**DOKUMENTACJA PROJEKTOWA**

**,,Programu przeprowadzającego analizę porównawczą cen w sklepach e-commerce"**

**Autorzy projektu:**

Paulina Rzeszutek

Karolina Czoska

**Opis projektu:**

Stworzony program ,,Porównywarka cenowa w sklepach e-commerce" jest narzędziem powstałym na bazie porównywarki cenowej. Ułatwia konsumentowi porównanie cen i asortymentu, a także wybór najtańszego sklepu spożywczego, w którym najkorzystniej jest robić zakupy pod względem kosztów (podstawowy koszyk produktów). Analizie zostały poddane sklepy: Barbora, Delio oraz Megasam24. Do utworzenia porównywarki cenowej wykorzystano różne rodzaje produktów spożywczych: cukier, kawę rozpuszczalną, mleko, majonez, wodę, olej oraz masło.

**Wymagania biznesowe:**

Celem projektu jest zachęcenie potencjalnego klienta do skorzystania z oferty sprzedażowej danego sklepu. Projekt przyniesie korzyść zarówno klientom poszukującym oszczędności oraz sklepom internetowym - zwiększając ich przychód jak i ruch na stronie.

**Wykorzystane narzędzia do stworzenia projektu (architektura technologiczna):**

* Web scraping
* Jupyther
* Visual Studio Code
* Microsoft Excel
* SQL Server Managnement Studio
* Python
* Html
* CSS

**Prawa autorskie**

Otwarte prawa autorskie.

**Architektura oraz zakres projektu:**

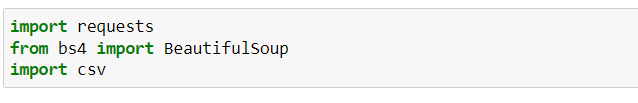
Informacje dotyczące wybranych produktów (z koszyka podstawowego) zostały pobrane za pomocą Jupythera i web scrapingu. Następnie dane zostały oczyszczone z błędów powstałych w wyniku scrapowania, przy pomocy Excela np. usunięcie produktów, które trafiły do złej kategorii. Uporządkowane pliki z informacjami o typach, nazwach, pojemności i cenach produktów zostały zaimportowane w formie tabel do SQL Server Management Studio. Utworzone zapytania w SQL Server Management Studio pozwoliły na dokonanie analizy porównawczej sklepów: asortymentu i cen. Następnie w Visual Studio Code został stworzony kod, dzięki któremu za pomocą flaska, pythona i html najważniejsze wybrane wnioski są wyświetlane na stronie. Użytkownik ma możliwość przeniesienia się do strony i z najtanszą ofertą.

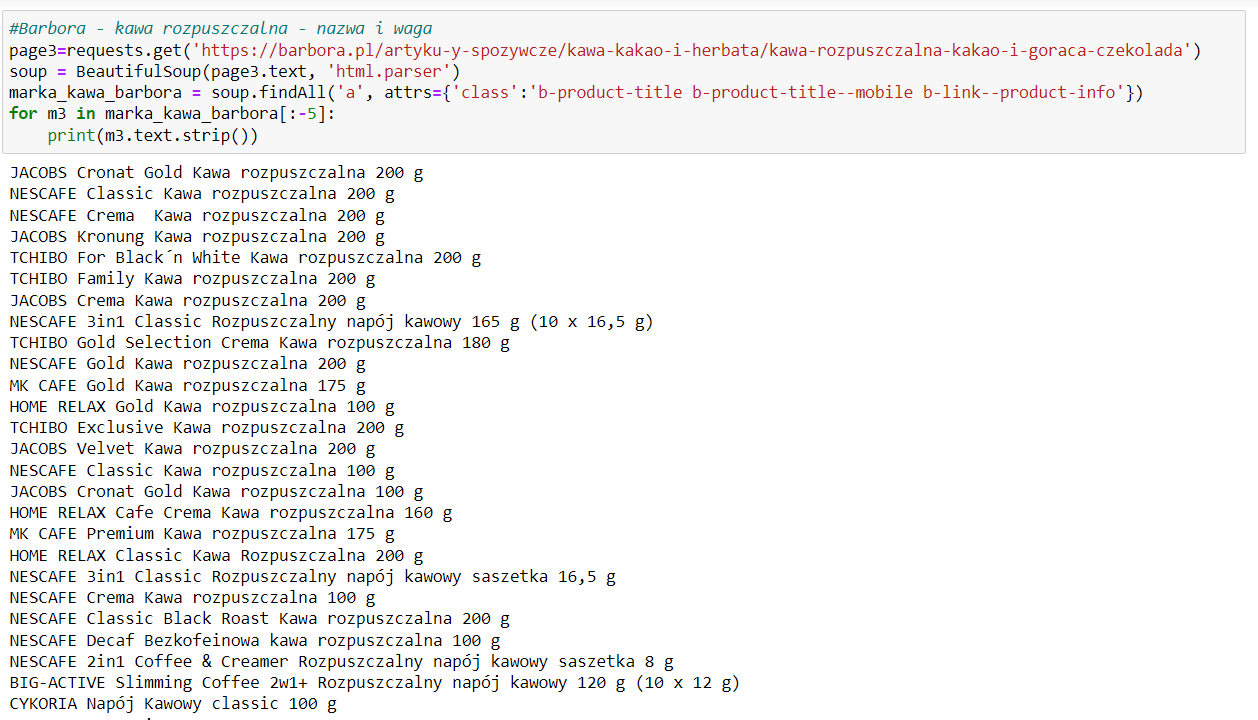
**Napotkane problemy podczas tworzenia projektu:**

