N15/3/ITGSX/HP3/ENG/TZ0/XX/M

Diploma Programme
Programme du diplôme
Programa del Diploma

# Markscheme

## November 2015

## Information technology in a global society

## Higher level

## Paper 3

10 pages

Examiners should be aware that in some cases, candidates may take a different approach, which if appropriate should be rewarded. If in doubt, check with your Team Leader.

If candidates answer more than the prescribed number of questions:

- In the case of an "identify" question read all answers and mark positively up to the maximum marks. Disregard incorrect answers.

- In the case of a "describe" question, which asks for a certain number of facts *eg* "describe two kinds", mark the **first two** correct answers. This could include two descriptions, one description and one identification, or two identifications.

- In the case of an "explain" question, which asks for a specified number of explanations *eg* "explain two reasons", mark the **first two** correct answers. This could include two full explanations, one explanation, one partial explanation *etc*.

**1.** (a) Define “data mining”. **[2]**

*Answers may include:*
- discovering patterns in sets of data
- obtaining useful data from sets
- representing data found in data sets
- collecting data from different sources and different databases
- may include any analysis of data using business intelligence software
- obtaining information from data which is used for decision-making
- dealing with large quantities of data.

*Award **[1]** for each point identified up to a maximum of **[2]**.*

(b) Identify **two** ways customer details on a database can be made anonymous. **[2]**

*Answers may include:*
- remove names, addresses *etc* from the main customer data / exported data
- using identifiers and not using personal direct identifiers
- remove redundant or unnecessary data fields (such as post code or known name *etc*)
- avoid data being linked in combination of fields which together might identify a person
- use of pixilation on images (such as Google maps for faces and license plates)
- limit the personal fields used to only those which are necessary
- randomize IDs and do not make them sequential or based on a formula linked to name or date of birth *etc*
- remove metatags and other identifiable information from images.

*Award **[1]** for each point identified up to a maximum of **[2]**.*

**2.** (a) Explain **one** way that data visualization can be beneficial for store managers. **[2]**

*Answers may include:*
- different learning styles – some managers may be visual
- can represent a large amount of data in an image
- easily understandable by humans rather than by machines
- analysis can be made from an image whereas it is difficult to see patterns within raw data
- large data sets can be compared side by side (*eg* this year's and last year's sales)
- store managers can manipulate the data to see how the image can change
- can be used to support marketing
- management boards can use visualizations in promotional and shareholder information
- can show productivity of store quickly
- data can be broken down into different focus areas and graphically represented
- management can examine an aspect of the data holistically and quickly make decisions.

*Award **[1]** for identifying a method that data visualization may be beneficial to store managers and **[1]** for valid explanation of that point up to a maximum of **[2]**.*

(b) Distinguish between clustering analysis and hypothesis testing of data. **[4]**

*Answers may include:*
**Hypothesis testing**
- answering questions by looking into the data
- finding evidence in the data to prove/disprove ideas
- creating what-if scenarios
- allowing data to be manipulated to produce theoretical or artificial situations for testing
- using non-real or imaginary data/scenarios to test output results
- business intelligence software may take data and manipulate to incorporate unfamiliar scenarios with real data
- use deductive reasoning.

**Clustering**
- grouping together sets of data using similar values or trends
- discovering different types/behaviours of real-world entity (*eg* customers) in the data
- discovering real-life similarities/dependencies through exploring data
- using mathematical methods such as k-means/artificial neural network/nearest neighbour to automatically group data
- grouping data based on distances from the mean and standard deviation of values
- linking raw data by similar or desired attributes
- representing data graphically or numerically using appropriate methods
- business intelligence software may be used to represent data
- use inductive reasoning.

*Award **[1]** each for definition/identification of a characteristic of hypothesis testing and clustering (Maximum of **[2]**). Award additional **[2]** for comparison of both methods.*

**3.** The *Asociación de Supermercados Independientes* (ASI) will collect additional data such as the brand of the product or personal information about the customer every time they make a purchase.

