

CZECH UNIVERSITY OF LIFE SCIENCES PRAGUE

FACULTY OF ECONOMICS AND MANAGEMENT

DEPARTMENT OF INFORMATION TECHNOLOGIES



DIPLOMA THESIS

**Sales Analysis of E-Commerce Websites Using Data Mining Techniques & Power
BI Tool for Reporting**

Author: B.Sc. Nooshin Assadi Karam

Supervisor: Merunka Vojtěch, Doc. Ing., Ph.D.

© 2020 CULS Prague

CZECH UNIVERSITY OF LIFE SCIENCES PRAGUE

Faculty of Economics and Management

DIPLOMA THESIS ASSIGNMENT

B.Sc. NOOSHIN ASSADI KARAM, BSc

Systems Engineering and Informatics
Informatics

Thesis title

Sale analysis of e-commerce websites using data mining techniques and Power BI tool for reporting

Objectives of thesis

The goal of this thesis is to discover data patterns, increase their core value and use data to gain knowledge for the concrete international company. There are several areas in which the volume of data is stored in centralized or distributed databases. Data mining effects in marketing and wholesale will be summarized as follows: Customer identification, Purchase rate, Purchase volume, Identification and differentiation of customer groups, Calculate customer lifetime, Looking ahead, Review the success or failure of marketing plans, and Sales forecast.

Methodology

Methodology of the thesis will be based on study and specialized sale and marketing in a concrete international company. This is the first part of the thesis. In the second part, there will be practical analysis and documentation of data mining in Sportissimo Co. using standard business-process modelling tools based on the UML or BPMN standards.

The proposed extent of the thesis

80 – 120 pages

Keywords

Data mining; E-Commerce; Web mining; sale analysis; data mining techniques, Power BI

Recommended information sources

JAMES, G. *An introduction to statistical learning : with applications in R*. ISBN 978-1461471370.

REYNOLDS, Janice. *The complete e-commerce book: design, build & maintain a successful Web-based business*. 2nd ed. Berkeley, Calif.: Distributed to the Book trade in the U.S. by Publishers Group West, c2004. ISBN 9781578203123.

SMYTH, P. – HAND, D.J. – MANNILA, H. *Principles of data mining*. Cambridge, Mass.: MIT Press, 2001. ISBN 026208290.

Expected date of thesis defence

2019/20 SS – FEM

The Diploma Thesis Supervisor

doc. Ing. Vojtěch Merunka, Ph.D.

Supervising department

Department of Information Engineering

Electronic approval: 5. 4. 2020

Ing. Martin Pelikán, Ph.D.

Head of department

Electronic approval: 5. 4. 2020

Ing. Martin Pelikán, Ph.D.

Dean

Prague on 06. 04. 2020

Declaration

I declare that I have worked on my diploma thesis titled "Sales analysis of e-commerce websites using data mining techniques and Power BI tool for reporting" by myself and I have used only the sources mentioned at the end of the thesis. As the author of the diploma thesis,

I declare that the thesis does not break the copyrights of any third person.

In Prague on April of 2020

NOOSHIN ASSADI KARAM

Acknowledgment

First, I would like to acknowledge the enormous contribution of my lecturer and supervisor Doc. Ing. Vojtěch Merunka, Ph.D. for guidance, useful comments, remarks and engagement through the learning process and preparation of this work.

Second, I would like to thank Mr. Petr Kraus, IT Director of SPORTISIMO CO., for his advice, help, and encouragement. And, this work could not have been completed without the support, help, and knowledge which I received from Mr. Michal Procházka (Analytic E-commerce of SPORTISIMO CO.). I would like to thank the following people, Mr. Martin Bambas (IT Delivery Manager of SPORTISIMO CO.), Mr. Petr Jiříček (ERP Manager of SPORTISIMO CO.), Mr. Aleš Ulrych (BI&DWH consulting of SPORTISIMO CO.) for their help.

And wish to acknowledge the support and great love of my family, my father, Mohammad; my mother, Sima; my sisters, Nassibeh and Nazanin; my brother, Ali; and my little angel, Elsa. Special thanks to Mr. Arash Goudarzi for his advice and kindness in my life. I am deeply indebted to Mr. Vincent Clapdorp for his support, encouragement, and patience in this difficult journey. They kept me going on and this work would not have been possible without their input.

Sales Analysis of E-Commerce Websites Using Data Mining Techniques and Power Bi Tool for Reporting

Abstract

The use of modern scientific services in everyday life can eliminate the inevitable needs. Watching videos and doing things through reviewing can be restricted from being reorganized on your internet network and able to share your information with others. With the development of an e-commerce store, you can increasingly play an important role in the life of an actor. Using the data mining technique, the creator of these sites can be preferred to you and allow you to shop online through custom product sites, management, and analysis. Data mining tools help you deliver services tailored to your needs.

Business intelligence products have become a necessity for organizations, offices, and corporations. In recent years, this type of product has been moving towards simplifying the workflow and procedures and improving the user experience of its users. One of the most important applications that can help in analyzing data in Power BI. Power BI is a set of services that make it easy for business experts to analyze business information.

Keywords

Data mining, E-Commerce, Web mining, Sales analysis, Data mining techniques, Power BI

Abstrakt

Využití moderních vědeckých služeb v každodenním životě může eliminovat nezbytné potřeby. Sledování videí a provádění věcí prostřednictvím opakovaného prohlížení může být omezeno na reorganizaci ve vaší internetové síti a na sdílení informací s ostatními. S rozvojem elektronického obchodu můžete hrát stále důležitější roli v životě herce. Pomocí technologie data mining vám může být tvůrce těchto stránek upřednostněn a umožní vám nakupovat online prostřednictvím vlastních produktových webů, správy a analýzy. Nástroje pro data mining vám pomohou poskytovat služby šité na míru vašim potřebám.

Produkty business intelligence se staly nezbytností pro organizace, kanceláře a korporace. V posledních letech tento typ produktu směřuje ke zjednodušení pracovního postupu a postupů a ke zlepšení uživatelského dojmu jeho uživatelů. Jednou z nejdůležitějších aplikací, která může pomoci při analýze dat, je Power BI. Power BI je sada služeb, které obchodním odborníkům usnadňují analýzu obchodních informací.

Klíčová Slova

Data mining, E-Commerce, Web mining, Analýza prodeje, Techniky data mining, Power BI

Table of Contents

1.	Introduction.....	14
1.1	The necessity & importance of customer relationship management.....	14
1.2	Goals of customer relationship management in e-commerce.....	15
1.3	Purpose of the thesis.....	17
2.	Literature Review	18
2.1	Definition of e-commerce	19
2.2	Businesses that sell directly to consumers (B2C).....	20
2.2.1	B2C Marketing.....	20
2.2.2	Businesses that sell directly to businesses (B2B)	20
2.2.3	B2B Marketing.....	21
2.3	E-commerce & sales	21
2.3.1	Gain proper insight.....	22
2.4	Understanding customer lifetime value.....	27
2.5	Online shopping	28
2.5.1	Compare prices.....	29
2.5.2	Unlimited selection:	30
2.5.3	Access to Consumer Reviews:	30
2.5.4	Get access to your company services within 2 hours	30
2.5.5	Analyze the accuracy of visitors:	30
2.6	Data	30
2.7	Information	31
2.8	Data Mining.....	31
2.9	Data mining algorithms.....	34
2.9.1	Clustering.....	34
2.9.2	Association Rules Algorithm.....	35
2.9.3	Classification of data	36
2.9.4	Regression and Time Series Analysis.....	37
2.9.5	Neural Networks.....	37

2.10	Business Intelligence	38
3.	Practical Part	39
3.1	Craft.CASE.....	40
3.1.1	Why choose Craft. Case:.....	40
3.2	Business Diagram (Online shopping).....	41
3.2.1	Order products	41
3.2.2	Login system.....	42
3.2.3	Ordering products	43
3.2.4	Check and prepare the products.....	44
3.2.5	Pay for products	45
3.2.6	Send the products to the customers.....	46
3.2.7	Pay by credit card	47
3.3	Draw.io	48
3.3.1	Powerful & Easy to Use	48
3.4	UML (Online shopping)	49
3.4.1	CLASS MODEL	50
3.4.2	Dynamic Model	52
3.5	Interaction Model.....	54
3.5.1	Use Case Diagram.....	54
3.5.2	Sequence Diagram.....	58
3.6	Activity Diagram	60
3.7	SPORTISIMO	64
3.8	Keboola.....	65
3.8.1	Keboola is a secure data innovation platform for your data needs.....	65
3.8.2	All the tools you need in one place	65
3.8.3	Collaboration breeds success	65
3.9	Data preparation in Keboola for the SPORTISIMO company	66
3.9.1	Daily Traffic.....	67
3.9.2	Daily Transactions	73
3.9.3	Structure of Traffic, Transactions, Products and Customers tables of SPORTISIMO	76

3.10	RStudio	77
3.11	Sales forecast of 2016 in RStudio	78
3.11.1	Load Data.....	78
3.11.2	Convert Date From List to Date type.....	79
3.11.3	The graph shows sales in 2016.....	82
3.11.4	Final Script in RStudio:.....	84
3.11.5	Forecast result by RStudio.....	85
3.11.6	Comparison of the ARIMA model and Time series	86
3.12	Power BI	87
3.12.1	Why use Power BI?.....	87
3.12.2	The workflow in Power BI	88
3.12.3	Create a report for sales in 2016 for the SPORTISIMO company:.....	88
4.	Conclusion	89
5.	References.....	91

Table of Figures

Figure 1 Artificial intelligence in e-commerce.....	27
Figure 2 Craft.CASE.....	40
Figure 3 Business diagram - Ordering products in online shopping	41
Figure 4 Business diagram - Login customer in system.....	42
Figure 5 Business diagram - Ordering products by customers in system	43
Figure 6 Business diagram - Check and prepare the products in system.....	44
Figure 7 Business diagram – Pay for products in the system.....	45
Figure 8 Business diagram - Send the products to the customers.....	46
Figure 9 Business diagram - Pay by credit card	47
Figure 10 draw.io.....	48
Figure 11 Class diagram for online shopping system.....	51
Figure 12 STATE DIAGRAM for online shopping	53
Figure 13 USE CASE DIAGRAM - All actors in the online shopping system	54
Figure 14 USE CASE DIAGRAM - Customers	55
Figure 15 USE CASE DIAGRAM - Manager.....	56
Figure 16 USE CASE DIAGRAM - Employee	57
Figure 17 USE CASE DIAGRAM - Transformer	58
Figure 18 SEQUENCE DIAGRAM - Login.....	59
Figure 19 SEQUENCE DIAGRAM - Ordering products and save in database.....	60
Figure 20 ACTIVITY DIAGRAM - Ordering Products.....	62
Figure 21 ACTIVITY DIAGRAM - Login in system	63
Figure 22 SPORTISIMO	64
Figure 23 Data preparation in Keboola	66
Figure 24 Data preparation in Keboola - Transactions.....	66
Figure 25 Data preparation in Keboola - Daily Traffic.....	67
Figure 26 Data preparation in Keboola - Final query	68
Figure 27 Data preparation in Keboola - Create a sandbox for writing query	69
Figure 28 Data preparation in Keboola - Login into Snowflake	69
Figure 29 Data preparation in Keboola - Offline daily traffic's table	70

Figure 30 Data preparation in Keboola - Online daily traffic's table.....	71
Figure 31 Data preparation in Keboola - Daily traffic's table.....	71
Figure 32 Data preparation in Keboola - Daily traffic's data.....	72
Figure 33 Data preparation in Keboola - Run transformation.....	72
Figure 34 Data preparation in Keboola - Daily transaction.....	73
Figure 35 Data preparation in Keboola - Final query.....	74
Figure 36 Data preparation in Keboola - Writing Daily transaction query.....	75
Figure 37 Structure of Traffics, Transactions, Products and Customers tables of SPORTISIMO.....	76
Figure 38 RStudio.....	77
Figure 39 RStudio - Load data.....	79
Figure 40 RStudio - Check data.....	79
Figure 41 RStudio - Convert Date from List to Date type.....	80
Figure 42 RStudio - Load data.....	81
Figure 43 RStudio - Group by based on date and sum of.....	81
Figure 44 RStudio - The graph shows sales in 2016.....	82
Figure 45 RStudio - The ARIMA Model for preliminary analysis.....	83
Figure 46 RStudio - Sales forecast of upcoming few days.....	83
Figure 47 Comparison of ARIMA model and Time series.....	86
Figure 48 Power BI report.....	88

1. INTRODUCTION

With virtual e-shops, the webspace of enthusiasts who can offer you new models can show you and be able to choose for themselves through the possibility of change. By providing all the services of a website and accessing it as a new operating system, you can provide different companies and organizations with more data and information on your site using new resources and tools of your own. The website's information and scientists can be downloaded and provided with exploration and data mining. Websites can play an important role in providing limited information and knowledge.

With the advent of the internet and the development of e-commerce, the experience of business has taken on a new look. Enterprises using e-commerce models, companies are facing heavy competition. By referring to specialists and customer satisfaction and loyalty, organizations can manage customer relationship management as a tool for their profitability services.

Web technologies enable each other's limitations concerning each other. Being a Service Provider is an exceptional service and the opportunity to create a solution opportunity for the website.

1.1 The necessity & importance of customer relationship management

Web mining techniques to retrieve, extract, and automatically evaluate information to discover the knowledge of web documents and services, and analyzes web structural information such as links between web pages and information on how users use the web. Data mining can be used in areas such as market analysis and management, market portfolio analysis, target marketing, understanding customer behavior, risk analysis, and management. Delivering superior quality products and delivering excellent customer service continually creates a competitive advantage for companies and organizations and can also leverage customer retention and long-term retention.

Customers who are dissatisfied with the products or services being offered are sharing their experiences with others and providing the opportunity to reduce the company's revenue and profits. Changing expectations, needs and demands imply that continually Monitoring and assessing customer satisfaction is essential.

Undoubtedly, the most important asset of most organizations is their customers. Customers are a valuable source of opportunities, threats, and operational questions relevant to the industry because of their direct relationship to the actions of an organization. Today, for growth and survival in the field of economic competition, companies and organizations must pay particular attention to customer orientation and further enhance their relationship with buyers of goods.

1.2 Goals of customer relationship management in e-commerce

In new business trends, customer satisfaction has been a vital part of corporate goals, and top executives understand their success in achieving corporate goals is dependent on customer satisfaction. The purpose of customer relationship management is to empower the organization to deliver better customer service by creating automated and integrated processes for collecting and processing customer information.

In sum, achieving the following benefits is one of the goals of customer relationship management in the field of e-commerce:

- Increase awareness of the company's online presence;
- Increase the trust and credibility of the company in the online environment;
- Enhance perceptions of the company as a business using high technology;
- Increased customer support during and after the purchase process;
- Increase the total number of users and reduce the cost of attracting new customers;
- Increasing the number of customers who generate more revenue and benefits;

- Increase the rate of handling and resolving customer complaints and answering users' questions;
- Increase and improve the use of advanced technologies compared to competitors;
- Deploying business processes much faster and more effectively;
- Increase the capability of business processes used in the company;
- Improve site content management and increase profitability;
- Increase the variety of services offered on the site such as email, essential phones, quick-response system.

Good and effective customer relationship management requires a clear understanding of customer needs. Management needs to be attentive to the day-to-day needs of customers and act accordingly.

The most important application of web mining in the field of e-commerce and economics is to identify many of the facts and factors affecting customer relationship management and customer service through behavior and system communication.

So, web mining helps to quickly identify the customer and reduce the cost of customer reviews. Web mining is one of the elements of customer relationship management in the field of e-commerce and can help companies move towards the customer.

In the process of web mining in customer relationship management, raw data is collected from various sources and becomes useful knowledge and information. Web mining by data mining techniques can predict the potential profitability of potential customers who can become actual customers. Web mining can then determine the value of customers, predict their future behavior, and make informed decisions.

1.3 Purpose of the thesis

Due to the fierce competition in e-commerce, the decisions made to attract and retain better customers are essential for the success and survival of the business and create an effective customer relationship management plan and a well-designed website using data mining techniques. It seems to be the only way forward, and data mining techniques can greatly contribute to the success of e-commerce.

In the present study, the aim is to develop and present an improved approach to existing methods in the field of e-commerce and customer orientation in the virtual world, thus focus more on increasing customer satisfaction and loyalty, increasing the profitability of companies and organizations can achieve the goals.

2. LITERATURE REVIEW

2.1 Definition of e-commerce

Despite the spectacular dot-com bust of a few years ago, the internet has markedly changed the way we do business, whether it's finding new streams of revenue, acquiring new customers, or managing a business's supply chain. E-commerce is mainstream - enabling businesses to sell products and services to consumers on a global basis. As such, e-commerce is the platform upon which new methods to sell and to distribute innovative products and services electronically are tested. The Web's influence on the world's economy is truly astonishing. The business world knows that the Web is one of the best ways for businesses such as manufacturers to sell their products directly to the public, brick-and-mortar retailers to expand their stores into unlimited geographical locations, and for entrepreneurs to establish a new business inexpensively. Thus, the executive in the 21st Century must know 1) where technology stands in the business processes of his or her company, 2) how technology relates to the company's strategies, 3) how rapidly technology changes and evolves, and 4) how the company and its business partners will respond to the changing technology. (Reynolds, 2004)

E-commerce means the production, marketing, sales and delivery of goods and services using electronic tools. E-commerce is divided into four main groups based on the fact or legality of the parties to the transaction and the type of relationship they have: B2B (Business-to-Business) - B2C (business-to-consumer) - C2C (consumer-to-consumer)- Consumer to Business (C2B). Although e-commerce hasn't been around for a long time, it has played a role in our daily lives, so avoiding it is not easy.

Buying and selling products through credit cards are one of the simplest and most effective e-commerce roles in the everyday life of internet users. E-business is the application of all information systems to enhance and control business processes.

Today, these processes are developed using web-based technologies. E-business in simple terms means doing business using telecommunications and related tools. E-commerce in this sense has a smaller and more limited concept than e-business.

In e-commerce, exchanging money between two parties is necessary to get to the point of business, but in e-business there is no need to exchange money. For example, a business meeting via the Internet is considered an e-business.

2.2 Businesses that sell directly to consumers (B2C)

The purpose of B2C marketing is to push potential customers to the product and maximize the value of the transaction through activities that provide incentives to buy even more products. The relationships between companies and potential customers are based on customer experience during the buying process. The following describes the consumer market buying process:

- It cares about the benefits of using the product and how it relates to the person.
- It wants to do a simple, preferably with just a few steps/clicks shopping processes.
- It can be affected by the market.
- It primarily makes emotional purchasing decisions based on status, desire, and price.

2.2.1 B2C Marketing

In B2C marketing, marketing activities should be built so that people can make their purchases fast and easy. B2C companies can base their business activities on such things as shopping bins, exhibitions, stores (both real and online) and so on so that they can finally reach the marketplace for services. And encourage their products.

2.2.2 Businesses that sell directly to businesses (B2B)

The purpose of B2B marketing is to maximize the value of the relationship through the marketing efforts that your target audience offers to your value. If successful, your marketing activities will lead to a relationship that is based on trust and credibility. The value of the relationship is determined by what is important to the customer.

Business customers tend to be more logical and rational when making purchase decisions. Your marketing should target the decision-maker in a business and then help them find solutions to their specific problems and challenges.

There is usually a longer sales cycle due to a multi-stage buying process where the B2B customer is conducting further research on the right solution for their company. When a business buyer sees the benefits of your solution, they look for deep insights into their business reputation.

2.2.3 B2B Marketing

B2B marketing refers to the tactics and strategies in which a company promotes its products and services to individuals: producing, advertising, and selling products to customers for use in their everyday lives.

2.3 E-commerce & sales

Every business about customers is spent a lot of effort to attract new customers and maintain the current customer. Due to intense competition, attract new customers is not an easy task. However, there is little doubt that e-commerce companies' success depends on how they interact with their customers.

How can we reach the hearts of consumers? There are various ways to use what is called artificial intelligence marketing. What is Intelligent Marketing? While this may sound complicated, most marketers are nothing more than a data entry and analytics device and ultimately give us an intelligent output. Finally, artificial intelligence marketing can help us as a marketer to personalize suggestions beyond the reach of a human being.

Strangely, this means that instead of having to manually sort segmented data and messaging each user one by one, you rely on a machine to perform this heavy-duty operation. While dedicating your time to valuable work. That's why the first "way" to create e-commerce

experiences is to know more about customers. Artificial intelligence marketing is the result of a mix of artificial intelligence (building systems that can make intelligent decisions) and marketing. It's the connection between marketing and artificial intelligence - and the only way to scale personal automation.

2.3.1 Gain proper insight

To create personalized experiences, we need to have complete information about the people you provide your services or products to Who are they, what do they like, where are they, what is their gender, how old are they, and so on.

