Marcel Bauer 5570300
Duc Viet Danny Pham 5567583
Joshua Hanheiser 5568227

Tutor: Brüchert, Jonah Benedict

Due: 18.07.2023

Database Systems - Assignment 10

Task 1: Develop a project idea

1. Have a look at the data sets Fahrraddiebstahl.csv1 , lor_planungsraeume.csv2 , and bezirksgrenzen.csv3 (also to find under resources in KVV/Whiteboard).

   Done.

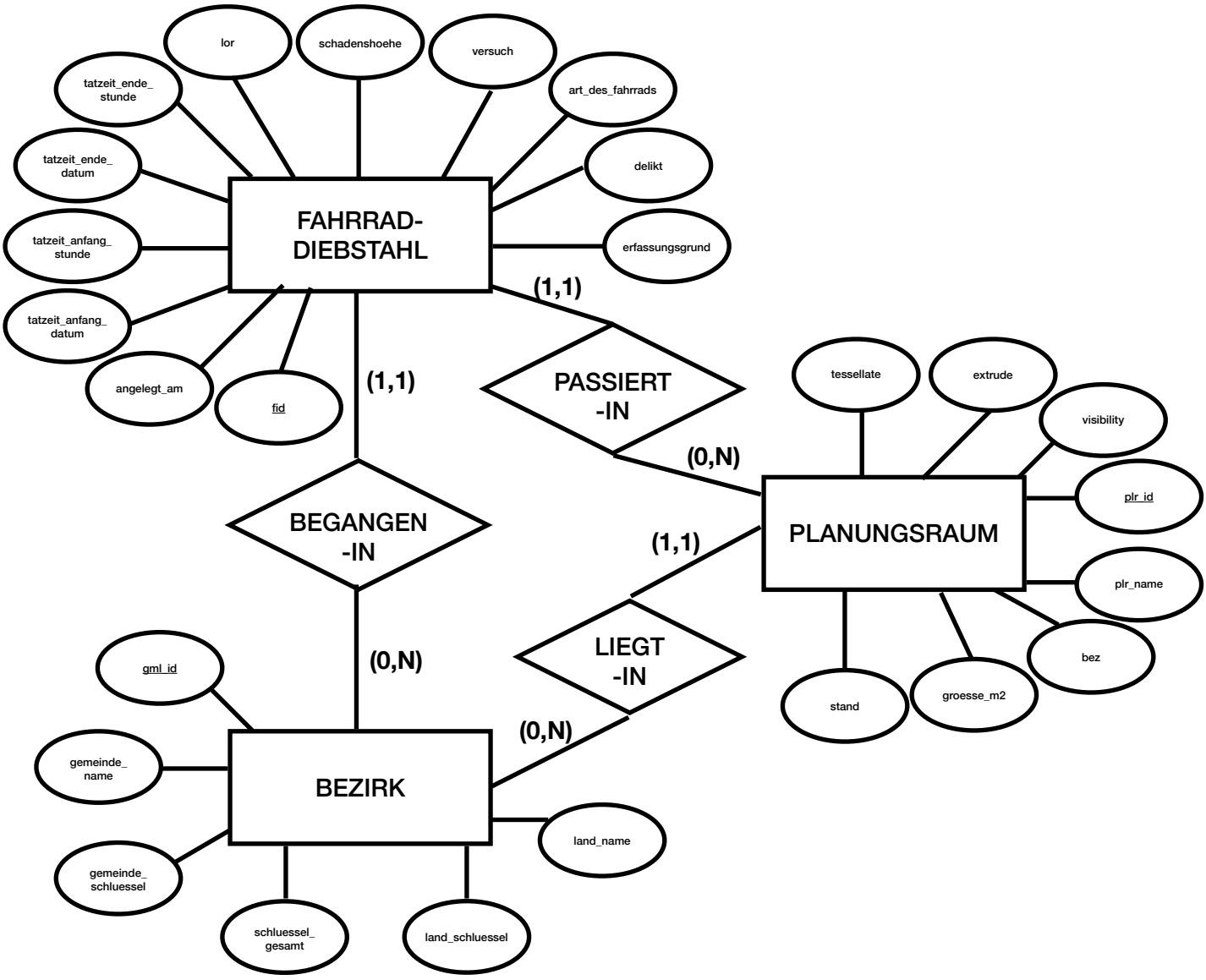
2. Develop a concept for data visualization that integrates and effectively presents information from the given datasets. Prepare a brief written description of your idea. Hint: Not all data within the data sets need to be processed. A map visualization is to be preferred. Optional: You are also encouraged to incorporate additional open data sources, to enrich the visualization and insights obtained from the original datasets.

