

Markscheme

May 2025

Computer science

Higher level

Paper 2

© International Baccalaureate Organization 2025

All rights reserved. No part of this product may be reproduced in any form or by any electronic or mechanical means, including information storage and retrieval systems, without the prior written permission from the IB. Additionally, the license tied with this product prohibits use of any selected files or extracts from this product. Use by third parties, including but not limited to publishers, private teachers, tutoring or study services, preparatory schools, vendors operating curriculum mapping services or teacher resource digital platforms and app developers, whether fee-covered or not, is prohibited and is a criminal offense.

More information on how to request written permission in the form of a license can be obtained from <https://ibo.org/become-an-ib-school/ib-publishing/licensing/applying-for-a-license/>.

© Organisation du Baccalauréat International 2025

Tous droits réservés. Aucune partie de ce produit ne peut être reproduite sous quelque forme ni par quelque moyen que ce soit, électronique ou mécanique, y compris des systèmes de stockage et de récupération d'informations, sans l'autorisation écrite préalable de l'IB. De plus, la licence associée à ce produit interdit toute utilisation de tout fichier ou extrait sélectionné dans ce produit. L'utilisation par des tiers, y compris, sans toutefois s'y limiter, des éditeurs, des professeurs particuliers, des services de tutorat ou d'aide aux études, des établissements de préparation à l'enseignement supérieur, des fournisseurs de services de planification des programmes d'études, des gestionnaires de plateformes pédagogiques en ligne, et des développeurs d'applications, moyennant paiement ou non, est interdite et constitue une infraction pénale.

Pour plus d'informations sur la procédure à suivre pour obtenir une autorisation écrite sous la forme d'une licence, rendez-vous à l'adresse <https://ibo.org/become-an-ib-school/ib-publishing/licensing/applying-for-a-license/>.

© Organización del Bachillerato Internacional, 2025

Todos los derechos reservados. No se podrá reproducir ninguna parte de este producto de ninguna forma ni por ningún medio electrónico o mecánico, incluidos los sistemas de almacenamiento y recuperación de información, sin la previa autorización por escrito del IB. Además, la licencia vinculada a este producto prohíbe el uso de todo archivo o fragmento seleccionado de este producto. El uso por parte de terceros —lo que incluye, a título enunciativo, editoriales, profesores particulares, servicios de apoyo académico o ayuda para el estudio, colegios preparatorios, desarrolladores de aplicaciones y entidades que presten servicios de planificación curricular u ofrezcan recursos para docentes mediante plataformas digitales—, ya sea incluido en tasas o no, está prohibido y constituye un delito.

En este enlace encontrará más información sobre cómo solicitar una autorización por escrito en forma de licencia: <https://ibo.org/become-an-ib-school/ib-publishing/licensing/applying-for-a-license/>.

Subject details: Computer science HL paper 2 markscheme

Mark allocation

Candidates are required to answer **all** questions in **one** Option. Total 65 marks.

General

A markscheme often has more specific points worthy of a mark than the total allows. This is intentional. Do not award more than the maximum marks allowed for that part of a question.

When deciding upon alternative answers by candidates to those given in the markscheme, consider the following points:

- Each statement worth one point has a separate line and the end is signified by means of a semi-colon (;).
- An alternative answer or wording is indicated in the markscheme by a “/”; either wording can be accepted.
- Words in (...) in the markscheme are not necessary to gain the mark.
- If the candidate’s answer has the same meaning or can be clearly interpreted as being the same as that in the markscheme then award the mark.
- Mark positively. Give candidates credit for what they have achieved and for what they have got correct, rather than penalizing them for what they have not achieved or what they have got wrong.
- Remember that many candidates are writing in a second language; be forgiving of minor linguistic slips. In this subject effective communication is more important than grammatical accuracy.
- Occasionally, a part of a question may require a calculation whose answer is required for subsequent parts. If an error is made in the first part then it should be penalized. However, if the incorrect answer is used correctly in subsequent parts then **follow through** marks should be awarded. Indicate this with “**FT**”.

General guidance

Issue	Guidance
Answering more than the quantity of responses prescribed in the questions	<ul style="list-style-type: none"> • 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 <i>etc.</i>

Option A – Databases

- 1. (a) (i) *Award [1 max]*
StaffID;
- (ii) *Award [1 max]*
RoleID
DepartmentID

- (b) *Award [2 max]*
Award [1] for the correct relationship between role and staff tables
Award [1] for the correct relationship between department and staff

ROLE and STAFF: one to many

STAFF and DEPARTMENT: one to many //accept many to many

Note: if a diagram is seen ignore the symbols and award marks for the relationships only



- (c). *Award [2 max]*
The query provides a virtual representation of the database;
This shows the relationship between the different attributes;

A query searches/provides a filter view of the database;
based on the criteria set in the query;

A query helps to manipulate the data in the database;
Using INSERT/ DELETE/UPDATE statement;

A query is used to aggregate / summarize data;
Using functions e.g. SUM(), AVG(), COUNT() etc.;
And enables grouping / sorting of data;

A query can alter the schema / structure of the database;
By modifying fields/data types of relations;
By creating / altering the relationship between tables/relations;

Note: Do not award marks across different clusters

- (d) **Award [4 max]**
Award [1] for selecting FirstName, Surname and PayGrade
Award [1] for correct JOIN (or mentioning the joining of) the tables STAFF and ROLE)
Award [1] for the correct JOIN condition (STAFF.RoleID = ROLE.RoleID)
Award [1] for identifying both conditions (Surname and PayGrade condition)

Example 1

```
SELECT STAFF.Firstname, STAFF.Surname, Role.PayGrade
FROM STAFF INNER JOIN Roles ON STAFF.RoleID = ROLE. RoleID
WHERE STAFF.Surname = 'Waters' AND ROLE.PayGrade = 17;
```

Example 2

```
SELECT STAFF.Firstname, STAFF.Surname, ROLE.PayGrade
FROM STAFF, ROLE
WHERE STAFF.RoleID = ROLE. RoleID
AND STAFF.Surname = 'Waters' AND ROLE.PayGrade = 17;
```

Note: Accept equivalent in Structured English

- (e) **Award [2 max]**
The pay grade values are whole numbers;
The use of an integer reduces the size of the field;
Listing/sorting the staff in order of the PayGrade / increasing or decreasing of PayGrade;
Allows calculation and mathematical functions;

