

Einführung in QGIS 3: Datenanalyse (Kapitel 4)

Ein Arbeitsblatt

In den letzten Kapiteln hast du erfahren, wie Geodaten erfasst und in einem GIS dargestellt werden. Somit ist alles vorbereitet für die „Spezialität“ des GIS, nämlich das Ableiten von neuen Informationen aus bestehenden raumbezogenen Daten. Dieses Ableiten von neuen Informationen nennt man **Analyse**.

Lernziele

Dieses Kapitel beinhaltet folgende Lernziele:

- den Unterschied zwischen Abfragen und Manipulationen in eigenen Worten erklären,
- aufzählen, welche Abfragearten es gibt und diese mit Hilfe je eines Beispiels erklären,
- für eine thematische Abfrage eine korrekte SQL-Abfrage formulieren.

Theorie

Aufgabe 1

Löse als Einstieg in das Kapitel Datenanalyse folgende Aufgaben:

1. *Suche 1: Rufe die Homepage <http://tel.local.ch/de/> auf und gib den Namen und falls bekannt auch den Wohnort einer Person ein, die du kennst und die einen Eintrag im Telefonbuch haben könnte (z.B. deine Eltern oder deine Grosseltern). Klicke anschliessend auf den `_Suchen`-Button. Was ist das Resultat deiner Suche? In welcher Form sind die Eingabe und die Ausgabe der Suche, wenn du die Suche auf der Seite tel.local.ch allgemein betrachtest?_*
2. *Suche 2: Rufe als nächstes die Homepage <http://www.sbb.ch/> auf und suche nach der nächsten Verbindung vom Bahnhof der am nächsten bei deinem Wohnort liegt nach Chur. Betrachte wieder die Eingabe- und Ausgabewerte. Dabei sind nicht so sehr die Resultate selber wichtig sondern die Form, in welcher die Resultate ausgegeben werden. Mache dir dazu wieder einige Notizen.*
3. *Vergleich: Was fällt dir auf, wenn du nun die Art der Resultate der Suche 1 und der Suche 2 vergleichst? Schreibe deine Feststellungen auf.*

Behalte deine Lösungen im Hinterkopf. Wir werden später in diesem Kapitel nochmals auf diese Aufgabe zurückkommen.

Wie bereits in der Einleitung angedeutet, ist das Ableiten von neuen Informationen aus bestehenden raumbezogenen Daten eine der wertvollsten und wichtigsten Funktionen eines GIS.

Allgemein werden zwei Arten von Analysen unterschieden: **Abfragen** und **Manipulationen**. Bei

den Abfragen werden die Daten nicht verändert, bei den Manipulationen werden entweder die Daten verändert oder es werden neue Daten erzeugt.

Abfragen

Es gibt drei Möglichkeiten, nach welchen Gesichtspunkten eine Abfrage durchgeführt werden kann.

1. Abfragen anhand von Sachdaten. Beispiel: Selektiere alle Seen, die eine Wassertiefe haben, die grösser als 100 m ist. Diese Abfragen werden **thematische Abfragen oder Selektionen** genannt.
2. Abfragen anhand von Beziehungen zwischen verschiedenen geometrischen Objekten. Beispiel: Selektiere alle Seen, die vollständig innerhalb eines Kantons liegen. Diese Abfragen werden **topologische Abfragen oder Selektionen** genannt.
3. Abfragen über die Geometrie der Objekte. Beispiel: Welche Hauptorte liegen in einem Abstand von weniger als 20 km um den Vierwaldstättersee. Diese Abfragen werden **geometrische Abfragen oder Selektionen** genannt.



Aufgabe 2

Überlege dir, um welche Arten von Abfragen es sich bei der Einstiegsübung handelte?

Datenbanksprache SQL

Um thematische Abfragen durchzuführen, wird die Datenbanksprache **SQL (Structured Query Language)** verwendet. Falls du bereits mit Datenbanken gearbeitet hast, wirst du die Sprache bestimmt bereits kennen. Dann ist für dich der nächste Abschnitt Repetition. Für alle anderen ist es eine kurze Einführung über die wichtigsten Ausdrücke einer Abfrage mit SQL.