2.3.1.1 Customer intelligence

The idea here is to get clear and practical insights from data that you can use. To do this we need to:

- Collect and execute analysis data ever-increasing data volumes
- Use the information and turn them to actionable outputs communicate with your audience as to where our customer is located most

2.3.1.2 Product Offers

Website product predictions are one of the simplest applications and one of the most effective personalization tactics for B2C brands. Based on data mining techniques, personalized offers show customers that you care about them because these offers are different for everyone online. But these tactics need to be aware of what the person is searching for, what they have purchased and what has been reviewed and purchased by similar customers. No matter how you present the products to the customer... what is more important is that you present relevant offers to everyone in a particular situation.

2.3.1.3 "Recommended for You" / "You May Like It"

Amazon uses interesting complexity "inspiration" (recommendations based on the location of your recent behavior inspired), the product offers its marketing strategy has achieved. Amazon can obtain customer visitation data and intelligently present information to its users promptly. This tactic allows you to show the customers you know what they might want.

2.3.1.4 "Other Customers Liked"

Successful shopping sites use this technique to market their products to new users based on the opinions of previous users and buyers who have bought or reviewed that product.

This tactic is one way to gain a customer's confidence, and you are saying to your customer, "We are here to help you have a great shopping experience, thinking about what we know about you so far. You may also like this product based on reviews"

When dealing with hundreds of thousands or even millions of contacts in your database, it is impossible to manually display relevant suggestions for everyone.

So, intelligent systems and algorithms should be used to follow each popular quote and to examine each user's behavior and opinions and provide intelligent suggestions based on this and other information. Using this you will have much more loyal customers and you will gain more trust.

2.3.1.5 "Recently Viewed"

One sure way to gain customer trust to get back to your website is to show products that have been viewed or searched by the user with a section called "Recently Viewed".

2.3.1.6 Share what you viewed

You can even share what is viewed by many of your website's audience in the form of a section that has been viewed by others recently, which can be important and trusting. Use photos, descriptions of any product, price, and availability of the product, and display it to your customer while the customer is at your site (customer lifecycle).

2.3.1.7 "Continue Search" / "Resume Your Last Search" / "Recent Activities"

Similarly, you can encourage users to "continue your search" (or purchase) by reminding them of recent user activity. This can work very well in both product sales and especially in tourism and hotel-related matters. Depending on your strategy for re-recruiting yourself, this can be as a notification in the center and on your home page or wherever you and your customers alike.

2.3.1.8 Know the customer's location

Unique offers based on a customer's location are more personalized and customized than standard personalization tactics. This can reinforce previous tactics to personalize product offerings and gain customer confidence if used correctly.

2.3.1.9 Personal motivations

Personal discounts apply based on past behavior of customers (whether they have previously accepted or rejected them). The best way is to let artificial intelligence and a pre-written algorithm Decide a discount for our customers. In this case, each customer will be satisfied with the discount and discounted product that is offered and will, therefore, trust your website.

2.3.1.10 Clickstream flow analysis

Click flow means the order of web pages visited by the user. When a visitor accesses the website, the server keeps track of all the activities the visitor has done in the log. Where he/she clicked, how long he watched a certain webpage, and so on., and the user session arranges web pages from multiple sites that the user has viewed in a given period time. Each mouse click corresponds to a page request and the order of clicks corresponds to a sequence of links. The analysis of sequential clickstreams can be used to understand user behavior. This analysis shows what pages a user is interested in or has been interested in for a long time, and which pages are quickly rejected. Two types of analytics are generally used for clicking streams:

- Traffic analysis
- E-Commerce analysis

➤ Traffic analysis

Traffic analysis examines user traffic by collecting click-through data stored on the server. The analysis also records the number of times a user has loaded pages during navigation.

➤ E-Commerce analysis

In e-commerce analytics, the flow of data clicks is used to determine effective marketing by evaluating user behavior during site visits and reviewing sales transactions. This analysis is an indicator that determines the degree of user convenience in using user interface forms, using shopping cards, payment methods, and so on.

Clickstream analysis is an important aspect of online business and advertising because it captures how users move from one page to and can help you to better understand potential customers. If the user lays down pages, they are likely to be interested in the content of that page.

By creating user profiles for specific sites, clickstream analysis can be effective in designing a website and increasing customer satisfaction as well as identifying the right places to place ads. Referrals from the home page to the other pages are promoted by the user, and by analyzing the clickstreams, you can determine the type of pages and sites the user is interested in or direct the user to where they can go. The main purpose of promoting effective advertising is to show the link between advertising and web pages that ultimately leads to buying.

Clickstream analysis can also be used for personalization, by combining the shopping cart analysis to bring together the goods that are most customer-focused. So, it can do the layout of the web pages and the design of the virtual store according to customer's preferences and not bring discounted items together. This approach can be used to attract and persuade customers to buy, in addition to improving the status of sellers.

In general, the benefits of this method are paying attention to the customer and how he or she navigates the site, paying attention to the appropriate design of the site to attract the attention of the customer, and how to present and arrange the products and display the appropriate advertising. While attracting customers, we can take steps to maintain their satisfaction and loyalty.



Figure 1 Artificial intelligence in e-commerce

2.4 Understanding customer lifetime value

CLV is the profitability of a customer over their entire relationship with the business. Businesses need to look at long-term customer satisfaction and relationship management, rather than short-term campaigns and quick wins – this approach leads to increased value over the entire lifetime of a customer and means that CLV is a metric central to any CRM initiative. (Quirk, 2008)

It's important to look at your customer base and segment them according to how often they purchase and how much they spend with your company. Very often, customers who spend more cost more to acquire, but they might also stay with the company longer. Referrals made by a customer can also be included as part of the revenue generated by the customer. (Quirk, 2008)

For example, a potential customer looking to purchase a digital camera is likely to search on Google for cameras. As a company selling digital cameras, your excellent search advert and compelling offer attract the potential customer, who clicks through to your website. Impressed with your product offering, the user purchases a camera from you, and signs up to your email newsletter as part of the payment process. (Quirk, 2008)

Analyzing the amount spent on your search campaign against the sales attributed to the campaign will give the cost per acquisition of each sale. In this case, this is the cost of acquiring new customer. (Quirk, 2008)

As the user's now signed up to your newsletter, each month you send her compelling information about products she might be interested in. These newsletters could be focused on her obvious interest in photography and highlight additional products she can use with her new camera. The costs associated with sending these emails are the costs of maintaining the relationship with the customer. When she purchases from you again, these costs can be measured against the repeat sales likely to be made for the customer's lifetime. (Quirk, 2008)

Assuming that a customer buys a new camera every three years, moves up from a basic model to a more expensive model, perhaps buys a video recorder at a certain point – all of these allow a company to calculate a lifetime value and ensure that their spending on a particular customer is justified. (Quirk, 2008)

2.5 Online shopping

An online store is a website where consumers buy products or services. This type of site is most referred to as an e-commerce site or a "B2C" (Business to Consumer) site. In addition to most, if not all, of the content found in a brochureware site, an online store displays products/services along with detailed information (e.g. specifications and pricing) usually from a database with search features and a method for online purchase. An online store must also provide extensive information about the products/services offered that not only aids in attracting consumers but gives them enough confidence in the seller and the products/services to take the

next step making an online purchase. One question the author is often asked is "what should an e-commerce site offer? A process to order online? A toll-free number? Or both?" The answer is: Offer. You choose to take online payments, you must provide a secure, reliable, cost-effective system for authorizing payment and managing transactions. The best systems are based on the Secure Socket Layer (SSL) and/or Secure Electronic Transactions (SET) encryption technology, which provides the encryption of data and generates and displays a "results page" to the customer following the transaction. Further, a successful online store must be designed with the ability to store orders in a database or as tab-delimited text files so the data can be imported into an invoicing system. Then the website must be able to intelligently route encrypted email to the order fulfilled. (Reynolds, 2004)

Online shopping has become a part of our lives today. To do online shopping, internet businesses need to think about how easy the online shopping process is for users. They must tailor their online sales strategy so that the online shopping process can be done simply. In digital marketing, the importance of the online sales process and the user experience in this process is very high. It should also be borne in mind that today's customers are much more aware than before and compare the price and quality of the product or service with the price and quality of other brands before making a purchase.

Here are some of the most important benefits of an online store:

2.5.1 Compare prices

You can easily go from one online store to another online store, browse, compare prices and choose the gender you need. You can even use relevant software developed to compare prices. Online stores usually offer good discounts on occasion to protect their customers. By subscribing to the newsletters of these stores you can find special discounts.

2.5.2 Unlimited selection:

There are many choices in an online store. Even if you haven't found the product you want in one store, you can easily go to another store with one click.

2.5.3 Access to Consumer Reviews:

You can easily read the comments of other consumers and customers about a product in the comments section and make informed decisions.

2.5.4 Get access to your company services within 2 hours

One of the most important benefits of the online store is that it allows customers to register their orders anytime, anywhere, and on any device (phone laptop, tablet, etc.). Imagine how much this could benefit your business.

2.5.5 Analyze the accuracy of visitors:

In internet marketing with tools such as Google Analytics and so on, top products could be observed by customers based on other clients' search histories and make appropriate executive decisions to convert more and more visitors into customers.

2.6 Data

Data includes characters which are gathered and translated for some purpose, usually, they are used to be analyzed. If data is not put into context, it isn't worth anything to a human or computer. Here are some common types of data: Single character, Boolean (true or false), Text (string), Number (integer or floating-point), Picture, Sound and Video. Computers are called numbers, letters, and symbols for the common understanding of humans or computers.

In simple terms, data can be defined as a set of raw materials, such as bank documents. These days, data is emerging in newer forms that can be referred to as metadata. Zero and one occur, known as Digital data.

2.7 Information

Information refers to data that is used as raw data for someone else to use the information to accomplish their goals. Information is meaningful and, like data, is not ambiguous, the information allows users to understand the nature of the underlying data and to understand and use the data as a result. When researchers obtain data and establish relationships between data and variables, these relationships ultimately lead to a result called information. Information is the relationship that is created between these variables.

2.8 Data Mining

Data mining is the analysis of (often large) observational data sets to find unsuspected relationships and to summarize the data in novel ways that are both understandable and useful to the data owner. The relationships and summaries derived through a data-mining exercise are often referred to as models or patterns. Examples include linear equations, rules, clusters, graphs, tree structures, and recurrent patterns in time series. (by David Hand, 2001)

Nowadays, companies gain a lot of information through continuous customer service and communication, which can be of great benefit if they know how to use this data. Data Mining is simply a problem-solving method that extracts iterative patterns by analyzing large volumes of data. It then offers solutions to challenges by finding connections between different events and these patterns. In fact, data mining discovers valuable results and makes them usable from information that may be of no use.

Data mining is a powerful science that can guide everything and answer many of our questions unknowingly. Nowadays, the importance of this science in large companies is well known so

that prior to making decisions and planning for specialized campaigns or designing costly products, they first attempt to obtain public data.

Although organizations strive to maintain customer and sales relationships, many of the world's technology flagships are still not easily accessible to their competitors. One of the great mysteries of this success may be the benefit of data mining knowledge. In some of these companies, data mining is so important that they even create data collection campaigns.

Advanced business analytics includes diagnostic analytics, prospective analytics, and prescription analytics. Some examples of questions that will be answered in advanced business analytics are:

- a. Who are the different segments of the market and our real customers?
- b. How to identify customer buying patterns?
- c. How much is the loyalty or turnover of current customers?
- d. Is it possible to predict the number and amount of customer purchases based on past purchasing behavior?
- e. Is the profitability trend of the organization appropriate to the market size?
- f. Is the market trend towards our new product better than that of our competitors?
- g. Is the warehouse appropriate to current and future production needs and trends?
- h. How to prevent financial fraud and money laundering in the organization?
- i. How is the optimal balance between costs and profits created?
- j. Do the financial indicators of the organization go hand in hand with strategic goals?
- k. Is it possible to adjust the production process according to customers' tastes and market elasticity?

Data mining is the process of discovering the patterns and trends of regular and hidden data in large, distributed data, using a broad set of mathematical and statistical algorithms. These algorithms are usually applied to numerical and non-textual values and for textual data, text mining algorithms are used. Data mining utilizes science such as artificial intelligence, machine

learning, statistics, operational research, and database management to build models and answer questions.

Extracting and analyzing an organization's information from the data available to employees is a process that has been carried out for many years and is not a new task in organizations. The first algorithms to identify regular trends and patterns in the database originate from the science of statistics and probability theories.

In recent years, with the increasing computing power of computers and the ability to achieve the results of complex computations in a short time-period, advanced mathematical algorithms have attracted attention. These algorithms refine and analyze the data by considering different dimensions of the data, extracting and presenting complex and undetectable patterns by traditional methods. Computers have helped the process of extracting, refining, pre-processing and modeling data as well as validating the findings with extraordinary accuracy and speed.

2.9 Data mining algorithms

Based on the relationships that data mining has with the various sciences, it uses a variety of algorithms and methods. Different classifications are provided for the algorithms. According to the most popular classifications, these algorithms can be divided into two categories: descriptive algorithms and predictive algorithms. Accordingly, some of the most important descriptive algorithms are as follows:

2.9.1 Clustering

Clustering refers to a very broad set of techniques for finding subgroups, or clustering clusters, in a data set. When we cluster the observations of a data set, we seek to partition them into distinct groups so that the observations within each group are quite similar to each other, while observations in different groups are quite different from each other. (Gareth James, 2013)

One of the unobservable approaches to data knowledge exploration is that by default, it separates and categorizes the data based on the significant distance. The members within a cluster are very similar, but the members of each cluster are less like those in the other clusters. Clustering algorithms such as K-Means, K-Medoids, AGNES, DIANA, DBSCAN, and OPTICS are among the algorithms. Examples of data clustering are as follows:

- Market segmentation based on customer referrals and contacts with the organization
- Determining marketing strategies based on customer buying behavior in different clusters
- Determining the types of contracting contracts based on the process of paying clients
- Customer segmentation based on the costs and revenues they generate for the organization.

Partition-based clustering typically places more emphasis on the score function than on any formal notion of a model. Often the centroid or average of the points belonging to a cluster is a representative point for that cluster, and there is no explicit statement of what sort of shape of the cluster is being sought. For cluster representations based on the notion of a single "center" for each cluster, however, the boundaries between clusters will be implicitly defined. (by David Hand, 2001)

2.9.2 Association Rules Algorithm

Association rules are among the most popular representations of local patterns in data mining. An association rule is a simple probabilistic statement about the co-occurrence of certain events in a database and is particularly applicable to sparse transaction data sets. For the sake of simplicity, we assume that all variables are binary. (by David Hand, 2001)

Dependency rules are a convenient and powerful way to find attractive relationships between variables in large databases. According to this method, the degree of dependency between attributes or data fields is identified as a set of rules, and these rules determine the possibility of two or more attributes occurring side by side. Some dependency calculation algorithms include Apriori and FP-Growth. For example, dependency rules are used in analyzing the following:

- Analyze the customer buying process in new shopping bases based on previous purchases
- Finding the relationship between machinery failure and parts and their causes
- Analyzing the Relationship between Corporate Marketing and Branding Activities with Increasing or Decreasing the Market Reaction to Company Products
- Grouping or bundling products or services based on customers' interest in their group purchase.
- Communicate the quality of the after-sales service experts' response to customer needs based on their expertise, product type, product maturity, and similar characteristics

The most important predictive algorithms can be broken down as follows:

2.9.3 Classification of data

One of the supervised learning methods for predicting data classes is to identify new data classes based on predefined classes. Some data classification algorithms include Decision Tree, Naive Bayes, and K-Nearest Neighbors. Some examples of classification applications are as follows:

- Classify resellers based on sales and marketing capabilities.
- Decision tree to decide on the process and how to sell to large or specific customers.
- Learn the types of customer grouping based on contract, warranty or support history and determine the organization's approach and attitude to new customers before the start of the support or warranty period.
- Separate new customers based on the duration and type of service or product received, based on the records learned from current customers and customers separated from the organization.

In classification, the variable being predicted is categorical, while in regression the variable is quantitative. The term "prediction" is used here in a general sense, and no notion of a time continuum is implied. So, for example, while we might want to predict the value of the stock market at some future date, or which horse will win a race, we might also want to determine the diagnosis of a patient, or the degree of the brittleness of a weld. A large number of methods have been developed in statistics and machine learning to tackle predictive modeling problems, and work in this area has led to significant theoretical advances and improved understanding of deep issues of inference. The key distinction between prediction and description is that prediction has as its objective a unique variable. (by David Hand, 2001)

2.9.4 Regression and Time Series Analysis

If the organization's data is recorded over time and according to a specific trend, it is possible to predict the future trend of the organization based on time series and data regression functions. Regression is one of the simplest data mining methods that provide estimates of future values based on past trends. Some examples of regression applications are as follows:

- Analysis of the effect of economic indices on the trend of rising or decreasing stock prices in the stock market
- Compare the financial performance of a company compared to similar companies in the same industry
- Analyzing revenue growth trends based on marketing and sales organization policies
- Identifying factors and time frames affecting stock price volatility
- Comparison of compliance of production line performance with organizational policies and goals in previous years

2.9.5 Neural Networks

Artificial neural networks are a sophisticated model of continuously trained regressions that can predict the future based on past events. The data prediction in this algorithm is similar to the classification algorithms, in that it aims to predict continuous data values (such as the number of sales, income, profit or stock prices) but in the classification, the prediction results are based on the current classes that the algorithm uses. It is trained, provided and the output has a discrete spectrum. (In fact, the number of classes is limited to a few classes or groups and has little variation). Some types of neural networks include MLP (Multilayer perceptron), RBF (Radial Basis Functions), SVM (Support vector machines), SOM (Self-organizing map) and LVQ (Learning vector quantization) networks. Examples of neural network applications are as follows:

- Learn and predict the behavior of current and future customers based on feedback from the bidding system
- Optimize sending and receiving of support products and contracts based on past trends
- Adjust salespeople's approach to potential and new customers based on learning from current customer behavior
- Understand the quality of work of senior staff in responding to customer requests and identify the causes of noncompliance with the normal and everyday pattern of organization.
- Predict stock prices based on country economic indices and stock market financial indices and make offers to buy or sell portfolio stocks to optimize total dividend yields.

2.10 Business Intelligence

Business Intelligence helps companies and organizations to identify weaknesses and strengthen their strengths, facilitating process optimization, and having in-depth knowledge of all the factors such as customers, competitors, economic environment, operations and organizational processes for better decision making.

Business Intelligence helps consolidate, analyze, and access large amounts of data to make business decisions. Online analytical processing supports categories of software which allows users to analyze information from multiple database systems at the same time and multidimensional analysis and performs complex queries on a large volume of data, data mining and web mining techniques identify customer behavior patterns across a large volume of information.

Business Intelligence (BI) refers to technologies, applications, and practices for the collection, integration, analysis, and presentation of business information. The goal of Business Intelligence is to support better business decision making. And, Business Intelligence systems are data-driven Decision Support Systems (DSS).

3. PRACTICAL PART

3.1 Craft.CASE



Figure 2 Craft.CASE

Craft.CASE is a business process analysis (BPA) tool based on a sound C.C method. The core of this method explains how to progress in a BPA project without forgetting anything. The C.C method consists of small steps, sequences of which are tested and validated as soon - and as often - as possible. Following this method allows your processes to remain consistent even if the problem is complex. Craft.CASE leads its users step-by-step according to the C.C method. This means that the notation of the tool is rigid enough to discover, understand, and consistently analyze processes. With Craft.CASE you can easily get your BPA done. The tool leads you according to a complex and powerful method. Even though it may seem difficult, it is very simple to use.

3.1.1 Why choose Craft. Case:

- Analysis of business and conceptual phases
- Ability to set the user-defined properties to individual objects
- Generating reports and specifying their content

3.2 Business Diagram (Online shopping)

3.2.1 Order products

Craft.CASE can help explain more about the different parts of the online shopping business, in the pictures below you can find the business diagram for ordering products. It includes two parts. The first part is “External Process” and the second part is “Internal Process”. “External Process” is related to customers and all processes which customers should do to order products in the system. “Internal Process” can show all steps which are related to the system and how to check products or how customers can pay for orders. This diagram shows how customers can order a product and how the process works.

