

**Big Data generated by  
Organizations:**

**Structured but often  
siloed**

# After this video you will be able to..

- Describe the types of data organizations collect
- Explain the term structured data
- Cite the reasons for integrating structured data in different silos

# How organizations produce data

Commercial  
Transactions

Government  
Open Data

...

Banking/Stock  
Records

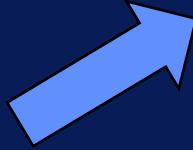
E-Commerce

Credit  
Cards

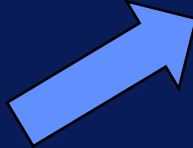
Medical  
Records

# Sale transaction data





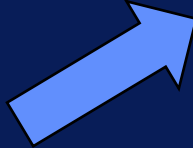
Detect correlated  
products



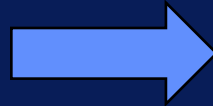
Detect correlated  
products



Estimate demand



Detect correlated  
products



Estimate demand



Capture fraudulent  
activity

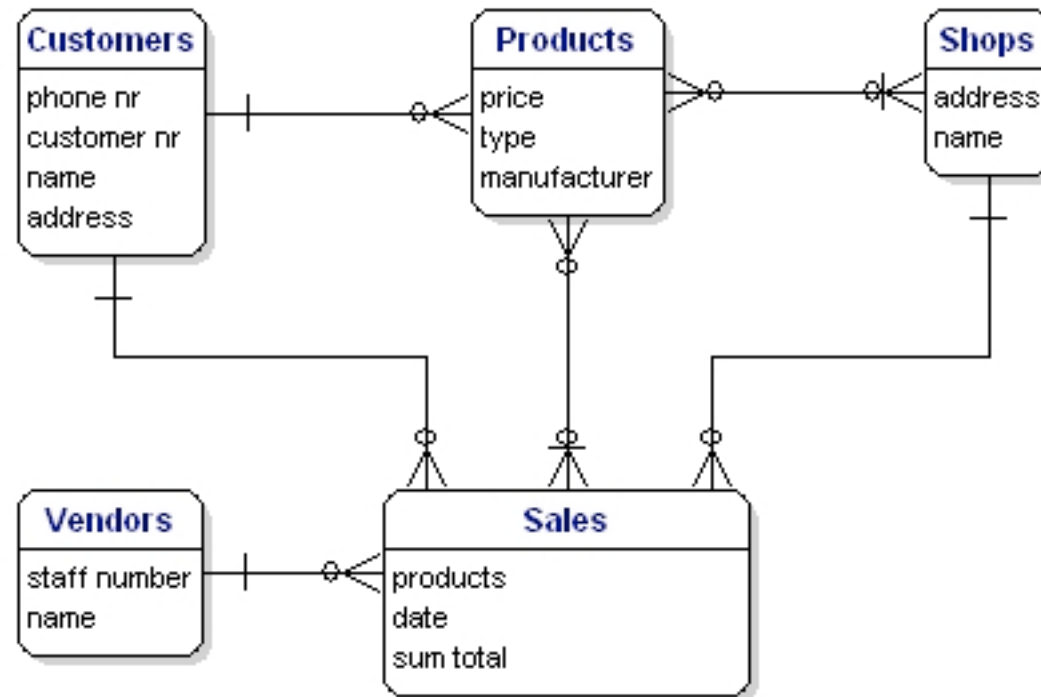


Analytics

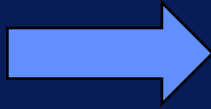
Better predictions



# Highly structured data



# Sales Transaction Records



Sale ID	Time	Customer	Product ID	Quantity
S00001	12/1/2012 9:00:00 AM	C0001	P025	1
S00002	12/1/2012 9:05:58 AM	C0025	P025	3
S00003	12/1/2012 9:11:33 AM	C0010	P001	2
S00004	12/1/2012 9:17:16 AM	C0017	P023	4
S00005	12/1/2012 9:23:04 AM	C0018	P016	5
S00006	12/1/2012 9:28:43 AM	C0011	P018	4
S00007	12/1/2012 9:34:07 AM	C0015	P006	1

# Sales Transaction Records



Sale ID	Time	Customer	Product ID	Quantity
S00001	12/1/2012 9:00:00 AM	C0001	P025	1
S00002	12/1/2012 9:05:58 AM	C0025	P025	3
S00003	12/1/2012 9:11:33 AM	C0010	P001	2
S00004	12/1/2012 9:17:16 AM	C0017	P023	4
S00005	12/1/2012 9:23:04 AM	C0018	P016	5
S00006	12/1/2012 9:28:43 AM	C0011	P018	4
S00007	12/1/2012 9:34:07 AM	C0015	P006	1

