

GIS Einführung mit QGIS

Christoph Rinne

03. Mai 2022

Inhalt

1 Raumabfragen und Quellenkritik	1
1.1 Vorbemerkung	1
1.2 Daten	1
1.3 Räumliche Abfragen	2
1.4 Übung zum Vorangehenden	3
1.5 Karte mit Diagramm - Die Bevölkerungsentwicklung	3
1.6 Polygone auflösen: Die Landschaftszonen	6
Literatur	7

1 Raumabfragen und Quellenkritik

1.1 Vorbemerkung

Ziel ist die Einführung in Abfragen, in einfache räumliche Abfragen, die grafische Umsetzung und eine resultierende Deutung aufgrund einer Fragestellung. Auch hier gilt: es geht nicht um die umfassende Darstellung aller Funktionen unter dem Menüpunkt “Vektor”, vielmehr möchte ich mit weitgehend typischen Fragestellungen und möglichen Lösungswegen auf die Arbeitsweise in GIS eingehen. Auch ist der Weg bisweilen das Ziel und das Ergebnis durchaus anders zu erreichen.

Die Fragen beginnen mit der Verteilung der Talaiots in den Gemeinden. Das ist so schlicht archäologisch natürlich uninteressant und muss mindestens gegen die Fläche geprüft werden. Für die Frage des Denkmalerhaltes müssten wir noch gegen die überbaute Fläche oder alternativ die Anzahl der Einwohner prüfen und erst dann könnten wir darüber nachdenken, ob die aktuelle Verteilung erkennbare Regeln und Bezüge zu geomorphologischen Parametern zeigt. Hieraus können wir dann abschließend eine Aussage über menschliches Verhalten in der Vergangenheit ableiten.

1.2 Daten

Es wird vorrangig mit dem Punktdatensatz der Fundplätze (mallorca-sites.shp) und dem Geländemodell (DGM 200) gearbeitet. Hinzu kommen die Gemeinden als kleine regionale Verwaltungseinheiten (recintos_municipales_inspire_peninbal_etr89). Ergänzen Sie diesen letztgenannten Layer bitte, nennen Sie ihn “Gemeinden” und setzen Sie die Symbolik auf “keine Füllung”. Später ergänzen wir noch Daten zur Bevölkerungsentwicklung.

1.2.1 Attribut filtern

Ziel ist es, aus der großflächigen Menge der Autonomen Gebiete nur die Balearn (Illes Balears) und aus den sehr zahlreichen Gemeinden nur die von Mallorca zu filtern und darzustellen. Für die Autonomen Gebiete Spaniens geht das natürlich am schnellsten über den Namen oder ein anderes spezifisches Feld der Attributtabelle (hier CODNUT2: ES53). Rufen Sie die Eigenschaften des Layers “Autonome Gebiete” auf: Register “Quelle”, Bereich “Objektfilter auf Datenanbieter” und bestätigen Sie [Abfrage erstellen]. Tragen Sie unter “Datenanbieter spezifischer Filter” folgenden Ausdruck ein "CODNUT2" = 'ES53'. Alternativ können Sie sich den Ausdruck mit den angebotenen Listen und Schaltern auch zusammenklicken. Wählen Sie [Testen] und bei Erfolg bestätigen Sie [OK] für die Abfrage und [OK] für die Eigenschaften. Zoomen Sie auf den Layer “Autonome Gebiete”.

Erstellen Sie auf die gleiche Weise einen Attributfilter für die Fundplätze mit folgender Bedingung: "Tipo_yacim" LIKE '%Talaiot%'. Und passen Sie den Layernamen an: "Talaiots".

Gefilterte Datenbestände haben mehrere Besonderheiten bzw. Vorteile:

- reduzierte Daten ohne die Datenquelle zu ändern oder einen neuen, redundanten Datensatz zu erzeugen und
- spezifische Informationen einer Datenquelle können als getrennte Layer mehrfach eingefügt und in der Abfolge der Darstellung eindeutig priorisiert werden.
- Gefilterte Datenbestände werden für das Editieren gesperrt.