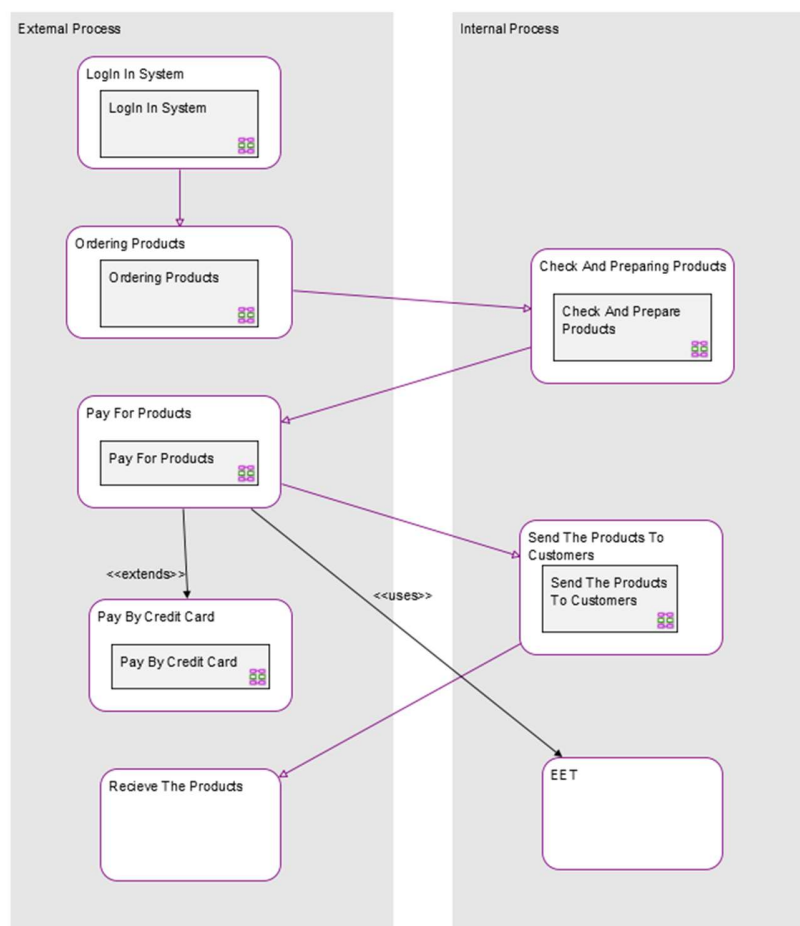


Figure 3 Business diagram - Ordering products in online shopping

3.2.2 Login system

Each customer should register and login to the system to order products. Here is the process that shows how customers should log in to the system and how the system should validate the login.

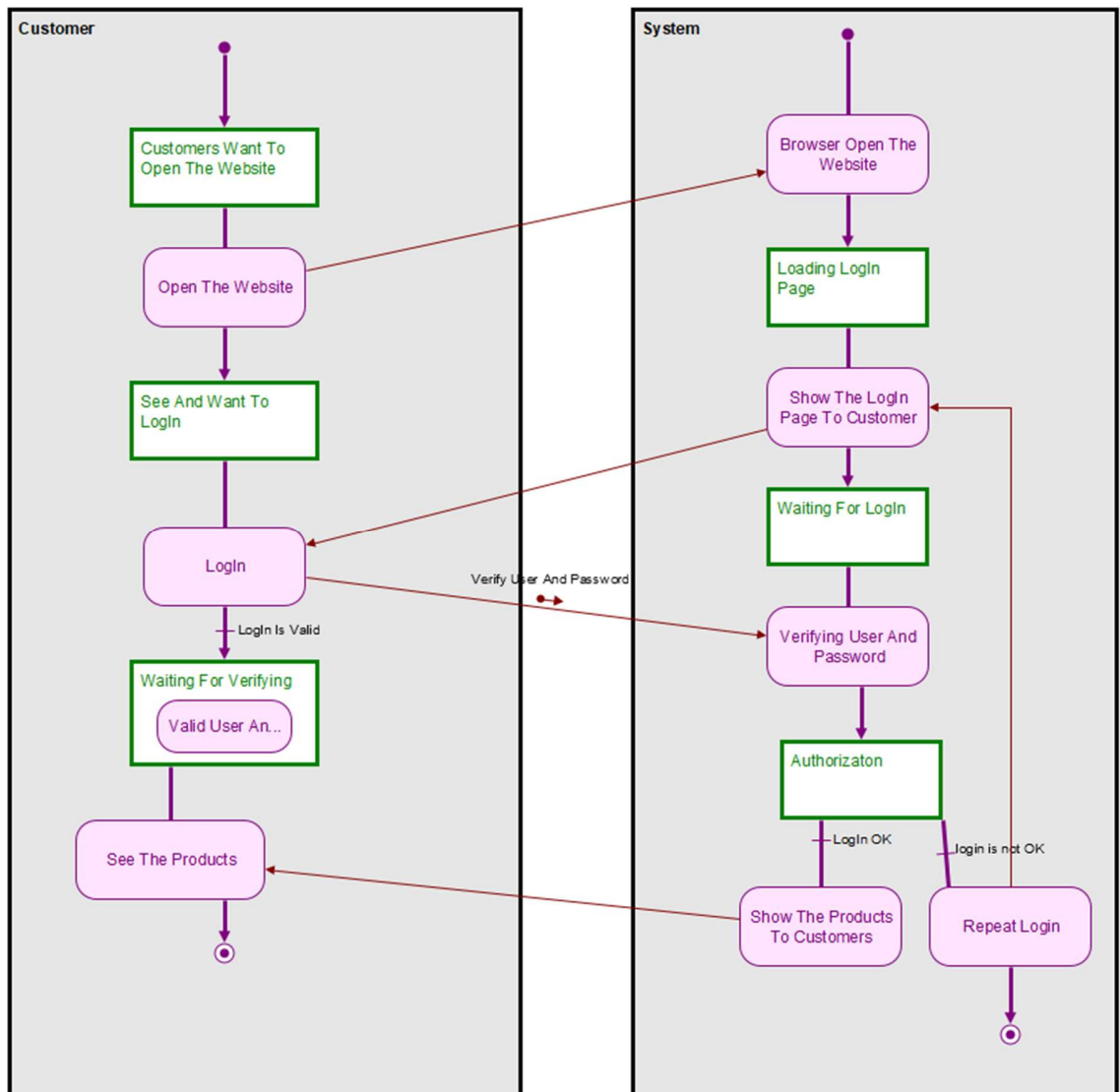


Figure 4 Business diagram - Login customer in system

3.2.3 Ordering products

After login customers can start ordering in the system, they can see all the products and check those which are available, or which are not. The system should show these products to the customers so they can choose and order them.

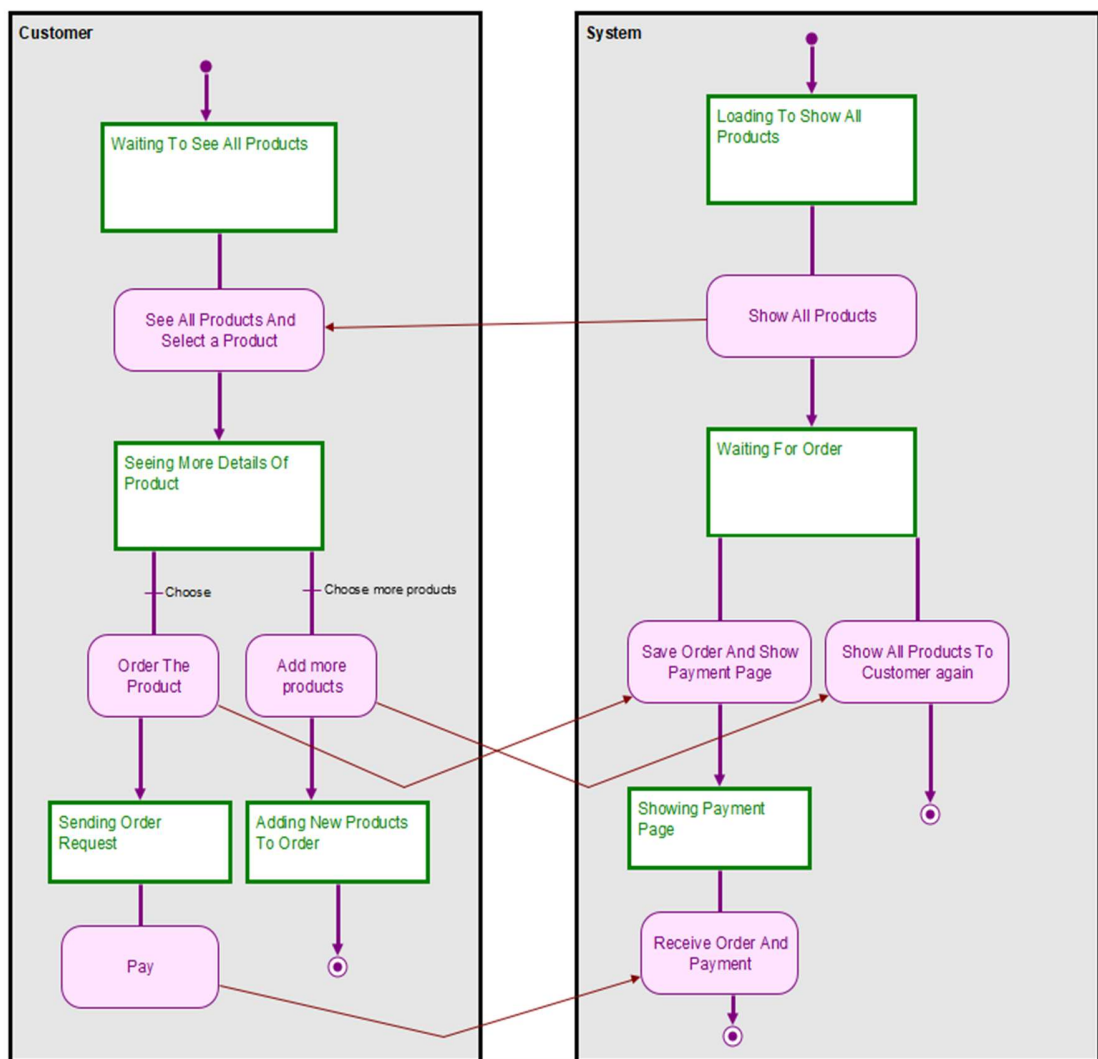


Figure 5 Business diagram - Ordering products by customers in system

3.2.4 Check and prepare the products

After the customers order the products, the system approves and sends the orders to the employees so they can prepare the products for the customers.

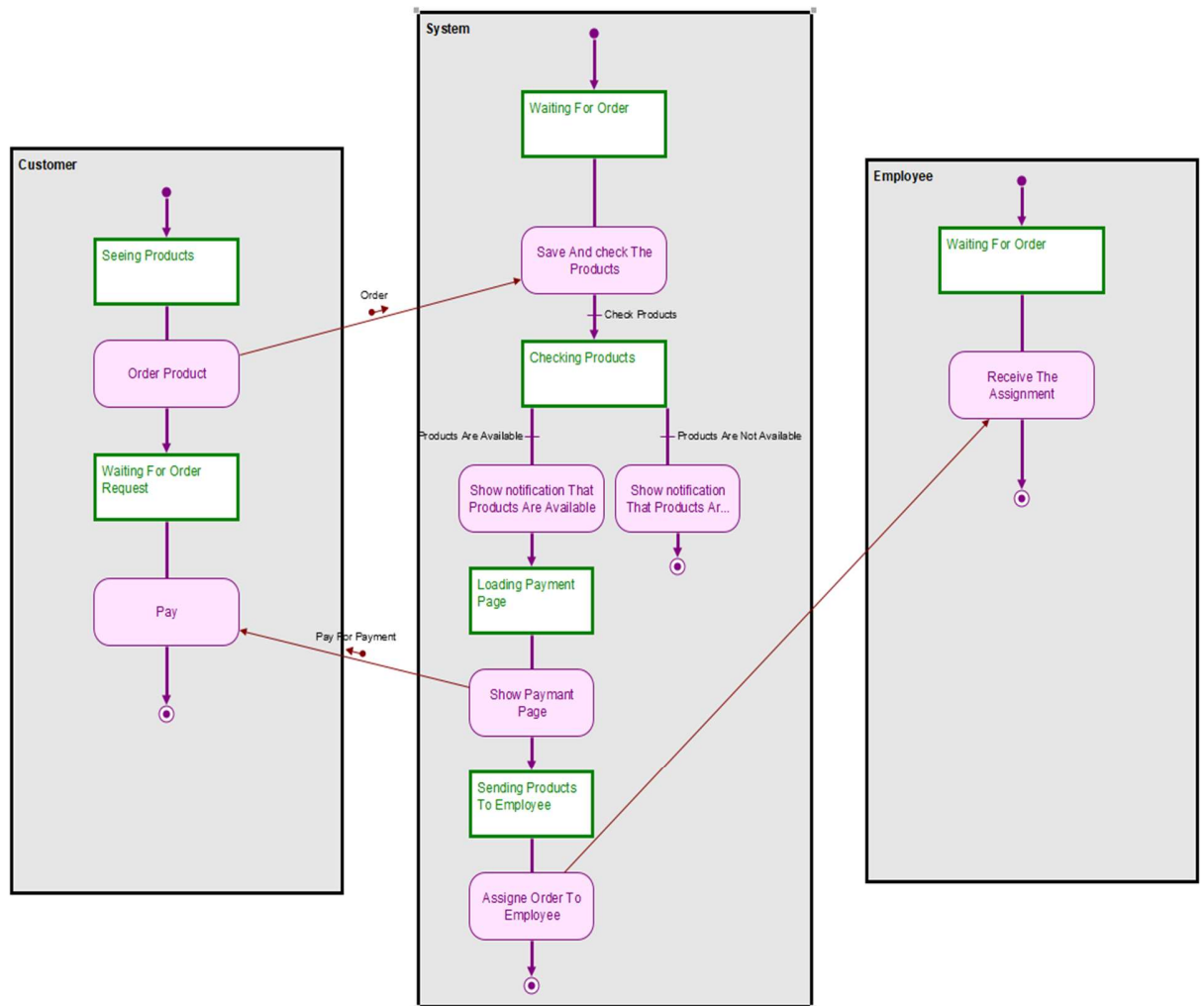


Figure 6 Business diagram - Check and prepare the products in system

3.2.5 Pay for products

Here is the process of paying by cash. In this process the transformer delivers the products to the customer and then the customer can see the bill and pay by cash. Before this step, the system shows the bill to the customers after the products are ordered.

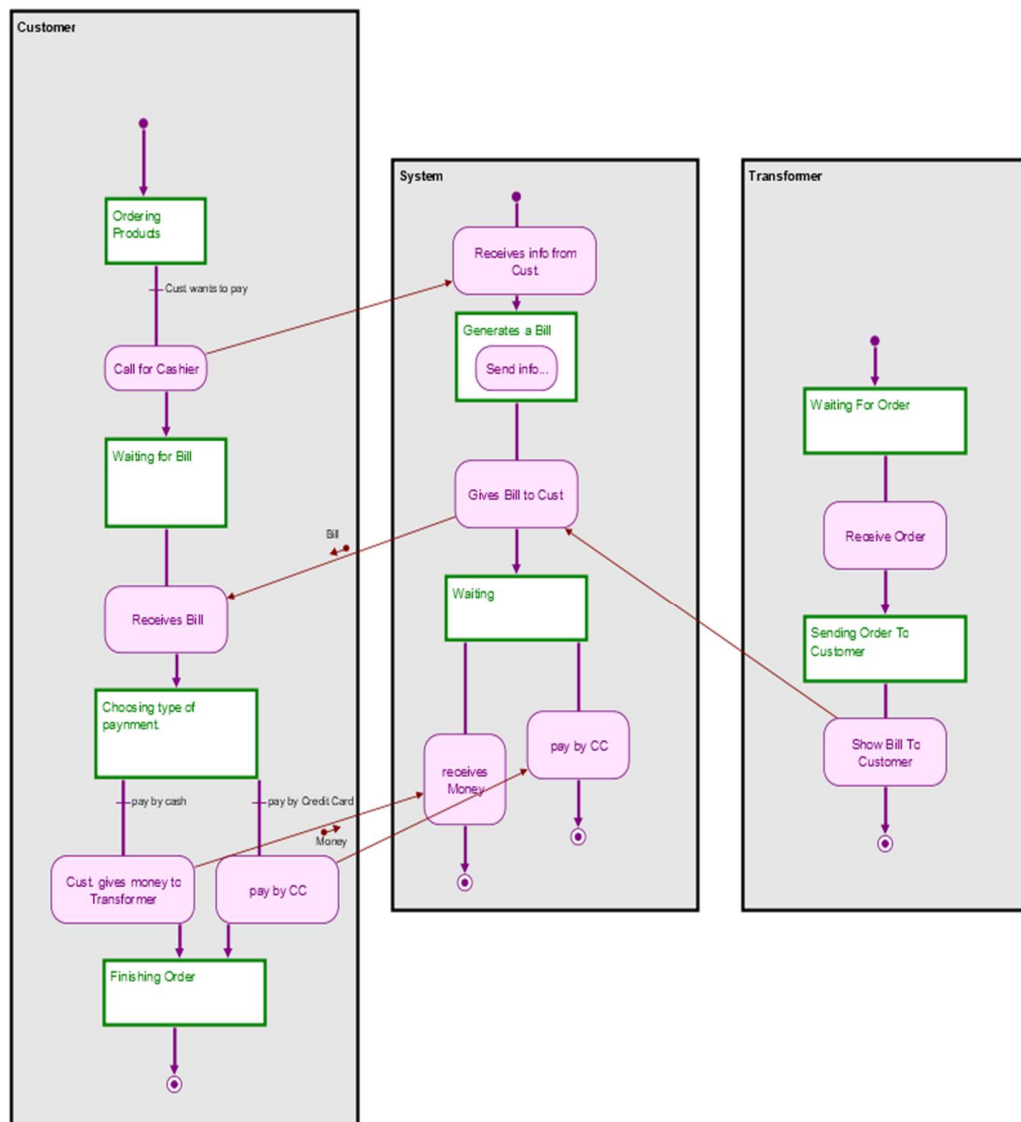


Figure 7 Business diagram – Pay for products in the system

3.2.6 Send the products to the customers

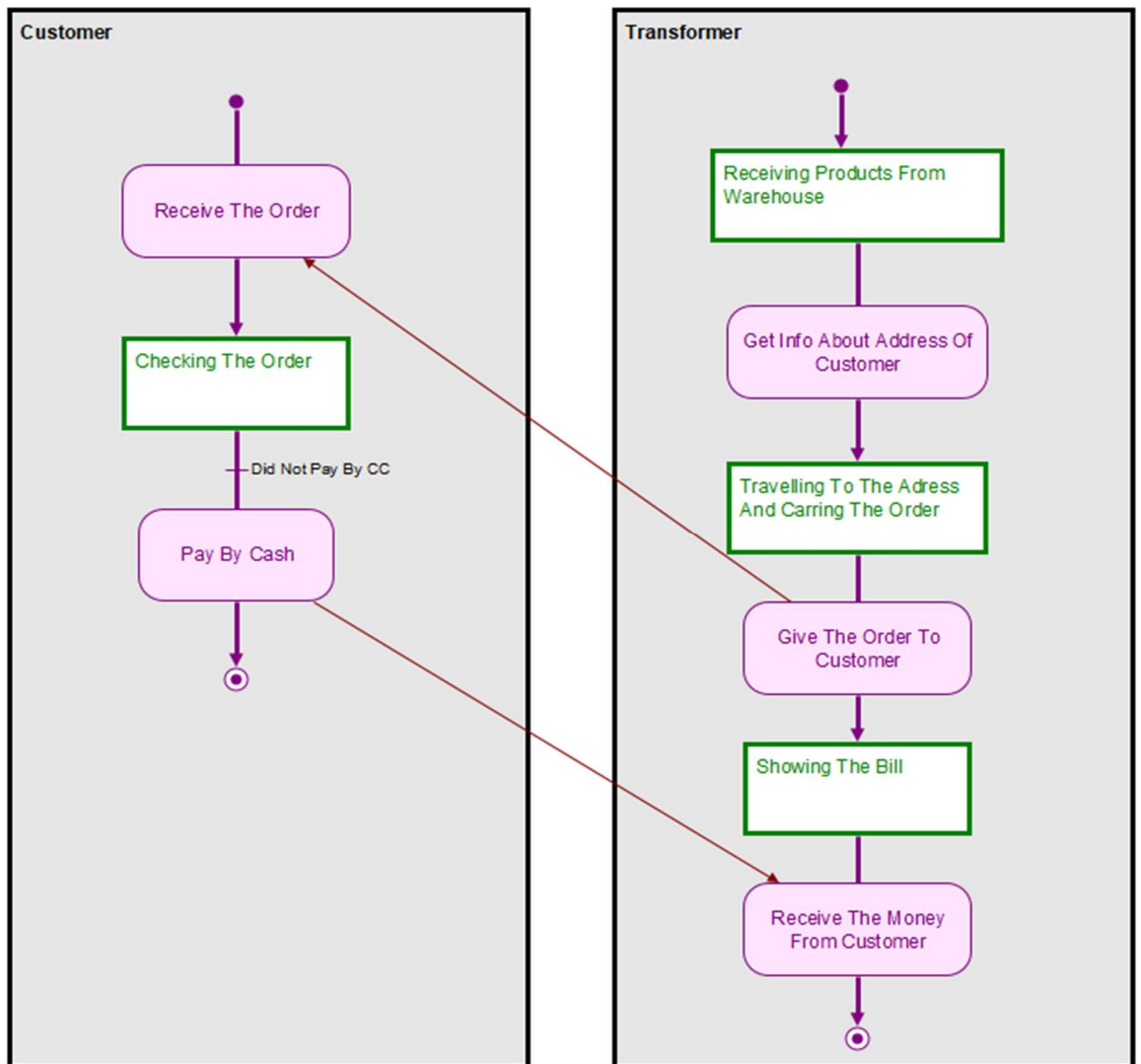


Figure 8 Business diagram - Send the products to the customers

3.2.7 Pay by credit card

The system can also offer a method so that customers can pay by credit card.