Niektóre sklepy zablokowały możliwość scrapowania danych, co uniemożliwiło analizę cen i asortymentu. Pozostałe dane zostały wykorzystane zgodnie z Robots Exclusion Protocol.

**Etapy projektu:**

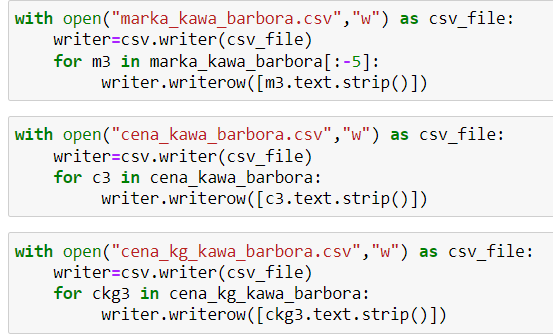
1. Pierwszym korkiem w wykonaniu projektu było zescrapowanie danych przy pomocy Jupyter Notebook. Poniżej został przedstawiony przykładowy kod, pobierający dane o kawie ze sklepu Barbora. Zostały pobrane informacje o nazwie, cenie, wadze produktu i cenie za kilogram.



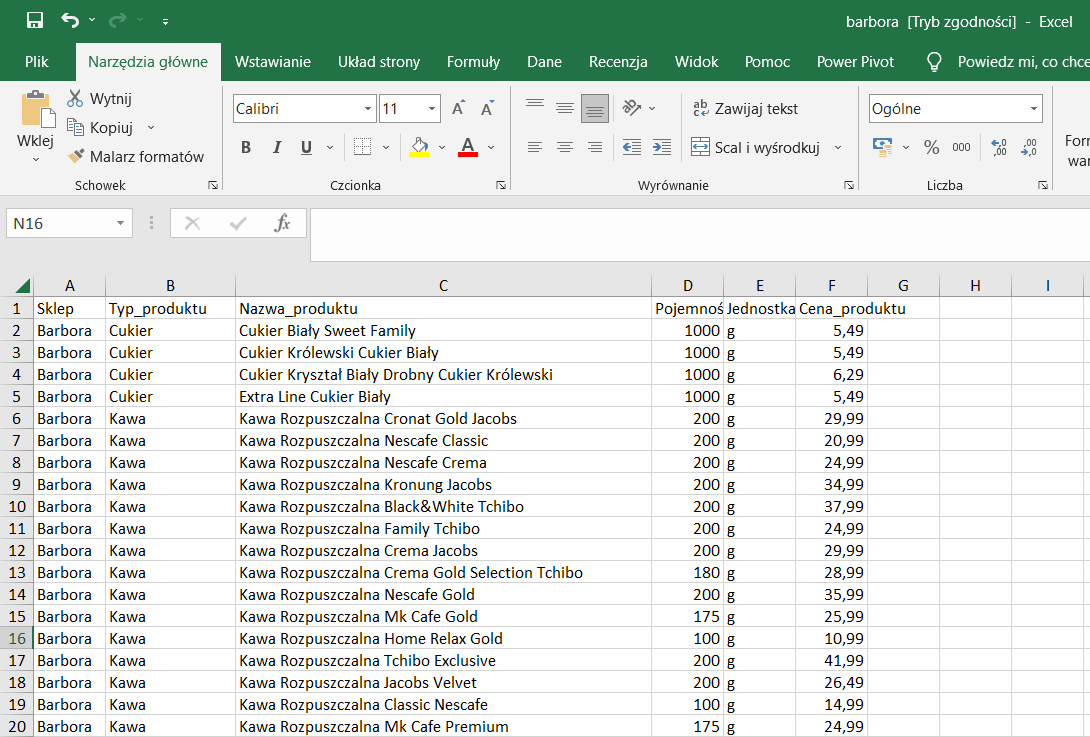


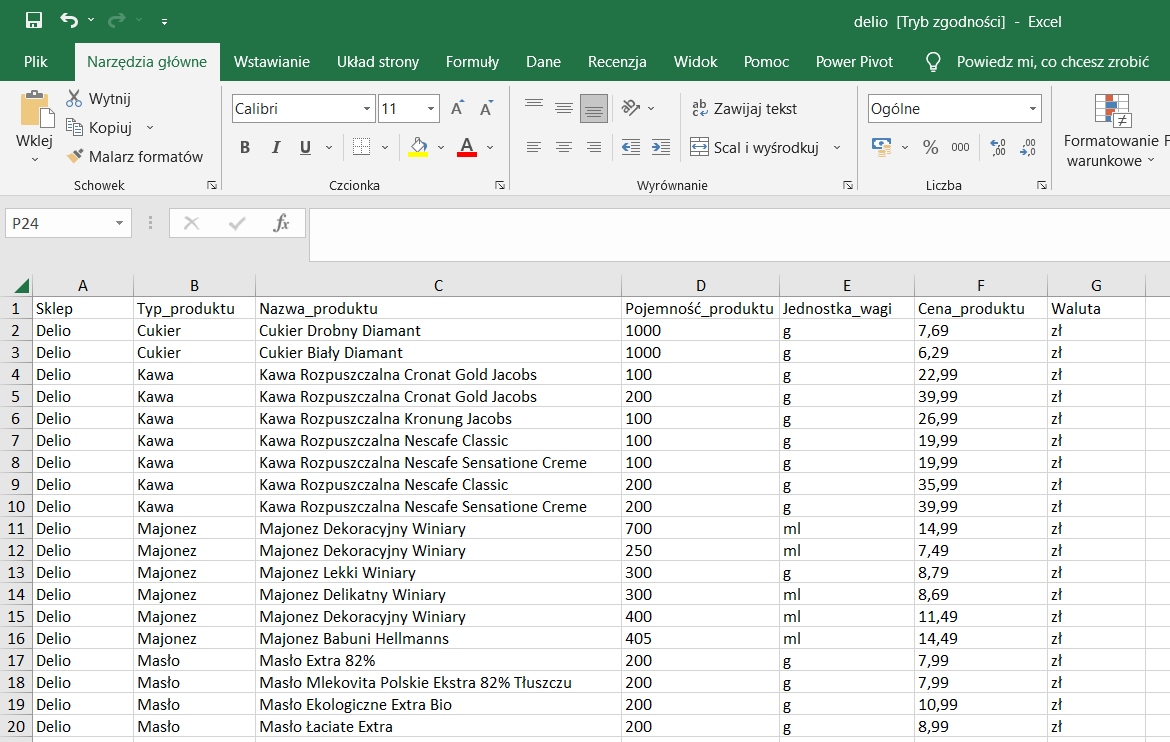


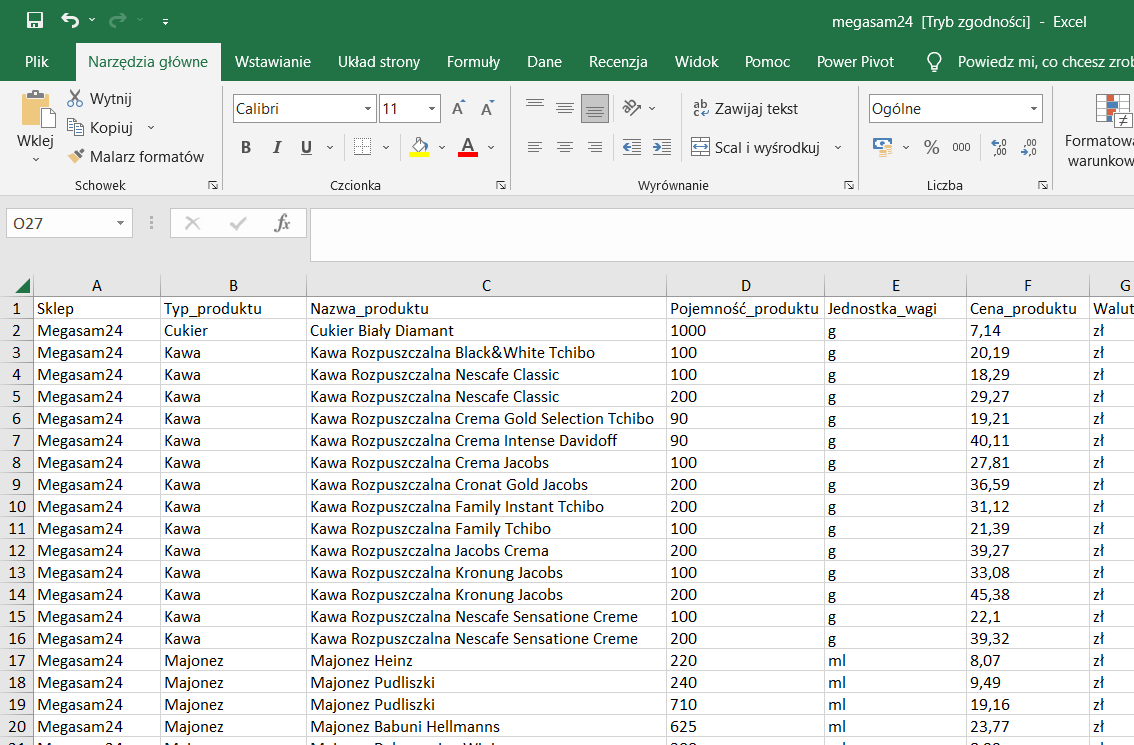




1. Zescrapowane dane zostały złączone w trzy tabele, po jednej tabeli przypadającej na dany sklep, i oczyszczone przy pomocy Excela. Z zescrapowanych 418 wierszy