1.3 Räumliche Abfragen

1.3.1 Vektoren zu Vektoren

Räumliche Abfragen sind eine Kernaufgabe eines GIS, für Vektordaten finden Sie dies unter "Vektor -> Forschungswerkzeuge -> Nach Position selektieren". Ziel ist es, nur die Gemeinden der Balearen zu erhalten, also alle Gemeinden innerhalb der entsprechenden Polygone des Layers "Autonome Gebiete".

Es gibt zwei mögliche Wege: 1. "Autonome Gebiete (gefiltert)" "enthält" "Gemeinden" und 2. "Gemeinden" "sind innerhalb" "Autonome Gebiete (gefiltert)". Das Ergebnis ist ein gefilterter Datenbestand oder verändert die vorhandene Auswahl. Beides führt zum Ziel, Variante 1 nach > 200 Sekunden, Variante 2 nach < 4 Sekunden (Zeit variiert nach Rechnerleistung). Diese Info finden Sie im Protokoll auf dem zweiten Reiter des zugehörigen Fensters. Der Unterschied ist gravierend, Gründe kenne ich nicht, vermute aber folgendes: Bei 1. wird für jedes Polygon aller Gemeinden und der vielen kleinen Inseln nachgesehen, ob dies zutrifft. Bei 2. wird für jede Gemeinde geprüft, ob es in den drei Polygon des gefilterten Autonomen Gebietes liegt, für die Masse der Gemeinden ist dies schnell erledigt. Merke: Prüfe erst mit einem kleinen Datensatz und dann mit dem großen.

Für die weitere Arbeit könnten wir diese Daten in eine neue shp-Datei exportieren, das schafft aber redundante Daten. Zudem sind da noch Formentera, Eivissa und Menorca dabei. Ich filter lieber den originalen Datenbestand der Gemeinden auf "CODNUT3" = 'ES532'.

1.3.2 Werkzeugkiste

Öffnen Sie nun die Werkzeugkiste (*toolbox*) mit "Verarbeitung -> Werkzeugkiste" oder <strg>+<alt>+<t>. Hier finden Sie alle Funktionen und ganz oben die "kürzlich verwendeten". Ich persönlich "verlaufe" mich in den Menüs regelmäßig. Mit den "kürzlich verwendeten", der Suche oder der angebotenen hierarchischen Ablage komme ich viel schneller zum Ziel.

1.3.3 Punkte in Polygon

Wir zählen zuerst die Talaiots je Munizipium: "Vektor -> Analyse Werkzeuge -> Punkte in Polygon zählen" oder in der *toolbox* 'zählen' suchen.

In **Punkte in Polygon zählen** sind die Angaben für Polygon = Gemeinden und Punkte = Talaiots klar. **Gewichtungsfeld** ist interessant, statt Punkte zu zählen wird die Summe des angegebenen Feldes berechnet. Als Gedankenspiel: wir haben Gräberfelder und gewichten (summieren) die jeweils aufgeführten Gräber oder Bestatteten. **Klassenfeld** ist auch interessant, dann wird die Vielfalt in dem ausgewählten Feld gezählt, nicht die Menge. In der Archäologie kennen wir das z.B. als einen Index für die Bewertung von 'Reichtum' bei Gräbern (u.a. Siklósi (2013), 31-40). Den Zählfeldnamen ändern wir in "Talaiots". Das Ergebnis wird als temporärer Layer erzeugt und kann bei Bedarf als neue Datei gespeichert werden.

Anmerkung zur Datenhaltung: An diesem Punkt wird die Arbeit mit shp-Dateien aus der Sicht der Datenbank unelegant oder sogar gefährlich. Will ich das Ergebnis sichern, muss ich speichern und schaffe so einen neuen, also größtenteils redundanten Datensatz. In einer SQL Datenbank würde die gesamte bisherige Arbeit als Abfrage (Anweisung) gespeichert, nicht das Ergebnis. Daraus resultiert kein neuer Datenbestand und ein stets aktuelles Ergebnis auf Grundlage der aktuellen Ausgangsdaten.

