information systems - Assessment Task 2 – Report

Frank Sharpe l39222594

# Table of Contents

[Table of Contents 1](#_Toc185499661)

[Table of Figures 3](#_Toc185499662)

[ER Diagram 5](#_Toc185499663)

[Tables and relationships 6](#_Toc185499664)

[Students Table 6](#_Toc185499665)

[Enrolment Table 7](#_Toc185499666)

[Medical Info 7](#_Toc185499667)

[Attendance 7](#_Toc185499668)

[Student fee 7](#_Toc185499669)

[Students contacts 8](#_Toc185499670)

[Courses 8](#_Toc185499671)

[Campus 8](#_Toc185499672)

[Certificates 8](#_Toc185499673)

[Global personal ID 9](#_Toc185499674)

[Role permissions 9](#_Toc185499675)

[Database implementation 11](#_Toc185499676)

[Views 13](#_Toc185499677)

[Manager view 13](#_Toc185499678)

[Teacher view 14](#_Toc185499679)

[Section 1: Code solutions to the following query and reporting problems as SQL statements – 16](#_Toc185499680)

[Show a list of names of all students enrolled on a course, showing student names and course names only. 16](#_Toc185499681)

[Show ID’s and names of students who have not enrolled on any course. 17](#_Toc185499682)

[Show all students who started a course between two specific dates. 18](#_Toc185499683)

[Show each course, and the number of students enrolled onto that course. 20](#_Toc185499684)

[Find the teacher with the lowest number of students. 21](#_Toc185499685)

[Show the total amount of turnover (funds generated) from the sales of course. 22](#_Toc185499686)

[Show all students living in a particular city. 24](#_Toc185499687)

[Show all certificates issues between two dates. 25](#_Toc185499688)

[Show all staff with their staff type (teacher, support or management) 27](#_Toc185499689)

[Section 2: Code solutions to the following manipulation problems as SQL statements 28](#_Toc185499690)

[Create a new course called ‘Virtual Dance’ and assign 5 new students to the course. 28](#_Toc185499691)

[Remove a course and all data associated with it from the system. 30](#_Toc185499692)

[Delete any students with a forename of ‘Adam’. 31](#_Toc185499693)

[Delete all courses with the word ‘creative’ as part of the course name. 33](#_Toc185499694)

[Change the contact number of a Student’s Contact 35](#_Toc185499695)

[Change the start date of any students with a forename of ‘Adam’. 36](#_Toc185499696)

[Section 3: Testing, exporting and profiling the data within the database. 38](#_Toc185499697)

[Perform an integrity check on each table within your database. 38](#_Toc185499698)

[Explain what an integrity check covers and why it is important when developing a database and describe other methods of testing a relational database. 40](#_Toc185499699)

[Export students into a csv file. 41](#_Toc185499700)

[Import the data from the courses table SQL database into Microsoft Excel and perform profiling using Power Query 42](#_Toc185499701)

[Section 4: Statistical Analysis 0](#_Toc185499702)

[Data source 0](#_Toc185499703)

[Duration of courses to subscribers 3](#_Toc185499704)

[Subscribers to cost of course 4](#_Toc185499705)

[Sum of subscribers to level of course 5](#_Toc185499706)

[Course name to subscribers 5](#_Toc185499707)

[Lectures vs the Subjects 6](#_Toc185499708)

[Subject to subscribers 7](#_Toc185499709)

[Mean number of subscribers 7](#_Toc185499710)

[Standard devation 8](#_Toc185499711)

[Conclusion 9](#_Toc185499712)

[References 11](#_Toc185499713)

[Bibliography 15](#_Toc185499714)

[Appendix 17](#_Toc185499715)

# Table of Figures

[Figure 1 ( ER diagram) 5](#_Toc185578068)

[Figure 2 – trigger for logging access attempts. 10](#_Toc185578069)

[Figure 3 - sample query used by a staff id = 1 10](#_Toc185578070)

[Figure 4 ( DB SQLite) 11](#_Toc185578071)

[Figure 5 ( DB SQLite) 11](#_Toc185578072)

[Figure 6 ( students table) 12](#_Toc185578073)

[Figure 7 - Manager view table query 13](#_Toc185578074)

[Figure 8 - Query result from the view 13](#_Toc185578075)

[Figure 9 - Teacher Query 14](#_Toc185578076)

[Figure 10 - Query result 14](#_Toc185578077)

[Figure 11 – list of students and course names. 16](#_Toc185578078)

[Figure 12 - no course enrolled. 17](#_Toc185578079)

[Figure 13 - enrolment table. 18](#_Toc185578080)

[Figure 14 - select from date. 18](#_Toc185578081)

[Figure 15 - show number of students on course. 20](#_Toc185578082)

[Figure 16 - teacher with least students. 21](#_Toc185578083)

[Figure 17 - student fee table. 22](#_Toc185578084)

[Figure 18 – calculation of sum 23](#_Toc185578085)

[Figure 19 - student table. 24](#_Toc185578086)

[Figure 20 – find where city = London 24](#_Toc185578087)

[Figure 21 - enrolment table. 25](#_Toc185578088)

[Figure 22 - course end date. 26](#_Toc185578089)

[Figure 23 - select teachers and role. 27](#_Toc185578090)

[Figure 24 – insert the table and 5 new students. 28](#_Toc185578091)

[Figure 25 - added into courses. 29](#_Toc185578092)

[Figure 26 - added into students. 29](#_Toc185578093)

[Figure 27 - on update set null. 30](#_Toc185578094)

[Figure 28 - delete course using where clause. 30](#_Toc185578095)

[Figure 29 - course id set to null. 30](#_Toc185578096)

[Figure 30 - Adam student. 31](#_Toc185578097)

[Figure 31 - delete student named Adam. 31](#_Toc185578098)

