Need for Data Warehouse:

The concept of data warehousing is not hard to understand. The notion is to create a permanent storage space for the data needed to support reporting, analysis, and other BI functions. On the surface, it may seem wasteful to store data in more than one place.

Brazos University is looking to attain the highest quality of students into their programs for which they require strategic information to make critical decisions. The historic data is available but not being used in the right manner. By using a data warehouse, BU would be able to make on the fly decisions regarding admissions, course capacity increase and new facility buildings which would give them competitive edge over other competing universities.

The best data warehouses do some predigesting of the raw data in anticipation of the types of reports and inquiries that will be requested. This is done by developing and storing metadata (i.e., new fields such as averages, summaries, and deviations that are derived from the source data). There is some art involved in knowing what kinds of metadata will be useful in support of reporting and analysis. The best data warehouses include a rich variety of useful metadata fields. It is important for BU to create the requirements of their data warehouse and the types of reports they require which would help the design team in incorporating accurate data that can be analyzed.

Once a data warehouse is made operational, it is important that the data model remain stable. If it does not, then reports created from that data will need to be changed whenever the data model changes. New data fields and metadata need to be added over time in a way that does not require reports to be rewritten. Therefore, it is highly essential for the implementation team to make note of measures which would help BU create precise reports.

There is a need to integrate many different sources of data in near real-time. This will allow for better business decisions because BU users will have access to more data.  Plus this will save users lots of time because they won’t waste precious time retrieving data from multiple sources. Tons of historical data that is needed to gather would be easily accessible and will have common formats, common keys, common data model, and common access methods. Data can be restructured, tables and fields renames so that it makes more sense to users. In traditional database systems, users are running reports directly against operational systems, causing performance problems.  Instead, a data warehouse would help users run reports off of that.  Data warehouse is optimized for read access, resulting in faster report generation.

There is a risk that BI users might misuse or corrupt the transaction data. Having an easy to use data warehouse allows users to create their own reports without having to get IT involved.  Leading to “Self Service BI”. A data warehouse is a convenient place to create and store metadata. Improve data quality by cleaning up data as it is imported into the data warehouse (providing more accurate data) as well as providing consistent codes and descriptions. Having a data warehouse makes it easy to create business intelligence solutions on top of it, such as SSAS cubes.

Business Requirements:

BU has provided a list of requirements for the data warehouse in relation to the type of data, measures and reports to be obtained. The requirements are from various domains.

The first step is to define the various business dimensions that need to be tracked along with their relevant hierarchies. Business dimensions form the underlying basis of the new methodology for requirements definition. The business dimensions and their hierarchical levels form the basis for all further phases.

Business Dimensions:

BU has identified important domains of the university that need data to be tracked on a daily basis. These domains require dimensions to be defined in order to track measures. The dimensions are listed below along with their usage.

Applicant: The core information of BU is about its applicants, who apply from various countries and during different times in the year. The important attributes of the applicant that need to be tracked are:

* Applicant Name – First name and last name of the applicant are required as per federal regulations. Applicants can fill out applications through online web portal and at the university. The first and last name of the applicant can be retrieved from their application form.
* Applicant Address – All details related to the physical and contact location of the applicant are required. Details like country, state, city, street address and zip code need to be entered into the DW.
* Applicant Phone – The contact number of the applicant at which he/she can be contacted
* Applicant Birthday – The age can of the applicant can be helpful for planning coursework and more importantly understanding the age group of applicant pool.
* Marital Status and Gender – Demographic information is key to the university DW and this information can be retrieved from the applicant form.
* Examination Scores – Entrance exam scores of the applicant which would be used to make the final decision of admission is important to determine the quality of applicants.

Course: BU is very interested in understanding patterns related to course registration and the kind of students that register for a particular course. Reports related to tuition generated by a course and which professor takes that course are important for planning course enrollments. The important attributes of the course that need to be tracked are:

* Course Size: The maximum capacity/registrations that a course can have to be tracked as room requirements, professor availability and exam schedule all depend on this attribute.
* Course Fee: The cost of attending this course to be an important dimension.

