Integrating Crisp DM Methodology for a Business Using Tableau Visualization

Research · July 2020

DOI: 10.13140/RG.2.2.36619.31520

Akshata Parate
Private University of Applied Sciences

Integrating Crisp DM Methodology for a Business Using Tableau Visualization

E-Business Term Paper for Second Semester

within the framework of the study course “General Management" with the degree "MSc. in General Management" at the PFH - Private University of Applied Sciences Göttingen

submitted on: 10th July 2020

by: Parate, Akshata

from: Göttingen

Examiner: Prof. Hubert Schüle

I. Table of Contents

II. List of Figures

III. Index of Tables

IV. List of Abbreviations

ACE:	Analytics Centre of Excellence (Verizon)
Airbnb:	Air Bed and Breakfast
BI:	Business Intelligence
CRISP-DM:	Cross Industry Standard Processing for Data Mining
ISL:	Integral Solutions Limited
KPIs:	Key Performance Indicators
PERT:	Program Evaluation and Review Technique
NCR:	National Cash Register (Corporation)
SAS:	Statistical Analysis System

1. Introduction

1.1 Relevance and Scope of Research

Today, data science and analytics is used worldwide to empower organizations to make informed decisions and most profitable choices. (Stan, 2020), There are various models used for the purpose of structuring business data by companies. One such model is the CRISP-DM model. This paper focuses on this data mining model and studies how Tableau is used as a part of the CRISP-DM process.

The inspiration of the research topic was taken from the fact that many companies are integrating the CRISP-DM model and also using visualisation tools like Tableau to better understand the past and recent business scenario and also to forecast the future of the sales and revenue conditions. The research was carried out by studying the concepts of the CRISP- DM methodology and how each phase of the cycle is relevant for every step of a project. Then, the benefits of using Tableau for the purpose of the modelling and evaluation phase was studied based on the literature provided by professors and PhD students. Towards the end of the literature review, the limitations of the CRISP-DM model are outlined.

The paper is backed by two case studies in which the CRISP-DM model was applied in their business processes and Tableau was used as a visualisation tool in the modelling phase. Finally, the paper concludes with a summary of the arguments discussed in the paper.

For most businesses, structuring and prioritizing their business projects and gets tedious. By using CRISP-DM as a business analysis strategy, a project follows a clear road map which helps in achieving the desired goal. Therefore, this paper highlights the benefits of using this model and explains how Tableau is the easiest and cheapest way to visualise the otherwise complex data so that the executives or CEOs of the company can make their business decision based on the end results.

1.2 Limitation of Research

The usage of real-world data sets is prohibited by many corporations, because of which the examples used in this paper are limited to historical data provided by companies and student research papers. To explain the process of how Tableau can be integrated with the CRISP-DM model, the paper focuses on case study examples. An advanced research of this topic is possible by business executives and data scientist who have access to the confidential data sets of their companies. Platforms to find data repositories like GitHub, Kaggle, Google DataSet Search, Shopware, etc. are recommended if an individual needs information about well-known companies to use the CRISP-DM model for research or for business projects and has the ability to use Python or R to arrange the data to use the CRISP-DM method. The Tableau Professional software also extends the features and benefits of visualisation if one possesses a license key for the same.

2. Methodology

The aim of this term paper is to understand how the CRISP-DM Model works and how it is used as a data mining methodology in six phases by businesses for their data mining projects. Tableau can be used for modelling data in its process model by the business intelligence team in companies.

The methodology of research for this paper was structured in a horizontal manner as shown in Figure 1. First the background of the CRISP-DM model, the hierarchical structure it follows, and the different phases of the model are discussed. An overview of the Tableau visualisation tool is outlined within the discussion of the CRISP-DM model and an explanation is given of the different functions used, which can be applied in the "Modelling" phase of the CRISP-DM model.

There are two case studies which will be discussed in this paper that prove as examples of how the CRISP-DM model is used by globally known companies like Verizon and Airbnb. They are based on the articles on Tableau and reports given by some business analysts. The benefits of using this model will be discussed, based on their reflections on the model.

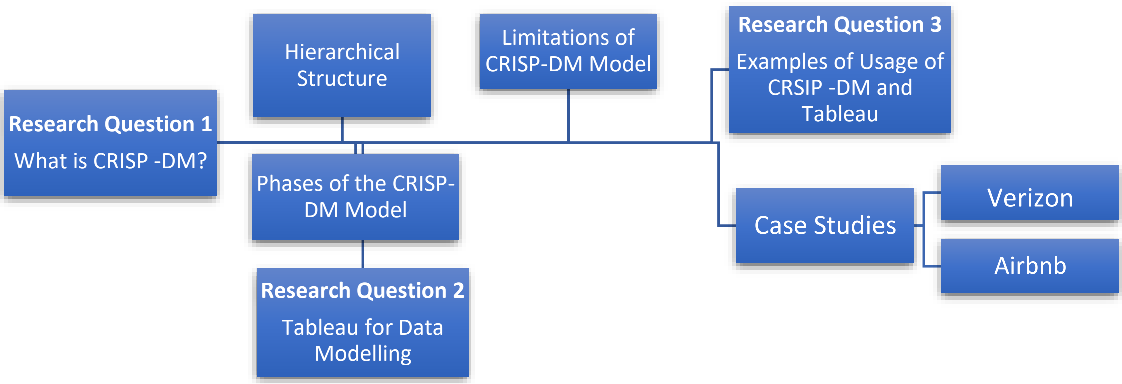

Figure 1: Research Methodology

3. Literature Review

3.1 CRISP-DM Methodology

The Gartner Group and SAS (Statistical Analysis System) define data mining as the process of discovering meaningful correlations, patterns, and trends by sifting through large data sets that are stored in the repository to find and forecast outcomes. The CRISP-DM or Cross Industry Standard Process for Data Mining is a process model within the scope of data mining that is used by data mining experts in businesses. This methodology was created as a collaborative project by 5 companies –Daimler AG, NCR Corporation, OHRA, ISL and Teradata. Today, it is the most widely used business analytics model in the world. (Berns, 2019)

