# Solutions\_Week\_01

## Task 1

In this exercise, you must define business objectives and hypotheses for the two use cases defined below. Prepare a short description (bullet points, short notes) by using the following structure:

1. Problem definition (identify one or more potential problem(s)).

2. Objective (define a clear objective based on the problem definition).

3. Research question (formulate a research question based on the objective).

4. Hypothesis (formulate a clear hypothesis based on the research question).

5. Expected added value of the data analysis (business value of data analytics).

**Use Case 1: Preisvorhersagemodell für Mietwohnungen im Kanton Zürich**

1. **Problemdefinition**:
   * Schwankende Mietpreise erschweren es Mietern und Vermietern, faire Preise zu bestimmen.
   * Mietpreise können durch viele Faktoren beeinflusst werden (Lage, Größe, Baujahr, etc.).
   * Unvorhersehbare Preisentwicklungen könnten zu Leerständen oder überteuerten Wohnungen führen.
2. **Zielsetzung**:
   * Entwicklung eines Modells zur Vorhersage der Mietpreise von Wohnungen im Kanton Zürich, basierend auf verschiedenen Faktoren (Lage, Ausstattung, Größe, etc.).
3. **Forschungsfrage**:
   * Welche Faktoren beeinflussen die Mietpreise von Wohnungen im Kanton Zürich am stärksten, und wie kann ein Modell entwickelt werden, das die Mietpreise präzise vorhersagt?
4. **Hypothese**:
   * Die Mietpreise im Kanton Zürich werden primär durch die Lage (Postleitzahl) und die Größe der Wohnung bestimmt, während zusätzliche Faktoren wie Baujahr und Ausstattung einen sekundären Einfluss haben.
5. **Erwarteter Mehrwert durch die Datenanalyse**:
   * Erhöhung der Markttransparenz für Mieter und Vermieter.
   * Unterstützung bei der Preisbildung für Immobilienunternehmen und private Vermieter.
   * Optimierung der Angebotsgestaltung basierend auf aktuellen Marktdaten.

**Use Case 2: Nachfrage und Angebot in Supermärkten in Schweizer Gemeinden**

1. **Problemdefinition**:
   * Diskrepanz zwischen der Nachfrage und dem Angebot an Supermärkten in verschiedenen Schweizer Gemeinden.
   * Mögliche Über- oder Unterversorgung in bestimmten Regionen.
   * Ineffiziente Standortplanung von Supermärkten basierend auf unzureichenden Daten.
2. **Zielsetzung**:
   * Analyse des Angebots und der Nachfrage in Supermärkten in Schweizer Gemeinden, um mögliche Versorgungslücken oder -überschüsse zu identifizieren und das Angebot gezielter anzupassen.
3. **Forschungsfrage**:
   * Wie korreliert die Nachfrage nach Supermarktprodukten mit der Anzahl der Supermärkte und deren Verteilung in Schweizer Gemeinden, und welche Optimierungen lassen sich daraus ableiten?
4. **Hypothese**:
   * In kleineren, ländlichen Gemeinden gibt es eine Unterversorgung an Supermärkten, während in städtischen Regionen eine Überversorgung besteht, was zu ineffizienter Ressourcenverteilung führt.
5. **Erwarteter Mehrwert durch die Datenanalyse**:
   * Effizientere Standortplanung für Supermärkte basierend auf tatsächlicher Nachfrage.
   * Bessere Auslastung bestehender Supermärkte und Reduzierung von Versorgungslücken.
   * Verbesserung der Kundenzufriedenheit durch eine bedarfsgerechte Versorgung in allen Regionen.