z informacjami o produktach, po oczyszczeniu pozostało 225 rekordów. Usunięciu zostały poddane dane, które były odmienne od oczekiwanych i nie pasowały do kategorii danych, co widać na powyżej zaprezentowanym przykładzie danych z kawą rozpuszczalną w sklepie Barbora, w której znalazły się też informacje o innych kawach lub produktach do picia. Poniżej zostały zaprezentowane oczyszczone trzy tabele z informacjami o asortymencie sklepów: typach, nazwach, pojemności i cenach produktów.



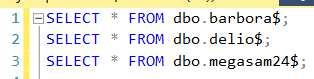


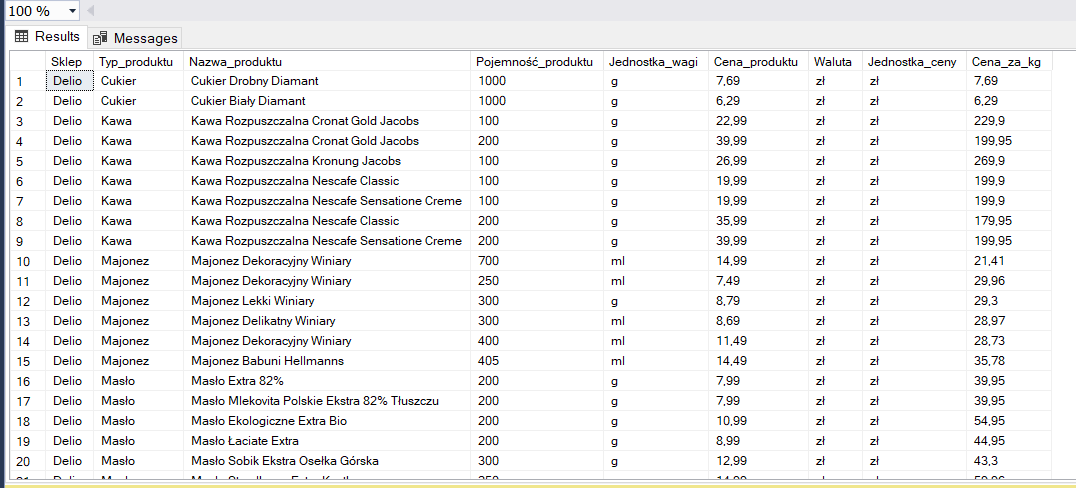


1. Uporządkowane pliki zostały zaimportowane do SQL Server Management Studio, w którym zostały poddane analizie. Poniżej zostały szczegółowo opisane zapytania użyte w bazie.

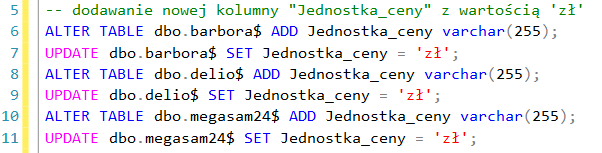