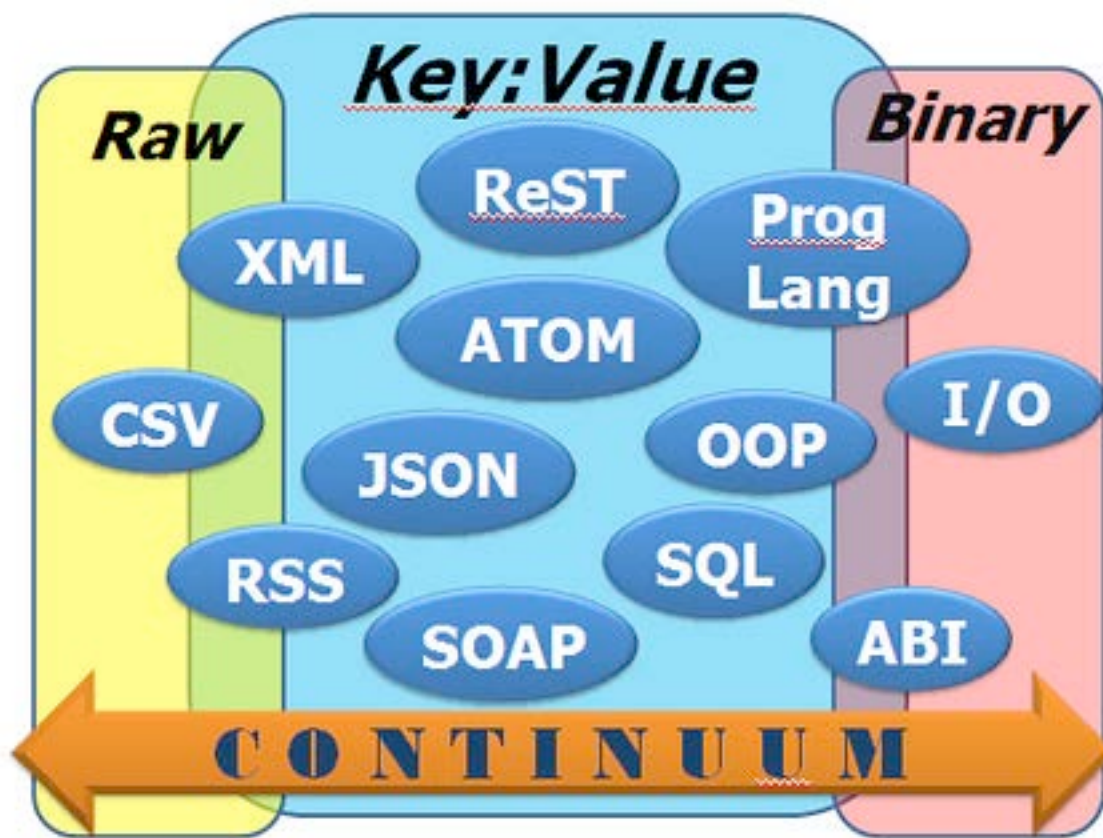
# Sales Transaction Records



Sale ID	Time	Customer	Product ID	Quantity
S00001	12/1/2012 9:00:00 AM	C0001	P025	1
S00002	12/1/2012 9:05:58 AM	C0025	P025	3
S00003	12/1/2012 9:11:33 AM	C0010	P001	2
S00004	12/1/2012 9:17:16 AM	C0017	P023	4
S00005	12/1/2012 9:23:04 AM	C0018	P016	5
S00006	12/1/2012 9:28:43 AM	C0011	P018	4
S00007	12/1/2012 9:34:07 AM	C0015	P006	1

**S/W Produced Data**

## "Structured Data": A Continuum



**H/W Produced Data**

# Data silos within an organization!



# Data Silos



**Hindered  
opportunity  
generation**



**Unconnected  
information  
islands**

# Data Silos

Hindered  
opportunity  
generation

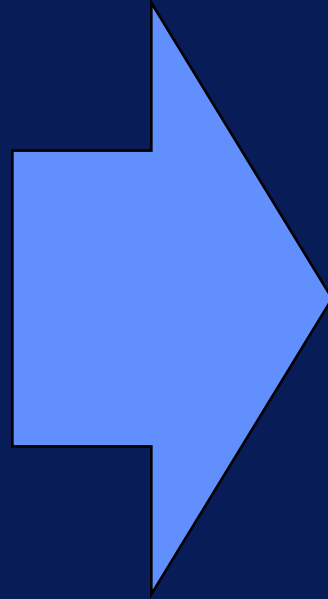
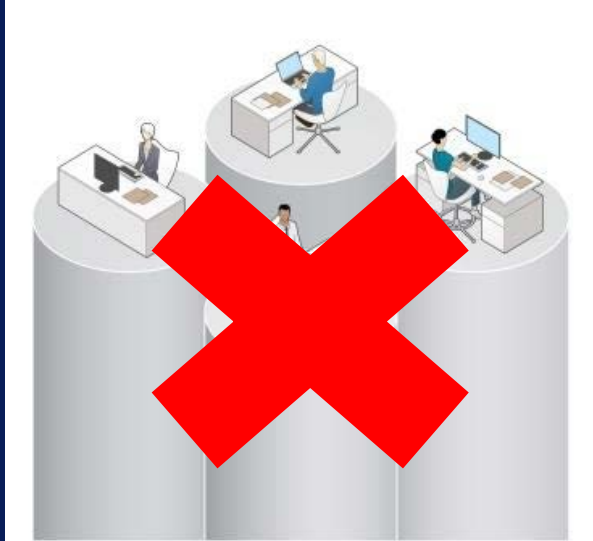
Unconnected  
information  
islands

Outdated,  
unsynchronized,  
even invisible data





# Highly-structured organizational data



## Value!