(f) **Award [6 max]**

Award [3 max] x [2]

Award [1] identifying the method

Award [1] for explaining how privacy is ensured using this method

Award [1] for application of the method to hospital staff data (accept other suitable examples)

By separating sensitive data (*paygrade and possibly age/DoB*) into a separate table;

Minimum number of people have access to sensitive data;

And only using the unique identifiers of the *staff* as a reference point;

The use of different levels of access /authorization;

Meaning the minimum number of people have access to this data;

For example, password protect/lock/restrict access to Role table;

Encrypting the stored data in the database (or specific columns in tables);

The data is scrambled/converted to cipher text (and cannot be understood without a key);

For example, only the employees with the key can decrypt the data;

Data anonymisation/masking/obfuscation;

This removes or transforms Personally Identifiable Information (PII) in the data;

For example, suppressing certain characters in the staff pay Grade etc;

Database views can be used to limit data access;

By exposing only specific columns or rows from a table;

For example a view can be created for a dept manager to see the data of staff in their dept only;

2. (a) *Award [2 max]*
Award [1 max] description of conceptual schema
Award [1 max] description of logical schema

The conceptual schema is a high level/least detailed representation of the database;
Conceptual is sometimes described as the “what” Model;
Involves identification of entities and (high level) relationships between them;

The logical schema is more detailed than a conceptual schema and shows details such as field names;
Logical schema is sometimes described as the “How” Schema;
Logical schema is developed from the Conceptual schema;

- (b) *Award [2 max]*
Award [1 max] explanation of why DDL is important
Award [1 max] example of use

A DDL is used to specify the schema of a database;
A DDL allows you to define the tables, fields, and set datatypes (e.g., CREATE TABLE);
A DDL allows you to specify keys/create relationships;
A DDL allows you to modify schema (e.g. DROP, ALTER);

Note: Award 1 mark for any valid DDL command term

- (c) *Award [4 max]*
Award [1] explanation and [1] for example x 2

It helps to identify the entities/tables in the database;
If incorrect, they will not support the purpose of the DB;

The attributes of the tables should be necessary and sufficient for the purpose;
To avoid inefficiencies in the operation of the database;

The keys for accessing the tables need to be identified;
To ensure the user can easily access the data;

The relationships between tables are identified;
To enable the user to perform complex queries across several tables;

Normalization during data modelling reduces data duplication;
Which reduces data anomalies;
And saves storage space;

Enables the database structure to be understood by other stakeholders/analyst/programmers;
Which makes database maintenance easier;

- (d) **Award [3 max]**
Data validation is an automated process that ensures input meets the data entry rules) / Data validation is an automated process that ensures input data is reasonable/valid;
Data verification is the checking of data to ensure it is the input intended;
Using both techniques will provide the optimal solution, data validation can be applied to fields where the values can be easily determined (for example, no date of birth before a specified date), then data verification allows checking where data validation is not possible/appropriate (for example, a person's surname);
Award one mark for each definition [max 2] and one mark for linkage [1]
- (e) **Award [2 max]**
Integrity is maintained by no changes being made until the transaction is complete;
If the transaction cannot be completed it is rolled back (to the original state);
Atomicity (all or nothing): change made or no changes made;

Consistency ensures a transaction does not put the database in an invalid state;
Rules and integrity constraints (validation, entity and referential integrity) are automatically enforced before a transaction is committed;
Preventing anomalies like orphan records and erroneous data;

Isolation prevents interference between concurrent transactions;
Which ensures intermediate/uncommitted values of one transaction are not visible to other transactions;
Through the use of locks and other concurrency control measures;
Preventing errors like the lost update, dirty reads, phantom reads etc;

Durability ensures that the effects of a committed transaction are permanent (irrespective of a subsequent system failure);
Transaction logs record all the changes made during transactions;
If a system failure occurs, the committed transactions in the log are reapplied to the database during recovery;
Data integrity is ensured by preventing data loss;

Note: Award 1 mark if a student only refers to ACID properties without any expansion
- (f) **Award [2 max]**
Referential integrity refers to relationships between tables in a database;
Referential integrity is maintained by the connection between the primary key in one table and the foreign key in another;
Which ensures records are appropriately updated (e.g. FK value cannot be changed to a value that does not exist as a PK);
It cascades updates and deletes from the PK table to the FK table;
To prevent update anomalies and orphan records;

3. (a) (i) **Award [2 max]**
A derived field is calculated by using data that exists within the database;
That means the data does not have to be input / does not take up storage space;
A derived field has currency as it is calculated/derived at the time so accounts for changes in date etc;

(ii) **Award [2 max]**
Award [1] for identifying the data from the database;
Award [1] for the correct formula to create the derived data;

```
[DateOfBirth] = SELECT DateOfBirth FROM Table  
Age = YEAR ([Date]) – YEAR([DateOfBirth])
```

```
SELECT DATEDIFF (YYYY, DateOfBirth, GetDate()) AS Age FROM Table
```

```
SELECT TIMESTAMPDIFF(YEAR, DateOfBirth, CURDATE()) AS Age FROM Table;
```

```
SELECT DATE_FORMAT(NOW(), '%Y') - DATE_FORMAT(DOB, '%Y') AS age  
FROM Table
```

Accept similar structure

Accept answers in structured English example:

Select Date of Birth from the patient table;

Subtract the year of the birth date from the current year to get the “Age”;

(b) **Award [2 max]**

The prerequisite for 2NF is 1NF;
1NF has no prerequisite;

The focus of 1NF is to eliminate repeating groups, ensure atomicity of data values and ensure entity integrity;

The focus of 2NF is to ensure full functional dependency (remove partial dependencies);

1NF applies to any data table;
2NF only applies to tables with composite keys;

- (c) **Award [6 max]**
Award [1] for identifying three tables with correct fields (without PK and FK)
Award [1] for the DEPARTMENT table (correct fields and PK identified)
Award [1] for the DOCTOR table (correct fields and PK identified, without FK)
Award [1] for including FK (DepartmentID) in DOCTOR/APPOINTMENT table **but not both**
Award [1] for identifying the PK in the APPOINTMENT table
Award [1] for including FK (PatientID, DoctorID) in APPOINTMENT table
Award [1] for all fields correct in APPOINTMENT table

