

**DATABASES**

**ASSESSMENT 2:  Web application with supporting documentation**

**Abstract:**

This report delineates the meticulous process of designing and implementing a normalized database for historical figures within STEM (Science, Technology, Engineering, and Mathematics), alongside a robust system for managing bookable events featuring talks on these luminaries. Beginning with the normalization of the database, the report elucidates the creation of tables, insertion of data, and the development of stored procedures. Subsequently, it delves into the construction of the event management system, detailing the requisite features and their implementation. Furthermore, the report presents a comprehensive array of SQL queries aimed at extracting pertinent information from the database.

**Purpose of the Report:**

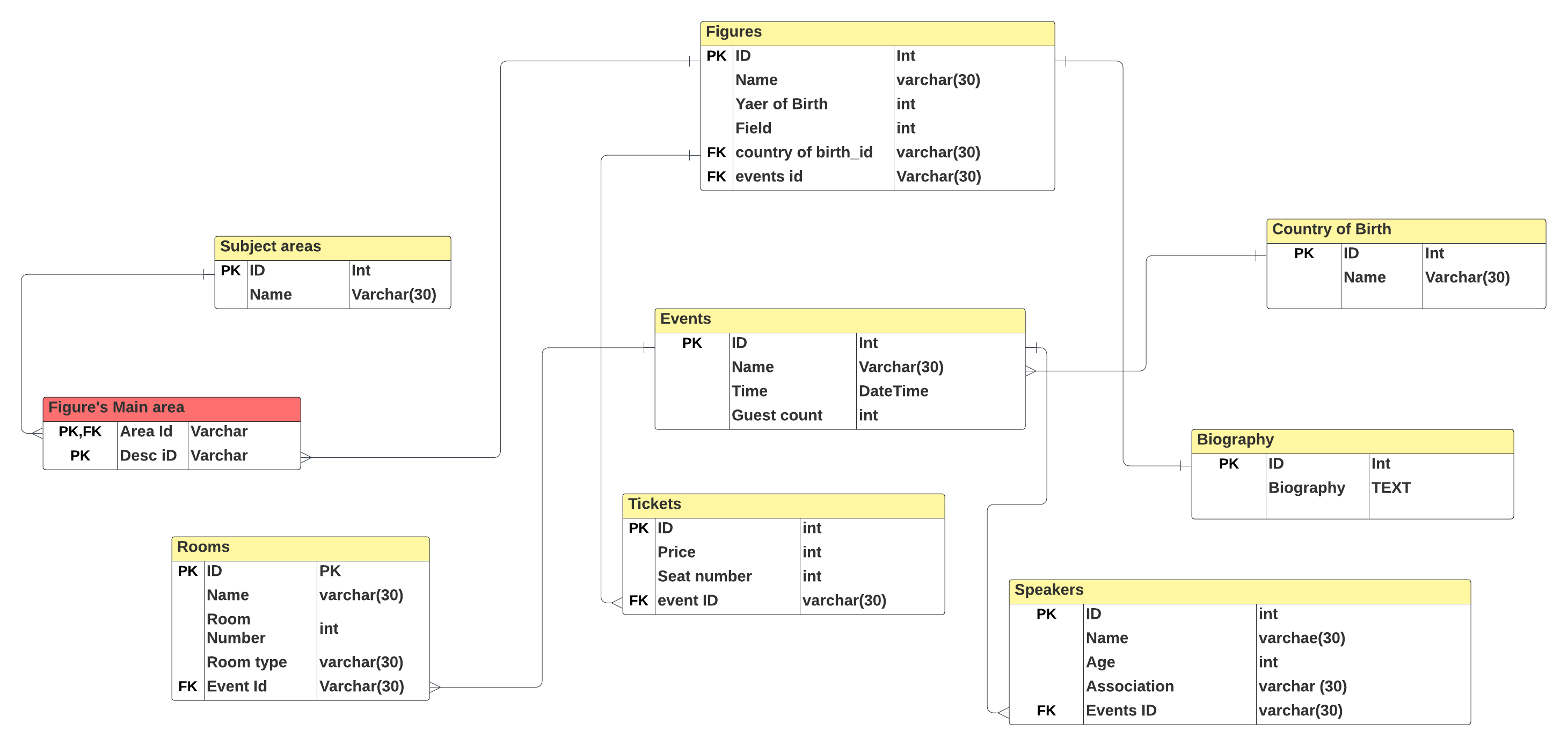
The primary objective of this report is to document the methodology employed in constructing a normalized database for historical figures in STEM and designing an event management system tailored to the domain. By elucidating the steps involved in database creation, data insertion, and system implementation, the report seeks to provide a clear and replicable framework for similar projects in academia and industry. Furthermore, the report aims to underscore the importance of preserving and celebrating the contributions of historical figures within STEM fields, fostering a deeper appreciation for their enduring legacy.

