.10, is possible to see that the high risk has the lowest medium price.

Medium risk could be priced higher due to increased manufacturing costs, additional regulatory compliance requirements, or greater demand for these specific components.

On the other hand, high-risk components may be priced lower, potentially because they represent older or outdated technologies, or due to factors such as market dynamics, competition, and variations in component quality.

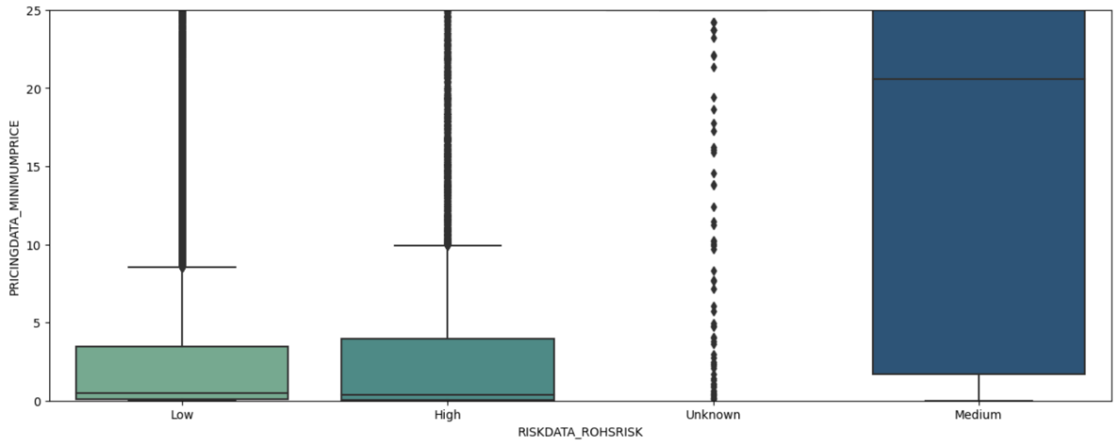


Figure 4.10 - Boxplot showing the distribution of Price across RoHS Risk

For Inventory risk, in Figure 4.11, around 140000 electronic parts are in low risk. This risk is based on the availability of the parts in the authorized distributors inventories, which means that these parts have low risk of not being in stock.

However around 120000 parts are on the high risk of no availability.

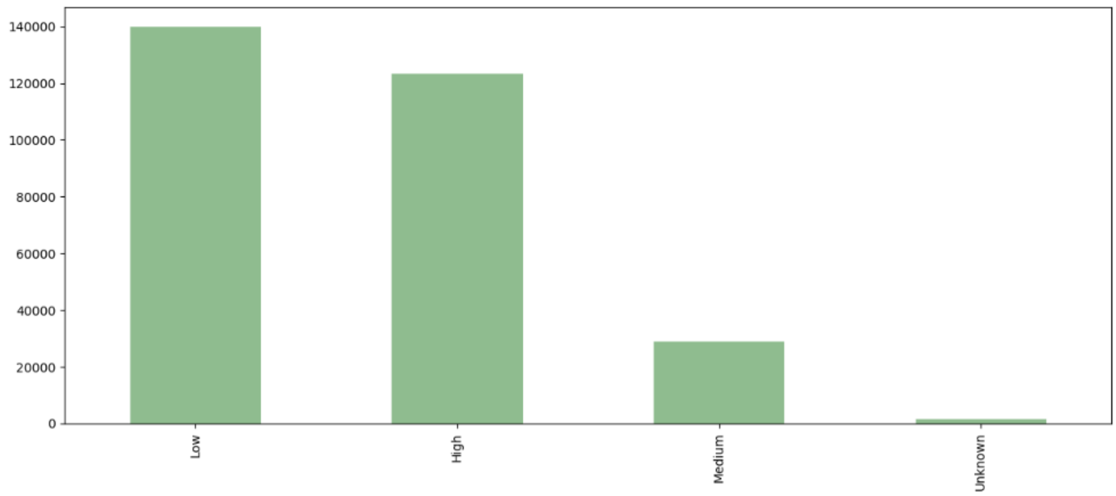


Figure 4.11 - EC count by Inventory Risk

In Figure 4.12, is possible to verify that the parts that have the probability high and medium of no availability have the higher price around 1 compared to low risk that has the medium price of around 0.25. Is possible to conclude that when rarer the part, more expensive it is.