Example 1

PATIENT (PatientID, FirstName, Surname, PreferredName, DateOfBirth)
DEPARTMENT (DepartmentID, departmentName)
DOCTOR (DoctorID, FirstName, Surname, DepartmentID*) //accept DoctorName
APPOINTMENT (PatientID*, DoctorID*, date, time)

Example 2

PATIENT (PatientID, FirstName, Surname, PreferredName, DateOfBirth)
DEPARTMENT (DepartmentID, departmentName)
DOCTOR (DoctorID, FirstName, Surname) //accept DoctorName
APPOINTMENT (PatientID*, DoctorID*, date, time, DepartmentID*)

Award the mark for FK if not explicitly indicated as long as the name matches the PK

Accept an additional field as PK in APPOINTMENT:

APPOINTMENT (AppointmentID, PatientID*, DoctorID*, date, time)

*Accept a **dateTime** field in APPOINTMENT instead of **date** and **time** separately.*

Accept LOCATION in place of DEPARTMENT (and LocationID in place of DepartmentID):

LOCATION (LocationID, LocationName)

Note: Apply the MS according to example 1 or example 2

Note: The relation names do not matter; check the fields used

4. (a) **Award [2 max]**
object-oriented;
network;
spatial;
multi-dimensional;
No-SQL (could be called Not Only SQL);
- (b) **Award [2 max]**
Subject-oriented;
Repositories of integrated data from multiple sources;
That stores both current and historical data/time-variant;
Consolidates/organises data for analysis;
Used in support of management decision making;
- (c) **Award [4 max]**
The data required for a data warehouse may require the aggregation of data from multiple sources;
This data may contain errors, have empty data fields, contain duplicated data or require simplification;
The format of the data needs to be revised so it is in a consistent/standardized format;
Once the data has undergone transformation it is ready for analytics to take place and is loaded into a data warehouse;
Extract can be transformative or as-is;
Extract will identify data for transfer;
Transformation could include – data cleaning, denormalization, normalisation, enrichment, aggregation etc;
Accept a valid example of transformation, e.g., changing all MM/DD/YYYY to DD/MM/YYYY;
- (d) **Award [2 max]**
The timestamping of data provides a snapshot/recorded representation of data at a particular date/time;
The usefulness of information is often time dependant;
This allows data from different times to be compared;
And enables trends to be established;
- (e) **Award [2 max]**
Data extracted in real time from transaction systems;
Data stream is transformed/standardised/augmented;
Data is time-stamped;
The system executes a transaction (in short intervals) to load the data into the data warehouse;
- Note:** award only 1 mark for generic answers about ETL processes that disregard the real time in the question
- (f) **Award [2 max]**
Data is collected, standardized and transformed into numeric values;
Statistical methods or rule-based systems used to build a model of normal behaviour;
Deviation thresholds are determined;
The model is applied to new data to determine outliers;
Results can be subjected to further investigation to determine error/fraud;
Or to discover new trends;

- (g) *Award [6 max]*
Award [4 max] for a contrasting of cluster analysis and association rules
Award [4 max] for a comparing of cluster analysis and association rules

Cluster analysis as a data mining function that performs analysis of abstract objects;
Which are grouped into classes of similar/dissimilar objects / finds groups of objects based on their similarity/dissimilarity;
Cluster analysis could be used to determine relationships that are not detectable using association;
The input data points are mostly numerical;
The output is groups (clusters) of data points;
Cluster analysis is applied in areas like anomaly detection, customer segmentation, document classification etc;

Association is rule based;
And determines relationships / the co-occurrence between large sets of data items;
The greater the confidence value the more reliable the (association) rule is;
The input data points are categorical/transactional;
The output is represented by rules (e.g., $A \Rightarrow B$);
Association technique is applied in areas like market basket analysis, recommender systems, medical diagnosis, etc;

Similarities

Both techniques use unsupervised learning techniques;
The techniques do not rely on predefined target variables or class labels;
Both depend on feature similarity or co-occurrence of data points;
Association techniques use the frequency of co-occurrence to define confidence, and clustering uses distance between data points to determine the degree of similarity;
Both are data exploratory techniques;
Whose outputs show discovered relationships between the data points;

Option B – Modelling and simulation

5. (a) **Award [1 max]**
TemperatureReading;
DailyMaximumTemperature;
TotalTemperature;
Accept variables where the meaning is similar to the above.

Note: Only one answer is required to achieve the one mark.

- (b) **Award [3 max]**
Identifying data for mean daily maximum temperatures **and** years / selecting the cells for temperature and year data;
Selecting a line graph;
Including the correct chart heading;
Labelling both axes as 'Temperature (°C)' and 'Year';

- (c) **Award [2 max]**
Displays the data visually to allow trendlines to be identified;
This trend displays an increase in mean daily temperatures over time;
Data can be compared with other similar charts (for different regions or time frames);
Complicated data is easier to understand if displayed visually;

Accept other meaningful answers.

- (d) **Award [2 max]**
24 Recordings over a day rather than 1440;
That will require less processing / less storage;

The minor changes per minute;
Should not affect the study;

- (e) **Award [5 max]**
Create two (parallel) arrays of size 24 (TIME and TEMP);
Initialise the time array (with 00:00, 01:00 to 23:00 for the full 24-hour clock);
Use a loop counter (N);
Loop 24 times;
Show TIME(N);
Input TEMP(N);

(f) **Award [8 max]**

Initialise variables for totalling temperatures and nighttime temperatures;
 Initialise variables for MIN and MAX; // MIN to a high number or first element of TEMP array
 // MAX to a low number or first element of TEMP array

Use a loop with 24 iterations;
 Totalling all hourly temperatures from TEMP array;
 Selection statement to test if current TEMP value is lower than MIN and if so, assign it to MIN;
 Selection statement to test if current TEMP value is higher than MAX and if so, assign it to MAX;
 Selection statement for adding a nighttime temperature to the nighttime total;
 Calculation of mean temperature and mean nighttime temperature;
 Both mean temperature calculations only carried out once, outside the loop;

Example answer (accept any form of pseudo-code)

