Data Security & Ethics

O Level Computer Science

Compiled By: Engr. Fahad Khan

Data Security

Data security is about keeping data safe. Many individuals, small businesses and major companies rely heavily on their computer systems.

If the data on these computer systems is damaged, lost, or stolen, it can lead to disaster.

Key threats to data security

Data may get:

- lost or damaged during a system crash - especially one affecting the hard disk
- corrupted as a result of faulty disks, disk drives, or power failures
- lost by accidentally deleting or overwriting files
- lost or become corrupted by computer viruses
- hacked into by unauthorized users and deleted or altered
- destroyed by natural disasters, acts of terrorism, or war
- deleted or altered by employees wishing to make money or take revenge on their employer

Keeping data secure

Steps/measures that can be taken to keep data secure include:

- making regular backups of files (backup copies should be stored in fireproof safes or in another building)
- protecting yourself against viruses by running anti-virus software
- using a system of passwords so that access to data is restricted
- safe storage of important files stored on removable disks, e.g. locked away in a fireproof and waterproof safe
- allowing only authorized staff into certain computer areas, e.g. by controlling entry to these areas by means of ID cards or magnetic swipe cards
- always logging off or turning terminals off and if possible locking them
- avoiding accidental deletion of files by write-protecting disks
- using data encryption techniques to code data so that it makes no apparent sense

Understanding the Internet Risks

1. **Virus is a program designed to copy itself and propagate, usually attaching itself to applications.** It can be spread by downloading files, exchanging CD/DVDs and USB sticks, copying files from servers, or by opening infected email attachments.

2. **Spyware is often secretly installed without users consent when a file is downloaded or a commercial pop-up is clicked. Spyware can reset your auto signature, monitor your**

keystrokes, scan, read and delete your files, access your applications and even reformat your hard drive. It constantly streams information back to the person that controls spyware.

3. **Trojan might appear harmless and even useful at first, but it leaves your PC unprotected, enabling hackers to steal sensitive information.**

4. **Adware is a malware which launches advertisements, mostly in the form of pop-ups.** These are customized to you as a user, based on your behavior on the Internet, which may be monitored by spyware.

5. **Malware is short form for "malicious software,"** malware refers to software programs designed to damage or do other unwanted actions on a computer system.

6. **A worm can be injected into a network by any types of means, like an USB stick or an email attachment.** Email worm tends to send itself to all email addresses it finds on the infected PC. The email then appears to originate from the infected user, who may be on your trusted senders' list, and catch you off guard.

7. **Spam may be defined as unwanted emails. Most users are exposed to scam, which is more than 50% of all Internet emails.** Though spam is not a direct threat, it can be used to send different kinds of malware.

8. **Phishing is the fraudulent acquiring of sensitive personal information such as passwords and credit card details.** This is accomplished by sending official-looking emails impersonating a trustworthy sender. Users of online banking and auction sites are most likely to become a target.

9. **Pharming is a technique through which one can create a fake website that looks like a real one for instance web bank page, and then collect the information users think they are giving to their real bank.**
10. **Denial-of-Service (DoS) attack, a type of attack on a network that is designed to bring the network performance down by flooding it with useless traffic (data).**

11. **Hacking is the process of gaining unauthorized access to data in a system or computer.**

Protection against Internet Risks (Threats)

1. **Antivirus software is a type of utility software used for scanning and removing viruses from computer.** While many types of antivirus (or "anti-virus") programs exist, their primary purpose is to protect computers from viruses and remove any viruses that are found.

2. **A firewall is a program or hardware device that filters the information coming through the Internet connection into your personal computer or into a company's network.**
   A firewall establishes a barrier between a trusted, secure internal network and another network (e.g., the Internet) that is assumed not to be secure and trusted. Many hardware-based firewalls also offer other functionality to the internal network they protect.

Many personal computer operating systems include software-based firewalls to protect against threats from the public Internet. Many routers that pass data between networks contain firewall components and, conversely, many hardware based firewalls can perform basic routing functions.

3. **A proxy server may act as a firewall by responding to input packets (connection requests, for example) while blocking other packets containing suspicious data.** A proxy server is a gateway from one network to another for a specific network application, in the sense that it functions as a proxy on behalf of the network user.

