

Explorative Datenanalyse: Schweizer Immobilien

Alexander Schilling, Marius Läubli, Arian Iseni, Oliver Pejic

12. Januar 2024

1 Einleitung

In dieser Studie befassen wir uns mit einem Datensatz, der Informationen über Schweizer Immobilien enthält. Der Datensatz, gesammelt zwischen August und Dezember 2022, umfasst diverse Attribute, darunter URL und verschiedene Website-Attribute der Immobilien. Zusätzlich wurde der Datensatz mit geopolitischen Daten angereichert, um eine umfassendere Analyse zu ermöglichen. Ziel dieser explorativen Datenanalyse (EDA) ist es, ein besseres Verständnis der Datenstruktur zu gewinnen und Einblicke in die Beziehungen und Abhängigkeiten zwischen verschiedenen Attributen zu erlangen, insbesondere im Hinblick auf die Attribute `price_cleaned` und `type`.

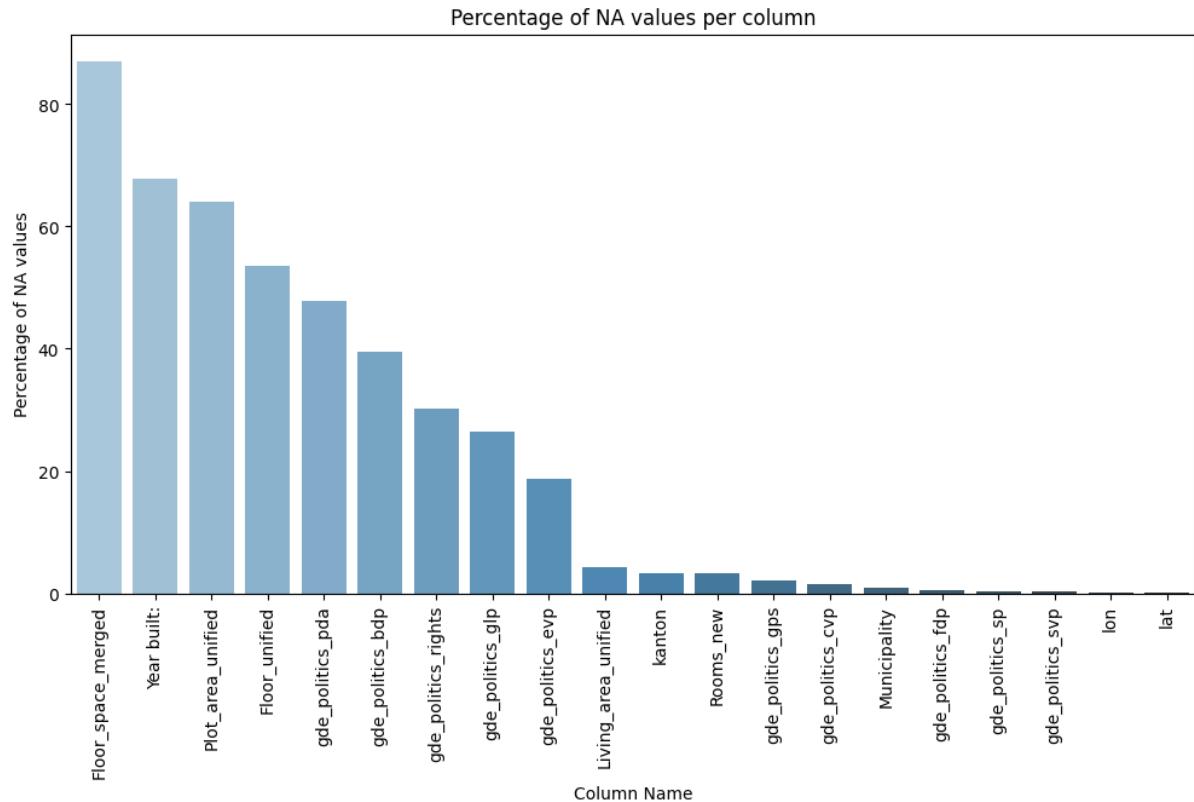
2 Datenaufbereitung

Zu Beginn wurde der Datensatz sorgfältig aufbereitet, um die Konsistenz und Qualität der Daten zu gewährleisten. Nicht relevante und fehlende Daten wurden identifiziert und bereinigt. Zudem wurden einige Attribute vereinheitlicht und bereinigt, um die Analyse zu erleichtern. Beispielsweise wurden Spalten mit der Anzahl der Zimmer und dem Preis vereint und bereinigt, und einige Spalten mit hohen Fehlwerten wurden entfernt.

3 Explorative Datenanalyse (EDA)

Die explorative Datenanalyse dient dazu, tiefere Einblicke in den Datensatz zu erhalten und etwaige Muster, Anomalien oder Beziehungen zwischen den Daten zu identifizieren.