// both array TEMP (used to store hourly temperatures) and array TIME
 (used to store times in 24-hour format) are being used but do not have
 to be declared and initialised

```
TOTAL = 0
NIGHT_TOTAL = 0
MIN = 1000
MAX = -1000
loop T from 0 to 23
  NEXT = TEMP[T]
  TOTAL = TOTAL + NEXT
  if TIME[T] >= 19:00 AND TIME[T] < 07:00 then
    TOTAL_NIGHT = TOTAL_NIGHT + NEXT
  end if
  if NEXT < MIN then
    MIN = NEXT
  end if
  if NEXT > MAX then
    MAX = NEXT
  end if
end loop
MEAN_TEMP = TOTAL / 24
MEAN_TOTAL_NIGHT = TOTAL_NIGHT / 12
```

6. (a) **Award [2 max]**
A model is a mathematical representation / abstraction (of a real-life situation / system);
A simulation is the running of a mathematical model over time (on a computer);

(b) **Award [3 max]**
Collect / analyse data (for any pair of temperature, sea level, flooding);
Suggest rules (for the relationship);
Check the (suggested) rules with actual data;

*Accept for [1 mark] any reference to data validation rules.
Do not accept inputting data into the simulation.*

(c) **Award [4 max]**
Find historical data;
That is not impacted by external factors / that is consistent with the model;
Input the (observed) data;
Verify the simulation against the known results;
Modify the algorithms / model (if necessary);
Repeat if necessary;

*Award [2 max] for an evaluative statement such as:
This use of test cases will help to improve the model;
And (therefore) make the simulation more accurate;*

*Mark as [4] + [2]
Accept other valid statements.*

(d) **Award [5 max] overall**

Award [3 max] advantages
This is a simplified version of a far more complex system;
That can be set up with minimal time and cost factors;
It can produce measurable visual results;
It can help predict floodings / to be prepared;

Award [3 max] disadvantages
It is a crude solution / only testing one or two factors;
It does not consider social/cultural factors;
Makes the decision based on (potentially) inaccurate / unchecked facts;

Accept any other valid answer

7. (a) **Award [1 max]**
Visualization is a graphical representation of data;
- (b) **Award [2 max]**
The image in memory is stored as a mathematical model;
Images (in memory) are rendered;
To create a 3D visualization;

- (c) **Award [5] max**
Award [2 max] for time consideration
Example answer:
3D animation is very complex (in terms of programming);
It requires a lot of time for processing / rendering;

Award [2 max] for memory consideration

Example answer:
Rendering different layers / transitions;
Which requires a lot of RAM;
May (also) require the use of secondary memory / GPU;
Which has the issue of a different processing speed;

Award [1 max] for an evaluative statement

Example answer:
Therefore, 3D animation requires sufficient, fast primary and secondary memory.

Mark as [2] + [2] + [1]

Note: the memory marks require some specificity – do not award for generic statements such as “it needs lots of memory”.

Note: the concluding mark can be gained for any sensible concluding statement that addresses time and memory.

8. (a) **Award [3 max]**
Establish initial population;
Define the fitness function;
Selection;
Crossover;
Mutation;
Evaluation;
Replacement;
Reproduction;
Accept other valid answers
- (b) **Award [2 max]**
Input layer;
(One or more) hidden layers;
Output layer;
- (c) **Award [5 max]**
Teach the rules for moves allowed in the game;
Teach typical moves based on previous moves;
Allow the neural network to predict moves (based on previous moves);
Learn from the results;
Play against human players (of differing abilities);
Play against itself;
Use feedback loops to modify its decision making;
Link back to the value network to judge how successful it has been;
Accept other valid answers.
- (d) **Award [2 max] for statements about supervised learning**
Supervised learning uses labelled data;
To teach known outcomes based on given inputs;
Which can make the results more predictable;
- Award [2 max] for statements about unsupervised learning**
Unsupervised learning uses unlabelled data;
The AI learns by looking for patterns and relationships (on its own / with minimal human input);
Unsupervised learning can lead to unpredictable outcomes;
- Award [1 max] for a concluding statement**
Unsupervised learning is more random with potentially deeper learning than supervised learning;
Learning Go may benefit more from unsupervised learning because it may lead to more interesting gameplay.

Mark as [2] + [2] + [1]

(e) *Award [5 max] overall*

Award [3 max] from any cluster

Pre-trained Language Models / Transferred Learning;

Are pre-trained on large data sets;
And can be fine-tuned for specific tasks;
With a smaller set of labelled data;

Transformer Architecture;

Use self-attention mechanisms;
To effectively model long-range dependencies in text;
Processing sequences in parallel;
Allowing faster and more accurate training on large data sets;

Contextual Embedding;

Words are represented with a context (for example, school-building vs school-fish);
Which greatly enhances semantic understanding / reduces ambiguity;

Larger / More diverse data sets for training

Reduce bias;
Increase robustness;
Enhance understanding of nuances / idiom;

Option C – Web science

9. (a) *Award up to [2 max]*
Award [1] for describing the internet
Award [1] for describing the world wide web

the internet is a global network of inter-connected computers;
the internet uses internet protocols addresses or IP addresses;
the internet is the network of computers infrastructure that the world wide web is a service on;
the World Wide Web is a subset of the internet which uses protocols like HTTP or Hypertext transfer Protocol;
the World Wide Web is a service on the internet/a collection of information and resources that are accessed via the internet using web browsers.

- (b) *Award [2 max]*

Protocol - https;
Host - www.bbc.co.uk
Domain – bbc.co.uk;
Path - /index.htm or /index;
A URL uniquely identifies a resource on the web.

- (c) *Award [2 max]*

FTP establishes 2 processes one for control and one for data transfer;
FTP establishes two connections, one for control and one for data transfer;
FTP usually uses port 21 for control and port 20 for data transfer;
FTP facilitates the transfer of files efficiently between a client and a server.