Degree and Department: Information regarding degrees offered and the relevant departments offering them are to be tracked and measured. Department Head are officials in charge and one department can offer multiple degrees. The important attributes of the course that need to be tracked are:

* Degree Duration: The minimum and maximum duration of the degree will provide information regarding the active years of education of students.

Placement Domain: BU requires details in relation to student placement, graduate placement rate, base salary etc… The information would help BU keep in touch with alumni and ensure a strong alumni presence. From a marketing point of view, the salary statistics could influence future students to join BU. The important attributes of the course that need to be tracked are:

* Domain Category: The domain of the company where the students get a job offer
* Role: For which position/role the student has taken the job.

Facility: Since BU is looking to expand its student base, it also needs to ensure that its facilities are not overcrowded. Another aspect is maximum utilization of existing facilities and allocation of resources to building new required facilities. BU has requested for reports related to facility usage in relation with amount of revenue and the number of students attending/using the facility. The important attributes of the facility that need to be tracked are:

* Facility Capacity: Maximum number of people that can be accommodated in a facility.
* Facility Type: Each facility at BU is categorized into a number of facility types.

Faculty: BU has been known for its reputed faculty and are looking to track measures using their faculty and the courses they teach. The important attributes of the faculty that need to be tracked are:

* Position: The designation of the faculty is an important attribute that determines the courses they take and also the salaries of the faculty.
* Experience: The number of years of experience will help BU determine the growth and reputation of each faculty and the number of students that want to take up research or another opportunities under the faculty.

Program: Each department in BU offers a variety of programs. The programs are offered across different degrees.

Date: An important dimension to be tracked is date and time. Different metrics need to be measured at a daily, weekly, semester and annual level. The date dimension would be the foundation dimension along which all metrics would be measured.

Key Business Metrics

The approach for designing the data warehouse is such that each domain of the university would be modeled into a data mart with connecting dimensions, and an overall centralized schema would be designed by merging all data marts.

Admissions and Finance Metrics: As BU is looking for high quality students they require reports and analysis on the basis of applicants admission scores, acceptance rate based on scholarship offering, amount of federal loan dispersed etc… By understanding their applicant pool better, BU would be able to provide influence applicant decision by providing choice of semester enrollment, instate waiver for international and out-of-state applicants, and support of federal loan process. Important metrics to be tracked:

* Amount of Scholarship: This amount given to each applicant, tracked at a daily level and across programs would provide strategic information about the quality of the student and interest in different programs.
* Amount of Federal Loan: This is the amount borrowed by the student and can be compared against the amount of scholarship received per semester.

Facility Usage: Each facility needs to be tracked on a daily basis for the amount of revenue generated and the number of attendees. This would provide information regarding the utility of a particular facility and can give critical information to BU about which facility is overburdened and which is underutilized. Important metrics to be tracked:

* Number of Students: The total number of students (aggregated) that attend a particular event/stay at a facility.
* Revenue: Total amount of money received from ticket sales or rent of a particular facility.

Placement Statistics: Information regarding the companies that students get hired and their relevant roles would be an important measure from a marketing and administration perspective. Important metrics to be tracked:

* Total Salary: The amount of salary(aggregated) at a program level on a daily basis.
* NoofPlacements: The count of the number of student placements at a daily level across programs.

Course Registration: Since BU is intending to track course enrollment and the amount of tuition per course, measuring tuition received by a course taught by a particular faculty is a key measure.

* Total Tuition: The total amount of tuition received by a particular course.
* NumberOfStudents: The total number of students that have enrolled for a particular course.

Information Gathering Methodologies: Data would need to be collected through various sources like online web portals, existing OLTP databases, and conducting interviews with IT staff.