Chapter 5

CRM AND DATA MANAGEMENT

INTRODUCTION TO DATA:

♦ Data refers to collection of raw facts and figures. CRM data usually consists of personally identifiable information. Essentially, this means any data that can be used to specifically identify an individual (such as name, address, email, phone number).

♦ At the centre of a business' CRM is their customer database, which is a really helpful tool enabling them to; identify customer trends, create customer loyalty and increase their customer communication.

♦ Establishing and maintaining good relationships with customers provided few problems in the past when businesses were small and customers were identifiable by sight but now the scenario has changed where large enterprises knows their customers with Unique Id code allotted to them.

TYPES OF DATA:

(1) Transactional Data:

♦ Transactional data refers to the data that is created and updated within the operational systems. Examples of transactional data included the time, place, price, discount, payment methods, etc. used at the point of sale. These are the business documents that you create using the master data like purchase orders, sales orders etc.

♦ Transactional Data can change very often and are not constant. And this data is created/modified out of an application transaction.

♦ A company's operations are supported by applications that automate key business processes. These include areas such as sales, service, order management, manufacturing, purchasing, billing, accounts receivable and accounts payable. These applications require significant amounts of data to function correctly.

♦ This includes data about the objects that are involved in transactions, as well as the transaction data itself. For example, when a customer buys a product, the transaction is managed by a sales application. The objects of the transaction are the Customer and the Product.

♦ Transaction data is data which is frequently updated like sales quantity. It is changed within short period of time.

(2) Reference Data:

♦ Reference data is the data without which you cannot do any transactions and is mandatory for every organization. It describes the things that interact when a transaction occurs. For instance, Reference data that represents product and customer must be present before the transaction is fired to sell a product to a customer.

♦ Reference data are distributed throughout the company, they are often not standardized and often redundant. As a result it is very costly to offer efficient customer service, keep track of supply chains and make strategic decisions.

♦ Regardless of the industry, companies often work with different ERP and Legacy systems. The result: the

business processes are based on information about customers, partners and products which is displayed in different ways in the systems. If the data are recorded manually, there are more inconsistencies: some data sets are entered several times, others cannot be retrieved by all divisions of the company.

♦ As corporate applications are becoming increasingly complex and produce ever greater amounts of data, the problem is intensified further. Nevertheless, employees must work with the inconsistent data and make decisions on this basis. The lack of standardized reference data easily leads to wrong decisions, which restrict efficiency and threaten customer satisfaction and profitability.

♦ In order to save costs and ensure your company's success it is necessary to consolidate reference data about customers, partners and products, make them available to all employees beyond system boundaries and use attributes valid company-wide for the purpose of description.

♦ Reference Data is the data which will describe the type of object, like name, description, cost, length and dimensions of the object. It is a one-time data. for example - if there is material X. The description of X, attributes like what type of material, is it a finished goods, prices etc.

♦ Reference data is data which is frequently read, but rarely updated like: Name, Address. It remains unchanged for a period of time like client handling banking application. In the process the address,

account number of banking customers is reference data.

(3) Data Warehouse:

♦ A data warehouse is a subject-oriented, integrated, time variant and non-volatile collection of data in support of management decision making process.

♦ Data warehousing implements the process to access heterogeneous data sources; clean, filter, and transform the data; and store the data in a structure that is easy to access, understand, and use. The data is then used for query, reporting, and data analysis.

♦ The volume of data in data warehousing can be very high, particularly when considering the requirements.

♦ Data warehousing is a phenomenon that grew from the huge amount of electronic data stored in recent years and from the urgent need to use that data to accomplish goals that go beyond the routine tasks linked to daily processing.

♦ A data warehouse is a collection of data that supports decision-making processes.

Features of Data warehouse:

(a) It is subject-oriented: Data warehouses are subject-oriented because they hinge on enterprise-specific concepts, such as customers, products, sales, and orders. On the contrary, operational databases hinge on many different enterprise-specific applications.