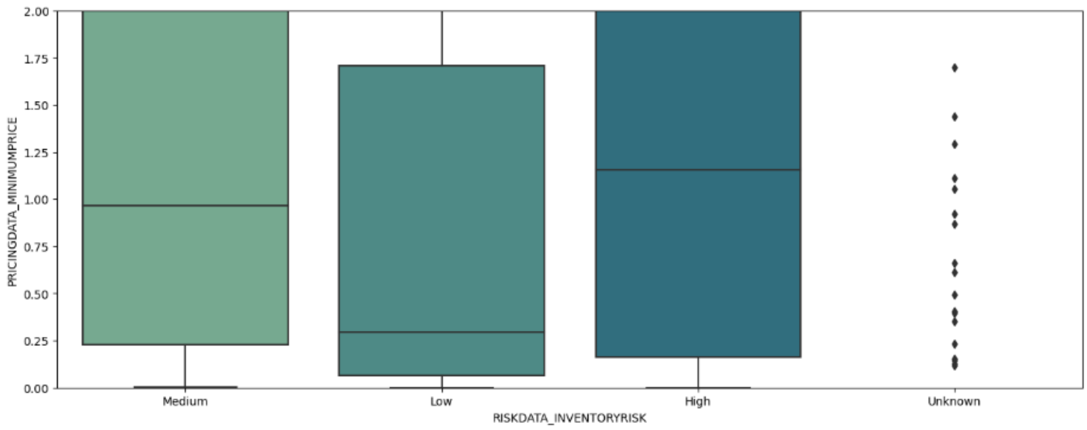


Figure 4.12 - Boxplot showing the distribution of Price across Inventory Risk

Regarding Multisource risk, this is related on the availability of cross-functional resources. In Figure 4.13, around 130000 parts are on Low risk what indicates that each part is sold only by one Manufacturer and 120000 on High risk.

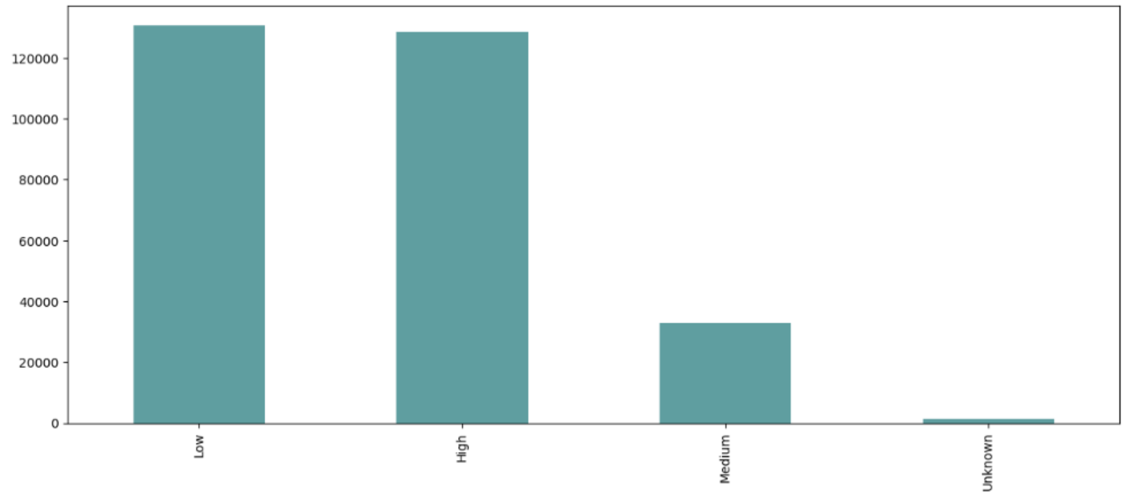


Figure 4.13 - EC count by Multisource Risk

In Figure 4.14, the High risk shows a higher price. This implies that the absence of these resources could have a significant negative effect, while low risk predicts a minor effect and price.