This additional data is not on the receipt and customers will not know it is being collected. Discuss whether this is acceptable. **[8]**

*Answers may include:*

- what is the nature of the invisible data that is being collected?
- how is it collected?
  - in person at the register?
  - online via survey at the checkout page
  - online data collected may include technical metadata such as browser information or IP address
- is the data collected of a personal nature such as name, address?
- what is the reason for ASI to collect this information?
- where will this information be used?
- does ASI have a privacy policy on the acceptable use of data?
- who has access to this information?
- does appropriate data security exist so that it is kept safe?
- how long is it kept for?
- how is it destroyed (wiped/disposed of safely)?
- will it be shared with 3rd parties?
- is the purpose and data registered with the legal authorities in the country of storage/collection?
- is there a procedure for customers to see their data and change it if incorrect?
- can customers opt out if they discover the data is being stored?
- can information about purchased products be used to derive/deduce sensitive information about customers (*eg* pregnancy testing kit, contraceptives, medicine, hobby magazines or books, DVDs)
- does ASI have a privacy policy on the acceptable use of data?
- if collected through the use of a loyalty card, there should be a clause in the user agreement that additional data would be collected?
- is the data anonymized for data mining purposes?
- collecting invisible data can bring about distrust with clients
- to improve a business, collecting as much data about their client base could improve the organization and help with strategic planning. This could lead to pattern analysis or clustering
- the data may be used for targeted marketing.

**SL and HL paper 1 part (c) and HL paper 3 question 3 markband**

| ***Marks*** | ***Level descriptor*** |
|---|---|
| ***No marks*** | • *A response with no knowledge or understanding of the relevant ITGS issues and concepts.*<br>• *A response that includes no appropriate ITGS terminology.* |
| ***Basic***<br>***1–2 marks*** | • *A response with minimal knowledge and understanding of the relevant ITGS issues and concepts.*<br>• *A response that includes minimal use of appropriate ITGS terminology.*<br>• *A response that has no evidence of judgments and/or conclusions.*<br>• *No reference is made to the scenario in the stimulus material in the response.*<br>• *The response may be no more than a list.* |
| ***Adequate***<br>***3–4 marks*** | • *A descriptive response with limited knowledge and/or understanding of the relevant ITGS issues and/or concepts.*<br>• *A response that includes limited use of appropriate ITGS terminology.*<br>• *A response that has evidence of conclusions and/or judgments that are no more than unsubstantiated statements. The analysis underpinning them may also be partial or unbalanced.*<br>• *Implicit references are made to the scenario in the stimulus material in the response.* |
| ***Competent***<br>***5–6 marks*** | • *A response with knowledge and understanding of the relevant ITGS issues and/or concepts.*<br>• *A response that uses ITGS terminology appropriately in places.*<br>• *A response that includes conclusions and/or judgments that have limited support and are underpinned by a balanced analysis.*<br>• *Explicit references to the scenario in the stimulus material are made at places in the response.* |
| ***Proficient***<br>***7–8 marks*** | • *A response with a detailed knowledge and understanding of the relevant ITGS issues and/or concepts.*<br>• *A response that uses ITGS terminology appropriately throughout.*<br>• *A response that includes conclusions and/or judgments that are well supported and underpinned by a balanced analysis.*<br>• *Explicit references are made appropriately to the scenario in the stimulus material throughout the response.* |

**4.** The ASI is concerned that the benefits to customers of being involved in the loyalty card scheme may be outweighed by their desire for privacy and anonymity.

To what extent do you agree with this statement? **[12]**

*Answers may include:*

**Personal data security**

- a lot of personal data may be collected while registering for a loyalty card
- a clear plan of what specific pieces of personal data is collected should be made by ASI
- does ASI have a policy in place for access to this person data?
- ASI should have protocols in place for the security of data
- can clients opt out of supplying personal data in their profile?
- different levels of access should be in place for protecting the data
- train in store staff to support customers who ask questions about data security
- is data shared with 3rd parties, exported, transferred, or accessed?
- is there an appropriate level of security employed by the servers on which this data is stored, and all backup media?
- are customers allowed to see their own data and make corrections if it is wrong?
- policies on the third party access of data needs to be included in ASI company documentation
- when a customer leaves the loyalty programme, the data should be erased
- will customers become aware of the issues linked to data being shared, will this inform their decision making about whether to join the loyalty card scheme?

