

## Factor Analysis of Customer-Ranking Using Multiple Data Source for International Standard Audit Business

Kritchanut Chaimongkon<sup>1</sup> and Juggapong Natwichai<sup>2</sup>

<sup>1</sup> Data Science Program, Chiang Mai University, Chiang Mai, Thailand

<sup>2</sup> Faculty of Engineering, Chiang Mai University, Chiang Mai, Thailand

Kritchanut\_c@cmu.ac.th

**Abstract.** This independent study analyzes the factors affecting customer priorities using multiple data sources for the International Accreditation Services business with a conceptual representation of the importance of this independent study. As the current offering model does not prioritize and take into account a number of factors, this work can result in a more effective offering. Factor analysis by visualizing information is studied to see the overall picture of the data as a dimension. The model is developed into Microsoft Power BI, thus it is possible to find observations from the various type of data. We have found the following insight, i.e. the food factory business had the highest number of requests for certification, followed by electronic parts manufacturing and energy sector. In terms of financial viability factor, if we observed from those who are interested in purchasing the certification service, it was found that the ability to spend on investments was approximately four times higher than those of the non-interested group. These factors together with registered capital, income, and asset value can further prioritize the service offering in the future.

**Keywords:** Factor Analysis, Visualize Analysis , Data Visualizations.

### 1 Introduction

At the present, accreditation of international standards for factories is an important activity to increase credibility or in some industries it is a condition of trade between the factory and the buyer. At the same time, the competition is rather aggressive from the experience of the sale teams for certification business. Therefore, there is a need to use the alternative data to further benefit people who are working in the sale departments. Thus, this work pull together various source of data and integrate them to determine the insight for the sales team as the measurement pattern regularly. This can help improving efficiency of the sales process, i.e. shortening the time to contact to make an offer, focusing only on customers with high sales opportunities. The contribution can help studying on how to apply data science technology for business-to-business sales. This

will benefit both vendors selling international certification services, as well as businesses that have a business-to-business sales model in other types of products.

## **2 Method**

The proposed system is separated into three parts. The first part is design the type of data to be analyzed. The second process is prepare data for use in visualization process. The last part is using Microsoft Power BI to visualize the data.

### **2.1 Collect Data from multiple source**

The purpose of this process is to select the data for use in this factor analysis, starting from the 3 data types: Type 1 is data of group of customers that were interested in purchasing international certification services. Type 2 is the data of customers who are not interested in purchasing international certification services. Type 3 is the data of customers who may purchase international certification services. After that, the data from the sources were studied.

### **2.2 Prepare Data**

After the study and selection of factors by analyzing the exposed sources, the data from the revealed sources was extracted using the Selenium Web Driver tool with python language. The data was collected in CSV format. The main source of data is the business warehouses of the Ministry of Commerce (DBD). The structure of data is as follows: Industry type, Customer location area, registered capital, period of business operation, total revenue, net profit, selling and service expenses, assets, percentage change of all financial factors, and type of purchase decision. After the data is retrieved, the data is cleaned to prepare it for visualization on the Microsoft Power BI tool.

### **2.3 Visualization**

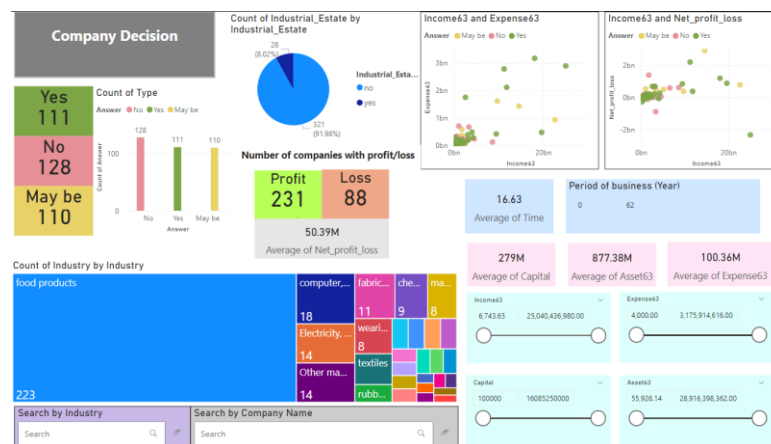
In this process, data visualization uses data from the master database and publish to web service. Microsoft Power BI is used to present visualized data from the dataset with experience from domain experts.