3.1 Fehlwerte in den Daten

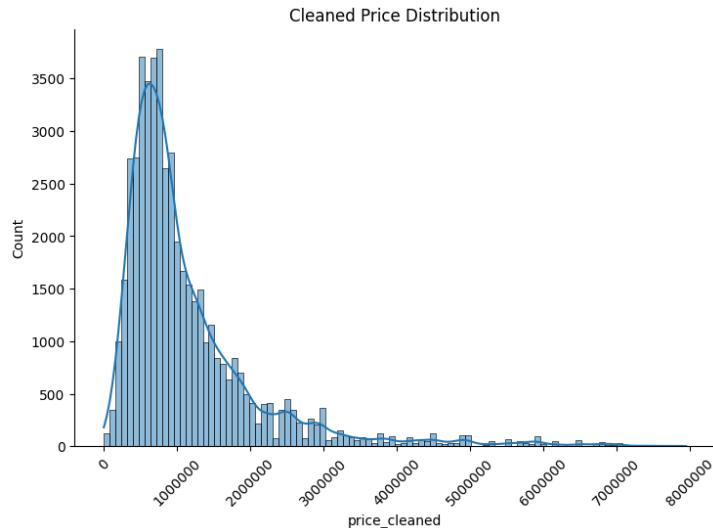


Der erste Schritt in der EDA war die Untersuchung von Fehlwerten in den Daten. Die Darstellung der Percentage of NA values per column zeigt deutlich, welche Spalten eine hohe Anzahl von fehlenden Werten aufweisen. Insbesondere Spalten wie "Floor_space_merged", "Year built:" und einige andere weisen über 50% Fehlwerte auf. Dies könnte darauf hindeuten, dass diese Informationen in den meisten Listings nicht verfügbar sind oder nicht erfasst wurden. Eine sorgfältige Überlegung ist erforderlich, ob diese Spalten für zukünftige Analysen behalten oder entfernt werden sollten.

3.2 Preisanalyse

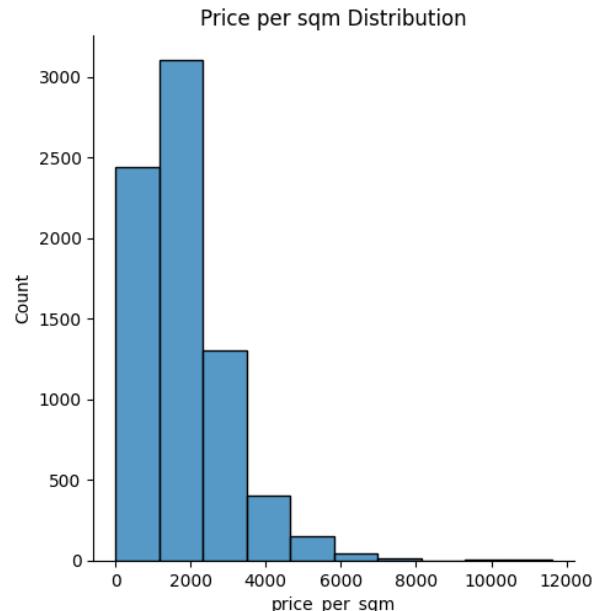
Die Immobilienpreise in der Schweiz variieren stark je nach Region, Gemeinde und Kanton. In der vorliegenden Analyse betrachten wir die durchschnittlichen Immobilienpreise sowohl auf Gemeinde- als auch auf Kantonsebene. Dies ermöglicht es uns, ein besseres Verständnis für die Preislandschaft im Land zu bekommen und die teuersten und günstigsten Regionen zu identifizieren.

