Reducing Food Waste by IT Systems

Case Sodexo's cafeteria

HAMK
HÄMEEN AMMATTIKORKEAKOULU
HÄME UNIVERSITY OF APPLIED SCIENCES

Bachelor's thesis

Degree Programme in Business Information Technology

Hämeenlinna University Center

Spring 2019

Yves Einfeldt

HÄMEEN AMMATTIKORKEAKOULU
HÄME UNIVERSITY OF APPLIED SCIENCES

Degree Programme in Business Information Technology
Hämeenlinna University Center

Author	Yves Einfeldt	**Year** 2019
Subject	Reducing Food Waste by IT Systems in Sodexo's Cafeteria	
Supervisor	Lasse Seppänen	

ABSTRACT

One-third of all originally created food will be thrown away before it even reaches the consumers. With an increasing population, this leads to a massive problem society has to deal with. Hotels, restaurants, and cafeterias are responsible for a significant part of the food waste generation. With the digital transformation a lot of opportunities emerge. With IT systems like a point-of-sale system and a tracking solution for food waste, the collection of data is possible. The data can be analysed to identify the origin of food waste.

Supported by literature research from multiple papers and textbooks the hospitality industry and gastronomy with their opportunities and challenges using IT systems were described. Also, a focus on food and the origin and reasons for food waste in the supply chain are mentioned and the current ethical movement to more sustainability is explained in the theory part of this thesis. The second part was a practical data analysis investigating how Sodexo's cafeteria at HAMK's campus in Hämeenlinna can reduce their food waste creation. With a cross-industry standardised process for data mining, the occurrences of food waste were identified. On the days of the week when fewer dishes were sold, more food waste was produced. Because of relying on a non-accurate forecast of the expected amount of guests, 30 kg of food waste is generated daily. The error margin of the previously used forecast was 30%. By creating a new calculation of the expected guests and comparing the historical data, the error margin could be improved to 15%.

With the improved accuracy of the prediction, the executive chef can perform better menu calculations and produce more efficiently the demanded amount of food. This case confirms that IT systems support the business processes of the cafeteria and thus reduce their food waste.

Keywords Food waste, Gastronomy, Digitalisation, Predictive Analytics, CRISP-DM

Pages 45 pages

CONTENTS

FIGURES

TABLES

LIST OF ABBREVIATION

API	Application programming interface
BI	Business Intelligence
CRISP-DM	Cross-industry standardised process for data mining
CSV	Comma-separated values
HAMK	Häme University of Applied Sciences
IT	Information Technology
POS	Point-of-sale system

1 INTRODUCTION

This thesis creates a sustainable value by helping Sodexo's cafeteria at HAMK, Häme University of Applied Sciences in Hämeenlinna Finland, to reduce their food waste. The theoretical and practical parts are aiming to give cafeterias and restaurants insights in how IT systems might support their business processes and reduce their food waste.

The research will focus firstly on the gastronomy industry with its business processes in general. Secondly, a more detailed analysis based on research material about food waste will give an overview of the problems and opportunities in this industry. Furthermore, the current state of affairs related to the digitalisation and insights in Sodexo's cafeteria will be given to understand what kind of IT systems they need to support their business processes and reduce their food waste.

To reduce food waste, the cafeteria needs to know how many guests they are having each day. Currently, Sodexo is getting the numbers of students who should be at University according to the room reservations from HAMK's IT department. As this data is very inaccurate and often wrong, the kitchen does not know how much food they have to prepare. With an introduction in data science as the main topic for possible enablers such as predictive analytics, the theoretical part explains how more accurate forecasts will be possible to create.

As a practical outcome, a data analysis will be performed which supports the chef in calculation and purchase decision but also in the moment of preparing and producing the food. This project realisation and results will be described in this research as well and will answer the following questions:

- How can IT systems support business processes in a cafeteria and thus reduce food waste?
- How can a cafeteria perform better menu calculations and produce more efficiently through predictive analytics?

2 THE HOSPITALITY INDUSTRY

This chapter will give a general overview of the hospitality industry. In order to gain sufficient knowledge for understanding the possibilities and challenges of the industry, the business processes will be explained, based on literature and research.

The hospitality industry is understood as an industry that serves and accommodates guests. The industry is divided into three sectors: hotel, gastronomy, and para-hotel. The most known sector is the hotel business. There, guests are satisfied through comprehensive service which includes food, beverages, and overnight stays. The gastronomy, on the other hand, offers no accommodation and focuses on catering only. This includes restaurants but also cafeterias from hospitals or universities. The third sector, para-hotel, offers accommodation with limited service like holiday flats or camping. (Egli, 2016, pp. 20–22)

The value chain in the hospitality industry has its origin in agriculture where the raw material gets produced. After processing the raw material, it gets sold to the wholesaler and later to the regional dealer. The process ends in the kitchen where a meal for the guests gets created out of the raw material. Every value chain step requires an exchange between goods and money. Based on this value chain, various costs for a company in the hospitality industry arise which adds up to 75% of the products selling price. This costs include mainly goods and employee costs. With the remaining operating, financial and investment expenses, the profit in this industry is very little. (Egli, 2016, pp. 20–21)

The restaurant industry dollar is a report from Deloitte & Touche LLP and the National Restaurant Association (2010, p. 1) which contains feedback from 650 restaurant operators. There are various types of restaurant analysed in this report such as cheap or expensive full-service restaurants and limited service restaurant. Full-service restaurants are establishments with a wide offer of dishes, tables and waiters. (CHD Expert, 2019a)

Limited service restaurants on the other hand do not provide the same service as full-service restaurants. In a limited service restaurant the guests often pay before receiving the food, which is often taken out instead of eaten in the restaurant. (CHD Expert, 2019b) According to the National Restaurant Association (2010, p. 9) the median income before income taxes in limited-service restaurants is around 5.9%. This leads to an income of $ 5.90 when selling a meal (food and beverage) of $ 100.00 after deduction of all costs such as ingredients, salaries, the restaurant occupancy costs, and before taxes.

In full-service restaurants, the median income before income taxes varies between 1.8% and 3.5%, which makes it very hard to survive in this

market. Moreover, with the three main aspects of finances, real-estate value, and business operation the decision to invest in this industry is critical because it is very capital-intensive. (Barrows & Powers, 2009, p. 370)

The management needs to adjust their strategy towards the current market demand to increase their profit. Even though, it is changing from a fine-dining restaurant to a more relaxed atmosphere and more reasonable prices. In order to increase profits, operators need to maximise two basic approaches to being competitive. One is to increase sales and the other is to reduce costs. However, it needs to be considered that reducing cost might have a more significant effect on the profit instead of increasing the sales, as the increased sales affect the variable costs which will not lead to a full profit increase. With multiple common techniques restaurants control costs by their productivity standards such as how many guests are served per hour, breakage of the inventory like glass, china, and silver or by monitoring supplies. In the latter case, the aim is to purchase non-food supplies such as soap, paper, etc. in larger quantities in order to obtain a quantity discount. This method can also be used to buy larger quantities of ingredients that do not spoil quickly. (Barrows & Powers, 2009, pp. 72–117)

In order to distinguish a company from others and to remain successful on the market, the economic and social environment have to be well-known. This starts with the observation of the economy to evaluate how much money the guests are willing to spend in regard to the current economic situation and the stability of their real income. Furthermore, political changes do have an impact through various legal requirements. These may be in relation to labour law, construction law or hygiene requirements. Moreover, there are the social and ecological environment. In the social environment it is important to find out how the behaviour of consumers could change. In the ecological environment, however, is the focus on the sustainable behaviour of the guest. Lastly, the technological developments should also be considered which could have an influence on the various systems. (Egli, 2016, p. 23)

2.1 **The gastronomy as a subdivision**

The definition of gastronomy is difficult to define as it is used in various subjects and professions. In the hospitality industry, organisations with various business models define gastronomy differently. Thus, broad definitions such as gastronomy is “study of nutrition” or “the art, or science, of good food” are already useful definitions in general. Not only food is important in gastronomy but also beverages. (Gillespie & Cousins, 2001, p. 2)

According to Egli (2016, p. 22) gastronomy is the generic term for various restaurants. These can be bars, cafés, bistros, restaurants of a specific

kitchen like Chinese, Italian, vegetarian, and even catering organisations or different cafeterias. The gastronomy sector, which is given the greatest attention in this thesis, can be managed independently, by private tenants, but can also act regionally, nationally or even internationally. There are differences to such called system gastronomies which are systematically organised and transfer the standards from the main office to all their branches. This organisation model differs from classical gastronomy business but is widely used in the gastronomy and reflects in cafeteria chains, among other things.

2.2 **The structure of cafeterias**

The cafeteria's basic characteristic is that it is self-served. The guests come to where the food and beverage are offered and take according to the offer what they would like without being served by a waiter. The offer in a self-service is often limited in the choice and number of dishes for a fixed price. Furthermore, it is usually available during a specific time period. "A la carte" menu is the opposite where the guest can decide from a more extensive choice with a separate price for each dish. This is common in fine dining restaurants or other operations like brasseries, cafés or catering. (Gillespie & Cousins, 2001, pp. 71–74)

Cafeterias which are systematically managed and organised are seen as chains. (Egli, 2016, p. 22) Such chain restaurant systems have different strengths to create a more sustainable business. With their marketing and brand recognition, and their control and information systems which are organised centrally, they can scale very easily on the market. Chains which have a big business might even receive the capital from banks easier, than small independent operators. Since it is already known that it is a competitive market with high capital costs, a big chain might have advantages to run a cafeteria. (Barrows & Powers, 2009, pp. 128–132)

The organisational structure of cafeterias is different from that of hotels and is therefore similar to other gastronomy businesses such as catering or restaurants. The management, administration, and food and beverage departments are common in hotels, restaurants, and cafeterias. The hotel has a rooms division with tasks in housekeeping, reception, or concierge which are not needed in the gastronomy. Cafeterias still need employees in the cleaning department, as well as maintenance of the equipment, in particular, the kitchen. (Egli, 2016, p. 26)

2.3 **The kitchen and the menu planning**

In the food and beverage department of a gastronomy business, there is the kitchen team associated. The "chef de cuisine" (French for the executive chef) is in charge of the kitchen. He is responsible for the work schedules, menu planning, calculation of the dishes, and the apprentice

supervisor for chefs. The team is organised in specific functions which have a French name. The “saucier” creates sauces, meat, chicken, fish, and other warm meals. The “entremétier” is responsible for soups, vegetables, and potato meals. The “garde-manger” does the preparation of raw meat, fish, and chicken, but also the creation of starters like salad and dishes in the buffet. The “rôtisseur” produces dishes on the grill or roast them in the oven while the “pâtissier” is responsible for the desserts. (Egli, 2016, pp. 26–27)

Each organisation and its kitchen differ and have individual aspects to make the most of their operations. The objectives from the kitchen are smooth workflows in preparing and producing the food, proper use of the equipment and the premises, and reduce stress and disorder as much as possible. With the following tasks, the kitchen has to take care of various tasks during the production process: calculating, purchasing, storing, preparing, producing, refining and disposing of a meal.

All dishes must be created in the exact order and within the given time in order to present them appropriately to the guest. There must be no additional effort or unnecessary costs throughout the production process. To achieve the given goals and bear the responsibilities there are five different production systems like “cook and serve”, “cook and chill”, “cuisson sous vide”, “cook and hold”, and “cook and freeze” which are now described in more detail.

With “cook and serve” the food will be prepared with the needed preparation, production and the finishing in one flow to minimize the planning period. It is the traditional method of producing food and can easily handle individual wishes from guests but might cause stress situation if there are a lot of unexpected orders at the same time.

“Cook and chill”, on the other hand, needs intense planning because the food will be chilled after the production. After it is cooled and stored the food will be portioned, prepared and served. This type reduces stress and optimises the working hours but needs more effort in planning and consideration of hygiene standards.

“Cuisson sous vide” (English: steam cooking in a vacuum bag) is a cooking technique which is often used when the production and finishing are in geographically different locations. After short cooking, the food will be vacuumed, afterwards it will be gently cooked and then stored for up to 14 days in the refrigerator or cold rooms. Before serving, the food will be portioned and finished. It needs more material and energy than other techniques and it is also challenging to season the dish in the production.

“Cook and hold” is also energy-intensive while keeping the food, at the right temperature, warm after cooking. The hygiene standards must be

respected, especially for some ingredients that may not be suitable for keeping warm over a more extended time period.

For sweet dishes such as desserts or in large companies such as restaurant chains, the kitchen uses the technique of “cook and freeze”. With the efficient shock freeze, the food can be stored in the freezer after production for a longer time. The dish has to be thawed first before it can be portioned, finished, and then served to the guests. (Egli, 2016, pp. 45–54)

Not only the production but also the planning and the creation of the menu is an essential task from the kitchen. The modern food order has between three to six courses and is nutritionally in small portions compiled. To compose a harmony of taste and appearance in all dishes on a menu, it might even be necessary that the executive chef gets help by restoration professionals, nutritionists, and even physicians in some cases.