## **3 Result**

The result of the visualization for all of three data type for analysis were divided into seven components for analyzing. The first view is the overview of the number of all data types, reflection of the ratio of factories with areas in industrial estates and outside industrial estates, reflection the number of profit and loss companies as well as their average value, a scatter plot of company positions with three classifications of data

groups as in Fig 1. Such visualizations were presented to the domain experts. It was found that the food factory business had the highest number of requests for certification, followed by electronic parts manufacturing and energy sector. In terms of financial viability factor, if we observed from those who are interested in purchasing the certification service, it was found that the ability to spend on investments was approximately four times higher than those of the non-interested group.

**Fig. 1.** Overall Visualization from 3 data group



## 4 Conclusion

In this study, the factors that may affect the purchase decision of international accreditation services were compiled into a visual representation of the results by seven components for sales planning. We have found a few insight from the visualization, for example, food factory business is the most interesting in term of opportunities for sales campaign. In the future, such work can be applied to the more focused certificates, e.g. environmental, or IT security.

## References

1. International Organization for Standardization ISO Central Secretariat, ISO SURVEY 2019. [www.iso.org.2019.](https://www.iso.org/the-iso-survey.html) แหล่งที่มา: <https://www.iso.org/the-iso-survey.html>. [Accessed: 04-July-2021]
2. แกะรอย Data, *Web scraping ด้วย Python โดย Selenium <Beginner>*, [https://kaeroidata.medium.com/2019.](https://kaeroidata.medium.com/2019.แหล่งที่มา:https://kaeroidata.medium.com/web-scraping-%E0%B8%94%E0%B9%89%E0%B8%A7%E0%B8%A2-python-%E0%B9%82%E0%B8%94%E0%B8%A2-selenium-beginner-ad6369624ccd) แหล่งที่มา: <https://kaeroidata.medium.com/web-scraping-%E0%B8%94%E0%B9%89%E0%B8%A7%E0%B8%A2-python-%E0%B9%82%E0%B8%94%E0%B8%A2-selenium-beginner-ad6369624ccd>. [Accessed: 04-May-2021].

3. Pakhapoom Sarapat, Nontawit Cheewaruangroj, เลือกแผนภาพอย่างไรสำหรับการทำ Data Visualization. bigdata.go.th.2021. แหล่งที่มา: <https://bigdata.go.th/big-data-101/picking-chart-for-data-visualization/>. [Accessed: 04-July-2021].
4. Julie Zuckerman, *Data Visualization and Visual Analytics: Seeing the world of Data*. www.sisense.com. แหล่งที่มา: <https://www.sisense.com/blog/data-visualization-and-visual-analytics-seeing-the-world-of-data/>. [Accessed: 05-July-2021].
5. Saksit Srimarong, *Visual Analytics การวิเคราะห์ด้วยภาพ*. bigdata.go.th.2021. แหล่งที่มา: <https://bigdata.go.th/big-data-101/visual-analytics/>. [Accessed: 05-July-2021].
6. Duangjai Jitkongchuen, การเตรียมข้อมูลเบื้องต้นด้วย Power BI (Basic Data Preparation with Power BI). bigdata.go.th.2021. แหล่งที่มา: [https://bigdata.go.th/big-data-101/data\\_preparation\\_power\\_bi/](https://bigdata.go.th/big-data-101/data_preparation_power_bi/). [Accessed: 11-July-2021].