3.2.1 Verteilung des bereinigten Preises (Cleaned Price Distribution)



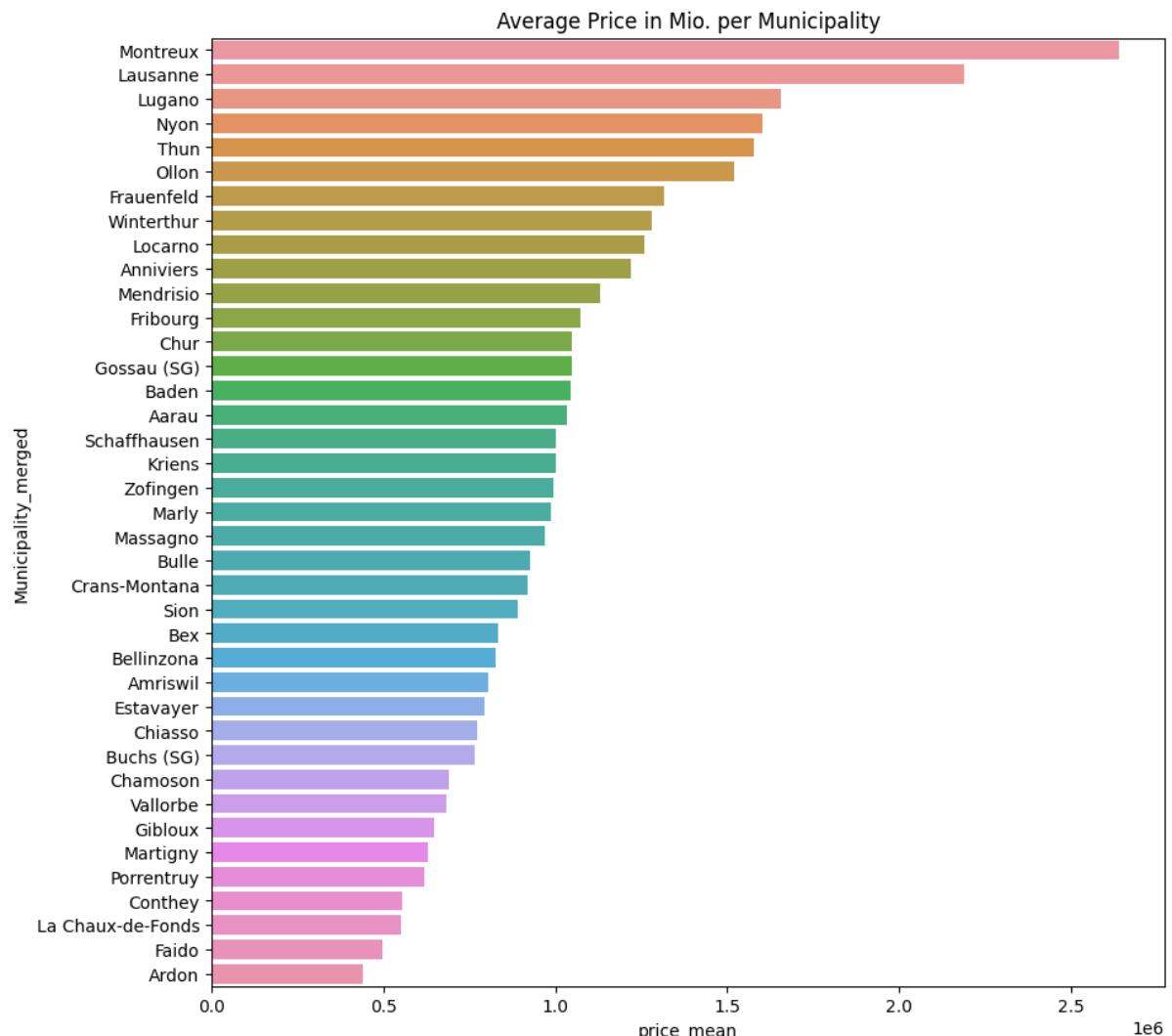
Dieses Histogramm zeigt die Verteilung der bereinigten Immobilienpreise. Der Grossteil der Immobilien liegt in einem Preissegment von 0 bis 2.000.000, wobei der Höhepunkt deutlich unter 1.000.000 liegt. Dies deutet darauf hin, dass der grösste Teil der Immobilien in diesem Datensatz zu relativ günstigen Preisen verkauft wird. Es gibt jedoch auch einige Immobilien, deren Preis bis zu 8.000.000 reicht, diese sind jedoch deutlich seltener.