4. **Secure Sockets Layer (SSL) is a secure protocol developed for sending information securely over the Internet.** Many websites use SSL for secure areas of their sites, such as user account pages and online checkout. Usually, when you are asked to "log in" on a website, the resulting page is secured by SSL.
   SSL encrypts the data being transmitted so that a third party cannot "eavesdrop" on the transmission and view the data being transmitted.

5. **Collectively, Username and password can also be used to provide protection.** When someone log onto your network at school, you have to type in your User ID and Password. This identifies you to the network as an authorised user.
   Any sensible company will ensure that staff need a User ID and Password to gain access to the system. This should reduce the risk of outsiders being able to get onto the system and damage data.

6. **Encryption is a method of scrambling data in such a way that only the people who have the 'secret key' to unlock the message can read it**
   OR
   Encryption is the process of converting data to an unrecognizable or "encrypted" form.

   You can encrypt a file, folder. Encryption is also used to secure data sent over wireless networks and the Internet. Many websites and other online services encrypt data transmissions using SSL. Any website that begins with "https://," for example, uses the HTTPS protocol, which encrypts all data sent between the web server and your browser. SFTP, which is a secure version of FTP, encrypts all data transfers.
   Cleartext is readable data transmitted or stored “in the clear” (i.e. unencrypted).
   Plaintext is the input to an encryption algorithm.
   Ciphertext is the unreadable output of an encryption algorithm.

   Examples of plain text are:
   Humpty Dumpty sat on a wall.
   Humpty Dumpty had a big fall.

   Examples of cipher text are:
   lj86ik,£lj)ay%9w2+m?lsild171724

jkd2f*hkdfh7$171kjfh7d1h4d

A key is a variable value that is applied using an algorithm to plaintext to produce encrypted text, or to decrypt encrypted text. Key size or key length is the size measured in bits. More number of bits in a key will ensure more security of data.

There are two basic techniques for encrypting information:

- Symmetric encryption (also called secret key encryption)
- Asymmetric encryption (also called public key encryption.)

Symmetric Encryption is a type of encryption where the same key is used to encrypt and decrypt the message.

Asymmetric encryption is a type of encryption which uses one key to encrypt a message and another to decrypt the message.

Data Verification

Verification is performed to ensure that the data entered exactly matches the original source.

There are two main methods of verification:

- **Double entry** - entering the data twice and comparing the two copies. This effectively doubles the workload, and as most people are paid by the hour, it costs more too.
- **Proofreading data** - this method involves someone checking the data entered against the original document. This is also time consuming and costly.

Data Validation

Validation and verification are two ways to check that the data entered into a computer is correct. Data entered incorrectly is of little use.

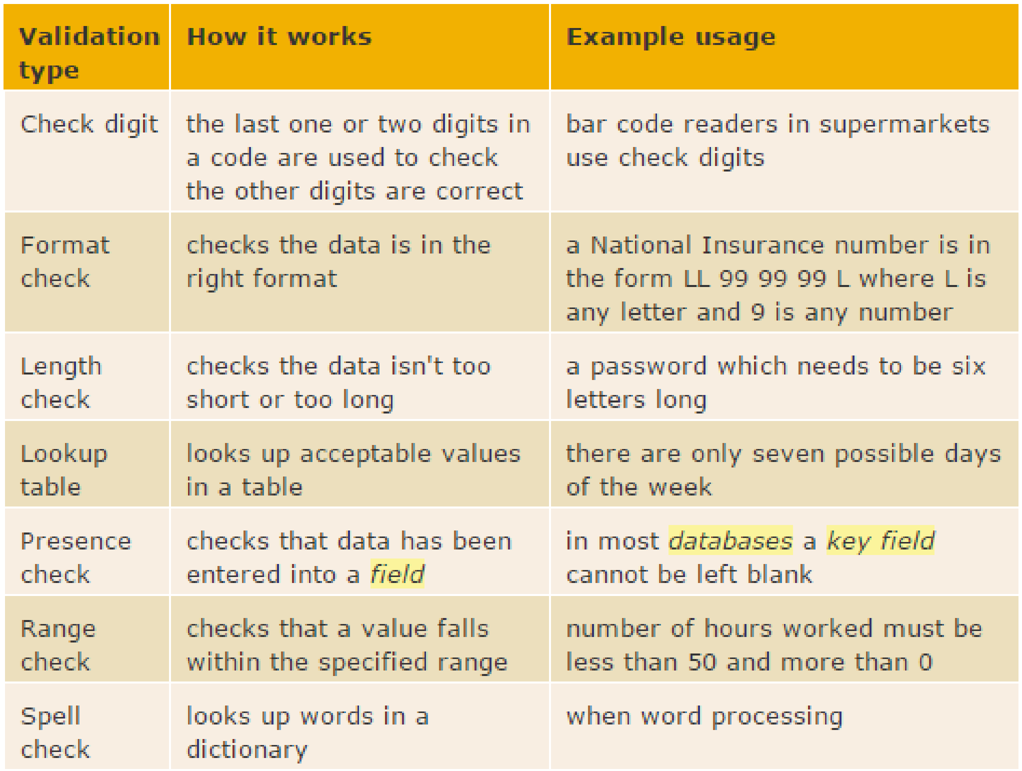
Validation is an automatic computer check to ensure that the data entered is sensible and reasonable. It does not check the accuracy of data.