Eine *SQL-Abfrage* ist vereinfacht dargestellt wie folgt aufgebaut:

```
SELECT Auswahlliste
FROM Quelle
WHERE Where-Klausel
```

Überblick über die Variablen, die du für einfache Abfragen kennen solltest:

Auswahlliste: Namen aller Attribute, die als Resultat angezeigt werden sollen. Ein * bedeutet, es werden alle Attribute angezeigt. Bei thematischen Abfragen innerhalb einer Attributtabelle ist nur die Abfrage nach allen Attributen möglich und auch die (Attribut-) Quelle ist gegeben. Deshalb konzentrieren wir uns zuerst auf die Where-Klausel.

Quelle: Namen der Tabellen, die Attribute enthalten, welche bei der Abfrage eine Rolle spielen. Bei einer Attributtabelle gibt es nur eine Tabelle.

Where-Klausel: Bedingungen, die erfüllt werden müssen.

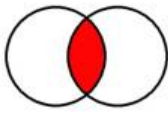
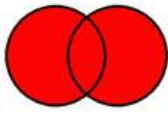
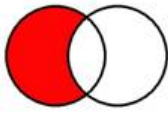
Überblick über die SQL-Ausdrücke, die in einer Where-Klausel verwendet werden können, und ihre

Bedeutung:

SQL-Ausdruck	Bedeutung
=	gleich
<	kleiner
<=	kleiner oder gleich
>	größer
>=	größer oder gleich
<>	nicht gleich, verschieden
LIKE	Vergleich von Zeichenfolgen. Dabei können folgende Zeichen als Platzhalter eingesetzt werden: %: Platzhalter für beliebig viele Zeichen. Z.B. LIKE 'F%' für alle Einträge, die mit dem Buchstaben F beginnen _: Platzhalter für genau ein Zeichen. Z.B. LIKE 'Me_er', wenn man nach einer Person sucht und nicht weiss, ob ihr Nachname Meier oder Meyer ist. Bei LIKE 'M_er' könnte er zusätzlich auch Mayer lauten.
IS NULL	Zelle enthält keinen Wert.
IS NOT NULL	Zelle enthält einen Wert.
BETWEEN	zwischen zwei Werten.

Es können auch mehrere Bedingungen hintereinander gesetzt werden. Durch logische Operatoren kann definiert werden, ob die *Bedingung 1* **und** (AND) *Bedingung 2* oder *Bedingung 1* **oder** (OR) *Bedingung 2* oder *Bedingung 1* **und nicht** (NOT) *Bedingung 2* gelten sollen.

Hier zur Veranschaulichung das Ganze nochmals mit Hilfe von Venn-Diagrammen dargestellt:

AND -Operator: Welches Element gehört zur Menge A und zur Menge B?	
OR -Operator: Welches Element gehört zur Menge A oder zur Menge B?	
NOT -Operator: Welches Element gehört zur Menge A aber nicht zur Menge B?	

Aufgabe 3

Probiere nun selber mal aus, folgende einfache Abfrage in SQL zu formulieren. Beachte dabei, dass

Attributnamen in doppelten Anführungszeichen und Textwerte wie 'Müller' in einfachen Anführungszeichen geschrieben werden:

1. Es sollen alle Seen selektiert werden, deren Fläche grösser als 20 km² sind und die tiefer als 400 m.ü.M. liegen.
2. Die Suche die du als Einstieg in die Theorie als Suche 1 auf der Homepage <http://tel.local.ch/de/> durchgeführt hast.

QGIS-Übung

Die "Spezialität" eines GIS soll natürlich auch nicht nur in der Theorie vermittelt werden. In diesem Kapitel wirst du selber einige thematische und räumliche Abfragen durchführen.

Lernziele

Am Ende dieser Übung beherrscht du folgende Punkte:

- Elemente anhand von thematischen und räumlichen Kriterien selektieren
- Aus bestehenden Werten für ein Feld einen neuen Wert berechnen

Thematische Selektion

Zu Beginn dieser Übung möchten wir eine einfache Abfrage durchführen. Du kennst sie bereits aus der SQL-Übung im Theorieteil:

Es sollen alle Seen selektiert werden, deren Fläche grösser als 20 km² sind und die tiefer als 400 m.ü.M. liegen.