Unsere Frage zielte auf die Verteilung und die Repräsentanz unseres Datenbestandes. Wir setzen deshalb die Zählung ins Verhältnis zur Fläche und Visualisieren das Ergebnis.

Nennen Sie den neuen Layer “**Talaiots je Munizipium**”. Da das Ergebnis noch keine Bedeutung hat, machen wir direkt weiter. Öffnen Sie die Attributtabelle und dort den Feldrechner (Abakus). Erstellen Sie ein neues Dezimalfeld (numeric) mit 3 Nachkommastellen (Genauigkeit), nennen es “Talaits_qkm” (Talaiots je Quadratkilometer) und verwenden folgende Anweisung: “Talaiots” / (\$area/1000000). Erläuterung: Die Einheit unseres Koordinatensystems ist Meter, die Fläche ist demnach m², durch die Division erhalten wir km², das Hundertfache des sonst so beliebten ha (Hektar). Wählen Sie bei der Eigenschaft des Layers -> Symbolisierung “Abgestuft”, als Wert das neue Feld “Talaits_qkm” und denken Sie bei der Wahl für den Farbverlauf an die Rot-Grün-Schwäche (ca. 9% der Männer). Wählen Sie “Gleiche Anzahl (Quantile)”, bei Klassen “4” und dann [Klassifizieren]. Wechseln Sie auf den Reiter “Histogramm”, bestätigen Sie [Werte laden] und betrachten Sie die Häufigkeitsverteilung.

Anmerkung zur **Statistik**: Vier Quantile (= Quartile) sind eine allgemeine (pauschale) Teilung in 25%-Schritte von mindestens ordinal skalierten Werten, um Verteilungen vergleichen zu können. Ich will nicht vergleichen sondern darstellen und finde, diese pauschalen Klassengrenzen trennen meine Daten nicht dem Datenbestand angemessen. Deshalb ändere ich diese Grenzen und trenne an den erkennbaren Brüchen in den Daten: <0.015 für eigentlich keine, <0.05 wenige, <0.1 ordentliches Mittelfeld, > 0.1 die “Reichen” und bestätige mit [OK]. Schieben Sie den Punktlayer “Talaiots” über den neuen Layer.

Eine **erste Deutung** einzelner Auffälligkeiten ist möglich. Die Gemeinden sind überwiegend kleine *a priori* politische Einheiten, Sie sind aber zugleich abhängig von einer traditionellen, teils naturräumlichen Gliederung (*comarcas*). Wechseln Sie dafür zwischen DGM (Höhenmodell) und Dichtekarte hin und her. Natürlich fällt das Hochgebirge mit über 1000 m Höhe um den Puig Major (1445 m) aus, deshalb verändere ich die Symbologie für das DGM auf maximal 800 und bekomme mehr Dynamik in den Rest. Am östliche Abhang der Tramuntana und im Übergang zur Ebene (Raiquer) sehe ich mehrere Agglomerationen, ich vermute einen Zusammenhang zur Hydrologie oder dem Baumaterial (als Fragen für später notieren!). Die nördliche Mitte (Pla de Mallorca) ist recht regelhaft belegt, wogegen der südliche Abschnitt (Migjorn), insbesondere die sehr große Gemeinde Llucmajor, trotz flacher Landschaft erstaunlich leer ist (Frage notieren: Moderner Landwirtschaft?). Palma hat erstaunlicherweise mehr Talaiots als Llucmajor: Die hohe Bevölkerungszahl und der hier erfolgte Landesausbau sind kein Problem?.

Die Frage zur Bevölkerungsentwicklung und -dichte wird als nächstes thematisiert.

1.4 Übung zum Vorangehenden

Sie sollen das Vorangehende wiederholen und mehr über Mallorca lernen. Deshalb kommen jetzt die *comarcas* (Landschaftszonen) ins Spiel. Lesen Sie dazu z.B. den entsprechenden Abschnitt bei [Wikipedia](#).