For example, a secondary school student is likely to be aged between 11 and 16. The computer can be programmed only to accept numbers between 11 and 16. This is a range check.

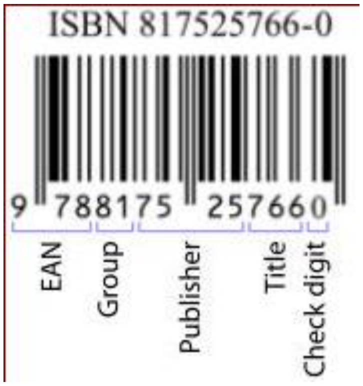
However, this does not guarantee that the number typed in is correct. For example, a student's age might be 14, but if 11 is entered it will be valid but incorrect.

Types of validation

There are a number of validation types that can be used to check the data that is being entered.



Validation type	How it works	Example usage
Check digit	the last one or two digits in a code are used to check the other digits are correct	bar code readers in supermarkets use check digits
Format check	checks the data is in the right format	a National Insurance number is in the form LL 99 99 99 L where L is any letter and 9 is any number
Length check	checks the data isn't too short or too long	a password which needs to be six letters long
Lookup table	looks up acceptable values in a table	there are only seven possible days of the week
Presence check	checks that data has been entered into a *field*	in most *databases* a *key field* cannot be left blank
Range check	checks that a value falls within the specified range	number of hours worked must be less than 50 and more than 0
Spell check	looks up words in a dictionary	when word processing



Verify that ISBN mentioned above with check digit is correct or not by using Modulo 11 system.

Types of Computer Misuses

Misuse of computers and communications systems comes in several forms:

- **Data misuse and unauthorised transfer or copying**
  Copying and illegal transfer of data is very quick and easy using online computers and large storage devices such as hard disks, memory sticks and DVDs. Personal data, company research and written work, such as novels and textbooks, cannot be copied without the copyright holder's permission.
- **Copying and distributing copyrighted software, music and film**

This includes copying music and movies with computer equipment and distributing it on the Internet without the copyright holder's permission. This is a widespread misuse of both computers and the Internet that breaks copyright regulations.

Attacks at Online Systems

- A **denial of service (DoS) attack** is an attempt to make a machine or network resource unavailable to its intended users.
- In a **distributed denial-of-service (DDoS)**, large numbers of compromised systems (sometimes called a botnet) attack a single target.
- **Phishing is an e-mail fraud method** in which the perpetrator sends out legitimate-looking email in an attempt to gather personal and financial information from recipients.
- **Pharming is a scamming practice in which users are misdirected to fraudulent Web sites** without their knowledge or consent.

Computer Ethics

Computer ethics is set of moral principles that regulate the use of computers. Some common issues of computer ethics include such as copyrights and plagiarism issues.

- **Copyright is the protection of someone’s already published work and prevention of this work from being used without prior permission.**
- **Plagiarism is the practice of taking someone else's work or ideas and passing them off as one's own.**

Types of Software Licensing

- **Free software** is a software that gives its users the freedom to share, study and modify it. It has no copyright or other restrictions for distributing, modifying and using the software in any way.
- **Freeware** is any software that is distributed for use at a price of zero. However, freeware may not be "free software".
- **Shareware** is a software that is distributed free on a trial basis with the understanding that the user may need or want to pay for it later.