3.2.2 Verteilung des Preises pro Quadratmeter (Price per sqm Distribution)



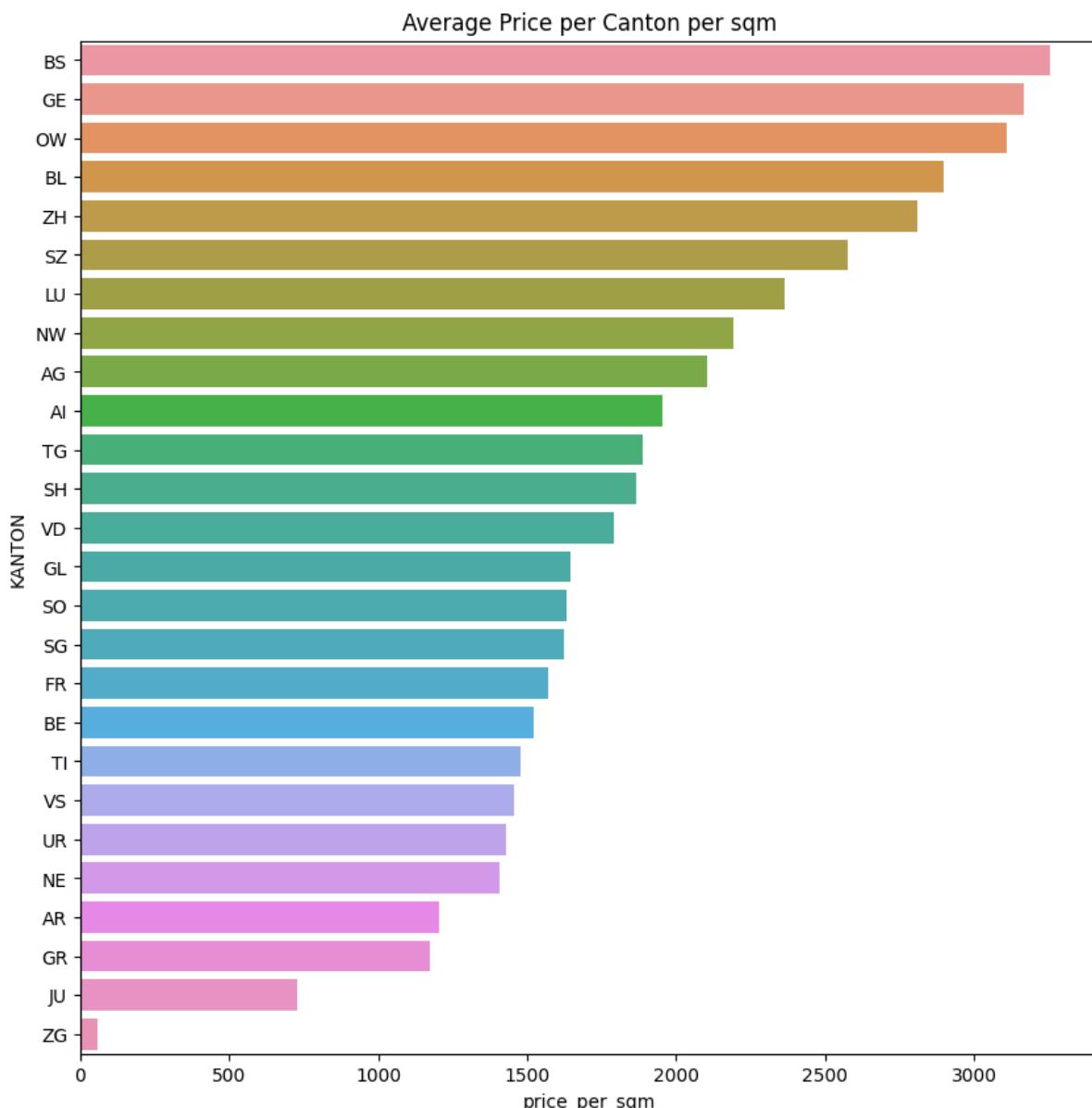
In diesem Histogramm wird der Preis pro Quadratmeter für Immobilien dargestellt. Die meisten Immobilien haben einen Quadratmeterpreis zwischen 0 und 4.000, wobei der Höhepunkt zwischen 2.000 und 4.000 liegt. Dies gibt uns einen Einblick in den durchschnittlichen Quadratmeterpreis der Immobilien im betrachteten Markt. Es gibt nur wenige Immobilien mit einem Quadratmeterpreis von über 8.000, was auf Premium-Immobilien oder Standorte mit hoher Nachfrage hinweisen könnte.

Durchschnittspreis in Mio. pro Gemeinde



- **Montreux:** Diese Gemeinde hat den höchsten durchschnittlichen Preis unter allen aufgeführten Gemeinden. Die Immobilien in Montreux sind im Durchschnitt am teuersten.
- **Ardon:** Im Gegensatz zu Montreux hat Ardon den niedrigsten durchschnittlichen Preis. Das bedeutet, dass die Immobilien in Ardon im Vergleich zu anderen Gemeinden günstiger sind.
- **Lausanne, Lugano und Nyon:** Diese Städte liegen im oberen Bereich des Diagramms, was bedeutet, dass ihre Immobilienpreise ebenfalls relativ hoch sind.

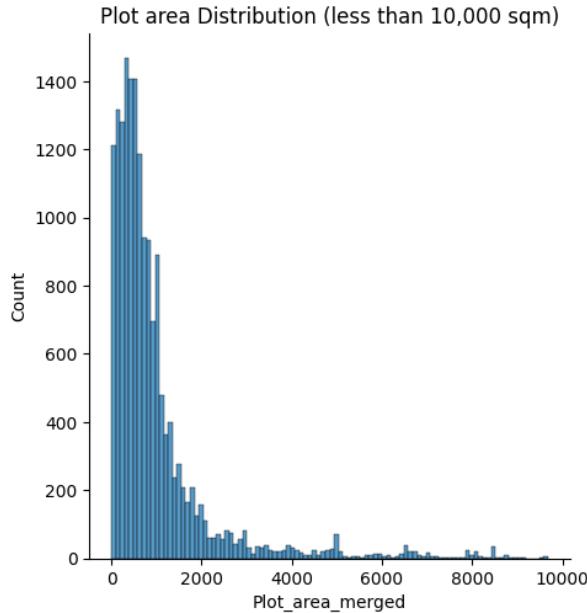
Durchschnittspreis pro Kanton pro Quadratmeter