(b) It is integrated and consistent: Data warehouses take advantage of multiple data sources, such as data extracted from production and then stored to enterprise

databases, or even data from a third party's information systems. A data warehouse should provide a unified view of all the data. Generally speaking, we can state that creating a data warehouse system does not require that new information be added; rather, existing information needs rearranging.

(c) It shows its evolution over time: Operational data usually covers a short period of time, because most transactions involve the latest data. A data warehouse should enable analyses that instead cover a few years. For this reason, data warehouses are regularly updated from operational data and keep on growing.

Fundamentally, data is never deleted from data warehouses and updates are normally carried out when data warehouses are offline. This means that data warehouses can be essentially viewed as read-only databases. This satisfies the users' need for a short analysis query response time and has other important effect.

(4) Business View Data:

♦ Business View Data is generated from data warehouse. Business views are calculations or summaries compared over period of time.

♦ This data focus on particular business area such as answering business questions like: How many complaints have been received from customers in last three years?

♦ Analysis tools run more efficiently against a consolidated business view than any other data included in data warehouse.

DATA MANAGEMENT:

*   In an organization, Data management plays a significant role to generate revenue, control costs and mitigate risks.

*   Data Management can bring about competitive advantage as business environment will be able to share, store, protect and retrieve the ever-increasing amount of data within a short span of time.

*   Management of data generally focuses on the defining of the data element, how it is structured, stored and moved.

*   Management of information is more concerned with the security, accuracy, completeness and timeliness of multiple pieces of data.

*   Most organizations today are flooded with data, the volume of which is increasing at an alarming rate. It is vital, therefore, to determine which data are most relevant and essential from an enterprise-wide perspective.

*   Identification and classification of the enterprise's critical data should be performed by a team of senior-level representatives from each line of business or department. These team members must have knowledge of the relevant contributing business systems and processes, and the requirements of their respective stakeholders, systems and processes, and the requirements of their respective stakeholders.

IDENTIFYING DATA QUALITY ISSUES:

*   Data accuracy is a troublesome issue. In an ideal world it would be wonderful to have 100 per cent accurate data. But data accuracy carries a high costs. Data are captured, entered, integrated and analysed at various moments. Any or all of these processes may be the source of inaccuracy.

*   Keystroke mistakes can cause errors at the point of data entry. Inappropriate analytical processes can lead to ill-founded conclusions. In CRM, data inaccuracy can lead to undue waste in marketing campaigns, inappropriate prospecting by salespeople and general suboptimal customer experience. It also erodes trust in the CRM system, thus reducing usage. This leads to further degrading of data quality.

Features of Qualitative Data:

*   **Accuracy:** Stored values should be compliant with real-world ones.

*   **Freshness:** Data should not be old.

*   **Completeness:** There should be no lack of information.

*   **Consistency:** Data representation should be uniform.

*   **Availability:** Users should have easy access to data.

*   **Traceability:** Data can easily be traced data back to its sources.

*   **Clearness:** Data can be easily understood.

Issues Related with Data Quality are as follows:

(1) Data Deterioration

(2) Lack of customer trust

(3) Improper sources of data collection

Explanation:

(1) Data Deterioration:

♦ 75 percent of the customer information held by most enterprises is inaccurate, out-of-date or ineffective for marketing purposes.

♦ Many customers address, contact details get changed which they hardly bothered to update it everywhere. More than 50% of database gets deteriorated by such inaccurate data, thus affecting the quality of data.

♦ A fully-functional relational database management system allows users to enter new information, update current records and delete outdated data. As an example, when a salesperson sells 1,000 units, that person will enter the transaction information into the relational database management system.

♦ The data can include the salesperson's name, the customer information, the product sold and the quantity sold. The relational database management system enters a new record in the customer table, updates the salesperson's record and subtracts 1,000 units from the inventory record.

(2) Lack of Customer Trust:

♦ Relevant data for CRM is captured both by employees and customers. Recently the use of self service applications has become of growing importance. If customers are suspicious about the use of their data, the willingness to provide information about themselves is often limited. This can lead to inconsistent or missing

data if faulty data is captured deliberately because of security or privacy requirements.