For example, a software developer may choose to make her software available for download and use on her website. This software may be freeware if downloaded for personal use but commercial use may require a fee. In either case, if it is prohibited to freely distribute (for any purpose) or modify the software, then this freeware is not free software.

Data Protection Act (DPA)

The Data Protection Act (DPA) is a law designed to protect personal data stored on computers or in an organised paper filing system.

The 1998 Data Protection Act was passed by Parliament to control the way information is handled and to give legal rights to people who have information stored about them.

The basic way it works is by:

- setting up rules that people have to follow
- having an Information Commissioner to enforce the rules

The Roles/Components Involved

- The Information Commissioner is the person (and his/her office) who has powers to enforce the Act.
- A data controller is a person or company that collects and keeps data about people.
- A data subject is someone who has data about them stored somewhere, outside of their direct control.

The Eight Principles of Data Protection:

For the personal data that controllers store and process:

1. It must be collected and used fairly and inside the law.
2. It must only be held and used for the reasons given to the Information Commissioner.
3. It can only be used for those registered purposes and only be disclosed to those people mentioned in the register entry. You cannot give it away or sell it unless you said you would to begin with.
4. The information held must be adequate, relevant and not excessive when compared with the purpose stated in the register. So you must have enough detail but not too much for the job that you are doing with the data.
5. It must be accurate and be kept up to date. There is a duty to keep it up to date, for example to change an address when people move.
6. It must not be kept longer than is necessary for the registered purpose. It is alright to keep information for certain lengths of time but not indefinitely. This rule means that it would be wrong to keep information about past customers longer than a few years at most.
7. The information must be kept safe and secure. This includes keeping the information backed up and away from any unauthorised access. It would be wrong to leave personal data open to be viewed by just anyone.
8. The files may not be transferred outside of the European Economic Area (that's the EU plus some small European countries) unless the country that the data is being sent to has a suitable data protection law. This part of the DPA has led to some countries passing similar laws to allow computer data centres to be located in their area.

Types of Personal Data

There are two types of personal data

Personal data is about living people and could be:

- their name
- address
- medical details or banking details

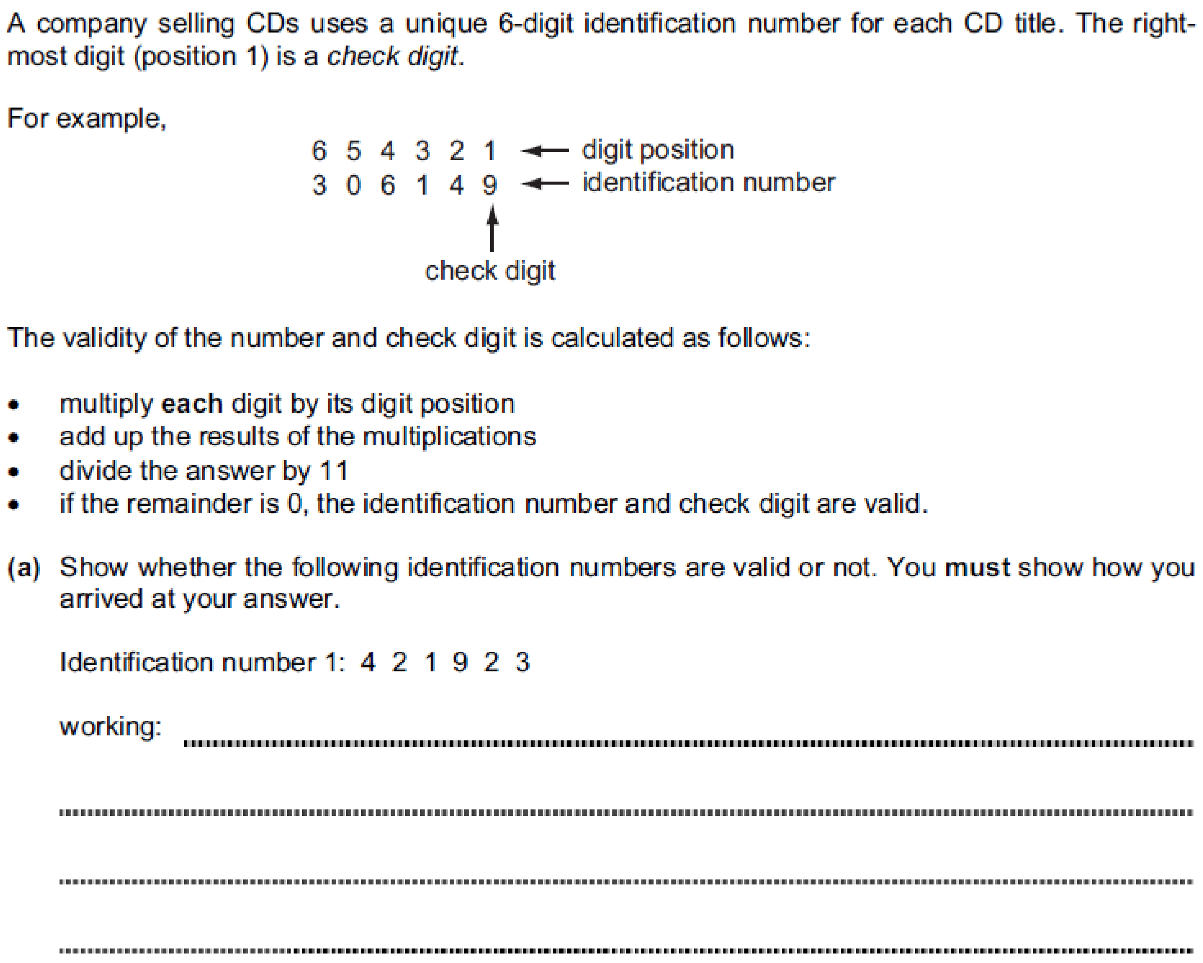
Sensitive personal data is also about living people, but it includes one or more details of a data subject's:

- racial or ethnic origin
- political opinions

- religion
- membership of a trade union
- health
- sex life
- criminal activity

Practice Questions

Question 1: (Specimen Paper 2015, Q1)



A company selling CDs uses a unique 6-digit identification number for each CD title. The right-most digit (position 1) is a *check digit*.

For example,

6 5 4 3 2 1 ← digit position
3 0 6 1 4 9 ← identification number
↑
check digit

The validity of the number and check digit is calculated as follows:

- multiply **each** digit by its digit position
- add up the results of the multiplications
- divide the answer by 11
- if the remainder is 0, the identification number and check digit are valid.

(a) Show whether the following identification numbers are valid or not. You **must** show how you arrived at your answer.

Identification number 1: 4 2 1 9 2 3

working: ....

....

....

....

valid or not valid? ....

Identification number 2: 8 2 0 1 5 6

working: ....

....

....

....

valid or not valid? ....

(b) Find the check digit for this identification number.

5 0 2 4 1 __

working: ....

....

....

....

check digit: ....

(c) Describe, with examples, **two** different types of data entry errors that a check digit would detect.

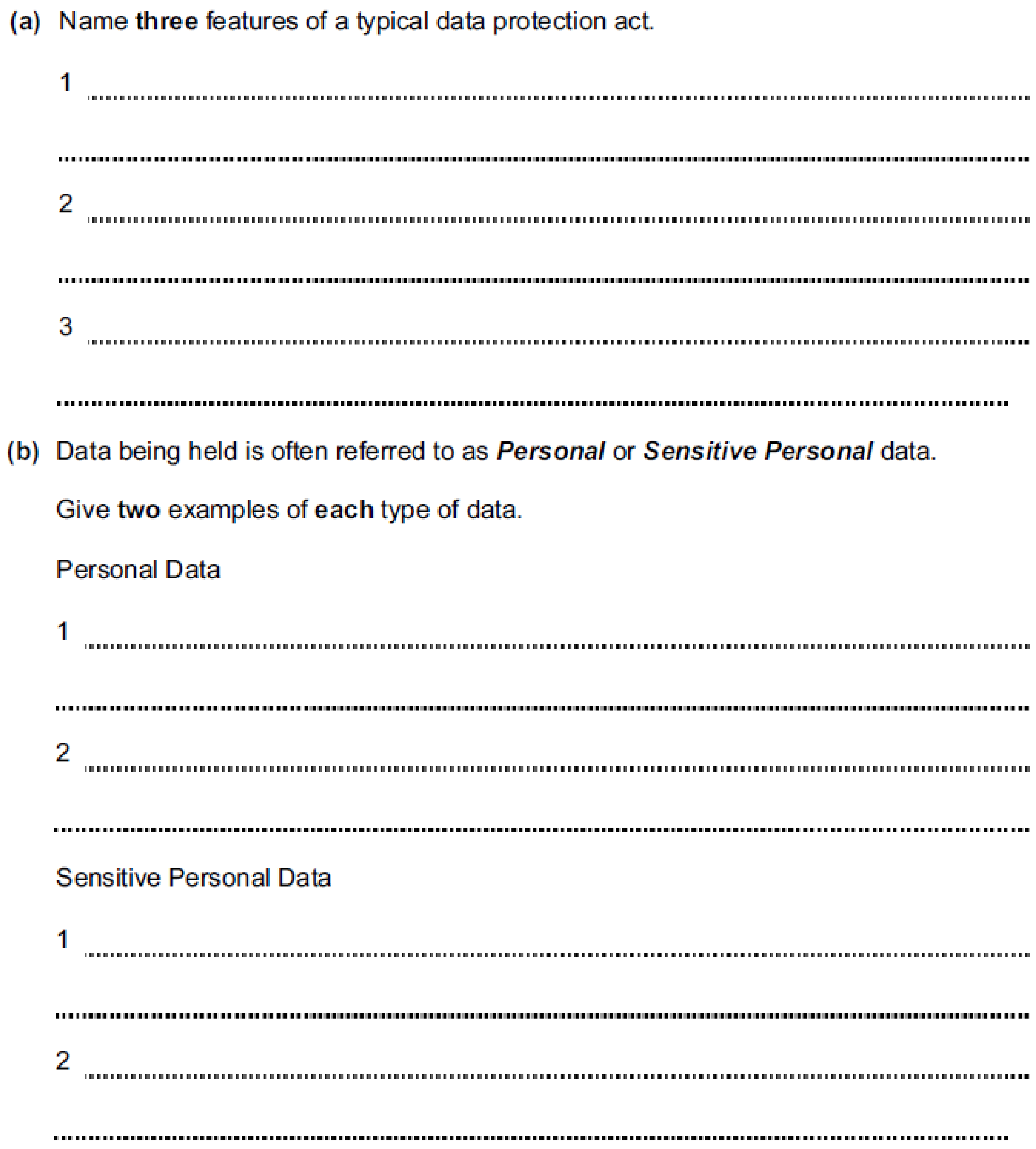
1 ....

....

2 ....

....

Question 2: (Oct/Nov 2013, P13, Q1)



(a) Name **three** features of a typical data protection act.

1 ....

....

2 ....

....

3 ....

....

(b) Data being held is often referred to as ***Personal*** or ***Sensitive Personal*** data.

Give **two** examples of **each** type of data.

Personal Data

1 ....

....

2 ....

....

Sensitive Personal Data

1 ....

....

2 ....

....

Question 3: (May/June 2014, P11, Q3)

A hospital holds records of its patients in a database. Four of the fields are:

- date of visit (dd/mm/yyyy)
- patient's height (m)
- 8-digit patient ID
- contact telephone number

The presence check is one possible type of validation check on the data. For each field, give another validation check that can be performed. Give an example of data which would **fail** your named validation check.