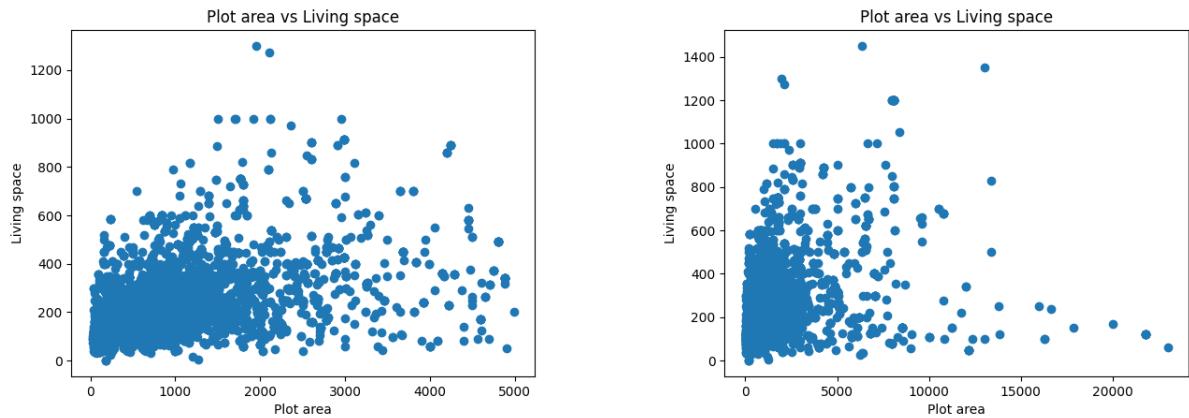
- **BS (Basel-Stadt):** Von allen Kantonen hat Basel-Stadt den höchsten Durchschnittspreis pro Quadratmeter. Das bedeutet, dass der Quadratmeter Wohnfläche in diesem Kanton im Durchschnitt am teuersten ist.
- **GE (Genf) und OW (Obwalden):** Diese beiden Kantone folgen Basel-Stadt in der Rangliste. Das bedeutet, dass sie auch relativ hohe Quadratmeterpreise haben, obwohl sie nicht so teuer sind wie Basel-Stadt.
- **ZG (Zug):** Zug liegt am unteren Ende des Diagramms. Dies bedeutet, dass der Durchschnittspreis pro Quadratmeter in Zug im Vergleich zu anderen Kantonen niedriger ist.

3.3 Raumanalyse (Space)

Raum beeinflusst unseren Alltag, unser Wohlbefinden und unsere Interaktionen. Mit wachsenden Städten und sich verändernden Lebensstilen ist es wichtig, den Raum genauer zu betrachten. In dieser Analyse untersuchen wir Raumdaten, um Trends und Muster zu erkennen und besser zu verstehen, wie Raum in unserem Leben verwendet wird.

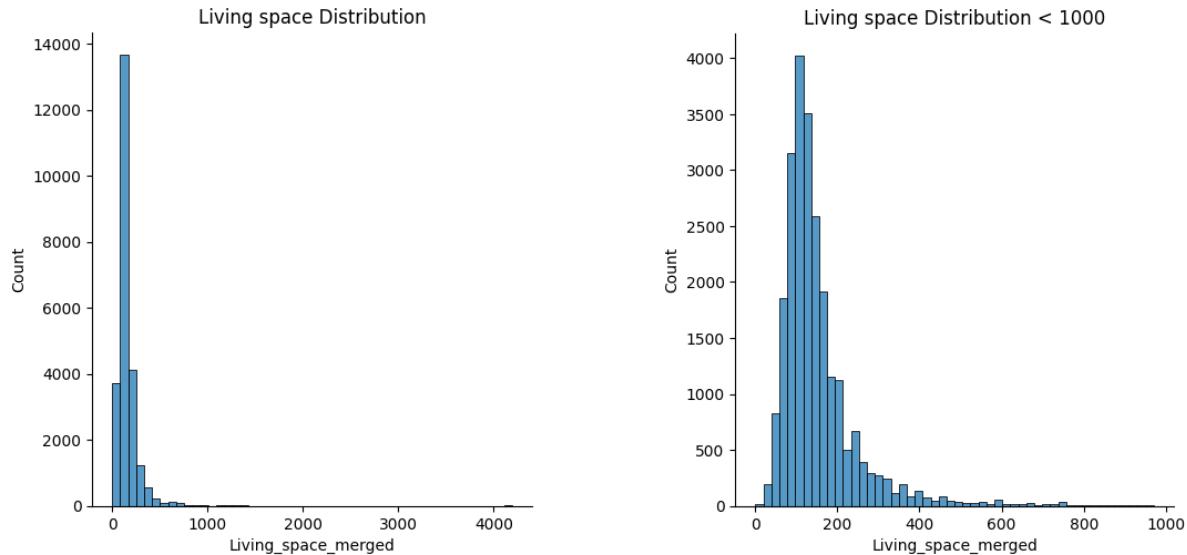


Der Plot zeigt die Verteilung von Immobilien nach ihrer Grundstücksfläche für Grundstücke mit einer Fläche von weniger als 10.000 Quadratmetern. Die meisten Immobilien scheinen eine Fläche von weniger als 2.000 Quadratmetern zu haben.