♦ If the employees are not aware of the relevance of the customer information captured, the quality of this data is likely to be inadequate for further usage. An employee may fill an attribute, e.g. the age of the customer, with default data only to fulfil integrity checks, if he is not aware of the relevance of a specific information captured. This can often be hardly detected and corrected during data analysis causing wrong analytical results.

(3) Improper Sources of Data Collection:

♦ Data from different internal and external sources is integrated in a customer data warehouse to realize a holistic view on the customer and its needs and behaviour. The influence on the quality of an external data supply and the information about the processes behind are limited. Finally, the design of the data exchange processes and the data quality assurance while integrating external data, influence the quality perceived by the data user.

♦ Within a company the management sets guidelines, like the availability of financial and human resources and incentive schemes. These guidelines limit resources available for data quality assuring activities, for designing business processes, for IT systems and later during the execution processes: e.g. the time, a call centre agent can spend on assuring data during a customer contact, is restricted, the possibility to check and refresh data is restrained.

(4) Internationalization and untrained staff:

♦ The organization has to face internationalization when they have to maintain global customer relationship. Here the organization may have to deal with foreign customers and would have to maintain customer information in other languages and other character set.

♦ There can also be difference in the format in maintaining data when department varies. As data quality is a new concept in the IT automation, the problem comes when there are very less number of trained and skilled staffs to handle the application system.

PLANNING AND GETTING INFORMATION QUALITY:

♦ Accurate information and reports are the life blood of an effective sales force. Without it, management doesn't have the data to make good decisions, sales force don't have the tools to turn leads into customers, and the company will find it difficult to reconcile CRM data with data in other systems.

♦ To ensure consistently high data quality, company need to train its users, create and implement a data quality process, and use available technologies to automate the process whenever possible.

Following are six stages which help to plan and get qualitative information:

(1) Data Profiling: This is all about understanding your data. Company should be well aware with the source of data which could be spreadsheets, backend systems, or

sticky notes all over manager's desks. It includes listing all data sources and the names of the fields in which data is stored. To identify potential problems with current data. To check automated quality before a new record could be saved. To avoid duplicate information between objects (such as Accounts, Opportunities, and Contacts) and fields (such as State, City, and so on).

(2) Data Control: This is all about achieving data accuracy and ensuring the right users have access to the right information, which also means blocking access, as needed. To control the data, you first need to "clean" it by removing duplicates and errors, and then set up processes and use technologies to keep it clean. Automated routines or tools should be used to clean data. Fix data that's highly visibly and frequently used, such as addresses and emails. Fix business-specific information next, such as Opportunity types and stages.

(3) Data Integration: Most organizations have data in more than one system. Whenever there's duplicated data in several systems, it's easy for information to get out of synch. One way to avoid this problem is to integrate your systems so that updates in one system automatically update the others, resulting in a single source of "truth" and making it easier for end users to access information.

(4) Data Augmentation: To make CRM system even more valuable, augment your data with information that will give salespeople and managers an edge. For example, a number of third-party organizations such as Dun & Bradstreet and Hoover's provide valuable information to

sales representatives can use to prepare for sales calls. Also make the most of internal market intelligence, such as purchasing patterns or competitive analyses. To understand what data is valuable, sales and marketing users frequent used information should be surveyed. Internal information about your customers' behaviour and buying patterns should be noted so as to evaluate whether this information would be helpful to sales representatives or not. Data augmentation adds value to base data by adding information derived from internal and external sources within an enterprise. Data augmentation can help reduce the manual intervention required to developed meaningful information and insight of business data, as well as significantly enhance data quality.

(5) **Monitor Data:** Monitoring ensures data conforms to rules that define good data quality in business. Data monitoring checks important performance indicators in the database system, such as the database size, quality of the database buffer, and the database indexes. Data monitor uses statistics that are provided by the database system with which companies are working.