Um diese Seen zu selektieren, könntest du nun jede einzelne Zeile der Attributtabelle durchschauen und so bestimmen, welche Objekte dazu gehören und welche nicht. Wenn du aber viele Objekte hast, dann würde diese Art von Suche sehr schnell sehr viel Zeit in Anspruch nehmen.

Einfacher geht es, wenn die vordefinierten Suchmasken verwendet werden. Wenn du nur nach einem Namen oder einer Zahl suchst, dann ist es am Einfachsten, den *Felderfilter* in der Attributtabelle unter *Alle Objekte anzeigen* zu benutzen.

Probiere diese Suche an einem kleinen Beispiel aus. Öffne die Attributtabelle der Seen. Im unteren linken Bereich siehst du eine Zeile auf der *Alle Objekte anzeigen* steht und ein Menü, das du aufklappen kannst. Suche in diesem Menü, im *Felderfilter*, nach dem Attribut, nach welchem du sortieren möchtest. Zusätzlich musst du den Begriff noch eingeben, nach dem du filtern möchtest. Du möchtest nun z.B. nach allen Seen suchen, in denen im Namen „Lago“ vorkommt. Wähle das Attribut „name“ im Menü unten links aus und suche nach dem Begriff "Lago".

Damit du sofort siehst, wie viele Elemente ausgewählt wurden, kannst du die Option *Filter...* aktivieren.

Aufgabe 4

Wie viele Lösungen erhältst du?

Kommen wir nun aber zum Anfangsbeispiel zurück. Diese Abfrage ist etwas komplizierter. Öffne durch Klicken auf den Button *Objekte über Ausdruck wählen* den Abfrageeditor. Hier kann die Abfrage in Form einer SQL-Abfrage formuliert werden. Es muss jedoch nur die *Where-Klausel* der SQL-Abfrage eingegeben werden. Von "*SELECT * FROM Seen WHERE max_Tiefe > 100*" muss also nur der Teil "*max_Tiefe > 100*" ins Abfragefeld eingetippt werden.

Die einzelnen Komponenten der Where-Klausel können durch Doppelklick auf die Felder im rechten Fenster zusammengesetzt werden.

! Aufgabe 5

Formuliere nun die Where-Klausel der SQL-Abfrage für die Einstiegsaufgabe und führe die Abfrage durch. Wie viele Seen erfüllen diese Bedingungen?

In anspruchsvolleren Analysen möchte man die ganze SQL-Sprache nutzen. Dazu reicht der *Feldfilter* nicht aus. Für eigene SQL-Abfragen kann man in QGIS den *DB Manager* nutzen, den man im QGIS-Menü unter *Datenbank* findet. Gegeben, dass im Projekt schon Daten geladen sind, kann man dort *Virtuelle Layer* wählen. Wenn man es aufklappt, sieht man "Projektlayer" und darunter die im Projekt abfragbaren Layer, d.h. in unserem Falle Seen und Kantonsgrenzen. Um eine eigene SQL-Abfrage durchzuführen, wählt man *Projektlayer* und klickt oben auf das Icon *SQL-Fenster*. Es erscheint nun ein neuer, zweigeteilter Dialog mit einem leeren Textfeld oben und einem Ausgabefeld unten. Im oberen Feld steht der Cursor bei Zeile 1 und man kann dort SQL schreiben.

Die SQL-Abfrage kann man mit Klick auf den Button *Ausführen* ausführen. Wenn die Abfrage erfolgreich war, wird das Ergebnis im Ausgabefeld angezeigt. Falls man vor der Ausführung der SQL-Abfrage die Checkbox *Als neuen Layer laden* anwählt, wird das Resultat als neuer Layer geladen.

The screenshot shows the QGIS interface with the DB Manager dialog open. The SQL query is as follows:

```

1 SELECT *
2 FROM Kantonsgrenzen, Seen
3 WHERE Kantonsgrenzen.Beitritts_j > 1600
4 AND Seen.Flaeche > 50
5 AND st_intersects(Seen.geometry, Kantonsgrenzen.geometry)

```

The results table shows the following data:

fid	Name	Kuerzel	Beitrittsj	E
1 8	Genf	GE	1815.0	4461
> 13	Neuenburg	NF	1815.0	1709

! Aufgabe 6

Formuliere dieselbe Anfrage wie in vorheriger Aufgabe mit der Where-Klausel nun aber als eigenständige SQL-Abfrage. Wie lautet die SQL-Abfrage?

