

THE INTERNATIONAL JOURNAL OF BUSINESS & MANAGEMENT

Applications of Business Intelligence in e –Commerce Websites – A Theoretical Overview of Data Warehouse Schema Design & Usage

S. N. Vivek Raj

Assistant Professor, Department of MBA, Panimalar Engineering College, Bangalore Trunk Road,
Varadharajapuram, Nazarethpettai, Poonamallee, Chennai, Tamil Nadu, India

Abstract:

In the last few years we have seen a tremendous shift in shopping patterns of the Indian customers. Customers are changing their shopping preference from brick and mortar companies to e-commerce companies. People of different age groups and various backgrounds are buying in the e commerce websites. Many brick and mortar companies are now modifying their business models to suit the new age internet economy. The e- commerce websites must measure their business periodically to dynamically modify the strategies and form ad- hoc strategies to suit the needs and wants of the various customers. These internet oriented businesses generate millions and millions of 'data' based on the volume of business. Only when these businesses tap these data and analyse to generate useful patterns about the customers, partners, delivery etc., they can sustain in the market and run the business profitably. Business intelligence is the science of converting data in to actionable information and presenting it to the decision making authorities usually managers to make informed decisions. To this cause, in this paper a well known technique of data warehousing has been demonstrated using schema design to highlight the use of this technology to store data and later use it for taking informed decisions.

Keywords: Data, Data-warehousing, Schema design, measures, business intelligence, decision making, e-commerce.

1. Introduction

Online shopping is one of the important trends of the decade that many will finger point. Traditionally Indian customers are used to shop in the physical stores. But due to the advent of e- commerce websites and good marketing techniques adopted by these firms, many customers have adopted shopping in e commerce websites. There are many statistical findings that state that the more number of Indians shop in popular e- commerce websites compared to the previous year. The Growth of the online business although a great potential for business fraternity, also presents an underling warning to the businesses running the e-commerce websites. The customers are well informed about the products for sale in the websites and information is readily available on the web. 'Word of Mouth' can easily spread through social media platform. Customers are instantly reviewing about the online shopping experiences and can easily bring a bad reputation to the company if something goes wrong. e-Commerce companies need to have systems with suitable control points wherever necessary to measure and control the operational performance 24 by7. Traditional measurement systems may fail since the entire business model differs for the e-commerce websites. E-Commerce websites themselves provide an inbuilt solution for the above cause. These websites generates huge amount of data(facts) in form of customer information, purchase information, website information, traffic information, device information which can be tapped for generating valuable patterns which can be used for strategy formulation, execution and control of the e commerce websites. But all the data generated from e commerce websites is not of use to us. The data warehouse – place where these data are stored can be suitably designed based on the business requirements and later the data stored in the warehouses can be mined for getting information and the same can be presented to the managers in form of reports, charts etc for decision making. In this paper a star schema has been suitably designed to give a testimonial to the use of data ware housing and applications of data mining and analytics for the internet based economy.

2. Review of Literature

There are various studies and researches that have been carried out in the area of data warehousing and business intelligence. A few of the important studies that relates to the present work have been discussed below. Kamal Nain Chopra (2014) has discussed concepts of business intelligence and its related tools like data warehousing, OLAP and their impact on improving enterprise effectiveness. Roy Mathew (2012) in his communication has highlighted the use of data warehouse and data mining for improving the quality of educational institutions and applying the same to knowledge management. Lynette Ryals and Adrian Payne (2001) has described the application of data warehouse and data marts for tapping, storing and managing complex customer data in financial services for effective management of relationship marketing. Abid Sohail, P.D.D. Dominic (2015) explains an approach of transformation by which ER model can be transformed to star model for the use of novice designers. Zack Jourdan, R. Kelly Rainer, and Thomas E. Marshall (2008) in their research analyzed information from a list of methodologically collected articles published from 1997 to 2006

and perceived that only a few research has been done in the domain of business intelligence by survey method and insisted that survey research can be done to a greater extent for adding more benefits to this area. Veronica Rozalia Rus, Valentin Toader (2008) explained the advantages of deploying business intelligence tools and techniques like Data warehouse, Data mart, OLAP, ETL (Extract, Transform, Load) for aiding the decision making process of the managers in hotel management. Juhnyoung Lee, Mark Podlaseck, Edith Schonberg, Robert Hoch and Stephen Gomory (2000) proposed a data mart architecture with selected metrics for electronic commerce. WANG Jian-bo, FAN Chong-jun (2012) described the importance of a data warehouse for an airport decision making, and they explained different data warehouse architectures and put forward a data warehouse architecture for airport.