[Figure 32 - Adam has been removed. 32](#_Toc185578099)

[Figure 33 – course names. 33](#_Toc185578100)

[Figure 34 - delete name based on condition. 33](#_Toc185578101)

[Figure 35 - creative course has been removed. 34](#_Toc185578102)

[Figure 36 – student. 35](#_Toc185578103)

[Figure 37 - phone update. 35](#_Toc185578104)

[Figure 38 - update. 35](#_Toc185578105)

[Figure 39 – find first names called Adam. 36](#_Toc185578106)

[Figure 40 – SQL logic to change start date. 37](#_Toc185578107)

[Figure 41 - updated. 37](#_Toc185578108)

[Figure 42 - integrity test. 38](#_Toc185578109)

[Figure 43 - check foreign keys no errors returned. 39](#_Toc185578110)

[Figure 44 - file export table as CSV 42](#_Toc185578111)

[Figure 45 – Courses to be exported as CSV 43](#_Toc185578112)

[Figure 46 – excel import the csv file. 43](#_Toc185578113)

[Figure 47 - power query editor. 44](#_Toc185578114)

[Figure 48 - power query editor results 0](#_Toc185578115)

[Figure 49 - results show all data unique and distinct. 0](#_Toc185578116)

[Figure 50 - in depth results of columns. 0](#_Toc185578117)

[Figure 51 - in depth results of column 0](#_Toc185578118)

[Figure 52 - Top Udemy online courses for students (Kraggle,2023) 0](#_Toc185578119)

[Figure 53 - remove duplicates to help clean data 1](#_Toc185578120)

[Figure 54 - removing duplicates 2](#_Toc185578121)

[Figure 55 - unique data sets 2](#_Toc185578122)

[Figure 56 - duration of course to number of subscribers 3](#_Toc185578123)

[Figure 57 - price to number of subscribers 4](#_Toc185578124)

[Figure 58 - top 10 subscribed courses and price 4](#_Toc185578125)

[Figure 59 - shows the number of subscribers to the course 5](#_Toc185578126)

[Figure 60 - shows the spread of subjects that are most popular 7](#_Toc185578127)

[Figure 61 - mean price of prices 7](#_Toc185578128)

[Figure 62 - standard deviation of price 8](#_Toc185578129)

[Figure 63 - normal distribution 8](#_Toc185578130)

# A diagram of a computer Description automatically generated with medium confidenceER Diagram

Figure ( ER diagram)

The ER Diagram contains 11 tables with data types and the entities inside the tables, It will complete most of the requirements by leaf village, gathered In the requirement phase. The diagram is used to model and design the database, it displays the logic of the data and how it flows and the relationship between tables (Lucidchart, 2024). Primary keys set entity apart and foreign keys define relationships(Belcic and Stryker, 2024).

## Tables and relationships

### Students Table

**Student\_id** is the primary key it auto increments helping keep ID unique, it also contains information such as name and address.

**Foreign keys**

* **Course\_id** - connects to the course table, the relationship is a many to one as students can only be on one course ,but courses have many students assigned to their courses.
* **Balance\_id** - connects to the student fee table , the relationship between the tables is a one to many, as students only have one year cost of the course to pay, but student fee has many students paying their courses.
* **Contacts\_id** – many to many relationships as students can have many contacts and contacts can have many students.
* **Attendance\_id** - attendance is a many to one and only one relationship as students can only have one attendance per student but attendance can have lots of students.
* **Med\_id** – it is a many to one relationship as students can have one medical record, but medical information can have lots of student’s medical information.
* **Cert\_id** – certificate links to students as a many to one relationship, it is not visually showed in the ER diagram, but students can only have one certificate, but certificate table can have many students.
* **Enrolment\_id** – many to one and only one relationship as student can only be enrolled once but enrolment can have many students.

### Enrolment Table

Enrolments table is used to track students start and end date of enrolment and the course onto which they are enrolled.

Enrolment\_id is the primary key which automatically incriminates helping keep new students\_id unique.

**Foreign keys**

* **Student\_id** – this is used to track student id relationship to enrolment id.
* **Course\_id –** this is a many to many as enrolment can have many courses and courses can have many students enrolled.

### Medical Info

Medical info table is used to keep records of student’s medical records it has med\_id as a primary key and has only one foreign key student\_id which links as a one-to-many relationships as one student\_id can only have one medical record.

### Attendance

Atten\_id is the primary key for this table which keeps it unique per student assigned to, this table is used to track attendance for students. Student\_id is the foreign key which has a one-to-many relationship you can have only one and only one attendance per student but lots of student’s attendance.

### Student fee

Student fee has Balance\_id as the primary key used for the table. The table is used to track course fee balance due and balance paid per student with a due date to make sure fee are not paid past deadline with student\_id being the only foreign key, this has a one and only one to many relationship with students only able to have one balance id but balance have lots of students.

### Students contacts

Contacts\_id is the primary key of this table, it is used to store student contracts such as parents used in emergencies such as name and phone number. It has a foreign key to link to student\_id so it can be accessed easily. It has a many to many relationships as you can have many contacts per student and many students per contact.

### Courses

Course table is used to store all the courses that are on offer and the cost, Course\_id is the primary key for this and is unique to each course. It has 2 foreign keys connect to other tables.

**Cert\_id –** this is a one and only one to many relationships as course can only have one certificate, but courses can have many certificates.

**Staff\_id –** this connects to the global\_personal\_id table it has a one-to-one relationship as you can only have one staff member per course and a course can only have one staff member.

### Campus

Campus table is the names of campus courses can be assigned to with campus\_id as its primary key then, it has Course\_id as its foreign key connecting to course table with a many to one relationship with courses having only one campus, but a campus can have many courses.