The CRISP-DM portrays a project as a cyclical process, in which all phases are tuned towards achieving the main business goal. The model can be divided into four levels as given in Fig. 1, namely phases, generic tasks, specialised tasks, and process instances, in a specific hierarchical structure that defines each level of abstraction, depending upon the level of effort in each phase. The levels are differentiated according to a specific sequence of events. It should be designed in such a way that in case of repetition of the process in the future, the same model could be applied. (Chapman et al, 2000)

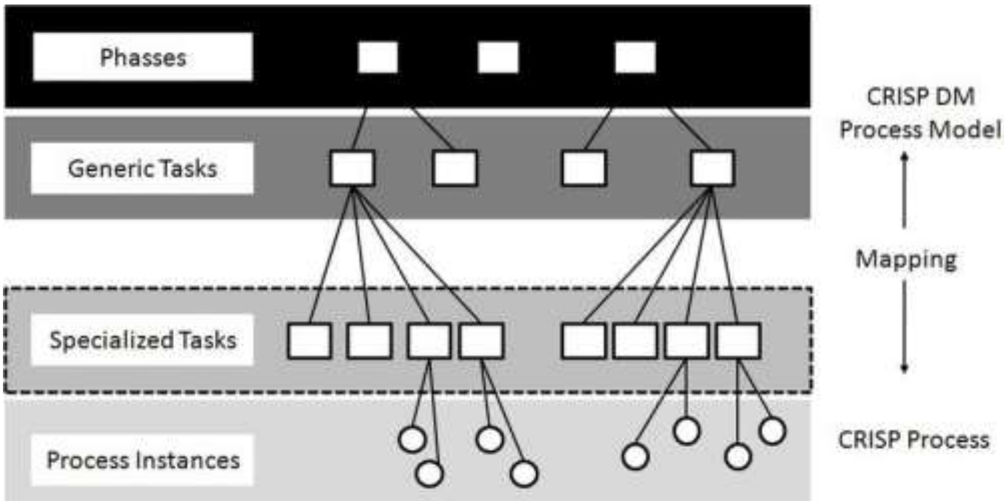

Figure 2: Hierarchical Structure of CRISP-DM Methodology with Four Levels

Source: Pete Chapman et al, The CRISP-DM Process Model

First, the “Phases” consists of the main tasks, which is then followed by “Generic tasks” level, wherein the phases are consisting of second-level or generic tasks that are defined across the data mining process and stable in case of new developments in the model. Third, the “Specialized Tasks” under generic tasks are differentiated according to the specificity of the situation. For example, if the generic task is to arrange the data of a product, one of the specific tasks will be to arrange the data with respect to “quantity sold” or sales of that product in a definite time duration. (Chapman et al, 2000) Fourth is the “Process Instance” level which consists a record of decisions, actions, and results of the data mining process. It is a detailed arrangement of tasks that were defined in the earlier levels. (Wirth et al)

3.2 CRISP-DM Model

Many businesses will already have a set of data regarding the products, production costs, shipping costs, revenue generated from online selling platforms, customer data, etc. If not, the data is extracted from various sources like GitHub or Shopware and then stored in a data repository such as data warehouse or data lake and is further analysed. This is the initial stage of data mining. The CRISP-DM model now comes into the picture, where the company assigns a task to a team, explains the problem it needs to solve or a goal it wishes to achieve. (Chapman et al, 2000)

The model is aimed at making data mining project implementation less expensive, simple, reliable, reusable, and fast. (Wirth et al) It consists of six phases namely, business understanding, data understanding, data preparation, modelling, evaluation, and deployment. The Figure 2 below shows how each phase is followed by the previous phase in a cyclic manner. It is also possible to backtrack to previous phases and repeat actions if required. (Chapman et al, 2000)

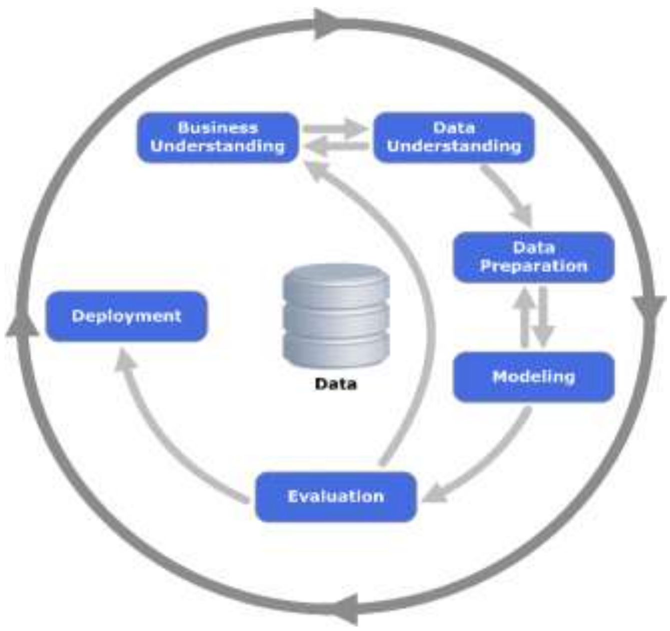

Figure 3: Phases of the CRISP-DM Model

Source: Riepl, Statistik Dresden, 2012

3.2.1 Phase 1 – Business Understanding

When starting a new data mining project, the business intelligence team needs to understand the business aspect of the project. The first phase of the model is where the business objectives are discussed, and goals are set. The next phases and the end outcome should always be in sync with the first phase. (Gupta, p.13, 2014)