1. Importieren Sie die Tabelle “municipio-comarca.txt”, nachdem Sie sich die Struktur der Datei angesehen haben.
2. Stellen Sie eine Verbindung (join) zwischen dem Layer “Gemeinden” und der zuvor importierten Tabelle her (NAMEUNIT, municipio).
3. Verändern Sie die Symbologie für den Layer auf “Kategorisiert”, Wert ist das neue Feld “municipio-comarca_comarca” (hätten Sie beim *join* umbenennen können!) und zufällige Farben.

Wiederholen Sie das ganze für die Tabelle zu den Bevölkerungszahlen (pop-balears.txt). Betrachten Sie erst die Rohdaten, importieren Sie und finden Sie selbständig die Felder für den *join*. Verwenden Sie diesmal die Option “Benutzerfeldnamenpräfix” (“pop_”).

1.5 Karte mit Diagramm - Die Bevölkerungsentwicklung

1.5.1 Anmerkung zur Forschungsgeschichte

Der Erhalt archäologischer Denkmale, zumal dieser monumentalen, obertägig gut sichtbaren Talaiots, ist vor allem durch den Menschen gefährdet. Die Bevölkerungszahl und deren Entwicklung ist demnach durchaus ein wesentlicher Faktor bei der Bewertung des Denkmalbestandes. Daneben ist es aber auch das Bewusstsein für den Denkmalcharakter und die Verknüpfung zur eigenen Identität, die den Erhalt beeinflusst. Auf den Balearen beginnt eine erste systematische Erfassung der Archäologie in der ersten Hälfte des 20. Jh. Es sind z.B. die Arbeit von Josep Colomines und Luis Amorós. Das politische Umfeld in dieser Zeit ist von gewaltigen Umbrüchen geprägt, der 2. Spanischen Republik, dem Bürgerkrieg und der frühen Franco-Diktatur. Ab der Mitte des Jahrhunderts prägt dann die zentralistische Organisation der etablierten Diktatur auch die Archäologie und die Entwicklung eines zentralen Registers, Denkmal- und Grabungsamtes. Die zentrale Figur dieser Zeit ist Rosseló Bordoy. Der Tourismus entwickelt sich

ab den 1950er Jahren deutlich, er hat auch Auswirkung auf einen zunehmenden Bauboom bis in die 1970er Jahre. Ein massiver Straßenausbau geht damit einher, denn die Eisenbahn des ausgehenden 19. Jh. setzt sich nicht durch. Beachten Sie hierzu die publizierten Fotos der Ausgrabung des DAI auf der Talaïot-Siedlung von S'Illot in den Madrider Mitteilungen und den heutigen Zustand z.B. in Google Earth (39.5687,3.372), es ist nur ein Beispiel von vielen (Frey and Rosselló Bordoy (1964)). Alternativ können Sie im [IDEIB WebGIS](#) 1. nach S'Illot suchen und 2. bei den Layern (Lista de Capes) ergänzen Sie (Afegir dades) das "Ortofoto 1956". Warum das? Es ist wichtig für die Bewertung des Denkmalbestandes und unser ganzheitliches Verständnis von Geschichte, Mensch und Umwelt.

1.5.2 Die Bevölkerungsentwicklung

Die Bevölkerungsentwicklung der Balearen reflektiert den Landesausbau, ein wichtiger Aspekt für den Denkmalerhalt, und weist auf die zuvor nicht genannten Unterschiede in der Entwicklung zwischen den Inseln hin, insbesondere Mallorca und Menorca. Beachten Sie die gewaltigen Zuwachsraten (%) und Unterschiede für die Zensus von 1940 bis 1981. Diese Entwicklung beeinflusst Mallorca noch heute und soll nachfolgend grafisch umgesetzt werden.

Insel	1900	1910	1920	1930	1940	1950	1960	1970	1981	1991	2001	2011
Mallorca	211263	4,6	6,5	9,0	12,6	5,1	7,8	25,5	23,4	6,4	14,8	26,8
Menorca	39178	4,1	2,7	5,7	2,9	0,3	-0,6	16,1	10,9	15,0	15,9	35,7

Tabelle zur Bevölkerungsentwicklung. Einwohner für 1900 und nachfolgende Steigerung (%) zum jeweils vorangehenden Zensus. Quelle der Rohdaten: [INE](#).