### Certificates

Certificates table is used to track the certification for each course its primary key is cert\_id, with only one foreign key being course\_id this is a many to one relationship as you can only have one certificate per course, but courses have lots of certificates.

### Global personal ID

This is where all staff information is stored such as role which can be teacher, support or manager. It also has information such as name and phone number. The primary key is staff\_id which automatically incriminates helping keep it unique. It has one foreign key being course\_id which is linked to courses as a one-to-one relationship as each course can only have one and only one teacher assigned, and a teacher can only be assigned one course.

### Role permissions

This table would be used injunction with PHP code or python as a way of completing the requirement of teachers and support staff can only see students on their course. Using the primary key role and the two attributes of can\_view\_all\_students which would be used for managers to view all and then can\_view\_assigned\_students would connect to a PHP code or python to run through a program. This would be needed as SQLite doesn’t support blocking select statements. . Since data security, privacy, and controlled access are given top priority in frameworks like GDPR and ISO 27001, controlling and restricting access to sensitive data in databases is essential for compliance(Freestone, 2023).   
GDPR: Places a strong emphasis on need-to-know access and data minimisation. Sensitive personal information should only be viewed by those with a legitimate reason(GDPR.eu, 2018).  
Implementing controls to guarantee the availability, confidentiality, and integrity of information is the main goal of ISO 27001. One important aspect of preserving secrecy is limiting access(ISO 27001, 2018).

A trigger will check if a SELECT query is allowed based on the role and course ID. SQLite doesn’t natively support "blocking" SELECT statements via triggers, so you'll enforce this using a helper table to log access attempts or within your application code.

A screenshot of a computer program

Description automatically generated

Figure – trigger for logging access attempts.

A white background with colorful text

Description automatically generated

Figure - sample query used by a staff id = 1

If the user is unauthorized, no results will appear, and you can log the attempt as needed.

# Database implementation

The database follows the logic and design of the ER diagram with 11 tables all connected using foreign keys.

A screenshot of a computer

Description automatically generated

Figure ( DB SQLite)

A screenshot of a computer

Description automatically generated

Figure ( DB SQLite)

The data types all match and primary and foreign keys to the ER diagram. The conditions such as UNIQUE keep the data from being duplicated helping with data integrity. The NOT NULL makes sure fields that need to be filled aren’t left blank and would connect to a front end to get error messages if they are (W3schools.com, 2024).

A close-up of a computer code

Description automatically generated

Figure ( students table)

As shown above the primary key student\_id AUTOINCREMENT helping keep all new ids unique. With the foreign keys they have the constraint ON DELETE SET NULL, this is used so if the ID of that foreign key is deleted the value with be updated and set null in the students table, This can help with unexpected errors having empty foreign keys can result in (Mysql.com, 2018).

## Views

### Manager view

A screenshot of a computer

Description automatically generated

Figure - Manager view table query

A screenshot of a computer

Description automatically generated

Figure - Query result from the view

### Teacher view

A computer screen shot of a code

Description automatically generated

Figure - Teacher Query

A screen shot of a computer

Description automatically generated

Figure - Query result

To expand on this further implementation for all courses, each teacher id would be assigned to a course id to show the students in that course based on the teacher’s id assigned.

# Section 1: Code solutions to the following query and reporting problems as SQL statements –

## Show a list of names of all students enrolled on a course, showing student names and course names only.

A screenshot of a computer

Description automatically generated

Figure 11 – list of students and course names.

SQL Select, which is used to choose particular values in tables, is required to get this outcome (W3schools.com, 2024). To take out the students table, first choose the values that need to be achieved and specify which table they originate from, such as "students.course\_id." The data will be obtained from the students' table by using the command From. The matched values in the table will then be combined and shown using an inner join (Tutorialspoint.com, 2024) . In order to acquire the correct course name for the student, the last command ON is used to match the course id in both databases.

## Show ID’s and names of students who have not enrolled on any course.

A screenshot of a computer

Description automatically generated

Figure - no course enrolled.

This is a straightforward select statement with a where clause. First, choose the values that should be retrieved: course\_id, student\_id, and firstn. Next, use from to obtain from the students database, and last, use where clause to determine which course\_id values match NULL (www.techonthenet.com, 2016).

## Show all students who started a course between two specific dates.

A screenshot of a computer

Description automatically generated

Figure - enrolment table.

A screenshot of a computer

Description automatically generated

Figure - select from date.

Selecting all values from the specified table (W3schools.com, 2024)—in this example, the enrolments table—is accomplished by using select \*. To find out what information is being gathered, it will consult where clause. It features a clause that allows you to choose from the Between dates in the start\_date column. In this instance, the data type is date, and the AND is used to filter records based on many conditions. The between operator is used to choose between two values (W3schools.com, 2024).

## Show each course, and the number of students enrolled onto that course.

A screenshot of a computer

Description automatically generated

Figure - show number of students on course.

The number of students who have the same course.id will be counted and retrieved using COUNT(students.course\_id), which is obtained using the SELECT SQL commands to obtain all the specified data with course and course\_id (Tang, 2022). Even if count is 0, left join will pull all student course IDs. From, will choose where the "course" information will originate from. The information is displayed using group by (GeeksforGeeks, 2017).

## Find the teacher with the lowest number of students.

A screenshot of a computer program

Description automatically generated

Figure - teacher with least students.

Using the same logic as last data retrieval follows the similar steps, using ON clause will find where the course id matches the students course id and then can use group by to display the teachers’ names and student count (GeeksforGeeks, 2018).

## Show the total amount of turnover (funds generated) from the sales of course.

A screenshot of a computer

Description automatically generated

Figure - student fee table.

A screenshot of a computer

Description automatically generated

Figure – calculation of sum

The entire sum of a numeric column is returned by the SUM() function (W3schools.com, 2024). The FROM function will be used to select which table to retrieve the data from with the data retrieval answer being £11,850.