   The web application we want to develop is an interactive map of Berlin that allows users to filter by districts and planning areas. The main focus is to provide important data on bicycle thefts in the respective areas.
   The application will provide a user-friendly interface on which the interactive map will be displayed. The map of Berlin will show the different districts or planning areas as desired. Users will have the option to filter the map according to their interests by selecting specific districts or planning areas.
   As soon as a user clicks on a district or planning area, detailed information on bicycle thefts in this area is displayed. This information can in turn be additionally sorted by the respective table headers.
   In order to obtain the required data on bicycle thefts in Berlin, we will use the given CSV files. Various technologies and frameworks are used for the technical implementation of the web application. Python, R and HTML can be used here, supported by frameworks such as Pandas or Dash to create the interactive map and user interface. A suitable database solution is ensured with PostgreSQL for storing and querying the bike theft data. APIs can be used to integrate publicly available data into the application.
   The user interface will be designed in an intuitive and user-friendly way to allow easy navigation on the map and use of the filter and click functions. A clear design and presentation of information will improve readability.
   It is important to update the bicycle theft data regularly to provide users with accurate and up-to-date information. Consideration will be given to how the data update process can be automated or manually managed to ensure that the data is always up to date.
   Before the web application is published, an extensive testing phase takes place to ensure that it works properly and delivers the expected results.
   The above concept serves as a starting point for the development of the web application. Depending on specific requirements and technical possibilities, further details and functions can be added.

Marcel Bauer 5570300 Tutor: Brüchert, Jonah Benedict Due: 18.07.2023
Duc Viet Danny Pham 5567583
Joshua Hanheiser 5568227

Task 2: Data schema and database set up

1. Create a meaningful data schema for your data set(s) and draw an according Entity-Relationship model (ERM).



2. Transform your ERM into a relational model.

FAHRRADDIEBSTAHL(fid, lor, angelegt_am, tatzeit_anfang_datum, tatzeit_anfang_stunde, tatzeit_ende_datum, tatzeit_ende_stunde, schadenshoehe, versuch, art_des_fahrrads, delikt, erfassungsgrund)
BEGANGEN-IN(lor, gml_id)
PASSIERT-IN(lor, plr_id)
BEZIRK(gml_id, gemeinde_name, gemeinde_schluessel, schluessel_gesamt, land_schluessel, land_name)
PLANUNGSRAUM(plr_id, tessellate, extrude, visibility, plr_name, bez, groesse_m2, stand)
LIEGT-IN(plr_id, gml_id)

Marcel Bauer 5570300
Duc Viet Danny Pham 5567583
Joshua Hanheiser 5568227
Tutor: Brüchert, Jonah Benedict
Due: 18.07.2023

3. Set up a PostgreSQL database with your relational model.

   CREATE DATABASE mydb WITH OWNER myuser ENCODING='UTF8' LC_COLLATE='en_US.UTF-8' LC_CTYPE='en_US.UTF-8';

   \c mydb;

   CREATE TABLE FAHRRADDIEBSTAHL(angelegt_am DATE, tatzeit_anfang_datum DATE, tatzeit_anfang_stunde INT, tatzeit_ende_datum DATE, tatzeit_ende_stunde INT, lor INT, schadenshoehe INT, versuch VARCHAR(32), art_des_fahrrads VARCHAR(32), delikt VARCHAR(64), erfassungsgrund VARCHAR(64), fid INT NOT NULL, PRIMARY KEY(fid));

   CREATE TABLE BEZIRK(gml_id VARCHAR(32) NOT NULL, gemeinde_name VARCHAR(32), gemeinde_schluessel INT, land_name VARCHAR(32), land_schluessel INT, PRIMARY KEY(gml_id));

   CREATE TABLE PLANUNGSRAUM(tessellate INT, extrude INT, visibility INT, plr_id INT NOT NULL, plr_name VARCHAR(32), bez INT, stand DATE, groesse_m2 FLOAT(32), PRIMARY KEY(plr_id));

Task 3: Pre-processing and import data

1. Pre-process (data cleansing) the data set(s).

   We have added a column for an id called „fid“ in the csv file of the bicycle thefts to have a primary key here.

2. Import the data set(s) into your database.

   SET datestyle = GERMAN, DMY; SELECT '16.04.2023'::DATE;

   COPY FAHRRADDIEBSTAHL(angelegt_am, tatzeit_anfang_datum, tatzeit_anfang_stunde, tatzeit_ende_datum, tatzeit_ende_stunde, lor, schadenshoehe, versuch, art_des_fahrrads, delikt, erfassungsgrund, fid) FROM '/Users/marcelbauer/Downloads/Fahrraddiebstahl.csv' DELIMITER ';' CSV HEADER;

   COPY BEZIRK(gml_id, gemeinde_name, gemeinde_schluessel, land_name, land_schluessel, schluessel_gesamt) FROM '/Users/marcelbauer/Downloads/bezirksgrenzen.csv' DELIMITER ',' CSV HEADER;

   COPY PLANUNGSRAUM(tessellate, extrude, visibility, plr_id, plr_name, bez, stand, groesse_m2) FROM '/Users/marcelbauer/Downloads/lor_planungsraeume_2021.csv' DELIMITER ';' CSV HEADER;

Task 4: Develop a Web application

1. - 3.

   See attachment.

Task 5: Present your solution in your tutorial

Task 6: Documentation