Anmerkung: Im Datenbestand der Übung finden Sie zwei Dateien, pop_balears.csv und pop_balears.txt. Letztere beinhaltet Angaben zu Quelle und die ursprünglichen Spaltennamen. Diese Namen sind nur das Jahr des Zensus und führen bei der späteren Arbeit zu irritierenden Fehlern. Das war ein Fallstrick, den ich Ihnen hier erspare.

Stellen Sie eine Verknüpfung (Join) zwischen den Gemeinden und der importierten Tabelle der Bevölkerungszahlen (pop_balears.csv) her. Öffnen Sie die Eigenschaften für den **Layer "Gemeinden"** mit angehängten Bevölkerungszahlen. Ziel ist die Darstellung des jüngsten Zensus (2011) als abgestufter Farbverlauf je Polygon mit 4 Klassen. Für die **Symbologie** starte ich wie im vorigen Beispiel mit den Quantilen, kontrolliere das Histogramm und bilde dann folgende Klassen: bis 7500, bis 20000, bis 54000, bis 403000. Ich kontrolliere erneut das Histogramm und erhöhe Anzahl der Balken (Histogrammkästchen) auf 100. Um die niedrigste Klasse gegen die Gemeinden ohne Werte abzugrenzen, erhalten diese einen schwachen Farbwert (10%).

Die **Beschriftung** aller Gemeinden in der Karte kann die Darstellung überfluten, deshalb will ich nur die größere Hälfte beschriften. Wählen Sie bei Beschriftung statt "keine Beschriftung" "Einzelne Beschriftung" und öffnen bei Wert mit [E] den Ausdruckeditor. Dort ergänzen Sie folgenden Ausdruck:

```
if(
  $area>
  (array_get(
    array_sort(array_agg($area)),
    round((count("NAMEUNIT")/2),0))),
  "NAMEUNIT", "")
```

"Die größere Hälfte", was hätten Sie von Hand gemacht? 1. Die Daten nach der Fläche aufsteigend sortiert und 2. ab der Mitte alle Werte beschriftet. Genau das passiert im Code. Ich erläutere den Inhalt der Klammern gemäß den Regeln von von links nach rechts und innen nach außen.

- `array_agg($area)` : schreibt alle Flächengrößen in eine Liste (array)
- `array_sort()` : sortiert unsere Liste ohne weitere Anweisung aufsteigend.
- `array_get()` : liefert uns aus einem array den Wert des n-ten Elementes
 - Der erste Teil ist die eben erklärte Liste.
 - Der zweite Teil ist: `round((count("NAMEUNIT")/2),0)`. Dies zählt die Anzahl der Namen (Gemeinden), teilt diese durch 2 und rundet das Ergebnis auf 0 Stellen nach dem Komma (Ganzzahl). Ein 1/2 Element kann es bei ungerader Zählung ja nicht geben.

- “NAMEUNIT”, “ ” : ist die Rückgabe der *if*-Bedingung, die das ganze klammert. Die Bedingung ist: die aktuelle Fläche (\$area) sei größer als der ausgezählte Grenzwert in der Mitte der Reihe. Wenn das wahr ist, folgt der erste Wert (NAMEUNIT), wenn nicht eine leere Zeichenkette.

Bestätigen Sie den Ausdruckseditor mit [OK]. Wechseln Sie noch zur Platzierung und wählen hier “Horizontal (langsam)”. Die vorangehende Regel hätten wir auch unter “Regelbasierte Beschriftung” definieren können, Rückgabewerte der *if*-Bedingung sollten dann 1 (wahr), 0 (falsch) sein. Hier habe ich aber nicht die Platzierungsoptionen.