Najdroższym sklepem, który jednocześnie oferuje najszerszy wybór produktów z oferty jest sklep Megasam24. Natomiast najtańszym sklepem z najmniejszą ilością produktów jest Delio. Megasam24 ma w swojej ofercie 95 produktów, które kosztują średnio za kilogram 49,64 zł, a sklep Delio 47 produktów w cenie 44,85 zł za kilogram. Uwzględniając produkty, które znajdują się we wszystkich trzech sklepach, najtaniej wypada Barbora z średnią ceną za kilogram 30,26 zł, a najdrożej Delio z średnią ceną za kilogram 38,12 zł.

* Zapytanie umożliwia wyświetlenie zawartości zaimportowanych tabel z produktami, które znajdują się w sklepach.

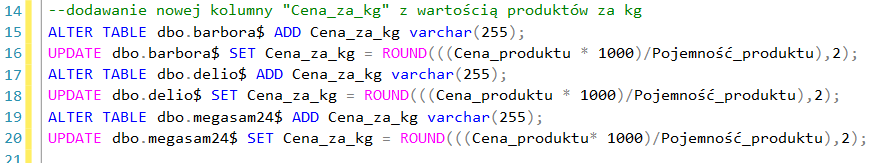




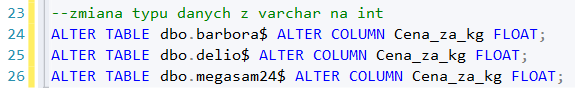
* Zapytanie tworzy nową kolumnę “Jednostka\_ceny”. Kolumna będzie zawierała wartości “zł”.