## Task 2

**Use Case 1: Preisvorhersagemodell für Mietwohnungen im Kanton Zürich**

**Notwendige Datenquellen:**

1. **Wohnungsinserate von ImmoScout24 (**[**https://www.immoscout24.ch**](https://www.immoscout24.ch)**)**:
   * **Relevante Datenfelder**:
     + Mietpreis
     + Wohnfläche (m²)
     + Anzahl Zimmer
     + Lage (Adresse, Postleitzahl)
     + Wohnungstyp (Wohnung, Haus, Loft etc.)
     + Baujahr des Gebäudes
     + Zustand der Wohnung (renoviert, neuwertig, etc.)
     + Stockwerk
     + Verfügbare Ausstattung (Balkon, Terrasse, Garage, etc.)
   * **Beschreibung**: Diese Daten bieten einen detaillierten Überblick über die verschiedenen Parameter, die den Mietpreis beeinflussen können. Sie ermöglichen es, Preisvergleiche anzustellen und Trends zu identifizieren.
2. **Gemeindedaten von der SFSO (Regionalporträts 2021: Kennzahlen aller Gemeinden)**:
   * **Relevante Datenfelder**:
     + Bevölkerungsdichte
     + Durchschnittliches Haushaltseinkommen
     + Arbeitslosenquote
     + Steuerbelastung
     + Infrastruktur (Schulen, öffentliche Verkehrsmittel, Krankenhäuser)
   * **Beschreibung**: Diese Daten auf Gemeindeebene können dabei helfen, die Attraktivität bestimmter Regionen zu bewerten, was sich direkt auf den Mietpreis auswirken kann.
3. **Zusätzliche Geodaten (optional)**:
   * **Relevante Datenfelder**:
     + Entfernung zum nächsten öffentlichen Verkehrsmittel (Bahnhof, Bus)
     + Entfernung zu wichtigen Einrichtungen (Einkaufsmöglichkeiten, Schulen)
   * **Beschreibung**: Die Verfügbarkeit von Verkehrsanbindungen und wichtigen Einrichtungen kann den Mietpreis signifikant beeinflussen.

**Use Case 2: Nachfrage und Angebot in Supermärkten in Schweizer Gemeinden**

**Notwendige Datenquellen:**

1. **Supermarktdaten von OpenStreetMap (**[**https://wiki.openstreetmap.org/wiki/DE:Key:shop**](https://wiki.openstreetmap.org/wiki/DE:Key:shop)**)**:
   * **Relevante Datenfelder**:
     + Standort der Supermärkte (Koordinaten, Adresse)
     + Supermarktgröße (falls verfügbar)
     + Öffnungszeiten
     + Verfügbare Dienstleistungen (Online-Shop, Lieferservice)
   * **Beschreibung**: Diese Daten bieten einen Überblick über die Verteilung von Supermärkten in der Schweiz und können genutzt werden, um das Angebot in verschiedenen Gemeinden zu analysieren.
2. **Gemeindedaten von der SFSO (Regionalporträts 2021: Kennzahlen aller Gemeinden)**:
   * **Relevante Datenfelder**:
     + Bevölkerungsdichte
     + Altersstruktur der Bevölkerung
     + Durchschnittliches Haushaltseinkommen
     + Einkaufsgewohnheiten (falls verfügbar)
     + Mobilität (Anteil der Bevölkerung, der auf öffentliche Verkehrsmittel angewiesen ist)
   * **Beschreibung**: Diese Daten helfen, die Nachfrage in verschiedenen Gemeinden zu quantifizieren. Zum Beispiel könnten ländliche Gemeinden mit einer alternden Bevölkerung eine geringere Nachfrage nach großen Supermärkten haben, während städtische Regionen möglicherweise eine größere Nachfrage haben.
3. **Zusätzliche Daten zu Konsumtrends (optional)**:
   * **Relevante Datenfelder**:
     + Regionale Konsumtrends (z.B. bevorzugte Einkaufstage, durchschnittlicher Warenkorb)
     + Anteil von Online-Einkäufen im Lebensmittelbereich
   * **Beschreibung**: Diese Daten helfen dabei, das Konsumverhalten besser zu verstehen und die Nachfrage in verschiedenen Gemeinden realistischer einzuschätzen.