

## DLMDSEBA02 Project Business Intelligence

Finalization Phase

---

### Optimal Data Mining Tool Selection for Acmeflex

---

Author: Ajaychandra Arekal Satishchandra  
Enrollment Number: 321149868  
Email: [ajay.chandra-a-s@iu-study.org](mailto:ajay.chandra-a-s@iu-study.org)  
Study Program: 120 ECTS M.Sc Data Science  
Date: 26.02.2024  
Place: Berlin, Germany

## Table of Contents

<b>Introduction .....</b>	<b>3</b>
<b>Data Mining Solution Scenarios .....</b>	<b>3</b>
Scenario 1 .....	3
Scenario 2 .....	4
<b>Data Mining Tools .....</b>	<b>5</b>
Weka .....	5
Rapidminer .....	6
KNIME .....	6
<b>Selecting the Optimal Data Mining Tool .....</b>	<b>6</b>
Success Story: How a Home Furnishing Retailer Leveraged KNIME .....	7
Suggested Solution Scenario .....	7
KNIME's Role in Data Mining Strategies at Acmeflex: .....	8
<b>Conclusion .....</b>	<b>8</b>
<b>References .....</b>	<b>9</b>

# Introduction

In the final phase, aligning with the previous stages, we focus on evaluating diverse data mining software solutions to determine the optimal toolset for fulfilling the requirements of Acmeflex Sleep Solutions. This assessment involves considering solution scenarios based on requirements, costs, functionality, usability, and the company's unique scenario. We will thoroughly weigh the pros and cons of each scenario, aiming to make an informed decision. The overarching goal of Phase III is to pinpoint the most suitable data mining software solution tailored to our CRM needs. This entails a comprehensive analysis, ensuring the chosen toolset meets specific requirements, provides robust functionality, remains cost-effective, and comprehensively addresses CRM processes. Ultimately, our objective is to recommend a singular solution scenario that optimizes our CRM strategy, enhances customer satisfaction, and propels business growth.

## Data Mining Solution Scenarios

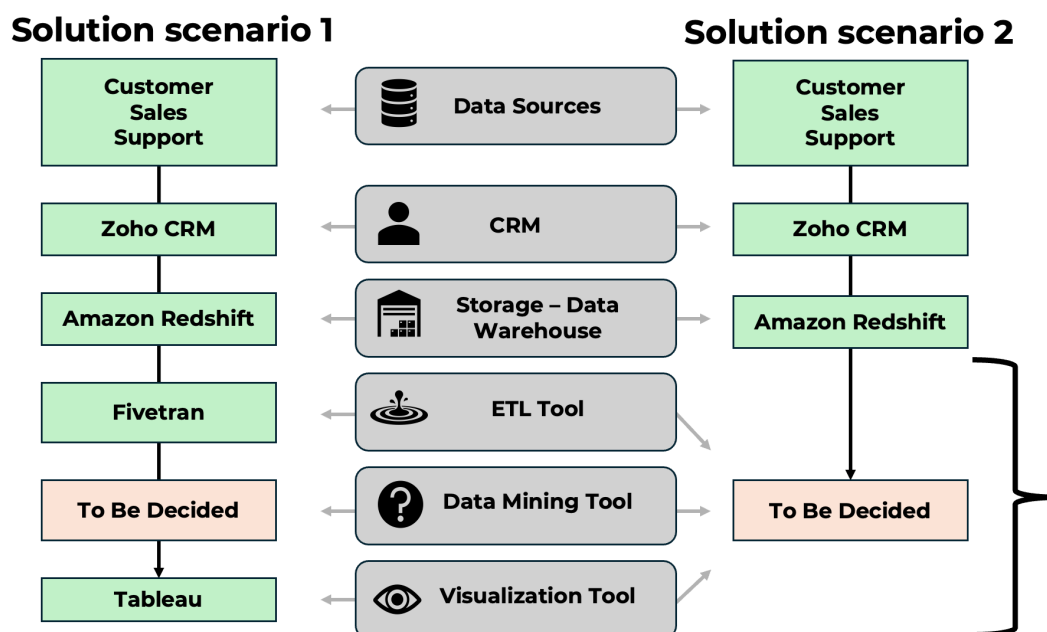


Fig: Solution Scenarios

## Scenario 1

In Solution Scenario 1, integration begins with linking Customer, Sales, and Support data sources to Zoho CRM. This CRM system is connected to Amazon Redshift Data Warehouse. The data stored in the warehouse undergoes extraction, transformation, and loading (ETL)

processes facilitated by the Fivetran ETL Tool. Following this, the data mining tool takes charge of analytical tasks, and the insights are visually presented through the Tableau visualization tool.