Note: Do not accept vague answer such as it is set of rules that governs transmission of data/files

- (d) *Award [3 max]*
Award [1] for stating the protocol required
Award [1] for stating what the web browser retrieves
Award [1] for further information

e.g. Applies the appropriate protocols to enable communication (HTTP, HTTPS etc) with the web server; / FTP runs over the TCP (Transmission Control Protocol)

e.g. Provide a way to navigate to, access and fetch/retrieve web pages and provide internet services. (HTTP Request and Response): A URL is sent to DNS server for corresponding IP address, which if found is sent back to the browser.

e.g. Support and enable multimedia features (using plug-ins);
provide security - Authentication and encryption services;
display web pages properly based on standards; / The data is typically written in a markup language, such as HTML, XML, or similar formats.
support and enable multimedia features (using plug-ins);
provide security - Authentication and encryption services;

Note: This question expects students to use key terms like request/response, render, protocol (HTTP/HTTPS), etc. Avoid vague answers like “it searches the internet” or “shows information.”

10. (a) *Award [1 max]*

a software system/program/application that searches the world wide web or a database;
for keywords that match the user's specification;
and filters, such as date, usage rights, size, currency, etc. if used;

(b) *Award [2 max]*

Note: Web crawler is also called a bot or spider
start at “seed”, this could also be a designated starting page:
review and categorise web pages based on criteria for information searched for;
looks for keywords, content, hyperlinks, metatags etc;
follows hyperlinks from page to page;
can move through a site depth-first or breadth-first;
review can be stopped by rules set in sites robots.txt file;
often copies part or all of the content of visited page:
the crawler continues this cycle, constantly updating the index with new or changed content.

(c) *Award [2 max]*

Web crawlers look for keywords from the meta keywords/ tags;
in the meta description, title and potentially the URL of the page;
Then determine how many times the keywords appear in the body of the page;
and then the ranking algorithm uses this as part of the ranking process. Keywords are important
for web indexing
Support SEO (Search Engine Optimization): Strategically using keywords enhances visibility
and helps websites attract more traffic from search engines.

(d) Award [6 max]

Command term is **Discuss** – requires balanced review with conclusion clearly supported.

black Hat Search Engine Optimisation is manipulating the search engine guidelines/rules to gain a higher search ranking and therefore potentially increased share of revenue or improved traffic;

Black hat techniques include:

keyword stuffing, poor/low quality content (duplicated content or rewritten content), hidden keywords, paid links and link farming, cloaking (Showing one piece of information to SE and different to users), “sneaky redirects”, blog comment spam and ranking stuffing;

Advantages:

brings increased traffic/visitors;
 directs user to content, developer wants them to see;
 gives a higher search engine ranking based on number of hits;
 potentially increases revenue;

Disadvantages:

reputational damage
 search engine penalties for inappropriate techniques;
 blacklisting
 lower search engine score / Initial ranking may improve, but long-term score can drop if detected and penalized by search engines.
 flagged as an unsafe site;
 ethical issues—inaccurate, unreliable, inappropriate content etc;

Conclusion:

candidates need to present a position and support their decision. Award marks for the support of the position rather than stating their position.

Note: Responses should focus on the organization (e.g., Xero) as the primary stakeholder when evaluating the use of black hat SEO techniques.

Marks	Level Descriptor
0	No relevant response. The answer is missing, completely off-topic, or shows no understanding of black hat SEO.
1–2	Limited. Shows minimal understanding of black hat SEO. Discussion is superficial, with little or no explanation of advantages or disadvantages. May mention a technique but lacks clarity or context. No valid conclusion.
3–4	Adequate. Some discussion of advantages and disadvantages, though, may be imbalanced or descriptive. Includes at least one clear example of a black hat technique. An attempt at a conclusion is made but may be underdeveloped.
5–6	Proficient. Balanced and well-developed discussion of both advantages and disadvantages. Includes multiple black hat techniques with explanations (e.g., keyword stuffing, cloaking, link farming). Presents a clear and well-supported conclusion that evaluates the organization’s decision.

- (e) **Award [2 max]**
Award [1 max] describing the deep web
Award [1 max] describing the surface web

deep web is a part of the world wide web that is not indexed by the search engines and therefore not discoverable by normal search engines, it includes databases and dynamic pages that requires authentication;
the surface web is indexed by the common search engines and are therefore accessible to most users and is searchable by “normal”/common search engines like Google, Bing etc;

Note: Do not accept answers referring to the dark web. For simple explanations, such as deep web is not discoverable, and the surface web is discoverable without pointing to the indexing will get a max of 1.

- (f) **Award [4 max]**
user selects a file and presses/click the submit button;
data is extracted from the file including file size, file type, name etc;
file is checked against criteria in conditional statements for file extension (pdf, doc and docx);
and file size (greater than 2097152);
if there are errors this is added to the error array;
if the error array is empty upload the file and print that the file upload was successful;
otherwise print the errors from the error array;

Output of the PHP code, for reference purpose.

Curriculum Vitae

Choose File No file chosen

Submit

- (g) **Award [3 max]**
Award [1 mark] for stating the concept clearly
Award [2 max] for implications/development

Concept:

the processing of the script occurs on the web server rather than on the browser of the client;

Implications:

there is a consistent result regardless of the processing capacity of the client device;
only the processed result is seen;
and the processing is secure on the server;
underlying data is secure on the server;
the speed of processing will be consistent;
offers more consistent experience for end user;
and greater control for the site owner;

11. (a) *Award [2 max]*

allow the compressed media to be reconstructed perfectly/completely;
no data is deleted during the lossless compression.
can replace common non-essential data with a token or shorter form of itself;
uses a shorthand version to replace repeating elements;
some lossless algorithms replace more common sequences with shorter sequences used to represent them; example Huffman coding. Example a is encoded as "0" b is encoded as "01" c is encoded as "10" D is encoded as "11" rather than full bitstream;
replaces sequences of similar pixels/recurring patterns with a shorter representation in image or video files (e.g., run-length encoding).

(b) *Award [2 max]*

is a standard that is:
openly accessible;
usable by anyone;
not owned by any governing body or private entity.
can be adopted, implemented and updated;
jointly managed by stakeholders;
usually has rules around adjustments and updates;
designed to ensure interoperability between systems and platforms.