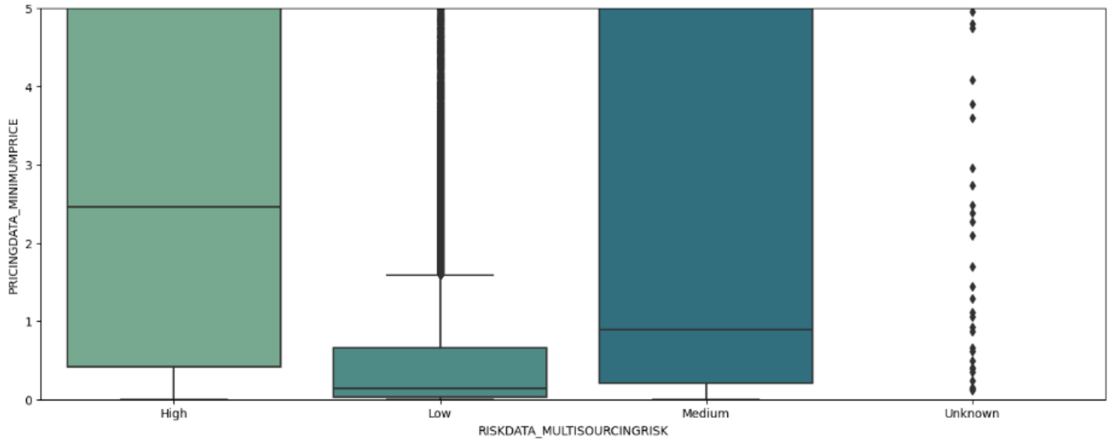


Figure 4.14 - Boxplot showing the distribution of Price across Multisource Risk

For the SE Grade, this represents on a scale from 1 to 5 the probability of the part to be counterfeit. The grades 2 and 5 have between 60000 and 70000 parts each, which is concerning to have so many to the higher scale, possible to check in Figure 4.15.

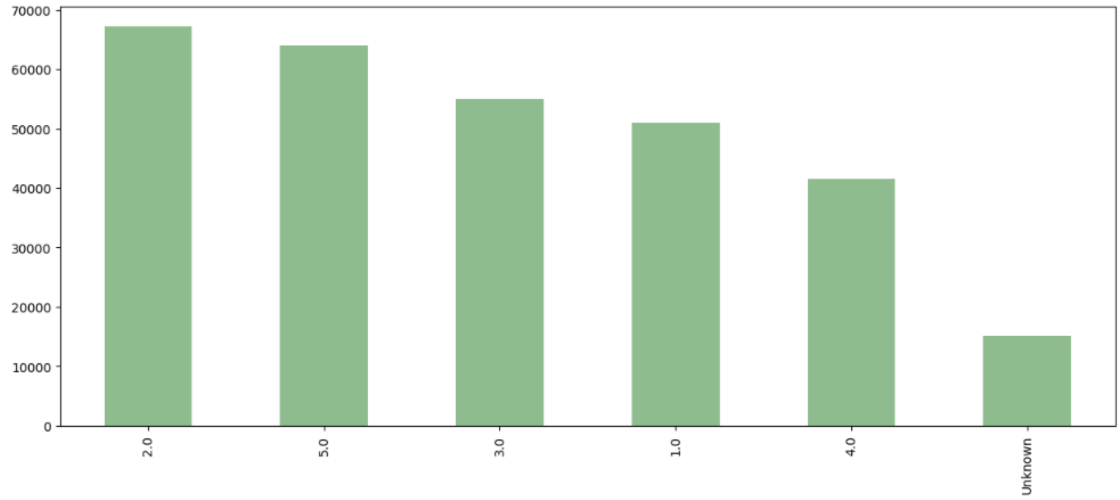


Figure 4.15 - EC count by SE Grade

When analysing with the price, is possible to conclude that the scale 5 have the minor medium price than the other grades, so easier to be counterfeit. Scales 2, 3 and 4 have almost the same medium price and scale 1 has the higher medium price, so concluding to be more difficult to be counterfeit. Figure 4.16 represents that.