The is critical for the management to allocate preliminary budget for the data mining project and appointing the team responsible for handling the project. It is important to identify the areas of the company that are to be analysed, design a business model if necessary, outline the expectations of the analysis and forecast the desired result, make a hypotheses, define the BI tasks of the project and find resources to gather data. An added criterion is to make a time structure plan for project management and task division using Gantt charts or using PERT of the Critical Path Analysis to set a timeline for each phase of the project. (Chapman et al, 2000)

3.2.2 Phase 2 – Data Understanding

In the second phase of the model, the required data is extracted from reliable sources like the GitHub or Shopware platform and by web scrapping. The right data is selected which will determine the accuracy of the hypotheses made in the initial phase. This includes recording the sources and problems while acquiring the data, so that it is reviewable if revisited again. This phase also involves verifying the data quality of the data collected from various sources and reflecting on the business goals to understand the data. Therefore, you can often move back and forth in these two initial phases to understand the data better. (Gupta, p.14, 2014)

3.2.3 Phase 3 – Data Preparation

The third phase of the CRISP-DM model is about data preparation, which includes three sub steps namely selecting, pre-processing and transformation. First is selecting the data for analysis by checking the uniformity, plausibility, and consistency of the data. Then, in the pre-processing step, the data is examined and based on its relevance, it is either included or excluded from the main data which will be analysed. This is also called cleaning the data. The Python data cleaning technique can be used for this purpose with the help of Pandas and NumPy. In Python, data cleaning is done by dropping columns in the "Data Frame", using ".str()" to clean columns and finally using "Data.Frame.applymap()" to clean the complete data set according to the elements in the data set. It also includes removing unrecognizable elements and skipping unnecessary rows in a CSV file. The CSV file is later imported in a data visualisation tool to model the data. After this comes, the data is constructed according to the derived attributes and the generated records. For example, if a product was never sold before but started selling on one of the e-commerce websites, this new information will need to be recorded. This is critical for the modelling purposes so that the change in sale scenario is clearly visible to the head of the project or the CEO. Following this step comes the transformation step, in which the data is converted into a consistent format and encoded for security purposes. (Wirth et al)

The data can be prepared using the Tableau visualisation tool. Every data source created on Tableau has an original data model. The data model is followed by Tableau to query the data in the connected database tables. For example, in the Figure 4 below, an example of unifying and modelling a data set is given where the customer information. The sales information is joined with the "Union Data set". As the four sources are identical, they can be made into a single data set. Different field names are pulled from the Table 1 and inserted in the model in Figure 4. The + icon is simply clicked, and the rest of the data set is added to form a single model in the end. The Figure 4 is an example of joining similar data sets into one model.

earth_sales   Fields selected: 5 of 5

Select the fields to include in your flow. If you make changes to the data, the data source will be queried again.

☑		Field Name	Original Field Name	Filters	Sample Values
☑	#	TransactionID	TransactionID		1, 2, 3
☑	#	CustomerID	CustomerID		115, 280, 142
☑	Abc	ShipCode	ShipCode		BMI-23065643, KAV-60961002, YST-23100847
☑	Abc	SalePrice	SalePrice		$1322685.14, $1200575.02, $1465902.03
☑	Abc	SaleDate	SaleDate		5/15/3009, 5/1/2932, 5/1/2919

Table 1: Data Sources to Add in the Union Set

Source: Interworks (2018)

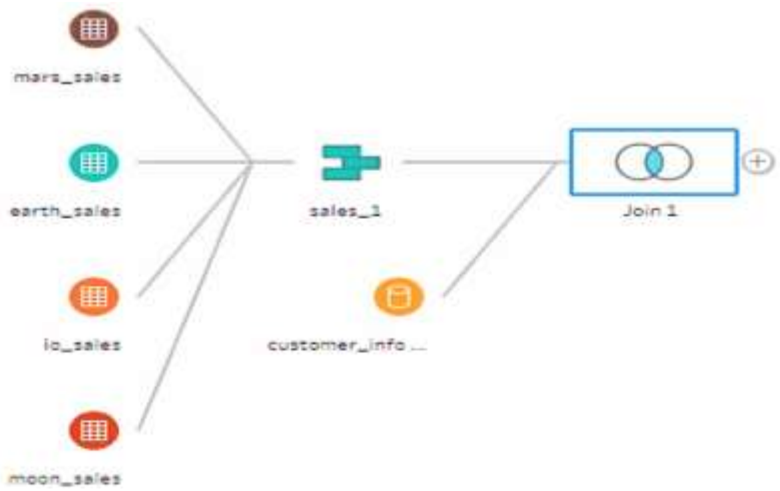

Figure 4: Joining Data Sets

Source: Interworks (2018)

3.2.4 Phase 4 – Modelling

This phase is of more relevance for the main objective of this paper. The modelling phase includes selecting a modelling technique after studying each parameter, generating a design template, building the model, and assessing the model. For selection of a model, you can use a decision tree analysis or a neural network generation (Weinberg, 2019). To check the validity and the quality of the model, a test design is important. After this step, the main purpose of the modelling phase begins, which is data visualisation (Gupta, p.14, 2014)

3.2.4.1 Tableau

In this phase of modelling we will discuss the importance and benefits of integrating Tableau as a data visualisation tool in the 4th Phase of the CRISP-DM model. The Tableau analytics platform is the most secure, powerful, and agile “end-to-end” platform for data analysis. It is a software company created by the Salesforce organization. The software is used by data scientists and analysts, business leaders and individual users.

