Why we study data?

Data Intuition

Data

- Collection of facts about something that can be used in calculating, reasoning, or planning

Information

- Processed data used to make decisions and take actions

Data Life cycle

- Planning:
  - Type of data, How it going to be managed, who is the responsible for it.
- Capture
  - Collecting data (outside resources, inside resources)
- Manage
  - Data caring: store and maintain data, where it going to be stored,
- Analyze
  - Make use of data, help in solving problems, gain insights.
- Archive:
  - Stored in a place (might or might not be used again)
- Destroy
  - To protect data privacy and users' privacy

So, why we study data?

- Improve people’s lives
- Make informed decision
- Find Solutions
- No more guessing
- Be strategic with a free experience
- Know your baselines and identify goals

Business Intelligence 01

Hady Khaled Hammad, MEng
Faculty of Engineering, University of Ottawa - Canada

What is Business Intelligence?

- Make use of firm’s data
- Create Dashboards and Reports
- Highlighting the problems and its causes

“Business strategy for more proactive decisions”

BI importance

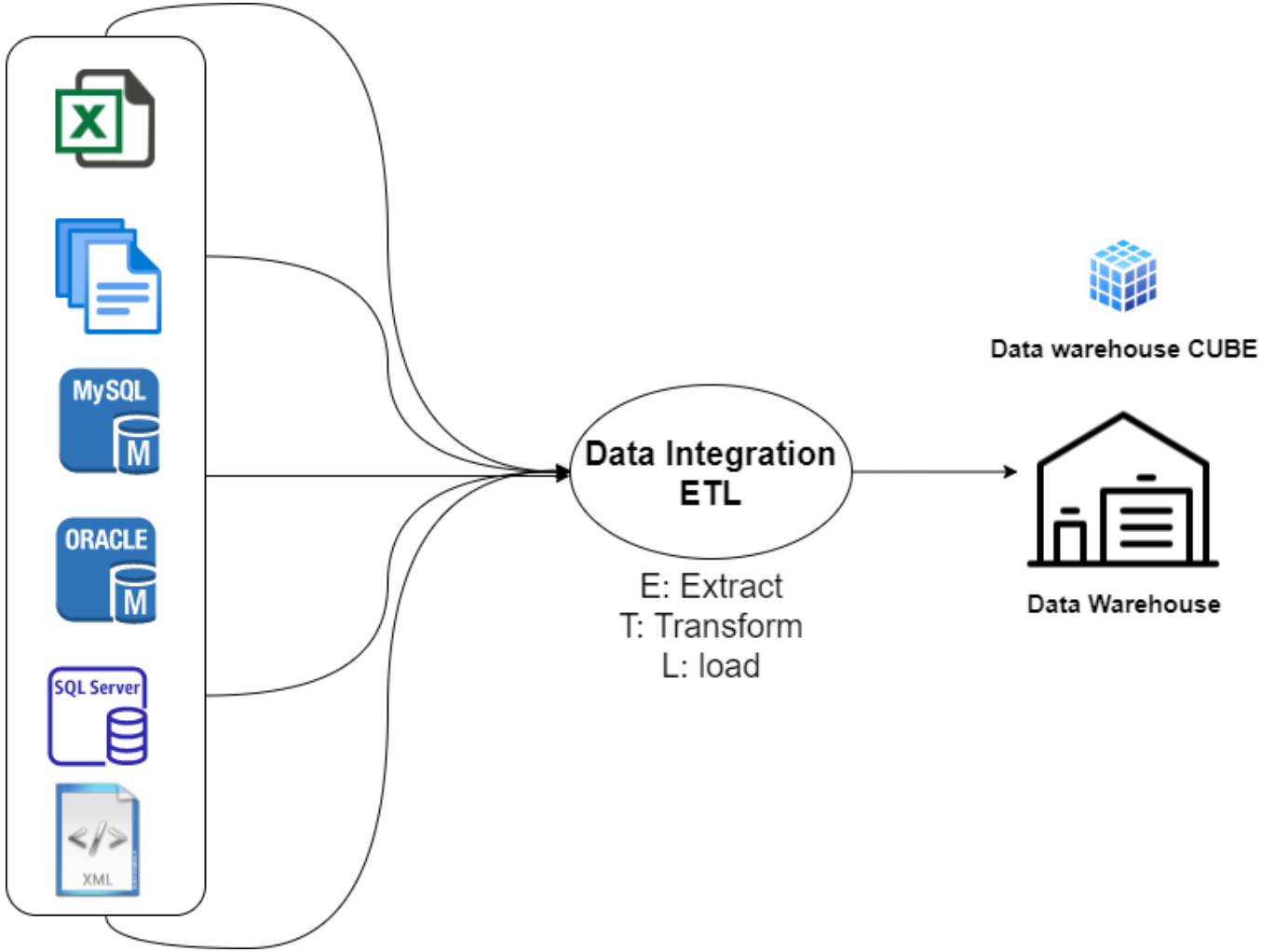
- Create KPIs and Measurements based on historic data
- Identify and set business baselines and goals
- Identify market trends and spot business problems
- Identify the areas that need attention
- Gives an overall view of the business’ process