During operating hours it is needed to get statistics about the sales and demand of the dishes, and personal customer feedback to understand if the guests like the menu and what can be improved. To create a menu with a high demand it is essential to use seasonal and different ingredients and take care of special dishes for children or seniors, or even vegan, vegetarian or gluten-free to satisfy the customer’s needs.

The menu is like a contract between the provider and the customer. Thus, the offer must be described as precisely as possible, since the legal basis is to protect the customer from deception and fraud. The declaration regulations must also be strictly observed, which are regulated by law. The defined price for the dishes on the menu is depending on the costs of the restaurant.

While a restaurant with waiters has typically more employees than self-service restaurants or cafeterias, the costs might be higher, and this affects the selling price. Additionally to the employee costs, the operations and ingredients costs need to be taken into account as well. Thus, calculation and purchase management is essential. (Egli, 2016, pp. 669–673)

2.4 **Purchase management and calculation of produced food**

The purchase management includes everything about the procurement of ingredients, beverages and all other goods needed in a gastronomy business. In most companies, the administration with their purchase team and leader is responsible for all the tasks around purchase management. The objectives for them is to purchase the needed amount of material in the required quality, at the right moment for the best price and the ideal service from the supplier. To achieve these goals the process of the purchase team’s is presented in Figure 1. (Egli, 2016, pp. 26–30)

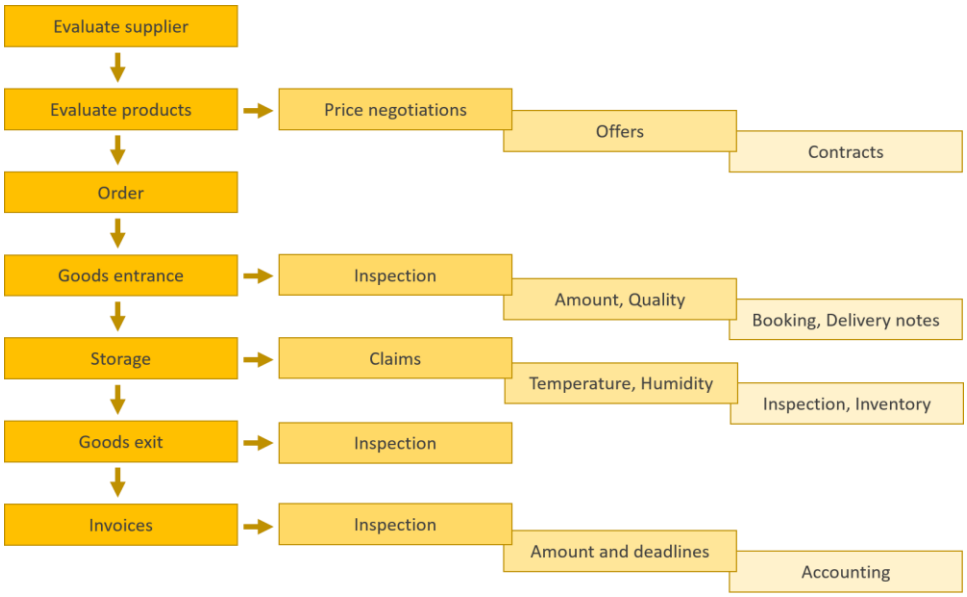

Figure 1. Tasks and process of the purchase team (Egli, 2016, p. 30)

As already mentioned it is essential to decrease the costs of the business as much as possible to be successful in long terms in this industry. Thus, the purchase management is a critical factor for this success. The mutual trust between the restaurant and its suppliers, as well as their professional advice, punctual delivery, and fair prices, are essential for the entire purchasing planning to work. Costs can be saved through appropriate product evaluation, price negotiations and clear and accurate inventory levels. However, it is not about the cheapest solution, but about the price-performance ratio which is decisive and also fulfils the necessary quality requirements. (Egli, 2016, pp. 30–31)

In addition to purchasing management, the preceding calculation is necessary to ensure that the correct quantity is purchased. Thus, it is required to calculate the number of ingredients used in the recipes and how much should be purchased in total, in order to offer a defined number of meals. A corresponding new calculation and control will be necessary at a later date to validate how high the effective costs were, in order to make more accurate calculations in the future that can reduce costs even further. It is also important to know how many meals have to be prepared and how many guests are expected. To simplify purchasing management and calculation and make it more efficient, electronic data processing can be of help. (Egli, 2016, pp. 35–43)

2.5 Innovation as a success factor

Nowadays electronic data processing in the gastronomy and also the hospitality industry, in general, is already used, but not as widespread as in other industries. (Gandhi, Khanna, & Ramaswamy, 2016) According to

Egli (2016, p. 44) the IT might support different workspaces like the purchase management, calculation, creating the work plan, storage and inventory, or menu planning. The Restaurant Technology Report from Toast (2017, pp. 4–15) have analysed the impact of restaurant technology to improve the customer experience in restaurants. “95% of restaurant owners agree that technology improves restaurant efficiency, while 73% of the guests say technology improves the dining experience” the key finding of the report confirms this claim. The results of the report show that guest Wi-Fi and online ordering are the most critical factors for a good experience according to the guests and also the operators. The guests are more focused on accessibility, they like to make the reservation online and even on social media platforms like Snapchat or Instagram instead of calling the restaurant. Furthermore, the payment methods are important for the guests, they prefer credit card in 90% of the cases over cash and almost 60% of all survey participants use mobile pay always or sometimes.

The factors mentioned by the guests as innovative are not the key to success. The most important success factor for innovations in the hotel industry is understanding and reacting to the market’s demand. (Ottenbacher, 2007, p. 445) A good approach is to look at the key figures on a daily basis in order to attract more guests. However, the restaurant should not forget the quality of the food, which is the most important factor for a positive restaurant experience for customers. Restaurants could underestimate this factor because, according to Toast's Restaurant Technology Report (2017, pp. 7–22), they believe that service quality is as important as the food itself.

Younger people are more interested in a good meal and even in the preparation on their own and in a fast and efficient process than receiving high quality service. Consequently, this is why the quality of service, in regards to the waiter and employees, is not the significant factor which restaurants claim. According to Ottenbacher (2007, pp. 432–436) customer satisfaction and employee feedback are important in measuring the success of innovations, not just financial aspects such as sales and profits. Interviews with German hotel managers also show that customer satisfaction, employee feedback and financial indicators are the most important indicators of innovation success.

3 FOOD WASTE

To answer the question “how can IT systems support business processes in a cafeteria and thus reduce food waste” it is important to know more about food waste and where it occurs. This chapter will give an overview of the food waste in general and the occurrences in the gastronomy.

3.1 Food pyramid and products

Before elaborating the topic food waste, it is necessary to understand the term food and how it is used in gastronomy. Food is important and has a high influence on our health. A healthy diet, with the right exercise, rest and relaxation is essential for a better quality of life. The food pyramid in Figure 2 recommends a balanced diet for adults. The foods in the lower levels may be consumed in larger quantities, the foods in the upper levels rather in smaller portions. With the right ratio, a balanced and healthy diet is guaranteed. (Egli, 2016, p. 90)

Figure 2. Food pyramid (Egli, 2016, p. 90)

One to two litres of water and other unsweetened drinks such as fruit or herbal teas should be consumed on a daily basis. Vegetables and fruits provide important vitamins, minerals, food fibres and secondary plant substances that are important for a healthy diet. With a varied selection of fresh vegetables and fruit, the body is supplied with healthy ingredients. Grain products, potatoes and pulses, together with carbohydrates, are

essential sources of energy for the body and its muscles, brain and other organs. Milk and other dairy products such as cheese, yoghurt provide the body with important proteins and other nutrients such as calcium. The daily protein requirement can be met with about three portions of such products. Fish supplies the body with omega-3 fatty acids. Furthermore, meat and eggs are necessary iron suppliers. Meat can also be replaced by tofu, seitan, or Quorn, which are valuable vegetable protein sources. Fat is not only an important flavour carrier, it is also the most energy-rich nutrient. Fats should be consumed sparingly as they provide twice as much energy as carbohydrates or proteins. They are important and provide the body with essential fatty acids and fat-soluble vitamins. Oils, fats and nuts should be consumed daily in small amounts. Sweet, salty and alcoholic foods provide a lot of energy in form of calories, but should only be consumed in small amounts. In all levels of the pyramid the ecological aspect is important and therefore it is essential to prefer seasonal and regional products before others. (Egli, 2016, pp. 90–92)

3.2 **Definition and classification of food waste**

Food wastage can be divided into food loss and food waste. (FAO, 2013, p. 9) According to the FAO (2018a) food loss is defined as any food that is lost in the supply chain. This might be caused by pre-harvest problems, during the storage, packing or transportation to the market. Food waste is caused in many ways, for example by sorting operations, because it is not of optimal shape, size or colour, or because food is created in large quantities and leads to an unused left over. The biggest food wastage occurs in agricultural production with 33%. This means, one-third of all originally created food will be thrown away before it reaches the consumers.

Cereals with wheat, rye, oats, and rice, but also vegetables, starchy roots and milk and eggs are the biggest causes of wastage in agricultural production. (FAO, 2013, p. 12) Especially the food waste caused by rice in Asia is enormous. The wastage of meat, on the other hand, is in all regions on the world comparatively low. (FAO, 2013, p. 58) More recent statistics by FAO (2016) claims that fruits and vegetables are wasted the most.

Moreover, the industrialised as well as the developing countries waste almost the same amount of food. However, the consumer in Europe, North America, Oceania and industrialised Asia waste much more food than the citizens from Africa, Central-, South- and Southeast-Asia, and Latin America. FAO (2018b) carried out a project in Finland with different institutions to analyse the food waste in the country. Day-care centres, hospitals and centres for the elderly are the most wasteful facilities in Finland. 27% of waste is food waste. According to the results, fast food restaurants are the most efficient because their waste consists of only 2% kitchen waste, 2% service waste and 3% scraps.

Figure 3 shows the food service process which was further analysed by Pirani & Arafat (2016, pp. 130–132) to analyse what happens to the raw ingredients which are produced and where the food waste occurs. Food waste can therefore be classified into two categories: edible and inedible food. Inedible food waste is created anyway, such as the bones of chicken or meat. Edible food waste is technically still eatable but not allowed to sell it anymore for reasons like safety and hygiene. So the waste from guest plates and serving dishes (the buffet) contains edible and inedible food waste. The preparation waste is only considered as inedible food waste which is caused in the kitchen during production through the preparation like peeling.

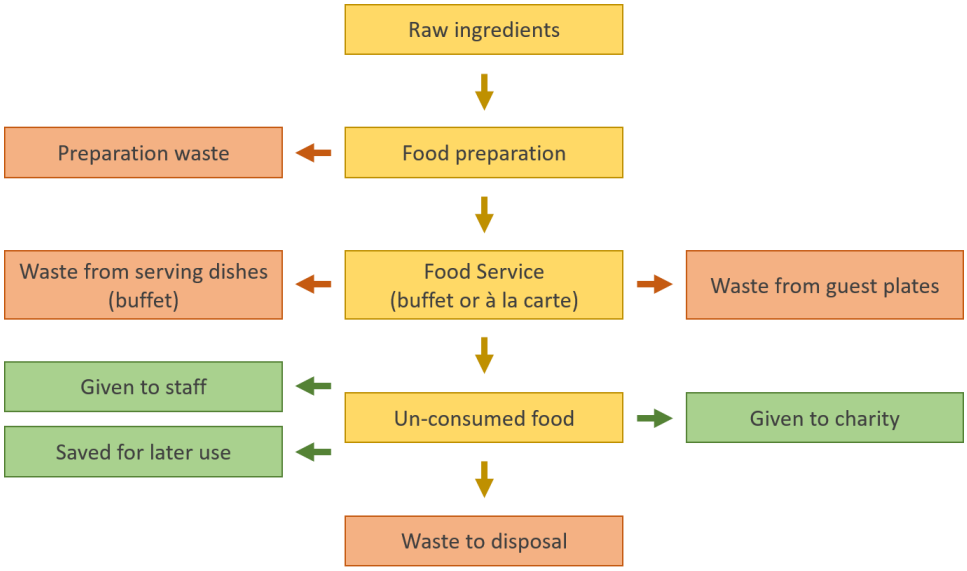

Figure 3. Food Service process (Pirani & Arafat, 2016, p. 130)

The next two figures show a Sankey diagram where the input food is displayed as a flow to the output. Figure 4 shows the generated food waste from the “à la carte” style of service and Figure 5 shows the flow of food from the “buffet” style of service at lunch. Both data are gathered in the same hotel. (Pirani & Arafat, 2016, p. 135)

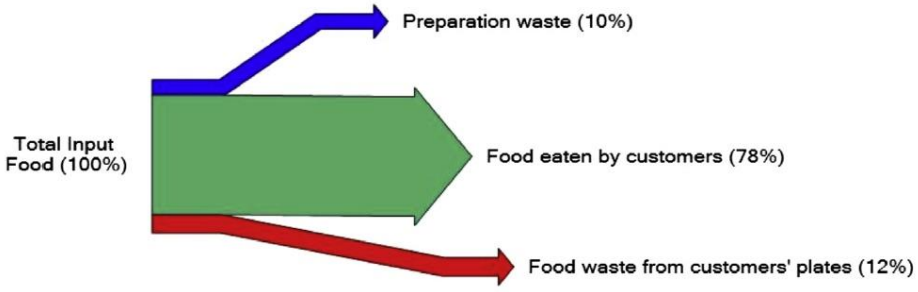

Figure 4. Sankey diagram a la carte (Pirani & Arafat, 2016, p. 136)

