



S. S Jain Subodh P.G. (Autonomous) College

SUBJECT - DATA WAREHOUSING

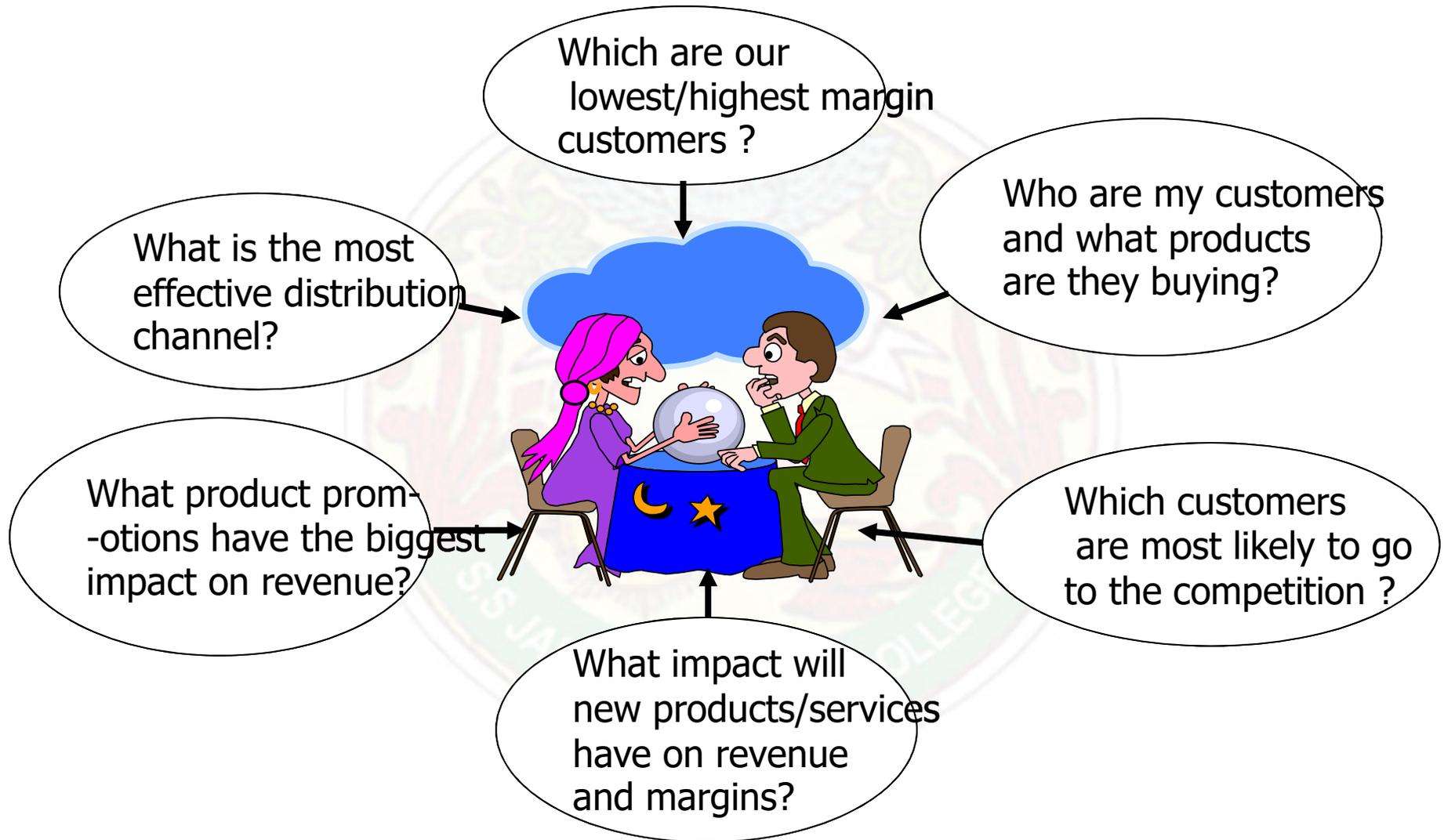
TITLE – INTRODUCTION TO DATA WAREHOUSING BY: Dr. VAIBHAV

