

TIPS TO ORGANIZING DATA

Using accurate, complete, and duplicate free data to make informed decisions

When evaluating how to collect your data, it is important to understand how the data will be stored.

Data Building Blocks

Entity

An **entity** is a person, place, object, or event about which we want to store and report information. For example, businesses need to keep track of customers, sales employees, inventory and payments.



Field

A **field** is an attribute of an entity. Think of a field as a category of data.

* There are 8 fields in the in the *Customer List*, including First, Last, and Birth Date.



Record

A **record** is a collection of related fields, each associated with the same entity.

* The top row in the *Customer List* is record for Ginny Braithwaite.



Table

A **table** is a collection of records, each having the same field structure.

* The *Customer List* is a table.



Database

A **database** is a collection of tables that help an organization manage its entities.



Customer List

	A	B	C	D	E	F	G	H	I
2	1	Ginny	Braithwaite	3 Which Way	Salem	MA	1970	1/10/60	
3	2	Virginia	Rodarmor	123 Main Street	Andover	MA	1810	1/6/70	
4	3	Kristen	Reis	4848 Ashley	Fontanelle	IA	50810	3/18/68	
5	4	Tom	Reis	3848 Ashley	Fontanelle	IA	50810	7/3/65	
6	5	Mark	Eagan	987 Lincoln	Schaumberg	IL	44433	9/3/85	
7	6	Peg	Fox	125 Maple	De Moines	IA	50625	10/18/98	
8	7	Ron	Fox	125 Maple	De Moines	IA	70524	8/28/87	
9	8	Amanda	Fox	125 Maple	De Moines	IA	81128	1/30/88	
10	9	Rebecca	Gross	123 Oak	Bridgwater	KS	91732	9/20/63	
11	10	Robin	Spencer	293 Serenity Dr	Concord	MA	102336	1/30/50	
12	11	Camilla	Dobbins	486 Inel Circuit	Rio Rancho	NM	112940	3/15/65	
13	12	Pip	Khalsa	1100 Vista Road	Santa Fe	NM	123544	4/16/69	
14	13	Kendra	Majors	530 Spring Street	Lenox	MA	134148	5/4/70	
15	14	Tasha	Williams	Purgatory Ski Area	Lenox	MA	144752	5/8/74	
16	15	Fred	Gonzales	11 River Road	Durango	CO	155356	6/10/60	
17	16	John	Black	72 Cubs Way	Brookfield	CT	165961	7/15/65	
18	17	Scott	Owen	3434 Hazelwood	Brookfield	CO	176565	3/15/65	
19	18	Ralph	Norton	42 Entrepreneur	Rio Grande	WY	187169	4/19/72	
20	19	Robert	Reid	9 Solutions Dr	Santa Clara	WY	197773	8/9/84	
21	20	Ned	Rierson	33 Dorsett Dr	Iona	ID	208377	9/30/78	
22									
23									
24									

Primary Key: In order to distinguish one record from any other record, each table should have a primary key defined. A primary key is the central field that uniquely identifies each each record in atable.

* The *Customer Number* (Cust #) is the primary key on the *Customer List*.

Foreign Key: When a primary key is also included in another table for the purpose of linking the fields together, it is called a foreign key.

* The *Invoice Number* (Inv #) is the primary key for the *Invoice List* and the *Customer Number* (Cust #) is the foreign key.

Invoice List

	A	B	C	D	E	F	G	H	I	J	K
1	Invoice #	Cust #	Invoice Date	EventID	First	Last	Street	City	State	ZIP	Birth Date
2	1	12	8/1/17	Bigs0197	Pip	Khalsa	1100 Vista Road	Santa Fe	NM	123544	4/16/69
3	2	8	8/2/17	Bigs0198	Amanda	Fox	125 Maple	De Moines	IA	81128	1/30/88
4	3	4	8/3/17	Bigs0199	Tom	Reis	3848 Ashley	Fontanelle	IA	50810	7/3/65
5	4	1	8/8/17	Road0696	Ginny	Braithwaite	3 Which Way	Salem	MA	1970	1/10/60
6	5	13	8/8/17	MTBI0997	Kendra	Majors	530 Spring Street	Lenox	MA	134148	5/4/70
7	6	18	8/9/17	MTBI0997	Ralph	Norton	42 Entrepreneur	Rio Grande	WY	187169	4/19/72
8	7	20	8/9/17	MTBI0997	Ned	Rierson	33 Dorsett Dr	Iona	ID	208377	9/30/78
9	8	3	8/11/17	Road0696	Kristen	Reis	4848 Ashley	Fontanelle	IA	50810	3/18/68
10	9	3	8/11/17	Road0696	Kristen	Reis	4848 Ashley	Fontanelle	IA	50810	3/18/68
11	10	11	8/12/17	Fun3333	Camilla	Dobbins	486 Inel Circuit	Rio Rancho	NM	112940	3/15/65
12	11	14	8/12/17	NUF777	Tasha	Williams	Purgatory Ski Area	Lenox	MA	144752	5/8/74
13	12	10	8/13/17	NUF777	Robin	Spencer	293 Serenity Dr	Concord	MA	102336	1/30/50
14	13	7	8/13/17	MTBI0997	Ron	Fox	125 Maple	De Moines	IA	70524	8/28/87
15	14	3	8/14/17	Road0696	Kristen	Reis	4848 Ashley	Fontanelle	IA	50810	3/18/68
16	15	6	8/14/17	Fun3333	Peg	Fox	125 Maple	De Moines	IA	50625	10/18/98
17	16	9	8/15/17	Fun3333	Rebecca	Gross	123 Oak	Bridgwater	KS	91732	9/20/63
18	17	5	8/16/17	Bigs0199	Mark	Eagan	987 Lincoln	Schaumberg	IL	44433	9/3/85
19	18	2	8/16/17	MTBI0997	Virginia	Rodarmor	123 Main Street	Andover	MA	1810	1/6/70
20	19	17	8/17/17	NUF777	Scott	Owen	3434 Hazelwood	Brookfield	CO	176565	3/15/65
21	20	19	8/17/17	Bigs0199	Robert	Reid	9 Solutions Dr	Santa Clara	WY	197773	8/9/84
22											
23											
24											
25											

Inv # = Primary Key

Cust # = Foreign Key



Data Duplication

When combining information from multiple tables into one table, we begin to see data duplication. Look at *rows 9, 10, and 15* on the *Invoice List*. Kristen Reis' customer information has been entered *3 times*.

30%

It is common for many small to medium-sized businesses to have up to 30% duplicate records in their business processes.

Duplicate data causes:



Inaccurate Reporting and Less Informed Decisions



Poor Business Process and Low Employee Confidence



Revision Control Challenges and Increased Overhead

If you plan on using your data to make informed decisions, you need to make sure your data is accurate, complete, and duplicate free.