



Contemporary Analysis
Work smart.



One Version of the Truth

Creating a Business
Intelligence Data
Warehouse

*For business intelligence
to succeed there needs
to be one version of the
truth.*

Smart companies are data driven. They use spreadsheets to analyze performance, budgets, calculate return on investments (ROI), and develop plans.

However, this business intelligence is difficult to coordinate, usually resulting in everyone creating separate reports. With different reports, there are different answers to the same questions, often with slightly different conclusions. This leads to confusion and a lack of trust in business intelligence. This is called having multiple versions of the truth.

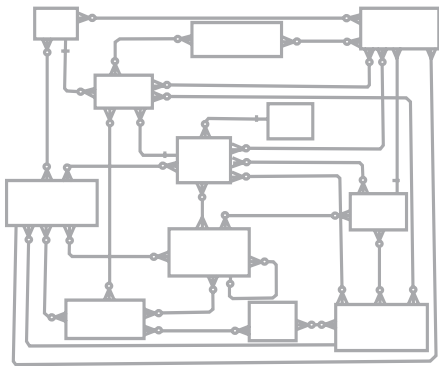
Companies can eliminate rework, confusion, and multiple versions of the truth by giving everyone access to the same data, the same basic tools, and the same platform for publishing, finding, and evaluating reports. *Employees will need different tools*, but for business intelligence to succeed employees need access to the same starting block. This starts by creating a Business Intelligence Data Warehouse.

What is a Business Intelligence Data Warehouse?

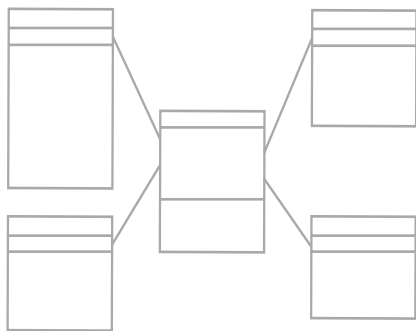
A Business Intelligence Data Warehouse allows users to ask complex questions by combining data from multiple internal sources, external sources, applications, surveys and studies to answer “*who did what, where, when, why, how, and how often?*”

The goal is to create simpler databases with fewer tables and table joins.

An OLTP Database:



vs. a BI Data Warehouse:



A Business Intelligence Data Warehouse could be used to find the answer: “What causes a customer to purchase our premium subscription?” The answer requires information about the demographics of previous, current and prospective customers, the marketing messages they receive, customer service interactions, previous usage, and checkout abandonments.

The best Business Intelligence Data Warehouses are designed around how people ask questions, instead of optimizing queries for computers. This is done by restructuring data around 7 questions: *who, what, when, where, why, how and how many/often*. The outcome is a simpler data structure, with fewer table joins, and simpler queries.

While a Business Intelligence Data Warehouse is focused on how people ask questions, by comparing historical events and finding context, most databases are designed for computers. They are called Online Transaction Processing Databases (OLTP), and are designed to quickly serve data to applications and websites. OLTP databases are optimized for speed and accuracy when creating, updating and deleting data, not answering questions.

An OLTP Database:

- **Designed to support applications**
- **Optimized for speed**
- **Support many concurrent users**
- **No duplication of data points**
- **Performs simple frequent tasks**
- **Is updated in near-real-time**

In contrast a Business Intelligence Data Warehouse:

- **Designed to support searches**
- **Optimized for volume and historical accuracy**
- **Has few concurrent users**
- **Performs infrequent and complex tasks**
- **Is updated periodically**

The critical difference between an OLTP and a Business Intelligence Data Warehouse is the intended user and their strengths. Humans are the intended users of Business Intelligence Data Warehouses, and computers are the intended users of OLTPs.

A case study.

Deciding if and where to go to college is one of the most important decisions people make. Attending university is time consuming and expensive. It also represents a significant sale for the University. The following is the average college published tuition for 2013:

Type of College source: collegeboard.org	Avg 2013 Published Yearly Tuition & Fees	Cost of a 4-year Degree
Public 2-year College (in-state students)	\$3,131	\$12,524
Public 4-year College (in-state students)	\$8,655	\$34,620
Public 4-year College (out-of-state students)	\$21,706	\$86,824
Private 4-year College	\$29,056	\$116,224

The college selection process is very complex. What university a student attends depends on their parents, parents income, friends, career goals, social goals, past educational performance, and many more variables. CAN was approached by a university that was interested in understanding their admissions process.

The problem was that the administration and faculty were *already* using data to make decisions, but that data was *corrupted*. Employees were *operating from multiple versions of the truth*.

Data was being delivered as spreadsheets. The IT staff was manually responding to requests for data and email users spreadsheets. This wouldn't have been a problem, but it took time to put together spreadsheets encouraging administrators and faculty to copy spreadsheets from past projects and do their own data manipulations without understanding the nuances of the data. This created an untraceable, uncontrollable and corrupt trail of data.

The solution.

The University already had a sophisticated OLTP data warehouse. It was well built and maintained by a dedicated database administrator. It was well designed and supported the applications that run the University. It delivered data quickly – as long as the request was standard and didn't need to combine multiple types of data. It wasn't designed for business intelligence.

Each custom data request required the database administrator to spend hours extracting and transforming data. The problem was that the database administrators were receiving more requests than they could fulfill.