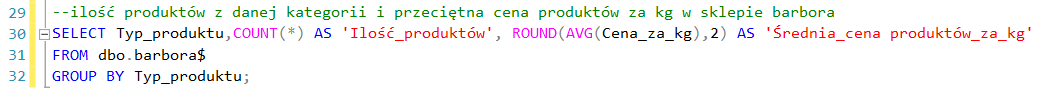
* Zapytanie tworzy nową kolumnę “Cena\_za\_kg “. W kolumnie będą znajdowały się informacje o cenie produktu w przeliczeniu na jeden kilogram.

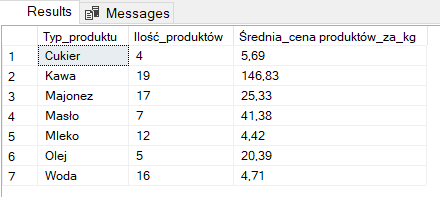


* Zapytanie zmienia typ danych z tekstowych na liczbowe.

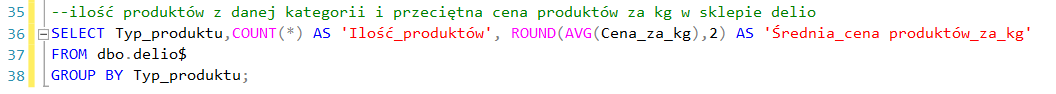


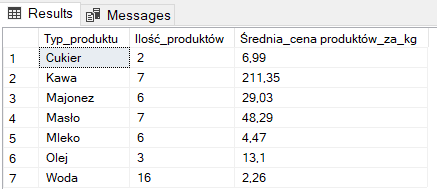
* Zapytanie wyświetla liczbę rodzajów produktów, znajdujących się w sklepie Barbora, a także przeciętną cenę za kg rodzajów produktów np. w sklepie Barbora znajduje się 19 produktów, które są kawą rozpuszczalną. Ich przeciętna cena za kg wynosi 146,83 zł.



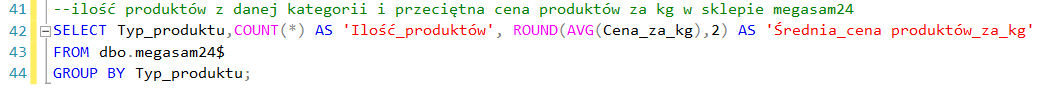


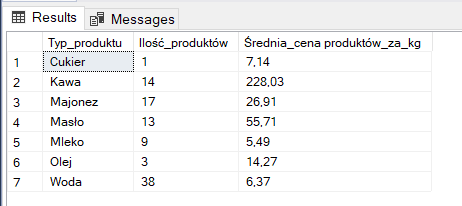
* Zapytanie wyświetla liczbę rodzajów produktów, znajdujących się w sklepie Delio, a także przeciętną cenę za kilogram rodzajów produktów np. w sklepie Delio znajduje się 7 produktów, które są kawą rozpuszczalną. Ich przeciętna cena za kg wynosi 211,35zł.



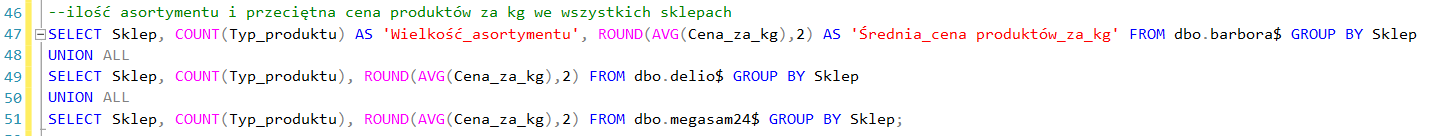


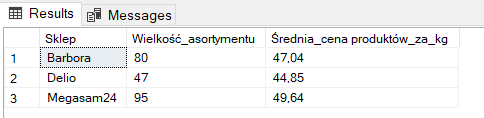
* Zapytanie wyświetla liczbę rodzajów produktów, znajdujących się w sklepie Megasam24, a także przeciętną cenę za kilogram rodzajów produktów np. w sklepie Megasam24 znajduje się 14 produktów, które są kawą rozpuszczalną. Ich przeciętna cena za kg wynosi 228,03zł.



