Assignment 2

Task 1: Relational Model I

1. **What is a key attribute?**

   An attribute set that uniquely identifies all tuples and is minimal. Minimal means that we can't take a subset of this attribute set that also identifies all tuples uniquely.

2. **Give an example relation of degree/arity 4.**

   `Student`($\underline{\text{Matr.Nr.}}$, FirstName, LastName, StudyProgram)

3. **Give an example of a non-recursive 1-to-N relationship.**

   `Caretaker`($\underline{\text{SSN}}$, Name)
   `Panda`($\underline{\text{Id}}$, Name, Birthdate, MainCaretakerSSN)

   Each Panda has one unique Main Caretaker. But a caretaker takes care of multiple pandas (or other animals).

4. **Give an example of a recursive 1-to-N relationship.**

   `Animal`($\underline{\text{Id}}$, Species, eatenFromId)

Task 2: Relational Model II

1. Create a relational model that corresponds to the Entity Relationship model given in Figure 1.

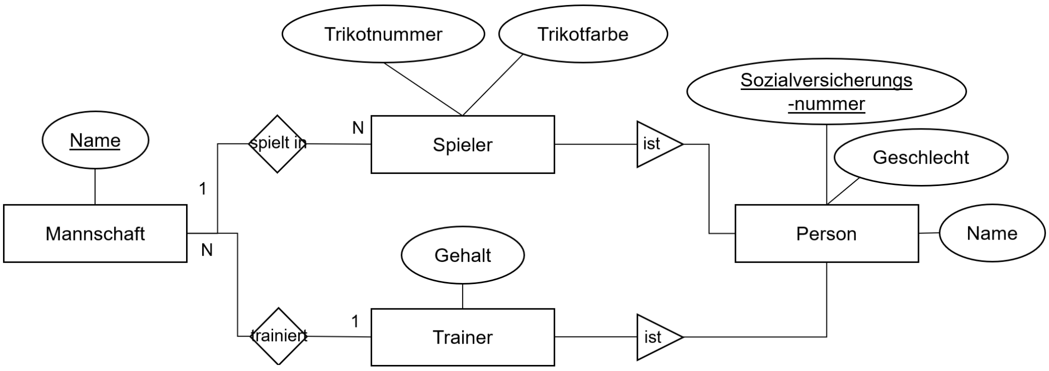

Figure 1: Entity Relationship model of a Soccer Team.

`Person`$(\underline{SSN}, Name, Geschlecht)$

`Spieler`$(\underline{SSN}, Trikotnummer, Trikotfarbe)$

`Trainer`$(\underline{SSN}, Gehalt)$

`Mannschaft`$(\underline{Name}, TrainerSSN)$

`spieltIn`$(\underline{SpielerSSN}, MName)$

2. Create a relational model that corresponds to the Entity Relationship model given in Figure 2.

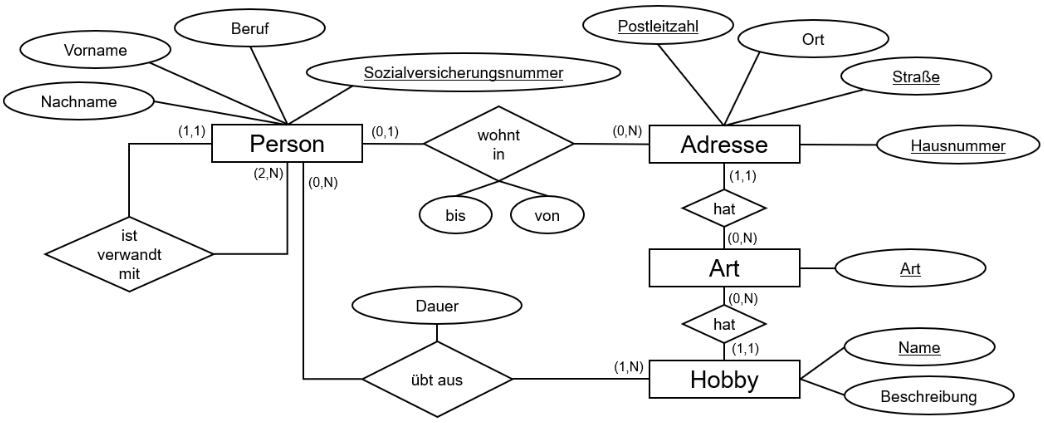

Figure 2: Entity Relationship model of a Person with Hobbies.