The benefits of using Tableau for data modelling and visualisation include ability to handle large data sets, easy integration and understanding of the software that does not require high IT skills, availability of online help on Tableau.com, and connecting to Python for advanced analysis. In the "Modelling" phase of the CRISP-DM model, Tableau can perform a wide range of tasks and make the process easy and precise for the later phase of data evaluation. You can connect, integrate, prepare, and structure data according to the needs of the business. (Tableau)

Now, let us see how we join charts and integrate it on one page for modelling the data. For example, a global car dealership company wants to know in which parts of the world its cars are sold most. In the CRISP-DM model, this will be the generic task. The specialised task will be to model the data based on country, state, region, and city where the cars are sold. First, there will be an analysis of each dataset in a "worksheet". The required sheets are then accumulated and connected in a "Dashboard". A dashboard is "a collection of several views" from the worksheet which is "updated with the most recent data source" (Tableau). Different charts or graphs can be compiled on one page and filters like city, country etc. can be applied to check the same for all cars. Hence, you can get a distinct picture of the sale scenarios.

Figure 5: Cars Sold – Country-wise

Source: Aimee Wiscomb, Tableau Public, 2018

The modelling tools and methods will be discussed more in detail further in the case studies.

3.2.5 Phase 5 – Evaluation

The next phase in the CRISP-DM model is the evaluation phase which involves evaluating results, reviewing the process, and determining the next steps. The results are evaluated after modelling to check the degree to which the model or the visualisation matches the objectives of the company. The tasks in this step are evaluating and assessing the results, comparing the results, interpreting the patterns about validity, comprehensibility, and interestingness. It is helpful to create a ranking of results in terms of business success criteria, checking the impact on the end goal, and stating conclusions for further data mining projects. (Chapman, 2000)

3.2.6 Phase 6 – Deployment

The final phase is deployment, where the data is translated into practical actions and used by the decision makers like the team leader or the company executive. The analysis should be comprehensible by the concerned departments so that they can make informed decisions. The CRISP-DM process must be automated if the company requires similar analysis in the future. (Gupta. p. 17, 2014)

3.3 Limitations of CRISP-DM Methodology

The CRISP-DM methodology has many advantages, because of which it used by businesses even today after 22 years of its creation. However, there are some limitations of this methodology. The model is easy to be replicated and repeated after it is used once if there is a similar project requirement. But there are four issues that can arise in some phases of the process as shown in Figure 5.

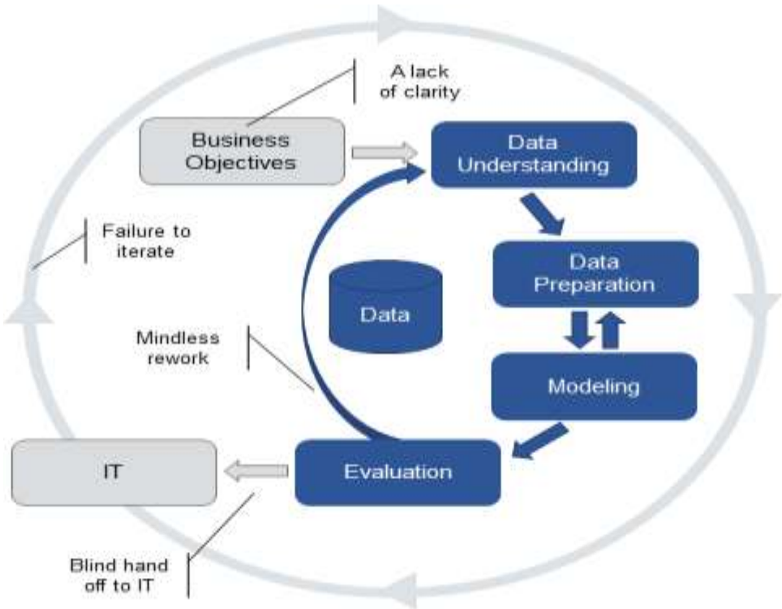

Figure 6: Where can the CRISP-DM Model go Wrong

Source: J. Taylor, KDnuggets News, 2017

One such issue is lack of clarity in the first phase. After the analytical team understands the business objectives, it focuses on gathering data and analysing it for the purpose of modelling. Many times, however, the business goal is lost in this process and in the pursuit of making the results interesting, they stray away from the business need. In the evaluation phase, it is possible to backtrack to Phase 2 for rework and this is where the next issue may arise. If the evaluation does not meet the objectives of the business, most analysts will restart the process of from the beginning and start gathering new data to model. This can be time consuming and would result in mindless rework. On finishing the evaluation phase of the methodology, a few analysts may ignore the deployment phase, where the data is to be presented to the IT team or executives. Sometimes the data is blindly handed over to IT and they are unable to judge the data, whether it is easy or difficult to execute. To make the CRISP-DM model reusable and ensure sustainability of the model, it must be revaluated and updated constantly. This is often seen as a future task. The analysts should have a clarity of the changing business goals and objectives to be able to update and keep it valuable. (J. Taylor, 2017)

It is critical that the team is able not only able to build an interesting analytical solution, but also add business value to it. Companies that want to sustain their businesses and reuse data mining techniques for further problems and goals cannot afford to have intrinsic problems related to the integration between business objectives and data modelling. (J. Taylor, 2017)

4. Case Studies

The following case studies will provide a deeper understanding of the CRISP-DM model and its integration with Tableau for data modelling.

4.1 Case Study 1: Verizon's ACE Team Task of Support Call Reduction

Verizon is an American telecommunications company that offers wireless products and services. The company's analytics team used the Tableau visualisation tool for sales analysis after which it was able to cut customer support calls by 43 percent and enhance customer experience by introducing chatbots to answer customer queries. (Tableau, 2018). Let us see how the CRISP-DM model can be integrated with Tableau and applied in the data mining project of Verizon's ACE team.

Phase 1 – Business Understanding

To understand the business, the team understood the goal of the business project, which was to reduce the number of customer calls by 43 percent and increase good customer experience.