Im linken Plot wird der Zusammenhang zwischen der Grundstücksfläche und der Wohnfläche für Immobilien mit einer Grundstücksfläche von bis zu 5.000 Quadratmetern dargestellt.

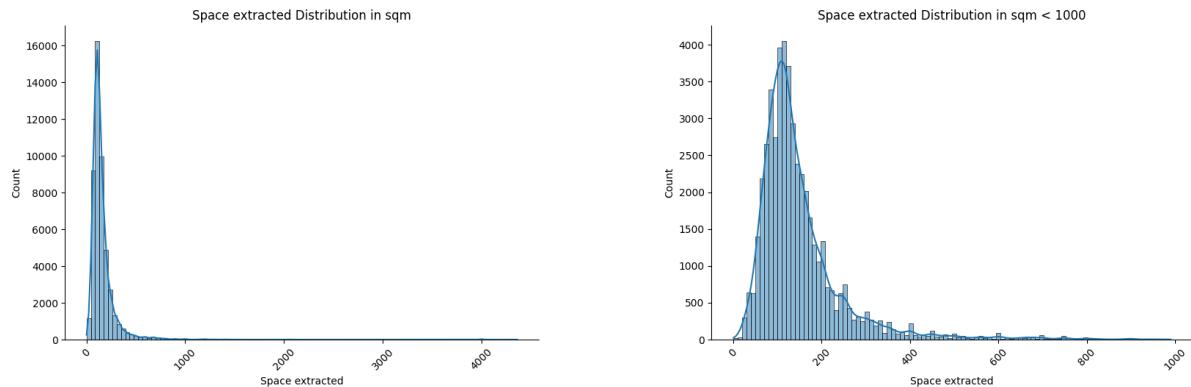
Der rechte Plot bietet eine erweiterte Ansicht des Zusammenhangs zwischen Grundstücksfläche und Wohnfläche, die auch Immobilien mit grösseren Grundstücksflächen umfasst.



Im linken Diagramm ist die Gesamtverteilung des Wohnraums dargestellt. Auffällig ist hier, dass ein grosser Teil der Daten einen Wohnraum von nahezu 0 Quadratmetern aufweist. Dies könnte auf fehlende oder fehlerhafte Daten hinweisen, da solch ein geringer Wohnraum in der Realität unwahrscheinlich ist. Jenseits dieses Peaks nimmt die Anzahl der Objekte mit zunehmender Wohnfläche deutlich ab.

Das rechte Diagramm konzentriert sich auf Objekte mit einer Wohnfläche von weniger als 1000 Quadratmetern. Die Verteilung zeigt hier einen deutlichen Peak im Bereich von 200 bis 300 Quadratmetern. Dies deutet darauf hin, dass ein signifikanter Teil der Objekte in dieser Wohnraumkategorie liegt. Danach nimmt die Anzahl der Objekte mit steigendem Wohnraum stetig ab.

Beide Diagramme bieten wertvolle Einblicke in die Verteilung des Wohnraums und können bei der weiteren Analyse von Immobilien und ihrer Preisbildung berücksichtigt werden.

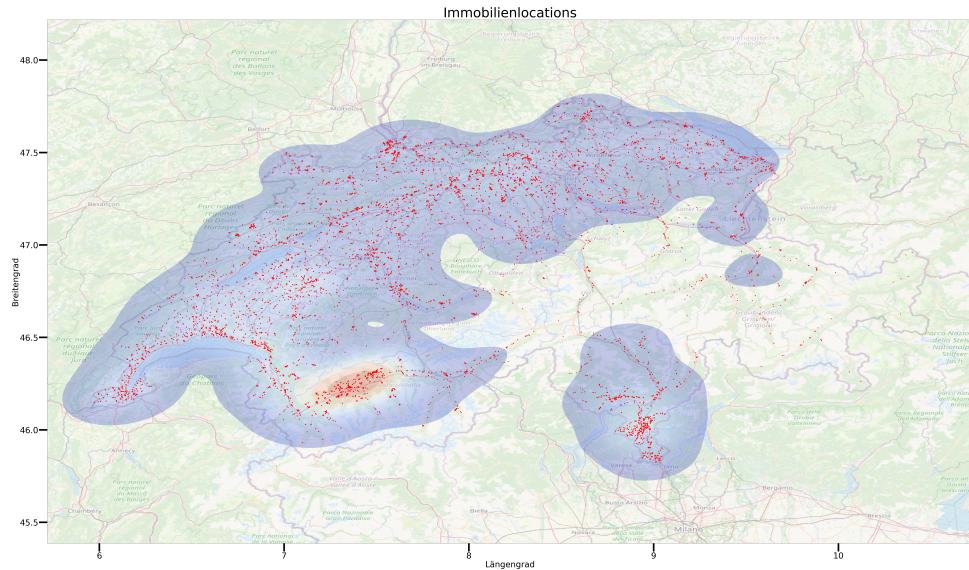


Im ersten Diagramm wird die gesamte Verteilung des extrahierten Wohnraums dargestellt. Ein markanter Peak ist nahe dem Wert von 0 Quadratmetern zu beobachten, was auf eine hohe Anzahl von Einträgen mit sehr kleinem oder fehlendem Wohnraum hinweist. Abseits dieses Peaks nimmt die Anzahl der Einträge mit steigendem Wohnraum rasch ab.