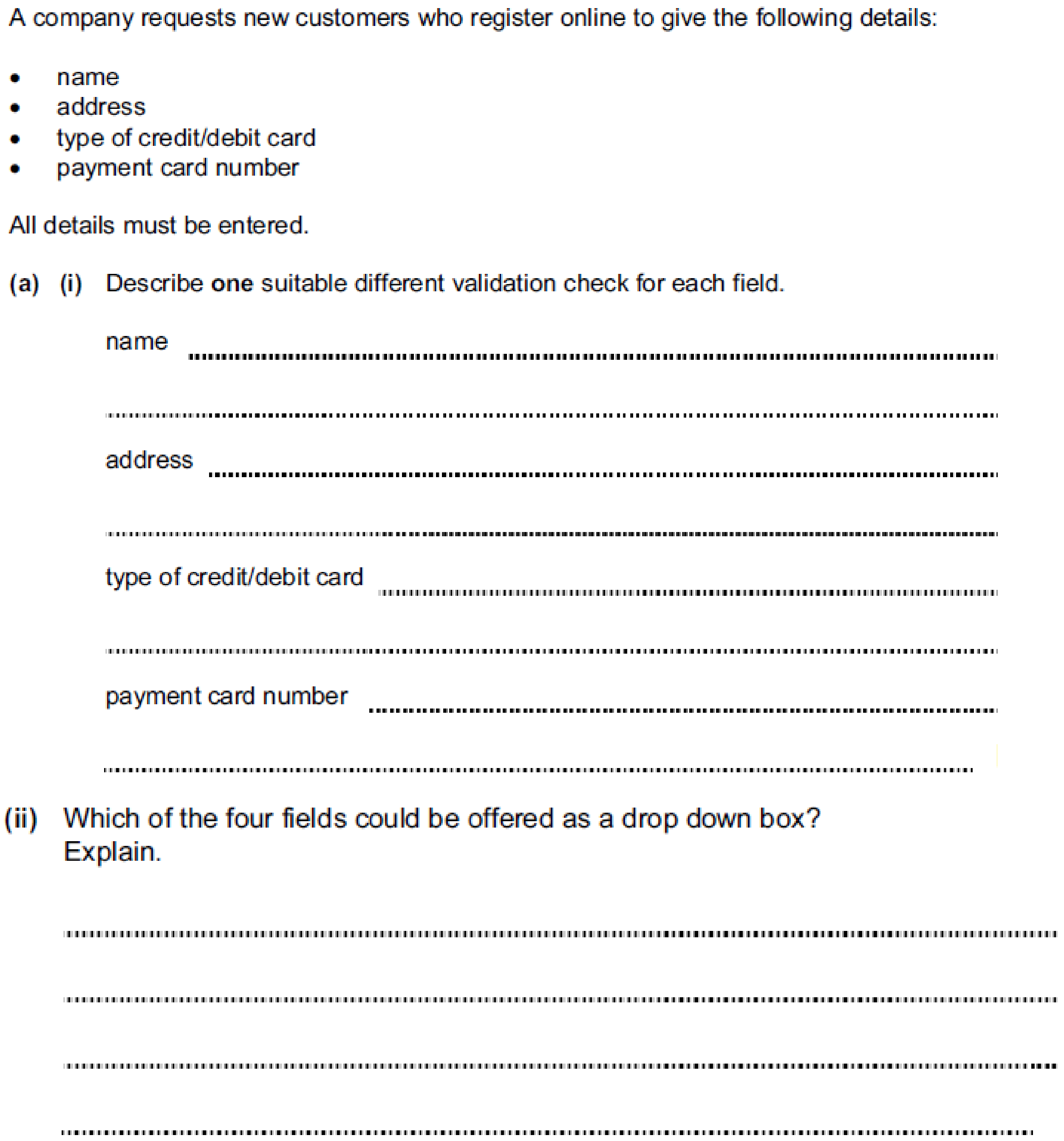
A **different** validation check needs to be given for each field.

field name	name of validation check	example of data which would fail the validation check
date of visit		
patient's height		
patient ID		
contact telephone number		

Question 6: Fill in the blanks with appropriate words.

1. _______________ is a software that gives its users the freedom to share, study and modify it.
2. _______________any software that is distributed for use at a price of zero but may not be "free software".
3. _______________ is a software that is distributed free on a trial basis with the understanding that the user may need or want to pay for it later.
4. _______________ is the practice of taking someone else's work or ideas and passing them off as one's own.
5. A ______________ attack is an attempt to make a machine or network resource unavailable to its intended users.
6. ________________is the protection of someone's already published work and prevention of this work from being used without prior permission.

Question 7: (May/June 2013, P12, Q13)



A company requests new customers who register online to give the following details:

- name
- address
- type of credit/debit card
- payment card number

All details must be entered.

(a) **(i)** Describe **one** suitable different validation check for each field.

name ....

....

address ....

....

type of credit/debit card ....

....

payment card number ....

....