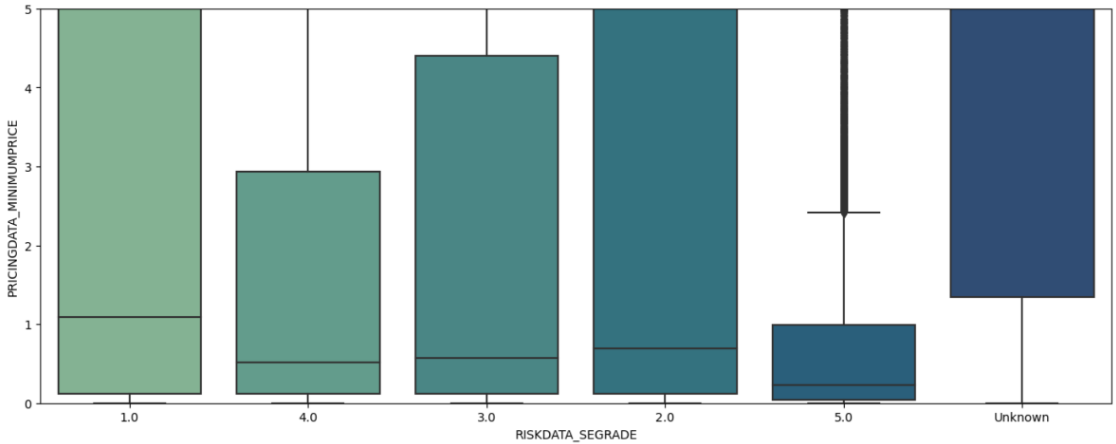


Figure 4.16 - Boxplot showing the distribution of Price across SE Grade

In the real world, product life cycle data can be used by companies to customize their marketing and pricing plans. These can benefit from early adopters by introducing more innovative products and charging a higher initial price during the debut phase and should concentrate on cost optimization and pricing competitiveness when they reach the mature stage.

Businesses must evaluate how their products affect the environment and think about using eco-friendly materials and production techniques, which could lower production costs. Reducing the RoHS risk classification of their constituents can confer a competitive edge in markets where customers prioritize environmental concerns.

In order to mitigate the danger of either overstocking or understocking components, businesses must effectively manage their inventory. In order to minimize supply chain interruptions, parts that are categorized as high-risk for availability should be regularly tracked and, if at all feasible, alternate sources or suppliers should be found.

A high multisource risk may indicate possible weaknesses in the supply chain. In particular for components with significant multisource risk, companies should evaluate the availability of alternative sources or develop backup plans to guarantee a consistent supply of components.

The SE Grade data is needed to determine how likely it is that fake parts will find their way into their supply chain. To stop fake parts from getting into their products, businesses should have strict quality control and verification procedures in place, especially for components with higher SE Grade numbers.

Businesses should take risk considerations into account when setting their prices. Components with increased inventory, multisource, or RoHS risks, for instance, could be worth a premium to cover potential issues with sourcing, production, and environmental compliance.

Perceived danger of counterfeiting can also affect pricing (SE Grade). Higher SE Grade component numbers may come with a higher price tag to reflect the additional assurance of authenticity.

In order to lower risk, companies can think about expanding the range of products they offer by adding additional components or buying from other vendors.

A company's brand reputation and competitiveness can be improved by investing in research and development to create safer and more environmentally friendly components.

4.3 Statistical Analysis - Risk on Price

When analyzing the statistics and metrics of the regression model, in Table 4.4, it is possible to verify that R-squared is 0.335, which shows that the included independent variables can account for around 33.5% of the variation in the price.

The adjusted R-squared in this instance is similarly 0.335, indicating that the model explains the same amount of variance in the dependent variable as the normal R-squared.

An overall statistically significant fit of the model is shown by a high F-statistic (6597) and a low probability value (Prob (F-statistic): 0.00).

Finally, lower AIC and BIC values denote better model fit. Because there are many observations in the dataset, the AIC and BIC values in this instance are relatively high.

Table 4.4 - Statistics and Metrics

Statistics and Metrics	Value
R-squared	0.335
Adj. R-squared	0.335
F-statistic	6597.
Prob (F-statistic)	0
AIC	9.907E+05
BIC	9.909E+05

Analyzing the regression results, from Table 4.5 is possible to verify that the coefficient for RISKDATA_SEGRADE_2.0 is around -0.12. This indicates that the forecasted price is anticipated to reduce by 0.12 units when the SEGRADE risk category is 2.0. For RISKDATA_SEGRADE_3.0 the coefficient is around -0.21. This indicates that the forecast price is anticipated to decline by 0.21 units.

RISKDATA_SEGRADE_4.0 indicates that the forecast price is anticipated to fall by 0.25 units.

Finally, RISKDATA_SEGRADE_5.0: This indicates that the estimated price is anticipated to decline by 0.83 units.

These coefficients are statistically significant since the p-values attached to them are all extremely modest (around zero). This indicates that the values of the risk categories significantly affect the regression model's anticipated price.

Overall, these coefficients indicate that lower anticipated prices are related with higher SE Grade risk categories (3.0, 4.0, and 5.0) when compared to the reference group (SE Grade risk category 1.0). Understanding how various risk levels affect the Electronic Components cost and making decisions based on these risk categories may both benefit from knowing this.

The other variables don't present a significant impact on price since the p-values are relatively high.