BI Process

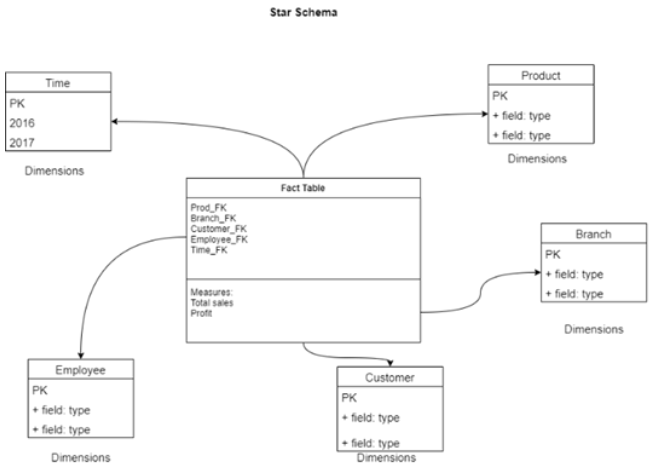
1. Data Sources
2. Data Integration
3. Data Analysis
4. Data Reporting

Data Warehouse

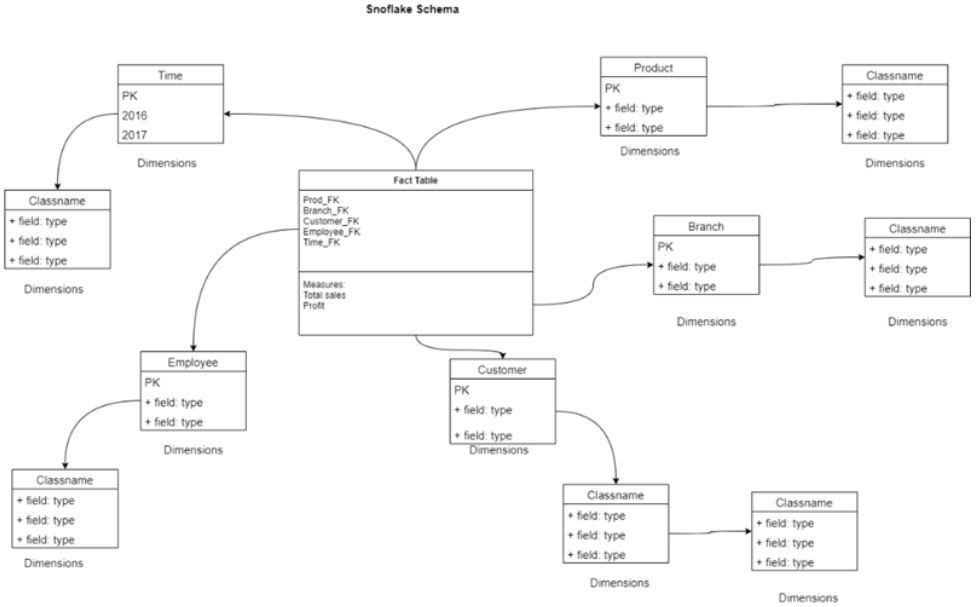
- Integrating data from multiple sources (Denormalized)
- Relational database that is designed for query and analysis (OLAP)
- Subject oriented, Time variant & Non-volatile
- Schemas: Star, Snowflake schemas, Galaxy schema (Fact constellation)
- Using cubes for analysis


MySQL
ORACLE
SQL Server
XML
Data Integration
ETL
E: Extract
T: Transform
L: load
Data warehouse CUBE
Data Warehouse