Räumliche Selektion

Im ersten Teil dieses Abschnittes wurde eine Abfrage anhand von Werten formuliert, die in der Attributstabelle abgespeichert sind. Nun soll eine Abfrage durchgeführt werden, bei welcher die Position der Objekte berücksichtigt wird. Es handelt sich dabei um eine räumliche Abfrage.

Das Werkzeug für die räumlichen Abfragen findet man, in dem man den Button *Vektor* klickt und aus dem neuen Fenster unter *Datenmanagement Werkzeuge* das Werkzeug *Attribute nach Position zusammenfügen* auswählt.

Es sollen alle Kantone selektiert werden, bei denen mindestens einer der erfassten Seen ganz innerhalb liegt (d.h. der Kanton teilt sich den See nicht mit einem anderen Kanton oder einem anderen Land).

Der *Eingabe*-Layer in der Maske ist der Layer, in welchem die Elemente selektiert werden sollen. Da im vorliegenden Fall die Kantone selektiert werden sollen, wählen wir hier den Layer *Kantonsgrenzen*. Somit sind die Seen der *Layer verknüpfen* Layer. Die Checkbox *Nur gewählte Objekte* muss ausgewählt sein, damit nur die Seen mit einer Tiefe von mindestens 100 Metern ausgewählt werden.

Die Seen sollen vollständig innerhalb eines Kantones liegen. D.h. die Kantonsflächen enthalten die Seeflächen. Deshalb lautet die topologische Operation *enthält*. Man geht bei der Wahl der topologischen Operation also immer vom *Eingabe*-Layer aus und überprüft, wie er im Verhältnis zum *Layer verknüpfen* steht.

! Aufgabe 7

Führe diese Abfrage durch. Wie viele Kantone erfüllen diese Bedingung?

! Aufgabe 8

Wie müsste die Abfrage lauten, wenn alle Seen selektiert werden sollen, die zu mehr als einem Kanton oder Land gehören?

! Aufgabe 9

Wie viele Seen erfüllen diese Bedingung? Schreibe die Lösung auf dein Arbeitsblatt.

Felder berechnen

Als nächstes soll mit Hilfe von bestehenden Attributen ein neuer Wert erzeugt werden. Genauer gesagt soll die Einwohnerdichte der Kantone berechnet werden. Diese geschieht durch folgende Formel:

$$\text{Einwohnerdichte pro km}^2 = \text{Anzahl Einwohner} / \text{Fläche in km}^2$$

Für diese Berechnung muss die Attributstabelle der Kantone geöffnet werden. Wähle die Option *Feldrechner öffnen*. Falls du den Dichtewert in eine bereits bestehende Spalte schreiben möchtest, kannst du in der Maske *Vorhandenes Feld erneuern* ankreuzen und im Dropdown-Menü deine gewünschte Spalte auswählen. In unserem Fall möchten wir jedoch ein neues Feld anlegen. Gib deshalb unter *Ausgabefeld Name* den Attributnamen *Dichte* ein. Wähle als *Ausgabefeldtyp* die Option *Ganzzahl (integer)* und als *Ausgabefeldlänge* 10.

Wenn du den Dichtewert für alle und nicht nur für die selektierten Kantone berechnen möchtest, musst du darauf achten, dass das Kreuz bei *Nur x ausgewählte Objekte aktualisieren* nicht gesetzt ist.

Nun kannst du den *Feldrechner* entsprechend der Formel für die Einwohnerdichte formulieren. Dabei musst du nur den Ausdruck nach dem Gleichheitszeichen eingeben.

Wir möchten ausserdem die Dichte als ganze gerundete Zahl ohne Nachkommastelle berechnen (sog. Ganzzahl, englisch integer). Deshalb muss im Feldrechner-Ausdruck die Umwandlungsfunktion *to_int* (d.h. "wandle in Ganzzahl um") gewählt werden. Funktionen erwarten immer einen Ausdruck in Klammern. Setze also deinen Ausdruck in die Klammer nach *to_int*.

