Big Data

Three characteristics

- **Volume**
- **Velocity**
- **Variety**

Terms like value (the potential worth of the data) and veracity (whether the data is trustworthy) have also been added to describe big data.

Big Data

Sources of Big Data include:

- Media (e.g., social media)
- Cloud data
- Web pages
- Internet of Things (IoT)
- Traditional databases

To harness the value of big data, organizations must:

- Integrate data.
- Store it.
- Analyse it.

ERD

using

Unified Modelling Language (UML) Notation

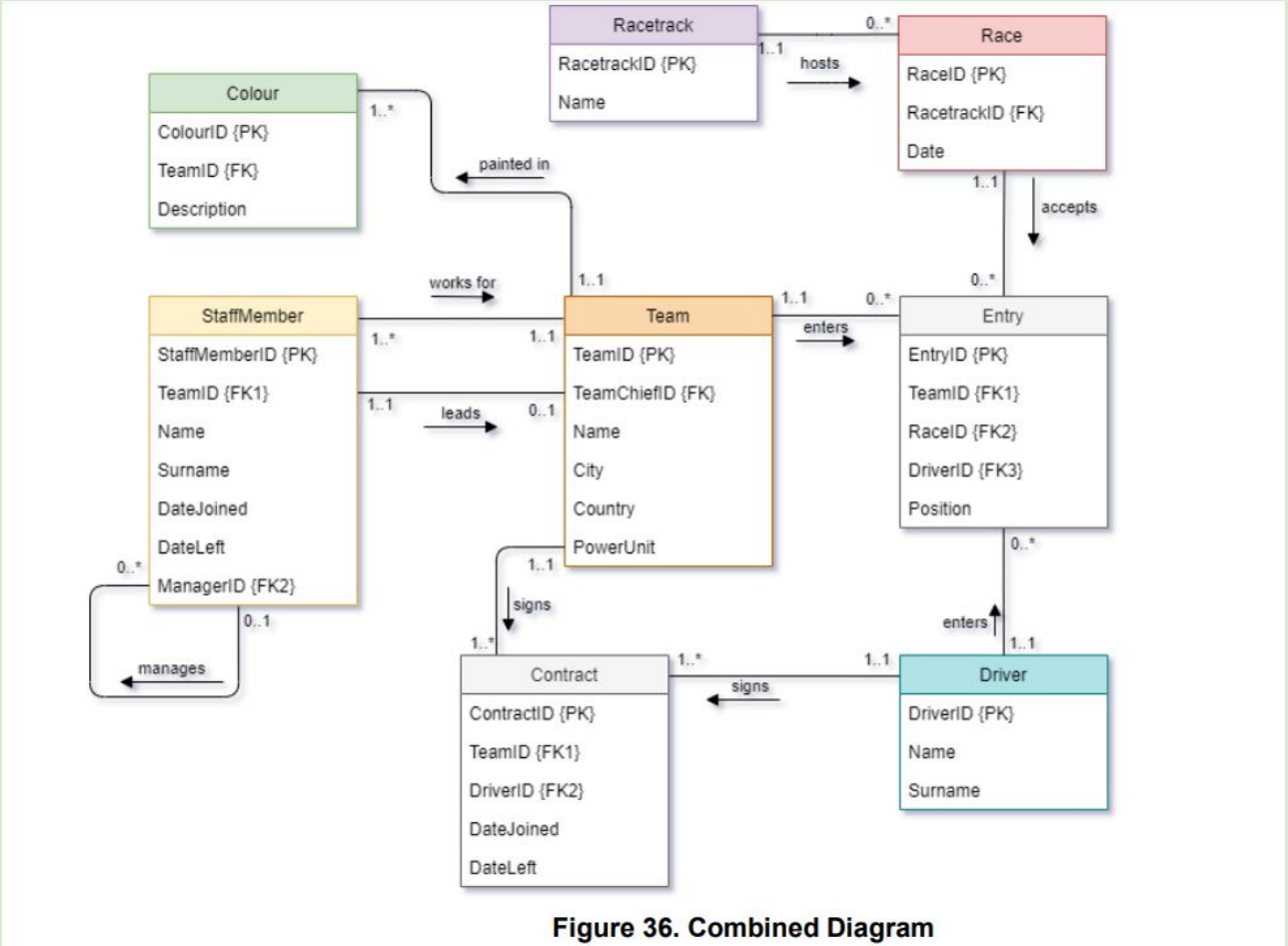

Figure 36. Combined Diagram

ERD

- Business Rule 1 -When a team enters, the following information needs to be provided: team name, location of where the team is based, the power unit used by the team and the team colours.

- Business Rule 2 - A racetrack hosts multiple races over time. A racetrack can be stored even before the first race is scheduled to take place there. Each race takes place at a specific racetrack, on a specific date.

- Business Rule 3 - Each team has several staff members (that are not drivers). A staff member is only allowed to ever work for one team. The name and surname of each staff member needs to be stored in the database, as well as the date that they joined and (optionally) the date that they left the team. One of the staff members is the team chief.