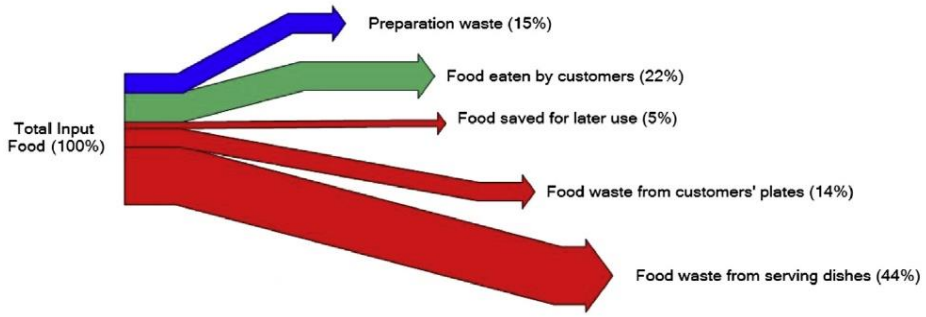

Figure 5. Sankey diagram lunch buffet (Pirani & Arafat, 2016, p. 136)

The material flow analysis clearly indicate that the “buffet” style of service generates more waste than the “à la carte”. With 44% food of the total input as raw ingredients wasted from serving dishes (the buffet), the buffet is generating a significant amount of waste compared to the other outputs or even other style of service. (Pirani & Arafat, 2016, p. 135)

Consumers in a buffet-style restaurant like a cafeteria should be made aware of only take as much food as they can eat with good-visible information signs. Instead of let the customers take the amount they want, make portioned plates to make sure they do not take too much. Even removing the style of “all you can eat”-buffet with “pay by weight”-systems have a positive effect of reducing food waste. (Priefer, Jörissen, & Bräutigam, 2016, p. 163) The results from Pirani & Arafat (2016, p. 144) shows that the amount of served dishes and the accuracy of the forecasts, how many guests will attend, has a significant effect on the amount of generated food waste.

3.3 **An ethical movement to more sustainability**

The gastronomy is a raising industry. For example in urban areas of China the industry is developing in a rapid pace and therefore produce even more food waste each year. The Ministry of Commerce in China expects millions of tonnes per year of commercial food waste. All over the world regulations as such from the European Union’s Waste Framework Directive 2008/98/EC are adopted to reduce food waste. These regulations specify an order of preference of waste. The one from the European Union for example recommend prevention, preparing for re-use, recycling, other recovery and treatment. The adapted Belgium standard propose even more stringent hierarchy in which they mention that it should first be considered as human nutrition or animal feed, or that the waste should be used as fertiliser by anaerobic digestion or composting.

The production of renewable energy is also mentioned as a possible use of food waste. With anaerobic digestion heat, power, and natural gas can be produced, which can be used in the agricultural process or even from

the restaurant or hotel itself. For such a treatment to produce biogas, specific knowledge is needed and the food waste has to be processed more efficient to make a successful amount of output.

The government needs to establish suitable policy frameworks and because they are concerned about the food waste, the future for such treatments looks good. (Wen, Wang, & De Clercq, 2016, pp. 88–95) According to Pirani & Arafat (2016, pp. 133–134) who analysed the hospitality industry in the United Arab Emirates shows that the treatment of food waste in all kind of hotels is also a matter of costs. 57% of the hotels which do not compost its food waste, do not have a composting machine because it is too expensive. Besides that, almost 50% of all hotels work together with charities donate the uneaten food to them. This seems, that companies in the hospitality industry prefer to safe money than do something with the waste. But, it is not financial attractive to throw the waste on landfills. 18 European Member States applies taxes on waste sent to landfills. Some European countries even provide financial support for the production of energy from waste, which makes the right treatment of food waste more attractive. (Priefer et al., 2016, pp. 162–163)

Future trends show there is potential for food waste reduction. With the increasing population and the impacts of climate change increase the economic value of food. Thus, more efficient processes are tried to be implemented, shared logistics with collaborative warehousing, and better demand forecasting may lead to less food waste. Not only these technological innovations are needed, but also the cultural shift, that retailers, food services and consumers value food differently. (Parfitt, Barthel, & MacNaughton, 2010)

4 SODEXO

This chapter is about the cafeteria at HAMK's campus in Hämeenlinna run by Sodexo. It will give an overview of the cafeteria chain and its business processes. Additionally, some opportunities will be explained how Sodexo's business processes can be supported by IT systems.

4.1 Overview of the cafeteria chain

The operator of the cafeteria is the Finnish company Sodexo Oy which is part of the international Sodexo Group. With 11,000 employees in Sweden, Finland, Norway and Denmark and an annual turnover of € 800 million, it is one of the leaders in service management in the Nordic countries. (Sodexo, 2019a) With over 120 restaurants Sodexo is operating in different companies like HTC, at the Airport Helsinki-Vantaa, in universities and universities of applied sciences (Sodexo, 2019c). Three of the seven Häme University of Applied Sciences campuses have a cafeteria operated by Sodexo (HAMK, 2019). According to HAMK's risk manager Mikko Tuomas (interview 17 January 2019) the campus in Hämeenlinna has issued a tender for the cafeteria operation, which Sodexo has won.

Now they operate the largest cafeteria on the campus since 1 January 2018. The cafeteria at HAMK's campus in Hämeenlinna offers healthy and varied breakfast, lunch and snack. Also, cold and hot beverages are offered by the cafeteria. Six different dishes are normally offered every weekday, two popular lunch meals, one vegetarian, one vegan, one soup and one special dish which is a bit more expensive. With a side salad, some bread and water, milk or juice the students and other guests get a wide variety of food every weekday. (Sodexo, 2019b)

4.2 The cafeteria and their business processes

Sodexo's executive chef Janne Vierimaa (interview 11 January 2019) and his team in the kitchen is responsible for the food offered in the cafeteria. The planned menu has a six week cycle. Every six weeks, the dishes on a weekday are the same. This menu is provided by the cafeteria chain which is systematically organised and managed. Mr. Vierimaa has the opportunity to make adjustments in the menu but most of the time this is not done, since the menu has the needed variety for a balanced and healthy diet.

The kitchen produces the food always one day beforehand to minimize stress situation. To produce beforehand they need a good planning, how much food is needed on each day. Therefore they work together with HAMK's risk manager Mikko Tuomas (interview 17 January 2019) to get the expected amount of students each day. The provided data give the

amount of students based on the enrolled modules and the reserved rooms each day. If additionally a company event or another event held by HAMK takes place at the campus the kitchen will be informed as early as possible, to take the additional amount of guests into account in the food calculation.

All the needed ingredients for the production have to be ordered approximately one week before. This is caused by the supply dates on Monday, Wednesday and Friday and the needed calculation and purchase beforehand. Figure 6 shows the process in days of a normal purchase for Thursday's Lunch.

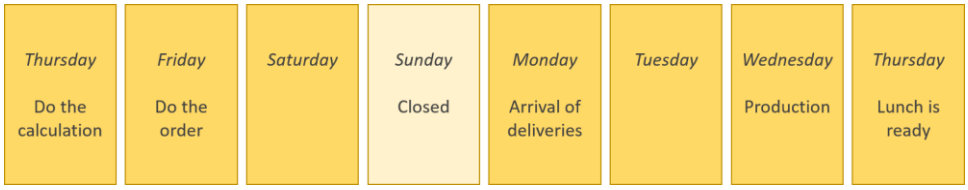

Figure 6. Timeline of a Thursday's lunch production

4.3 Law of nutrition – the government and subvention

The lunch for enrolled students costs € 2.60 and for other guests the not reduced price is € 8.90. This includes the side salad, a menu, bread and butter and something to drink like milk, juice or water. (Sodexo, 2019b) For each student Sodexo receives subsidy by Kela. Kela is a Finnish organisation, which offers financial assistance for students. They also offer different services for families with children, unemployed citizens, pensioners and more. (Kansaneläkelaitos - The Social Insurance Institution of Finland, 2018b). The independent social security institution is supervised by the Finnish parliament (Kansaneläkelaitos - The Social Insurance Institution of Finland, 2018a). The meals subsidy for each student's meal is € 1.94 which the cafeteria operator of higher education studies receives from Kela (Kansaneläkelaitos - The Social Insurance Institution of Finland, 2018c, p. 2).

4.4 Opportunities

The hospitality industry in general is not very well developed in terms of digitalisation. This creates a lot of opportunities for IT systems to fill this missing gap and offer different solutions to support the business. (Gandhi et al., 2016) The following opportunities shows how Sodexo's cafeteria in Hämeenlinna can benefit from such IT systems.

The cafeteria needs to produce in low costs as possible to sustain on the market, but still needs to meet the guest's needs for a good and hungry-

satisfying lunch. To achieve this, the executive chef Janne Vierimaa (interview 11 January 2019) needs to know how many guests will come to eat each lunch. The information they get from HAMK how many students should be at university are very inaccurate and it is needed to improve this forecast to produce the right amount of food. Barrows & Powers (2009, p. 200) claims, that universities have a fair opportunity to predict how much food needs to be produced, as the number of students should not vary so much every day. Many chefs prefer to just rely on their experience to predict the amount of food which has to be produced. However, computers with predictive analytics can support them even by considering aspects like past guest patterns, size of the offered menu, external factors such as other restaurants located nearby. To under-estimate the needed food, and have no more dishes ready for guests, is a risk no one in this business is ready to take. (Pirani & Arafat, 2016, p. 137) According to Janne Vierimaa (interview 11 January 2019) this happened two times in the year 2018 and is unacceptable to happen again.

With gathered data from a point-of-sale system (POS), like Oracle is providing one to Sodexo's cafeteria, reports and analysis can be made. The POS is a cashier system where the guests pay for the lunch. It offers the needed functionalities to enter what menu the guest took and is handling the payment with cash or other supported payment methods like credit cards. (Oracle, 2019) With the historical data of the POS, weather conditions and other key parameters the menu planning can be improved. For example it can be analysed whether the temperature and season has impact on the guests food selection. (Priefer et al., 2016, p. 163)

The cafeteria could offer different specials to shift the demand from a high volume weekday to a low volume weekday. With two-for-one coupons or loyalty cards, where the fifth meal is half price, the cafeteria has different opportunities to directly influence the behaviour of the guest. (Tranter, Stuart-Hill, & Parker, 2014, p. 100) As most of the guests are students, the cafeteria should be able to deal with them. Students might be very demanding and have different opinions which they might share with the cafeteria. With gathered feedback from them or even monthly meetings with student's organisation an open communication can help the operators to improve their offer. (Barrows & Powers, 2009, pp. 202–203)

It does not make economic sense to fully automate the restaurant and all processes from serving the guests, payment, and the whole food production. But, technology might support this business in different ways and each operating business has to consider their opportunities and risks in this topic. (Barrows & Powers, 2009, p. 253) And in the matter of reducing food waste the training of staff and the improvement of internal routines for all processes like purchasing, storing, freezing is essential to success. With the collection of food waste data, vulnerable points of the restaurant can be identified and therefore prevented. (Priefer et al., 2016, p. 163) The cafeteria already uses a food waste solution to weight the

waste where it occurs daily. The waste will be measured and with a tablet entered into the system, which offers different reports and analysis about the amount of wasted food. (Lassila & Tikanoja Oy, 2017)

Executive chef Janne Vierimaa (interview 11 January 2019) claims, once food is produced in a too big amount, the leftovers can currently still be sold one day after, according to the Finnish hygiene and safety regulations. It is also possible to take some food additionally home in the evening for a fair price. Sodexo might consider trending mobile applications similar like “Too good to go” (Too Good To Go ApS, 2018), where restaurants can participate to offer their food to the community for a reduced price. That would bring Sodexo more customers and thus more revenue for food that is otherwise thrown away.

5 THE APPROACH TO ACCURATE PREDICTIONS

As claimed by Pirani & Arafat (2016, p. 144), a key factor in the creation of food waste is the accuracy of predictions about how many guests will attend. Due to the possibilities that Sodexo is able to take advantage of, special attention are given to predictive analytics and the related topics in the current chapter.

5.1 Data science

With data science, quantitative and qualitative methods are applied to solve relevant problems and predict future outcomes. The enormous growth of data which are available today to make different data analysis is an enabler to achieve accurate analysis. The domain knowledge, on the other hand, is needed for effective analysis. This approach is not easy and especially not learned fast. So the balance between the analytical capabilities and domain experience is essential for success stories. (Waller & Fawcett, 2013, p. 78) And the skills a data scientist needs for this are illustrated in Figure 7.