Im Register **Diagramme** wähle ich Histogramm und bei den Attributen ergänze ich die Felder pop_1900 bis pop_2011 mit dem [grünen Plus] bei “Zugewiesene Attribute”. Die Farbe und der Legendeneintrag sind jetzt leidige Handarbeit. Die Farbe wird automatisch zufällig zugewiesen und stört, ich setze alles auf ein mittleres Grau (40%) und übertrage es mit der Pipette auf die anderen Felder. Die Legendenbeschriftung kürze ich auf das Jahr des Zensus. Unter Darstellung setze ich die Balkenbreite auf 1,5. Wichtig wird jetzt der Bereich **Größe**. Anders als bei Kreisdiagrammen (Durchmesser) kann das Histogramm keine feste Größe bekommen sondern wird anhand des größten Wertes eines Attributes skaliert. Wählen Sie das Attribut pop_2011 und wichtig(!) klicken Sie auf [suchen], weisen Sie der Balkenlänge “100” zu. Wegen der großen Unterschiede im Datenbestand ist diese Skalierung insgesamt schwierig. Wir könnten eine logarithmische Skalierung erwägen, das ginge bei der Zuweisung der Attribute zu Beginn mit dem [E] über dem grünen Plus. Den Ausdrucksgenerator haben Sie schon kennengelernt, der Ausdruck `log10("pop_2011")` bereitet Ihnen also keine Probleme. Das muss aber für alle Attribute wiederholt werden. Zudem ist unsere Sehweise für logarithmische Skalierung nicht gut ausgeprägt und zerstört obendrein die Dramatik der Zahlen, deshalb nehme ich davon Abstand.

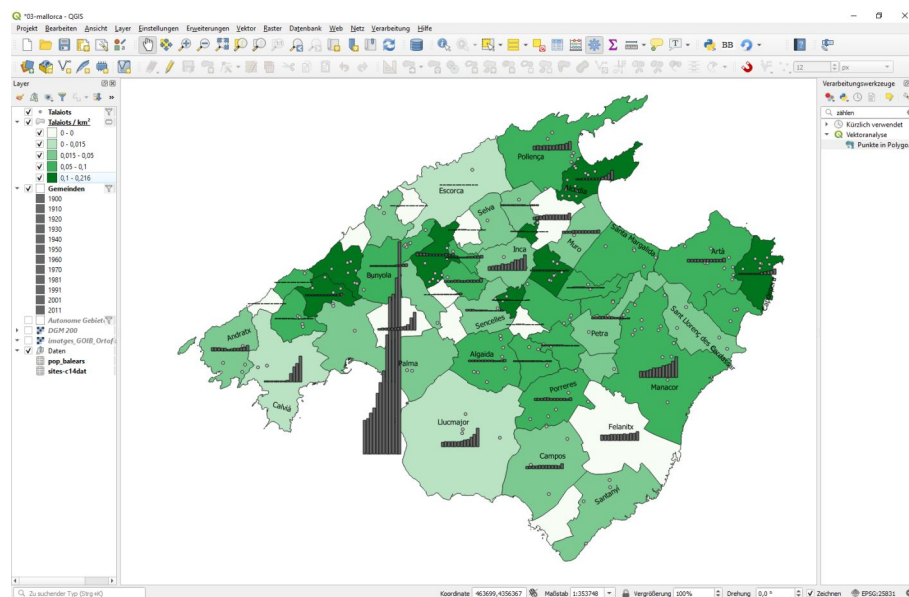


Abb. 1: Mallorca. Talaiots je km² pro Gemeinde und jeweilige Bevölkerungsentwicklung von 1900 bis 2011.

Einige Gemeinden haben keine Bevölkerungsdaten und werden deshalb nicht dargestellt. Ich blende das DGM aus, die Autonomen Gebiete (Küstenlinie) ein und ergänze auch noch den Punktlayer der Talaiots. Leider liegen die Balkendiagramme immer oben drauf und überdecken einige Punkte der Talaiots. Die Grafik visualisiert Bevölkerungsentwicklung und Denkmalbestand ganz brauchbar, ein statistischer Beleg ist sie nicht.