3. Need for Data Ware House

Data ware house is a collection of historical data which are updated periodically. Data warehouse is used to store and analyze the key performance indicators. By means of periodically updating the key measures, the website can find out where it is performing optimally and where it is lagging.

Some of the applications and potential uses of data warehouse are as follows

- To improve the measures/metrics/KPI's which are lagging behind.
- To provide the executives with a snap shot of the website performance and updating it periodically
- To store the summarized data in a well organized data model
- To generate ad-hoc reports to the managers
- To assists further analysis like predictive analytics, data mining to unravel most interesting patterns about the customers online shopping behavior
- To store the customer profile data.
- To support decision making by providing data for generating reports, charts and dashboard
- Analysing data in various dimensions to get a multidimensional perspective of the data.
- To target the most profitable customers and make them feel special by giving special discounts.
- To understand why people are visiting the webpage and leaving without making a purchase
- To find the frequent user and customize the screen accordingly.
- To find the mostly accepted mode of purchase and use it accordingly.
- To Find the Traffic Rate/load in various time intervals and increase the server performance if necessary.
- To find the most frequently accessed page and page which is frequently bounced
- To find most preferred payment type for users.
- To Find the Server Response Characteristics
- To place Service level agreements (SLA'S) for the 3rd party companies who are maintaining the websites.

Data warehouse is the repository where all these measures are stored. Suitable schema design has to be formulated so that the facts are stored and updated periodically and can be used later.

4. Star Schema

In this paper a sample star schema for constructing a data warehouse and a few of the applications of the data stored in the warehouse would be discussed. For certain practical difficulties, a data warehouse schema is designed for a hypothetical e-commerce website which books flight tickets online. A Star schema contains two types of tables a fact table and dimension tables. In the Figure 1 the tables in the centre which are shaded in the both blue and yellow are called fact tables. The tables surrounding the fact tables (tables which are completely shaded in blue) are called as the dimension table. The fact table contains two different components. The first one is called Measures/ Metrics (shaded in yellow) for storing the numerical facts of the business. The measures can even be stored in an aggregate form. The next component of the fact table is foreign key (shaded in blue) containing references to the dimension tables.