Figure 7. Skills of a data scientist (Kelleher & Tierney, 2018, p. 20)

A main task in data science is the extraction of different patterns. Not patterns which humans can create easily in their mind, but patterns which are too complex to identify manually and which are worth the time and effort invested to discover them. Identifying groups is called clustering in data science terminology, and it is applied with an association-rule mining process. With the search for patterns among hundreds or even more attributes the goal is to get relevant information about the problem that is not obvious. (Kelleher & Tierney, 2018, pp. 1–5)

To achieve this, the technologies for big data needs to be mastered. Not the amount of the data is essential, but the right and important data and the most efficient method to get results in a reasonable time are the key factors to get a meaningful data analysis outcome. (Said & Torra, 2019, p. 2)

5.2 **Big data**

Big data is a large amount of data which are stored and processed in order to analyse them. Since the innovation in the information technology (IT), the investment in such trends is tremendous, because of the impact on productivity and competitiveness. The companies are outperforming each other. Even startups have to familiarise themselves with the advantages of big data to compete and attain real value out of them. It is not only about the big amount of data, but it is also about the high frequency and real-time nature of the data. So the management has to concern to get the right data as fast as possible. This can be well illustrated by the example of sensors that evaluate various conditions in a production plant. But also other industries like healthcare, real-estate or government use big data to create value by improving their operational efficiency. (Zhang, 2017, pp. 20–22)

5.3 **Data mining**

Data mining is the process of discovering useful patterns and trends in large data sets (Larose & Larose, 2015, p. 4). The data mining analysis follows the purpose of commercial applications (Kelleher & Tierney, 2018, p. 1). With describing, estimating, predicting, classifying, clustering, and associating data mining accomplish a broad variety of tasks. The tasks estimation, classification and prediction are very similar and are part from predictive analytics. (Larose & Larose, 2015, pp. 10–17)

5.4 **Predictive analytics**

With predictive analytics as a subset of data science, the main focus is on statistics to forecast or optimise. Additionally to forecasting, predictive analytics is about adding questions regarding how past events could have been influenced by different conditions. In the subject of optimising

predictive analytics concerning system characteristics that were not operating optimally. The growth of data volume and diversity leads to data sets that are larger than those that can be managed with traditional, practical management tools. To manage such potentially valuable data sets, new methods in form of predictive analytics have been developed. (Waller & Fawcett, 2013, pp. 77–80)

Predictive analytics refers to other disciplines like statistics, machine learning and artificial intelligence which are used to discover such patterns. It is data-driven which means that the algorithm generates models from the data. They are not influenced by assumptions made by analysts or data scientists. So, for example, the decision tree which is an often used predictive analytics algorithm identifies the inputs to predict the target variable. (Abbott, 2014, p. 3) To get this description short, Larose & Larose (2015, p. 4) has a very precise definition of predictive analytics. Predictive analytics is the process of extracting information from large data sets in order to make predictions and estimates about future outcomes.

To identify which model should be used for data analysis, the data needs to be understood first. The data can be numeric, ordinal or nominal. Numeric values are integers or real values which are measurable on an interval scale or a ratio scale. The ratio value indicates that the value 0 does not mean there is nothing, like the interval scale does. So for example temperature is a ratio value, because 0 degrees does not mean zero heat. Nominal, also known as categorical, attributes have values from a specific category like an author, or a specific dish and can therefore not be ordered. It can be sorted alphabetically, but this is a distinct operation from ordering. Ordinal attributes are similar to nominal, except they can be ordered. So if the weather is sunny or rainy is an attribute based on a category but it can be sorted and thus have a weighted influence on the result. (Kelleher & Tierney, 2018, pp. 42–46)

Two different approaches are possible to do predictions. One is the classification where the goal is to predict a state, for example by a method like a decision tree. The second approach is to predict a specific number by regression. The prediction model of classification returns a label or a category as a result. This requires historical data where each instance indicates if the target event has happened. Therefore the model considers two different time periods. One is the observation period, where the input values have been created, and the other is the outcome period where the target is calculated. And based on the data from the observation period this model calculates the likelihood the event will happen in the outcome period. With different machine learning algorithms, which trains to build a classification model, the most accurate one can be tested on a not considered subset of the dataset to validate the suitability for the business needs. The second model, regression analysis, is structurally similar to the first explained classification model. The difference is, that it is not a value

from a categorical attribute that is estimated, but a value from a continuous attribute. Therefore, the estimated target attribute must be part of the historical data and with this corresponding data, the regression algorithm defines the influence of each attribute on the output. (Kelleher & Tierney, 2018, pp. 151–180)

5.5 Cross-industry standardised process for data mining

The cross-industry standardised process for data mining (CRISP-DM) is a comprehensive data mining methodology and process model which is based on best practices. It supports organisations to realise better and faster results from data mining. The model was conceived in 1996 while data mining was not widely used. With this standard, the industry developed a non-proprietary, documented and free available model which got improvements over the years and can be specialised for any industry or company where it is used. (Shearer, 2000, p. 13) The process, shown in Figure 8, will here be described with the containing phases and will be applied in the realisation of the Sodexo's cafeteria case.

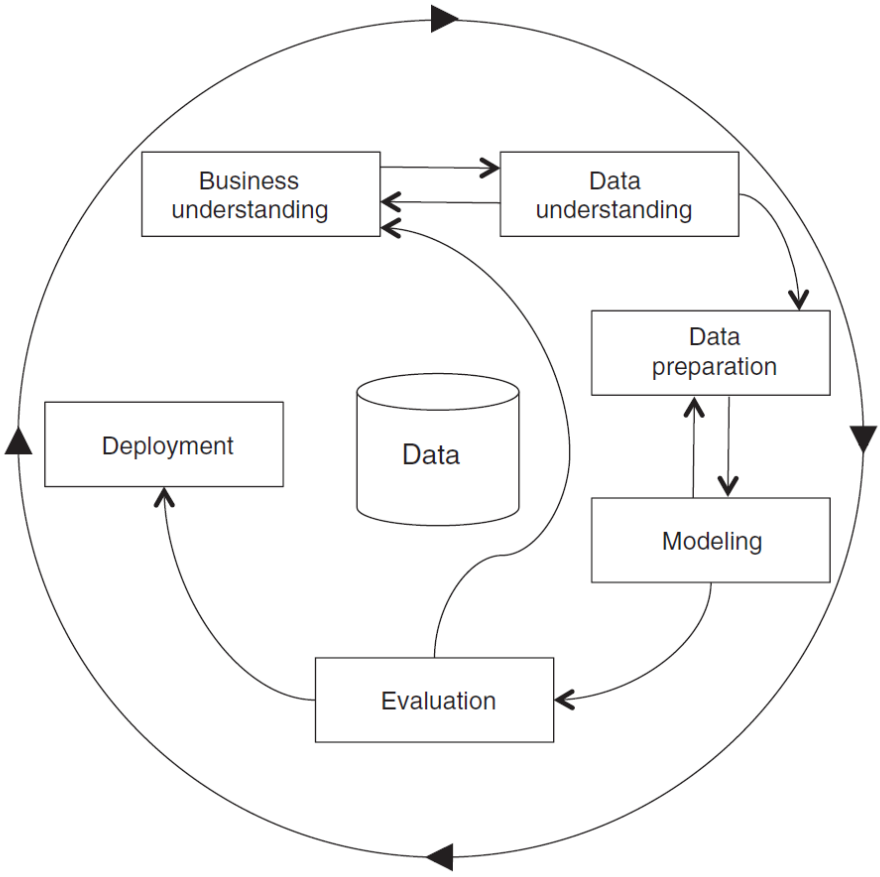

Figure 8. CRISP-DM model (Shearer, 2000, p. 14)

According to Shearer's CRISP-DM model (2000, pp. 14–15), the first phase in the process is the business understanding, and it focuses on the project

objectives and requirements from a business point of view. These goals will be specified, and the plan to achieve those will be described within this process as well.

In the phase of the data understanding the goal is to get familiar with the data, identify possible quality problems and already discover some first insights. Additionally, some subsets from the data could already be defined. (Shearer, 2000, pp. 15–16)

In the data preparation of the CRISP-DM process, the data has to been prepared for the modelling afterwards. Which means they are getting transformed, cleaned, merged or formatted to satisfy the requirements for the specific modelling tool. (Shearer, 2000, pp. 16–17)

The modelling phase is the main part of the data analysis. The data will be analysed to achieve the defined project goals. Thus a model technique will be established and described how the data are analysed. Since not every goal can be achieved with the same model, it is also possible to use different model techniques which are supported by the available tools and general conditions. (Shearer, 2000, p. 17)

After the creation of the models, it is now time to evaluate them and figure out if the business objective can be properly achieved. In this evaluation phase of the process, the main goal is to determine that the defined goals are achieved and how it is done. Therefore the models itself and the findings out of them are used to answer the initial questions. (Shearer, 2000, pp. 17–18)

The last phase of the CRISP-DM process is the deployment. The goal is to define the actions to make use of these models and how the analysis can be done again to another date of time (Shearer, 2000, p. 18).

6 METHODS IN REALISATION

During the implementation of the project, a data analysis for better menu calculation and more efficient production by predictive analytics will be performed. Therefore, the following methods from project management, and the style of communication will be described in this chapter. Findings, conclusions or outputs from these methods have an influence on the realisation and the theory and will therefore flow into this thesis paper.

6.1 A lean approach in agile software development

For the implementation, a lean approach in agile software development will be applied. There are multiple agile techniques like Scrum, eXtreme Programming, Feature-Driven Development and even hybrids that are widely adopted in software development. Methods like Scrum have some restrictions which cannot be fulfilled in this realisation, like a team of multiple people is needed, but also agile methods might fail to satisfy on an organisational level. Hybrids like “Scrumban”, which is based on Scrum and Kanban suits better on projects with frequent and unexpected user stories. So to understand the lean approach in agile software development, the two terms agile and lean have to be defined. According to most research, the two different names are the same thing. Which means there is no meaningful distinction to be made. However, there are other literature considering a difference. Initiated by the top of the organisation the agile methods are by tactical nature but do not scale without change. So agile methods suit perfectly for project management that surrounds software development. Lean on the other hand has the objective to optimise activities across the whole organisation. Thus, lean approaches are needed if changes have to be done on a different level than the specific practice of developing software. Lean can be applied to any scope, even the entire enterprise where software development is only a small part. (Wang, Conboy, & Cawley, 2012, pp. 1287–1289)

With lean thinking, it is possible to specify a value and line up their actions in the best sequence whenever needed to perform effectively. With the following principles of lean software development, the objectives of this approach should be clear: eliminating waste, build quality, create knowledge, defer commitment, deliver fast, respect people and optimise the whole.

6.2 Kanban

Kanban consists of the two Japanese words “kan” (English: signal) and “ban” (English: card). The basic idea is to draw attention at an early stage to the dangers of simply organised work and to obtain an overview of the progress of work, in order to achieve the highest possible business value.

(Epping, 2011, pp. 23–25) David J. Anderson who published first about Kanban in 2007 summarises the claim of Kanban as follows: “Value first, then flow, then waste reduction/elimination” (Anderson, 2008). It is about the order of the three main focuses of Kanban. The first one, the value, is about the business value which has to be achieved from every single work. The flow indicates the uniform of work progress. This is important because it can maximise the business value with continuous work progress. But it is subordinated to value, which means that if another work creates more value than adhering to uniform work progress, the other work comes first. And the third part: any kind of waste, which reduces the business value or slow the progress down, must be removed. (Epping, 2011, pp. 25–27) This summary comes from the lean principle of creating value and eliminating waste which is concerned in Kanban by prioritised features and limiting the work in progress. With a visible number of maximum allowed tasks which are in progress and without a fix duration of them, there is a lean approach in this agile method. (Wang et al., 2012, p. 1297)

With Kanban there is an existing approach which combines the agile and lean method. By visualising the workflow, limiting the current work in progress, using explicit process policies, and improve collaboratively some similarities are existing from the lean perspective. With the primary goal to create value and eliminating waste, in terms of work which creates no value, the target is to improve the agile process. With Kanban which has the core strategy “from agile to lean” a suitable approach is existing for the implementation. (Wang et al., 2012, pp. 1288–1297)

Kanban is used in the different phases of a software development project. During the requirements engineering while requirements are collected, documented, validated and managed the Kanban cards are defined by the requirements. These requirements represent a business value that must be achieved through the realisation. With the properties of being independent of others, negotiable, valuable, estimable, small and testable each request has a priority on the basis of which the order of implementation is determined. (Epping, 2011, pp. 69–74)

In the phase of development, the defined requirements will be realised. With small isolated steps, the developers are self-organised and self-responsible to work on the defined feature which creates value. If problems occur, they will be discussed in the team, and all information is transparent. The requirement which was transformed as a feature in the application during the development phase will be validated during the quality assurance. During the quality assurance, the realisation will be reviewed, and it is ensured that the requirements are entirely fulfilled. (Epping, 2011, pp. 74–77)

For tracking the status of each requirement, the collaboration tool Infinity (Infinity, 2019) will be used during the project implementation. With the open, in progress and finished tasks, the progress can be seen all the time

on a Kanban board. The tasks can be further categorised and specified with due dates to make sure that the right requirement will be realised first.