### *Pros*

- Zoho CRM: It provides a centralized platform for customer data, streamlining customer relationship management and enhancing data accuracy.
- Amazon Redshift Data Warehouse: It offers scalable and high-performance data storage, enabling efficient analysis of large datasets for informed decision-making.
- Fivetran ETL Tool: It simplifies data integration by automating the extraction, transformation, and loading (ETL) process, reducing manual effort and ensuring data consistency.
- Data Mining Tool: *To Be Decided*
- Tableau Visualization Tool: It provides powerful and interactive data visualization capabilities, enhancing the presentation of insights for better decision-making.

### *Cons*

Leveraging five separate tools—Zoho CRM, Amazon Redshift, Fivetran, a Data mining tool, and Tableau—may introduce challenges in terms of integration complexity, potential data inconsistencies, the need for specialized expertise, and licensing costs, posing hurdles for seamless data flow and analysis.

## Scenario 2

Solution Scenario 2 initiates integration by linking Customer, Sales, and Support data sources to Zoho CRM. The CRM system is connected to Amazon Redshift Data Warehouse. Unlike the first scenario, our approach in this scenario entails selecting a data mining tool directly connected to the data warehouse, capable of handling ETL, data mining, and visualization processes in line with the company's requirements.

Adopting a data mining tool to handle ETL, data mining, and visualization not only simplifies integration but also presents opportunities for cost efficiency. Serving as a comprehensive solution, it substantially diminishes the necessity for additional licensing expenses linked to multiple tools, eliminating separate costs for ETL, data mining, and visualization solutions.

This consolidated approach aligns with the company's budget considerations, providing a more cost-effective solution while maintaining comprehensive functionality.

Therefore, based on the specified criteria, we choose Solution Scenario 2, employing Zoho as the CRM, Amazon Redshift as the Data Warehouse, and intending to identify a data mining tool capable of managing ETL, data mining, and visualization tasks.

## Data Mining Tools

Acmeflex requires a versatile data mining tool to enhance customer focus, satisfaction, and retention which should include the following:

- The tool should offer intuitive functionality for implementing the suggested data mining algorithms.
- Cost considerations should involve a scalable licensing model, ensuring affordability as the company expands.
- Functionality-wise, seamless integration with the data warehouse is crucial.
- Usability demands an easy-to-learn interface for the users, while the tool's scenario adaptability should meet the company's evolving data mining needs.
- Additionally, it should be capable of handling ETL, data mining, and visualization processes.

**The data mining tools which can fulfill the suggested requirements include the following:**



It provides an extensive array of features and algorithms designed for data mining and machine learning tasks, making it well-suited for addressing diverse business scenarios ("WEKA," n.d.). The tool is proficient in handling all the necessary data mining and machine learning algorithms required by Acmeflex. However, it lacks deployable, user-friendly visualization capabilities. Considering our need for a data mining tool that can handle ETL, data mining, and visualization comprehensively, WEKA is not an ideal fit for our project.



It offers a comprehensive set of features and algorithms for end-to-end data science and machine learning workflows, making it suitable for addressing diverse business questions(Arnaldo, 2021) related to customer focus, satisfaction, and retention. RapidMiner is capable of handling all the suggested data mining and machine learning algorithms required by Acmeflex(Hofmann & Klinkenberg, 2016). Additionally, it can integrate smoothly to the company's existing data warehouse.