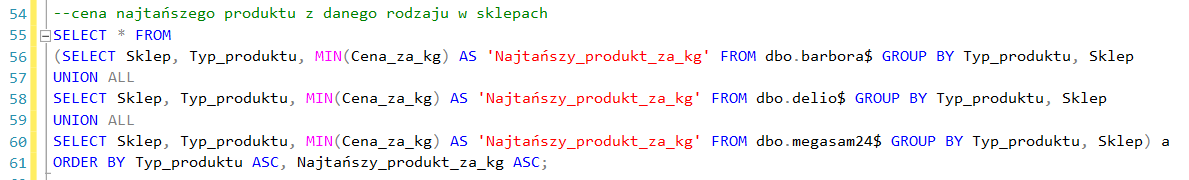


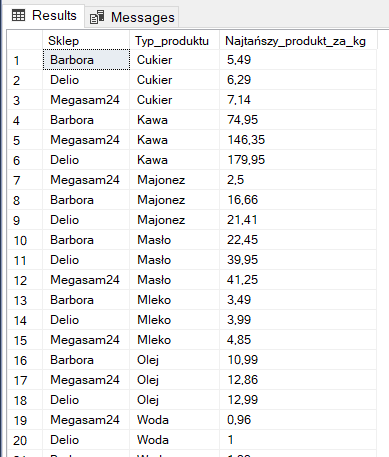
* Zapytanie wyświetla łączną ilość produktów z kategorii: woda, mleko, kawa rozpuszczalna, majonez, masło, cukier, olej, które znajdują się w sklepie Barbora, Delio, Megasam24. Najwięcej produktów znajduje się w ofercie sklepu Megasam24, który ma przeciętnie najwyższą cenę produktów, a najmniej w sklepie Delio, wypadającego najtaniej.



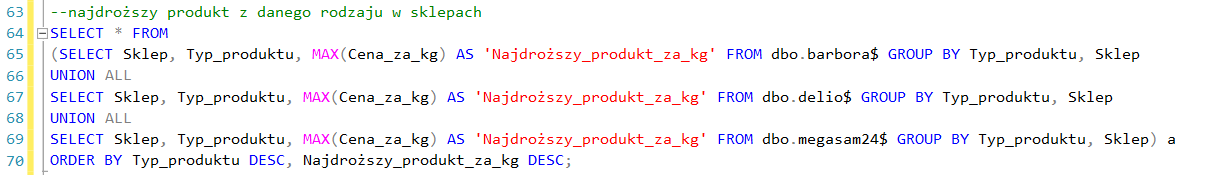


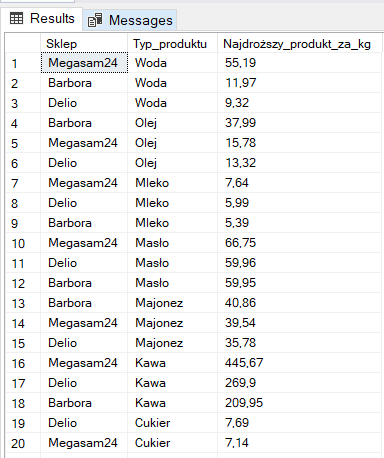
* Zapytanie wyświetla cenę najtańszych produktów w przeliczeniu na kilogram z danego typu produktu we wszystkich sklepach np. najtańszy cukier za kg znajduje się w sklepie Barbora, a najtańsza woda w Megasam24.