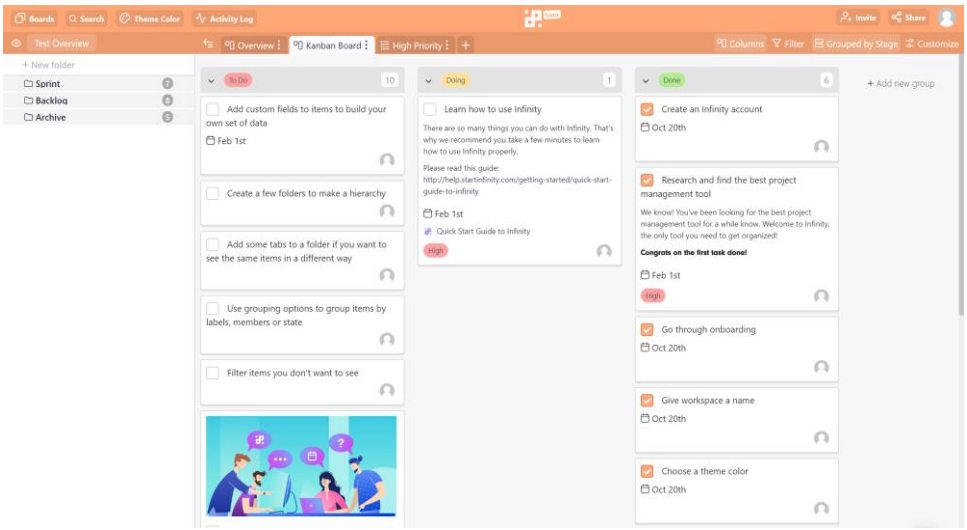

Figure 9. Example of a Kanban board in Infinity (Infinity, 2019)

6.3 **Meetings**

By meetings, a face-to-face meeting is meant. This is distinguished from meetings that could be held via video conferencing because the participants must be in the same geographical position. According to Oshri, Kotlarsky, & Willcocks (2007, p. 28) face-to-face meetings are important for the team development and the positive effects on the collaboration, and team performance enhanced through the interactions between team members.

Such meetings will be held with the commissioning party, Sodexo and also HAMK as the university which helps to provide information. The goal and agenda of each meeting are clear and will be defined beforehand, to have an efficient and successful meeting. Notes are taken continuously and if some agreements have been made, they are distributed to all participants after the meeting.

6.4 **Email**

If team members are involved in the same project and are in different geographical locations and might even be facing different time-zones and cultural differences including different languages, value and norms of behaviour, the term of globally distributed software development team is very appropriate. In such occasions, the communication through face-to-face interactions is challenging. It might be caused by a non-existing high

speed connection. In such cases the most commonly suggested collaborative technologies are email, instant messaging, or conferences through phone or video. (Oshri et al., 2007, p. 28)

During the realisation phase, the development team consists of one single person. There are other people involved, like the representative from the client, the point of contact from the supplier with their IT department. However, these people will not be included in the core team, but information will be shared by email.

6.5 **Diary**

To retrace past activities, a diary is written every day during the realisation. The project started on 11 January 2019 with the kick-off meeting with the client, Sodexo's cafeteria at HAMK, Hämeenlinna. The end of the project was the presentation day of the data analysis to the client on 15 March 2019. During this period, each progress or decision made, got daily a new entry in the diary. These include similar to Scrum's daily scrum meeting the following three questions: what has been done today, what kind of problems were faced today, and what will be done next. According to the agile framework Scrum and how Koch (2004, p. 77) claims the daily scrum is a short meeting around 15 minutes to share the status and concerns of each team member briefly. According to Cohn (2010, p. 382) it is possible to write the meeting down, and do not hold it in a phone call or in a physical meeting, if all team members agree on this form. Since in this implementation only one team member is actively working, the written diary is used as a form of a daily scrum.

7 TECHNOLOGIES AND TECHNIQUES FOR THE ANALYSIS

According to Stodder's research (2018, pp. 3–15) around 82% of the 197 surveyed companies are investing in new technologies for their enterprise business intelligence and business analytics. 78% even invest in advanced analytics like predictive analytics. Different types of technologies are already in use or will be invested in. Typical types are spreadsheets, business intelligence platform, which are on premise or in the cloud or even data warehouses. Spreadsheets are with 85% of the companies the most widely spread technology. Also on premise business intelligence tools are often applied, since it is the second most used technology. In this data analysis two tools from this two technologies will be applied and in this chapter introduced. Additionally needed techniques, like special formulas and features will be described, because they are needed in the data analysis.

7.1 Excel

A spreadsheet can be used for almost everything. It can be needed for business or personal purposes. Most often, budget, financial reporting, inventory or other lists are tracked by grouping text and numbers in a grid or table. Microsoft's Excel offers this functionality in an electronical way, where a workbook, the document, can handle spreadsheets. (Parsons, Oja, Carey, & DesJardins, 2017, p. 39) Excel provides mathematical and statistical functionalities in all versions. Only small changes in the formulas were made since the first version Excel 97 until the most recent Excel 2016. (Carlberg, 2018, p. 1)

Since the most data preparation will be done in Power BI's Power Query Editor only one specific Excel formula will here be described. It is the following index-match formula which is needed to get a value from one column by multiple criteria. The value from the range1 is searched by filtering range2 and range3 by the given attributes A1 and B1. Important in this formula are the curly brackets in the beginning and end. They are applied by entering the formula with Ctrl + Shift + Enter. Because the result is an array, multiple values as a result, and this entering method will therefore handle the result correctly. (Bruns, 2019)

{=INDEX(range1, MATCH(1, (A1=range2)*(B1=range3), 0))} (1)

7.2 Power BI

Microsoft's Power BI is a business intelligence collection to do data analysis. The desktop application, known as Power BI Desktop, is designed to create, share and consume business insights. (Sparkman, 2019) There

are other services, like the Report Server, provided by Power BI which will not further be described, because it is not used in this thesis.

In the data preparation phase of the CRISP-DM the Power Query Editor from Power BI will be used. The Query Editor is needed to load the data from different data sources, like CSV, Excel or even the database directly. Default transformation actions like changing the data type or adding custom columns with formulas will be needed to prepare the data to make them analysable. Other features like un-pivot columns will be applied as well to make multiple records with additional columns based on a matrix. (Iseminger, 2018b)

In the report mode of Power BI different chart types can be used to analyse any attributes of the data. Instead of creating a new column in the source of the data, it makes sometimes sense to use a measure in Power BI. A measure is like a column, but it is used when values from a table should be aggregated instead of adding a new attribute to each record. To create such measures, and solve basic calculations or data analysis problems the Data Analysis Expressions (DAX) is used. It is a collection of functions to calculate or use in any kind of expressions. (Iseminger, 2018a)

8 CRISP-DM

In this chapter the cross-industry standardised process for data mining will be applied to analyse the cafeteria and support them to make better menu calculation and produce more efficient to reduce their creation of food waste.

8.1 Business understanding

The Sodexo cafeteria on the HAMK campus in Hämeenlinna produces daily food waste. This is caused by the fact that the number of guests per day is not exactly known. The chef is currently receiving daily information from the university on how many students could be in the school. This data is based on the room reservation system, in which the number of enrolled students occupying the rooms is known. This forecast is inaccurate, because it varies from the effective sales data.

The aim of this data analysis is to find out how food waste can be reduced. An analysis of historical sales and waste data will identify dishes that are more likely to be thrown away. This allows adjustments to be made to the offered menu. With the analysis of the historical data, a more accurate forecast will be developed to determine how much lunch should be produced in future. The Kanban board in Figure 10 shows all necessary requirements for data analysis with the status before the start of the data understanding phase of the CRISP-DM process.

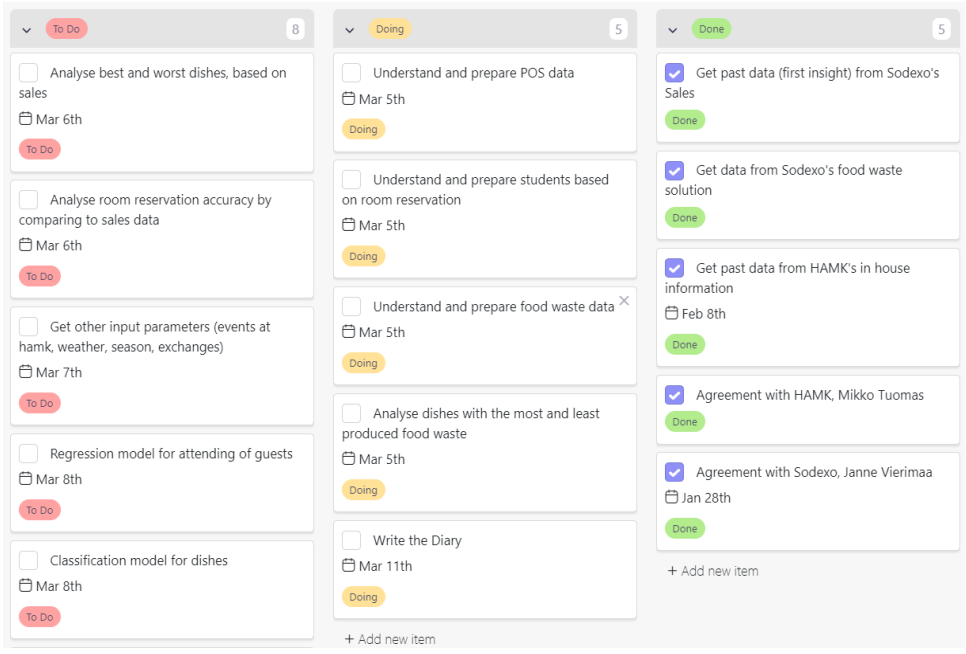

Figure 10. Kanban board of the data analysis realisation

8.2 **Data understanding**

The data required for the analysis are collected from various sources. The sales information from the cafeteria's POS system, the food waste information, and HAMK's student's information based on the room reservation.

The point-of-sale system contains information from all sold dishes each day. If the guest takes food and goes to the cashier to pay, the cashier enters the type of menu in the system. These information will be needed for this analysis. Because the POS system does not have the daily updated menu plan, only the type of menu is known. Therefore the menu from the six week cycle is needed to link the dishes with the amount of sale. Because the system does not offer an export function in a supported file format, which can be used by a business intelligence tool like Power BI, the data are manually entered in an Excel-file. The data are from the time period 29 October 2018 until 21 December 2018.

The food waste solution Hävikkimestari is used by Sodexo's employees to track the generated food waste. With the online solution, analysis can be made on the website with the corresponding login credentials. Thus, the information about the amount of wasted food, and all information around it like the category of food, can be exported in an Excel file. The data are in Finnish language. The solution offers an export function from all entered data by a specific time period. Data from 10 October 2018, when the tool was first used, to 6 February 2019 will be exported to analyse the most current information at the time of implementation.

The university uses a room reservation system to reserve a classroom for the lectures. A module will be attached to this room reservation. And according to the amount of students enrolled in such a module, HAMK makes an assumption about how many students could be on campus. The historical data are saved in a comma-separated values (CSV) file from the internal database to do the analysis without needed access to the database system. The exported data are available from 6 May 2018 until 6 May 2019.

8.3 **Data preparation**

To use the exported CSV or Excel files in the business intelligence tool Power BI, the data have to be cleaned and transformed first. The POS data were manually exported and needs some adjustments. Not all records had the date as a value, because they were entered with the week number and weekday. Therefore the date has to be added to all records first. The Table 1 shows a subset of the data before any other transformation.

Table 1. POS data before transformation

Week	Weekday	Date	Home 1	Home 2
1	Monday	10/29/2018	402	342
1	Tuesday	10/30/2018	635	185

Additionally the data have to be transformed by un-pivot the columns of the sales information. This is needed to get a format of data which can be analysed properly. The un-pivot is done by Power BI's Power Query Editor. The Table 2 shows the same data from Table 1 after the un-pivot.

Table 2. POS data after un-pivot columns

Week	Weekday	Date	Type	Sales
1	Monday	10/29/2018	Home 1	402
1	Monday	10/29/2018	Home 2	342
1	Tuesday	10/30/2018	Home 1	635
1	Tuesday	10/30/2018	Home 2	185

The data do not have the information which dish was sold on the day. Therefore the POS data needs to be connected with the dishes from the six week cycle menu. With the information from the provided menu in Table 3 from Sodexo, the dish name can be connected with the sales information.

Table 3. Un-grouped dishes from menu cycle

Week	Weekday	Type	Dish name
1	Monday	Home 2	Boiled Frozen Peeled
1	Monday	Home 2	Red Bulletin Eh
1	Monday	Bread	Lunching Tomato-Mozzarella Eh
1	Tuesday	Home 1	Mashed Potatoes
1	Tuesday	Home 1	Baked Sausage

Before the dish can be connected. The data from the menu cycle needs some transformation. By grouping the week number, week day, and the food type the multiple row values can be concatenated. Table 4 shows the output data after grouping the dishes and concatenate the names by a comma. The Power Query Editor from Power BI offers the group by function with different aggregation methods for numeric values, like the sum of these. With the advanced adjustment of the aggregation method to a text concatenation method of the grouped values, the menu data contains identical entries only. Therefore the dish can be successfully connected with the POS data.