It delivers a comprehensive platform for end-to-end data science, encompassing the creation, and deployment of analytic models within the organization. Its visual interface facilitates the design and execution of data workflows, supporting various data manipulation and analysis tasks. With a user-friendly UI, KNIME ensures accessibility and ease of use. Its flexibility and scalability make it well-suited to tackle diverse challenges in data mining(*Software Overview | KNIME*, n.d.) related to customer focus, satisfaction, and retention. KNIME is capable of handling all the suggested data mining and machine learning algorithms required by Acmeflex(Feltrin, 2015). Moreover, it can seamlessly integrate with the company's existing data warehouse.

To execute the mentioned tasks, the required plan would be the RapidMiner Enterprise plan, priced at 15,000 euros per user per month(*RapidMiner Pricing 2024*, 2022). Alternatively, the KNIME Standard plan, costing 1,042 euros per user per month(*KNIME Hub Pricing*, n.d.), is also suitable.

## Selecting the Optimal Data Mining Tool

Below is the table illustrating the points fulfilled by the selected data mining tools:

	Requirements	Costs	Functionality	Usability	Scenario	User-Friendly UI
KNIME	✓	✓🚀	✓	✓	✓	✓🚀
Rapidminer	✓	✓	✓	✓	✓	✓

Both RapidMiner and KNIME prove to be comprehensive solutions, effectively meeting the project requirements. However, in our assessment, the better user interface of KNIME and the availability of a more cost-effective licensing plan became decisive factors, ultimately leading us to choose KNIME for Acmeflex's data-related tasks.

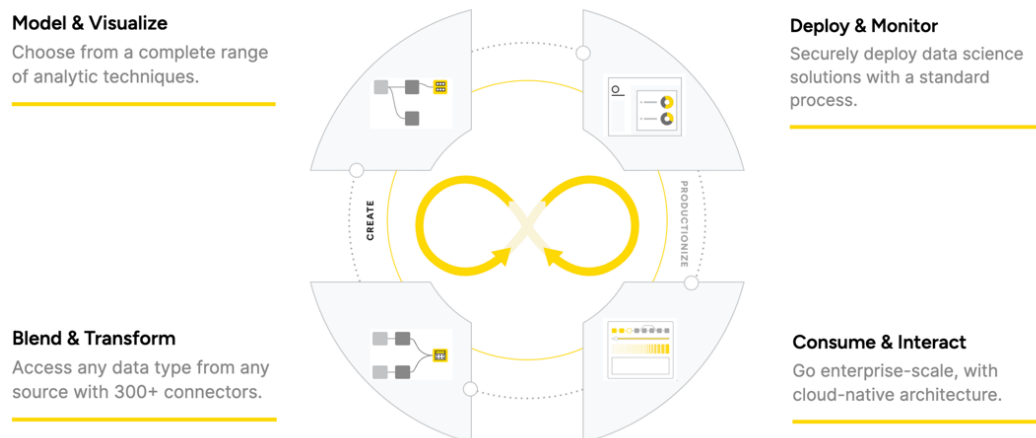


Fig: KNIME - A complete platform for Data Science (*Software Overview* | *KNIME*, n.d.)

## Success Story: How a Home Furnishing Retailer Leveraged KNIME

JYSK, an international home furnishing retailer, successfully transformed into a data-driven organization using KNIME for predictive analytics and overcoming various challenges. By saving hundreds of work hours monthly, optimizing decisions on store locations, product allocation, and discount strategies, KNIME demonstrated its flexibility, seamless integration, and low-to-no code usability. JYSK's success story highlights KNIME's potential in creating a data-driven culture, making it a compelling choice for Acmeflex's advanced analytics needs(*How JYSK Optimized Product Sales with KNIME for Predictive Analytics* | *KNIME*, n.d.).