(c) *Award [6 max]*

Award [2 max] for conceptual understanding
Award [2 max] for positive impacts on privacy
Award [2 max] for negative impacts on privacy

Conceptual understanding

the decentralised web is where operation, ownership and control of the internet is decentralised to the users away from the major organisations like Google, Facebook, Microsoft, etc;
the decentralised web uses peer-to-peer. No one entity has control of the web;
decentralisation means that there is less opportunity for censorship; people can say what they want, but also removing data is potentially harder as control is not from a single point;

Positive impacts on privacy

individual control means that the user can allow or restrict information sharing, thus potentially improving privacy by reducing third- party access to the data/information;
ownership of the data remains with the user;
decentralisation removes the middleman in data transfers, enabling direct transfer of data and therefore greater control and privacy. Middle man could be organisations/entities that act as mediums for data transfer;
cryptocurrencies are an aspect of the decentralised web and bring privacy and security to financial transactions. Cryptocurrencies are usually in decentralised control. Cryptocurrencies use strong encryption to secure the transaction records;
decentralisation reduces the surveillance on user activities, as the activity may be direct and not monitored or logged. This increases privacy;

Negative impacts on privacy

since there is no overall control, there is the ability to publish any information by anyone; the decentralised web operates on a trust model. This makes it harder for an individual to take down a page and harder for an organisation or government to restrict access; decentralized systems may lack consistent privacy standards, exposing users to platforms with weak or nonexistent protections; because data is stored across multiple nodes, there is a greater surface area for potential breaches or leaks; accountability is often unclear in decentralized systems, making it difficult for individuals to seek recourse for privacy violations; decentralisation may mean that some of the data is stored in systems that are not as secure and do not conform to industry practices, thus making them more vulnerable to illicit activity; cryptocurrencies are an aspect of the decentralised web and bring privacy and security to financial transactions. Cryptocurrencies are usually in decentralised control. Cryptocurrencies use strong encryption to secure the transaction records;

Mark	Level Descriptor
0	No response, or the response shows no understanding of the decentralized web or its relation to privacy.
1–2	Limited. Minimal understanding shown. May describe the decentralized web vaguely or mention privacy in a general sense. No clear link between the two. Argument is one-sided or unclear. No valid conclusion.
3–4	Adequate. Some relevant points on how decentralization can affect privacy (positively or negatively). May lack balance, depth, or examples. Conclusion is attempted but may be simplistic or only loosely supported.
5–6	Strong. Clearly discusses both the benefits and challenges of decentralization regarding privacy. Specific examples or implications are given (e.g., peer-to-peer control, censorship resistance, challenges to the “right to be forgotten”). A well-reasoned and supported conclusion is provided.

- (d) **Award upto [2 max] for each cluster**
Award [1] for identifying a relevant development from the cluster
Award [1] for explaining its significance or impact

Command term is "explain"—give a detailed account including causes and reasons.

Economic and Operational Efficiency

no need to develop physical software packages or large infrastructure;
lowers upfront capital investment for both Xero and its subscribers;
Infrastructure-as-a-service (IaaS) and platform-as-a-service (PaaS) models reduce operational overhead;
subscription-based revenue provides predictable income and scalability;
digital ads, social media campaigns, and SEO are more cost-effective than traditional media;
cloud storage protects data from loss; Data is backed up and secured against accidental deletion or system failure;
web platforms allow easy deployment; Services can be launched and updated quickly without physical distribution, enabling global access;

Accessibility and Scalability

users from any region can access the platform without physical distribution;
time zones and geographic distance no longer limit service delivery;
widespread smartphone and internet access increases the user base;
cloud-based platforms allow instant access to financial data and services anytime, anywhere;

Customer Convenience and Ecosystem Integration

customers can subscribe and pay through secure web portals;
multi-currency and third-party API support enhances utility for global users;
users benefit from continual improvements without needing to manually install updates;
helpdesks, bots, and user forums offer scalable customer support;

Note: Candidates need to show a clear link to how these developments have an impact and enable the small companies like Xero to gain global access, promote their product or service, and offer scalable, viable products globally.

- 12.(a) (i) *Award [1 max]*
Node— a point in the network where connections intersect or branch out— usually represented as a website or page, sometimes called a vertices;
- (ii) *Award [1 max]*
Edge is the link or connection between nodes/websites (or vertices);
- (b) *Award [1 max]*
The Strongly Connected Core (SCC) is a subset of nodes which have multiple connections between the nodes each node is connected to every other node through one or more direct links ;
The *in* and *out* components represent nodes that have connections to nodes in the SCC;
- (c) *Award [3 max]*
Award [1] for a correct definition of Metcalfe's law or its formula/idea
Award [1] for identifying how the SCC fits (i.e., high connectivity within).
Award [1] for providing a reasoned limitation of using Metcalfe's Law for the whole web.

metcalfe's law assumes full connectivity; each node can connect to every other node in the network;

simple models that broadly show general trends;

as the diameter of SCC increases, the model becomes a better approximation. It models the possible growth. / The network's value grows with the square of the number of nodes;

the bowtie structure shows only SCC is fully connected; IN and OUT components exist but are not fully connected to each other or the entire network; / e.g., In both 2000 and 2012, large portions (IN and OUT) cannot reach or be reached by all nodes;

IN increased from 21% to 32%, meaning more nodes can reach SCC but not each other. OUT decreased from 21% to 6%, showing fewer nodes can be reached from SCC;

metcalfe's Law overestimates overall connectedness; it does not account for directionality or disconnected parts of the web;

does not account for the Deep web (but neither does the bowtie model/structure);

it's not fully applicable to the real-world structure of the web;

- 13.(a) **Award [2 max]**
Award [1] mark per point

semantic web is developed using well-structured data and tags/meta-data (ontology);
the semantic web uses data in web pages that is structured and tagged. (ontology);
in such a way that it can be read directly by computers.;
the Semantic Web aims to allow different systems, platforms, and applications to understand and share data meaningfully;
semantic web is meant to improve human-machine co-operation;

- (b) **Award [4 max] Award one mark per point**

ontologies use a common formal language;
this allows a more consistent and deeper understanding of the relationships;
formal structure better suits computer systems than informal structure of folksonomy;
formal structure reduces the ambiguity;
formal structure improve the search;
formal structure can be used by search engines more easily;
examples include Dublin core;

- (c) **Award [2 max]**
Award [1] mark for ambient intelligence
Award [1] mark for collective intelligence

ambient intelligence is an electronic system like Alexa or Siri that is sensitive to and responsive to the presence of people. (Ambient relates to the immediate surrounding of something);
collective intelligence is sometimes called group intelligence where via collaborations, collective efforts and consensus contributes to understanding;
an example of this could be social media, search engines and crowdsourcing;

(d) *Award [6 max]*

collective or group intelligence uses collective efforts to produce a greater product;
many people contribute to the collective intelligence of search engines through the actions of web crawlers indexing many sites;
search engine algorithms like PageRank and HITS uses the links between pages to give a value to the importance of the page;
and potentially to the authority of the page;
some search engine algorithms also use the popularity of the page (number of visits, click-through time, etc.) Hence collective visiting of a page influences the result;
using ontologies, the formal structured meta data / website meta tags and folksonomies, the informal use of popular tags, contributes to page ranking and therefore the result;
open databases like Wikipedia which require user contribution and maintenance, contribute to the results in searches and thus contribute to the product.