Table 4. Grouped dishes from menu cycle

Week	Weekday	Type	Dish name
1	Monday	Home 2	Boiled Frozen Peeled, Red Bulletin Eh
1	Monday	Bread	Lunching Tomato-Mozzarella Eh
1	Tuesday	Home 1	Mashed Potatoes, Baked Sausage

With a combination of the index-match formula in Excel, the dish name can be found by multiple criteria and associated with the sales information. Table 5 shows a subset of the final prepared data from the POS for the analysis.

Table 5. Final prepared sales data

Week	Weekday	Date	Type	Dish name	Sales
1	Monday	10/29/2018	Home 2	Boiled Frozen Peeled, Red Bulletin Eh	342
1	Monday	10/29/2018	Bread	Lunching Tomato-Mozzarella Eh	19
1	Tuesday	10/30/2018	Home 1	Mashed Potatoes, Baked Sausage	635

The food waste data exported from the Hävikkimestari, the food waste solution used by Sodexo, contains the data about the generated wastage each day. All values are in Finnish and have to be translated to English with the help of a Google Sheet Add-on from Joey Bronner (2018). With the help of this add-on all information can automatically be translated by Google Translator's API. In the Power Query Editor of Power BI the date has to be transformed to the data type date, and filtered for only records which contains a value, because multiple records does not contain a value. Those records looks suspicious because also other attributes contains no values, therefore they will not be considered in the analysis. The used attributes in the analysis of the data is the date, food category, dish, reason, origin, amount of waste in kg, and the waste value in euro.

The student's information data in the CSV contains the date, campus, room and the amount of students which are enrolled in the module which reserved the room. Because the data analysis only focus on the campus in Hämeenlinna, the data can be filtered by the corresponding campus. Additionally a filter for the existence of an amount of students is applied, because multiple records contains the value "null". The data contains all room reservations, even some with a short period only in the morning or in the afternoon. Because students may not stay for lunch if they finish

class earlier or will not come for lunch if they have class at 2 PM. Therefore, and after checking the data calculation, which Sodexo already receives, only room reservation which lasts over lunch will be considered. By applying a filter, that the start time is less than 1 PM and the end of the reservation greater than 12 PM, the amount of students in house during lunch time is more accurate. A new column with the student amount divided by two will be created. This is by considering a show up rate of 50%, which means, even if 100 students are in school, only 50 will come to eat lunch.

The difference from the student's information data and the amount of sales will be created as a measure in the sales data set. Additionally a measure with the absolute value of this difference will be created, which indicates the accuracy of the forecast.

8.4 **Modelling**

In the modelling phase the answers for the defined goals will be searched and found. How food waste can be reduced and how much lunch should be produced on future days?

To answer the first question, the occurrences of food waste will be identified to make changes in the offered menu. Therefore, the amount of food waste in the pie chart in Figure 11 shows the wastage by category. With around 33% of all wastage in the analysed time period, meat and fish from the main courses generates the highest amount of waste, followed by the category for potato, rice and pasta.

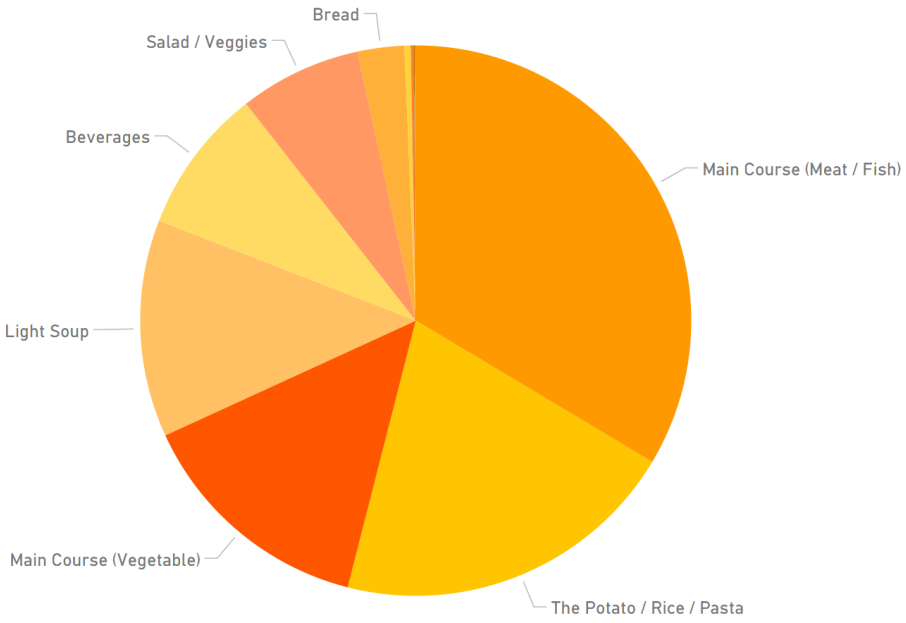

Figure 11. Food waste by category

To get a better insight in the dishes, which generates the waste, the Figure 12 shows a stacked bar chart with the dishes which are thrown away the most. The colour indicates the category. It is not the dishes from the meat and fish category that are thrown away the most, but rice and potatoes.

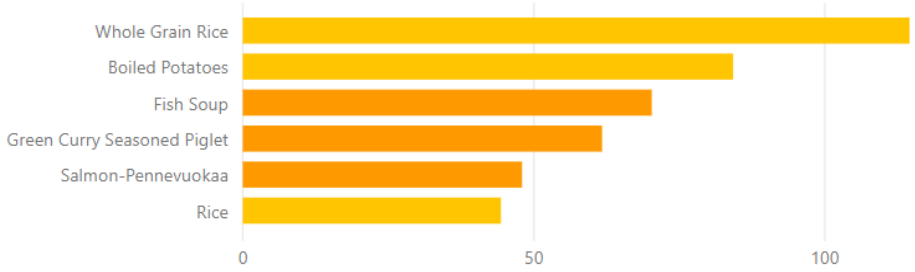

Figure 12. Food waste (in kg) by dishes with category

To understand the amount of food waste, the sales information with the sold dishes needs to be analysed as well. Figure 13 shows the sold dishes per food type. With 75% the two home foods are the most popular dishes from the cafeteria's guests. The insight in the records shows, that potatoes and rice is more used as a side than pasta in the dishes from home 1 and home 2.

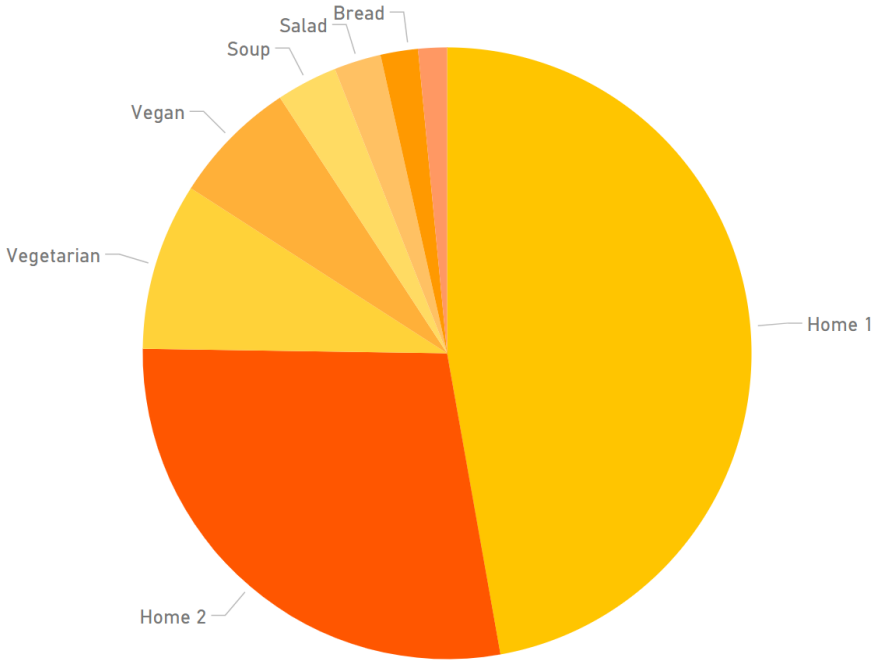

Figure 13. Sales by food type

To get another view on the generated waste. The reason for throwing away indicates to be interesting. The Figure 14 shows the food wasted by origin and reason. The biggest amount is the food which ends in the trash from the guest's plates or from the buffet which is shown as "the line".

Otherwise only around 9% of food waste is caused in the kitchen because of outdated raw material, or because it is not made for sale anymore. The insight in the data records showed, that waste in the kitchen happens not often, but if they do, they have a big impact. So if for example 140 litres of milk has to been thrown away on one day, because it is outdated, the deviation from the average waste per day is significant.

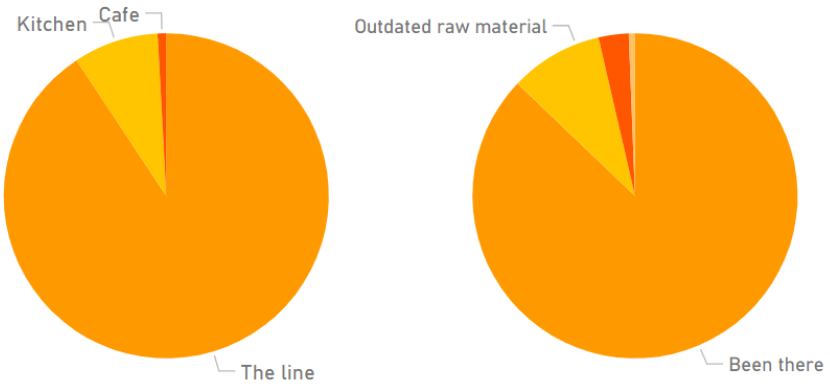

Figure 14. Food waste by origin and reason

The Figure 15 shows a line chart with the food waste per day. The deviation from the mentioned outdated milk is good visible end of October. The average food waste per day is 32.13 kg. The green dotted trend line indicates, that the food waste is decreasing over time.

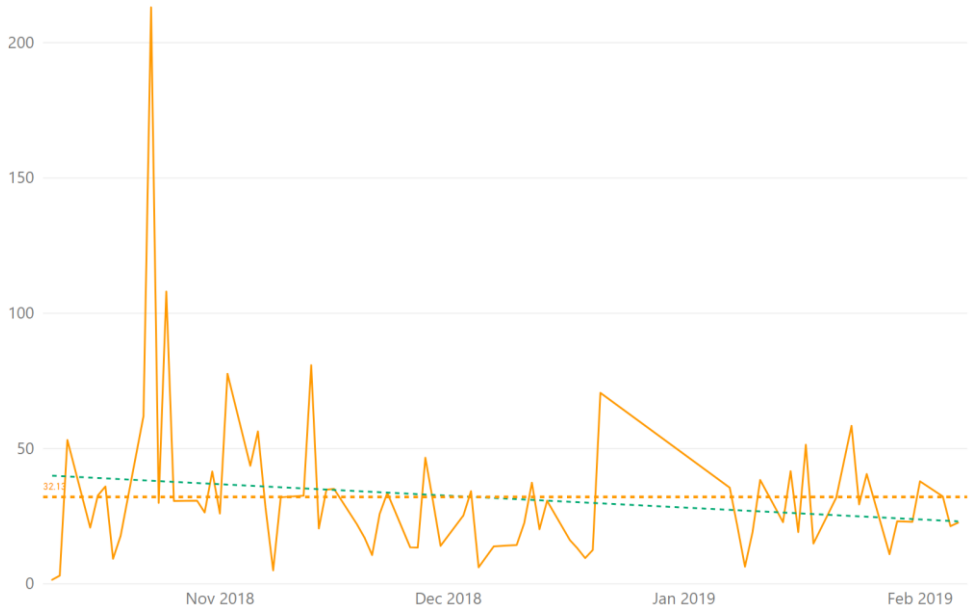

Figure 15. Food waste (in kg) per day with the average and trend line

By a detailed insight in the food waste generated per day an interesting finding is shown in Figure 16, which shows the average food waste in kg

per week day. On Mondays and Fridays the amount of food waste is significant higher than on the other days.

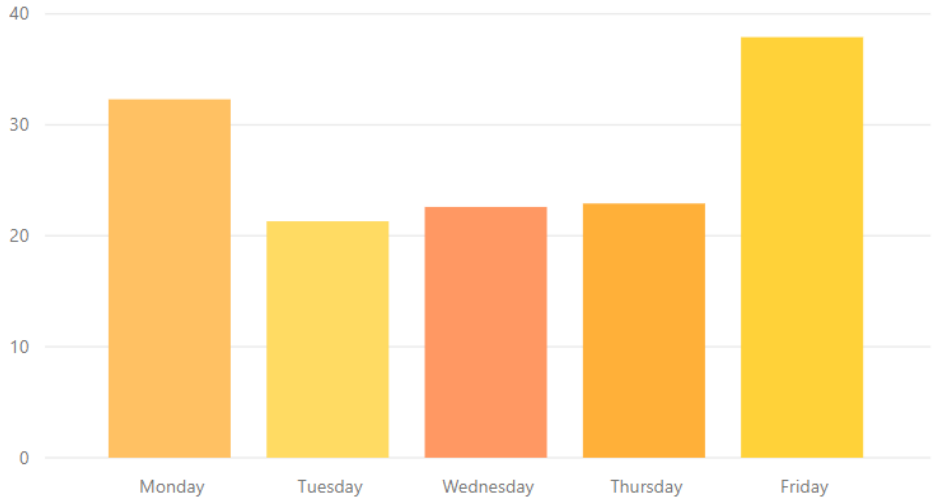

Figure 16. Average food waste (in kg) per week day

By analysing the average amount of sold dishes per week day, the pattern looks different, as seen in Figure 17. It seems, there might be a correlation between the average food wastes with the amount of sales. On Friday, which generates the most food waste, the average sold dishes is the lowest, compared to the other week days. On Tuesday, on the other hand, with the least food waste generation, the most dishes were sold in average. But on Monday there seems to be the exception. On Monday the second most dishes were sold, but the generated food waste is the second highest.