(6) **Assign Ownership, Train Users, and Commit to a Data-Quality Process**: Users need to know the importance of data integrity and how to do their part in any data-quality initiative. Training should be provided to the users to show them how data quality directly affects their work. It's also a good idea to assign ultimate responsibility for each region's data to a super user, geographic lead, or other business owner.

TOOLS TO MANAGE DATA:

♦ CRM tools should be integrated into your systems as seamlessly as possible, making them a natural part of the customer service interaction.

♦ CRM tools will help company team to identify precious members of customer mix and prompt team members to notice and value the extended relationship with them.

♦ CRM tool will help to track each transaction, spot new trends as customer expectation change, and identify opportunities to expand the service relationship.

♦ The most commonly used tools in data mining are:

(a) **Artificial neural networks**: Non-linear predictive models that learn through training and resemble biological neural networks in structure.

(b) **Genetic algorithms:** Optimization techniques that use processes such as genetic combination, mutation, and natural selection in a design based on the concepts of natural evolution.

(c) **Decision trees:** Tree-shaped structures that represent sets of decisions. These decisions generate rules for the classification of a dataset. Specific decision tree methods include Classification and Regression Trees (CART) and Chi Square Automatic Interaction Detection (CHAID). CART and CHAID are decision tree techniques used for classification of a dataset. They provide a set of rules that you can apply to a new (unclassified) dataset to predict which records will have a given outcome. CART segments a dataset by creating 2-

way splits while CHAID segments using chi square tests to create multi-way splits. CART typically requires less data preparation than CHAID.

(d) Nearest neighbour method: A technique that classifies each record in a dataset based on a combination of the classes of the k record(s) most similar to it in a historical dataset (where k 1). Sometimes called the k-nearest neighbour technique.

(e) Rule induction: The extraction of useful if-then rules from data based on statistical significance.

(f) Data visualization: The visual interpretation of complex relationships in multidimensional data. Graphics tools are used to illustrate data relationships.

DATA ANALYSIS:

There are many ways in analysing the data. Here are few of them are explained below:

Data Mining:

♦ The Data mining can be used to recognize the relationship from the gathered data, model and rules. It can produce the information about the model without knowing about that model. So, it is a great tool for obtaining knowledge from it. Data mining can be classified into three types.

(a) Predicting analysis: It can analyse the future prospects of the product using the historical data.

(b) Sequence analysis: It recognizes the combination of actions that are placed in a particular order. From this, it can analyse whether the consumers are doing something in certain order.

(c) Association analysis: It recognizes the items which can combine similar items. This analysis is regularly used for some profits to the organization by giving discounts to some products which are the combination of two items.

Online Analytical Processing (OLAP):

♦ Online analytical processing (OLAP) OLAP is an advanced data reporting tool. It is not strictly a data mining tool because it provides summary data rather than identifying patterns in data.

♦ OLAP tools are powerful and quite easy to use. They can make a significant contribution to extracting value from customer databases, adding to the value of data mining applications, rather than replacing them.

♦ OLAP tools have advanced graphical interfaces that make it possible for users with little statistical knowledge to explore large volumes of data. Underlying this interface is a new database containing data from the data repository that has been stored using a special structure to make the 'slicing and dicing' of data quicker and easier.

♦ Whereas conventional reporting tools can hours to gather data, OLAP tools can provide reports in only a few seconds. However, that achieving this fast response comes at the expense of losing some precision in the storage of certain types of data.

◆ Analytical tools are instrumental in sorting data and extracting meaning from it to guide the development of management strategies. In identifying customer and market trends, techniques such as data mining can help to clarify budget inefficiencies and the most useful allocation of resources.

◆ On-line analytical processing (OLAP) tools provide the ability to rapidly manipulate data containing a large number of table look-ups or dimensions and are particularly useful for performing trend analyses and forecasting. Where a very large number of variables are present and the goal is to determine an appropriate mathematical algorithm to determine likely outcomes, data mining tools can be leveraged.