Introduction to Data Warehousing





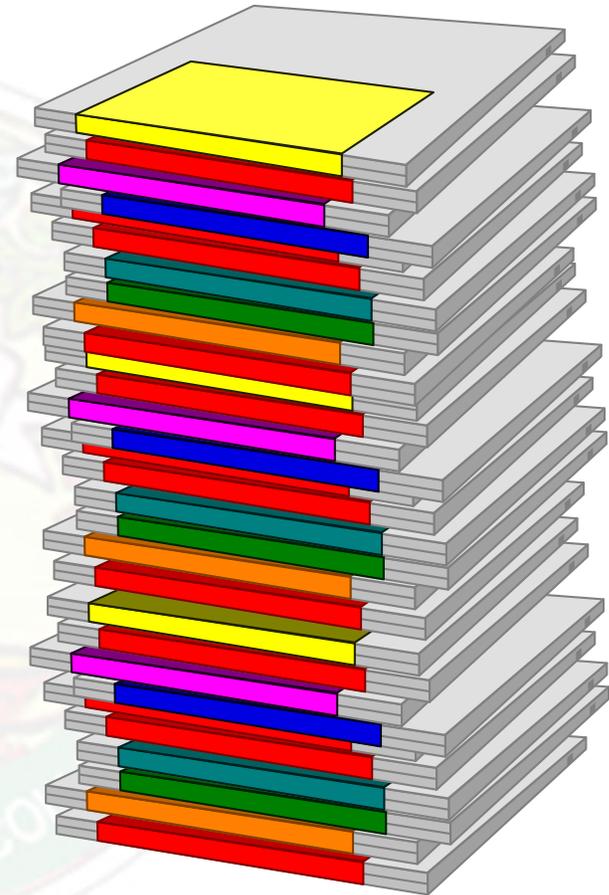