The solution was to build a Business Intelligence Data Warehouse specifically designed to answer complex and custom requests and could be easily queried by non-technical users. Users could then access data using a query tool or accessing data through a data visualization tool such as Tableau.

The goal was to make getting data from the Business Intelligence Data Warehouse had to be easier than getting data from co-workers. Administrators and faculty already understood the importance of uncorrupted data, but the benefits had to outweigh the costs of accessing data.

Creating a Business Intelligence Data Warehouse also freed the database administrator from having to respond to custom requests, and instead provided a more predictable, manageable work-flow.

- **Front-load work:** During development CAN and the University created standard rules for extracting, transforming and loading data from the OLTP data warehouse into the Business Intelligence Data Warehouse.
- **One vs. Many Jobs:** Instead of having to respond to custom requests the database administrator loads data on the 15th of each month from the OLTP data warehouse into the Business Intelligence Warehouse.

- **Time for Strategy and Innovation:** The database administrator used to spend nearly all his time responding to requests for data and repairing the data warehouse. Now, no longer a “data monkey”, he has time to focus on developing strategies, data governance and bringing new data sources on-line.
- **Training and Supporting Users:** While the Business Intelligence Data Warehouse is easy to use it is still important to train users and answer their questions. Also, training and supporting users helps the database administrator understand how people use the system, system bugs, and user experience issues. This makes understanding how to improve the system easier.

The process

The first step in the project was to appoint a Project Champion. CAN needed input from the University to help design and test the BI Data Warehouse. The Project Champion represented the University, facilitated discussions, reduced waste, and helped optimize people’s time. The Project Champion was selected because he was a subject matter expert, aware of the available resources and people, and had the ability to connect CAN with those resources and people. *The Project Champion was responsible for reviewing CAN’s plans and work, and helping CAN overcome roadblocks by making introductions to people and resources.*

The project started by creating a blueprint for the Business Intelligence Data Warehouse. This was done before any data collection or coding began. The process of developing the blueprint consisted of interviews with faculty and students. These interviews helped CAN determine what questions the Business Intelligence Data Warehouse needed to answer, what data needed to be collected, and how different types of data that needed to be combined together to provide answers.

The result was a non-technical document that is easy to read and understand. CAN presented the blueprint to the administration and faculty to make sure the plans reflect the needs and wants of the University. Our goal is to make sure that we can answer the right questions the first time. The presentation included a demonstration of how the Business Intelligence Data Warehouse would answer specific questions.

Once the Business Intelligence Data Warehouse has been laid out, CAN identified the data sources that met the requirement of the blueprint. Data sources included both internal and external data. As necessary, CAN helped the University collect new data by adding fields to applications, changing storage requirements, developing surveys, and conducting focus groups.

Once the data had been identified, CAN used a variety of techniques and processes to extract data from outside sources, transformed it to meet operational and quality needs, and loaded it into the Business Intelligence Data Warehouse. This step is called ETL or Extract, Transform and Load. CAN also created an automated process for all new data to be loaded into the system going forward. The goal was to make updating the Business Intelligence Data Warehouse as easy to use as possible.

Once the data has been loaded into the Business Intelligence Data Warehouse *it now can be used to create data visualizations, reports, dashboards, predictive models, and conduct research.* University administration and faculty are now able to connect to a variety of Microsoft and other tools to run reports and create data visualizations. When they do, they can be certain that they have the latest version of the data, and that reports will be automatically updated as new data becomes available.

The last step is training. This is a valuable and important step as CAN does not want our clients to slip back into creating multiple versions of the truth simply because staff does not know how to use the system. CAN created a marketing campaign to build excitement and early adoption by communicating the value of the new Business Intelligence Data Warehouse that informs administration and faculty on key features.

In addition to creating a marketing campaign for the project, CAN provided training for faculty on how to use the BI Data Warehouse to ask questions. Training was provided using the Microsoft technology stack. In addition to training, CAN worked with Universities database administrator provided phone and email support as needed.

The result

The Business Intelligence Data Warehouse has reduced data and reporting errors. This was accomplished by increasing access to data instead of implementing new controls and roles. CAN was able to increase the confidence in business intelligence by reducing the incentive to re-distribute and use potentially corrupted data.

The University has been using the Business Intelligence Data Warehouse for a variety of projects but the primary use of the Business Intelligence Data Warehouse has been to inform and support marketing and recruiting at the University. It helps answer complex questions about how to improve recruiting, marketing, branding, messaging, culture, programming, enrollment, and student life.



Since 2008, Contemporary Analysis has used predictive analytics and data science to help companies of all sizes work smart.

Our solutions use data to help our clients improve their sales, marketing, customer service, management, and strategic plans.

Our solutions are used by fast-growing technology companies, Fortune 500s, as well as small- and medium-sized organizations. Our clients are in a variety of industries including construction, insurance, education, healthcare, government, not-for-profit, software and engineering.

Our vision is to make predictive analytics simple and affordable because all companies, not just the largest, should be able to benefit from predictive analytics and data science.

Our principles:

1. We care about business.

Each business deserves a custom solution. Problems are our passion.

2. We solve core business problems.

We make a big impact quickly. Value is our focus.

3. We don't have all the answers.

We help our clients make better decisions. Less wrong is the goal.

4. We are technology agnostic.

We focus on the solution. Technology is just a tool.

5. Our job is to solve problems, not introduce complexity.

Our solutions are simple because our clients are busy.