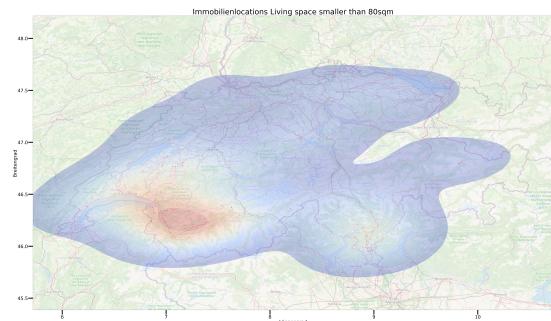
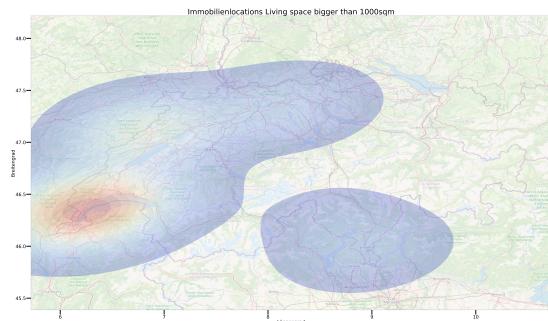
Das zweite Diagramm gibt einen detaillierteren Einblick in die Verteilung des Wohnraums für Flächen unter 1000 Quadratmetern. Es zeigt einen dominanten Peak im Bereich von 200 bis 300 Quadratmetern, was darauf hinweist, dass ein erheblicher Teil der Einträge in dieser Kategorie von Wohnraumgrösse liegt. Jenseits dieses Bereichs verringert sich die Anzahl der Einträge stetig.

Zusammenfassend geben diese Diagramme einen klaren Überblick über die Verteilung des extrahierten Wohnraums und können für weitergehende Analysen und Interpretationen herangezogen werden.

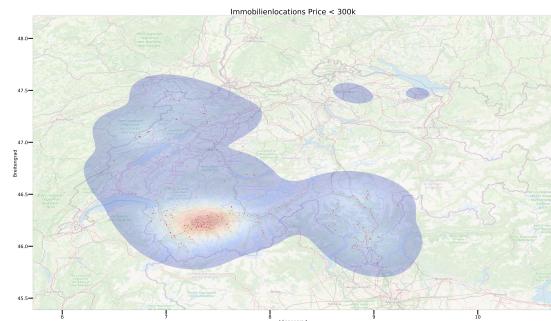
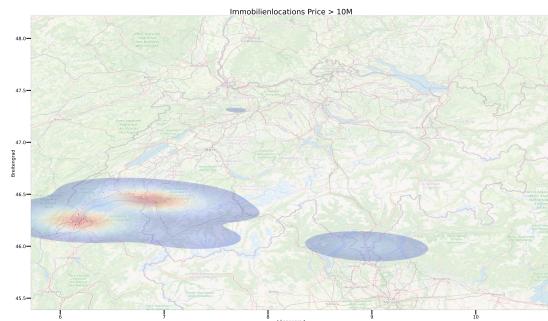
3.4 Standortanalyse



Die geografische Lage von Immobilien ist ein entscheidender Faktor, der ihren Wert und ihre Attraktivität beeinflusst. In dieser Analyse betrachten wir die Verteilung von Immobilienstandorten in Bezug auf ihre Wohnfläche und Preiskategorien in der Schweiz. Die Standortanalyse ermöglicht es uns, regionale Schwerpunkte und Muster im Immobilienmarkt zu identifizieren. Die Karte gibt einen allgemeinen Überblick über die Standorte aller erfassten Immobilien, unabhängig von ihrer Grösse oder ihrem Preis. Dies bietet eine umfassende Perspektive auf die Verteilung von Immobilien über die betrachtete Region.