ERD

- **Business Rule 4 - Each team can have two drivers at a time. Drivers may decide to change teams, and the history of that needs to be stored in the database. The date joined that a driver joined a team and (optionally) the date that they left needs to be stored. The name and surname of each driver needs to be recorded.**

- **Business Rule 5 - One staff member is designated as the crew chief. The crew chief manages other staff members that work in the pit crew.**

- **Business Rule 6 - A driver participates in a specific race, racing for a specific team. The position that the driver places in that race needs to be recorded in the database. A driver can compete in many races over time.**

Activity:

The university is looking to develop an online movie rental system that allows customers to browse and rent movies, track their rental history, and make payments for rented movies. The system will also allow the admin to manage movies, customers, and rental transactions.

Business Rules:

- **A Customer can place multiple Rentals, but each Rental is linked to only one Customer.**
- **A Rental can consist of multiple Rental Items, and each Rental Item references a specific Movie.**
- **A Movie can appear in multiple Rental Items (for different customers and rentals).**
- **A Rental will have a total amount calculated based on the price of the rented movies and the quantity of movies rented.**
- **A Payment is associated with a Rental, and the total payment amount must match the rental total amount.**
- **A Customer can have multiple Payments for the same or different rentals.**

ERD - Visual representation of data and its relationships within a database system. It is used in database design to model entities (tables), their attributes (columns), and relationships (connections between tables).

Key Components of an ERD

Entities – Represent objects or concepts in the database.

Attributes – Represent the properties of an entity.

Primary Key (PK) – A unique identifier for an entity.

Foreign Key (FK) – An attribute that establishes a relationship between two entities.

Relationships – Define how entities interact.

Notation Types

- Barker's Notation
- Chen Notation
- IDEF1X Notation
- Arrow Notation
- UML Notation
- Crow's Foot Notation

ERD Process:

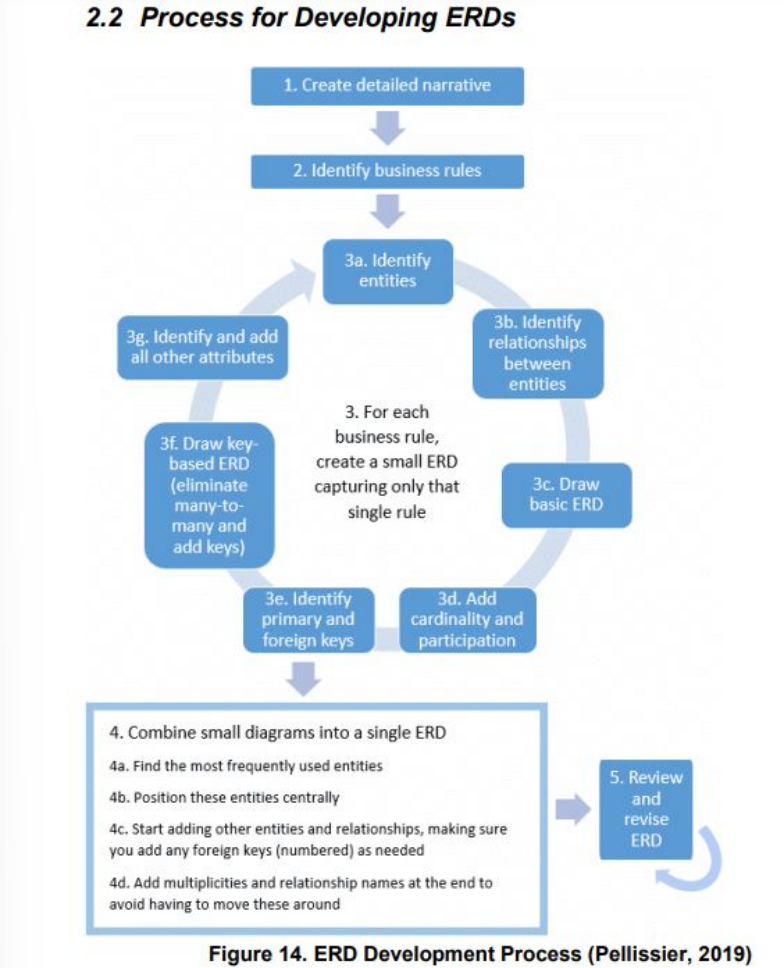

Figure 14. ERD Development Process (Pellissier, 2019)

- ## Unary Relationship (Recursive Relationship)

This occurs when an entity is related to itself.

Example: Employees in a company. One employee (a manager) supervises other employees. The entity "Employee" is related to itself through the "manages" relationship.

- ## Ternary Relationship

This involves three different entities that must all be considered together.

Example: A Supplier supplies a Part for a Project. The relationship only makes sense when all three entities are included. If any one is missing, the relationship is incomplete.

- ## Quaternary Relationship

This involves four different entities in a single relationship.

Example: A Doctor treats a Patient for a Disease using a Medication. All four entities are necessary to fully describe the relationship.

These relationships are rare in database design.

Activity:

You have been hired as a database designer for a new book review website. The website allows users to review books they have read. Users can manage their favorite books and keep track of their reading progress. Your task is to design an ERD based on the following business rules.

Business Rules:

1. **All entities should have surrogate primary keys.**
2. **Each user of the website must have their first name and last name stored in the database.**
3. **Users can add books to their personal reading list. A user can have multiple books in their list, but each book is specific to the user that added it.**
4. **Each book in a user's list must have a title, an author, and a publication year.**
5. **Users can write reviews for books they have read. A review must include a rating, an optional text description, and the date the review was submitted.**
6. **A book can have multiple reviews, but each review is written by a specific user for a specific book.**
7. **The system should store genres for books. Each book can belong to only one genre.**