A producer wants to know....





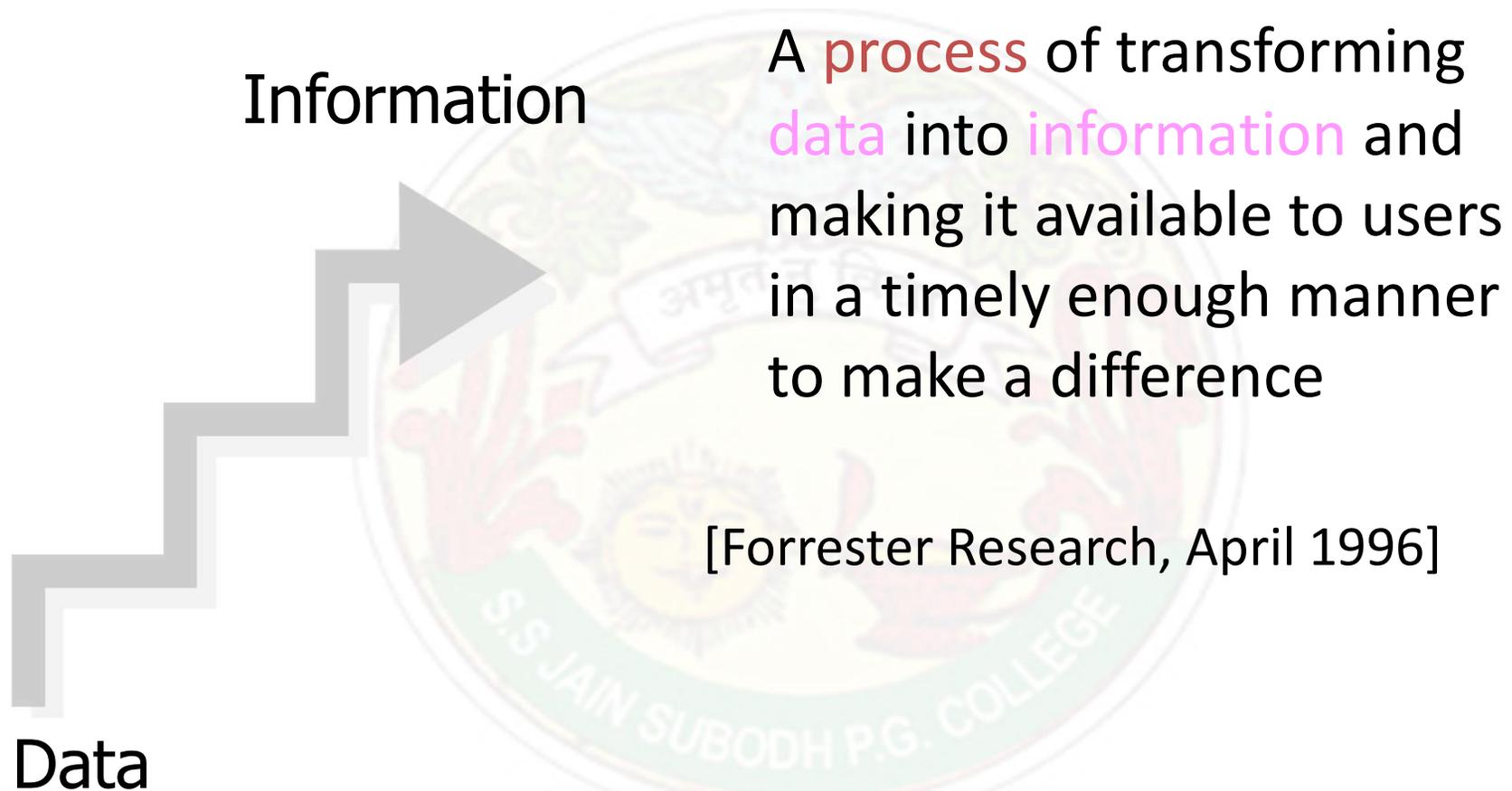
What is a Data Warehouse?