$$\texttt{Person}(\underline{SSN}, Beruf, Vorname, Nachname)$$
$$\texttt{Art}(\underline{Art})$$
$$\texttt{Hobby}(\underline{Name}, Beschreibung, \dashuline{Art})$$
$$\texttt{Adresse}(\underline{PLZ, Str., HNr.}, Ort, \dashuline{Art})$$
$$\texttt{wohntIn}(\underline{\dashuline{SSN}}, \dashuline{PLZ, Str., HNr.}, bis, von)$$
$$\texttt{uebtAus}(\underline{\dashuline{HName}, \dashuline{SSN}}, Dauer)$$
$$\texttt{istVerwandtMit}(\underline{\dashuline{SSN}}, \dashuline{SSN})$$

Relational Schema for Tasks 3, 4, 5, 6

$$\texttt{AIRLINE}(\underline{ID},\ Name)$$
$$\texttt{AIRPLANE}(\underline{ID},\ Name,\ Type,\ PassengerCapacity,\ Range)$$
$$\texttt{PASSENGER}(\underline{ID},\ FirstName,\ LastName,\ Age,\ Nationality,\ CreditCardNumber)$$
$$\texttt{CHARTER}(\underline{\dashuline{AirlineID}, \dashuline{AirplaneID}, FromDate, ToDate})$$
$$\texttt{FLIGHT}(\underline{\dashuline{AirlineID}, \dashuline{PassengerID}, \dashuline{Date}})$$
$$\texttt{WEATHER}(\underline{Date},\ Temp,\ PrecAmount,\ SunDuration)$$

Task 3: Relational Algebra I

Consider the given relational model on page 2. Express the queries below in the relational algebra.

1. Give the first name of all passengers whose last name is ”Johansen”.

$$\Pi_{FirstName}(\sigma_{LastName='Johansen'}(Passenger))$$

2. Give the credit card number of all passengers who booked a flight with ”Lufthansa” for ”01.01.2020”.

$$\begin{aligned}&specialFlights \leftarrow ((\sigma_{Date='01.01.2020'}(Flight)) \bowtie_{AirlineID=ID} (\sigma_{Name='Lufthansa'}(Airline))\\&\Pi_{CreditCardNumber}(Passenger \bowtie_{Passenger.ID=PassengerID} specialFlights)\end{aligned}$$

3. Give the first and last name of all passengers who booked a flight for April 2022.

$$\begin{aligned}&FlightsInApril \leftarrow \sigma_{Date<01.05.2020 \wedge Date>31.03.2020}(Flight)\\&\Pi_{FirstName,Lastname}(Passenger \bowtie_{ID=PassengerID} FlightsInApril)\end{aligned}$$

4. Give the name of all airlines that carried passengers on days when the temperature exceeded 30 degrees Celsius and the rainfall was less than 20 liters per hour and the sunshine lasted at least 8 hours. *Note: Assume that temperature is stored in degrees Celsius, rainfall in liters per hour, and sunshine duration in hours.*

$$\begin{aligned}&goodWeather \leftarrow \sigma_{Temp>30 \wedge PrecAmount<20 \wedge SunDuration \geq 8}(Weather)\\&\Pi_{Name}(Airline \bowtie_{ID=AirlineID} (Flight \bowtie goodWeather))\end{aligned}$$

Task 4: Relational Algebra II

Consider the given relational model on page 2. Express the queries below in the relational algebra.

