
ALT 1 -
Databases

ALT 1 - Learning Intentions

Students learn about:	Students should be able to:
Information systems User-centred design Web design File systems and **relational** databases Design process	• 3.1 - Understand and list user needs/requirements before defining a solution • 3.2 - Create a basic **relational** database to store and retrieve a variety of forms of data types • 3.3 - Use appropriate programming languages to develop an interactive website that can display information from a database that meets a set of users' needs

Databases

Data stored about you

Data on social media

Where have you seen / used / heard of a database before?

What data was being stored in the database?

Can you identify three databases that are storing information about you?

- A database is a collection of organised data that can be accessed and searched easily.
- Databases are used in many systems and information about you is stored on numerous databases already!
- Databases are useful because:
    - Store large amounts of data in an organised manner
    - Many users can access a database simultaneously
    - Can use a single database across numerous systems.

Databases

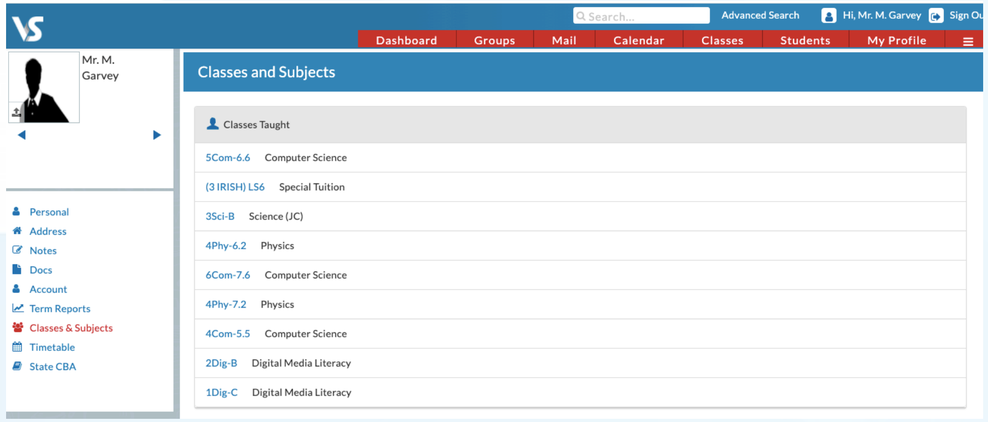
Ash	Nia	becky@bbc.com	02398 374927
02298 837492	James	Irfan	Becky
james@bbc.com	04972 048204	Mo	irfan@bbc.com
04972 048204	nia@bbc.com	Email	Parker

ID	Name	HomePhone	Mobile	Email
1	Irfan	02298 837492	0888 87492	irfan@bbc.com
2	James	02398 374927	098284 278213	james@bbc.com
3	Becky	03472 827492	098252 472911	becky@bbc.com
4	Nia	04972 048204	046482 472912	nia@bbc.com

- Databases need to be reliable, consistent and have structure.
- The structure must suit the data you collect.
- It is difficult to alter a database structure once it has been setup.

Information systems

- An information system is software that can organise and analyse data.
- Consider VSWare as a system in the school.
  - Teachers and students timetables
  - Teachers class groups
  - Student academic records
  - Facility for teachers to enter notes from parents
  - Bookings for different classrooms.



- Behind the front end, there are a number of different databases storing the information.
- Information systems allow companies to analyse and extract information from ‘big data’ easily.

Basic Database - Flat File

Customer Type DB

ID	Year_Birth	Education	Marital_Status	Income	Kidhome	Teenhome	Dt_Customer	Recency	MntWines	MntFruits	MntMeatProducts
5524	1957	Graduation	Single	58138	0	0	04-09-2012	58	635	88	546
2174	1954	Graduation	Single	46344	1	1	08-03-2014	38	11	1	6
4141	1965	Graduation	Together	71613	0	0	21-08-2013	26	426	49	127
6182	1984	Graduation	Together	26646	1	0	10-02-2014	26	11	4	20
5324	1981	PhD	Married	58293	1	0	19-01-2014	94	173	43	118
7446	1967	Master	Together	62513	0	1	09-09-2013	16	520	42	98
965	1971	Graduation	Divorced	55635	0	1	13-11-2012	34	235	65	164
6177	1985	PhD	Married	33454	1	0	08-05-2013	32	76	10	56

- A flat file database is the simplest type of database, it consists of a single table.
- It can be created in database software or in a simple spreadsheet.
- A table is a collection of related data made up of rows and columns.
- Each record should have a primary key, this is a unique identifier for each record.