## Show all students living in a particular city.

A screenshot of a computer

Description automatically generated

Figure - student table.

A screenshot of a computer

Description automatically generated

Figure – find where city = London

The WHERE clause is used to filter records. This approach only extracts data that satisfy a predefined requirement (W3schools.com, 2024). By using where clause to retrieve the table city where it matches London, this straightforward SQL query will choose all students who reside in London. presenting three students' responses.

## Show all certificates issues between two dates.

A screenshot of a computer

Description automatically generated

Figure - enrolment table.

A screenshot of a computer

Description automatically generated

Figure - course end date.

The And clause will choose the other time period for BETWEEN, while the select \* will choose all entries from enrolments that satisfy where requirement of BETWEEN with data type being date that can be extracted from datetime expression (W3schools.com, 2017). This criterion is only met by one student.

## Show all staff with their staff type (teacher, support or management)

A screenshot of a computer

Description automatically generated

Figure - select teachers and role.

The teacher’s table's staff\_id, first name, and role are retrieved using SELECT in this straightforward SQL query, which shows every employee's table role along with their role.

# Section 2: Code solutions to the following manipulation problems as SQL statements

## Create a new course called ‘Virtual Dance’ and assign 5 new students to the course.

A screenshot of a computer

Description automatically generated

Figure – insert the table and 5 new students.

The first section of the SQL code will insert a course named "Virtual dance" into the courses table with the charge off "3000" and size of "14." INSERT is used when generating new entries to enter into tables.

In the second section of the SQL code, five new students are added to the students table, and the newly established course, course\_id 11, is assigned.

A screenshot of a computer

Description automatically generated

Figure - added into courses.

A screenshot of a computer

Description automatically generated

Figure - added into students.

## Remove a course and all data associated with it from the system.



Figure - on update set null.

A screenshot of a computer

Description automatically generated

Figure - delete course using where clause.

Making use of The DELETE statement will be used to remove the clause that matches the value of healing with the arts (W3schools.com, 2024).



Figure - course id set to null.

## Delete any students with a firstname of ‘Adam’.



Figure - Adam student.

A screenshot of a computer program

Description automatically generated

Figure - delete student named Adam.

Students whose table first name matches "adam" will be chosen using the DELETE command (W3schools.com, 2024).

A screenshot of a computer

Description automatically generated

Figure - Adam has been removed.

## Delete all courses with the word ‘creative’ as part of the course name.

A white text on a white background

Description automatically generated

Figure – course names.

A screenshot of a computer

Description automatically generated

Figure - delete name based on condition.

The records in the `courses` table that include the term "Creative" in the `course` column as part of its value are deleted by this SQL code. The `LIKE` operator is used in the condition with the pattern `'%Creative%'`, which looks for any value that contains "Creative" regardless of where it ap (W3schools.com, 2024)pears in the text (at the start, middle, or end). The wildcard `%` stands for zero or more characters. In the `course` column, for instance, records containing the values "Creative Writing," "Team Creativity," or "Intro to Creative Thinking" would be removed.

A white sheet with black text

Description automatically generated

Figure - creative course has been removed.

## Change the contact number of a Student’s Contact

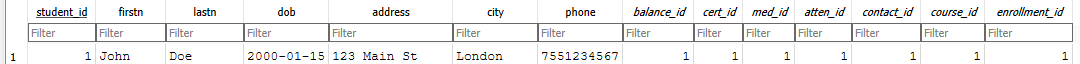


Figure – student.

A screenshot of a computer program

Description automatically generated

Figure - phone update.

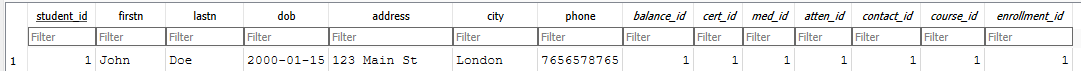


Figure - update.

Set is used to change particular values in columns; in this example, the phone will be updated to the new value set. change is used when changing table records (W3schools.com, 2024). Where is utilised to determine which student's value will be changed (W3schools.com, 2024).

## Change the start date of any students with a forename of ‘Adam’.

A screenshot of a computer

Description automatically generated

Figure – find first names called Adam.

A screenshot of a computer program

Description automatically generated

Figure – SQL logic to change start date.

The start date will be set to "2025-02-13" using the set statement as previously said. Where clause's reasoning is to choose the student ID from the student table whose first name is Adam.



Figure - updated.

# Section 3: Testing, exporting and profiling the data within the database.

## Perform an integrity check on each table within your database.

A screenshot of a computer

Description automatically generated

Figure - integrity test.

For low-level formatting and consistency tests of a database, PRAGMA integrity check looks for:   
• Out-of-sequence table or index entries  
• Records that are misformatted   
• Inaccurate, CHECK, and NOT NULL constraint errors;  
• Missing pages;  
• Excessive or missing index entries;   
• Inaccurate freelist integrity;   
• Database sections that are utilised repeatedly or not at all   
However, PRAGMA foreign key check is used to check for foreign key mistakes, but Pragma does not. No important restrictions are broken, as can be shown below (Sqlite.org, 2017).

A screenshot of a computer

Description automatically generated

Figure - check foreign keys no errors returned.

## Explain what an integrity check covers and why it is important when developing a database and describe other methods of testing a relational database.

Data integrity testing validates the accuracy and reliability of the data stored in database. This is to ensure data is not corrupted or lost during storage or retrieval or processing (Niv Sluzki, 2024) .

Data testing as 4 types.

**Functional testing –** verifying operations such as add ,delete and update.

**Compatibility testing –** checks how data acts around different platforms , operating systems and with other software integrated in.