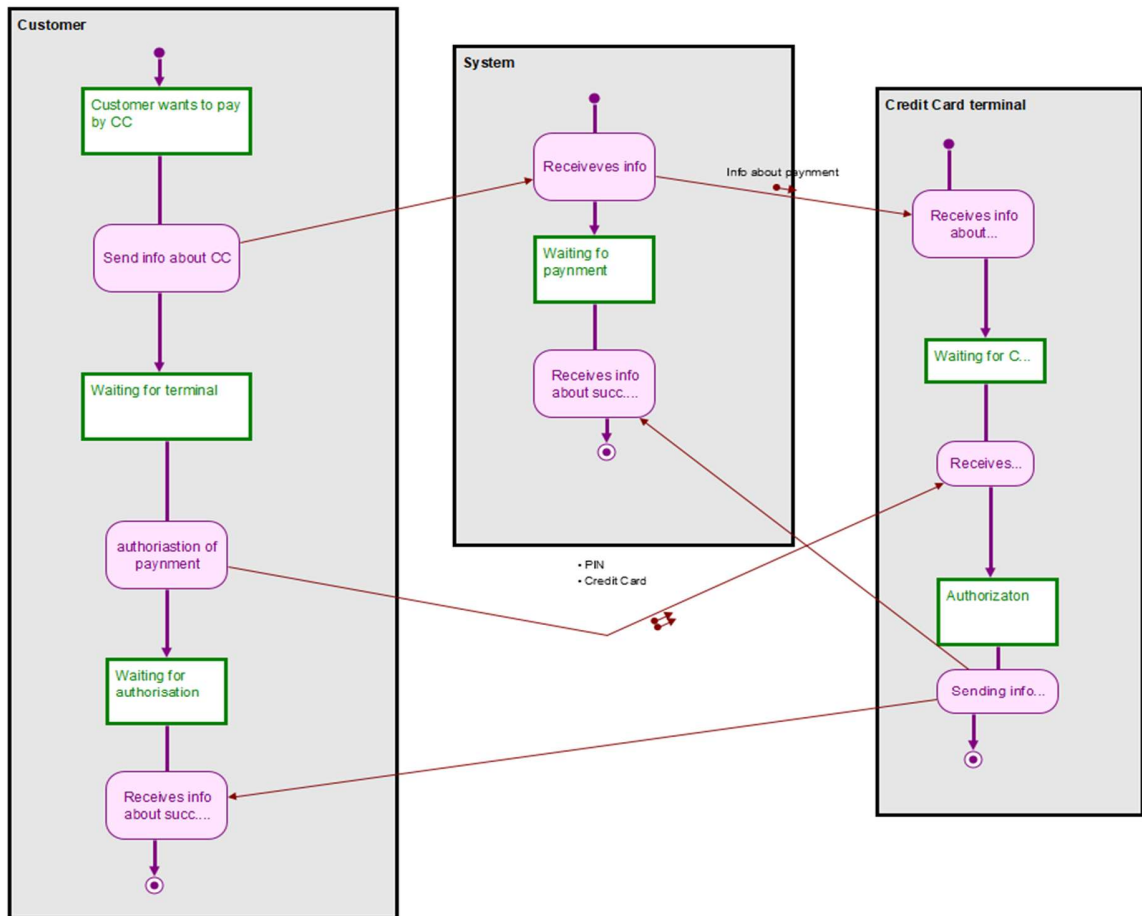


Figure 9 Business diagram - Pay by credit card

3.3 Draw.io



Figure 10 draw.io

Draw.io is an open-source technology stack for building diagramming applications, and the world's most widely used browser-based end-user diagramming applications.

3.3.1 Powerful & Easy to Use

Create flowcharts, process diagrams, charts, UML, ER diagrams, network diagrams and much more. Use the power of the automatic layout function, create own custom shape libraries or use a large collection of shape libraries that offer hundreds of visual elements. Work with metadata to pack more useful information into diagrams.

draw.io is a diagramming application that allows users to create and share diagrams within a web browser. In this application, there are possibilities to create and edit a variety of diagrams including flowcharts, charts, process diagrams, ER diagrams, UML, network diagrams, and more.

BPMN diagrams help to present complex processes and workflows, useful for onboarding new employees, or when a business process needs to be audited. Mind maps help visually brainstorm concepts or problems by organizing information into related topics, useful for making decisions or preparing a complex document.

3.4 UML (Online shopping)

Draw.io can help to explain and show all steps of online shopping, this part shows all process of online shopping and how it works, UML can help us explain all the steps. Here is the list of class names and their description as a data dictionary table.

Data Dictionary

Class Name	Description
User	User class Contains attributes and methods in the system. This class is a base class that inherits other classes of user-class properties and methods.
Customer	This class keeps all information about the system's permanent customers. The person who visits the site to buy the products you need and prepare them...
Transformer	This class keeps all the information about the courier. After purchasing via the internet by the customer and storing the customer's address information, the ordered goods are delivered to the shipper to enable the company to take the necessary steps to send the products.
Employee	This class keeps info about the staff of the system.
Manager	This class maintains all the management information. The manager is responsible for the entire site and has the task of overseeing the system, overseeing the work of employees, supervising sales affairs, making changes, and managing employees...
News	News helps the user to make a purchasing decision news stories have the potential to provide useful supporting information. The manager can write latest the latest news on the website
Address	In this class, we are describing the shipping address of the place where the product needs to be delivered.
Product	It explains the name of the product, its identity and the price of the product (currency).
Orders	This class tells about the order, the customer and the product
Payment	Payment is the operation where the customers pay for their purchase in the online shopping process.
Receipt	It explains about the number, the name of the product and the type of currency used.
Type	It describes the name, the gender that relates to the products.
Category	It describes the name and gender.

3.4.1 CLASS MODEL

A class diagram is a tool for accurate and easy expression of an object model. There are two types of object diagrams in UML: class diagram, in general, an instance diagram. In this project, instance diagrams will not be used. A class diagram consists of classes and associates. Classes contain the class name, attributes, and operations, while links connect those classes and express their cardinality. List of classes:

- User
- Customer
- Transformer
- Employee
- Manager
- News
- Address
- Product
- Orders
- Payment
- Receipt
- Type
- Category

Each of them has its own attributes. Customer, Transformer, Employee and Manager Classes are inherited from User Class.

Relationships are as follows: One customer can make zero or many orders. One order can “order” one or many products. One product can have one or many *Types*. One customer can have *addresses for shipment*. Any payment has one or two receipts. One manager can write one or many news on the website and has one or many employees. Each transformer brings one or many orders. The class diagram for Online-Shopping purchasing system can be seen in *Figure 11*.

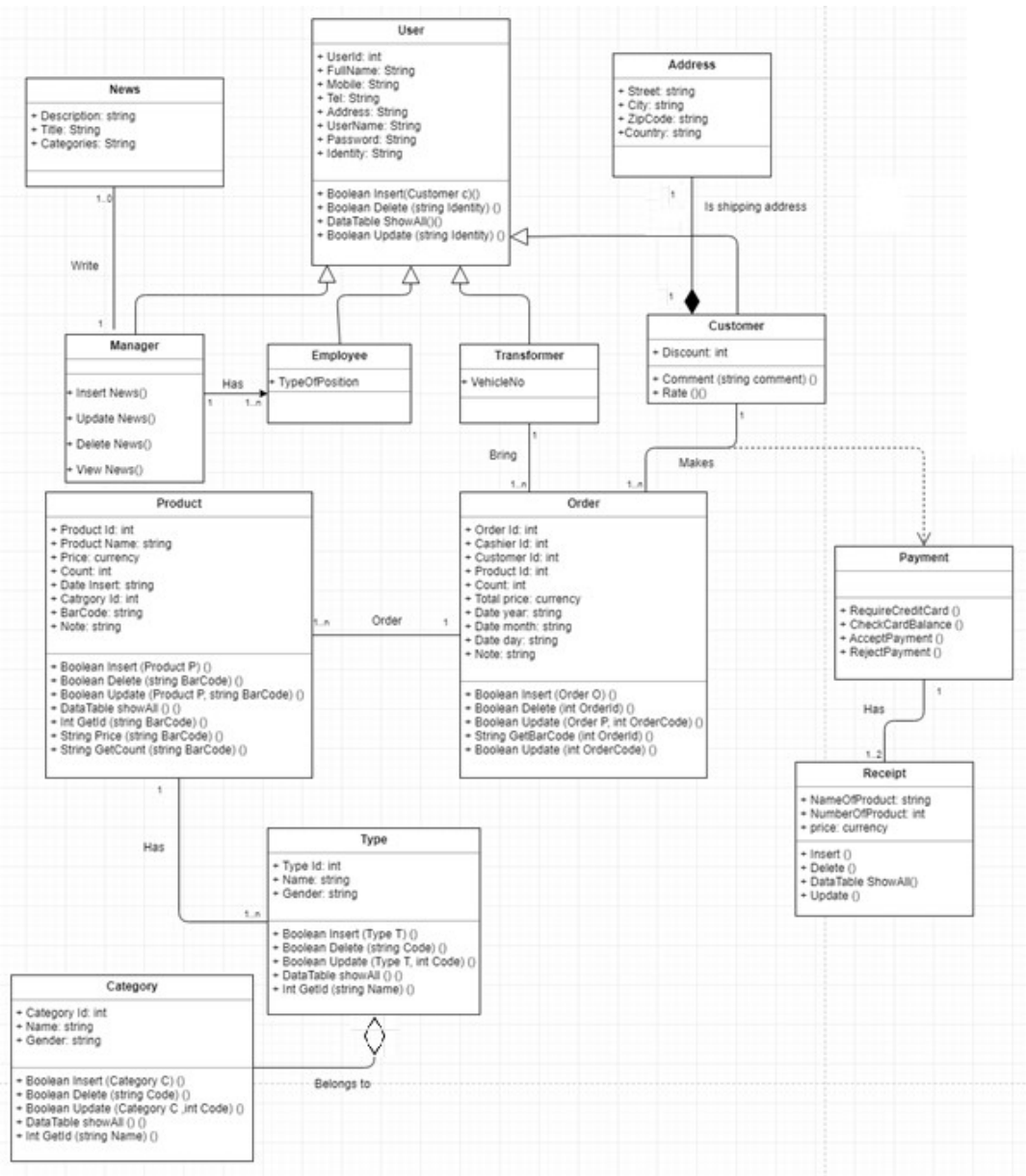


Figure 11 Class diagram for online shopping system

3.4.2 Dynamic Model

3.4.2.1 STATE DIAGRAM

State diagram (SD) shows the relationship between events and states and is one of the main concepts of dynamic modeling. When an event occurs, the next state depends on this event and a present state. Events and states are in a dual relationship: an event separates two states and a state separates two events. The state is often characterized by the fact that some attribute satisfies a certain condition. Change of a state caused by an event is transition. Activity is performed as a response to an event received in a state. Activities are performed on a transition, at the beginning, end or inside a state. The state diagram for Online-Shopping Purchasing System includes the customer's login to the website, the ordering and the payment of the purchase. The state diagram for Online-Shopping Purchasing System can be seen in *Figure 12*.

The dynamic model is a collection of the state diagram of all objects, which mutually affect each other using shared events. The state is an abstraction of attribute values and object links. In defining a state, we ignore attributes, which do not affect the behavior of the object. The state defines a response of the object at the input event. The state is often characterized by the fact that some attributes satisfy a certain condition.

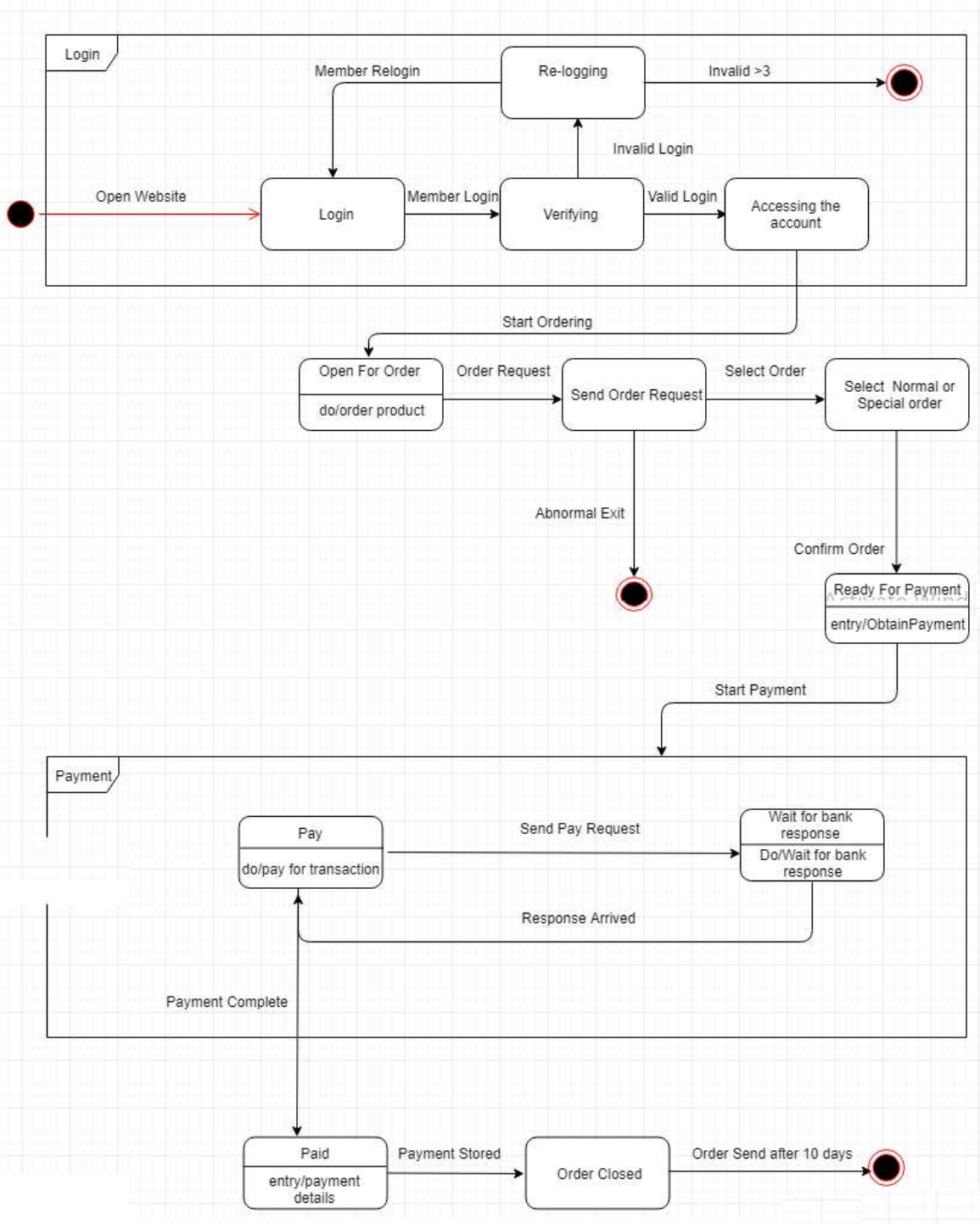


Figure 12 STATE DIAGRAM for online shopping

3.5 Interaction Model

3.5.1 Use Case Diagram

Use case diagram is used to express system roles (Actors) and different scenarios those actors can perform within the system.

In our case we have five actors:

- User
- Customer
- Manager
- Transformer
- Employee

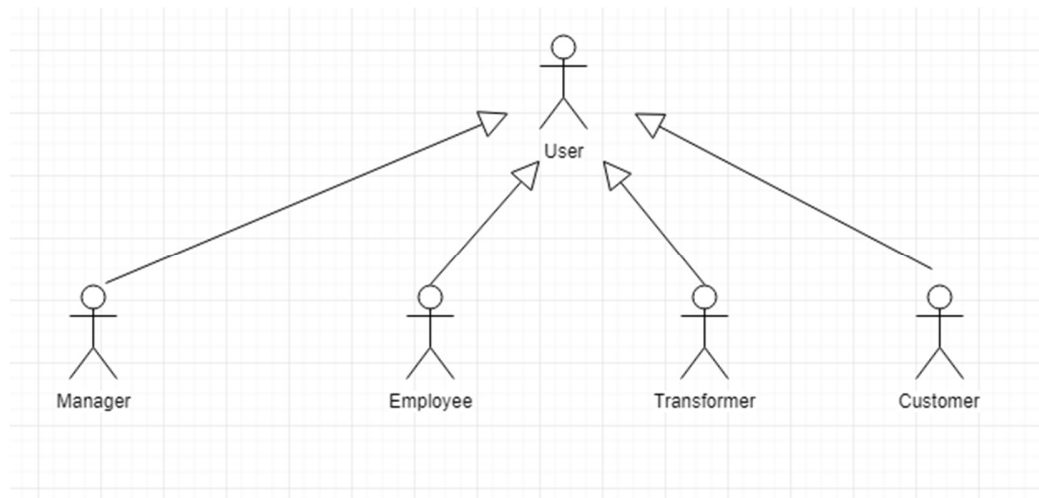


Figure 13 USE CASE DIAGRAM - All actors in the online shopping system

Manager, Employee, Transformer and Customer actors are inherited from User actor.

Registered Customer - have all rights to perform the main scenarios of the application like Login, Ordering or Payment for goods.

Use Case diagram is shown in *Figure 14*

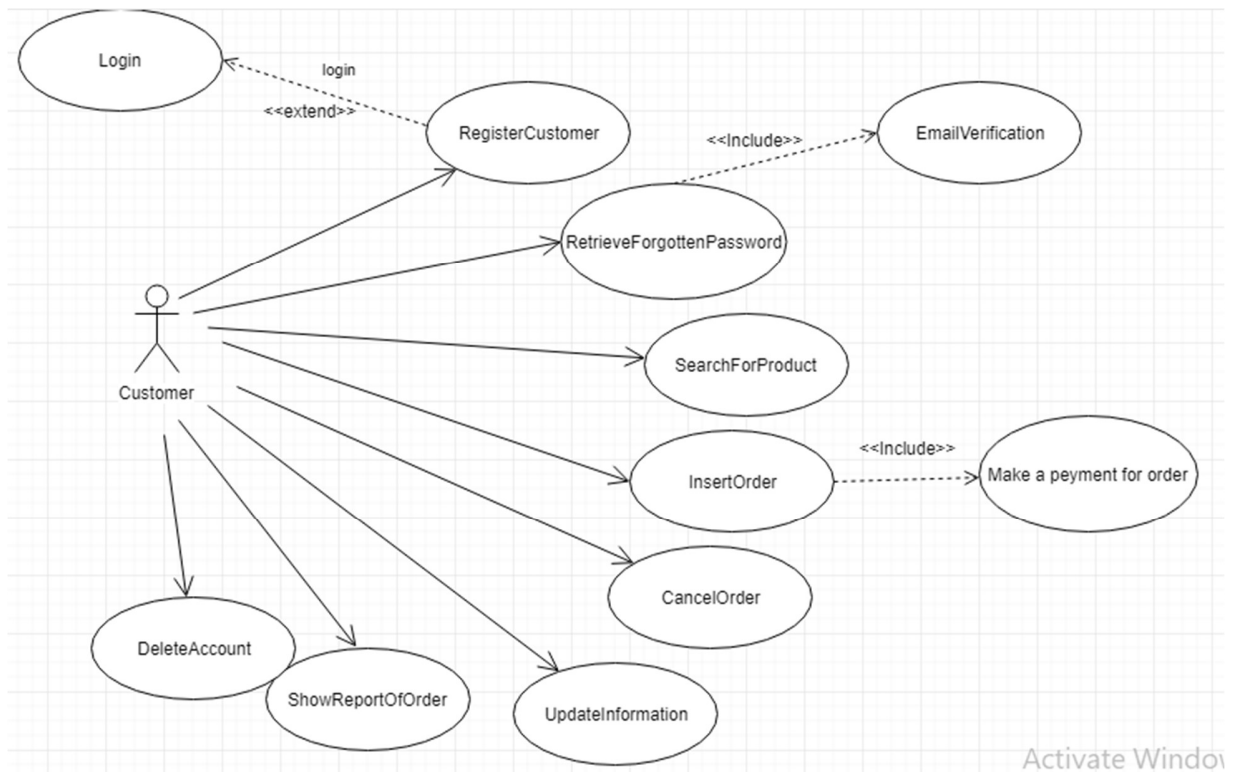


Figure 14 USE CASE DIAGRAM - Customers

Manager's scenarios are the most complex including managing Employees, Reports, Orders, and products as well as application configuration.