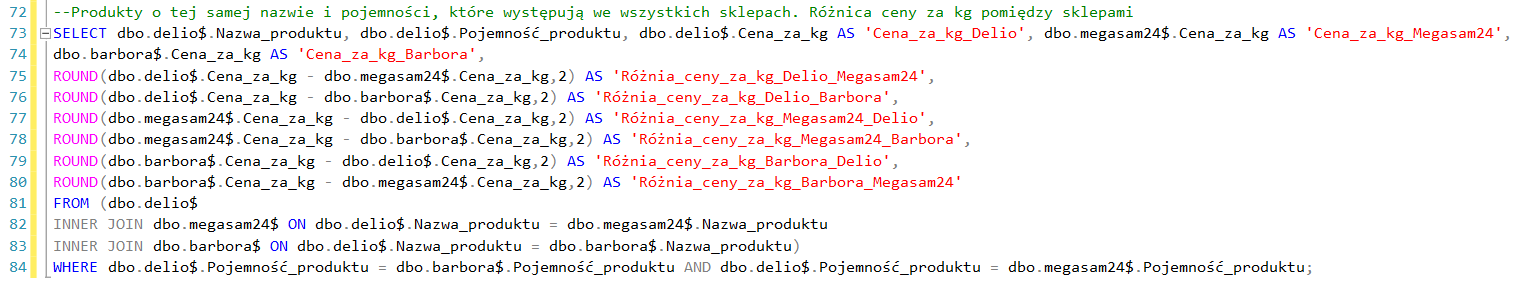


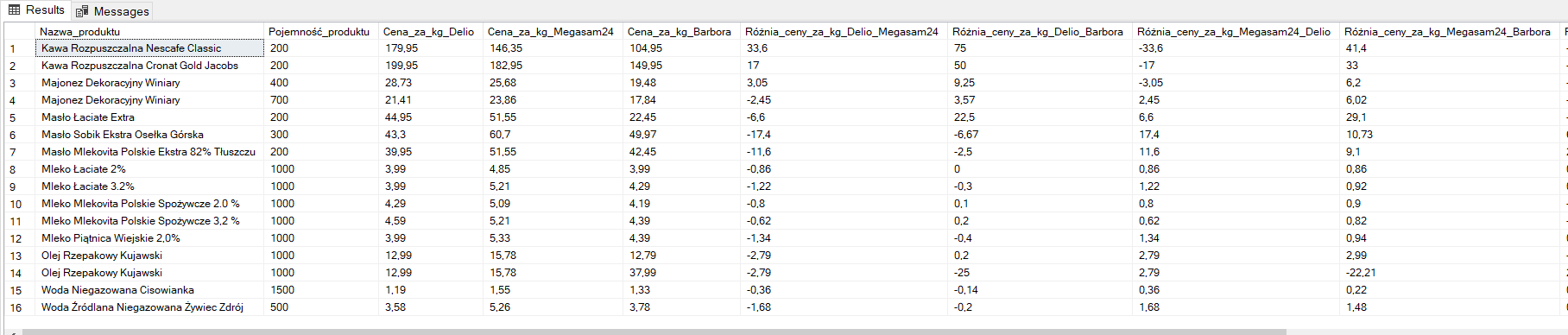
* Zapytanie wyświetla cenę najdroższych produktów w przeliczeniu na kilogram z danego typu produktu we wszystkich sklepach np. najdroższy cukier za kg znajduje się w sklepie Delio, a najdroższa woda w Megasam24.



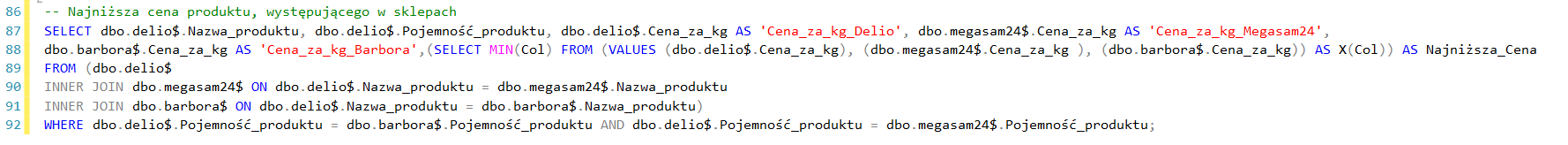


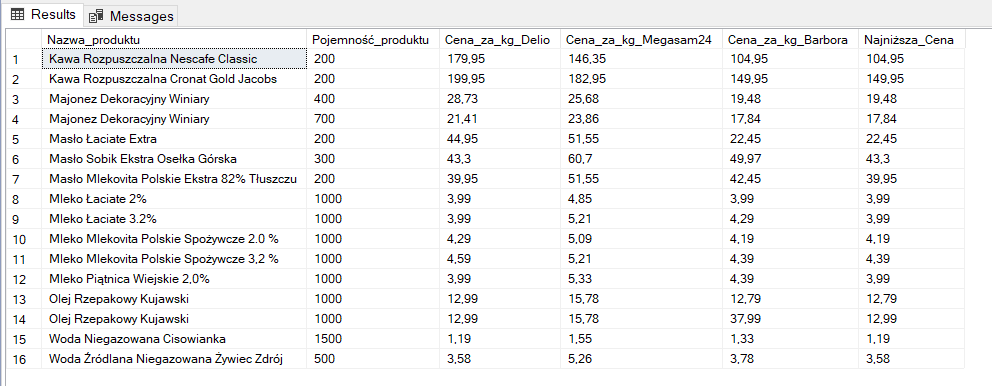
* Zapytanie wyświetla produkty o tej samej nazwie i pojemności, które znajdują się we wszystkich sklepach oraz różnice w cenie produktów występujących pomiędzy produktami np. mleko Łaciate 3,2% jest najtańsze w sklepie Delio, a najdroższe w sklepie Megasam24. Różnica w cenie mleka Łaciatego 3,2 pomiędzy sklepem Delio a Barbora wynosi 30 gr.