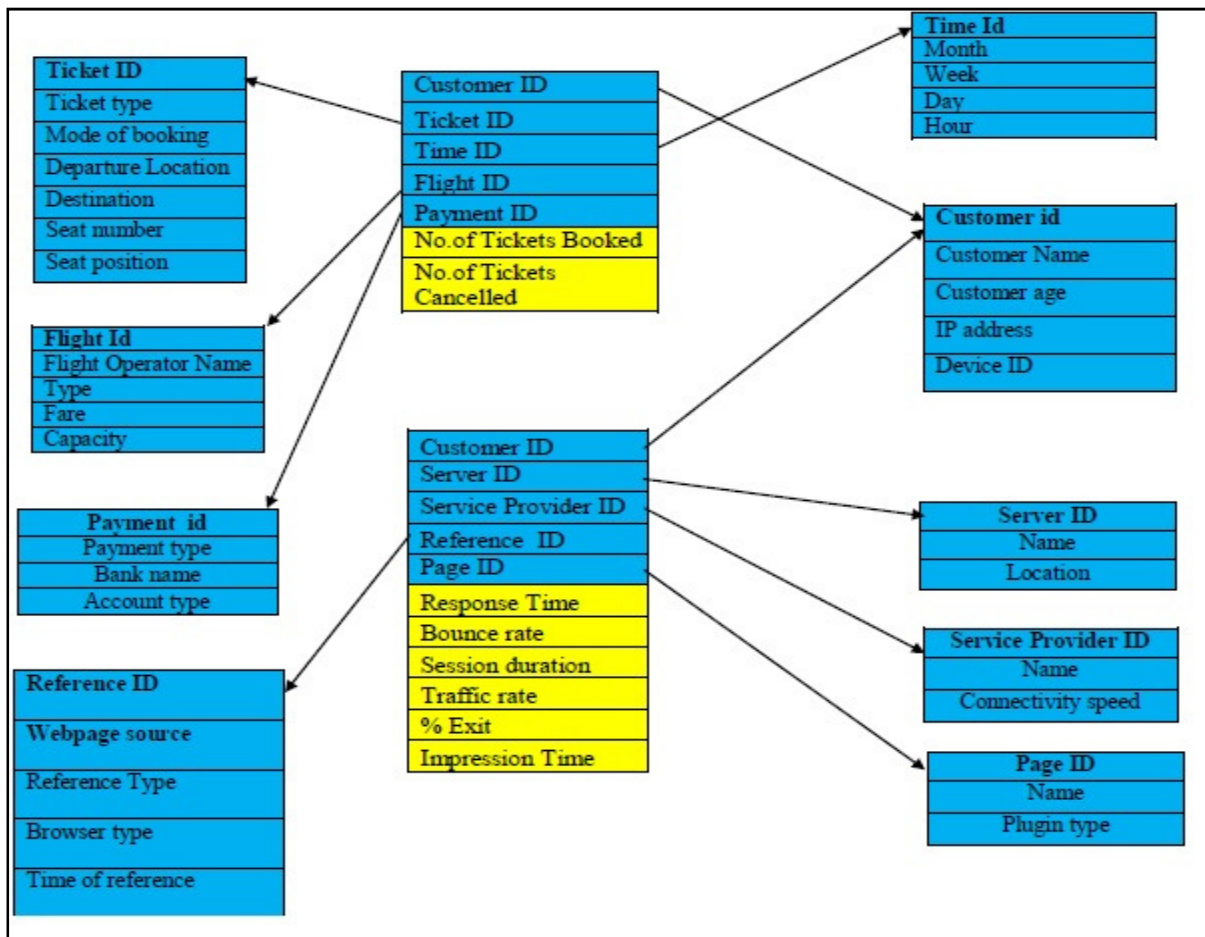


Figure 1: A Sample Star schema Design for a Flight Booking e-commerce Website.

5. Measures

The measures are the numerical facts that are stored in the central fact table which is used for the analysing the performance of the business enterprise, in this case an e-commerce website. There are various measures that can be used for measurement of the performance of an ecommerce website. In this present communication a few of the measures have been used and are listed below.

- Response Time
- Bounce rate
- Session duration
- Traffic rate
- Page View Duration
- % Exit
- Impression Time

6. Applications & Analysis of Data for Decision Making

A few of the applications and analysis that are possible are discussed below.

1. Using the time information stored, an analysis can be carried out to find the duration of the day or days of the week during which most number of tickets are booked. After finding out the time period during which there are frequent visitors online, the managers can place more number of people on the support team during that particular time, to help out the people booking tickets.
2. Using the payment information stored, most preferred payment mode like credit card, debit card, internet banking, most preferred bank used for booking ticket could be found and appropriate tie ups can be made with those financial institutions like providing bonus points based on the amount of booking, cash backs, EMI options, discounts etc. This would encourage more customers to do repeated/more purchase.
3. Analyse the profile of the customers booking ticket and finding out the most profitable customer group, frequent fliers etc. Based on the analysis of the customer profile/behaviour, the company can provide appropriate discounts, bonus points (which can be redeemed in next flight booking), targeted advertising etc.,
4. Finding out the demand of flights i.e. finding out the most favoured flight among the customers and fix the price of the flights accordingly.