◆ The use of OLAP tools to create performance indicators is usually sufficient and is supported by many dedicated churn management packages. These tools enable churn rates to be correlated with geographic areas, dealers, service plans and so on. High correlations will indicate, for example, that some dealers have much higher churn rates than others. These tools also support the identification of customer segments that have both high churn rates and high potential value. These segments can then be targeted with customer retention campaigns.

Clickstream Analysis:

◆ Click stream analysis (sometimes called click stream analytics) is the process of collecting, analysing, and reporting aggregate data about which pages visitors visit, in what order and which are the results of the succession of mouse clicks each visitor makes.

Click stream analysis is considered to be most effective when used in conjunction with other, more traditional, market evaluation resources.

◆ Click stream analysis can provide valuable insights to help enterprises know more about the parties with which they do business and thus act more proactively toward meeting their objectives.

◆ There are two levels of click stream analysis,

(a) Traffic analysis.

(b) E-commerce analysis.

(a) **Traffic analysis**: Traffic analysis operates at the server level by collecting click stream data related to the path the user takes when navigating through the site. Traffic analysis tracks how many pages are served to the user, how long it takes pages to load, how often the user hits the browser's back or stop button, and how much data is transmitted before a user moves on.

(b) **E-commerce analysis:** E-commerce- based analysis uses click stream data to determine the effectiveness of the site as a channel- to-market by quantifying the user's behaviour while on the Web site. It is used to keep track of what pages the user lingers on, what the user puts in or takes out of their shopping cart, and what items the user purchases.

Personalization:

◆ Personalization and customer relationship management (CRM) are complementary to each other. CRM is viewed as a strategy to attract, grow and retain customers. Personalization is an approach that can aid in bringing,

staying and returning visitors/customers. They both have the capability of providing the right information or content (e.g. products, services and data) to the end-user/customer at the right time and right place. Both provide a customized outcome for customers that can lead to competitive advantage for firms using either personalization or CRM. The challenge is to incorporate personalization with CRM to maximize the benefits. The concept of personalization and CRM has expanded in scope to emphasize a much broader notion of benefits to a company.

- Personalization ensures that the users get and receive the information in the format most suitable for them.

There are three ways to define personalization:

- **Personalization at the administrator level:** Administrators can define personalization for each user by setting the design of the portal structure for different users. Administrator can define roles, work sets, portal layout, and access methods for different users.

- **Personalization at the user level:** Users can personalize their content within the control limits set by the administrator.

- **Automatic personalization through predictive technology:** Predictive technology allows for automatic personalization based on user type, browser type, device type, user location (whether inside or outside the firewall), connection bandwidth, and the type of event being handled.

Collaborative Filtering:

- Collaborative CRM covers aspects of a company's dealings with customers that are handled by various departments within a company, such as sales, technical support and marketing. Staff members from different departments can share information collected when interacting with customers.

- Collaborative filtering (CF) is the method of making automatic predictions (filtering) about the interests of a user by collecting taste information from many users (collaborating).

- The underlying assumption of CF approach is that, those who agreed in the past tend will agree again in the future also. For example, a collaborative filtering or recommendation system for cooking oil could make predictions about which oil should like given a partial list of that user's likes or dislikes.

Data Reporting:

- Data reporting measures performance, and analyses other key elements that can then be shared within the organization or with the public.

- Reporting translates raw data into **information**.

- Reporting helps companies to monitor their online business and be alerted to when data falls outside of expected ranges.

- Good reporting should **raise questions** about the business from its end users.

- Through the process of performing analysis one may raise additional questions, but the goal is to identify

answers, or at least potential answers that can be tested.

- In short, reporting shows you **what is happening** while analysis focuses on explaining **why it is happening** and **what you can do about it.**

- If most of the team's time is spent on activities such as *building, configuring, consolidating, organizing, formatting,* and *summarizing* – that's reporting.