**Regression testing –** verifying that data remains unaffected after updating or changing data.

**Stress testing** **–** confirm how the system manages under normal or abnormal loads of data.

(Luxequality.com, 2024)

It is important to run through for many major organizations relying on integrity to predict accurately consumer behaviour, market activity and potential data and security risk (Fortinet, 2023).

## Export students into a csv file.



## Import the data from the courses table SQL database into Microsoft Excel and perform profiling using Power Query

A screenshot of a computer

Description automatically generated

Figure 44 - file export table as CSV

A screenshot of a computer

Description automatically generated

Figure 45 – Courses to be exported as CSV

A screenshot of a computer

Description automatically generated

Figure – excel import the csv file.

A screenshot of a computer

Description automatically generated

Figure - power query editor.

A computer screen shot of a black screen

Description automatically generated

Figure - power query editor results

A screenshot of a computer

Description automatically generated

Figure - results show all data unique and distinct.

Figure - in depth results of columns.

A screenshot of a computer

Description automatically generated

Figure - in depth results of column

# Section 4: Statistical Analysis

## Data source

A screenshot of a computer

Description automatically generated

Figure - Top Udemy online courses for students (Kraggle,2023)

One of the leading suppliers of online courses is Udemy, which will benefit Leaf Village by offering a variety of enormous datasets that will assist track future trends and give insight into industry patterns for both its B2C and B2B (Linkedin.com, 2024).

A screenshot of a computer

Description automatically generated

Figure - remove duplicates to help clean data

A screenshot of a phone

Description automatically generated

Figure - removing duplicates

A screenshot of a computer screen

Description automatically generated

Figure - unique data sets

The data set shows a big portion of subjects are business finance with market trends could have a gap in other areas which would help benefit leaf village who provide creative courses.

## Duration of courses to subscribers

A graph with blue dots

Description automatically generated

Figure - duration of course to number of subscribers

Shows the that shorter courses do get the most subscribers with a few anomalies being the 10 min course with 270,000.

To aid with the highest odds of most subscribers, Leaf Village should deliberately plant their courses within the times indicated by the triangle, which symbolises the most populated data for duration, this would help grow the B2C business.

## Subscribers to cost of course

A graph with blue dots

Description automatically generated

Figure - price to number of subscribers

A screenshot of a table

Description automatically generated

Figure - top 10 subscribed courses and price

This illustrates the minimum and maximum pricing condition for Leaf Village. With the price at 0 having the most subscribers and 200 having the second-highest, the graph indicates that the majority of the points are equal. The data dispersion may indicate that Leaf's pricing strategy would be tailored to their population and tendencies. in information. With a market value of £1.5 trillion, digital marketing is growing in popularity for unique courses like Leaf Offer (Nitin Gurmukhani, 2024).

## Sum of subscribers to level of course

A graph of a number of blue and white bars

Description automatically generated with medium confidence

The graph indicates that the majority subscribers finish beginner courses, with relatively few completing expert ones, making beginner courses the most effective strategy to develop for leaf’s B2C.

## Course name to subscribers

A graph with text on it

Description automatically generated

Figure - shows the number of subscribers to the course

With 260,000 users, the trend indicates that learning HTML 5 is the most popular course. With 100,000 fewer students, the second most popular course is Coding for Entrepreneurs Basic.   
Leaf may provide comparable courses to draw in customers and business and increase revenue during subscription losses.

## Lectures vs the Subjects

A graph with blue and white text

Description automatically generated  
Depending on how many users Leaf Village has in comparison to Udemy, this can help with expenses by reducing the number of lecturers they will require for the courses.

## Subject to subscribers

A graph with blue and white bars

Description automatically generated with medium confidence

Figure - shows the spread of subjects that are most popular

According to the data, graphic design and musical instruments are among the least popular courses. Since Leaf Village offers classes in these areas, this might be an issue. The trends in the market demonstrate this. The continued reduction in the popularity of creative subjects is confirmed by the 3% drop in the percentage of arts submissions since 2023 (Arts Professional - News, articles, jobs & events, 2024).

## Mean number of subscribers

A screenshot of a computer

Description automatically generated

Figure - mean price of prices

## Standard devation

A screenshot of a computer

Description automatically generated

Figure - standard deviation of price

The standard deviation for the price is 60.9, standard deviation is set of data points' degree of variance is measured by the standard deviation (Pannell, 2023). The mean price being 66 shows a large standard deviation this mean that they take a higher risk with these prices setting some high and others at much lower, With Udemy being a high market player they can set prices higher with Consolidated Net Dollar Retention Rate (NDRR): 101% in the second quarter of 2024 shows that the strategy is working for them (Earnings Presentation, 2024).

In order to reduce risk while growing, Leaf should concentrate on increasing their market share and maintaining a standard deviation near one (Pannell, 2023).

A close up of a blue and white striped background

Description automatically generated with medium confidence

Figure - normal distribution

# Conclusion

In the past, Leaf Village Training Ltd, a supplier of creative arts education, has managed important data, such as personnel, student, financial, compliance, and management records, using unstructured techniques including papers and spreadsheets. There were key issues with data accessibility, accuracy, and efficiency because of this disorganised data structure. In order to centralise and simplify the organization's data management procedures, The project created and deployed an SQLite database with eleven interconnected tables after seeing the necessity for a scalable and organised solution (www.javatpoint.com, 2024).

The majority of Leaf Village's operational needs are met by the database's seamless integration (Sqlite.org, 2015). Its strong functionality and dependability are demonstrated by the execution of many queries and thorough integrity tests (Niv Sluzki, 2024). By facilitating effective data retrieval and analysis, these procedures guarantee data consistency, increase security, and streamline operational operations (Linkedin.com, 2024). For example, compared to the prior system, personnel may now access and handle comprehensive datasets more efficiently (Newman, C ,2024).