Use Case diagram is shown in *Figure 15*

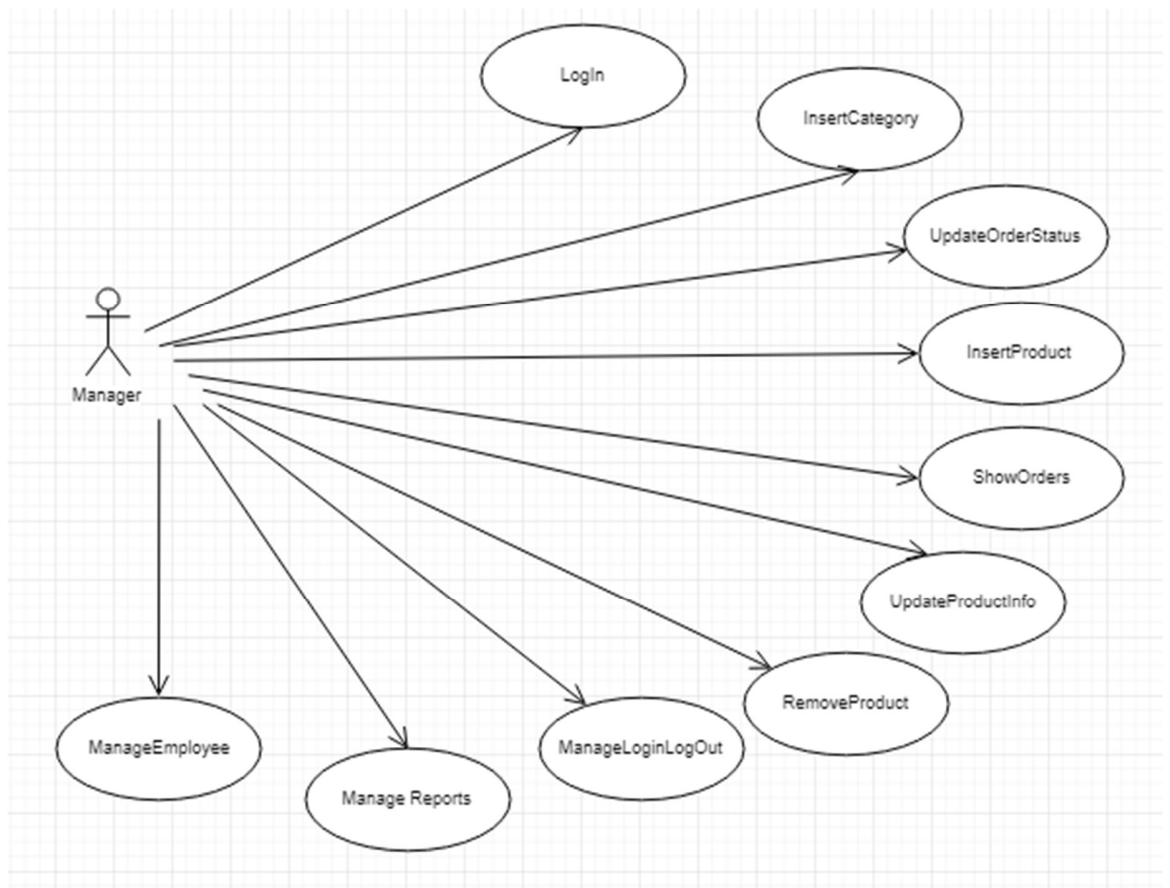


Figure 15 USE CASE DIAGRAM - Manager

Employee's scenarios are including Request, Edit, Cancel Orders and Payment in online-shopping system.

Use Case diagram is shown on *Figure 16*

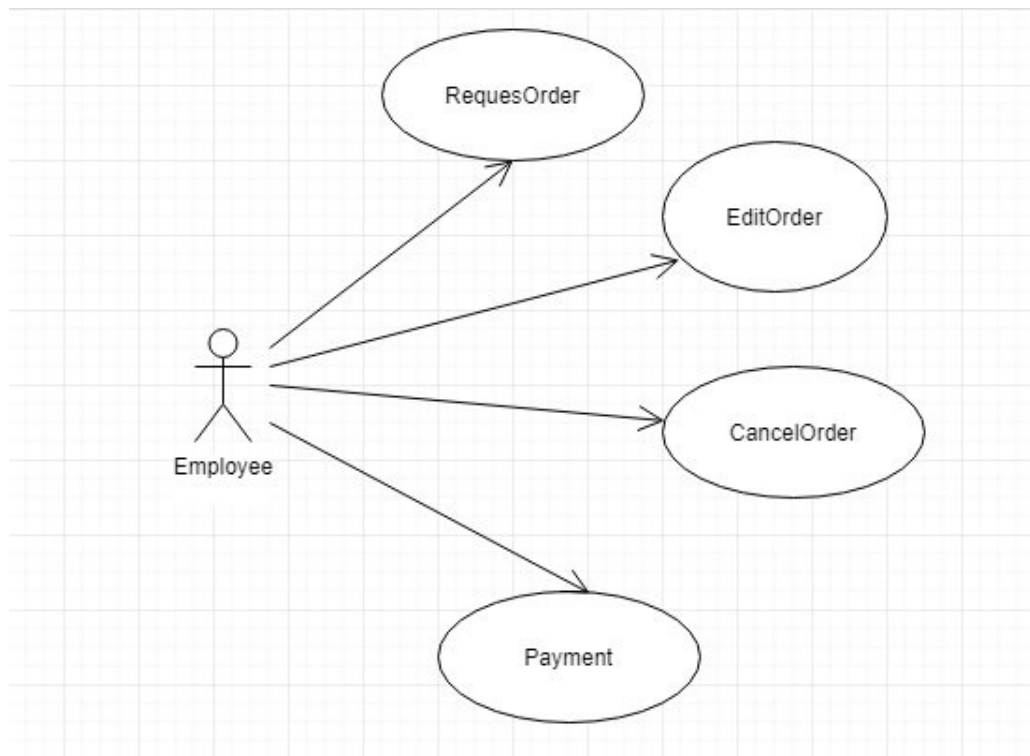


Figure 16 USE CASE DIAGRAM - Employee

Transformer's scenarios are including getting/Deliver Orders and Get money from customers in the online-shopping system.

Use Case diagram is shown in *Figure 17*

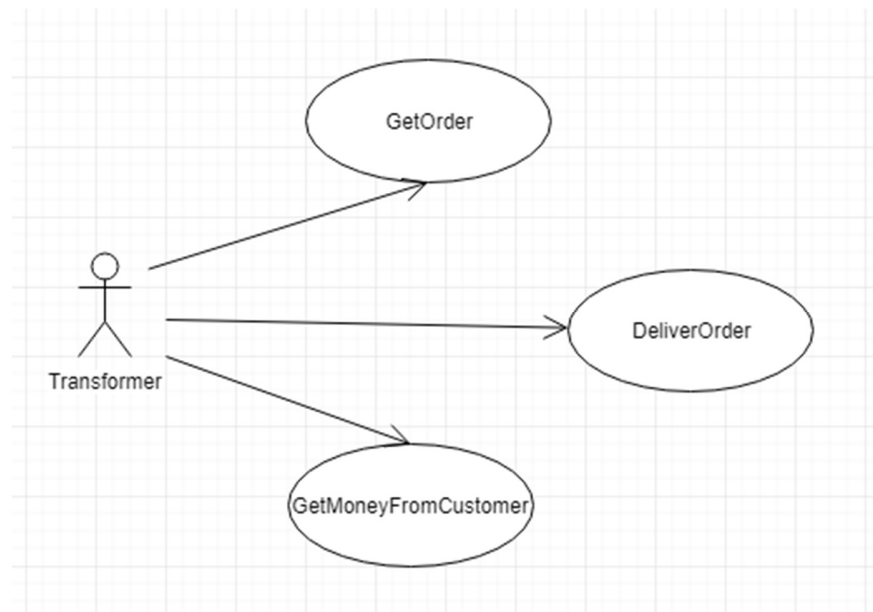


Figure 17 USE CASE DIAGRAM - Transformer

3.5.2 Sequence Diagram

The sequence diagram is a part of the interaction model. It is a more detailed way of describing an exchange of messages in time within a set of objects. It is suitable to describe sequences of behavior from the perspective of a system-user. Sequence diagrams show participants of interaction and messages among them. Each actor and a system is represented by a column – a vertical bar called lifeline and each message is depicted by a horizontal arrow from a sender to a receiver of a message. Time progresses from the top towards down.

The sequence diagram in *Figure 18* captures the scenario of a password change. The user has to insert the old and the new password. Because all passwords in the system are stored as a hash, the User Controller must send a raw password to the Password Encoder to get the hash.

If there is a match, the User Controller hashes the new password and saves it. In the end, the result of this procedure is sent to the user.

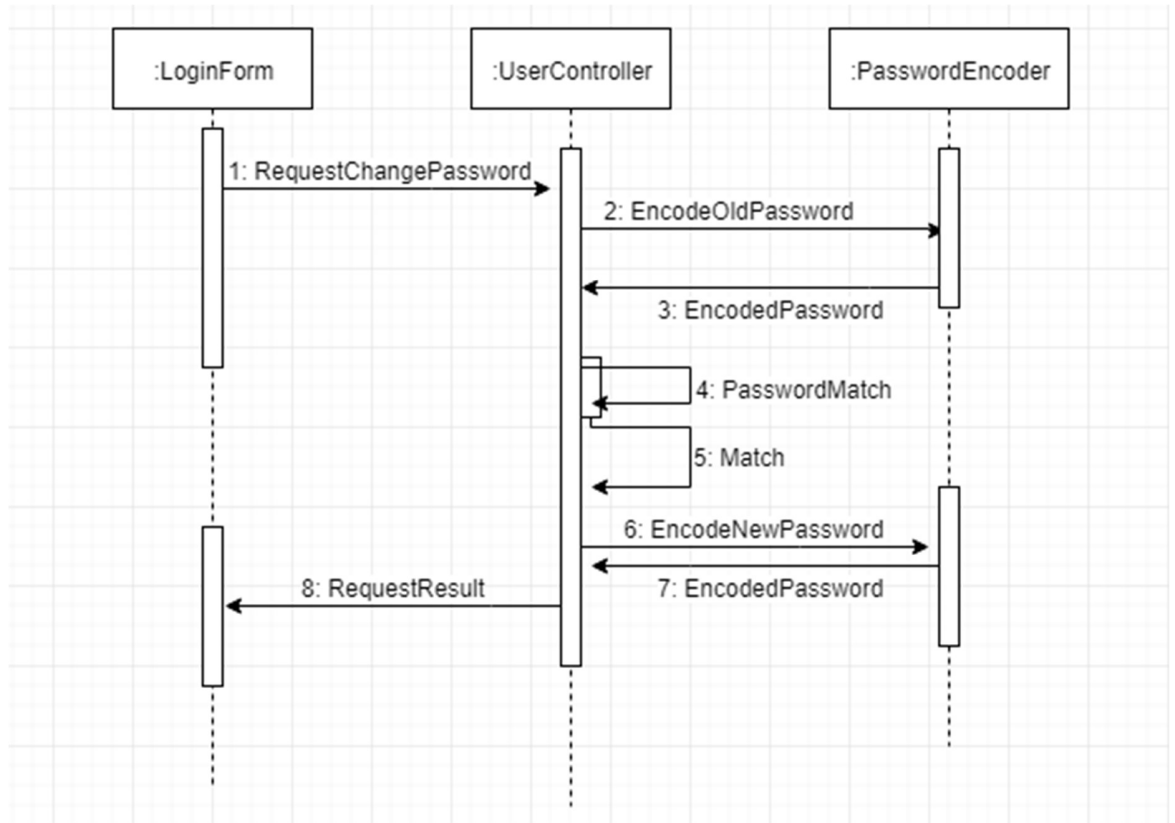


Figure 18 SEQUENCE DIAGRAM - Login

The sequence diagram in *Figure 19* captures the scenario of Create Category of the product. The user (Manager) must log in request on the website. The UserHandler checks the username and password in the system. The manager can insert product and create a new category of products after inserted products, the system saves all the products in the Database.

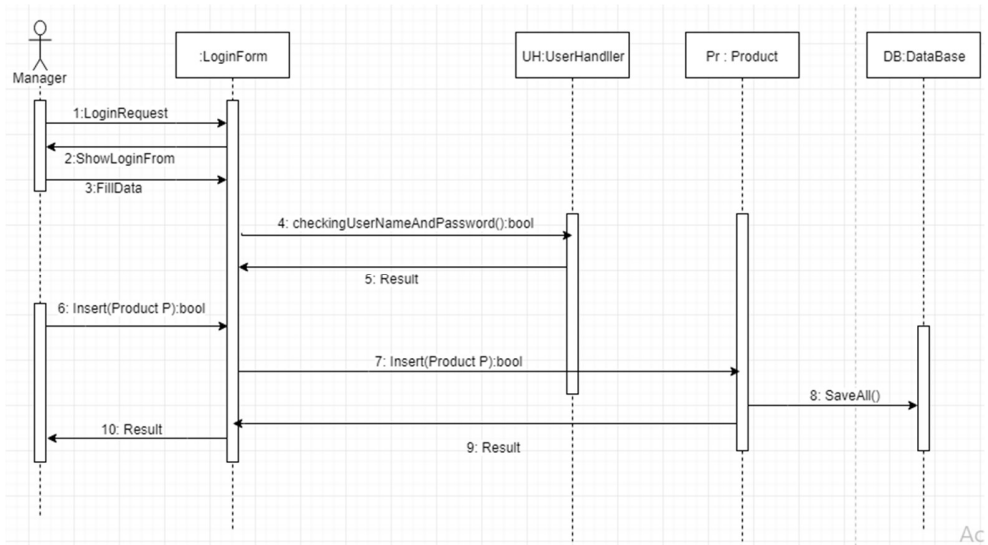


Figure 19 SEQUENCE DIAGRAM - Ordering products and save in database

3.6 Activity Diagram

The activity diagram is also a part of interaction modeling. It shows a detailed view of the control flow between individual steps of consumption. It shows both, data flow and control flow. Order Activity diagram for Online-Shopping Purchasing System is divided into two “lanes”, including customer and Online-Shopping purchasing Systems. It describes a process from the moment when customers check for a product, till the moment when Online-Shopping ships the product to the customer. The order Activity diagram for Online-Shopping Purchasing System can be seen in *Figure 20*. Activity is performed as a response to an event received in a state. Activities are performed on a transition, at the beginning, end or inside a state. If there is an activity on input transition, this activity is executed first. Input activities are indicated inside a state symbol and have a form entry/activity name. Output activities are indicated inside a state symbol and have a form exit/activity name. If a state is terminated by an output transition, the output activities are executed first. The object can perform an Activity by sending a signal to another object it is object interaction. An activity diagram shows steps for building a complex process, for example, algorithm or workflow. An activity diagram shows a control flow the same as a

flow chart. Terminating of the activity is terminating the event and it usually means that the next activity can begin. The activity can be decomposed into finer activities. If the activity has several successors, each branch should be labeled by a condition in square brackets. Moreover, the conditions are tested after completing the activity.

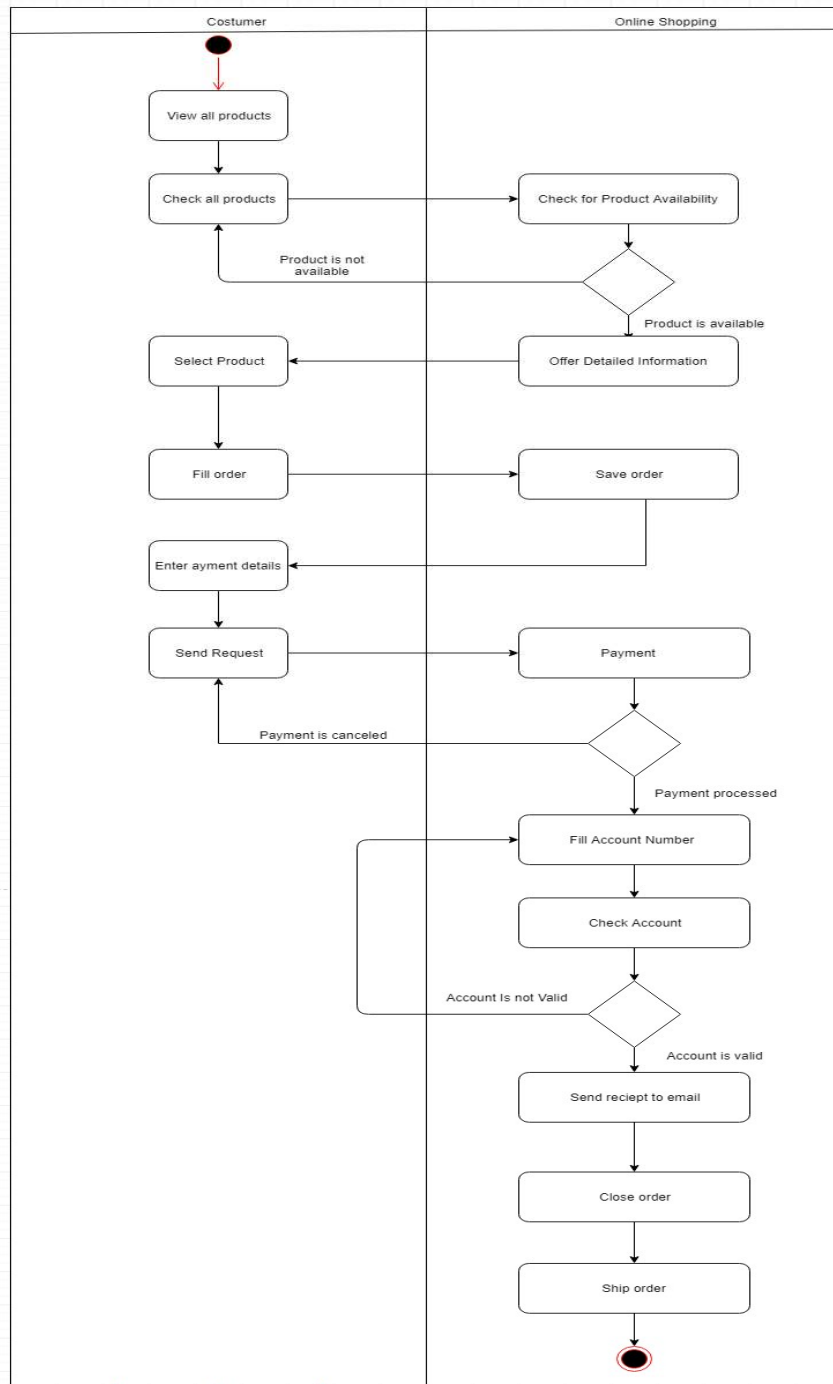


Figure 20 ACTIVITY DIAGRAM - Ordering Products

Login Activity diagram for Online-Shopping Purchasing System includes one part: The user wants to login into the system. It describes a process from the moment when the user logs in into the system, until the moment when the login ends. The login Activity diagram for Online-Shopping Purchasing System can be seen in *Figure 21*.

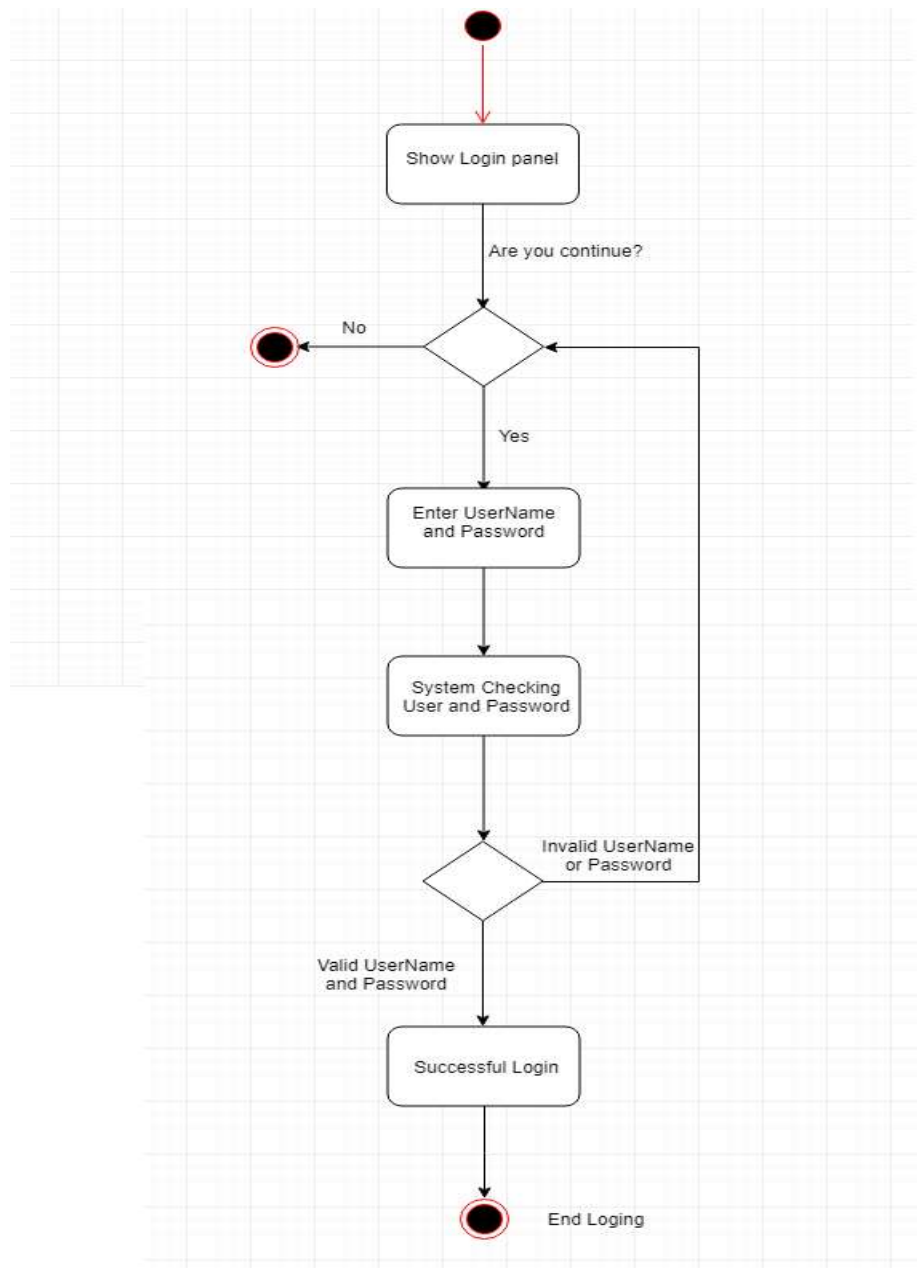


Figure 21 ACTIVITY DIAGRAM - Login in system

3.7 SPORTISIMO



Figure 22 SPORTISIMO

SPORTISIMO HAS MORE THAN 10 MILLION PRODUCTS IN STOCK

With over 10 million products in stock, we are far ahead of our competition. For each product, you will find the necessary information to help you make decisions. If you still have any additional questions, we are at your disposal by e-mail or telephone.