A single, complete and consistent store of data obtained from a variety of different sources made available to end users in a way that they can understand and use in a business context.





What is Data Warehousing?



A **process** of transforming **data** into **information** and making it available to users in a timely enough manner to make a difference

[Forrester Research, April 1996]

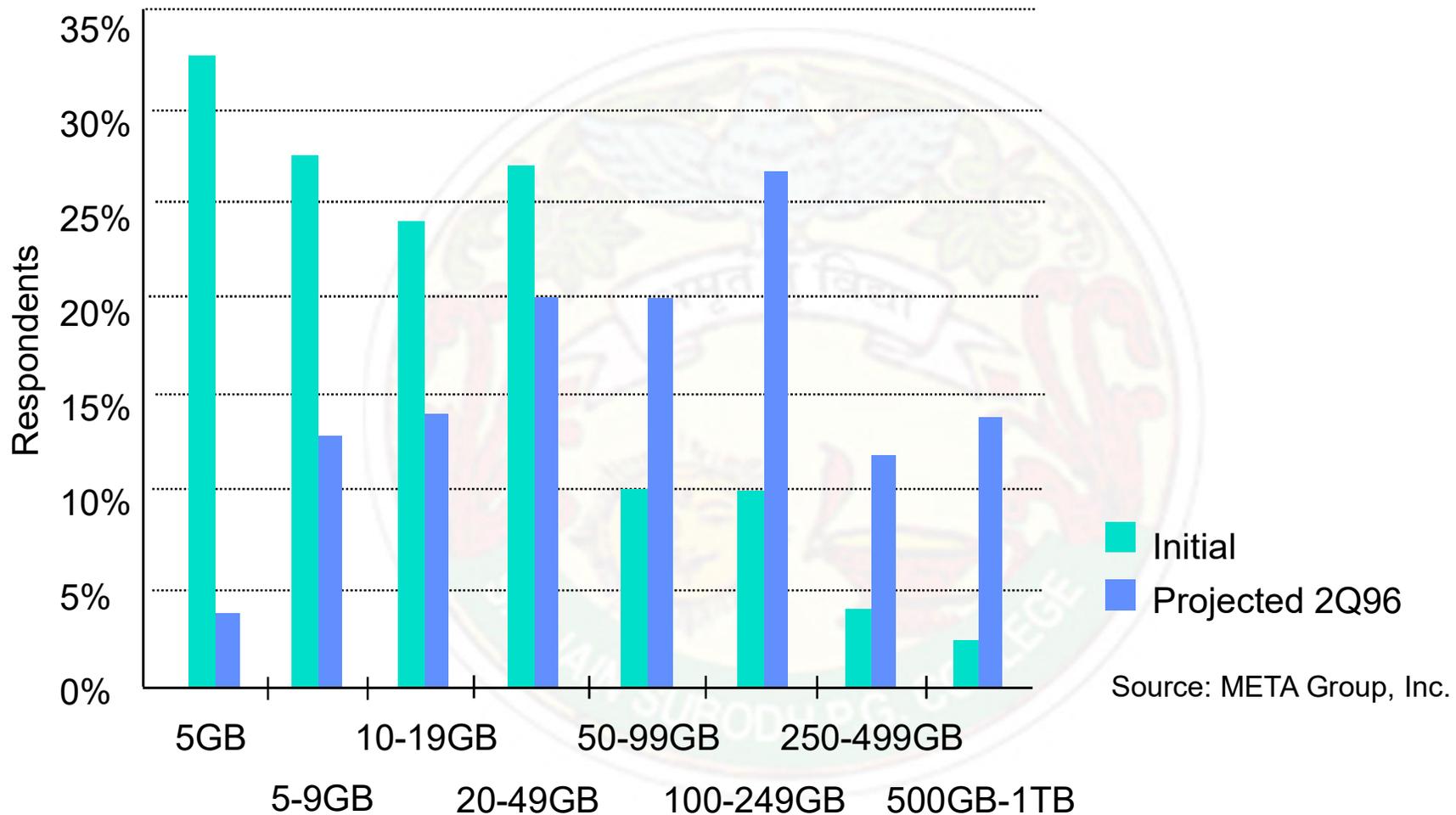


Evolution

- 60's: Batch reports
 - hard to find and analyze information
 - inflexible and expensive, reprogram every new request
- 70's: Terminal-based DSS and EIS (executive information systems)
 - still inflexible, not integrated with desktop tools
- 80's: Desktop data access and analysis tools
 - query tools, spreadsheets, GUIs
 - easier to use, but only access operational databases
- 90's: Data warehousing with integrated OLAP engines and tools



Warehouses are Very Large Databases





Very Large Data Bases

- Terabytes -- 10^{12} bytes: Walmart -- 24 Terabytes
- Petabytes -- 10^{15} bytes: Geographic Information Systems
- Exabytes -- 10^{18} bytes: National Medical Records
- Zettabytes -- 10^{21} bytes: Weather images
- Zottabytes -- 10^{24} bytes: Intelligence Agency Videos



Data Warehousing --

It is a process

- Technique for assembling and managing data from various sources for the purpose of answering business questions. Thus making decisions that were not previous possible



A decision support database maintained separately from the organization's operational database



Data Warehouse

- A data warehouse is a
 - subject-oriented
 - integrated
 - time-varying
 - non-volatile

collection of data that is used primarily in organizational decision making.