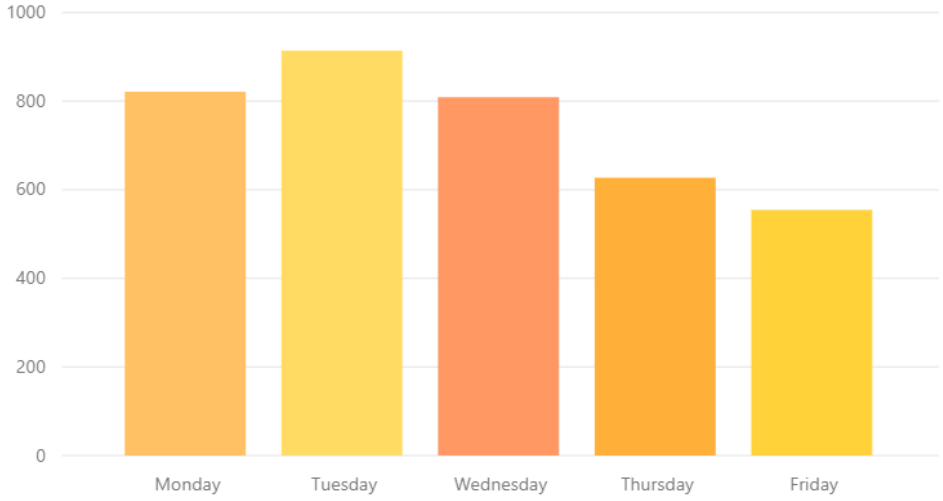

Figure 17. Average amount of sold dishes per week day

The second goal of the analysis is to predict the amount of guests which will attend on a specific day for lunch. Sodexo already got a prediction

from HAMK about how many students will be at the campus during lunch time. But, because this data is only based on the amount of enrolled students which attend a lecture and has a reserved room, the prediction might not be as accurate. For example exchange students, teachers, employees, students who do not attend a class or even other guests are not considered in the data. Figure 18 shows the total amount of sold dishes compared to the new calculated amount of students which should be at the campus. The difference between the prediction and the effective sold dishes is high and might still lead to food waste.

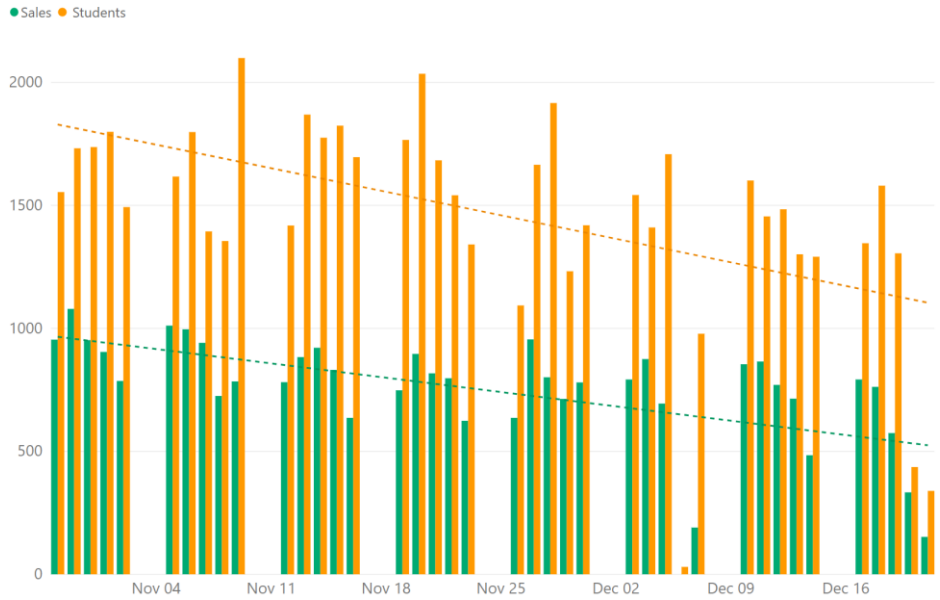

Figure 18. Sales and students per day with trend line

By applying a 50% show up rate to the student's amount, Figure 19 shows the prediction is not as inaccurate as previously assumed.

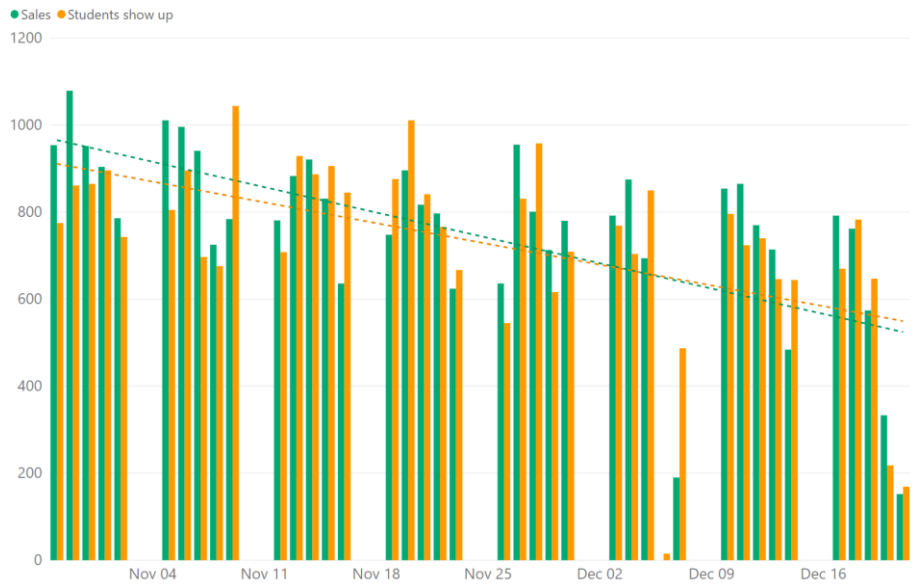

Figure 19. Sales and students (incl. no shows) per day with trend line

The original forecast, how many students should be at lunch, which Sodexo receives from HAMK will now be analysed. An error ratio will be calculated, which is the percentage of the deviation from the predicted value to the effective sales amount. So if the forecasted value is 1,000 guests which will attend and only 800 or even 1,200 dishes were sold on this day an absolute error margin from 200 is given. This would be an error ratio from 20%. The Figure 20 shows this error ratio from the origin forecast which Sodexo received already and the new forecast which is calculated in this data analysis including the 50% show up rate. The old forecast's average error ratio is 30%, while the new one's is 15%. The in this data analysis calculated forecast is more accurate and should be considered in the future.

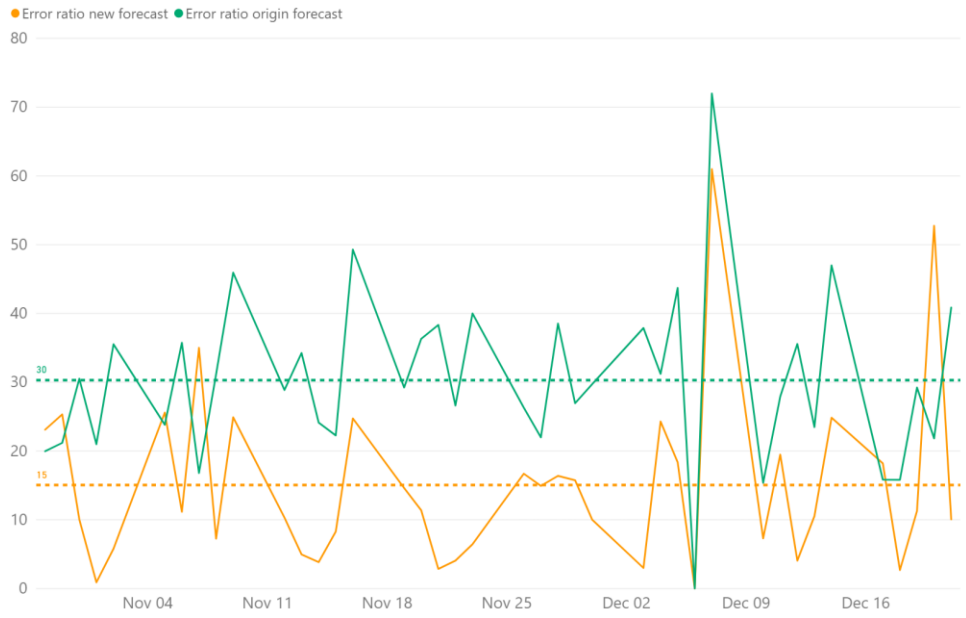

Figure 20. Error ratio of the origin and new prediction

With Figure 21 the linear regression model indicates a negative correlation of the accuracy of prediction to the generated food waste. Which means if the absolute difference from the forecast to sales (x-axis) increases, the amount of food waste (y-axis) decreases. And in more understandable words, this means the more accurate the forecast was, the more food waste was created. This seems confusing, because it actually does not make sense compared to the statements from the executive chef. Mr. Vierimaa claims (interview 11 January 2019), that he needs to know how many guests will come to eat to produce the right amount and avoid the food waste by leftovers.

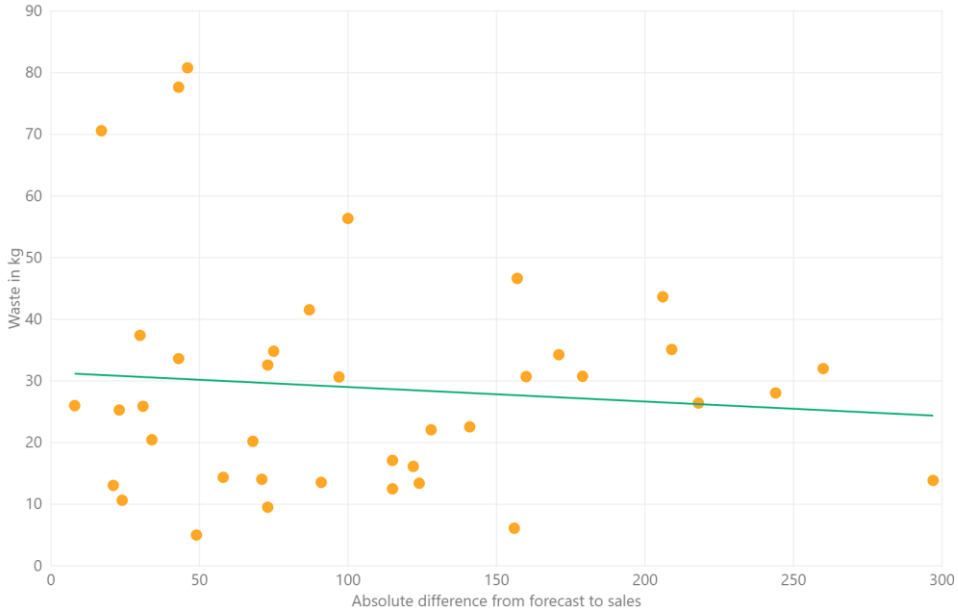

Figure 21. Linear regression of prediction accuracy and food waste

8.5 **Evaluation**

In general the trend of the generated food waste at the cafeteria is reducing. This is a good sign and might be caused by using the solution for tracking the food waste and the therefore already taken actions. On the other hand it might be caused by the short time which was analysed. Only around 20 weeks of the food waste data were analysed and the sales data from even a shorter period. So the question, if the data are accurate and true enough, is justified but nonetheless the data analysis shows impressive insights and the following mentioned actions should be taken to reduce the food waste even more.

The charts in the beginning claims, that the most food waste is generated by the main course from meat and fish. Followed by the sides like rice and potatoes, the waste is produced where it is sold the most. And the home 1 and home 2 are simply the most sold dishes. To mention here as well, on Mondays and Fridays the average food waste is significant higher than on the other week days. The produced amount for the home 1 and home 2 on Monday and Friday should therefore always be considered carefully. Monday's left overs might be sold on Tuesday, as long as the safety and hygiene regulations are guaranteed. On Friday the cafeteria might consider to just produce less or place a special offer, to sell the food before it ends in the trash.