Table 4.5 - Regression Results

Features	Coefficient	Std. Error	t-value	p-value
const	-1.889E+11	1.75E+11	-1.08	0.28
RISKDATA_SEGRADE_2.0	-0.12	0.016	-7.79	0
RISKDATA_SEGRADE_3.0	-0.21	0.019	-11.05	0
RISKDATA_SEGRADE_4.0	-0.25	0.019	-12.61	0
RISKDATA_SEGRADE_5.0	-0.83	0.023	-36.69	0
RISKDATA_MULTISOURCINGRISK_2	5.075E+10	7.04E+10	0.72	0.47
RISKDATA_MULTISOURCINGRISK_3	1.015E+11	1.41E+11	0.72	0.47
RISKDATA_INVENTORYRISK_2	2.855E+10	5.91E+10	0.48	0.63
RISKDATA_INVENTORYRISK_3	5.71E+10	1.18E+11	0.48	0.63
RISKDATA_ROHSRISK_2	2.674E+11	4.05E+11	0.66	0.51
RISKDATA_ROHSRISK_3	5.347E+11	8.1E+11	0.66	0.51

RISKDATA_LIFECYCLERISK_2	2.119E+10	6.83E+10	0.31	0.76
RISKDATA_LIFECYCLERISK_3	4.238E+10	1.37E+11	0.31	0.76

Overall, in Table 4.6, the VIF value shows that the Multisourcing risk, Inventory risk, and RoHS risk features have severe multicollinearity.

This may be an issue for the regression model since it could result in incorrect coefficient estimates and make it challenging to identify the precise influence of each risk category on the price data.

Table 4.6 - VIF Values

Features	VIF
const	0
RISKDATA_SEGRADE_2.0	1.838761
RISKDATA_SEGRADE_3.0	1.941636
RISKDATA_SEGRADE_4.0	1.901441
RISKDATA_SEGRADE_5.0	2.685822
RISKDATA_MULTISOURCINGRISK_2	3.51E+10
RISKDATA_MULTISOURCINGRISK_3	3737.355
RISKDATA_INVENTORYRISK_2	3.98E+09
RISKDATA_INVENTORYRISK_3	621.7886
RISKDATA_ROHSRISK_2	1.88E+10
RISKDATA_ROHSRISK_3	310215.2
RISKDATA_LIFECYCLERISK_2	1437.569
RISKDATA_LIFECYCLERISK_3	16240.43

It is important to rerun the regression model and limit its scope to just the variables that show a detectable impact on price after identifying and assessing the extent to which specific risks fail to exert a significant influence on price.

A R-squared of 0.317 suggests that the independent variables in the model can account for about 31.7% of the price variance.

The model is statistically significant, as shown by the F-statistic (1.032e+04) and the p-value (Prob (F-statistic)), which is normally used to assess if the F-statistic is significant, must also be taken into account while analyzing this statistic. A p-value of 0 means statistical significance for the model.

In table 4.7 is possible to check the statistics and metrics mentioned before when rerun of the regression model.

Table 4.7 - Statistics and Metrics

Statistics and Metrics	Value
R-squared	0.317
Adj. R-squared	0.317
F-statistic	1.032e+04.
Prob (F-statistic)	0
AIC	9.967E+05
BIC	9.967E+05

Table 4.8 shows how various risk grades affect the pricing in relation to the price category. The negative coefficients show that lower costs are related with higher risk grades (from 2.0 to 5.0). These coefficients are all statistically significant, indicating that the price is significantly influenced by these risk levels.

The price tends to decrease as the risk grade increases, and this association is statistically significant. A recurrent finding in financial and economic models is that lower asset prices or returns are frequently the outcome of higher perceived risk.

Rerunning the model with only the significant variables is possible to achieve more trustable results.

Table 4.8 - Regression Results

Features	Coefficient	Std. Error	t-value	p-value
const	-0.13	0.012	-10.40	0
RISKDATA_SEGRADE_2.0	-0.10	0.015	-13.15	0
RISKDATA_SEGRADE_3.0	-0.22	0.017	-16.98	0
RISKDATA_SEGRADE_4.0	-0.31	0.019	-12.61	0
RISKDATA_SEGRADE_5.0	-0.93	0.018	-50.40	0

The SE Grade variable show critical information for analyzing the prices of electronic components. Therefore, by including this variable in the model, the conclusions and usefulness for the companies of the statistical analysis can impact in several ways.

For pricing relevance, the probability of a component being counterfeit is important for the company, as counterfeit products can be of inferior quality and pose a risk to customers. Therefore, by including this variable in the model, the company can make pricing decisions that are more in line with the real risk associated with the components and the probability of counterfeiting is having a measurable impact on the prices of electronic components. This is fundamental to understand how risk affects the company's economy.