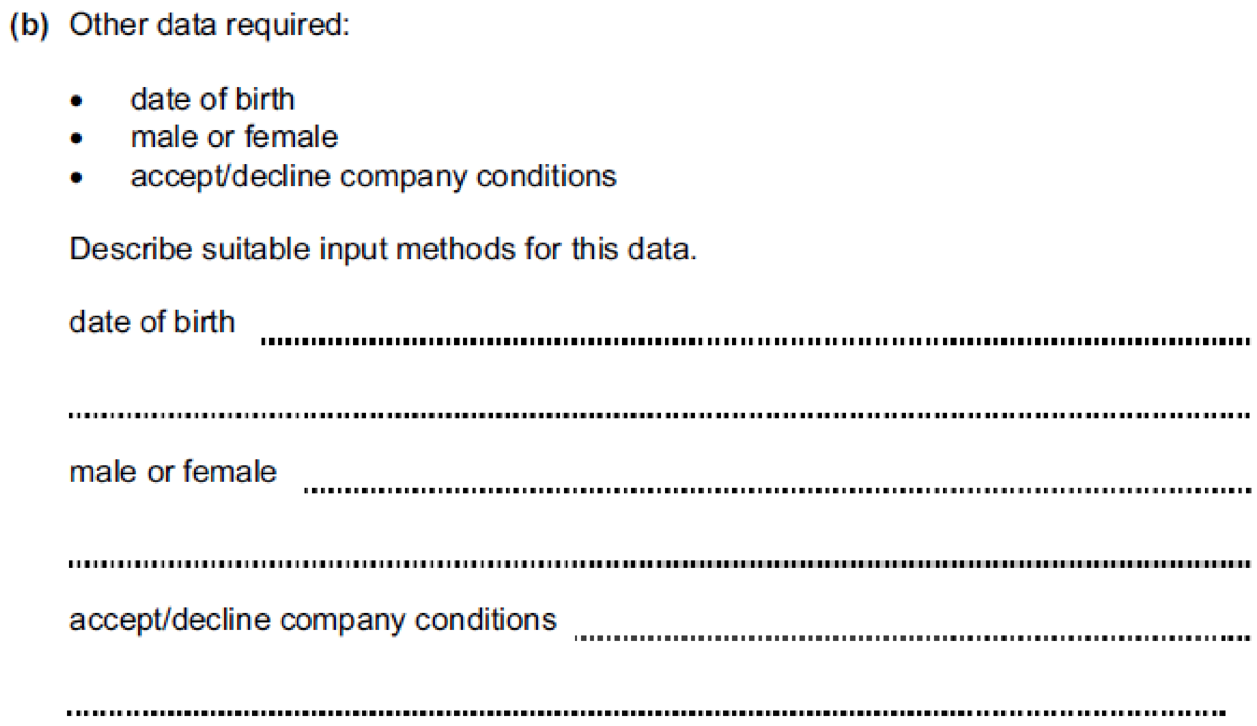
(ii) Which of the four fields could be offered as a drop down box?
Explain.

....

....

....

....



(b) Other data required:

- date of birth
- male or female
- accept/decline company conditions

Describe suitable input methods for this data.

date of birth ....

....

male or female ....

....

accept/decline company conditions ....

....

Question 8: List down three key threats to data and steps that may be taken against these threats.

Threat 1:

Step to be taken:

Threat 2:

Step to be taken:

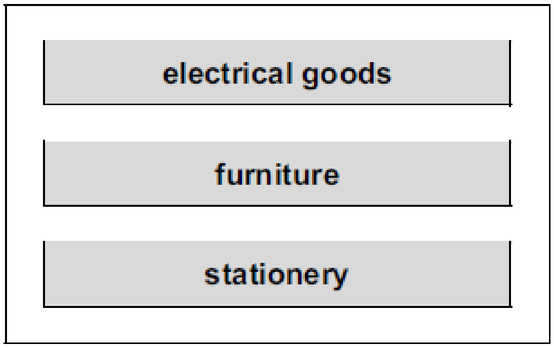
Threat 3:

Step to be taken:

Question 9: (Oct/Nov 2013, P12, Q12)

A shop uses an information screen linked to a computer to allow customers to order goods directly.

The first screen shows three options:



(a) What is the best input device to allow customers to choose one of the three options?

.... [1]

(b) The customer is then sent to another screen where they have to input:

- the goods reference number which is 8 digits long
- today's date which must be in the form dd/mm/yyyy
- the customer's telephone number

For each input give **one** validation check that should be performed.
A **different** type of check must be given in each case.

goods reference number ....

today's date ....

telephone number .... [3]