1. Give the ID of all passengers who have never flown (*Passenger Customer*).

$$\Pi_{ID}(Passenger) - \Pi_{PassengerID}(Flight)$$

2. Give the age of all passengers who have never flown with "Lufthansa" on days below 10 degrees Celsius.

$$\begin{aligned}&L10Flights \leftarrow (\sigma_{Name='Lufthansa'}(Airline)) \rtimes_{ID=AirlineID} Flight \ltimes (\sigma_{Temp<10}(Weather))\\&\Pi_{Age}(Passenger - (Passenger \ltimes_{ID=PassengerID} L10Flights))\end{aligned}$$

Task 5: Relational Algebra III

Consider the given relational model on page 3. Which of the queries below can be expressed in relational algebra? In each case, either write "Cannot be expressed in relational algebra." or write the corresponding expression in relational algebra. (5 P each)

1. Give the nationality of the passengers who flew the most.

   cannot be expressed in relational algebra

2. Give the name of all aircraft that never flew in 2010.

   Assuming that flown means has been chartered:

$$\begin{aligned}&2010Charters \leftarrow \sigma_{FromDate<01.01.2011 \wedge ToDate>31.12.2009 \wedge FromDate \leq ToDate}(Charter)\\&\Pi_{Name}(Airplane - (Airplane \ltimes_{ID=AirplaneID} 2010Charters))\end{aligned}$$

3. Give the number of airlines that never had passengers older than 50 years.

   cannot be expressed in relational algebra

Task 6: SQL - Basics

Consider the given relational model on page 3. Give the following queries in SQL.

1. Give the names of all the airlines.

```sql
SELECT Name
FROM Airline;
```

2. Give the first and last names of all passengers who are younger than 30 years old.

```sql
SELECT FirstName, LastName
FROM Passenger
WHERE Age < 30;
```

3. Give the first name and last name of all passengers who flew on days with more than 20 °C.

```sql
SELECT P.FirstName, P.LastName
FROM Passenger P, Weather W, Flight F
WHERE P.ID = F.PassengerID
AND F.Date = W.Date
AND W.Temp > 20;
```

4. Give the number of all passengers who flew with “British Airways” in 2020.

```sql
SELECT COUNT(PassengerID)
FROM Flight
JOIN Airline ON ID=AirlineID
WHERE Name = "British Airways"
AND Date BETWEEN 01.01.2020 AND 31.12.2020;
```

5. Give the names of all airlines that have “Airline” in their name (use LIKE).

```sql
SELECT Name
FROM Airline
WHERE Name LIKE "%Airline%";
```

6. Give the names of all airlines that chartered at least one “Airbus A380” in 2020.

```sql
SELECT Name
FROM Airline
JOIN Charter ON Id=AirlineID
WHERE FromDate < 01.01.2011
AND ToDate > 31.12.2019
AND AirplaneID IN (SELECT Id
                    FROM Airplane
                    WHERE Type = "Airbus A380");
```

7. Give the names of the top 5 airlines with the most passengers in 2020.

```sql
SELECT Name
FROM Airline
WHERE Id IN (SELECT AirlineID
            FROM Flight
            WHERE YEAR(Date) = 2020
            GROUP BY AirlineID
            ORDER BY COUNT(PassagierID) DESC
            LIMIT 5);
```

Company scheme for Tasks 7, 8

$\texttt{EMPLOYEE}(FName, LName, \underline{SSN}, BDate, Address, Salary, \dashuline{BossSSN}, \dashuline{NumDept})$

$\texttt{DEPENDENT}(\underline{Name}, Gender, BDate, Relation, \dashuline{ESSN})$

$\texttt{DEPARTMENT}(DName, \underline{DNumber}, \dashuline{MGRSSN}, MGRStartDate)$

$\texttt{DEPTLOCATION}(\underline{DNumber}, Location)$

$\texttt{PROJECT}(PName, \underline{PNumber}, Location, \dashuline{DNumber})$

$\texttt{WORKSON}(\underline{\dashuline{ESSN}, \dashuline{PNO}}, Hours)$

Task 7: SQL - Joins & Nested Queries

Consider the given relational model on a company. Provide the SQL statements below and, where requested, a brief description of the join type.