Database vocabulary

Customer Type DB

Attributes

Record

ID	Year_Birth	Education	Marital_Status	Income	Kidhome	Teenhome	Dt_Customer	Recency	MntWines	MntFruits	MntMeatProducts
5524	1957	Graduation	Single	58138	0	0	04-09-2012	58	635	88	546
2174	1954	Graduation	Single	46344	1	1	08-03-2014	38	11	1	6
4141	1965	Graduation	Together	71613	0	0	21-08-2013	26	426	49	127
6182	1984	Graduation	Together	26646	1	0	10-02-2014	26	11	4	20
5324	1981	PhD	Married	58293	1	0	19-01-2014	94	173	43	118
7446	1967	Master	Together	62513	0	1	09-09-2013	16	520	42	98
965	1971	Graduation	Divorced	55635	0	1	13-11-2012	34	235	65	164
6177	1985	PhD	Married	33454	1	0	08-05-2013	32	76	10	56

- An entity is a category of object, person, event or thing about which data is recorded.
- Examples of entities are employees, movies, ingredients, shoppers etc.
- Entities are represented in databases as records, so each customer in the table shown is an entity.
- A record is used to refer to a single row in a database.
- A column name and the data it contains e.g Education in the database shown, is called an attribute.

Task 1

- A flat file database consists of a single file and might be suitable to hold the names and address of members of a club but most databases are concerned with more than one entity and the relationship between those entities.

A dentists surgery employs several dentists, and an appointments system is required to allow patients to make appointments with a particular dentist.

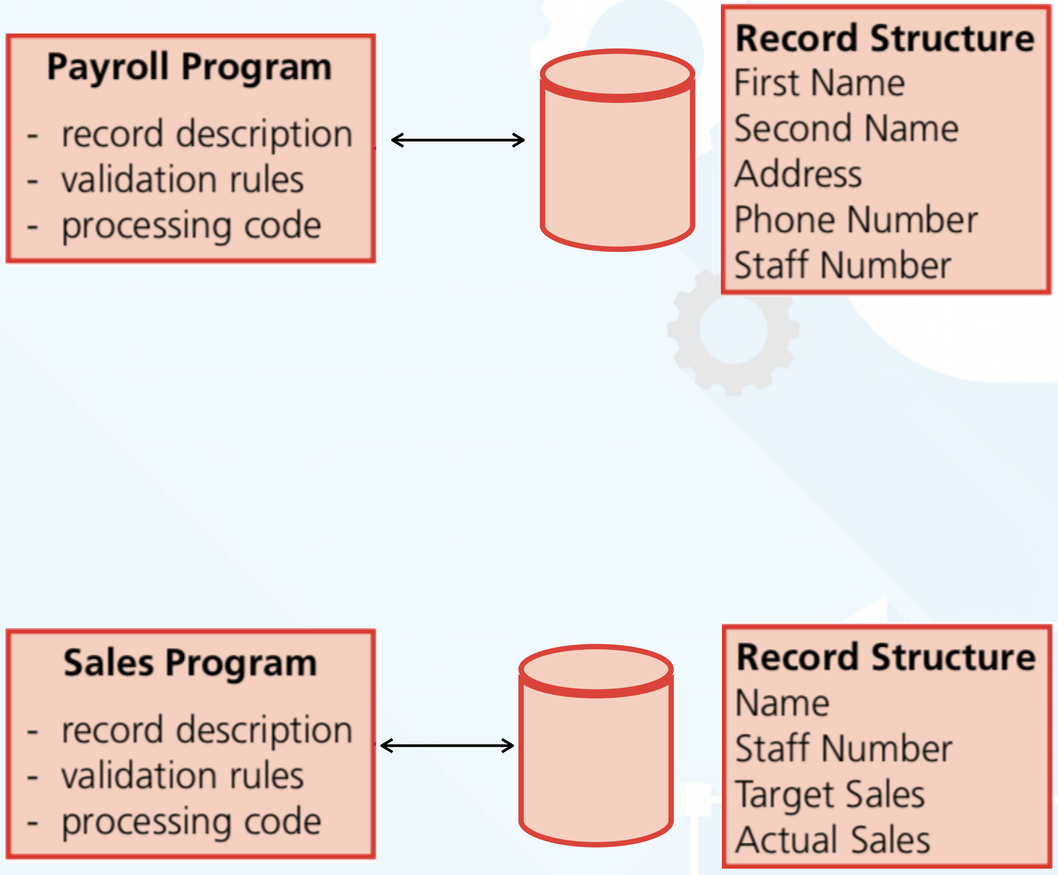
Entities in this system include Dentist, Patient and Appointment.