Note: Conclusion drawn on the points made. Do not award marks for the statement of position; rather, award marks for the supporting argument/statements in the conclusion.

Marks	Criteria
1–2	Limited or one-sided answer. May describe collective intelligence but lacks detail or connection to search engines.
3–4	Reasonable explanation of how collective intelligence contributes. Some examples may be included. May lack balance or evaluation.
5–6	Well-balanced, thoughtful discussion with clear examples (e.g., PageRank, user clicks, Wikipedia). Shows understanding of both benefits and limitations. Provides a reasoned conclusion.

Option D – Object-oriented programming

14. (a) **Award [1 max]**

aggregation;
has a;

Note: As long as the relationship is clearer, award mark.

(b) **Award [2 max]**

instantiation is the process of creating a new object of the class;
involves calling the constructor of the class;
object's attributes are initialized / object gets memory during instantiation;

(c) **Award [3 max]**

Award [1] for `Artist A // (accept any name)`

Award [1] for `new Artist()`

Award [1] for correct arguments `"Thomas Lucas", "Ireland";`

`Artist A = new Artist ("Thomas Lucas", "Ireland");`

(d) **Award [2 max]**

Award [1] mark for -

private attributes can't be accessed outside the class / private attributes need accessor/mutator methods to be accessed/modified;

Award [1] mark for -

using the reference from the Artist/Artwork class.

Note: A candidate can score both marks in one sentence.

Example - The attribute `artworkTitle` in the `Artwork` class can only be accessed by using the appropriate accessor method in the `Artist` class.

(e) **Award [4 max]**

Easier / faster to debug;

because there are far fewer mistakes in the smaller/individual modules;

Speedier / faster completion of the project;

because different teams work on different modules (simultaneously);

Collaborative / qualitative development;

as different teams work on different modules at the same time / teams (programmers) with expertise work on specific modules;

Facilitates reusability of the code;

as the existing modules can be reused across other modules;

Improves code readability / organisation;

smaller manageable modules leading to better logical organization;

Reduces the coupling effect;

leading to easier/faster maintenance of module not affecting others;

Mark as [2] and [2]

Note: Do not allow mixing the points from different clusters. Do not allow easier/faster without any further description in context of modularity.

- (f) **Award [3 max]**
Award [1] for three boxes and correct class name;
Award [1] for correct attributes and methods (at least one method (constructor / accessor / mutator / other methods);
Award [1] for correct symbols '-' and '+' (accept 'private'/'public');

Artist
- artistName:String - artistLocation:string - noOfArtworks:int - theArtworks[]:Artwork
+ getName() : String + getLocation() : String + getArtwork(x : int) : Artwork + addArtwork(x : Artwork): void + commissionToPay() : double + toString() : String

15. (a) (i) **Award [1]**
String

(ii) **Award [1]**
int;
float;
double;

(iii) **Award [1]**
boolean;

Note: Award the mark as long as the data type is stated correctly.

(b) (i) **Award [4 max]**

Award [1] for each advantage and [1] for expanding it

Reuse of code;
through inheritance sub classes can use methods of the superclass / less code needs to be written / saves time in coding;

Code flexibility/extensibility;
extending the parent's action and data, without redefining them (polymorphism);

child class redefines the base class methods (Overriding);
to provide a different functionality to existing method of the parent class;

Easy to debug;
As the existing code (base class methods) are already tested / less code needs to be debugged;

Easier to maintain;
as the changes in the parent class are automatically reflected in the child class;

Mark as [2] and [2]

(ii) **Award [3 max]**

All the of the common attributes related to the art works could be in a superclass
(Artwork);

Specific attributes related to the different types of artwork could be added to the sub
classes (Painting/Sculpture/Performance);

When instantiating an artwork the specific artwork would be chosen rather than "Artwork";

Note: Reponse must refer to Artwork as base class and Painting / Sculpture / Performance as subclass, to achieve the full marks.

(c) Award [6 max]

Award [1] for using a `boolean` flag or similar variable to stop either of the loops;
 Award [1] for correct outer loop to run through the `theArtists` array;
 Award [1] for correctly finding the required Artist;
 Award [1] for correct inner loop to run through the `theArtworks` array;
 Award [1] for an attempt to search for the artwork;
 Award [1] for correct use of appropriate accessors and `.equals()`; // accept `'=='`
 Award [1] for an attempt to set the `isSold`;
 Award [1] for correctly calling `issold()`;

Example 1:

```
public void Sold( String artistName, String artworkTitle)
{
    for (int i = 0; i < theArtists.length; i++)
    {
        if(theArtists[i].getName().equals(artistName))
        {
            for (int x = 0; x < theArtists[i].getnoOfArtworks(); x++)
            {
                if(theArtists[i].getArtwork(x).getArtworkTitle().equals( artworkTitle))
                {
                    theArtists[i].getArtwork(x).isSold();
                }
            }
        }
    }
}
```

Example 2:

```
public void Sold(String artistName, String artworkTitle)
{
    boolean found=false;
    int i=0;
    while(i<theArtists.length && found ==false)
    {
        if(theArtists[i].getName().equals(artistName))
        {
            int j=0;
            while(j<theArtists[i].getnoOfArtworks() && found == false)
            {
                if(theArtists[i].getArtwork(j).getArtworkTitle().equals(
                    artworkTitle))
                {
                    theArtists[i].getArtwork(j).isSold();
                    found=true;
                }
                j++;
            }
        }
        i++;
    }
}
```

Note: Accept any name for `getnoOfArtworks()` as long as it returns `noOfArtworks`.