1. First + last names of all employees using an Inner Join. Describe the characteristics and purpose of an Inner Join.

```sql
SELECT FName, LName
FROM Employee;
```

In der Ergebnismenge des inner joins sind nur Zeilen enthalten, die sowohl in der linken als auch in der rechten Tabelle Einträge haben, die mit der Joinbedingung übereinstimmen.

2. Names of all departments with their managers' SSNs using a Left Join. Describe the characteristics and purpose of a Left Join.

```sql
SELECT DName, MGRSSN
FROM Department;
```

In der Ergebnismenge des left (outer) joins sind alle Zeilen der linken Tabelle enthalten und die dazu passenden Einträge der rechten Tabelle.

3. Names and genders of all dependents using a Right Join. Describe the characteristics and purpose of a Right Join.

```sql
SELECT Name, Gender
FROM Dependent;
```

In der Ergebnismenge des right (outer) joins sind alle Zeilen der rechten Tabelle enthalten und die dazu passenden Einträge der linken Tabelle.

4. First + last names of employees with their department names using a Full Join. Describe the characteristics and purpose of a Full Join.

```sql
SELECT FName, LName, DName
FROM Employee
FULL JOIN Department ON NumDept = DNumber;
```

Alle Einträge aus beiden Tabellen enthalten.

5. First + last names of all employees who work in at least one project with John Smith, using nested queries (no JOIN).

```sql
SELECT FName, LName
FROM Employee
WHERE Ssn IN (SELECT Essn
                FROM WorksOn
                WHERE Pno IN (SELECT Pno
                               FROM WorksOn
                                WHERE Essn IN (SELECT Ssn
                                                FROM Employee
                                                WHERE FName = "John"
                                                AND LName = "Smith")));
```

6. Same as previous, but now with JOIN.

```sql
SELECT e1.FName, e1.LName
FROM Employee e2
JOIN WorksOn w2 ON e2.Ssn = w2.Essn
JOIN WorksON w1 ON w2.Pno = w1.Pno
JOIN Employee e1 ON e1.SSN = w1.Essn
WHERE e2.FName = "John"
AND e2.LName = "Smith";
```

Task 8: Relational Algebra & SQL

Consider the given relational model on a company. Express each of the following in relational algebra **and** SQL.

1. Give the department location and date of birth of all employees older than 60 years.

```sql
SELECT E.BDate, D.Location
FROM Employee E, DeptLocation D
WHERE E.NumDept = D.DNumber
AND E.BDate < 1965;
```

$$\Pi_{BDate,Location}((\sigma_{BDate<'1965'}(Employee)) \bowtie_{NumDept=DNumber} (DeptLocation))$$

2. Give the salary of all employees who work on all projects hosted by the “research” department.

$$researchProjects \leftarrow \rho_{PNO\leftarrow PNumber}(Project \bowtie (\sigma_{DName='Forschung'}(Department)))$$
$$\Pi_{Salary}(Employee \bowtie_{SSN=ESSN} ((\Pi_{ESSN,PNO} WorksOn) \div (\Pi_{PNO}(researchProjects))))$$

```sql
SELECT Salary
FROM Employee
WHERE Ssn IN (SELECT Essn
                FROM WorksOn W, Project P, Department D
                WHERE P.DNumber = D.DNumber
                AND P.PNumber = W.Pno
                AND D.DName = "Research"
                GROUP BY Essn
                HAVING COUNT(DISTINCT(Pno)) = (SELECT COUNT(Pno)
                                                FROM Project P, Department D
                                                WHERE P.DNumber = D.DNumber
                                                AND D.DName = "Forschung"));
```