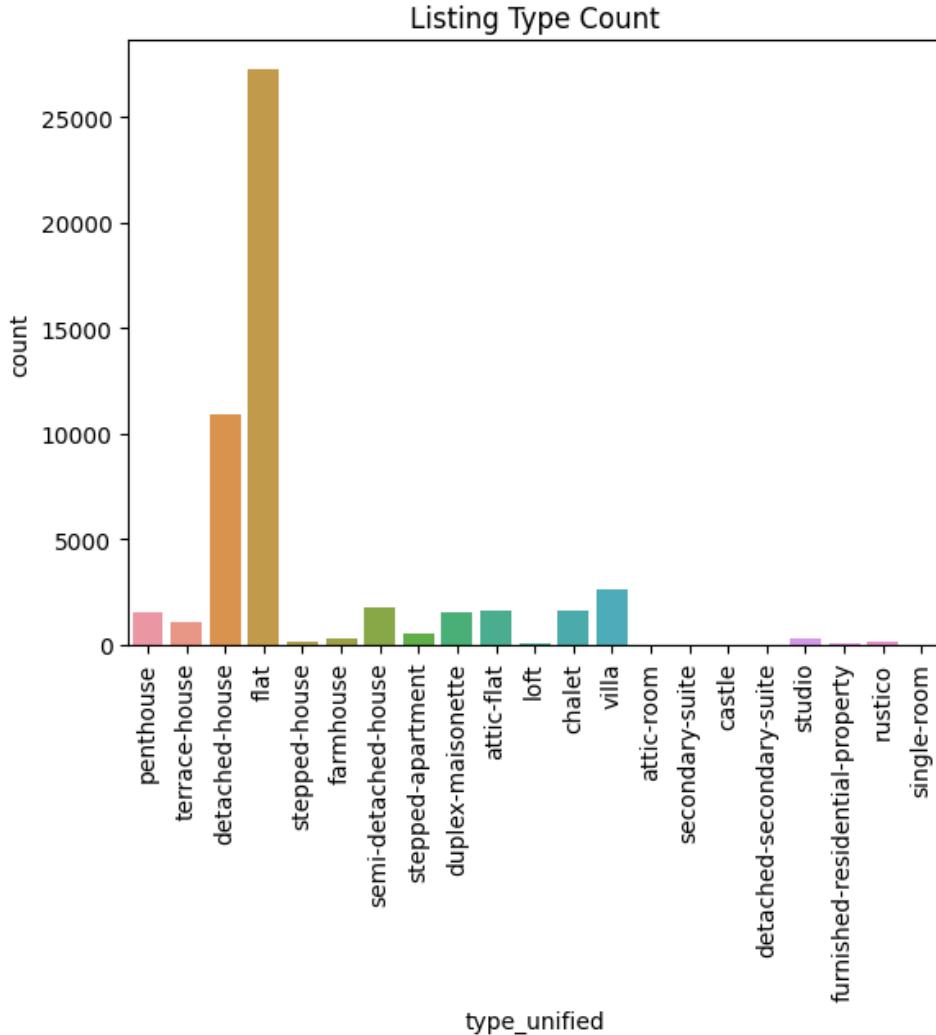
Die obigen beiden Karten zeigen die Verteilung von Immobilien mit einer Wohnfläche von mehr als 1000 Quadratmetern bzw. weniger als 80 Quadratmetern. Dunklere Farbtöne auf der Karte können auf eine höhere Dichte von Immobilien in diesen Größenkategorien hinweisen, was möglicherweise auf urbane Zentren oder auf Regionen mit spezifischen Wohnformen hindeutet.



Die obigen zwei Karte stellen die Standorte von Immobilien dar, deren Preis über 10 Millionen bzw. unter 300.000 liegt. Diese visualisierten Preisbereiche können wichtige Informationen über die Verteilung von Luxusimmobilien im Vergleich zu eher erschwinglichen Wohnobjekten liefern.

3.5 Analyse des Immobilientyps (Listing Type)

Die Klassifizierung von Immobilien nach ihrem Typ ist ein wesentlicher Faktor für das Verständnis des Marktes. Die folgenden Diagramme zeigen die Verteilung der verschiedenen Immobilientypen und deren Zusammenhang mit kommunalen Steuersätzen.



Das Balkendiagramm zeigt die Anzahl der Einträge für jeden Immobilientyp im Datensatz. Es wird deutlich, dass bestimmte Typen wie Wohnungen und Einfamilienhäuser am häufigsten vertreten sind, während andere, wie Bauernhäuser oder Villen, seltener vorkommen.

4 Schlussfolgerung

Die vorliegende explorative Datenanalyse liefert wertvolle Einblicke in den Schweizer Immobilienmarkt. Durch die Untersuchung von Preisverteilungen, Quadratmeterpreisen, Wohnflächen und Immobilientypen konnten wir signifikante Muster und regionale Unterschiede aufdecken. Es wurde deutlich, dass die meisten Immobilien zu verhältnismässig erschwinglichen Preisen angeboten werden, mit einigen wenigen hochpreisigen Ausnahmen. Die Analyse der Grundstücks- und Wohnfläche zeigte eine Konzentration kleinerer Immobilien, was auf eine dichtere Bebauung in städtischen Gebieten hindeutet.

Die Standortanalyse unterstreicht die Bedeutung der Lage, die sowohl durch die Dichte bestimmter Immobilientypen als auch durch Preisunterschiede auf Kantonsebene belegt wird.

Insgesamt bietet der Datensatz eine solide Grundlage für weiterführende Analysen, die zur Entwicklung von Marktprognosen und zur Beratung von Investitionsentscheidungen dienen können. Zukünftige Studien sollten insbesondere die hohe Rate fehlender Daten berücksichtigen, um die Genauigkeit und Aussagekraft der Ergebnisse zu verbessern.