Although the database satisfies most of the organization's requirements, a significant drawback is that SQLite does not enable blocking SELECT queries, which makes it difficult to implement real-time changes in some situations, such as the student viewing interface. A front-end solution that enhances the present system and offers a better user experience while preserving database performance will be needed to address this (Anastasiia Soktoeva, 2023). In order to unlock more features and scalability, future updates could think about connecting with more sophisticated relational database systems.

The recently deployed database system greatly improves operational efficiency and cost-effectiveness when compared to the fragmented and prone to errors spreadsheets that were previously in use(Sqlite.org, 2017). Centralised data storage minimises mistakes and cuts down on redundancies, and management decision-making is enhanced by quicker enquiries. Industry sources claim that implementing a well-designed database system may result in significant cost savings and enhanced productivity, making it a crucial first step for businesses moving towards digital solutions (NI Business Info, 2024).

The social problems associated with digital solutions include equality and accessibility, data security with the political problems of data regulations, funding, and digital divide with technological problems in other countries (Solix, 2023) . Unfortunately, these problems can impede education in developing nations due to technological and skill gaps such as UK According to 27% of UK workers, they do not possess the necessary digital skills for their position (Oxford Learning College, 2023).  
In summary, the SQLite database not only fixes Leaf Village's current data management problems but also establishes a solid framework for expansion and innovation in the future. This solution is a crucial first step in changing the organization's data practices, guaranteeing long-term operational success and a competitive edge in the educational sector, even though there is still opportunity for improvement(Linkedin.com, 2024).

# References

Arts Professional - News, articles, jobs & events. (2024). Entries for arts A-levels down nearly a third since 2010 - Arts Professional. [online] Available at: <https://www.artsprofessional.co.uk/news/entries-arts-levels-down-nearly-third-2010#:~:text=Today's%20results%20reaffirm%20a%20continuing,falling%20by%203%25%20since%202023>. [Accessed 6 Dec. 2024].

Belcic, I. and Stryker, C. (2024). Entity relationship diagram. [online] Ibm.com. Available at: <https://www.ibm.com/think/topics/entity-relationship-diagram> [Accessed 28 Nov. 2024].

Earnings Presentation. (2024). [online] Available at: <https://investors.udemy.com/static-files/284b4ccb-91cd-4de3-8efe-42b9f69c7619#:~:text=range%20for%20Q2.-,%E2%97%8F,NM%20=%20Not%20Meaningful> [Accessed 6 Dec. 2024].

Fortinet. (2023). What Is Data Integrity? Why Is It Important? | Fortinet. [online] Available at: <https://www.fortinet.com/resources/cyberglossary/data-integrity#:~:text=It%20is%20important%20to%20ensure,personally%20identifiable%20information%20(PII)>. [Accessed 22 Nov. 2024].

Freestone, T. (2023). Understanding Key Aspects of Data Compliance. [online] Kiteworks | Your Private Content Network. Available at: <https://www.kiteworks.com/regulatory-compliance/data-compliance/#:~:text=Data%20security%20compliance%20refers%20to,disclosure%20of%20protected%20health%20information>. [Accessed 3 Dec. 2024].

GeeksforGeeks (2017). SQL | GROUP BY. [online] GeeksforGeeks. Available at: <https://www.geeksforgeeks.org/sql-group-by/> [Accessed 2 Dec. 2024].

GeeksforGeeks (2018). SQL | ON Clause. [online] GeeksforGeeks. Available at: <https://www.geeksforgeeks.org/sql-on-clause/> [Accessed 2 Dec. 2024].

GDPR.eu. (2018). What is GDPR, the EU’s new data protection law? - GDPR.eu. [online] Available at: <https://gdpr.eu/what-is-gdpr/> [Accessed 3 Dec. 2024].

ISO 27001 (2018). ISO 27001. [online] Itgovernance.co.uk. Available at: <https://www.itgovernance.co.uk/iso27001#:~:text=ISO/IEC%2027001%20is%20the,(information%20security%20management%20system)>. [Accessed 3 Dec. 2024].

Javatpoint www.javatpoint.com. (2024). SQLite Advantages and Disadvantages - javatpoint. [online] Available at: <https://www.javatpoint.com/sqlite-advantages-and-disadvantages> [Accessed 3 Dec. 2024].

Linkedin.com. (2024). How are databases used to improve education outcomes? [online] Available at: <https://www.linkedin.com/advice/3/how-databases-used-improve-education-outcomes-rx5pf#:~:text=Databases%20in%20education%20enhance%20outcomes,progress%2C%20and%20inform%20targeted%20interventions>. [Accessed 3 Dec. 2024].

Linkedin.com. (2024). What are the key performance indicators for managing an SQLite database? [online] Available at: <https://www.linkedin.com/advice/0/what-key-performance-indicators-managing-sqlite-wfafe#:~:text=To%20enhance%20data%20security%20in,auditing%2C%20logging%2C%20and%20firewalls>. [Accessed 3 Dec. 2024].

Linkedin.com. (2024). LinkedIn. [online] Available at: <https://www.linkedin.com/pulse/online-course-providers-market-share/> [Accessed 6 Dec. 2024].

Lucidchart. (2024). What is an Entity Relationship Diagram (ERD)? [online] Available at: <https://www.lucidchart.com/pages/er-diagrams#:~:text=An%20Entity%20Relationship%20(ER)%20Diagram,each%20other%20within%20a%20system>. [Accessed 28 Nov. 2024].

Luxequality.com. (2024). Data Integrity Testing: Techniques, Tools, and Best Practices. [online] Available at: <https://luxequality.com/blog/data-integrity-testing/> [Accessed 22 Nov. 2024].