16. (a) **Award [1 max]**
0;

(b) **Award [5 max]**

Award [1] for correctly creating a temporary and returning an array of `Artwork` type;

Award [1] for correct initialising and updating of `max_index`;

Award [1] for correct outer loop up to `< noOfArtworks-1` (or up to 5 times);

Award [1] for correct inner loop;

Award [1] for correct comparison of price;

Award [1] for correct implementation of swap;

Award [1] for transferring the five most expensive artworks to the temporary array;

Note: Award only MP1, MP7, and another [1] for implementing a different sorting algorithm. If no sorting used then only MP1 and MP3/MP7 to be awarded.

```
public Artwork [] sortArt()
{
    for (int i = 0; i < noOfArtworks-1; i++)
    {
        int max_index = i;
        for (int j = i+1; j < noOfArtworks; j++) // accept j=i;
        {

            if(theArtworks[j].getArtworkPrice() >
               theArtworks[max_index].getArtworkPrice())
            {
                max_index = j;
            }
        }
        Artwork temp = theArtworks[max_index];
        theArtworks[max_index] = theArtworks[i];
        theArtworks[i] = temp;
    }
    Artwork [] expensive = new Artwork[5];
    for (int i = 0; i < 5; i++)
    {
        expensive[i] = theArtworks[i];
    }
    return expensive;
}
```

Note: Accept `< theArtworks.length-1` for the outer loop. However, a condition must be placed inside the loop to check for the null values in the `theArtworks[]` to avoid the `NullPointerException`. Allow any name for the getter `getArtworkPrice()` as long as it returns the `artworkPrice`.

(c) **Award [6 max]**

Award [1] for looping through the Artwork array up to noOfArtworks;

Award [1] for an attempt to check if the artwork is sold;

Award [1] for correctly using the accessor `getIsSold()`;

Award [1] for correctly accessing the price of the sold artwork;

Award [1] for correctly calculating the total (summing of the prices for all the sold artwork);

Award [1] for correctly calculating and returning the commission;

```
public double commissionToPay()
{
    double total = 0;
    double commission = 0;
    for (int i = 0; i < noOfArtworks; i++)
    {
        if (theArtworks[i].getIsSold()==true)
        {
            total = total + theArtworks[i].getArtworkPrice();
            // accept other name for the accessor
        }
    }
    commission = total * 0.15;
    return commission;
}
```

Note: Accept any name for getters `getIsSold()` and `getArtworkPrice()` as long as they return `isSold` and `artworkPrice`.

(d) **Award [2 max]**

Award [1] for -

has 2^{16} (about 64,000) characters / has more bits/bytes (in compare to ASCII);
provides a standard encoding language and methods;

Award [1] for -

which enables characters from different alphabets/languages to be used/stored (promoting internationalization);

making it easier to use without language barriers (promoting internationalization);

17. (a) *Award [2 max]*
Award [1] for identifying the advantage, [1] for expanding the advantage

Simplifies Data Handling / Faster Development Time
library collections offer built-in methods to add, remove, and access elements;
reducing manual coding and making development faster and less prone to errors.

Improves Code Readability and Maintainability
using library collections lead to cleaner, more readable code than manual array handling;
making it easier for others to understand and maintain using familiar Java classes;

Access to Useful Built-in Methods
library collections offer methods like size(), isEmpty(), get(), remove() and etc;
that simplify common tasks and help reduce logical errors;

Note: The response must highlight the advantage of library collections (LinkedList or ArrayList).
Generic responses about library methods receive a maximum of 1 mark.

- (b) *Award [5 max]*
Award [1] for creating a new node nnode;
Award [1] for iterate through the list (starting at the head pointer);
Award [1] for identifying the correct position for nnode until a larger node is found or null;
Award [1] for correctly setting the next of nnode to the next of previous node;
Award [1] for correctly setting the next of previous node to the nnode;
Award [1] for updating the head pointer if nnode to be added at the start of the list;

Example:

Create a new node;
Iterate through linked list, starting at the head pointer;
Using the compareTo(), identify the correct position until a larger node is found or null;
Set the next of the new node to the next of the previous node;
set the previous node to the new node;
Update the head pointer if the new node is to be added at the start of the list;

Note: Do not accept the code as answer. The question specifically says not to write code.

- (c) (i) **Award [2 max]**
 is a technique/process where a method calls itself;
 each call involves updated parameters;
 until it reaches the base case;

- (ii) **Award [5 max]**

Award [1] for correctly checking base case and returning 0;
Award [1] for comparing the max with customer's maximum budget;
Award [1] to return 1 + the subsequent recursive call;
Award [1] to return the subsequent recursive call, when comparison results in the budget is not enough;
Award [1] for decreasing n for each recursive call;

```
public int countRecursive(LinkedList <Customer> customerList, int
max, int n)
{
    // Base case: size of list = 0
    if (n==0)
    {
        return 0;
    }
    else if (customerList.get(n-1).getMaximumBudget() > max)
    {
        return 1+ countRecursive(customerList, max, n-1);
    }
    else
    {
        return countRecursive(customerList, max, n-1);
    }
}
```

- (iii) **Award [4 max]**

Award [1] for identifying the reason, **[1]** for expansion

Memory Usage

recursive algorithms use the call stack for each call, leading to excessive memory usage;
 this is inefficient for simple tasks like counting through a list, especially when the list is large;

Processing Overhead

each recursive call adds overhead due to repeated function calls and return handling;
 an iterative loop processes the same task more efficiently with lower runtime cost;

Complexity

the task of counting elements based on a simple condition does not require backtracking or divide-and-conquer.;
 linear loop with a counter is easier to write, read, and debug for this scenario;

Stack Overflow

extensive recursion can exceed the maximum call stack size, especially with a large linked lists;
 this can cause runtime errors, which are avoidable with iteration;

Mark as [2] and [2]

- (d) **Award [2 max]**
Award [1] for identifying the advantage, [1] for expansion
improves code readability / maintainability;
consistent naming helps developers understand the purpose (variables/methods/classes) easily;
- faster collaboration;
when team members follow the same conventions, it reduces confusion and makes code integration smoother;