Aufgabe 10

Wie lautet der Feldrechnerausdruck?

Nach dem Berechnen muss der Editiermodus wieder ausgeschaltet werden. Wenn du mit der Berechnung zufrieden bist, dann speichere die Änderungen ab. Falls nicht, verwirf diese.

Kombinierte Selektion

Bestimme in einer weiteren Abfrage die Namen aller Hauptorte, deren zugehörige Kantone eine Bevölkerungsdichte von mehr als 200 Einwohner pro km² haben und die zwischen 1800 und 1900 der Schweiz beigetreten sind.

Es soll also nun eine Abfrage durchgeführt werden, in welcher sowohl thematische als auch die räumliche Komponente eine Rolle spielt.

Aufgabe 11

Überlege kurz, bei welchen Komponenten es sich um die thematische und bei welchen um räumliche Abfragen handelt.

Wir werden bei dieser Abfrage in zwei Schritten vorgehen. Im ersten Schritt wird die thematische Abfrage durchgeführt. Öffne dazu die Attributstabelle der Kantone und selektiere mit Hilfe der Option *Objekt über Ausdruck wählen* alle Kantone, bei denen die Bevölkerungsdichte grösser als 200 Einwohner pro km² beträgt und die zwischen 1800 und 1900 der Schweiz beigetreten sind.

Aufgabe 12

Wie lautet die Where-Klausel?

Jetzt sind alle Kantone selektiert, die diese Bedingung erfüllen.

Nun sollen noch die dazugehörigen Hauptorte herausgefunden werden. Dies geschieht mit Hilfe einer räumlichen Abfrage. Öffne dazu das *Attribute nach Position zusammenfügen*-Werkzeug.

Aufgabe 13

Welcher Layer ist denn nun der Eingabe-Layer und welcher der Layer verknüpfen-Layer? Welche topologische Operation muss gewählt werden?

Überprüfe vor dem Durchführen der Abfrage, dass die Einstellungen so sind, dass von den Kantongrenzen nur die selektierten berücksichtigt werden.

! Aufgabe 14

Um welche Hauptorte handelt es sich?

Kapiteltest 4

Frage 1

Im Kapitel 4 hast du gelernt, dass allgemein zwei Arten von Analysen unterschieden werden, die Abfrage und die Manipulation.

Welches ist der Unterschied zwischen diesen beiden Analysearten?

Schreibe deine Antwort in eigenen Worten. Füge bei jeder Art noch ein Beispiel hinzu.

Frage 2

Du hast gelernt, dass thematische Abfragen mit Hilfe der Datenbanksprache SQL formuliert werden.

Erstelle für die folgende Fragestellung eine SQL-Abfrage:

"Welche Kantone haben eine grössere Einwohnerdichte als 200 Einwohner pro km², traten der Schweiz vor dem Jahr 1600 n.Chr. bei und haben einen Namen, der nicht mit dem Buchstaben Z beginnt?"

Benutze bei der Formulierung den vereinfachten Aufbau einer SQL-Abfrage. Hebe die typischen Elemente des Aufbaus speziell hervor. Die Bezeichnungen der Tabellen und Attribute sollen von den Daten der Übung übernommen werden.

Frage 3

Es gibt neben den thematischen Abfragen auch noch weitere Abfragemöglichkeiten. Die verschiedenen Abfragen können auch kombiniert werden.

Erstelle für die folgende Fragestellung eine SQL-Abfrage:

"Selektiere alle Kantone, die nach 1600 der Schweiz beigetreten sind und Kontakt mit einem See haben, der grösser als 50 km² ist"

Bei der Beantwortung dieser Frage müssen nur die Seen berücksichtigt werden, welche im GeoPackage `qgis_einfuehrung_daten.gpkg` abgespeichert sind. Führe die Abfrage durch und schreibe Schritt für Schritt auf, wie du dabei vorgehst. Kleiner Tipp: Denke an die Aufgabe aus der Übung, in welcher die Namen aller Hauptorte selektiert wurden, deren zugehörige Kantone gewisse thematische Bedingungen erfüllen mussten.