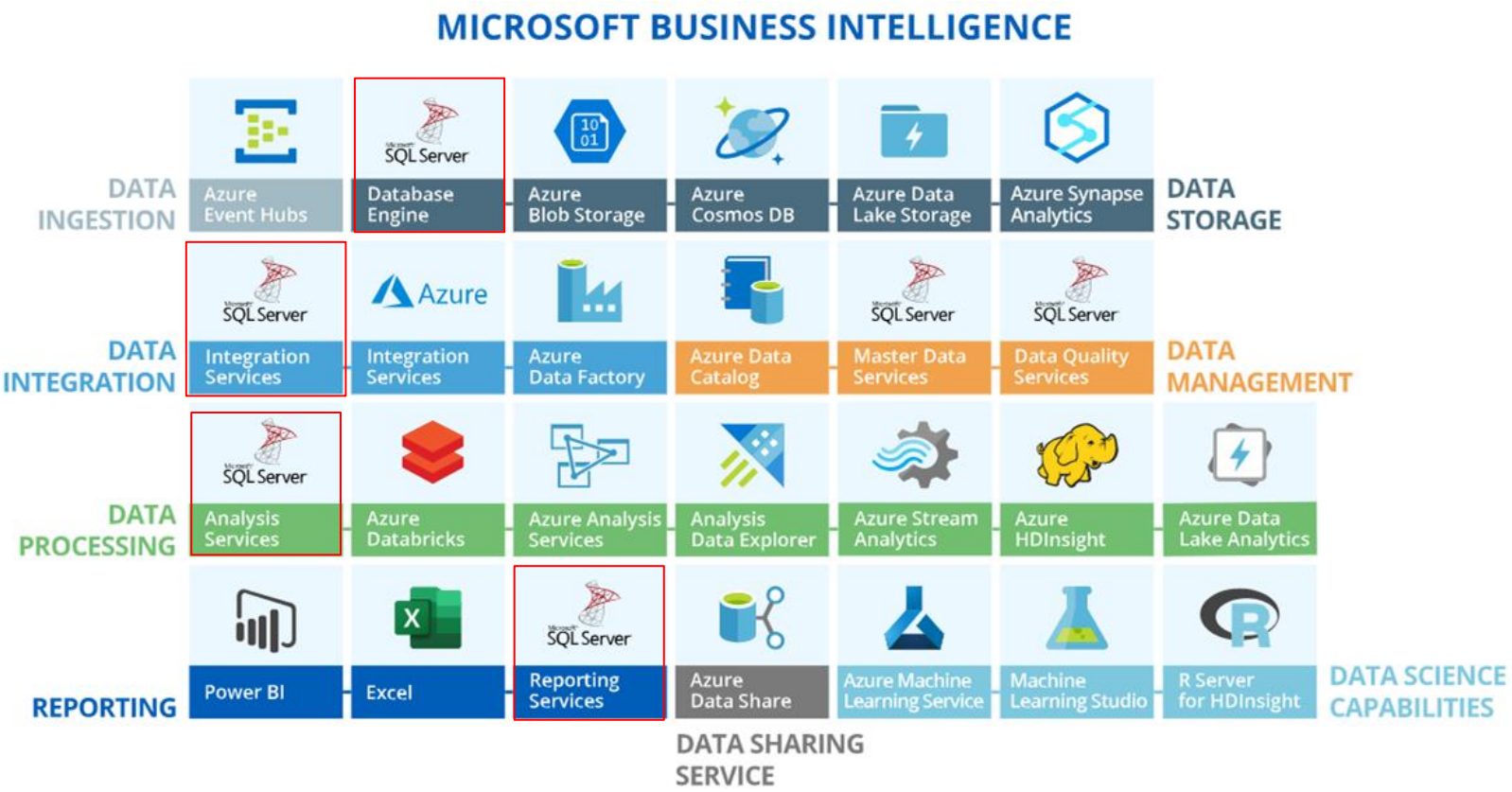
Star schema



Snowflake schema



Tools


MICROSOFT BUSINESS INTELLIGENCE
DATA INGESTION
Azure Event Hubs
SQL Server Database Engine
Azure Blob Storage
Azure Cosmos DB
Azure Data Lake Storage
Azure Synapse Analytics
DATA STORAGE
DATA INTEGRATION
SQL Server Integration Services
Azure Integration Services
Azure Data Factory
Azure Data Catalog
SQL Server Master Data Services
SQL Server Data Quality Services
DATA MANAGEMENT
DATA PROCESSING
SQL Server Analysis Services
Azure Databricks
Azure Analysis Services
Analysis Data Explorer
Azure Stream Analytics
Azure HDInsight
Azure Data Lake Analytics
REPORTING
Power BI
Excel
SQL Server Reporting Services
Azure Data Share
Azure Machine Learning Service
Machine Learning Studio
R Server for HDInsight
DATA SCIENCE CAPABILITIES
DATA SHARING SERVICE

OLTP Vs OLAP

- OLTP: Online Transactional Processing
  - Enable the execution of a large volume of database transactions, performed by multiple people.