MORE THAN 16 YEARS COOPERATION WITH SPORTS

SPORTISIMO helps athletes achieve higher goals. It does not distinguish between professional and amateur athletes. It is important for them to help you on your way to your goals and to motivate you. Our family has grown a lot during that time, and we will be glad if you become a part of it!

ONLINE AND PHONE CUSTOMER SUPPORT

Have a question? Not sure how to order? SPORTISIMO helps customers! With a clear conscience, it can promise you that it will be on your phone or e-mail until you are satisfied.

SPORTISIMO HAS THE WIDEST CHOICE OF SPORTS GOODS FROM MOST KNOWN BRANDS

In our offer, customers will find goods from the world's most famous brands. SPORTISIMO pride itself on quality, but at the same time, we need to have the goods at affordable prices. Customers will not find any imitations - everything is original!

3.8 Keboola

3.8.1 Keboola is a secure data innovation platform for your data needs

Keboola puts together computer engineering and data analytics on a single platform that can be used by anyone. Managed storage, data center, user provisioning, process automation by one click-all rolled up into one platform. Solutions focused on machine learning include predictive help, scaffolding one click and much more! No installation required, one button, and you're ready to go!

3.8.2 All the tools you need in one place

At Keboola we have fathomed the acquirement and organization bad dream that organizations as a rule face while looking for and actualizing new tools. Presently you can get moment access to many of your preferred information investigation instruments, servers, and information distribution centers across the board place in our application store.

3.8.3 Collaboration breeds success

Keboola permits individuals to cooperate on similar code in a similar pipeline. Your group can expand on one another's achievements progressively. Since everything is formed, nothing gets lost. We additionally make joining simpler by permitting your instruments to run in equal, so every client has their condition with access to all the significant information they need in only a single tick.

3.9 Data preparation in Keboola for the SPORTISIMO company

In this part, start preparing the data for daily traffic and daily transactions, there are two different parts for Traffic, the first one is online traffic and the second part is offline traffic. And there are two different parts for Transactions. The first one is the online transaction and the second part is an offline transaction. Then combine these two traffics and transactions and store the data daily in the “DataPreparation” bucket as you can see on the picture.

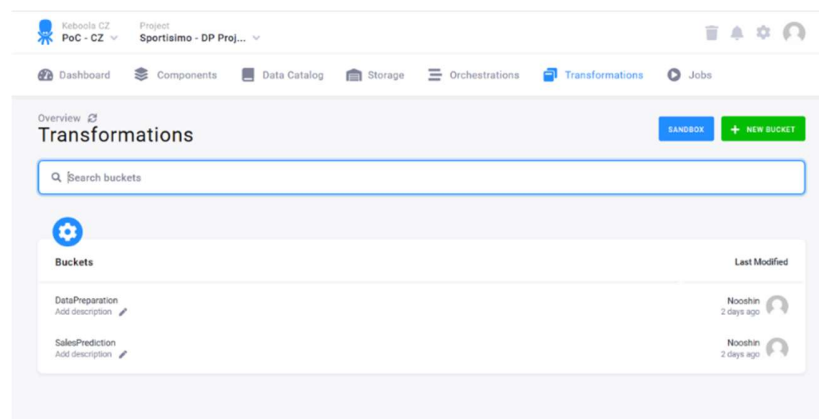


Figure 23 Data preparation in Keboola

In the part of the transaction, start creating different parts for storing data of Customers, Products, Traffics, and Transactions separately.

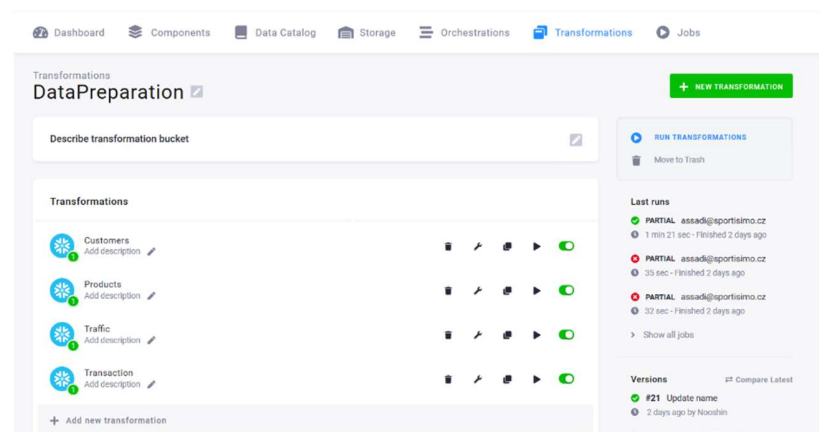


Figure 24 Data preparation in Keboola - Transactions

3.9.1 Daily Traffic

Steps on how to prepare the daily traffic. How to combine and store data in Keboola.

The screenshot displays the Keboola interface for configuring a transformation named "Traffic". The main configuration area includes:

- Describe transformation:** A text area for describing the transformation.
- Requires:** A dropdown menu to add required transformations.
- Dependent Transformations:** A section indicating no dependent transformations are present.
- Input Mapping:** A table showing two input transformations: "in.c-keboola-ex-storage-110747405.OfflineTraffic" mapped to "OfflineTraffic" and "in.c-keboola-ex-storage-110747405.OnlineTraffic" mapped to "OnlineTraffic".
- Output Mapping:** A table showing one output transformation: "DailyTraffic" mapped to "out.c-diplomathesis.DailyTraffic".
- Queries:** A code editor containing SQL queries:



```
1 CREATE OR REPLACE TABLE "OfflineDailyTraffic" AS (  
2 select  
3 "BranchNumber",  
4 "SnapDate" AS "Date",  
5 -- "SnapPeriodStartTime",  
6 -- "SnapPeriodEndTime",  
7 SUM("CountIncomingCustomers") AS "InUsers",  
8 SUM("CountOutgoingCustomers") AS "OutUsers",  
9 (SUM("CountIncomingCustomers")+SUM("CountOutgoingCustomers"))/2 as "TotalUsers"  
10 from "OfflineTraffic"  
11 GROUP BY 1,2)  
12 ;  
13  
14 CREATE OR REPLACE TABLE "OnlineDailyTraffic" AS (  
15 select  
16 "id",  
17 "idProfile",  
18 "date",  
19 "medium",  
20 "source",  
21 "campaign",  
22 "users",  
23 "newUsers",  
24 "pageviews"  
25 from "OnlineTraffic")  
26 ;  
27  
28 CREATE OR REPLACE TABLE "DailyTraffic" AS (  
29 select *  
30 from "OnlineDailyTraffic"  
31 );
```

On the right side, there are three panels:

- RUN TRANSFORMATION:** Shows the transformation is "Enabled (click to disable)". It includes links for Overview, Edit Phase, New Sandbox, SQLDep (BETA), Validate (BETA), Copy transformation, Delete transformation, and Documentation.
- Last runs:** A list of recent runs with status indicators (PARTIAL, SUCCESS) and completion times (e.g., "52 sec - Finished 2 days ago").
- Versions:** A list of version changes with status indicators and timestamps (e.g., "#8 Change Queries in Traffic 2 days ago by Nooshin").

Figure 25 Data preparation in Keboola - Daily Traffic

The final query for creating daily traffic that is saved in Keboola.

```
Queries  SAVED 
```

```
1 CREATE OR REPLACE TABLE "OfflineDailyTraffic" AS (  
2 select  
3 "BranchNumber",  
4 "SnapDate" AS "Date",  
5 --"SnapPeriodStartTime",  
6 --"SnapPeriodEndTime",  
7 SUM("CountIncomingCustomers") AS "inUsers",  
8 SUM("CountOutgoingCustomers") AS "outUsers",  
9 (SUM("CountIncomingCustomers")+SUM("CountOutgoingCustomers"))/2 as "TotalUsers"  
10 from "OfflineTraffic"  
11 GROUP BY 1,2)  
12 ;  
13  
14 CREATE OR REPLACE TABLE "OnlineDailyTraffic" AS (  
15 select  
16 "id",  
17 "idProfile",  
18 "date",  
19 "medium",  
20 "source",  
21 "campaign",  
22 "users",  
23 "newUsers",  
24 "pageviews"  
25 from "OnlineTraffic")  
26 ;  
27  
28 CREATE OR REPLACE TABLE "DailyTraffic" AS (  
29 select *  
30 from "OnlineDailyTraffic"  
31 );  
32  
33 INSERT INTO "DailyTraffic" ("id","idProfile","date","users")  
34 SELECT "BranchNumber"||"Date" AS "id",  
35 "BranchNumber","Date",  
36 "TotalUsers" AS "Users"  
37 FROM "OfflineDailyTraffic";
```

Figure 26 Data preparation in Keboola - Final query

Steps of creating a query for combining two online traffic and offline traffic tables.

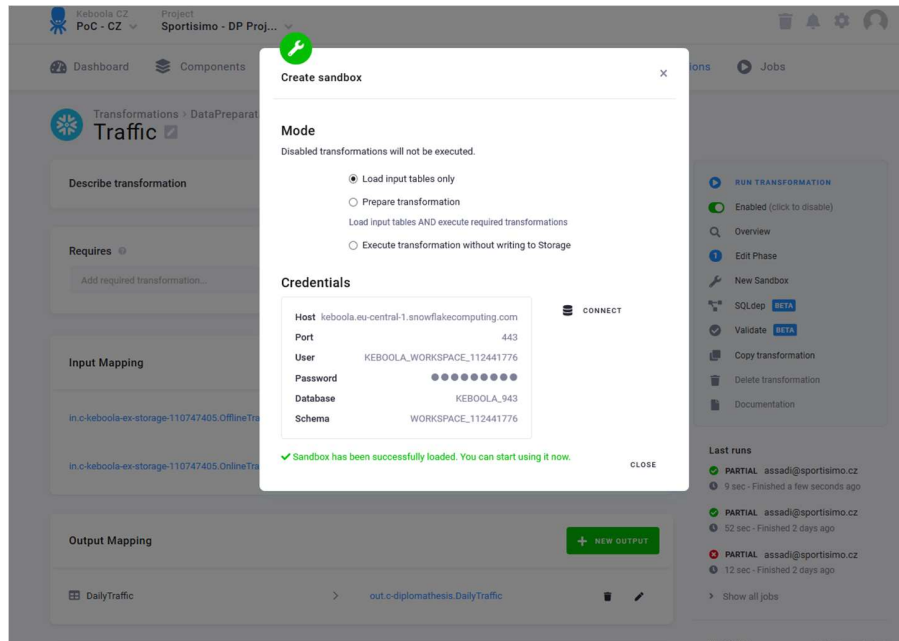


Figure 27 Data preparation in Keboola - Create a sandbox for writing query

1. Step 1: Log in to Snowflake

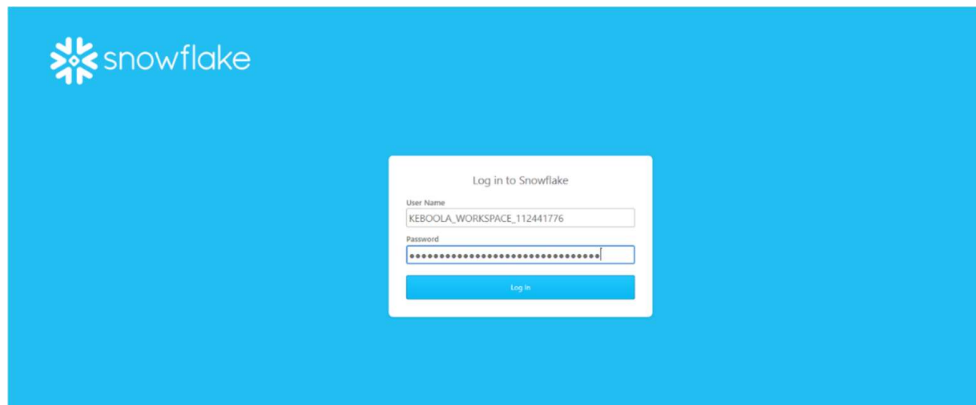


Figure 28 Data preparation in Keboola - Login into Snowflake

- Step 2: Start writing a query in this step. Creating OfflineDailyTraffic with specific columns which are BranchNumber, Date, Sum of InUsers, Sum of OutUsers and Totals of Users or Customers.

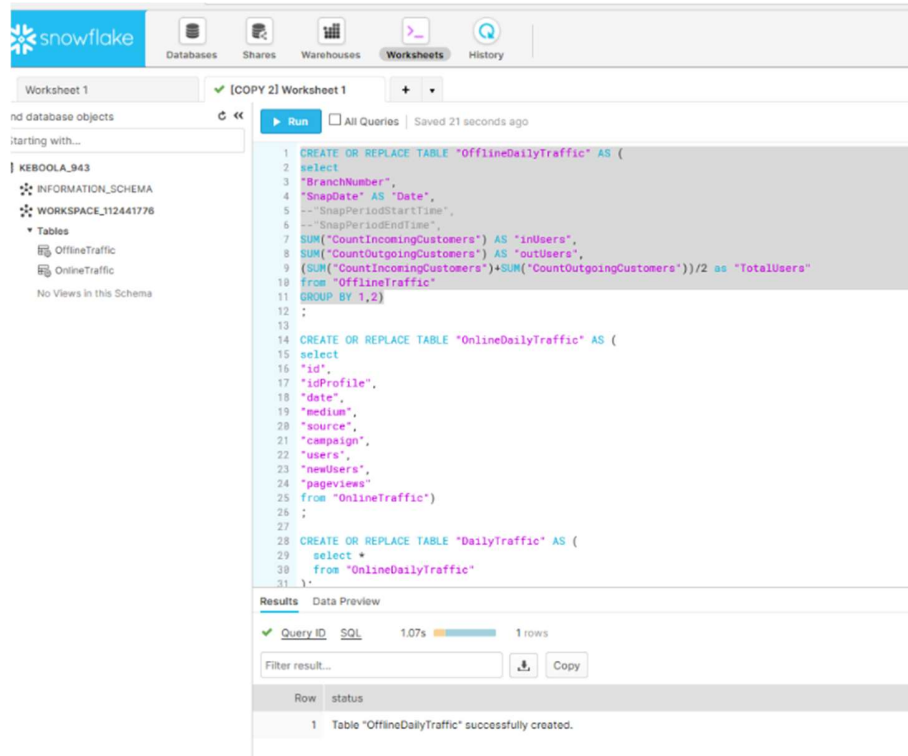


Figure 29 Data preparation in Keboola - Offline daily traffic's table

In this part, start creating OnlineDailyTraffic with specific columns.

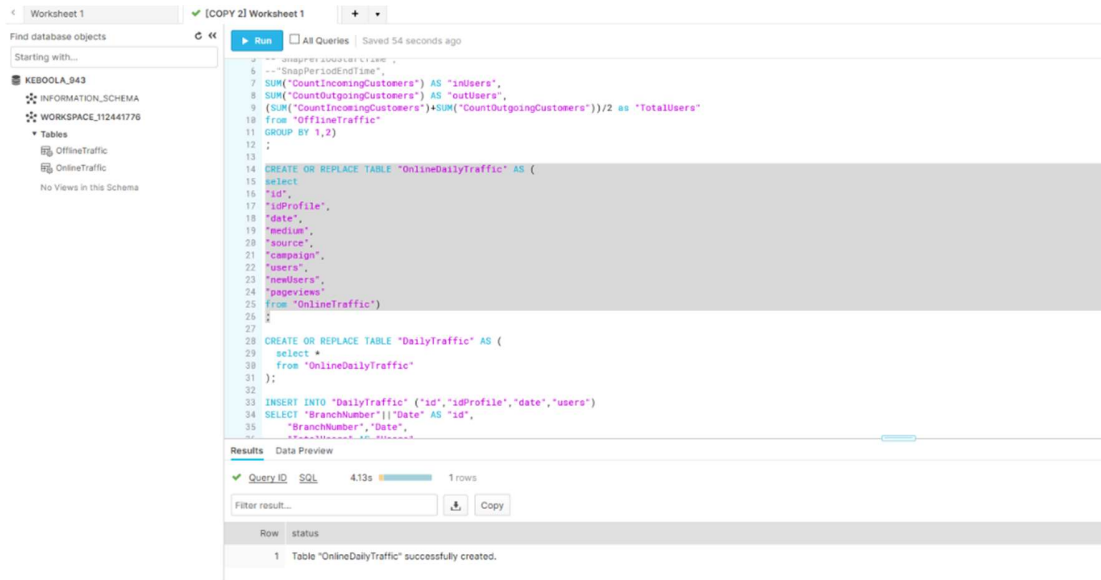


Figure 30 Data preparation in Keboola - Online daily traffic's table

After creating OfflineDailyTraffic and OnlineDailyTraffic, we can combine two tables and store all data in the DailyTraffic table.

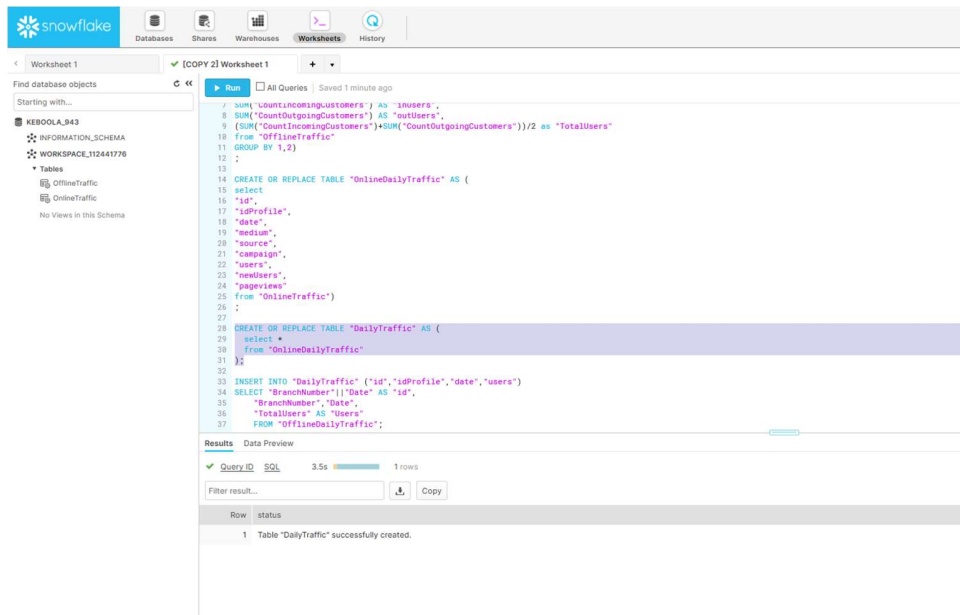


Figure 31 Data preparation in Keboola - Daily traffic's table

The screenshot shows the Snowflake SQL editor with the following query:

```

1: sum(CountDistinctCustomers) as 'newUsers',
2: sum(CountDistinctCustomers) as 'outUsers'
3: (sum(CountDistinctCustomers)+sum(CountDistinctCustomers))/2 as 'TotalUsers'
4: from 'OfflineTraffic'
5: GROUP BY 1,2
6:
7:
8:
9: CREATE OR REPLACE TABLE 'OnlineDailyTraffic' AS (
10: select
11: "id",
12: "idProfile",
13: "date",
14: "medium",
15: "source",
16: "campaign",
17: "users",
18: "newUsers",
19: "pageviews"
20: from 'OfflineTraffic')
21:
22:
23: CREATE OR REPLACE TABLE 'DailyTraffic' AS (
24: select *
25: from 'OnlineDailyTraffic'
26: )
27:
28:
29: INSERT INTO 'DailyTraffic' ('id','idProfile','date','users')
30: SELECT 'BranchNumber'||Date as 'id',
31: 'BranchNumber'||Date,
32: 'TotalUsers' as 'Users'
33: FROM 'OfflineDailyTraffic';

```

Below the query, a 'Data Preview' table is shown with the following data:

Row	id	idProfile	date	medium	source	campaign	users	newUsers	pageviews
1	382019-10-22	38	2019-10-22				602		
2	792019-08-03	79	2019-08-03				1058.5		
3	302019-05-15	30	2019-05-15				1619.5		
4	1012019-07-23	101	2019-07-23				350.5		
5	272019-09-03	27	2019-09-03				646.5		
6	382019-03-31	38	2019-03-31				549.5		
7	842019-03-14	84	2019-03-14				271.5		

Figure 32 Data preparation in Keboola - Daily traffic's data

- Step 3: After creating the table for daily traffic, we should Run Transformation for the final part and store the data.

