

# Kamień milowy 1

z projektu Hurtowni Danych

Maciej Szpetmański, Franciszek Szczepaniak

## Cel biznesowy

**Nasza hurtownia danych łącząca w sobie dane o restauracjach razem z danymi o zamożności i stanie zdrowia mieszkańców poszczególnych miast i dzielnic może mieć wielorakie zastosowania biznesowe. Do tych, jak nam się wydaje, posiadających największy potencjał biznesowy należą m.in.:**

**Optymalizacja lokalizacji restauracji:** Wykorzystując dane o zamożności i stanie zdrowia mieszkańców poszczególnych miast, hurtownia danych może pomóc restauratorom w identyfikacji najlepszych lokalizacji dla nowych restauracji, uwzględniając preferencje żywieniowe oraz dostępność potencjalnych klientów.

**Segmentacja rynku:** Analizując dane o zachowaniach konsumenckich w różnych miastach oraz ich związek z zamożnością i stanem zdrowia, można wyodrębnić segmenty klientów i dostosować ofertę restauracji do ich potrzeb i preferencji.

**Rozwój zdrowych opcji menu:** Na podstawie danych o stanie zdrowia mieszkańców oraz ich preferencjach żywieniowych, można opracować zdrowsze opcje menu dla restauracji, które będą lepiej dopasowane do lokalnych potrzeb i trendów.

**Profilowanie klientów:** Analiza danych może pomóc restauratorom w lepszym zrozumieniu swoich klientów, ich preferencji żywieniowych, wydatków oraz nawyków konsumenckich, co pozwoli na bardziej skuteczną personalizację oferty i marketingu.

**Prognozowanie trendów gastronomicznych:** Śledzenie zmian w konsumpcji żywności i preferencjach kulinarnych wraz ze zmianami w zamożności i stanie zdrowia populacji może pomóc w prognozowaniu trendów gastronomicznych i odpowiednim dostosowaniu oferty restauracji.

**Partnerstwa z firmami ubezpieczeniowymi lub programami zdrowotnymi:** Analiza danych o stanie zdrowia mieszkańców może być wartościowa dla firm ubezpieczeniowych lub programów zdrowotnych, które mogą chcieć nawiązać współpracę z restauracjami w promowaniu zdrowych nawyków żywieniowych.

**Optymalizacja dostaw:** Wykorzystując dane o lokalizacji restauracji i charakterystyce populacji, hurtownia danych może pomóc w optymalizacji tras dostaw, co przyczyni się do obniżenia kosztów operacyjnych restauracji.

# Architektura Systemu

Pierwszym krokiem było znalezienie odpowiadającym nam danych, które w łatwy sposób można by wykorzystać zgodnie z celem biznesowym.

Następnie za pomocą SSIS w Visual Studio przetwarzamy dane, które są ładowane do modelu hurtowni w SSMS. Do wizualizacji użyjemy PowerBI.

## Opis zbiorów danych

W projekcie używamy dwóch zbiorów danych:

**'Restaurant Data with Consumer Ratings'** -

<https://www.kaggle.com/datasets/uciml/restaurant-data-with-consumer-ratings>

**'mexico\_housing\_datasets – Mexico-real-estate-clean'** -

<https://www.kaggle.com/datasets/sasumwen/dataset?select=mexico-real-estate-clean.csv>

### Restaurant Data with Consumer Ratings:

Restaurants:

- 1 chefmozaccepts.csv – jakie metody płatności akceptuje restauracja – 1314 wierszy
- 2 chefmozcuisine.csv – jakiego rodzaju kuchni są potrawy w restauracji – 916 wierszy
- 3 chefmozparking.csv – jaki parking ma restauracja – 702 wiersze
- 4 geoplaces2.csv – dane dotyczące lokalizację restauracji – 130 wierszy

Consumers:

- 6 usercuisine.csv – preferencje użytkowników co do kuchni – 330 wierszy
- 7 userpayment.csv – jakimi metodami płatności dysponuje użytkownik – 177 wierszy
- 8 userprofile.csv – cechy użytkowników (zamieszkanie, czy pali, jak dużo pije w restauracjach, preferowany rodzaj ubioru, czym przyjeżdża do restauracji, status związku) – 138 wierszy

User-Item-Rating:

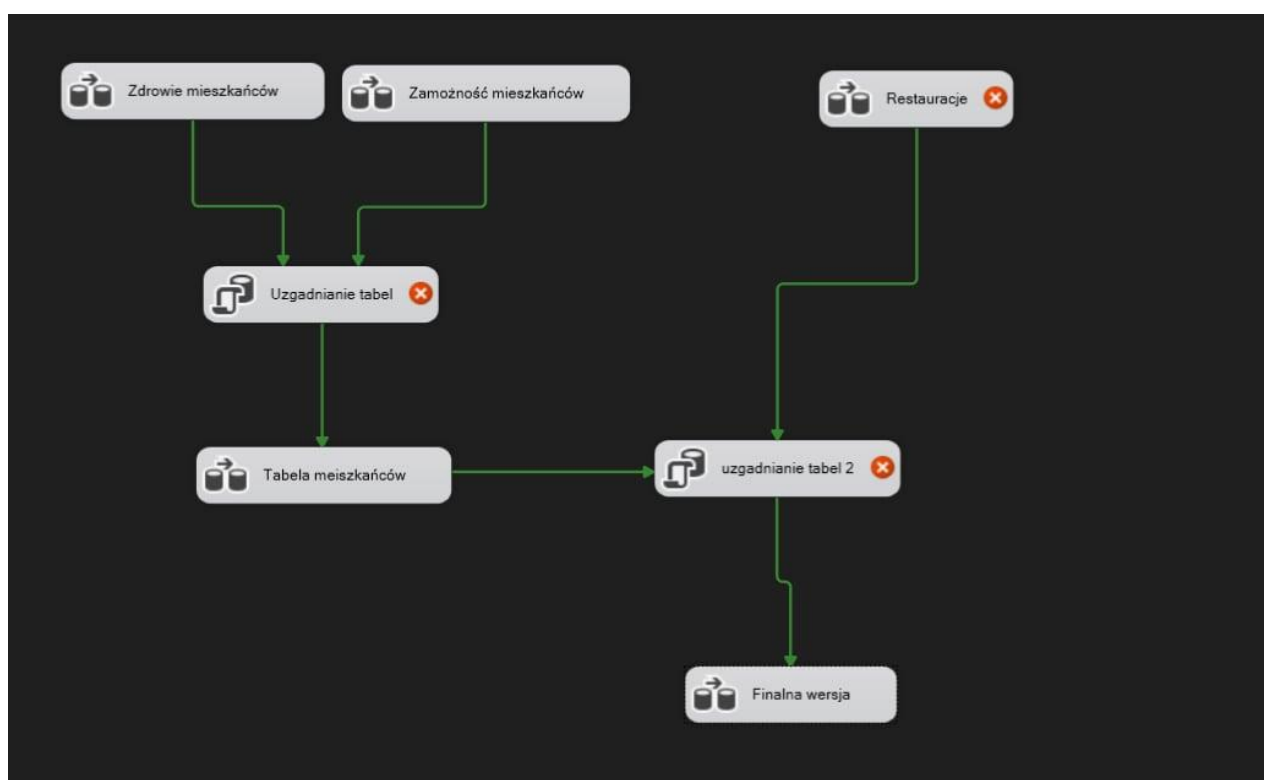
- 9 rating\_final.csv – opinie użytkowników zostawione w restauracjach (ogólna ocena, obsługa, jedzenie, data opinii) – 1161 wierszy

Opinie są u nas tabelą faktów, łączącą tabele wymiarów użytkowników i restauracji. Mają kolumnę z datą. Można odświeżać opinie co weekend, patrząc na nowe recenzje. Dzięki temu możemy zrobić tygodniowy ranking restauracji lub patrzeć na sezonowe trendy.

### Mexico- real-estate-clean:

Jedna tabela zawierająca położenie nieruchomości, jej cenę oraz typ (czy to dom, apartament, blok itd.)