Reporting follows a **push approach**, where reports are pushed to users who are then expected to extract meaningful insights and take appropriate actions for themselves (i.e., self-serve). I've identified three main types of reporting: *canned reports, dashboards,* and *alerts.*

(1) Canned reports: These are the out-of-the-box and custom reports that you can access within the analytics tool or which can also be delivered on a recurring basis to a group of end users. Canned reports are fairly static with fixed metrics and dimensions. In general, some canned reports are more valuable than others, and a report's value may depend on how relevant it is to an individual's role (e.g., SEO specialist vs. web producer).

(2) Dashboards: These custom-made reports combine different KPIs and reports to provide a comprehensive, high-level view of business performance for specific audiences. Dashboards may include data from various data sources and are also usually fairly static.

(3) Alerts: These conditional reports are triggered when data falls outside of expected ranges or some other pre-defined criteria is met. Once people are notified of what

happened, they can take appropriate action as necessary.

Reporting provides no or **limited context** about what's happening in the data. In some cases, the end users already possess the necessary context to understand and interpret the data correctly.

QUESTIONS

(1) Multiple Choice Questions:

(a) _____ plays a significant role to generate revenue, control costs and mitigate risks.
   (i) Data Profiling
   (ii) Data Control
   (iii) Data management
   (iv) Data Integration

(b) _____ is generated from data warehouse. Business views are calculations or summaries compared over period of time.
   (i) Transactional Data
   (ii) Reference Data
   (iii) Business view Data
   (iv) Data Integration

(c) _____ is the data without which you cannot do any transactions and is mandatory for every organization.
   (i) Transactional Data
   (ii) Reference Data
   (iii) Business view Data
   (iv) Data Integration

(d) _____ refers to the data that is created and updated within the operational systems.
   (i) Transactional Data
   (ii) Reference Data
   (iii) Business view Data
   (iv) Data Integration

(e) _____ resulting in a single source of "truth" and making it easier for end users to access information.

  (i)   Data Profiling
  (ii) Data Control
  (iii) Data management
  (iv) Data Integration

[**Ans.:** *(a – iii), (b – iii), (c – ii), (d – i), (e – iv)*]

(2) **Fill in the blanks:**

  (a) _____ refers to the collection of raw facts and figures.

  (b) _____ refers to the data that is created and updated within the operational systems.

  (c) _____ is the data which will describe the type of object, like name, description, cost, length and dimensions of the object. It is a one-time data.

  (d) _____ is generated from data warehouse. Business views are calculations or summaries compared over period of time.

[**Ans.:** *(a) Data, (b) Transactional data, (c) Reference Data, (d) Business View Data*]

(3) **Match the Following:**

Group 'A'	Group 'B'
(a) Data Profiling	(i) achieving data accuracy and ensuring the right users have access to the right information
(b) Data Control	
(c) Data Integration	(ii) Data created and updated within the operational systems
(d) Data Augmentation	
(e) Database Marketing	(iii) adds value to base data
(f) Transactional Data	(iv) Integrate data in one system
(g) Reference Data	(v) identify potential problems with current data
	(vi) Data without which you cannot do any transactions and is mandatory for every organization
	(vii) unique segments in the database

[**Ans.:** *(a – v), (b – i), (c – iv), (d – iii), (e – vii), (f – ii), (g – vi)*]

(4) **True or False:**

  (a) Accurate information and reports are the life blood of an effective sales force.

  (b) Data control is all about understanding your data.

  (c) Data profiling is all about achieving data accuracy and ensuring the right users have access to the right information, which also means blocking access, as needed.

  (d) Data integration results in a single source of "truth" and making it easier for end users to access information.

  (e) Data Augmentation can help reduce the manual intervention required to developed meaningful information and insight of business data.

[**Ans.:** *(a) True, (b) False, (c) False, (d) True, (e) True*]

(5) What is Data Management? Discuss different types of Data. State various challenges in maintaining data quality?

(6) Describe various stages so as to plan and get quality information.

(7) Elaborate various data analysis techniques.

(8) **Write short notes on:**

  (a) Data Warehouse.
  (b) OLAP.
  (c) Data Reporting.
  (d) Collaborative Filtering.

☑ ☑ ☑