Die schnelle **Erläuterung und Deutung**. Wir erkennen nun leicht, dass nicht alle bevölkerungsarmen Gemeinden viele Denkmale haben, es gibt auch bevölkerungsreiche Gemeinden mit vielen Denkmalen (z.B. Alcúdia im Norden). Nun könnten wir ja behaupten, Alcúdia ist als bekannter Touristenort erst ab den 1970er stark gewachsen, da wurde auf Denkmale bereits geachtet. (Letzteres ist übrigens eine unbewiesene Behauptung, bei der einige Archäologen auf der Insel nicht emotionslos bleiben werden.) Wir haben alternativ auch Calvià im Südwesten mit wenigen Talaiots und einer ebenfalls ab den 1970ern explodierenden Bevölkerung. Die Gemeinden Felanitx, Campos und Porreres liegen dicht beieinander

und zeigen eine eher moderate Bevölkerungsentwicklung bei sehr unterschiedlichem Denkmalbestand. Mein Fazit ist klar, die Bevölkerungsentwicklung und damit vermutlich der Landesausbau seit 1900 ist kein wesentliches oder bestimmendes Element für den Denkmalerhalt und die heutige Verteilung. Palma möchte ich ausklammern, hier reicht die Entwicklung weiter zurück, der Verlauf ist sicher anders.

Anmerkung

Harry Graf Kessler, ein illustrierter Zeitzeuge des frühen 20. Jahrhunderts, weilt ab November 1933 für ca. 2 Jahre auf Mallorca und beschreibt u.a. eine dekadent abgehobene Gesellschaft von Zugezogenen und die frühen Anfänge der zuvor dargestellten Entwicklung (Kessler (2010)).

1.6 Polygone auflösen: Die Landschaftszonen

Ich möchte nun die Umweltaspekte betrachten und beginne mit den *comarcas* als Landschaftseinheiten. Da mir eine entsprechende Karte fehlt, generiere ich diese aus den Gemeinden durch das Auflösen nach Attribut. Diese Attribute haben Sie als Übung im vorangehenden Kapitel selbstständig verknüpft (s.o., Tabelle municipio-comarca).

Wählen Sie “Vektor -> Geoverarbeitungswerkzeuge -> Auflösen”, Eingabelayer: Gemeinden, Feld(er) auflösen: “com_comarcas” (das Präfix “com_” hatte ich gesetzt), temporär erzeugen und mit [Starte] ausführen. Das Ergebnis ist erneut nur temporär vorhanden, zeigt überflüssige Spalten mit falschen Werten (jeweils den Wert des 1. Datensatzes), müsste bereinigt und bei Bedarf als eigener, eigentlich redundanter Datensatz (shp) gespeichert werden. Nennen Sie den Layer “comarcas”, das dt. Wort ist mir zu lang, wählen Sie eine klassifizierte Symbologie und eine Beschriftung aller Polygone mit dem Feld “com_comarcas”. Die neue (technische) GIS-Aufgabe ist damit erledigt.

Die zweite (aktuelle) **Frage** bezog sich auf einen allgemeinen Zusammenhang zwischen der Geomorphologie und der Verteilung der Talaiots. Das will ich eben noch prüfen und zwar mit einem Chi²-Test. Also erneut “Punkte in Polygon zählen” (Zählfeldname: Talaiots) und die Fläche (km²) ergänzen (Feldrechner, neues Feld: a_qkm, \$area/1000000). Wir brauchen nachfolgend nur drei Spalten: com_comarca, Talaiots, a_qkm. Das geht über viele Wege. Ich lösche alle überflüssigen Spalten, kopiere die Attributtabelle in die Zwischenablage (<strg>+<c>), füge diese in einer Office-Tabelle (Excel) ein und lösche die erste Spalte und kopiere das Ergebnis erneut.

com_comarca	Talaiots	a_qkm
Llevant	42	580
Pla de Mallorca	47	738
Palma	6	209
Serra de Tramuntana	42	832
Migjorn	13	811
Raiguer	29	472

Exkurs Statistik:

Wenn Sie dies wie nun dargestellt selbst nacharbeiten wollen, müssen Sie **R** und **R-Studio** als Nutzeroberfläche installieren. Es geht aber auch in Excel, übertragen Sie hierfür die bei Siegmund (2020) erläuterten Rechenschritte in Ihre Datentabelle.

