

Appendix 3.06.3 Normalisation of data

Of all the concepts it is perhaps this one which will live with us in accounting systems as we modify them for multiple-entry – which we believe will be the main type of the more modern accounting system – especially with blockchains, which we come to next.

As a simplified example, let us imagine we are creating a database of the sales to Elon Musk’s Tesla car and Space X rocket companies. When starting to build this database, the first approach might be to create a simple table with all of the information in one place, and one row for each transaction. A so-called flat file. Among other issues, this has redundancy, data integrity issues and makes for inefficient processing. See Table A6.1.

Table A6.1 Sales example for a relational database system

Trans- action	Salesperson	Address	Product	Units	Unit price	Customer	Location
T001	Heather	Freemont	Manifold	5,000	£250	Tesla	Freemont
T002	Rachel	Los Angeles	Battery	6,000	£3,000	Tesla	Freemont
T003	Jimmy	Hawthorne	Sheet steel	3,000	£1,000	Tesla	Hawthorne
T004	Lola	Hawthorne	Fairing	12	£25,000	Space X	Hawthorne
T005	Donald	Freemont	Manifold	75	£250	Tesla	Freemont
T006	Alex	Freemont	Sheet steel	25	£1,000	Space X	Hawthorne
T007	Rachel	Los Angeles	Battery	40	£3,000	Space X	Hawthorne
T007	Heather	Los Angeles	Battery	20	£3,000	Space X	Hawthorne

The problem is that there is redundancy; the same data items appear in numerous places. There are one to-many type relationships, such as salesperson to customer. Rachel supplies batteries to Tesla and Space X. (Heather deals with Tesla and Space X but with two different products.). In this case there are four attributes which can act as keys. The salesperson, the customers, and the products. These three tables are shown in Tables A6.2, A6.3 and A6.4 respectively.

Table A6.2 Salesperson Table		
Salesperson ID	Salesperson	Address
11	Heather	Freemont
12	Rachel	Los Angeles
13	Jimmy	Hawthorne
14	Lola	Hawthorne
15	Donald	Freemont
13	Alex	Freemont

Table A6.3 Customer Table		
Customer ID	Customer	Location
101	Tesla	Freemont
102	Space X	Hawthorne
Product Table		
Product ID	Product	Unit price
301	Manifold	£250
302	Battery	£3,000
303	Sheet steel	£1,000
304	Fairing	£25,000

Finally, we could have the transaction table, though this might be split as well in to a series of unique tables, but we will keep it in this form for ease.

Table A6.4 Transaction Table				
Transaction	Customer ID	Product ID	Salesperson ID	Units
T001	101	301	11	5,000
T002	101	302	12	6,000
T003	102	303	13	3,000
T004	102	304	14	12
T005	101	301	15	75
T006	102	303	16	25
T007	102	302	12	40
T008	102	302	11	20

Now we have a flexible and searchable structure in normalised form that can represent all the available information about each of the salespersons, customers, products and transactions, and the relationships among them. By using the select, join (adding or union) and project (take bits from several tables or intersections) functions one can construct any of the data you want on

Routledge Focus on Business and Management: Disruption in Financial Reporting
Disruption in Financial Reporting: A post-pandemic view of the future of corporate reporting
Appendix to Chapter 6: Disruption in Reporting and the New Technology
Appendix 3.06.3 Normalisation of data

any of the attributes. Of course, customers may also have balance and an aged of debt as well, which are artefacts from the individual transactions or events.