**Loyalty card**

- targeted marketing is a key component of loyalty card systems. Therefore a certain amount of personal data is required to establish a profile
- benefits of being a member means good deals advertised and pushed from the marketing system
- if recommended products are suggested based on this data, can the system explain why these recommendations were made and allow the customer to change assumptions (like Amazon does)
- does the loyalty card scheme allow the channel and frequency of contact with the customer to be chosen (*ie* can customers decide not to be part of a mail shot/spammed, cold called by phone *etc*).
- when making purchases online, the client can be identified based on their profile, therefore not needing to repeat personal information such as address and payment details
- not store financial information in the profile to avoid leaked/hacked data
- promotional material for clients stating what is collected, how information is used and how it is protected
- train in store staff to support customers who ask questions about data security
- potential savings or bonuses to the client by being a member of the loyalty programme
- by having customers belong to the loyalty programme, there is a potential for increase in sales for ASI through customer loyalty
- can be more competitive with the other market corporations by having a larger and loyal client base
- without processes or facilities in place, if the programme grows too quickly then it may have problems
- will ASI look to take the ethical lead in identity management for its customers?
- clients should really sign an agreement giving permission for their data to be collected.

**Business analysis**

- having a loyalty programme means the collection of data about customers and about sales
- using data mining and business intelligence software, analyzing this data can mean some improvements to the business plan/model. This may not have been obvious or known before
- more informed and financially viable decisions can be made with the use of data
- a more focused marketing campaign can be generated that would allow loyal customers to be rewarded and advertised to
- ASI can promote themselves as a community driven company by returning value back to the customers. Also they could use the community outreach angle.

**HL paper 3 question 4 markband**

| *Marks* | *Level descriptor* |
|---|---|
| ***No marks*** | • *A response with no knowledge or understanding of the relevant ITGS issues and concepts.*<br>• *A response that includes no appropriate ITGS terminology.* |
| ***Basic***<br>***1–3 marks*** | • *A response with minimal knowledge and understanding of the relevant ITGS issues and concepts.*<br>• *A response that includes minimal use of appropriate ITGS terminology.*<br>• *A response that has no evidence of judgments, conclusions or future strategies.*<br>• *No reference is made to the information in the case study or independent research in the response.*<br>• *The response may be no more than a list.* |
| ***Adequate***<br>***4–6 marks*** | • *A descriptive response with limited knowledge and/or understanding of the relevant ITGS issues and/or concepts.*<br>• *A response that includes limited use of appropriate ITGS terminology.*<br>• *A response that has evidence of conclusions, judgments or future strategies that are no more than unsubstantiated statements. The analysis underpinning them may also be partial or unbalanced.*<br>• *Implicit references are made to the information in the case study or independent research in the response.* |
| ***Competent***<br>***7–9 marks*** | • *A response with knowledge and understanding of the relevant ITGS issues and/or concepts.*<br>• *A response that uses ITGS terminology appropriately in places.*<br>• *A response that includes conclusions and/or judgments that have limited support and are underpinned by a balanced analysis.*<br>• *Explicit references to the information in the case study or independent research are made at places in the response.* |
| ***Proficient***<br>***10–12 marks*** | • *A response with a detailed knowledge and understanding of the relevant ITGS issues and/or concepts.*<br>• *A response that uses ITGS terminology appropriately throughout.*<br>• *A response that includes conclusions, judgments or future strategies that are well supported and underpinned by a balanced analysis.*<br>• *Explicit references are made appropriately to the information in the case study and independent research throughout the response.* |