Statistiker hätten *a priori* mit R als Programm mit vielen Paketen (vereinfacht Funktionssammlungen) gearbeitet und hier GIS als auch Statistik in einem Arbeitsumfeld gehabt. Wir müssen nun wechseln. Der Chi²-Test prüft Abhängigkeiten in einer Kreuztabelle. Mathematisch ist er einfach und bei entsprechenden Daten robust. Geprüft wird der Unterschied in den Feldern gegenüber einer unabhängigen (gleichen) Verteilung. In unserem Fall natürlich unter Berücksichtigung der jeweiligen Flächengröße (s. u.a. Shennan (2001), 65-70; Siegmund (2020), 201-207).

Wenn Sie RStudio gestartet haben, öffnen Sie ein neues R-script “File -> New File -> R-script” und fügen folgenden Code ein. Viel erklärt sich aus den engl. Namen der Funktionen, ergänzend nur folgende Hinweise:

- cbind = column bind,

- c = concatenate,
- p steht für die *probability* im Test, hier der Flächenanteil der *comarcas*,
- ds ist der frei wählbare Name des Datensatzes und
- ds\$Talaiots ist die Variable Talaiots in dem Datensatz ds.

```
# Einlesen der Daten
ds<-data.frame(
  cbind(
    c(42,47,6,42,13,29),
    c(580,738,209,832,811,472)
  )
)
# Zuweisen von Spalten- und Zeilennamen
colnames(ds)<-c("Talaiots","a_qkm")
rownames(ds)<-c("Llevant","Pla de Mallorca","Palma","Serra de Tramuntana",
               "Migjorn","Raiguer")
# Berechnen der Flächenanteile in Prozent
ds$a_percent<-ds$a_qkm/sum(ds$a_qkm)
# Ausführen des  $\chi^2$ -Tests
chisq.test(ds$Talaiots, p=ds$a_percent))
```

Das Ergebnis lautet:

```
Chi-squared test for given probabilities
data:  ds$Talaiots
X-squared = 30.918, df = 5, p-value = 9.724e-06
```

Der Prüfwert p ist sehr klein (0,000009724), die Nullhypothese für keine Unterschiede oder eine zufällige Verteilung dürfen wir ablehnen. Unser Ergebnis ist also kein Zufall, sondern es gibt einen Zusammenhang, den wir weiter hinterfragen können. Ab hier wird es spannend. Auf meinem Papier standen noch die Fragen nach einem Zusammenhang mit der Verfügbarkeit von Wasser und Baumaterial. Ergänzend möchte ich auch auf die runde und eckige Bauform bei den Talaiots hinweisen, auch hier gibt es in der räumlichen Verteilung Unterschiede.

Literatur

- Frey, O.-H., Rosselló Bordoy, G., 1964. Eine Talayot-Siedlung bei S'Illot, Mallorca. Madrider Mitteilungen 5, 55–71.
- Kessler, H.<Graf>., 2010. Das Tagebuch / Harry Graf Kessler. Hrsg. Von Roland S. Kamzelak Bd. 9 1926 - 1937, Veröffentlichungen der Deutschen Schillergesellschaft. Cotta, Stuttgart.
- Shennan, S., 2001. Quantifying Archaeology, Second. ed. Edinburgh Univ. Press, Edinburgh.
- Siegmund, F., 2020. Statistik in der Archäologie: eine anwendungsorientierte Einführung auf Basis freier Software. (Vers. 0.9) Pre-Print der Febr. 2020 im Druck erschienenen Fassung.
- Siklósi, Z., 2013. Traces of social inequality during the Late Neolithic in the eastern Carpathian Basin, Dissertationes pannonicae ex Instituto Archaeologico Universitatis de Rolando Eötvös Nominatae Budapestinensis provenientes. Eötvös Loránd University, Institute of Archaeological Sciences, Budapest.