Attributes of a Dentist entity may include: Title, First-name, Surname, Qualification

Attributes of a Patient entity may include: Title, First-name, Surname, Address

- Can you suggest any more attributes for Patient or Dentist?
- What attributes might the Appointment entity have?

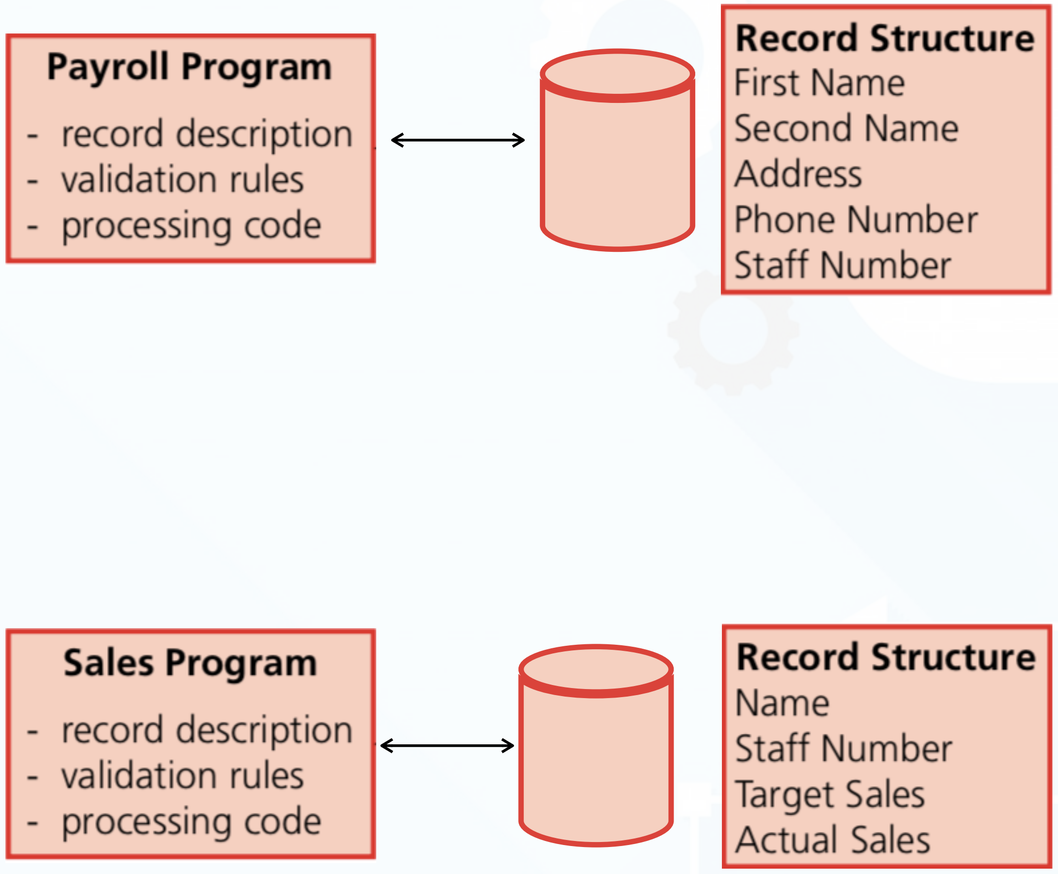
Problems with file based databases



A business keeps separate payroll file databases and sales file databases. Each databases is used by a different application.

What potential problems can you see with this system when it comes to sharing information between departments about an employee?

Problems with file based databases



- A file based approach is limited for a number of reasons.
- Storage space is wasted as some data items are duplicated.
- Data can be altered in one application or database and not the other, so it can become inconsistent.
- If you change data in one database you have to change it in every database.

Task 2

- Many file types can function as flat file databases e.g Excel, Numbers, CSV. We have used CSV (Comma Separated Values) files before for ALT 2 for data analysis so we will use them again to create a basic file database using Python.

Go to p247 in your book

Type up the code at the top of the page and make sure your program creates the new file ‘patients.csv’.