## ETL

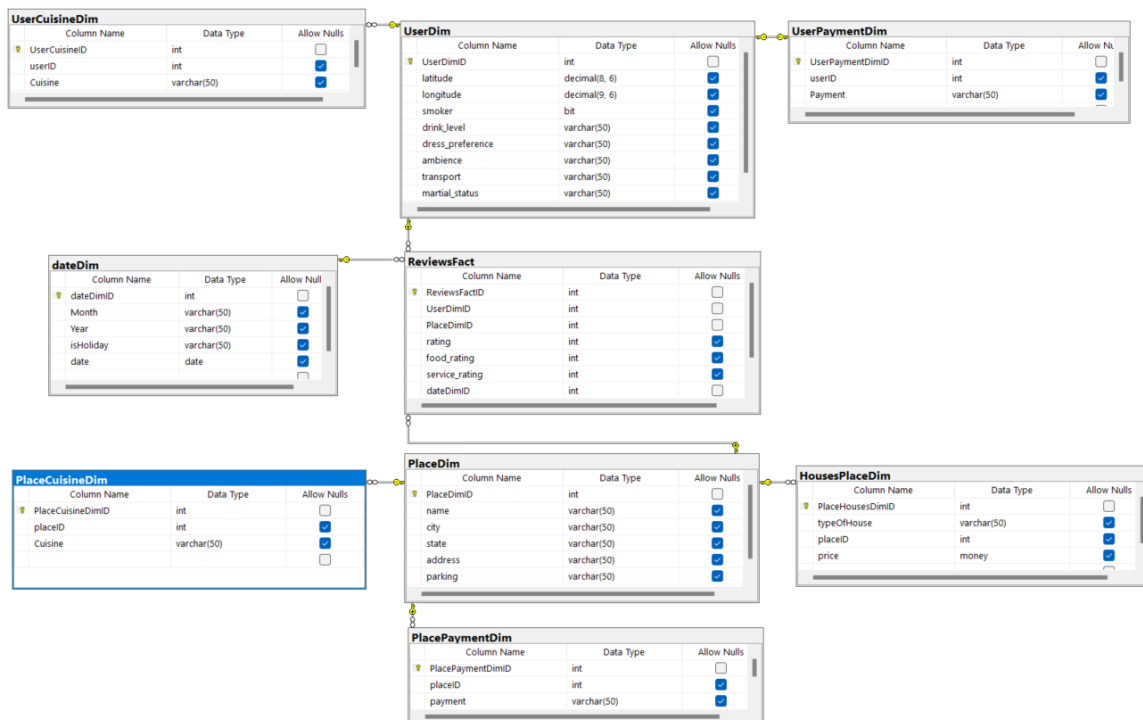


Na podstawie tabeli Mexico-real-estate-clean, będziemy znajdowali średnie ceny nieruchomości w pewnym wybranym promieniu od restauracji. Na podstawie tego zobaczymy stopień zamożności miast czy dzielnic w okolicach restauracji. Zamożność mieszkańców będzie również powiązana ze zdrowiem mieszkańców i ich nawykami (palenie, picie itd.).

Co do aktualizacji SCD1 możemy przyjąć, że każdy użytkownik może zmienić swoją opinię. Wtedy stara opinia jest zastąpiona przez nową.

SCD2 można wykonać, aktualizując ogólną ocenę restauracji w danym np. weekendzie. Wtedy można patrzeć na sezony, w których restauracja dostaje najlepsze opinie, a w których te gorsze.

# Model Hurtowni



## Pomysły na raport

### 1. Wykres kołowy z podziałem na rodzaje restauracji:

- Ten wykres może pokazać procentowy udział różnych rodzajów restauracji w danej lokalizacji. Możemy zobaczyć, ile restauracji jest typu fast food, fine dining, food trucki itp.

### 2. Wykres słupkowy średnich wydatków na posiłek w zależności od zamożności miasta:

- Ten wykres może porównać średnie wydatki na posiłek w różnych miastach, grupując je według poziomu zamożności mieszkańców. Może to pomóc restauratorom w dostosowywaniu cen do lokalnych warunków ekonomicznych.

### 3. Mapa cieplna z ocenami restauracji w różnych dzielnicach miasta:

- Taki wykres może zobrazować, które dzielnice miasta mają najwyższe lub najniższe oceny restauracji. Może to być przydatne dla osób planujących otwarcie nowej restauracji lub zmianę lokalizacji istniejącej.

4. **Wykres liniowy zmian w liczbie restauracji w czasie:**

- Ten wykres może przedstawić trend wzrostu lub spadku liczby restauracji w danym mieście w ciągu określonego okresu czasu. Może to być przydatne dla przedsiębiorców planujących otwarcie nowej restauracji lub inwestorów szukających trendów rynkowych.

5. **Wykres punktowy związku między stopniem zamożności a wskaźnikiem otyłości w danym mieście:**

- Taki wykres może zobrazować, czy istnieje związek między zamożnością mieszkańców a poziomem otyłości w danym mieście. Może to dostarczyć cennych wskazówek dla programów zdrowotnych lub strategii marketingowych restauracji.

6. **Histogram ocen restauracji:**

- Ten wykres może przedstawić rozkład ocen restauracji w danym mieście lub regionie. Może to pomóc w zrozumieniu ogólnego poziomu jakości usług gastronomicznych w danej lokalizacji.