- OLAP: Online Analytical Processing
  - Systems that are optimized for performing analyses on large volumes of data

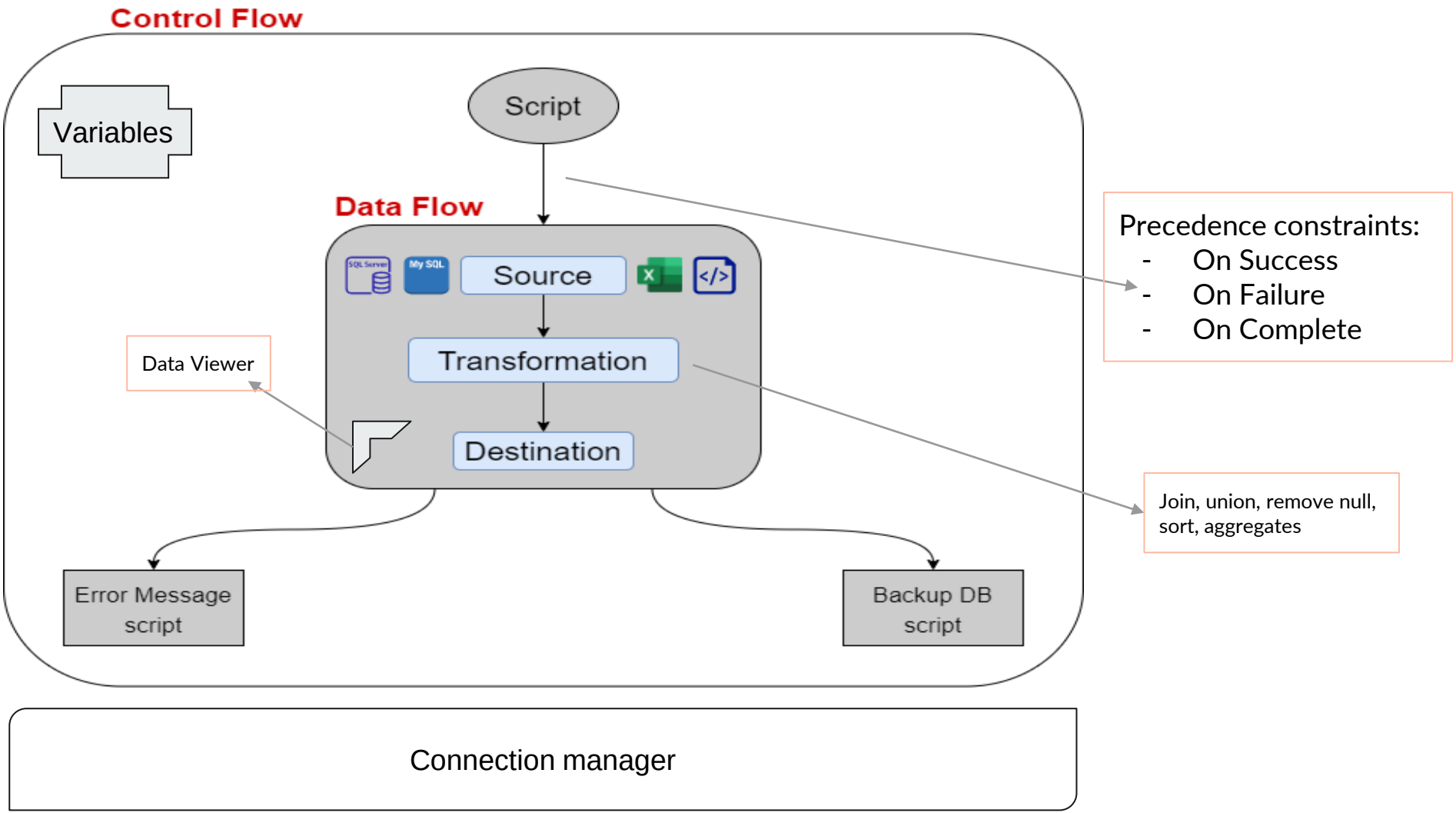
OLTP:

- Insert, Update, Delete for daily transactions
- Uptodate (Volatile data)
- Insert in real time
- Simple queries
- Should be highly available
- Less processing time (handle large amount of simple queries)
- Schema (Relational Database)
- Normalized
- Application oriented

OLAP:

- Report & analyze data, no daily transactions
- Historical data (Non-volatile data)
- Insert by a schedule
- Very complex Queries (Cube)
- Less availability requirement
- More processing time (handle small amount of complex queries)
- Schema (High Dimensional Schema)
- Denormalized
- Subject oriented

SSIS


Control Flow
Variables
Script
Data Flow
Source
Transformation
Destination
Data Viewer
Precedence constraints:
- On Success
- On Failure
- On Complete
Join, union, remove null, sort, aggregates
Error Message script
Backup DB script
Connection manager