Relational Databases

- A relational database is a type of database that stores and provides access to data points across different tables that are related to one another.
- Every record has a unique identifying attribute, this is called its ‘Primary Key’.
- The primary key is going to provide the relation between the different tables in a database.

client_id	client_name	address	industry
101	Big Business Federation	123 Falschungstraße, 10999 Berlin	NGO
102	eCommerce GmbH	27 Ersatz Allee, 10317 Berlin	Retail
103	AutoMaker AG	20 Künstlichstraße, 10023 Berlin	Auto
104	Banko Bank	12 Betrugstraße, 12345 Berlin	Banking
105	WeMoveIt GmbH	138 Arglistweg, 10065 Berlin	Logistics

Client table

Relational Databases

client_id	client_name	address	industry
101	Big Business Federation	123 Falschungstraße, 10999 Berlin	NGO
102	eCommerce GmbH	27 Ersatz Allee, 10317 Berlin	Retail
103	AutoMaker AG	20 Künstlichstraße, 10023 Berlin	Auto
104	Banko Bank	12 Betrugstraße, 12345 Berlin	Banking
105	WeMoveIt GmbH	138 Arglistweg, 10065 Berlin	Logistics

Client table

participant_id	first_name	last_name	phone_no	client
101	Marina	Berg	491635558182	101
102	Andrea	Duerr	49159555740	101
103	Philipp	Probst	49155555692	102
104	René	Brandt	4916355546	102
105	Susanne	Shuster	49155555779	102
106	Christian	Schreiner	49162555375	101

Participant table

CREATE A BASIC RELATIONAL DATABASE TO STORE AND RETRIEVE A VARIETY OF FORMS OF DATA TYPES

Relational Databases

- The Client attribute in the participant is called a Foreign Key.
- A Foreign Key is a column which is a Primary Key in another table.
- In this context the Client ID is a primary key on the Client table, so when it appears on the participant table it is a foreign key.
- The foreign key relates the tables (hence relational database)

participant_id	first_name	last_name	phone_no	client
101	Marina	Berg	491635558182	101
102	Andrea	Duerr	49159555740	101
103	Philipp	Probst	49155555692	102
104	René	Brandt	4916355546	102
105	Susanne	Shuster	49155555779	102
106	Christian	Schreiner	49162555375	101

Participant table

Relational Databases

TASK: Design a new database of your choosing from the list below, with at least 3 tables. Remember to include primary keys and foreign keys as relevant!

- Employee database for a shop
- Player database for a football team
- Talent database for a management company (musicians and other performers)

Past Paper Question - Sample Paper

Question 9

(a) Define what is meant by the term **relational database**.

(b) Explain the purpose of a **primary key** in a relational database.

Past Paper Question - 2021 HL

Question 12

The following selection of data is taken from a table in a database used to store information about dogs.

dog_name	breed	dob	microchip
rover	labrador	22/11/2011	Y
fido	poodle	02/02/2020	Y
fido	jack russell	15/06/2015	N
champ	greyhound	01/01/2010	Y
spot	dalmation	24/08/2007	N
buddy	rottweiler	21/10/2012	Y

(a) State why each of the following fields would **not** be good candidates for a primary key in the table shown above.

dog_name:
breed:

Past Paper Question - 2021 HL

(b) One owner can own many dogs. Explain how a foreign key could be introduced to improve the design of this database.

owner_id	owner_name	address	dog_name	breed	dob	microchip
1	Joe Murphy	1 main st.	rover	labrador	22/11/2011	Y
1	Joe Murphy	1 main st.	fido	poodle	02/02/2020	Y
2	Ada Traore	9 park ave.	fido	jack russell	15/06/2015	N
1	Joe Murphy	1 main st.	champ	greyhound	01/01/2010	Y
2	Ada Traore	9 park ave.	spots	dalmation	24/08/2007	N
3	James Tidy	7 bond st.	buddy	rottweiler	21/10/2012	Y

Relational Database Scenario

The International Language School is a language training school.

It offers language classes for corporate clients, which can be conducted at the School or at the offices of the client as they prefer. The School employs teachers, each of whom may teach multiple courses.

The school has clients, each of whom may offer multiple courses via the school. Clients offer courses to their employees, who have the option to participate.

Each course is offered by one client. Each course has one teacher at any given time.

Participants in the courses are employees of the client companies, i.e. they work for the client companies. Each participant can be employed by one company at a time. Participants may be enrolled in more than one course.