-- Bill Inmon, Building the Data Warehouse 1996

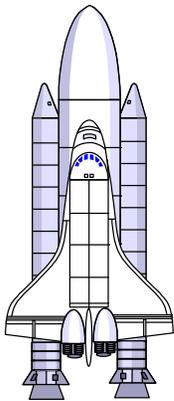


Explorers, Farmers and Tourists



Tourists: Browse information harvested by farmers

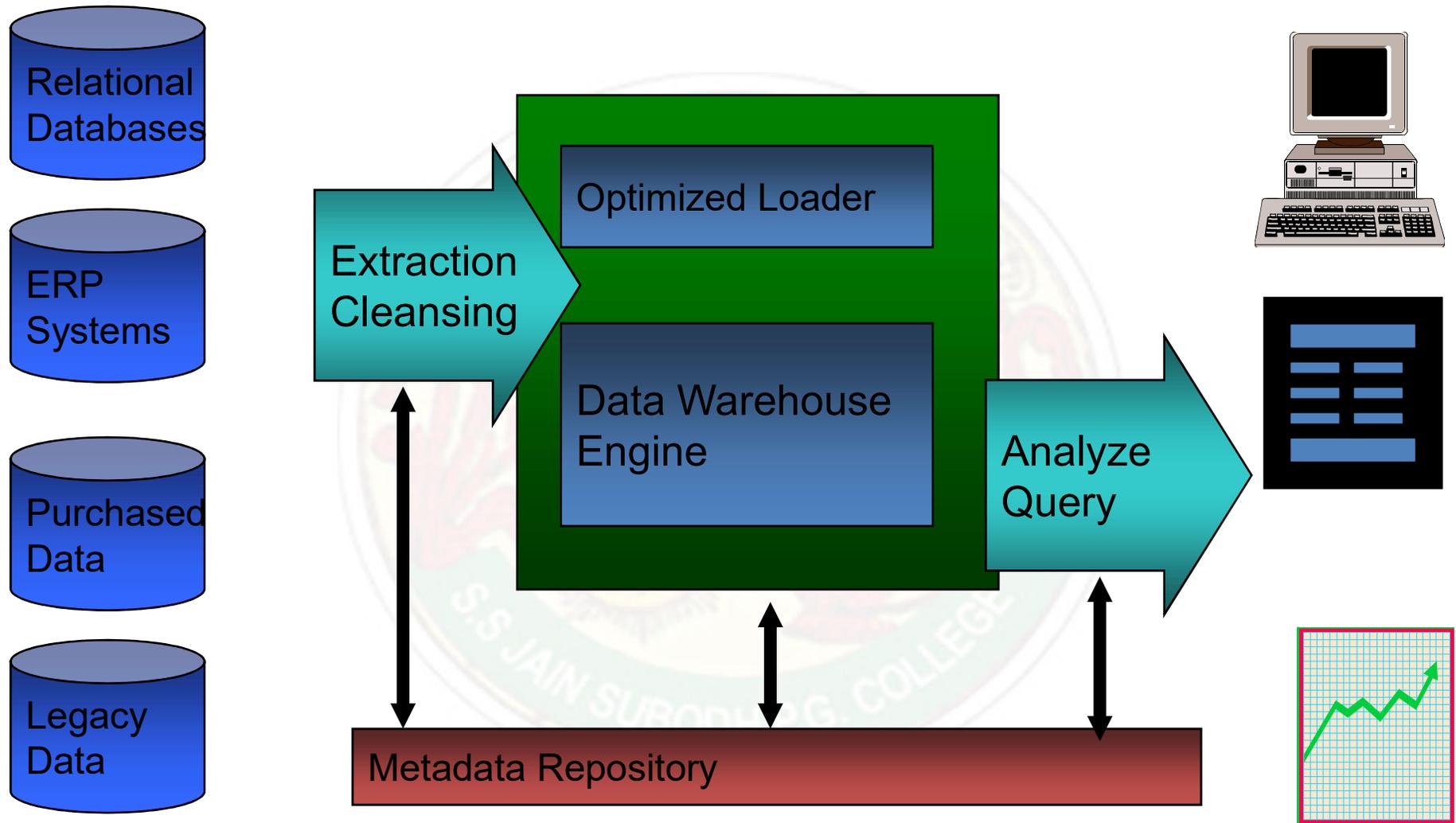
Farmers: Harvest information from known access paths



Explorers: Seek out the unknown and previously unsuspected rewards hiding in the detailed data



Data Warehouse Architecture





Data Warehouse for Decision Support & OLAP

- Putting Information technology to help the knowledge worker make faster and better decisions
 - Which of my customers are most likely to go to the competition?
 - What product promotions have the biggest impact on revenue?
 - How did the share price of software companies correlate with profits over last 10 years?



Decision Support

- Used to manage and control business
- Data is historical or point-in-time
- Optimized for inquiry rather than update
- Use of the system is loosely defined and can be ad-hoc
- Used by managers and end-users to understand the business and make judgements



S. S Jain Subodh P.G. (Autonomous) College