Phase 2 – Data Understanding

For data understanding, all data was extracted from various platforms like Oracle, Hadoop, and Teradata and aligned. Verizon's ACE team was able to reduce the large data set into smaller sets and analyse it on Tableau after integrating it with R.

Phase 3 – Data Preparation

In data preparation phase, the team used the Tableau Prep Builder to visually combine, reshape and clean the data from multiple sources. It structured the data correctly into union sets for it to be scalable and automated. So, each time a new information is added, because of the "live connection" option on the home page in Tableau Software, the data is automatically updated.

Phase 4 – Modelling

Now comes the modelling phase. The modelling process and the methods used in the process are structured in the Figure 7 below. To select a model, a decision tree analysis or a neural network generation can be used. For checking the validity and the quality of the model, a test design is important (Weinberg, 2019).

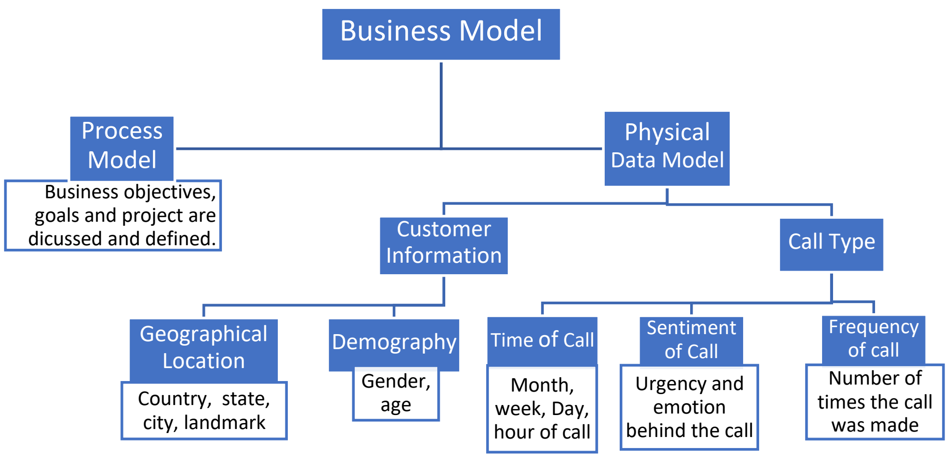

Figure 7: Data Modelling by ACE for Business Objectives and Representation of Customer Segmentation

The methods used for modelling include smart grouping algorithms to reorganize the list of customers, excluding and including customers based on the geographical location and call frequency, and filtration of the customers based on the sentiment and urgency of the call (Tableau).

The ACE team developed dashboards to get insights on different features, calculations, and field definitions on Tableau. After understanding the different groups of customers and their calling habits, ACE was able to build a routing solution for certain types of customers and use Tableau Dashboards to analyse customer engagement in call centres and monitor the type of call.

Phase 5 – Evaluation

To check the validity and the usefulness of the model, the dashboards on Tableau were checked to see if they gave a detailed view of the customer call scenario and the reasons behind it. The dashboard was able to record the past calling patterns of the customers and notify the customer support team when the call volumes would scale down. When this happened, an algorithm was activated to reveal high-request customers. The geospatial mapping in Tableau as shown in Figure 8, was used to map and monitor impact of service dispatches to homes.

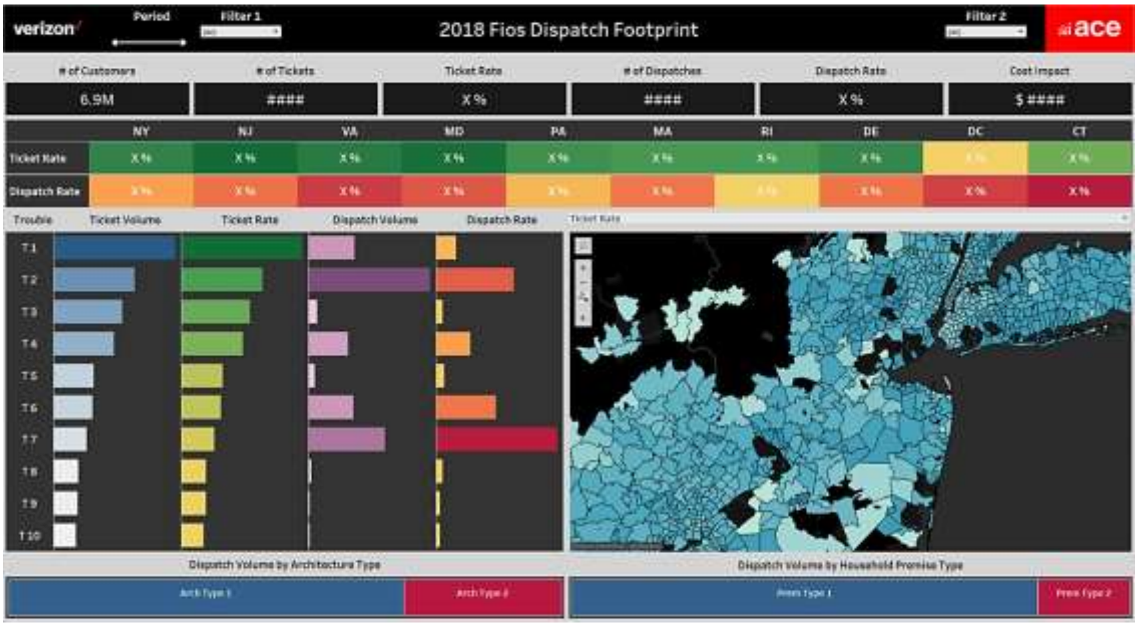

Figure 8: ACE Team Task – 2018 Fios Dispatch Footprint

Source: Tableau (2018)

As shown in the Figure 8 above, the dashboard included the number of customers requesting for at home service, tickets created, rate of times the tickets were generated by each customer, number of dispatches and the cost impact on the business. When required, the presentations were fixed on the dashboard by recalculation or refiltration of the data.