The screenshot shows a dropdown menu for 'Run Transformation' with the following options:

- RUN TRANSFORMATION** (active)
- Enabled (click to disable)
- Overview
- 1 Edit Phase
- New Sandbox
- SQLdep **BETA**
- Validate **BETA**
- Copy transformation
- Delete transformation
- Documentation

Figure 33 Data preparation in Keboola - Run transformation

3.9.2 Daily Transactions



Explain all the Steps to preparing daily transactions. As you can see, these steps are mostly the same as the daily traffic

The screenshot displays the Keboola Data Preparation interface for a transformation named "Transaction". The interface is organized into several sections:

- Describe transformation:** A text area for describing the transformation.
- Requires:** A dropdown menu to add required transformations.
- Dependent Transformations:** A section indicating no dependent transformations are present.
- Input Mapping:** A table showing two input sources: "in-c-keboola-ex-storage-110747405.OfflineTransactions" mapped to "OfflineTransactions" and "in-c-keboola-ex-storage-110747405.OnlineTransactions" mapped to "OnlineTransactions".
- Output Mapping:** A table showing one output destination: "DailyTransactions" mapped to "out-c-diplomathesis.DailyTransactions".
- Queries:** A code editor containing a SQL query to create or replace a table named "DailyTransactions" with columns from "OnlineTransactions".
- Right Sidebar:** Contains a "RUN TRANSFORMATION" button, a status indicator "Enabled", a menu with options like "Overview", "Edit Phase", "New Sandbox", "SQLdep", "Validate", "Copy transformation", and "Delete transformation", and a "Last runs" section showing three partial runs.

Figure 34 Data preparation in Keboola - Daily transaction

The final query for creating daily transactions that are saved in Keboola.

```
Queries  SAVED 
```

```
1 CREATE OR REPLACE TABLE "DailyTransactions" AS (  
2 select  
3 "TransactionID",  
4 "Date",  
5 "idProfile",  
6 "Domain",  
7 "DomainID",  
8 "Country",  
9 "Affiliation",  
10 "DeviceCategory",  
11 "Channel",  
12 "Medium",  
13 "Source",  
14 "Campaign",  
15 "Keyword",  
16 "LandingPage",  
17 "CurrencyISOCode",  
18 "ExchangeRate",  
19 "GALocalTransactionRevenueWithTax",  
20 "GALocalTransactionRevenueNoTax",  
21 "GACZKTransactionRevenueWithTax",  
22 "GACZKTransactionRevenueNoTax",  
23 "DWHVATRate",  
24 "ItemQuantity",  
25 "UserID",  
26 "TrafficId"  
27 from "OnlineTransactions")  
28 ;  
29  
30 INSERT INTO "DailyTransactions"  
31 ("TransactionID","Date","Domain","DomainID","CurrencyISOCode","DWHVATRate","GALocalTransactionRevenueNoTa  
x","GALocalTransactionRevenueWithTax","GACZKTransactionRevenueNoTax","GACZKTransactionRevenueWithTax","Us  
erID")  
32 SELECT  
33 "TransactionID","Date","Domain","DomainID","CurrencyISOCode",1-("DWHCZKPriceNoTax"  
/NULLIF("DWHCZKPriceWithTax",0)), "DWHLocalCurrencyPriceNoTax", "DWHLocalCurrencyPriceWithTax", "DWHCZKPrice  
NoTax", "DWHCZKPriceWithTax", "UserID"  
34 FROM "OfflineTransactions"  
35 ;
```

Figure 35 Data preparation in Keboola - Final query

Steps of creating a query for daily transactions.

1. Step 1: Log in to Snowflake.
2. Step 2: Start writing a query. Creating OfflineDailyTransaction and OnlineDailyTransaction with specific columns. In the final part, after creating OfflineDailyTransaction and OnlineDailyTransaction, we can combine the two tables and store all data in the DailyTransaction table and see all data.

The screenshot displays a Snowflake SQL query in a worksheet. The query uses a UNION ALL to combine data from 'OfflineTransactions' and 'OnlineTransactions' into a 'DailyTransactions' table. The query includes various columns such as TransactionID, Date, UserProfile, Domain, Country, Affiliation, DeviceCategory, Channel, Medium, Source, Campaign, Keyword, LandingPage, CurrencyISOCode, ExchangeRate, and GLocalTax. Below the query, the 'Results' section shows a preview of the data with 7 rows and 17 columns.

Row	TransactionID	Date	UserProfile	Domain	DomainID	Country	Affiliation	DeviceCategory	Channel	Medium	Source	Campaign	Keyword	LandingPage	CurrencyISOCode	ExchangeRate	GLocalTax
1	21238073	2019-07-02	13033466	sportsimo.cz	101	Czech Repu...	Rezervace	desktop	PPC	cpc	google	1-(cz)-ori-...	sportsimo	/ostupnost...	CZK	1	699
2	12709567	2019-11-26	100151435	sportsimo.c...	301	Global	sportsimo.c...	mobile	PPC	cpc	google	3-(uk)-pla-s...	brand+robu...	/shipping?p...	EUR	2704	74.48
3	12054437	2019-04-28	65188844	sportsimo.sk	201	Slovakia	sportsimo.sk	desktop	Mailing	email	newsletter	kampen-obc...	2016-apri-0...	/kampanelo...	EUR	27035	84.11
4	11890381	2019-02-21	13033466	sportsimo.cz	101	Czech Repu...	sportsimo.cz	tablet	Organic	organic	seznam	(not set)	sportsimo pl...	/mattoni-po...	CZK	1	2002
5	12086365	2019-05-10	65188844	sportsimo.sk	201	Slovakia	sportsimo.sk	desktop	CSE	referral	damske-plea...	(not set)	(not set)	/jadkaxi-3s...	EUR	2702	24.94
6	12430103	2019-09-19	65188844	sportsimo.sk	201	Slovakia	sportsimo.sk	mobile	PPC	cpc	google	4-(sk)-dsa-...	dynamic sea...	/panske-tric...	EUR	27025	20.94
7	12728325	2019-11-29	109843087	sportsimoro	501	Romania	sportsimoro	desktop	PPC	cpc	google	free_shipping...	luser vertica...	/jadkaxi-jen...	RON	5.991	171.88

Figure 36 Data preparation in Keboola - Writing Daily transaction query

3. Step 3: After creating a table for daily transactions, we should Run Transformation for the final part and store the data.

3.9.3 Structure of Traffic, Transactions, Products and Customers tables of SPORTISIMO

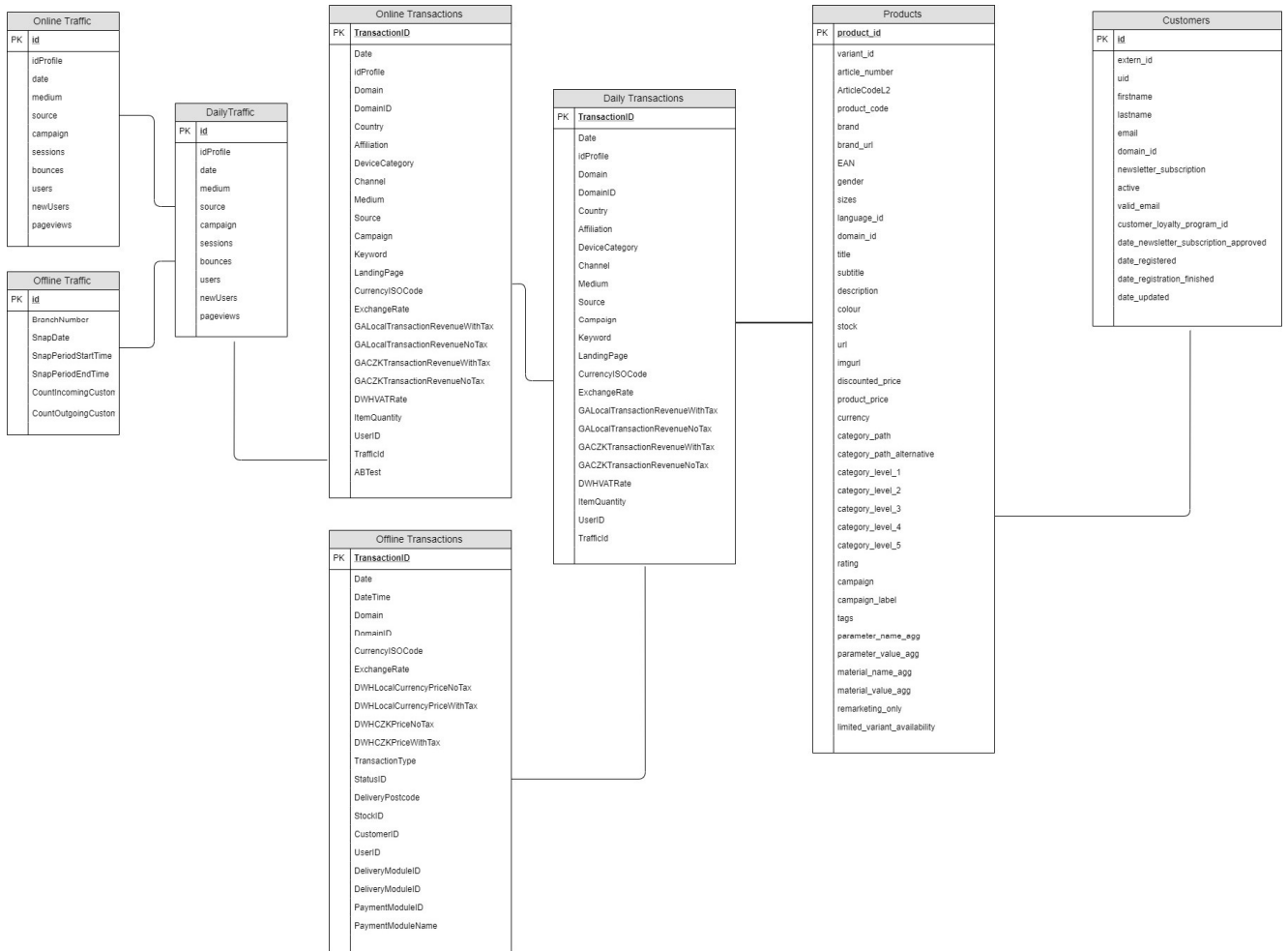


Figure 37 Structure of Traffics, Transactions, Products and Customers tables of SPORTISIMO

3.10 RStudio



Figure 38 RStudio

RStudio's main role is to make free and open-source programming for information science, logical research, and specialized correspondence. This permits anybody with access to a PC to take an interest openly in a worldwide economy that rewards information proficiency; improves the creation and utilization of information; and encourages coordinated effort and reproducible research in science, training, and industry

Maintainable Investment: The selection of the open-source information science programming makes interest for the business items, which empower groups to receive open-source information science programming at scale.

Demonstrated Commitment to Community: notwithstanding the help for information science instruction, RStudio widely bolsters the open-source information science network through the yearly RStudio:: conf, the RStudio Community site where countless open-source and expert clients discover help, Education assets for students and educators, and backing for industry-subsidized associations, for example, the R Consortium, NumFocus, and Ursa Labs. **Concentrate on Reproducible Data Science:** to keep up confidence in the center information science instruments, and permit answers for being reused, repurposed, shared and broadly utilized, information science results must be reproducible after some time and between various clients.

3.11 Sales forecast of 2016 in RStudio

Forecasting might be defined as a part of cognition theory relating to the future. There are three important methods of forecasting, these are event outcome forecasts, event timing forecasts, and time-series forecasts. We only use the time series forecast in this project. Time series forecast is a sequence of values usually recorded at equidistant time intervals. Time series analysis is used as an applied topic in all disciplines such as engineering, economics, geography, management, and so on. Past review of data to identify and determine the nature of a phenomenon is the first step in conceptual modeling that can be used to obtain a reliable prediction in the next step. Forecasting involves taking models to fit historical data and using them to predict future observations.

For forecasting sales, RStudio can help us. After storing the data in Keboola, we can extract the data from Keboola and start working on RStudio to forecast the upcoming few days. The data shows the sales of 2016.

3.11.1 Load Data

```
#####  
#Load Data  
#####  
alldata <- read.csv("C:/Users/assadi/Downloads/out.c-diplomathesis.DailyTransactions.csv")  
head(alldata)
```

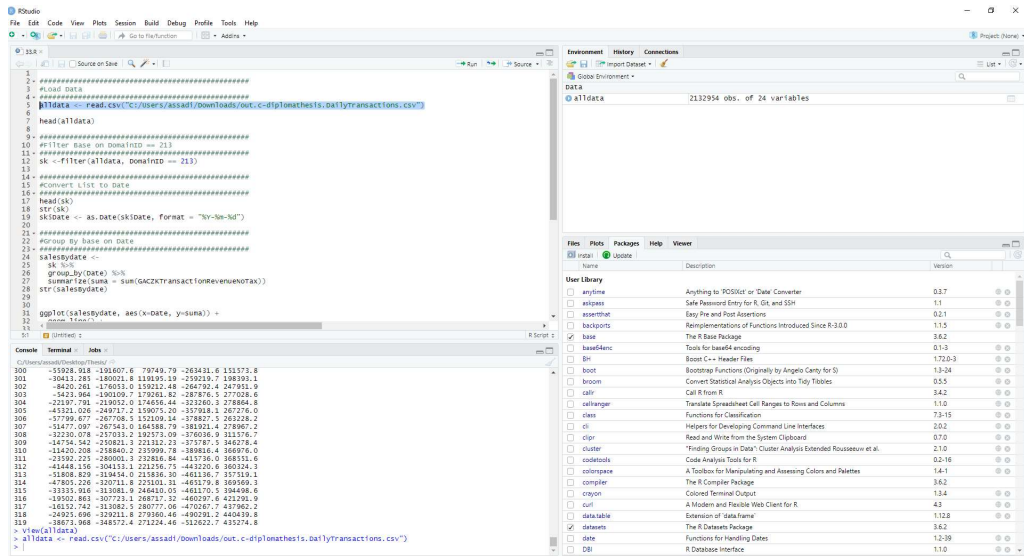


Figure 39 RStudio - Load data

TransactionID	Date	idProfile	Domain	DomainID	Country	Affiliation	DeviceCategory	Channel	Medium	Source	Campaign
1	12122671	NA	sportisimo.cz	111							
2	12463769	NA	sportisimo.sk	213							
3	12859253	NA	sportisimo.cz	101							
4	12044099	NA	sportisimo.sk	213							
5	12586707	NA	sportisimo.pl	601							
6	21274207	NA	sportisimo.cz	111							
7	21764169	NA	sportisimo.cz	111							
8	21402129	NA	sportisimo.cz	111							
9	22223079	NA	sportisimo.cz	101							
10	12753083	NA	sportisimo.cz	101							
11	12724495	NA	sportisimo.sk	201							
12	12718471	NA	sportisimo.cz	101							
13	11866597	NA	sportisimo.cz	111							
14	21350777	NA	sportisimo.cz	111							
15	21754543	NA	sportisimo.sk	213							
16	12628689	NA	sportisimo.cz	101							

Figure 40 RStudio - Check data

3.11.2 Convert Date From List to Date type

```

#####
#Convert List to Date
#####
head(sk)
str(sk)

```

```
sk$Date <- as.Date(sk$Date, format = "%Y-%m-%d")
```

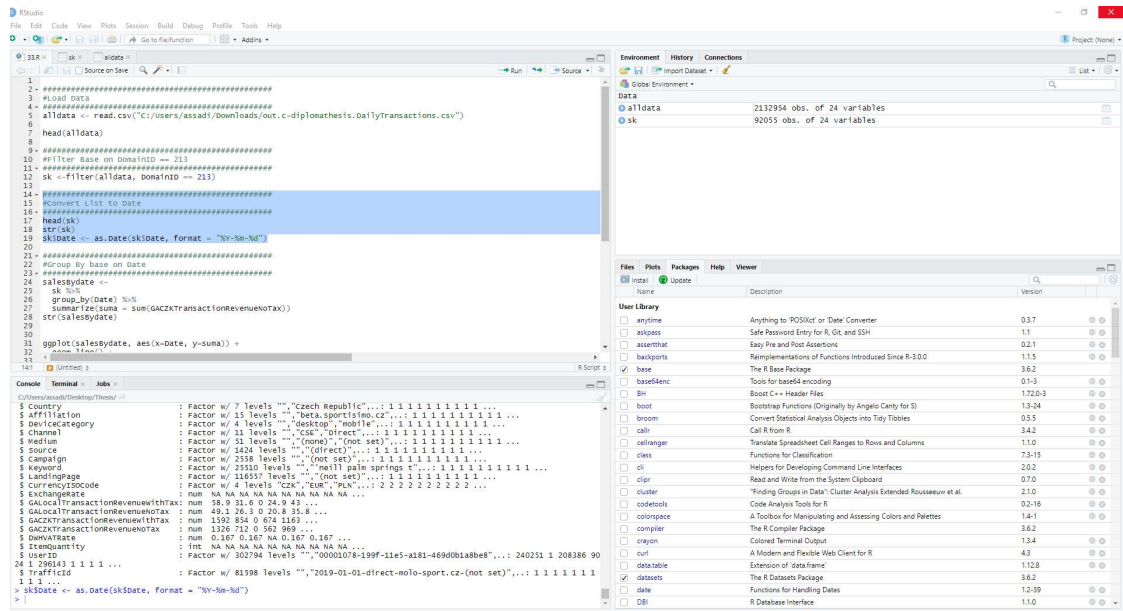


Figure 41 RStudio - Convert Date from List to Date type

Group by based on date and sum of GACZKTransactionRevenueNoTax

```
#####

#Group By based on Date

#####

salesBydate <-

sk %>%

group_by(Date) %>%

summarize(suma = sum(GACZKTransactionRevenueNoTax))

str(salesBydate)
```

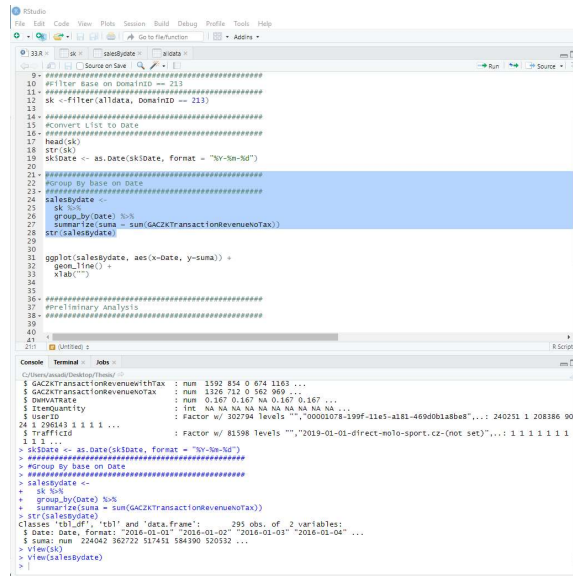


Figure 43 RStudio - Group by based on date and sum of GACZKTransactionRevenueNoTax