**Introduction:**

Historical figures in STEM have laid foundation for modern advancements. This project aims to preserve and organize their biographical information through comprehensive, normalized relational database adhering to Third Normal Form (3NF) to ensure data integrity and minimize redundancy.

The project also includes a system for managing bookable events, such as lectures and talks about these historical figures. This requires handling intricate relationship between and events, speakers, rooms, and historical figures, ensuring efficient and accessible management of all relevant information.

**ER Diagram:**

****

**Implementation of Database System Historical Figures**

**Historical Figures Database**

The database system for historical figures is meticulously designed to store comprehensive biographical and professional details of eminent scientists across various disciplines. The design process commenced with the identification of key entities: Scientists, Fields, Countries, and Scientist Fields, ensuring a normalized structure conforming to the principles of the Third Normal Form (3NF). Each scientist’s record encompasses their name, birth and death years, country of origin, and their contributions. The Fields entity categorizes domains of expertise (e.g., mathematics, physics, biology), while the countries entity records countries of birth. The scientist fields table captures the many-to-many relationships between scientists and their respective fields of expertise.

This normalized structure minimizes redundancy and enhances data integrity. By segregating data into these entities, we maintain a streamlined plan that can effortlessly scale with the addition of more historical figures. Each table is indexed appropriately to optimize query performance, particularly for join operations between theses related entities.

**Implementation of Extension for Events**

**Booking Lectures and Events**

To extend the databases for booking events and lectures, additional entities were introduced: Events, Speakers, Rooms, Event Topics, and Event Rooms. This extension enables the system to manage events cantered around historical figures, speakers, and manage room bookings.

* **Events:** This table captures the details of each event, Including the event name, date, and ticket requirement status.
* **Speakers:** This entity stores information about individuals who will speak at the events.
* **Rooms:** Rooms are classified into lecture room and event rooms, with each room’s details stored in this table.
* **Event Topics:** This many-to-many relationship table links events to the scientists who are the topics of discussion.
* **Event Rooms:** This table manages the many-to-many relationship between events and the rooms booked for hosting them.

This structure allows the system to efficiently schedule events, assign speakers, and allocate rooms, ensuring each event can have multiple speakers and topics, and can utilize various rooms as needed.

**SQL Queries:**

1. List all historical figures that were born in France

* **SELECT Name FROM Scientists WHERE CountryID = (SELECT CountryID FROM Countries WHERE CountryName = 'France');**

1. List all historical figures that contributed to Computing

* **SELECT s.Name FROM Scientists s JOIN Scientist\_Fields sf ON s.ScientistID = sf.ScientistID JOIN Fields f ON sf.FieldID = f.FieldID WHERE f.FieldName = 'Computing';**

1. List all historical figures that contributed to more than 1 field

* **SELECT s.Name FROM Scientists s JOIN Scientist\_Fields sf ON s.ScientistID = sf.ScientistID GROUP BY s.Name HAVING COUNT(sf.FieldID) > 1;**

1. Mary Jackson’s death is incorrectly listed. Update this date to 2005

* **UPDATE Scientists SET DeathYear = 2005 WHERE Name = 'Mary Jackson';**

1. List all historical figures who died before they were 40

* **SELECT Name FROM Scientists WHERE (DeathYear - BirthYear) < 40;**

1. Which countries have the most historical figures listed

* **SELECT c.CountryName, COUNT(s.ScientistID) AS ScientistCount FROM Scientists s JOIN Countries c ON s.CountryID = c.CountryID GROUP BY c.CountryName ORDER BY ScientistCount DESC;**