Most of the waste's origin is in the line, which means the buffet or the customer's left over. The kitchen generates food waste as well, but this has no significant impact in the total amount of waste. But, if something

has to be thrown away there, it is a higher amount than average wasted per day. This means, quickly decaying ingredients, such as milk should be purchased more often, but in smaller quantities. This might result in higher purchase costs, but the purchased goods can be fully utilised.

The second part of the data analysis considered the current prediction and gave a more accurate forecast. The linear regression model claims that predictive accuracy has a negative impact on the reduction of food waste. But given the short time span of the historical data and the chef's statements, this makes no sense. The accuracy of the prediction should have a positive effect on the reduction of food waste, as by knowing the number of guests attending, the right amount of food can be produced, thus avoiding waste. Therefore, it makes sense to have the most accurate forecast possible in order to produce less food waste. The new predictions by the room reservation system is fairly accurate by considering a 50% show up rate of the students and comparing the sales data. On average, the error ratio of the new forecast is 15%. After reviewing the analysed data in this data analysis with the current used forecast, which Sodexo receives, differences were found. Both information should be based on the same data, but different output values are the result. The output values in this analysis are more accurate and should be used in the future. A detailed research to find the discrepancy will not be part of this analysis, but will be arranged with the responsible persons.

8.6 **Deployment**

The recreation of this data analysis is at any point of time possible. The food waste solution and the room reservation system will be in operation for the foreseeable future and therefore a data export can be done without creating extraordinary expenses. The POS data on the other hand were created manually and therefore to update the data is time consuming. Before doing further development on the data analysis the data should be exported directly from the cash system. In this way, the data preparation process could be automated in the future.

The Power BI dashboard from this data analysis will not be further distributed nor deployed. The results and findings will be handed to the responsible persons, which will take the needed actions to benefit from this analysis.

9 CONCLUSION

Nowadays the amount of produced food waste is enormous. As mentioned in this thesis, the most wastage is created before customers even get the chance to buy it. Agricultural companies and the society follows regulations and values which leads to the waste of edible food. The hospitality industry with the high amount of restaurants, hotels, and cafeterias in all kind of public services like universities or hospital have a big impact in the creation of food waste. With today's opportunities of digitalisation, companies from all industries apply IT systems in their business to produce more efficient and sell more. The digitalisation in the hospitality industry is developing slower, because the business models are established over decades and the competitive market does not allow any high investments or expenses.

This thesis described the gastronomy industry with an example of the business process from Sodexo's cafeteria at HAMK's campus in Hämeenlinna. Sodexo is already in use of different IT systems like a point-of-sale system or a food waste solution, which is used to track the daily produced waste. A data analysis based on these data and HAMK's information how many students should be in school, is done to further reduce their food waste. Different actions based on the evaluation of the data models were mentioned to reduce the waste. Changes in the offered menu, improve the prediction of expected amount of guests are only one part of the findings. The second part would be to revise the business model and make the guests aware of the situation. By using a "pay per kg" model or to make the guest aware of taking less and going a second time instead of putting as much as possible on a plate, would reduce the customer's waste significantly.

Two questions were asked in this thesis and will now be answered. With systems like point-of-sale, a purchase management tool, tracking food waste and continuous data analysis, the occurrences of food waste can be identified, and therefore actions been defined and taken to reduce the waste generated in a cafeteria. By analysing the sales data, popular and non-popular dishes can be identified and should have influence on the creation of the menu. To use predictive analytics, the data analysis should be done continuously, which means, the data from all systems should be accessible at any point of time. Thus, and with more input parameters, like student's preferences and habits, the forecast gets more accurate over time to make better menu calculations and produce more efficiently.

The practical part in this thesis was a partial success. By analysing the historical data, origin and reasons of the generated food waste were identified and the prediction's accuracy could be improved. However, a sustainable benefit can only be achieved, if the data are available at all times and are continuously analysed in the future.

REFERENCES

Abbott, D. (2014). *Applied Predictive Analytics*. *John Wiley & Sons, Inc.* https://doi.org/10.1017/CBO9781107415324.004

Anderson, D. J. (2008). Providing Value with Lean | DJAA. Retrieved February 3, 2019, from http://www.djaa.com/providing-value-lean

Barrows, C. W., & Powers, T. (2009). *Introduction to Management in the Hospitality Industry*. John Wiley & Sons, Inc.

Bronner, J. (2018). Translate My Sheet. Retrieved from https://chrome.google.com/webstore/detail/translate-my-sheet/ikbffjfknllcipddomhgdhcedikhiaga

Bruns, D. (2019). INDEX and MATCH with multiple criteria. Retrieved March 11, 2019, from https://exceljet.net/formula/index-and-match-with-multiple-criteria

Carlberg, C. (2018). *Statistical Analysis Microsoft Excel 2016 - Rough Cuts*. Que.

CHD Expert. (2019a). Full Service Restaurants Data. Retrieved February 26, 2019, from https://www.chd-expert.com/products/foodservice-data/databases/commercial/full-service-restaurant/

CHD Expert. (2019b). Limited Service Restaurant Data. Retrieved February 26, 2019, from https://www.chd-expert.com/products/foodservice-data/databases/commercial/limited-service-restaurant/

Cohn, M. (2010). *SUCCEEDING WITH AGILE Software Development Using Scrum*. Addison-Wesley.

Egli, C. (2016). *Pauli Lehrbuch der Küche*. (C. Pauli & P. Pauli, Eds.) (14th ed.). Pauli Fachbuchverlag AG.

Epping, T. (2011). *Kanban für die Softwareentwicklung*. Springer-Verlag Berlin Heidelberg. https://doi.org/10.1007/978-3-642-22595-6

FAO. (2013). *Food wastage footprint. Impacts on natural resources. Food wastage footprint Impacts on natural resources*. https://doi.org/ISBN 978-92-5-107752-8

FAO. (2016). Key facts on food loss and waste you should know! Retrieved February 18, 2019, from http://www.fao.org/save-food/resources/keyfindings/en/

FAO. (2018a). Food Loss and Food Waste. Retrieved February 18, 2019, from http://www.fao.org/food-loss-and-food-waste/en/

FAO. (2018b). The Wastestimator project in Finland: a new tool to estimate and monitor food waste. Retrieved February 18, 2019, from http://www.fao.org/save-food/news-and-multimedia/news/news-details/en/c/1094416/

Gandhi, P., Khanna, S., & Ramaswamy, S. (2016). Which Industries Are the Most Digital (and Why)? Retrieved February 13, 2019, from https://hbr.org/2016/04/a-chart-that-shows-which-industries-are-the-most-digital-and-why

Gillespie, C., & Cousins, J. (2001). *European Gastronomy into the 21st Century*. Elsevier Inc.

HAMK. (2019). Campuses and Maps - Häme University of Applied Sciences. Retrieved February 18, 2019, from https://www.hamk.fi/tietoa-hamkista/kampukset-ja-kartat/

Infinity. (2019). Startinfinity. Retrieved February 16, 2019, from https://startinfinity.com/

Iseminger, D. (2018a). DAX basics in Power BI Desktop. Retrieved March 11, 2019, from https://docs.microsoft.com/en-us/power-bi/desktop-quickstart-learn-dax-basics

Iseminger, D. (2018b). Query overview in Power BI Desktop. Retrieved March 11, 2019, from https://docs.microsoft.com/en-us/power-bi/desktop-query-overview

Kansaneläkelaitos - The Social Insurance Institution of Finland. (2018a). Kela as an organisation. Retrieved February 28, 2019, from https://www.kela.fi/web/en/organization

Kansaneläkelaitos - The Social Insurance Institution of Finland. (2018b). Kela benefits. Retrieved February 28, 2019, from https://www.kela.fi/web/en/operations-kela-benefits

Kansaneläkelaitos - The Social Insurance Institution of Finland. (2018c). Student Financial Aid 2019. Retrieved from https://www.kela.fi/documents/10180/681919/Optuinfo_engl.pdf/e1aba890-9bc0-4d77-a7ee-b1a7453a374e

Kelleher, J. D., & Tierney, B. (2018). *Data Science*. The MIT Press.

Koch, A. S. (2004). *Agile Software Development: Evaluating The Methods For Your Organization*. Artech House Publishers.

Larose, D. T., & Larose, C. D. (2015). *DATA MINING AND PREDICTIVE ANALYTICS* (2nd ed.). John Wiley & Sons, Inc.

Lassila & Tikanoja Oy. (2017). Hävikkimestari. Retrieved February 28, 2019, from https://havikkimestari.lassila-tikanoja.fi/palvelu/

National Restaurant Association, & Deloitte. (2010). *Restaurant Industry Operations Report*. National Restaurant Association and Deloitte & Touche LLP.

Oracle. (2019). What is a Restaurant POS System? | Oracle. Retrieved February 28, 2019, from https://www.oracle.com/industries/food-beverage/what-is-restaurant-pos.html

Oshri, I., Kotlarsky, J., & Willcocks, L. P. (2007). Global software development: Exploring socialization and face-to-face meetings in distributed strategic projects. *Journal of Strategic Information Systems*, *16*(1), 25–49. https://doi.org/10.1016/j.jsis.2007.01.001

Ottenbacher, M. C. (2007). Innovation Management in the Hospitality Industry: Different Strategies for Achieving Success. *Journal of Hospitality and Tourism Research*, *31*(4), 431–454. https://doi.org/10.1177/1096348007302352

Parfitt, J., Barthel, M., & MacNaughton, S. (2010, September 27). Food waste within food supply chains: Quantification and potential for change to 2050. *Philosophical Transactions of the Royal Society B: Biological Sciences*. The Royal Society. https://doi.org/10.1098/rstb.2010.0126

Parsons, J. J., Oja, D., Carey, P., & DesJardins, C. A. (2017). *Microsoft Office 365 & Excel 2016*. Cengage Learning.

Pirani, S. I., & Arafat, H. A. (2016). Reduction of food waste generation in the hospitality industry. *Journal of Cleaner Production*, *132*, 129–145. https://doi.org/10.1016/j.jclepro.2015.07.146

Priefer, C., Jörissen, J., & Bräutigam, K. R. (2016). Food waste prevention in Europe - A cause-driven approach to identify the most relevant leverage points for action. *Resources, Conservation and Recycling*. https://doi.org/10.1016/j.resconrec.2016.03.004

Said, A., & Torra, V. (2019). *Data Science in Practice*. Springer International Publishing.

Shearer, C. (2000). The CRISP-DM Model. *Journal Of Data Warehousing*, *5*(4), 13–22. https://doi.org/10.1016/0022-4073(81)90019-4

Sodexo. (2019a). Company. Retrieved February 18, 2019, from https://www.sodexo.fi/yritys

Sodexo. (2019b). HAMK Hämeenlinna Visis. Retrieved February 18, 2019, from https://www.sodexo.fi/hamk-hameenlinna

Sodexo. (2019c). Our Lunch Restaurant. Retrieved February 18, 2019, from https://www.sodexo.fi/ravintolat/kaikki-ravintolamme

Sparkman, M. (2019). What is Power BI? Retrieved March 11, 2019, from

https://docs.microsoft.com/en-us/power-bi/power-bi-overview

Stodder, D. (2018). BI and Analytics in the Age of AI and Big Data. *Transforming Data With Intelligence*.

Toast. (2017). *Restaurant Technology in 2017*. Retrieved from https://pos.toasttab.com/hubfs/Content_+_Assets/Restaurant Technology in 2017.pdf?t=1540484682666&utm_campaign=Restaurant Technology in 2017 Report&utm_source=hs_automation&utm_medium=email&utm_content=57710954&_hsenc=p2ANqtz-_XB1N9rs-KF_0L5

Too Good To Go ApS. (2018). JOIN OUR FOOD WASTE REVOLUTION | Too Good To Go. Retrieved February 28, 2019, from https://toogoodtogo.co.uk/en-gb

Tranter, K. A., Stuart-Hill, T., & Parker, J. (2014). *An introduction to revenue management for the hospitality industry : principles and practices for the real world*. Pearson Education Limited. https://doi.org/10.2307/3211488

Waller, M. A., & Fawcett, S. E. (2013). Data science, predictive analytics, and big data: A revolution that will transform supply chain design and management. *Journal of Business Logistics*, *34*(2), 77–84. https://doi.org/10.1111/jbl.12010

Wang, X., Conboy, K., & Cawley, O. (2012). "Leagile" software development: An experience report analysis of the application of lean approaches in agile software development. *Journal of Systems and Software*, *85*(6), 1287–1299. https://doi.org/10.1016/j.jss.2012.01.061

Wen, Z., Wang, Y., & De Clercq, D. (2016). What is the true value of food waste? A case study of technology integration in urban food waste treatment in Suzhou City, China. *Journal of Cleaner Production*, *118*, 88–96. https://doi.org/10.1016/j.jclepro.2015.12.087

Zhang, A. (2017). *Data Analytics: Practical Guide to Leveraging the Power of Algorithms, Data Science, Data Mining, Statistics, Big Data, and Predictive Analysis to Improve Business, Work, and Life*. CreateSpace Independent Publishing Platform.

INTERVIEWS

Tuomas, M. (2019). Risk Manager, HAMK. Interview 17 January 2019.

Vierimaa, J. (2019). Executive Chef, Sodexo Oy, Hämeenlinna. Interview 11 January 2019.