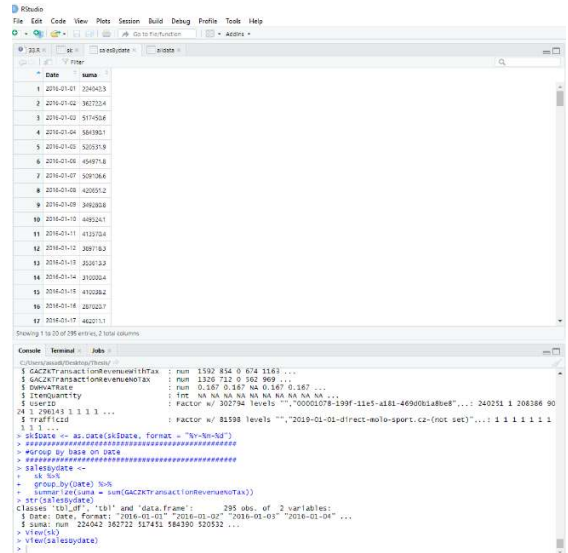


Figure 42 RStudio - Load data

3.11.3 The graph shows sales in 2016

```
ggplot(salesBydate, aes(x=Date, y=suma)) +  
  geom_line() +  
  xlab("")
```

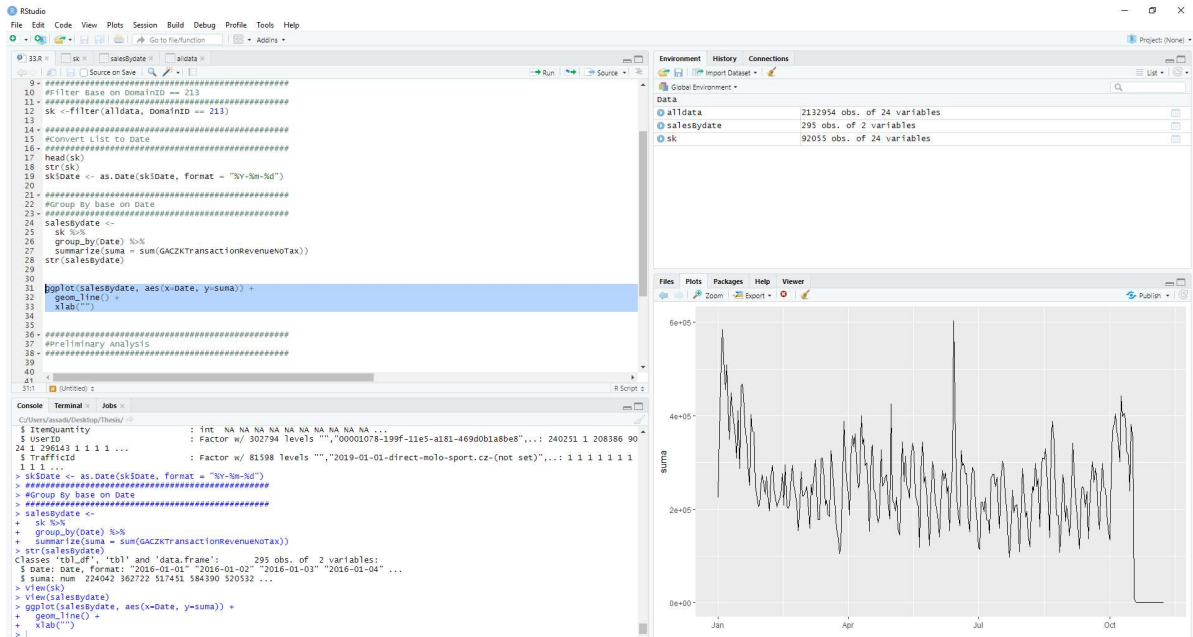


Figure 44 RStudio - The graph shows sales in 2016

1. Here we use the ARIMA Model for preliminary analysis:

ARIMA models are, in theory, the most general class of models for forecasting a time series. For creating the ARIMA model we need to check if the data is the correct data we want to use to forecast.

```
#####  
#Preliminary Analysis  
#####  
  
#fit on ARIMA model  
fit_arima <- auto.arima(salesBydate[,2],stepwise = FALSE, approximation = FALSE, trace = TRUE) #Residual SD= Sqrt() fit_arima <-  
auto.arima(Y,d-1,D-1,stepwise = FALSE, approximation = FALSE, trace = TRUE)  
print(summary(fit_arima))  
checkresiduals(fit_arima)
```

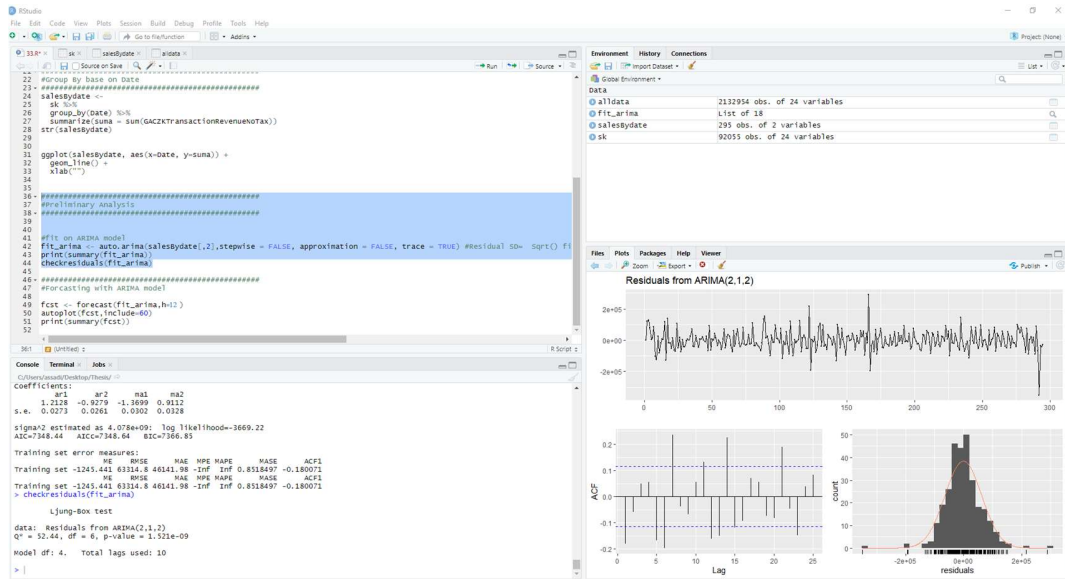


Figure 45 RStudio - The ARIMA Model for preliminary analysis

2. The last step is to sales forecast of upcoming few days. Sales have been showed by these following results:

```

#####
#Forecasting with ARIMA model

fcst <- forecast(fit_arma,h=12)
autoplot(fcst,include=60)
print(summary(fcst))

```

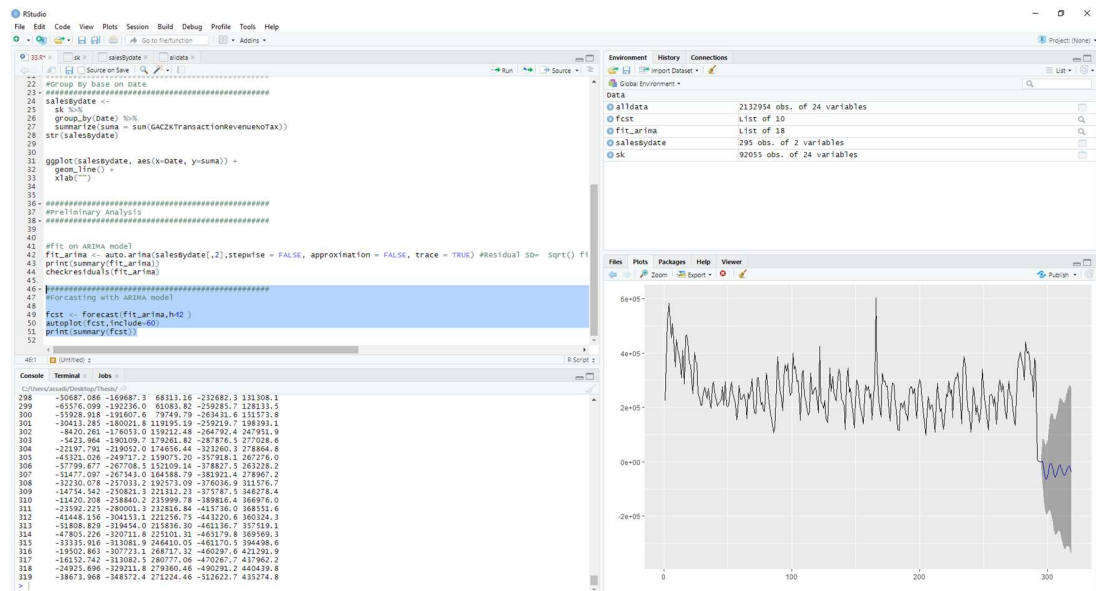


Figure 46 RStudio - Sales forecast of upcoming few days

3.11.4 Final Script in RStudio:

```
#####  
#Load Data  
#####  
alldata <- read.csv("C:/Users/assadi/Downloads/out.c-diplomathesis.DailyTransactions.csv")  
head(alldata)  
#####  
#Filter Based on DomainID == 213  
#####  
sk <- filter(alldata, DomainID == 213)  
#####  
#Convert List to Date  
#####  
head(sk)  
str(sk)  
sk$Date <- as.Date(sk$Date, format = "%Y-%m-%d")  
#####  
#Group By based on Date  
#####  
salesBydate <-  
  sk %>%  
  group_by(Date) %>%  
  summarize(suma = sum(GACZKTransactionRevenueNoTax))  
str(salesBydate)  
ggplot(salesBydate, aes(x=Date, y=suma)) +  
  geom_line() +  
  xlab("")  
#####  
#Preliminary Analysis  
#####  
#fit on ARIMA model  
fit_arma <- auto.arima(salesBydate[,2], stepwise = FALSE, approximation = FALSE, trace = TRUE)  
#Residual SD= Sqrt() fit_arma <- auto.arima(Y,d-1,D-1, stepwise = FALSE, approximation = FALSE, trace  
= TRUE)  
print(summary(fit_arma))
```

```

checkresiduals(fit_arma)
#####
#Forecasting with ARIMA model
fcst <- forecast(fit_arma,h=12)
autoplot(fcst,include=60)
print(summary(fcst))

```

3.11.5 Forecast result by RStudio

Here are the values of forecasting for the next 12 days. This table includes four different parts:

- Upper values with 80% and 90%
- Lower values with 80% and 90%.

	fcst[["upper"]] Time Series: Start = 296 End = 307 Frequency = 1 80%	fcst[["upper"]] Time Series: Start = 296 End = 307 Frequency = 1 90%	fcst[["lower"]] Time Series: Start = 296 End = 307 Frequency = 1 80%	fcst[["lower"]] Time Series: Start = 296 End = 307 Frequency = 1 90%
296	83983.47	127305.7	-79691.9	-123014.1
297	86199.08	142857.2	-127860.4	-184518.5
298	68313.16	131308.1	-169687.3	-232682.3
299	61083.82	128133.5	-192236	-259285.7
300	79749.79	151573.8	-191607.6	-263431.6
301	119195.19	198393.1	-180021.8	-259219.7
302	159212.48	247951.9	-176053	-264792.4
303	179261.82	277028.6	-190109.7	-287876.5
304	174656.44	278864.8	-219052	-323260.3
305	159075.2	267276	-249717.2	-357918.1
306	152109.14	263228.2	-267708.5	-378827.5
307	164588.79	278967.2	-267543	-381921.4

3.11.6 Comparison of the ARIMA model and Time series

By comparison of the Arima model with Time Series, we have concluded that both data values are close to each other and that means our chosen model has been created correctly.

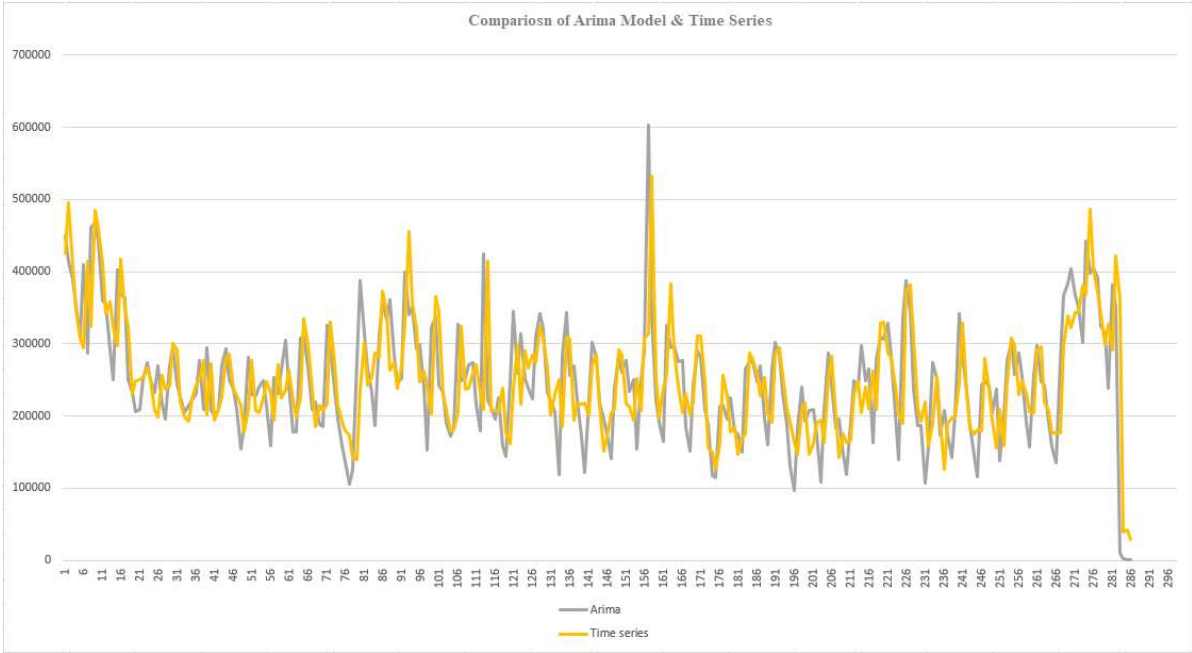


Figure 47 Comparison of ARIMA model and Time series

3.12 Power BI

Power BI is a cloud-based software for business analytics and business intelligence provided by Microsoft. A set of different software services and applications that work together to transform your data into coherent, interactive, and visual information. Your data may be a simple Excel file or a set of cloud data warehouses. Power BI lets you easily connect to your data, visualize important information, and share it with anyone or anywhere you want.

Power BI can quickly and easily provide you with the insights you need for data on an Excel spreadsheet or local database. It can also powerfully model your data to be used to design and develop personalized reports. So, this tool can be used as a tool for visualizing and preparing your report or decision-making engine and analyzing a large group project or organization.

3.12.1 Why use Power BI?

There are several reasons why we need Power BI software and why we should use it as an important tool nowadays. Real-time analytics in Power BI can be provided by connecting directly to data sources. After refreshing, your data will be updated to the latest available data. You can use customized visualization tools created by others or create your own visualization tools. These tools are available in several categories with different settings.

You can easily create a live or non-live connection to your local data sources, such as Scalable Server, and use a secure channel to access data through the data gateway. You can connect to other services such as Scalable Server Analysis Service (SSAS), Microsoft, and more.

Power BI is a new application with the latest technologies such as HTML5, column storage database, cloud computing, mobile application and more. This makes the tool popular.

3.12.2 The workflow in Power BI

The usual workflow in Power BI starts with connecting to a data source and building a report with Power BI Desktop. The report is then transferred from the desktop to Power BI service and released in a way that users of the Power BI service and mobile can use and interact with it. At times you may want to give your colleagues licenses similar to those you have (creator permissions). Then your colleagues can also modify reports, create dashboards and share their work using the Power BI service.

3.12.3 Create a report for sales in 2016 for the SPORTISIMO company:

One of the reporting tools which the company uses is Power BI. Many different reports in each department can help managers to track all events and make decisions for the future, this report shows the sales of 2016 in the Power BI report:

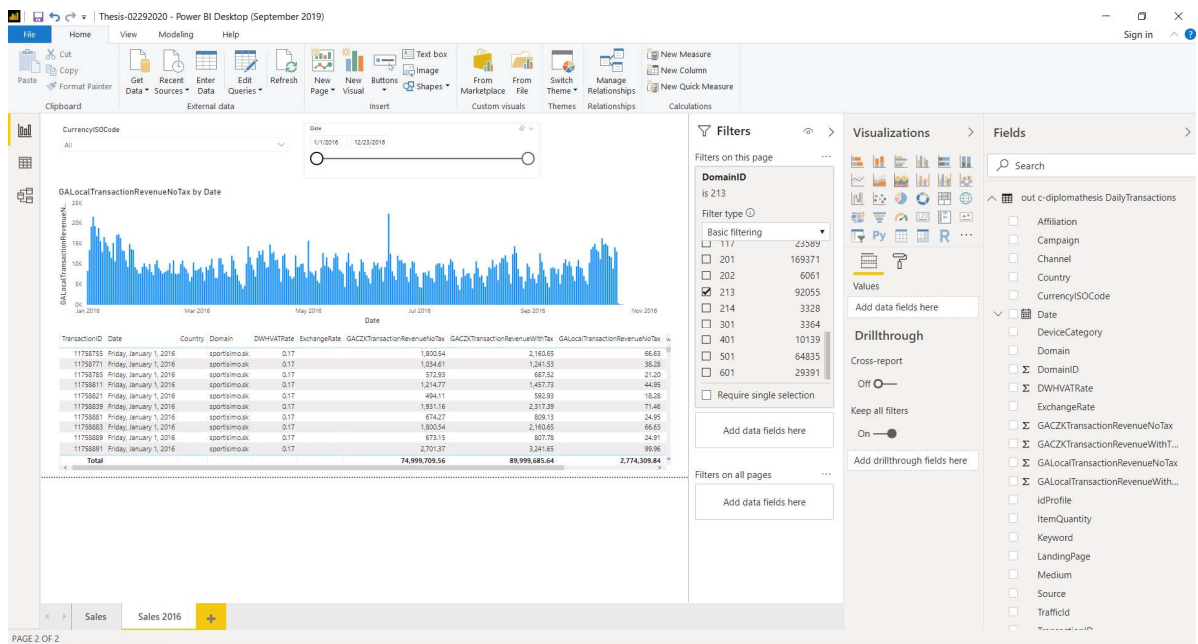


Figure 48 Power BI report

4. CONCLUSION

This analysis focuses on online/offline traffic and online/offline transaction of customers' purchases and it is predicting the sales of products a few days in advance. Use this analysis to determine the strategy to predict what will happen a few days in advance. This analysis shows the maximum and minimum sales in the future so it can help the company to make a marketing decision and how they can use it to improve sales. There are many methods to improve sales, one of them is forecasting.

The first step of the practical part shows the analysis of online shopping and the processes. It shows different tools used in analyzing the business and it explains all the steps more detailed. Draw.io and Craft.CASE are the tools that helped this analysis to show these processes of online shopping. The second step of the practical part shows the specific analysis of the company on how it stores data and analyzes the sales part to improve marketing, Keboola and RStudio are tools that help to store data and forecast sales. In the end the Power BI tool is used for reporting and for showing the result of the sales and to track the sales.

When it comes to sales forecasts, consider three different cash flows: realistic, optimistic, and pessimistic. In this case, you know how much capital you need in the worst-case scenario to be able to handle the situation. The sales forecast determines how much the seller, sales team, or firm will make in each period (weekly, monthly, quarterly or annually). Managers use sales forecasts to estimate what their entire team will achieve. They also predict the total sales of that part of the organization using forecasts made by the sales teams.

A company has a lot of data to deal with. most of this data comes from customer behavior, and its variety and size are largely dependent on the scale of the business. General customer information (age, gender, education, place of residence, etc.), frequency of purchase, amount of purchase, complaint data and customer contact are all such data. Usually, all companies record a lot of this data but having this data alone will not help any business. The main part is to find customers' behavior patterns by analyzing what they call data mining.

In fact; through data mining, a gold mine worth of data is obtained, and this can be used to help the company improve its results significantly.

5. REFERENCES

JAMES, G. An introduction to statistical learning: with applications in R. ISBN 978-1461471370.

REYNOLDS, Janice. The complete e-commerce book: design, build & maintain a successful Web-based business. 2nd ed. Berkeley, Calif.: Distributed to the Book trade in the U.S. by Publishers Group West, c2004. ISBN 9781578203123.

SMYTH, P. -- HAND, D J. -- MANILA, H. Principles of data mining. Cambridge, Mass.: MIT Press, 2001. ISBN 026208290.

eMarketing: The essential guide to marketing in a digital world By Rob Stokes and the Minds of Quirk, First published 2008 by Quirk eMarketing (Pty) Ltd, ISBN: 978-0-620-56515-8.

<https://searchcio.techtarget.com/definition/e-commerce>

<https://www.shopify.com/encyclopedia/business-to-consumer-b2c>

<https://www.oberlo.com/ecommerce-wiki/b2c>

<https://neilpatel.com/blog/connecting-with-millennial-shoppers/>

<https://devrix.com/10-easy-e-commerce-strategies-help-reach-customers/>

https://www.tutorialspoint.com/computer_fundamentals/computer_data.htm

<https://www.computerhope.com/issues/ch001629.htm>

<https://www.datasciencecentral.com/profiles/blogs/the-7-most-important-data-mining-techniques>

<https://www.talend.com/resources/data-mining-techniques/>

<https://data-flair.training/blogs/data-mining-techniques/>

<https://olap.com/learn-bi-olap/olap-bi-definitions/business-intelligence/>

<https://www.guru99.com/business-intelligence-definition-example.html>

<https://vantageonesoftware.com/top-five-benefits-business-intelligence-bi/>

<https://drawio-app.com/tutorials/beginners-guide/>

<http://www.craftcase.com/>

<https://powerbi.microsoft.com/en-us/what-is-power-bi/>

<https://us.hitachi-solutions.com/blog/8-reasons-why-you-should-shift-reporting-from-excel-to-power-bi/>

<https://radacad.com/introduction-to-power-bi-what-is-power-bi>

<https://data-flair.training/blogs/power-bi-components/>

<https://rstudio.com/about/what-makes-rstudio-different/>

<https://www.keboola.com/>

<https://www.sportisimo.cz/>