1. Are any historical figures still alive? If so who

* **SELECT Name FROM Scientists WHERE DeathYear IS NULL;**

1. Which historical figures worked for NASA

* SELECT Name FROM Scientists WHERE Description LIKE '%NASA%';

1. Which historical figures worked with Einstein?

* **SELECT Name FROM Scientists WHERE Description LIKE '%Einstein%';**

1. Which countries have the most fields of research contributions

* **SELECT c.CountryName, COUNT(DISTINCT sf.FieldID) AS FieldCount FROM Scientists s JOIN Countries c ON s.CountryID = c.CountryID JOIN Scientist\_Fields sf ON s.ScientistID = sf.ScientistID GROUP BY c.CountryName ORDER BY FieldCount DESC;**

**Reflection on the process and Changes Made**

**Initial design and iterative Refinement**

The initial design focused on capturing comprehensive biographical details of the scientists. It became evident during the process during the process that storing fields of expertise and countries of birth separately would enhance data integrity and facilitate better querying capabilities. This led to the creation of the Fields and Countries tables and the intermediary Scientist Fields table to manage the many-to-many relationships.

For the event booking, the design aimed to be flexible enough to accommodate various event types and configurations. This required introducing multiple new entities and ensuring their relationship were correctly established.

**Changes made in implementation.**

During implementation, the need for additional indexes to improve query performance was identified, especially for join operations between Events, Speakers, Rooms, and Scientists tables. Constraints were added to ensure data integrity.

**Reflection on the role of Historical Figures and Data Expansion**

**Coverage and Missing Figures**

The database currently includes a diverse range of historical figures from various fields and countries, emphasizing those who have made significant contributions to science and technology. However, there are notable gaps, particularly in representing scientists from non-western countries and fields that are underrepresented such as Economics, history etc.

**Expanding the data**

To provide a more comprehensive view, the database should be expanded to include figures from these underrepresented areas. This can be achieved by:

* Identifying and Researching Additional Historical Figures
* Collaborating with Historians and Subject Matter Experts
* Enhancing Accessibility and Engagement

**NOSQL**

Creating a MongoDB database for historical figures involves structing data into various collections, each representing different entities such as scientists, countries, and fields of expertise, each document within this collection encapsulates comprehensive details about a single historical figure, including their name, birth and death years, country of birth, and notable contributions.

1. List all historical figures that were born in France

* **db.scientists.find({ country: "France" });**

1. List all historical figures that contributed to Computing

* **db.scientists.find({ fields: "Computing" });**

1. List all historical figures that contributed to more than 1 field

* **db.scientists.find({ $where: "this.fields.length > 1" });**

1. Mary Jackson’s death is incorrectly listed. Update this date to 2005

* **db.scientists.updateOne({ name: "Mary Jackson" },{ $set: { death\_year: 2005 } });**

1. List all historical figures who died before they were 40

* **db.scientists.find({ $expr: { $lt: ["$death\_year", { $add: ["$birth\_year", 40] }] } });**

1. Which countries have the most historical figures listed

* **db.scientists.aggregate([{ $group: { \_id: "$country", count: { $sum: 1 } } },{ $sort: { count: -1 } },{ $limit: 5 }]);**

1. Are any historical figures still alive? If so who

* **db.scientists.find({ death\_year: { $exists: false } });**

1. Which historical figures worked for NASA

* **db.scientists.find({ contributions: /NASA/i });**

1. Which historical figures worked with Einstein?

* **db.scientists.find({ contributions: /Einstein/i });**

1. Which countries have the most fields of research contributions

* **db.scientists.aggregate([{ $unwind: "$fields" },{ $group: { \_id: "$country", fields\_count: { $sum: 1 } } },{ $sort: { fields\_count: -1 } }]);**

**Conclusion**

Creating a normalized database for historical figures and managing events has provided a robust system for storing and organizing biographical and event-related data. The process involved meticulous design to adhere to normalization principles, ensuring minimal to no redundancy and data integrity.

The implementation underscored the importance of structured data relationships, particularly in handling complex many-to-many relationships. The booking and event management capabilities demonstrated the database’s flexibility to accommodate various configurations and requirements.

In conclusion, this assessment provides a robust system for managing historical data and highlighted areas for future growth and improvement. By adding more data and refining the database, we can create a more comprehensive system that honours the contributions of a diverse range of historical figures.