Big Data - AQA Computer Science Cheat Sheet Cheat Sheet

by ehe via cheatography.com/212604/cs/46225/

5 Vs of Big Data

Variety	The range of data formats and data types collected
Value	How useful data is to an organisation
Veracity	Accuracy and quality of data
Volume	The amount of data (if the volume is large enough, it is considered big data)
Velocity	How quickly the data is generated

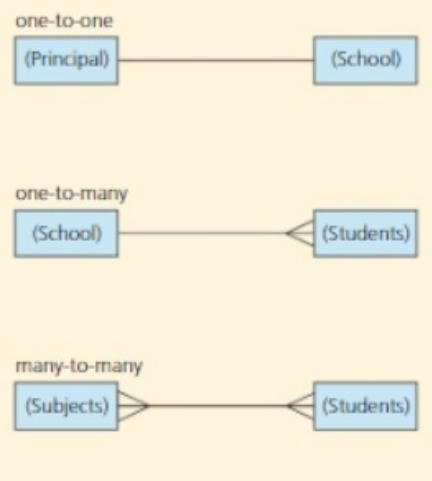
Example Big Data Applications

Healthcare	Predict disease outbreaks and personalize treatment plans
Entertainment	Recommend content and analyze audience preferences
Transportation	Improve traffic flow and predict maintenance needs
Retail	Optimize inventory and personalize customer recommendations
Finance	Detect fraudulent transactions in real-time

Relationships in Relational Databases

One-to-one	One school has one principle
One-to-many	One school has many students
Many-to-many	Many students and take many subjects

Entity Relationship Diagram



Relational Databases

Table	Set of facts or figures that are set out in a column and row structure
Flat-file database	Database that stores all data items using one table
Data redundancy	When data is unnecessarily repeated in a database
Data-entry error	Error that occurs when data is being entered into a database
Relational database	Database that stores data using two or more linked tables
Entity	Person, place or object represented in a table in a relational database
Attribute	Heading for organising data in a relational database
Primary key	Field in a database table that provides a unique identifier for a record/entity

Relational Databases (cont)

Foreign key	When the primary key from one table appears in another table to establish a link between two entities

Query

Simple queries	Only a single search criterion is used to select data items from a database
Complex queries	More than one criterion is used to search a database, a query is used to combine data from more than one table, or calculations are performed using the data in a query or a report
Parameter queries	Queries where the end user provides the search criteria
Wildcard queries	Queries where special characters are used to stand in for unknown characters (this is useful when trying to find lots of data items that are similar but not exactly the same)
Multi-table queries	Use data from more than one data table
Multiple--criteria queries	Use more than one criterion to select data items from a database

By **ehe**
cheatography.com/ehe/

Not published yet.
Last updated 28th April, 2025.

Sponsored by **Readable.com**
Measure your website readability!
https://readable.com

Spreadsheet Model

Function	Sub-program that can exist as part of a bigger program
MIN function	Returns the lowest value in a specified range of cells in a spreadsheet
MAX function	Returns the highest value in a specified range of cells in a spreadsheet
IF statement	This evaluates a condition which determines the path of the program depending on whether the condition is true or false
COUNT function	Checks all the cells in a specified range in a spread-sheet and outputs how many contain a numeric value

Evaluating models

Evaluation: checking the suitability of a solution to a problem

Efficient: the efficiency of a program can be measured by how quickly it runs

User requirements: tasks a user expects of an application

Data type: classification applied to a data item specifying which type of data that item represents, e.g. in a spreadsheet some of the data types available include currency, text and number

Frameworks

Structured Query Language (SQL)	Specialised language for accessing data in relational databases
Query by Example (QBE)	Interface that allows users to select fields and criteria for use in a query in a database application

Sponsored by **Readable.com**
Measure your website readability!
https://readable.com