Phase 6 – Deployment

The decision makers of the company were also beginning to see the analysis of the queries answered by the chatbot and the most appeared keywords in the queries. The keywords were categorised under three topics – customer service, television programs and movies.

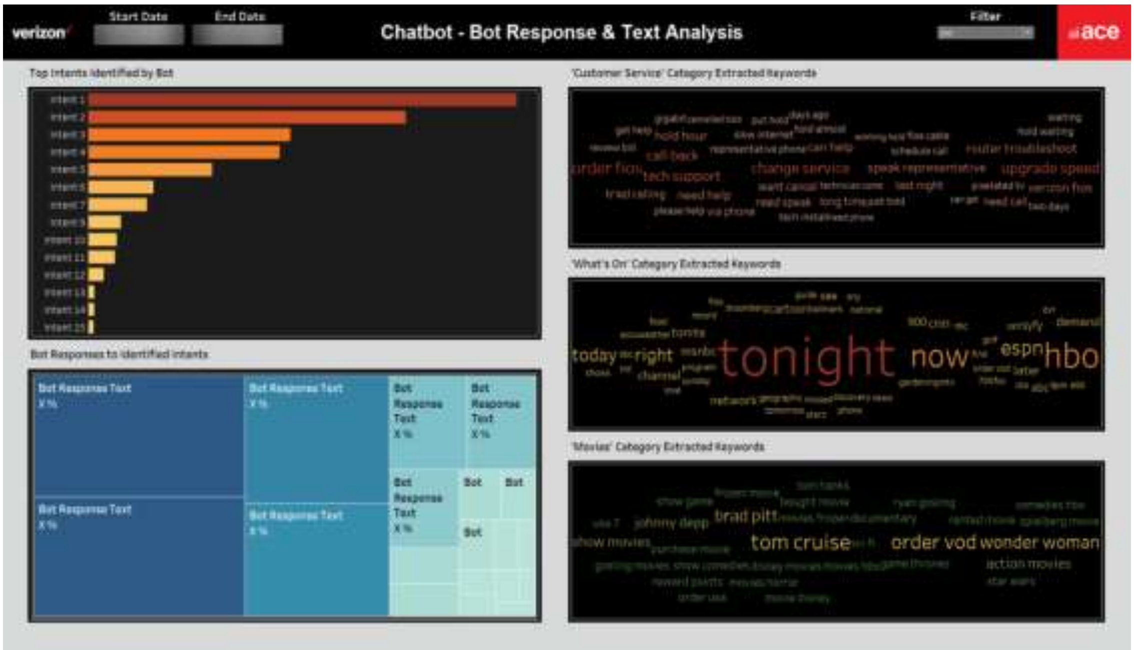

Figure 9: Chatbot Responses Analysis for Evaluation

To conclude, the features and ease of use with Tableau helped the ACE team of Verizon to examine the customer support calls and track responses of the chatbot used by the company that addressed the customer's queries. In the end, the benefit of using this model was proved when the calls were cut by almost half than before and the chatbots were able to prioritize the customer queries effectively.

4.2 Case Study 2: Using CRISP-DM for Airbnb Listings

This is a case study which is partly based on two analyses – one by Fan Yuan and another by Shinya Yaginuma, about Airbnb listings and what can be done to increase the transactions. The Tableau Visualisation tool was used for data preparation, modelling, and evaluation.

Phase 1 - To understand the business of Airbnb, let us discuss the background and objectives of the company. Airbnb is an American ecommerce company that offers lodging spaces after individuals rent it out for stay (Nath, Investopedia). By studying the business model below (Figure 10), we can see that the company's profits are mostly generated through service fees from bookings for guests and the hosts, depending on the size of the accommodation space and location (Ashworth, Lodgify, 2020) and the main source of revenue is the price for host (Yaginuma, Medium, 2019)

Now, it plans to generate more revenue from the locations in which people are booking rooms or houses the most. Additionally, it wants to know what days in a week the rooms are booked the most to set the price high. The new project requires an analysis of the type of customers, lodging locations, revenue generated with respect to the city, and the peak time of bookings.

By using the CRISP-DM model, we can study in which period of the year Airbnb has generated most profits and in what cities the listing rate is the highest.

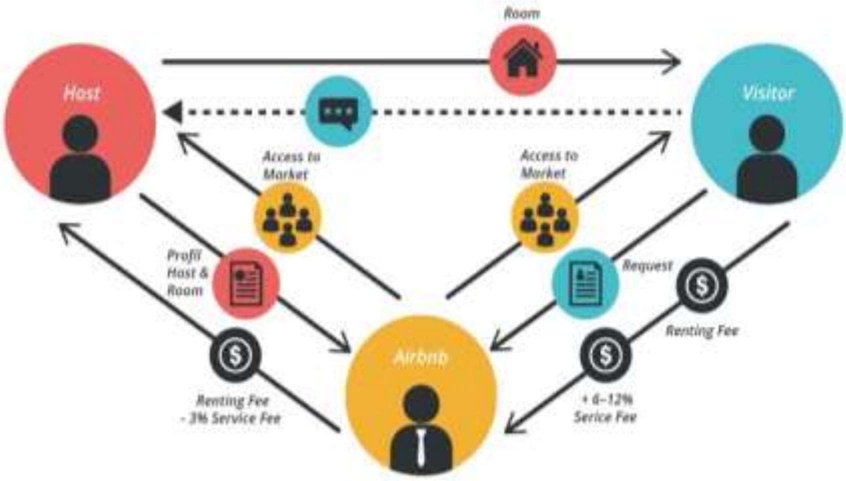

Figure 10: Airbnb Business Model

Source: Shinya Yaginuma, Medium, 2019

Phase 2 - After extracting the data from GitHub, the “listings.csv” file is arranged on python with respect to host, description of guest, location details like country, state, city, property type, number of rooms, beds, and amenities. It includes price factors like cleaning fee, cancellation fee, minimum nights, extra people. The outcome variables in the listings contain information about review ratings on stay, communication, and location.