5. Finding out the website response time during the different times of the day and finding out the time when the web server takes more time to respond. During peak hours web server may respond very slowly and this may make the potential customers to easily switch to the other website.
6. To find the out the service provider compatibility with the website and to sort out any technical difficulties if exists.
7. To find out compatibility with various mode of access. For example the tickets can be booked via various web browsers, various devices i.e. laptops, smart phones etc., and various operating systems. The web server characteristics like response time can be stored and analysed for various platforms. This is very important because the customers now have become very tech-savvy. They are accessing the websites through various devices with varied operating systems and browsers. The customers expect the websites to be reliable and inter operable. The technical team must make sure that the websites are interoperable.
8. To find out the percentage or count of visitors who jump from the website after seeing only a single page. This analysis can be further refined by drilling down the data further. i.e. grouping the number of visitors who navigate away from the website by age, location, marital status etc., and finding the out the group in which there are maximum percentage of bounce rate. By getting the online feedback from the particular group and finding out the reason which made them to navigate away from the website, the decision makers can eradicate the short comings.
9. The e-commerce administrator can find out how out the attractiveness of the website by finding out the time the users stay on the webpage. For example an ecommerce website having high number of website traffic but low session duration which denotes that the users are visiting the website but they are leaving it in a very short period of time due to various reasons . If it is the case then it is up to the website designers to find out the problems and make the website more attractive.
10. Many e commerce pages have a lot of plug-in installed. By analysing the average time taken for each page to load and respond , the website administrator or information systems team can find out the page/pages which are responding very slow or crashing frequently and modify the page by installing a fast loading plugin etc.,
11. Finding out the percentage of customers who have ordered the flight tickets in the shopping chart and abandoned the website without proceeding further. Using analytics we can further drill down to find out the exact group/cluster of customers that are having maximum number of abandonment.
12. Find out the number of times each page is viewed using impression metrics. This analysis helps to find the most viewed page and least viewed page.

The list above explains a few examples of the possible analysis that can be carried out from a data stored in a data warehouse. There are a lot of market ready on-the-shelf analytics tools which helps to generate ad-hoc reports ,build dynamic charts, track KPI's (Key Performance Indicators) by storing the data in multidimensional data bases (OLAP) and analysing the data in various dimensions e.g. slice and dice etc.. The Executives can be provided with the dashboard which updates periodically so that they can have a snap shot of what is happening in the website monthly, weekly, daily, even minute to minute. Ultimately the main aim is to provide the managers with right information at the right time.

7. Limitations & Future Directions

In this paper only a simple schema of a data warehouse is discussed. Not all the possible facts and dimensions have been shown. A real data warehouse would contain many detailed information than what has been discussed in the present communication. Only a few examples of the possible analysis of the data stored have been discussed. In reality many detailed data mining analysis and statistical techniques, predictive analysis can be used to find more interesting patterns out of the customer and website behaviour. The paper here concentrates only on a few common metrics for a flight booking e commerce website. The facts and dimensions would be different for different e commerce websites. There exists a huge scope for further research in this arena, where a real e-commerce website can be tracked based on various key performance indicators and predictive analytics can be used for predicting the buyer's online shopping behaviour.

8. References

- i. Lynette Ryals and Adrian Payne (2001). Customer relationship management in financial services: towards information enabled relationship marketing. *Journal of Strategic Marketing* 9(1), 3-27
- ii. Zack Jourdan, R. Kelly Rainer, and Thomas E. Marshall (2008). Business Intelligence: An Analysis of the Literature. *Information Systems Management*, 25(2), 121-131
- iii. Kamal Nain Chopra (2014). Modelling and Technical Analysis of Electronics Commerce and Predictive Analytics. *Journal of Internet Banking and Commerce*, 19(2), 1-10
- iv. Roy Mathew (2012).A Model for Organizing, Measuring, Analyzing Students' Knowledge and Performance. *International Journal of Research in Commerce, IT & Management*, 2(12), 32-38
- v. Abid Sohail; P.D.D. Dominic (2015).From ER model to Star model: a systematic transformation approach. *International Journal of Business Information Systems*, 18(3), 249-267
- vi. Veronica Rozalia Rus, Valentin Toader (2008). Business Intelligence for Hotels' Management Performance. *International Journal of Business Research*, 8(14), 150-154
- vii. Juhnyoung Lee, Mark Podlaseck, Edith Schonberg, Robert Hoch and Stephen Gomory (2000). Understanding Merchandising Effectiveness of Online Stores. *Electronic Markets*,10(1), 20-28
- viii. WANG Jian-bo, FAN Chong-jun (2012).Research on Airport Data Warehouse Architecture. *International Journal of Business, Humanities and Technology*, 2(4), 107-111