Mysql.com. (2018). MySQL :: MySQL 8.4 Reference Manual :: 15.1.20.5 FOREIGN KEY Constraints. [online] Available at: <https://dev.mysql.com/doc/refman/8.4/en/create-table-foreign-keys.html> [Accessed 29 Nov. 2024].

Nibusinessinfo.co.uk. (2024). 5 reasons why your business needs a good database | nibusinessinfo.co.uk. [online] Available at: <https://www.nibusinessinfo.co.uk/content/5-reasons-why-your-business-needs-good-database#:~:text=Streamline%20your%20data%20management&text=A%20strong%20database%20system%20can,as%20your%20decision%2Dmaking%20processes>. [Accessed 3 Dec. 2024].

Nitin Gurmukhani (2024). Top 15 Trending Courses Online: Explore Best Courses in 2025, Job Roles, Benefits of Online Courses,. [online] upGrad blog. Available at: <https://www.upgrad.com/blog/trending-courses-online/> [Accessed 12 Dec. 2024].

Newman, C. (2024). Introduction Welcome to SQLite. [online] O’Reilly Online Learning. Available at: https://learning.oreilly.com/library/view/sqlite/067232685X/pr06.html [Accessed 16 Dec. 2024].

Niv Sluzki (2024). Data integrity testing. [online] Ibm.com. Available at: <https://www.ibm.com/think/topics/data-integrity-testing> [Accessed 3 Dec. 2024].

Oxford Learning College. (2023). Skills Gap Statistics UK 2023 | Oxford Learning College. [online] Available at: <https://www.oxfordcollege.ac/news/skills-gap-statistics-uk/#:~:text=An%20estimated%2020%25%20of%20the,required%20for%20their%20job%20role>. [Accessed 10 Dec. 2024].

Pannell, R. (2023). Demystifying Standard Deviation: A Beginner’s Guide. [online] Leanscape.io. Available at: <https://leanscape.io/demystifying-standard-deviation-a-beginners-guide/#:~:text=If%20there's%20a%20low%20standard,consistency%20or%20predictability%20is%20desired>. [Accessed 6 Dec. 2024].

Sqlite.org. (2015). Appropriate Uses For SQLite. [online] Available at: <https://www.sqlite.org/whentouse.html> [Accessed 3 Dec. 2024].

Sqlite.org. (2017). 35% Faster Than The Filesystem. [online] Available at: <https://www.sqlite.org/fasterthanfs.html#:~:text=Furthermore%2C%20a%20single%20SQLite%20database,the%20blobs%20in%20individual%20files>. [Accessed 3 Dec. 2024].

Sqlite.org. (2017). Pragma statements supported by SQLite. [online] Available at: <https://www.sqlite.org/pragma.html#pragma_foreign_key_check> [Accessed 3 Dec. 2024].

Solix. (2023). What Are the Major Social Issues in Information Technology - Solix. [online] Available at: <https://www.solixbiofuels.com/what-are-the-major-social-issues-in-information-technology/#:~:text=Defining%20social%20issues%20in%20information%20technology&text=Some%20examples%20include%20online%20harassment,and%20exploitation%20of%20digital%20divides>. [Accessed 10 Dec. 2024].

Tang, T. (2022). COUNT() SQL FUNCTION. [online] Datacamp.com. Available at: <https://www.datacamp.com/tutorial/count-sql-function> [Accessed 2 Dec. 2024].

Tutorialspoint.com. (2024). SQL - Inner Join. [online] Available at: <https://www.tutorialspoint.com/sql/sql-inner-joins.htm> [Accessed 29 Nov. 2024].

W3schools.com. (2017). W3Schools.com. [online] Available at: <https://www.w3schools.com/sql/func_mysql_date.asp> [Accessed 2 Dec. 2024].

W3schools.com. (2024). W3Schools.com. [online] Available at: <https://www.w3schools.com/sql/sql_and.asp> [Accessed 2 Dec. 2024].

W3schools.com. (2024). W3Schools.com. [online] Available at: <https://www.w3schools.com/sql/sql_delete.asp> [Accessed 2 Dec. 2024].

W3schools.com. (2024). W3Schools.com. [online] Available at: <https://www.w3schools.com/sql/sql_like.asp> [Accessed 2 Dec. 2024].

W3schools.com. (2024). W3Schools.com. [online] Available at: <https://www.w3schools.com/sql/sql_notnull.asp> [Accessed 29 Nov. 2024].

W3schools.com. (2024). W3Schools.com. [online] Available at: <https://www.w3schools.com/sql/sql_ref_set.asp> [Accessed 2 Dec. 2024].

W3schools.com. (2024). W3Schools.com. [online] Available at: <https://www.w3schools.com/sql/sql_select.asp> [Accessed 29 Nov. 2024].

W3schools.com. (2024). W3Schools.com. [online] Available at: <https://www.w3schools.com/sql/sql_sum.asp> [Accessed 2 Dec. 2024].

W3schools.com. (2024). W3Schools.com. [online] Available at: <https://www.w3schools.com/sql/sql_update.asp> [Accessed 2 Dec. 2024].

W3schools.com. (2024). W3Schools.com. [online] Available at: <https://www.w3schools.com/sql/sql_where.asp> [Accessed 2 Dec. 2024].

‌Javatpoint [www.javatpoint.com](http://www.javatpoint.com). (2024). SQLite Advantages and Disadvantages - javatpoint. [online] Available at: <https://www.javatpoint.com/sqlite-advantages-and-disadvantages> [Accessed 3 Dec. 2024].

[www.techonthenet.com](http://www.techonthenet.com). (2016). SQL: IS NOT NULL Condition. [online] Techonthenet.com. Available at: <https://www.techonthenet.com/sql/is_not_null.php> [Accessed 29 Nov. 2024].