Phase 3 – For data preparation, data sets can be arranged into union sets just like it was done in the first case study. Tableau Prep Builder tool can be used to combine and clean the data. It is easy to rearrange in case of revaluation and automated because of the “live connection” option in the Tableau software.

Phase 4 - After understanding the business objectives, the data is modelled by removing irrelevant characters and words and arranging it in a specific format by keeping only 10 to15 most popular locations of Airbnb in the data set. The date format is also arranged in such a way that the exact day of booking is visible in the data set. To understand the data with regards to average price listing by date, it is aggregated into yearly data and daily data as shown in Figure 11 and Figure 12 respectively, using the Tableau Visualisation tool (Yaginuma, Medium, 2019). A further analysis by combining multiple data sets can be done to visualise the locations in which the Airbnb listings are highest. The results can be seen in Figure 13.

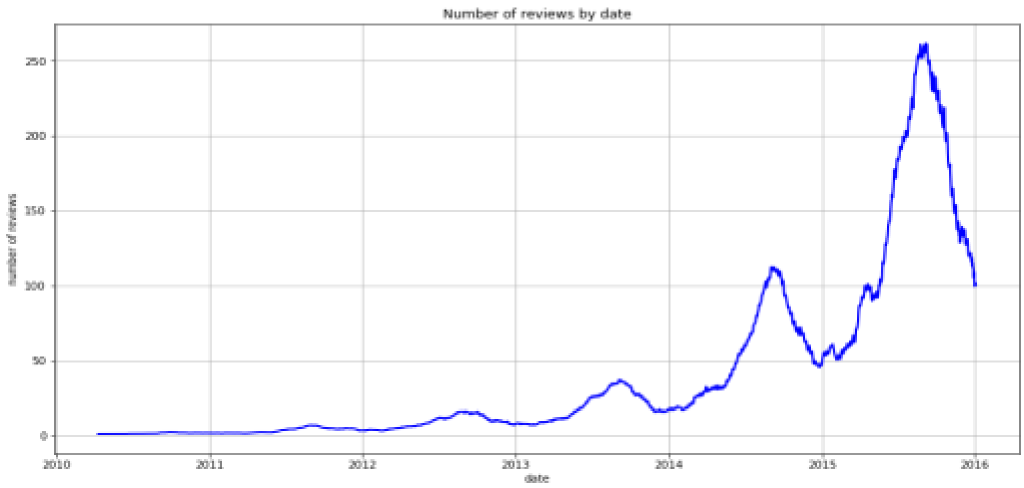

Figure 11: Number of Reviews per Year

Source: Shinya Yaginuma, Medium, 2019

Figure 12: Airbnb Price Trend by Date in 2 weeks in Year 2016.

Source: Shinya Yaginuma, Medium, 2019

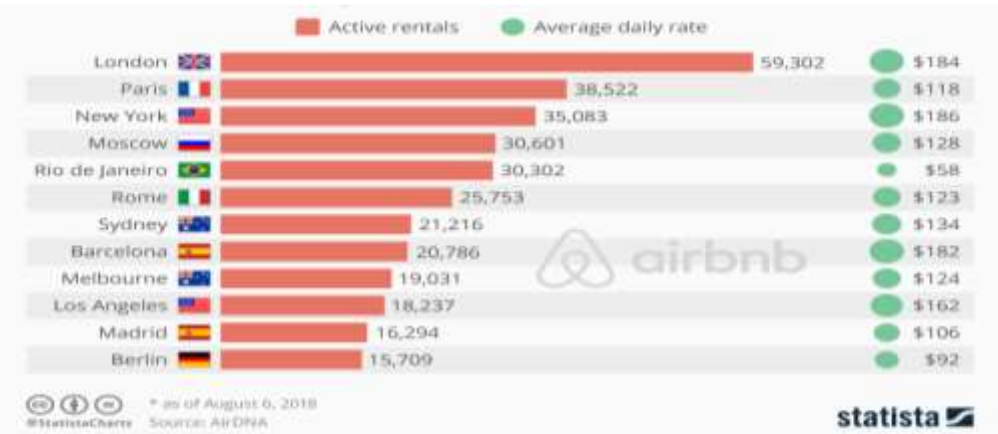

Figure 13: Airbnb - Number of Active Airbnb Listings in Major Cities Worldwide

Source: Felix Richter, Statista, 2018

Phase 5: The data is evaluated based on its validity and effectiveness in providing the solution to the problem. It was found out that Friday and Saturday are the days of high price listing for lodgers. Looking at the model, we can interpret that the peek time for booking at Airbnb is between July and September, because the available rooms are few and the room price is high.

Phase 6: The data can now be deployed for the sales team to view and they will be able to understand the parameters that led to this profit generation between July and September. They can increase their business potential by giving more preference and offers to the hosts in London after observing the outcome in Figure 13 and the decision makers can further increase the price for the lodgers.

So, the two case studies explained how the CRISP-DM model can be used with Tableau in a detailed manner.

5. Conclusion

Herewith, the paper concludes after attempting to give an insight into the CRISP-DM methodology and its integration with the Tableau data visualisation tool for a business. It is vital for companies to use a data mining model like the CRISP-DM to enhance their sales scenario and raise the profit margin. It is the most widely used data mining method in the world among all businesses as it works well in terms of planning, documentation, and communication of the project objectives, process, and results. Having six phases, it is well structured, easy to follow and shows a flexible process cycle for analysts to use it efficiently with ease (VK, 2019). The methodology is simple to reuse and it is also possible to repeat the algorithms that are originally based on the generic tasks.

