1. Your colleague from the marketing team needs some analysis.

a. You have to make him the following overview:

|  |  |
| --- | --- |
| Plz\_kurz | qty |

In plz\_kurz he needs just the first two digits of the plz.

To select the first two digits of an attribute you have to use …

… LEFT (attributename, n) …

n stays for the requested signs qty (e.g. 2 for two signs)

**SELECT LEFT(plz, 2) AS plz\_kurz, 2 AS qty FROM plz;**

b. Create an overview with all unique ort names with the qtys, how often you will find the ort in the plz table!

c. Create another overview (similar to b) where you get to see all ort which are only once in your table!

d. Your colleague saw in the last selection you gave him (Exercise 4.6) that he didn’t even have one plz from Frankfurt. Can you give him a feedback why and create a new selection where you will see all plz from:

**SELECT plz FROM plz WHERE ort IN ('Frankfurt‘); returns empty because IN returns only exact matches.**

Berlin **SELECT plz, ort FROM plz WHERE ort LIKE ('berlin');**

Frankfurt **SELECT plz FROM plz WHERE ort LIKE ('Frankfurt%'); return (F.a.M & F(Oder))**

Aachen **SELECT plz, ort FROM plz WHERE ort LIKE ('Aachen');**

Now it’s your decision which SQL command you are going to use!

e. Show your colleague the TOP three cities (highest).

To let you show just the first three results of your SQL statement you have to use the following command at the end of your statement …

… LIMIT n

n stays for the qty of datasets you want to see.

**SELECT DISTINCT ort FROM plz LIMIT 3;**

2. Write down the bullet points of a relational database model and explain them in your own words!

3. What is a primary key and which important property should it have?

**Every table must have a Primary Key (PK). A PK is an attribute or a combination of attributes which is unique for every dataset of the table.**

**A PK must be minimal or irreducible, meaning that if deleting a part of the combination that creates the PK, the uniqueness should not be lost.**

**There can be more than one attribute or combination of attributes that can be used. They are called Candidate Keys. The PK will be chosen from one of these.**

**The simplest way to make a PK is to add a new column that holds a unique identifier (a number would be wise and easy to increment, simple to program in code) for the entity (Ex. ID, Prod\_ID).**

4. Explain in your own words what is meant with „Irreducibility or minimalism in relation to the uniqueness” and make me an example.

**Irreducibility says that, if you remove any of the attributes from your key, it stops being unique. This means that an irreducible PK should contain the smallest set of attributes that can uniquely identify an entity/row.**

**Example:**

|  |  |  |  |
| --- | --- | --- | --- |
| **FirstName** | **LastName** | **Pet** | **FavColor** |
| Alice | Jones | Dog | Red |
| Alice | Smith | Dog | Green |
| Bob | Smith | Cat | Blue |

**PET cannot be a PK alone, because both of the entities poses a dog, the same goes for FirstName. We could say that (FirstName, LastName, Pet, FavColor) is our PK, because the combination will be unique for every row in our table. But it is still reducible, because if we throw away the “FavColor”, the uniqueness is still preserved. We can reduce our PK even further by removing the “Pet”. We arrive at the irreducible form (FirstName, LastName), because, together, they are unique but, taken separately, both of them have “duplicate entries” (FirstName: Alice, LastName: Smith).**

5. Look at the following table and write down three candidate keys.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ID | Nachname | Vorname | Geburtstag | Geburtsort | Emailadresse |
| 1 | Maffay | Peter | 01.01.1962 | Dortmund | pm@gmx.de |
| 2 | Türk | Andreas | 04.10.1976 | Köln | TA@web.de |
| 3 | Sciurti | Lucia | 24.06.1984 | Langen | sl@yahoo.de |
| 4 | Friedrich | Frank | 01.01.1962 | Köln | ff@arcor.de |

a. Discuss the three suggestions and write about their advantages and disadvantages.

**(ID, Nachname, Vorname) – unique but reducible**

**(Nachname) – unique, irreducible but hard to auto-increment**

**(Emailadresse) – unique, irreducible but hard to read**

b. Which candidate key is you favorite? Explain why?

**ID – because it is unique to every entity, it is irreducible, it is an integer which is easy to auto-increment, save and modify in the DBMS or programmatically.**

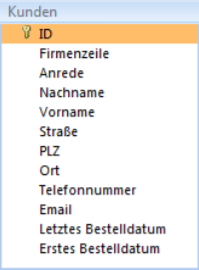
6. Why do we need primary keys?

**PKs are used to identify unique rows in a table, but also to reference an entity in other tables.**

7. Which advantages or disadvantages has the NULL value?

**Because NULL Values are placeholders for information that doesn’t exist yet, they are counted as existing, but are not used in manipulating commands like the function AVG (average). If using 0, the DBMS or third party Programing language wouldn’t be able to distinguish between a value entered by the user (0) or a placeholder.**

**An obvious disadvantage of a NULL Value is that it is a value and memory will be allocated accordingly. Another disadvantage is that a coding strategy must be employed to discover and treat it accordingly.**

8. The following table was sent to you. Now you have to change the structure so, that you won’t have NULL values in the attribute "Letztes Bestelldatum" and "Erstes Bestelldatum". Create a concept like the one you have seen in the lecture.