# Bibliography

Database Design Using Entity-Relationship Diagrams, 3rd Edition (2024). *Database Design Using Entity-Relationship Diagrams, 3rd Edition*. [online] O’Reilly Online Learning. Available at: https://learning.oreilly.com/library/view/database-design-using/9781000642704/?sso\_link=yes&sso\_link\_from=lancaster-university [Accessed 16 Dec. 2024].

Drkusic, E. (2020). *Learn SQL: SQL Triggers*. [online] SQL Shack - articles about database auditing, server performance, data recovery, and more. Available at: https://www.sqlshack.com/learn-sql-sql-triggers/ [Accessed 16 Dec. 2024].

GeeksforGeeks (2024). *SQL SELECT WHERE Field Contains Words*. [online] GeeksforGeeks. Available at: https://www.geeksforgeeks.org/sql-select-where-field-contains-words/ [Accessed 16 Dec. 2024].

General Data Protection Regulation (GDPR). (2024). *General Data Protection Regulation (GDPR) – Final text neatly arranged*. [online] Available at: https://gdpr-info.eu/ [Accessed 16 Dec. 2024].

How (2022). *How to use Power Query - Microsoft Excel Tutorial*. [online] YouTube. Available at: https://www.youtube.com/watch?v=MHIV0bYryiw [Accessed 16 Dec. 2024].

Jeremiah, O. (2024). *SQL Triggers: A Guide for Developers*. [online] Datacamp.com. Available at: https://www.datacamp.com/tutorial/sql-triggers [Accessed 16 Dec. 2024].

Kaggle.com. (2024). *Kaggle: Your Home for Data Science*. [online] Available at: https://www.kaggle.com/ [Accessed 16 Dec. 2024].

Loshin, P. and Steele, C. (2021). *ISO (International Organization for Standardization)*. [online] Search Data Center. Available at: https://www.techtarget.com/searchdatacenter/definition/ISO#:~:text=ISO%20(International%20Organization%20for%20Standardization)%20is%20a%20worldwide%20federation%20of,body%20representing%20each%20member%20country. [Accessed 16 Dec. 2024].

Meneghello, F., Chiara Di Francescomarino, Chiara Ghidini and Massimiliano Ronzani (2024a). Runtime integration of machine learning and simulation for business processes: Time and decision mining predictions. *Information Systems*, [online] 128, pp.102472–102472. doi:https://doi.org/10.1016/j.is.2024.102472.

Meneghello, F., Chiara Di Francescomarino, Chiara Ghidini and Massimiliano Ronzani (2024b). Runtime integration of machine learning and simulation for business processes: Time and decision mining predictions. *Information Systems*, [online] 128, pp.102472–102472. doi:https://doi.org/10.1016/j.is.2024.102472.

Newman, C. (2024). *Introduction Welcome to SQLite*. [online] O’Reilly Online Learning. Available at: https://learning.oreilly.com/library/view/sqlite/067232685X/pr06.html [Accessed 16 Dec. 2024].

Proquest.com. (2024a). *ProQuest Ebook Central - Book Details*. [online] Available at: https://ebookcentral.proquest.com/lib/lancaster/detail.action?pq-origsite=primo&docID=5986738 [Accessed 16 Dec. 2024].

Proquest.com. (2024b). *ProQuest Ebook Central - Reader*. [online] Available at: https://ebookcentral.proquest.com/lib/lancaster/reader.action?docID=588100&ppg=163&pq-origsite=primo [Accessed 16 Dec. 2024].

Proquest.com. (2024c). *ProQuest Ebook Central - Reader*. [online] Available at: https://ebookcentral.proquest.com/lib/lancaster/reader.action?pq-origsite=primo&ppg=17&docID=2028673 [Accessed 16 Dec. 2024].

Salesforce. (2022). *Salesforce*. [online] Available at: https://www.salesforce.com/uk/crm/what-is-crm/ [Accessed 16 Dec. 2024].

Schneider, J. and Smalley, I. (2024). *Transaction management*. [online] Ibm.com. Available at: https://www.ibm.com/topics/transaction-management [Accessed 16 Dec. 2024].

Sqlite.org. (2015). *CREATE VIEW*. [online] Available at: https://www.sqlite.org/lang\_createview.html [Accessed 16 Dec. 2024].

Sqlite.org. (2024). *SQLite Home Page*. [online] Available at: https://www.sqlite.org/ [Accessed 16 Dec. 2024].

Sqlitebrowser.org. (2024a). *DB Browser for SQLite*. [online] Available at: https://sqlitebrowser.org/ [Accessed 16 Dec. 2024].

Sqlitebrowser.org. (2024b). *DB Browser for SQLite*. [online] Available at: https://sqlitebrowser.org/ [Accessed 16 Dec. 2024].

Staff, C. (2024). *What Is Management Information Systems (MIS)? Your Career Guide*. [online] Coursera. Available at: https://www.coursera.org/articles/management-information-system [Accessed 16 Dec. 2024].

Team, W. (2019). *Feasibility Study Examples*. [online] Wallstreetmojo.com. Available at: https://www.wallstreetmojo.com/feasibility-study-examples/ [Accessed 16 Dec. 2024].

Tutorialspoint.com. (2024). *SQLite - Syntax*. [online] Available at: https://www.tutorialspoint.com/sqlite/sqlite\_syntax.htm [Accessed 16 Dec. 2024].

W3schools.com. (2024). *W3Schools.com*. [online] Available at: https://www.w3schools.com/sql/sql\_where.asp [Accessed 16 Dec. 2024].

Wilkins, J. (2021). *SQL Count – How to Select, Sum, and Average Rows in SQL*. [online] freeCodeCamp.org. Available at: https://www.freecodecamp.org/news/sql-count-how-to-select-sum-and-average-rows-in-sql/#:~:text=If%20you%20need%20to%20add,get%20data%20from%20the%20table. [Accessed 16 Dec. 2024].