There are some drawbacks of following this methodology, but it can be overcome by structured decisions, detailed focus on the business objectives, good communication and defined KPIs to control the progress of a project. The model must be updated on a regular basis to avoid future confusion when the analytical team plans to reuse it. Furthermore, the authors mentioned in this research used qualitative research for providing detailed explanation of how the CRISP-DM methodology is used, the different phases and the four drawbacks of using the methodology.

The case studies used in the paper showed real and detailed examples of how the model helps businesses. It can be used to ease the problems of high customer support costs and effort, as shown in the case of Verizon. The model also aids in the success of a project in a company to explore new potentials like in the case of Airbnb.

To conclude, the CRISP-DM methodology is a dependable method for data mining processes to solve business problems and implement their data mining projects. As data mining is an iterative process, the CRISP-DM model too, is an iterative process which makes it a future-oriented tool for many companies. Tableau helps in making the model easier to use. Using such a method of integration for any business project will bring an assurance and uniformity in the operational procedures of a company.

6. Bibliography

Ashworth, J., 2020. *How Much Is Airbnb Really Costing You?*. [online] Vacation Rental Owners & Property Managers Blog - Lodgify. Available at: <https://www.lodgify.com/blog/airbnb-vs-lodgify/> [Accessed 10 July 2020].

Berns, C., 2019. *Using CRISP-DM To Predict Car Prices*. [online] Medium. Available at: <https://medium.com/@christophberns/using-crisp-dm-to-predict-car-prices-f15eb5b14025> [Accessed 10 June 2020].

Gupta, G., 2014. *INTRODUCTION TO DATA MINING WITH CASE STUDIES*. [online] Google Books. Available at: <https://books.google.de/books?hl=en&lr=&id=fzB9BAAAQBAJ&oi=fnd&pg=PR3&dq=CRISP-DM+Case+Studies&ots=ylCMA8oNbg&sig=g2sbuNzs_m9xPhgZSYa4bHrABOw#v=onepage&q=CRISP-DM%20&f=false> [Accessed 1 July 2020]. 12-16

Inseaddataanalytics.github.io. 2011. *IBM SPSS Modeler CRISP-DM Guide*. [online] Available at: <https://inseaddataanalytics.github.io/INSEADAnalytics/CRISP_DM.pdf> [Accessed 1 July 2020]. 1-4

Itzhak Weinberg, A. and Last, M., 2019. *Selecting A Representative Decision Tree From An Ensemble Of Decision-Tree Models For Fast Big Data Classification*. [online] Journal of Big Data. Available at: <https://journalofbigdata.springeropen.com/articles/> [Accessed 10 July 2020].

Kumar, A., Kumar Sahu, S. and M. Baitule, S., 2018. *Data Discovery With Tableau: A Case Study Using Data Analytics Tool*. [online] Pdfs.semanticscholar.org. Available at: <https://pdfs.semanticscholar.org/dcda/d3152b40e2673da77e3ddcb86329b79da7e1.pdf?_ga=2.27028926.288580117.1593524077-349381089.1593524077> [Accessed 1 June 2020].

Lock, S., 2018. *Airbnb - Statistics & Facts*. [online] Statista. Available at: <https://www.statista.com/topics/2273/airbnb/> [Accessed 1 July 2020].

N. Milligan, J. and Guillevin, T., 2018. *Tableau 10 Complete Reference*. [online] Google Books. Available at: https://books.google.de/books?id=FNCBDwAAQBAJ&pg=PA466&lpg=PA466&dq=Integrating+Crisp+DM+Methodology+for+a+Business+Using+Tableau+Visualization&source=bl&ots=vbihAb22yP&sig=ACfU3U2S9yhnibxP1fW9z2BGfpLxYHgGBA&hl=en&sa=X&ved=2ahUKEwj0yc_E0L3qAhVQ2qQKHWOwB0EQ6AEwAHoECAkQAQ#v=onepage&q=Integrating%20Crisp%20DM%20Methodology%20for%20a%20Business%20Using%20Tableau%20Visualization&f=false> [Accessed 1 July 2020].

Public.tableau.com. 2020. *Interaktives Markt Prognose Tool*. [online] Available at: <https://public.tableau.com/profile/statista#!/vizhome/Statista-InteraktiesMarkt-Prognose-Tool_0/Statista-UmstzenachBranchen> [Accessed 4 June 2020].

Richter, F., 2018. *Infographic: London Is The World's Airbnb Capital*. [online] Statista Infographics. Available at: <https://www.statista.com/chart/14986/active-airbnb-listings-in-major-cities/> [Accessed 1 July 2020].

Smart Vision - Europe. n.d. *What Is The CRISP-DM Methodology?*. [online] Available at: <https://www.sv-europe.com/crisp-dm-methodology/> [Accessed 1 July 2020].

Tableau Software. n.d. *How Data Mining Works: A Guide*. [online] Available at: <https://www.tableau.com/learn/articles/what-is-data-mining> [Accessed 29 June 2020].

Wirth, R. and Hipp, J., n.d. *CRISP-DM: Towards A Standard Process Model For Data Mining*. [online] Cs.unibo.it. Available at: <http://www.cs.unibo.it/~danilo.montesi/CBD/Beatriz/10.1.1.198.5133.pdf> [Accessed 18 June 2020].

Statutory Declaration

I, **Akshata Satyawan Parate**, hereby declare that this paper titled **“How AI Is Changing Sales”** is my original work and has been authored in my own words. I have not used any sources other than the declared sources and have explicitly marked the material which has been quoted either literally or by content from used sources.

Name: Akshata Satyawan Parate

Date: 10.07.2020

Place: GÖTTINGEN

Signature