Based on the coefficients, the company can make more informed decisions about pricing strategies, negotiations with suppliers, and even the acquisition of components. For example, the company may be willing to pay more for components with a lower probability of being counterfeited, or it may implement additional quality assurance measures for components with a high probability of counterfeiting.

CHAPTER 5

Conclusions

5.1 Conclusions

Price is a key consideration when choosing electronic components (ECs) in the market of today. Finding suppliers who can provide components with the necessary quality level and give reasonable costs is a top priority for manufacturers. Cost effectiveness is a competitive advantage that enables companies to draw clients by offering premium goods at competitive prices.

Various elements, including as market trends and the volatility of common components, have an impact on EC pricing. Manufacturers examine relationships between the prices of various products to increase the precision of pricing predictions and customer value perception. Optimizing product pricing and market positioning involves understanding price trends and consumer preferences.

Manufacturers can achieve cost effectiveness and keep a competitive edge in the market by managing pricing strategies with care, taking risk into account, and adopting sustainable practices.

The examination of the electronic component dataset has yielded important insights into the variables affecting the availability and cost of these components, which can have a substantial effect on businesses operating in the electronics sector. There are some important details about the effects on businesses and what they can do to make wise decisions.

The information showed that the price points of various categories of electronic components vary. Businesses that specialize in expensive product categories, such chips and relays, might need to plan their manufacturing and marketing appropriately. To defend the increased costs, they can focus on specialized markets or highlight the features and functionality of their products.

The discovery that Infineon Technologies AG's electronic components are offered by the expensive company GSI Technologies emphasizes the significance of vendor comparison. To guarantee they are receiving the best value for their electrical components, businesses should continually assess and contrast their suppliers.

Some locations, such as West North America and Middle Europe, are high-volume producers of electronic parts, according to the study of geolocation data. Companies based in these areas may profit from potential cost savings and easy access to suppliers. They should, however, consider the higher median prices in areas like Sweden and France when determining their cost and pricing structures.

Companies should carefully plan their new releases, as evidenced by the finding that medium costs for electrical components are greater during the introduction stage than the mature stage. During the debut phase, they can spend money on marketing and promotion to generate early demand and possibly support higher prices.

Businesses should exercise caution when it comes to the distribution of SE grades, particularly grades 2 and 5, which show a high percentage of components. This suggests that there is a greater chance of fake parts. Strict quality control and verification procedures should be put in place by manufacturers to guarantee the legitimacy of the components they source.

Regression analysis is a useful tool for businesses to comprehend the connection between component costs and risk levels. The necessity of resolving this issue is highlighted by the discovery of multicollinearity among risk variables. Businesses can reduce multicollinearity by changing the way they analyze data or by looking at different ways to determine risk and price.

To sum up, the examination of the dataset provides significant perspectives for businesses involved in the electronics sector. Their decision-making on supplier selection, price tactics, geographic considerations, product life cycle planning, risk assessment, and quality control can be aided by these insights. Businesses can lower risks, increase supply chain efficiency, and boost their competitiveness in the electrical component industry by incorporating these insights.

5.2 Future Work

As concluded before, the regression model presented that several risk variables have multicollinearity issue.

For future work on this subject, I would propose a deepen analysis to identify multicollinearity causes among the risk variables. Meaning, focus on employing advanced statistical techniques, such as principal component analysis or ridge regression, to reduce multicollinearity and improve the accuracy of price prediction models.

Also, as the SE Grade was the risk variable that showed more impact on the price, according to the regression model, I propose further investigation on counterfeit detection methods, including advanced authentication techniques and blockchain-based solutions.

Detecting and preventing counterfeit components can help maintain product integrity and customer trust.

References

[1] C. Schröer, F. Kruse, and J. M. Gómez, "A systematic literature review on applying CRISP-DM process model," in *Procedia Computer Science,* 2021. doi: 10.1016/j.procs.2021.01.199.

[2] P. Chapman, Julian Clinton, R. Kerber, T. Khabaza, T. Reinartz, C. Daimlerchrysler, Rüdiger Shearer, Wirth., "Step-by-step data mining guide," *SPSS inc,* vol. 78, 2000.

[3] F.-E. Hamdani, I. A. Q. Quintero, M. Enjolras, M. Camargo, D. Monticolo, and C. Lelong, "Agile supply chain analytic approach: a case study combining agile and CRISP-DM in an end-to-end supply chain," *Supply Chain Forum: An International Journal,* pp. 1–15, Apr. 2022, doi: 10.1080/16258312.2022.2064721.