## Suggested Solution Scenario

Here is a diagram depicting the recommended solution scenario:

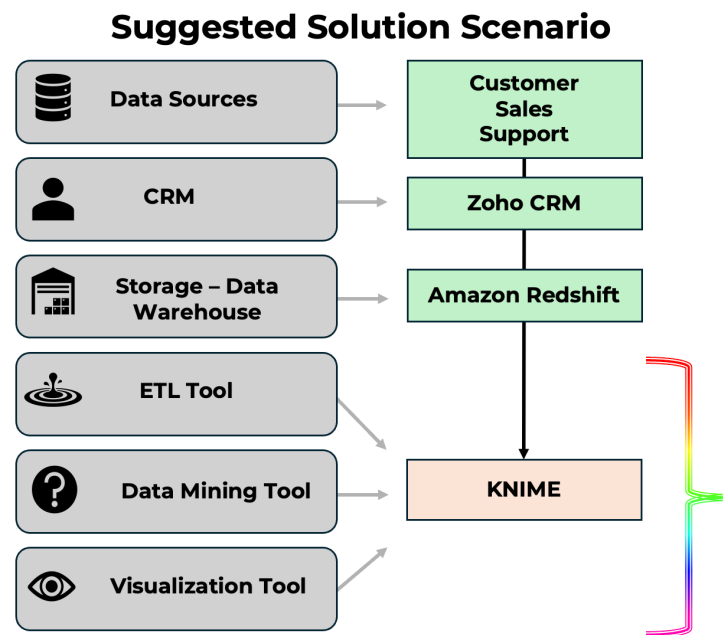


Fig: Suggested Solution Scenario

The proposed solution scenario entails linking the data sources with the existing CRM, establishing a connection between the data warehouse and the CRM. KNIME, the newly selected tool, is integrated with the data warehouse, managing ETL, data mining, and visualization processes.

### KNIME's Role in Data Mining Strategies at Acmeflex:

- Utilizing advanced techniques like Customer Segmentation and Association Rule Mining, KNIME empowers Acmeflex in refining customer focus.
- KNIME is vital in improving customer satisfaction at Acmeflex through advanced techniques like Predictive Analytics and Sentiment Analysis.
- KNIME maximizes customer retention at Acmeflex through advanced Predictive Modeling and Customer Lifetime Value Prediction techniques.

## Conclusion

In conclusion, improving Customer Relationship Management (CRM) at Acmeflex Sleep Solutions is crucial, and achieving this involves skillfully applying data mining methods. As the management of a company specializing in selling sleep-related products, including pillows,



mattresses, sleep aids, and accessories both in-store and online, we should efficiently address the targeted CRM challenges by strategically employing data mining techniques. KNIME, known for its versatility as a data mining tool, emerges as the ideal choice to effectively tackle the specified methods, covering customer focus, satisfaction, and retention. Moreover, KNIME seamlessly extends its capabilities to handle ETL and Visualization processes, making it a comprehensive solution tailored to the specific requirements of Acmeflex.

## References

- Arnaldo, M. (2021, October 4). Intro to Rapidminer: A No-Code Development Platform for Data Mining (with Case Study). *Analytics Vidhya*.  
<https://www.analyticsvidhya.com/blog/2021/10/intro-to-rapidminer-a-no-code-development-platform-for-data-mining-with-case-study/>
- Feltrin, L. (2015). KNIME an Open Source Solution for Predictive Analytics in the Geosciences [Software and Data Sets]. *IEEE Geoscience and Remote Sensing Magazine*, 3(4), 28–38. <https://doi.org/10.1109/MGRS.2015.2496160>
- Hofmann, M., & Klinkenberg, R. (2016). *RapidMiner: Data Mining Use Cases and Business Analytics Applications*. CRC Press.
- How JYSK optimized product sales with KNIME for predictive analytics | KNIME*. (n.d.). Retrieved February 23, 2024, from <https://www.knime.com/success-story/how-jysk-optimized-product-sales-knime-predictive-analytics>
- KNIME Hub Pricing*. (n.d.). Retrieved February 19, 2024, from <https://www.knime.com/knime-hub-pricing>
- RapidMiner Pricing 2024*. (2022, March 28). TrustRadius.  
<https://www.trustradius.com/products/rapidminer/pricing>
- Software overview | KNIME*. (n.d.). Retrieved February 21, 2024, from <https://www.knime.com/software-overview>
- WEKA. (n.d.). *WEKA*. Retrieved February 23, 2024, from <https://www.weka.io/get-started/>