[4] H. Liu, G. Li, and Y. Zhao, "A Sliding-Window Principal Component thermography reconstruction approach for enhancement and identification of electronic components internal structure," *Measurement (Lond),* vol. 184, 2021, doi: 10.1016/j.measurement.2021.109926.

[5] T. Zink, F. Maker, R. Geyer, R. Amirtharajah, and V. Akella, "Comparative life cycle assessment of smartphone reuse: Repurposing vs. refurbishment," *International Journal of Life Cycle Assessment,* vol. 19, no. 5, 2014, doi: 10.1007/s11367-014-0720-7.

[6] Z. Sharifi and S. Shokouhyar, "Promoting consumer's attitude toward refurbished mobile phones: A social media analytics approach," *Resour Conserv Recycl,* vol. 167, 2021, doi: 10.1016/j.resconrec.2021.105398.

[7] P. Cui, J. Dixon, U. Guin, and D. Dimase, "A Blockchain-Based Framework for Supply Chain Provenance," *IEEE Access,* vol. 7, 2019, doi: 10.1109/ACCESS.2019.2949951.

[8] S. Kadam, "Information Resources about Electronic Components," *International Journal of Engineering Research & Technology,* vol. 3, no. 8, 2014.

[9] G. Lucchese, W. Ketter, J. Van Dalen, and J. Collins, "Forecasting prices in dynamic heterogeneous product markets using multivariate prediction methods,"

in *ACM International Conference Proceeding Series*, 2011. doi: 10.1145/2378104.2378130.

[10] K. L. Chen, K. S. Chen, and R. K. Li, "Suppliers capability and price analysis chart," *Int J Prod Econ*, vol. 98, no. 3, 2005, doi: 10.1016/j.ijpe.2004.09.010.

[11] C. C. Hsieh and H. H. Lai, "Pricing and ordering decisions in a supply chain with downward substitution and imperfect process yield," *Omega (United Kingdom)*, vol. 95, 2020, doi: 10.1016/j.omega.2019.04.007.

[12] K. Yasuhiro and A. Keiji, "A simulation model of the pricing process for general-purpose electronic parts-The price-sales distribution observed from simulation-," *Journal of Japan Industrial Management Association*, vol. 60, no. 5, pp. 270–277, 2009.

[13] X. F. Shao and J. H. Ji, "Effects of sourcing structure on performance in a multiple-product assemble-to-order supply chain," *Eur J Oper Res*, vol. 192, no. 3, 2009, doi: 10.1016/j.ejor.2007.10.014.

[14] C. C. Lin and Y. C. Wu, "Optimal pricing for build-to-order supply chain design under price-dependent stochastic demand," *Transportation Research Part B: Methodological*, vol. 56, 2013, doi: 10.1016/j.trb.2013.07.011.

[15] P. Pasimeni, "Supply or Demand, that is the Question: Decomposing Euro Area Inflation," *Intereconomics*, vol. 57, no. 6, pp. 384–393, 2022, doi: 10.1007/s10272-022-1092-z.

[16] H. Allison, P. Sandborn, and B. Eriksson, "On the Applicability of Analytical Supply Chain Disruption Models for Selecting the Optimum Contingency Strategies for Electronic Supply Chain Disruption Management: A Comparison With Simulation," in *Volume 4: 19th Design for Manufacturing and the Life Cycle Conference; 8th International Conference on Micro- and Nanosystems*, American Society of Mechanical Engineers, Aug. 2014. doi: 10.1115/DETC2014-34512.

[17] D. M. Lambert and M. C. Cooper, "Issues in supply chain management," *Industrial Marketing Management*, vol. 29, no. 1, 2000, doi: 10.1016/S0019-8501(99)00113-3.

[18] K. L. Croxton, S. J. García-Dastugue, D. M. Lambert, and D. S. Rogers, “The Supply Chain Management Processes,” *The International Journal of Logistics Management*, vol. 12, no. 2, 2001, doi: 10.1108/09574090110806271.

[19] S. M. Wagner and C. Bode, “An empirical examination of supply chain performance along several dimensions of risk,” *Journal of Business Logistics*, vol. 29, no. 1, 2008, doi: 10.1002/j.2158-1592.2008.tb00081.x.

[20] J. Hou, A. Z. Zeng, and L. Sun, “Backup sourcing with capacity reservation under uncertain disruption risk and minimum order quantity,” *Comput Ind Eng*, vol. 103, 2017, doi: 10.1016/j.cie.2016.11.011.

[21] R. Alkhudary, M. M. Queiroz, and P. Féniès, “Mitigating the risk of specific supply chain disruptions through blockchain technology,” *Supply Chain Forum*, 2022, doi: 10.1080/16258312.2022.2090273.

[22] E. T. Gilmore, P. D. Frazier, I. J. Collins II, W. Reid, and M. F. Chouikha, “Infrared analysis for counterfeit electronic parts detection and supply chain validation,” *Environ Syst Decis*, vol. 33, no. 4, pp. 477–485, 2013, doi: 10.1007/s10669-013-9482-1.

[23] J. Stradley and D. Karraker, “The electronic part supply chain and risks of counterfeit parts in defense applications,” *IEEE Transactions on Components and Packaging Technologies*, vol. 29, no. 3, pp. 703–705, 2006, doi: 10.1109/TCAPT.2006.882451.

[24] A. Shrivastava and M. Pecht, “Counterfeit capacitors in the supply chain,” *Journal of Materials Science: Materials in Electronics*, vol. 25, no. 2, 2014, doi: 10.1007/s10854-013-1147-9.

[25] B. Bartels, U. Ermel, M. Pecht, and P. Sandborn, *Strategies to the Prediction, Mitigation and Management of Product Obsolescence*. 2012. doi: 10.1002/9781118275474.

[26] D. Li and N. Mishra, “The impact of parts obsolescence on contracts for durable goods with after-sales service,” *Int J Prod Res*, vol. 60, no. 16, 2022, doi: 10.1080/00207543.2021.1950934.

[27] K. Feldman and P. Sandborn, “Integrating technology obsolescence considerations into product design planning,” in *2007 Proceedings of the ASME International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, DETC2007*, 2008. doi: 10.1115/DETC2007-35881.

[28] V. J. Prabhakar, H. Allison, P. Sandborn, and B. Eriksson, “Optimizing part sourcing strategies for low-volume, long life cycle products using second sourcing and part hoarding,” in *Proceedings of the ASME Design Engineering Technical Conference*, 2013. doi: 10.1115/DETC2013-12464.

[29] K. Klein and K.-D. Thoben, “Industry 4.0 Technologies as an Obsolescence Mitigator for Testing of Mechatronic Systems in Aviation,” *Applied Sciences*, vol. 12, no. 21, p. 11142, Nov. 2022, doi: 10.3390/app122111142.

[30] J. Oliveira, M. Nunes, and P. Afonso, “New product development in the context of industry 4.0: Insights from the automotive components industry,” in *Springer Proceedings in Mathematics and Statistics*, 2019, pp. 83–94. doi: 10.1007/978-3-030-14973-4_8.

[31] Murad Salim Attiany, Sami Awwad Al-kharabsheh, lafie Saleh Al-Makhariz, Mohd Ahmad Abed-Qader, Sulieman Ibraheem Al-Hawary, Anber Abraheem Mohammad, Adeeb Ahmed AL Rahamneh., “Barriers to adopt industry 4.0 in supply chains using interpretive structural modeling,” *Uncertain Supply Chain Management*, vol. 11, no. 1, pp. 299–306, 2023, doi: 10.5267/j.uscm.2022.9.013.

[32] V. Koch, S. Kuge, R. Geissbauer, and S. Schrauf, “Industry 4.0 - Opportunities and challenges of the industrial internet,” *strategy& Formerly Booz & Company, PwC*, vol. 13, 2014.

[33] M. A. Villegas and D. J. Pedregal, “Automatic selection of unobserved components models for supply chain forecasting,” *Int J Forecast*, vol. 35, no. 1, 2019, doi: 10.1016/j.ijforecast.2017.11.001.

[34] R. Wirth, “CRISP-DM : Towards a Standard Process Model for Data Mining,” *Proceedings of the Fourth International Conference on the Practical Application of Knowledge Discovery and Data Mining*, no. 24959, 2000.

[35] "SiliconExpert - About," *https://www.siliconexpert.com/about/*. (accessed April 15, 2023)

[36] "SiliconExpert - Part Detail Outputs," *https://www.siliconexpert.com/apidocs/#part-detail-outputs*. (accessed April 15, 2023)

[37] Darya Rudenko and Arkady Borisov, "Blackboard architecture for product life cycle stage definition".

[38] https://environment.ec.europa.eu/topics/waste-and-recycling/rohs-directive_en. (accessed May 24, 2023)