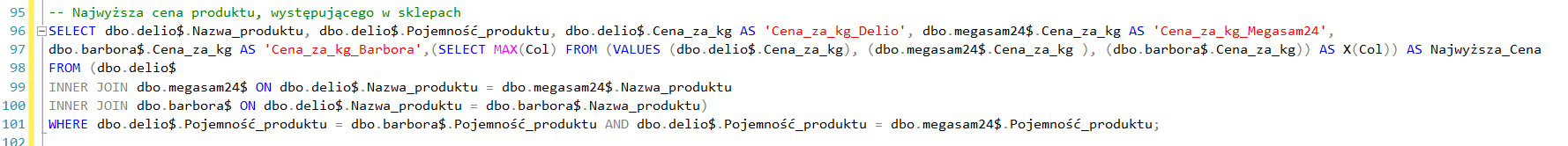


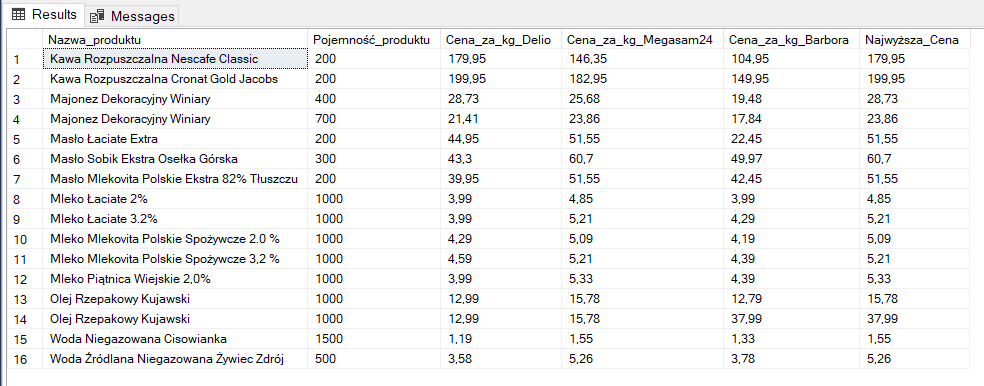
* Zapytanie wyświetla produkty o tej samej nazwie i pojemności, które znajdują się we wszystkich sklepach oraz najniższą cenę za kg danego produktu w nowo utworzonej kolumnie “Najniższa\_Cena” np. spośród wszystkich sklepów najniższa cena za kilogram Kawy Rozpuszczalnej Nescafe Classic wynosi 104.95 zł.



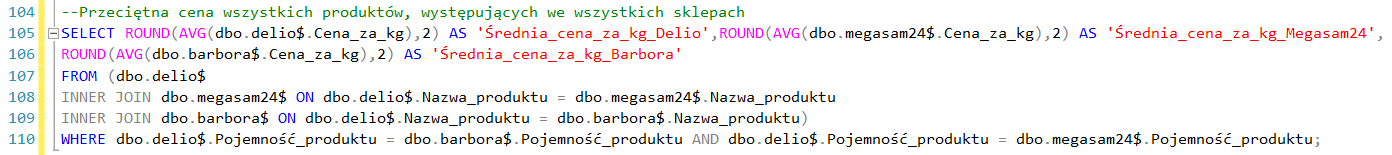


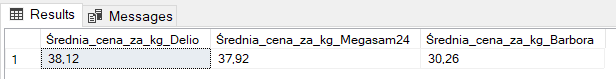
* Zapytanie wyświetla produkty o tej samej nazwie i pojemności, które znajdują się we wszystkich sklepach oraz najniższą cenę za kg danego produktu w nowo utworzonej kolumnie “Najwyższa\_Cena” np. spośród wszystkich sklepów najwyższa cena za kilogram Kawy Rozpuszczalnej Nescafe Classic wynosi 179.95 zł.



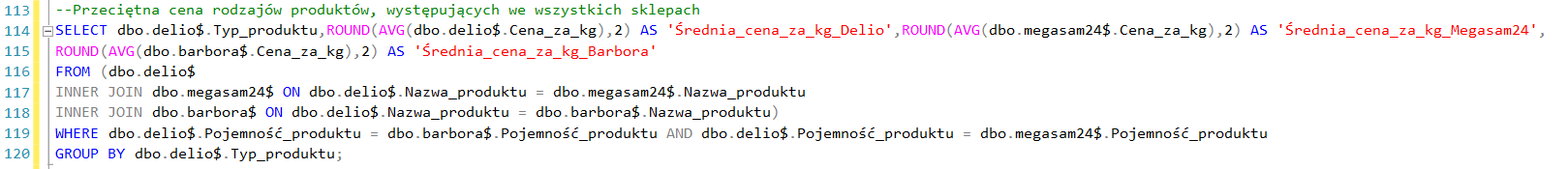


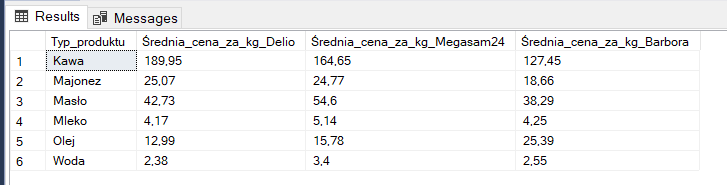
* Zapytanie wyświetla przeciętną, łączną cenę produktów o tej samej nazwie i pojemności, które można odnaleźć w asortymencie każdego z trzech sklepów. Najdroższe są produkty w sklepie Delio, a najtańsze w Barborze.



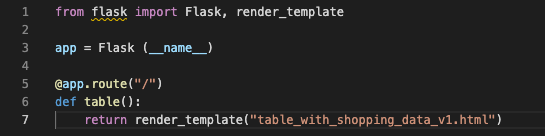


* Zapytanie wyświetla przeciętną cenę produktów z danej kategorii, które mają tą samej nazwę i pojemność oraz znajdują się w asortymencie każdego z trzech sklepów np. najwyższa średnia cena za kg kawy jest w Delio, a najniższa w Barborze.

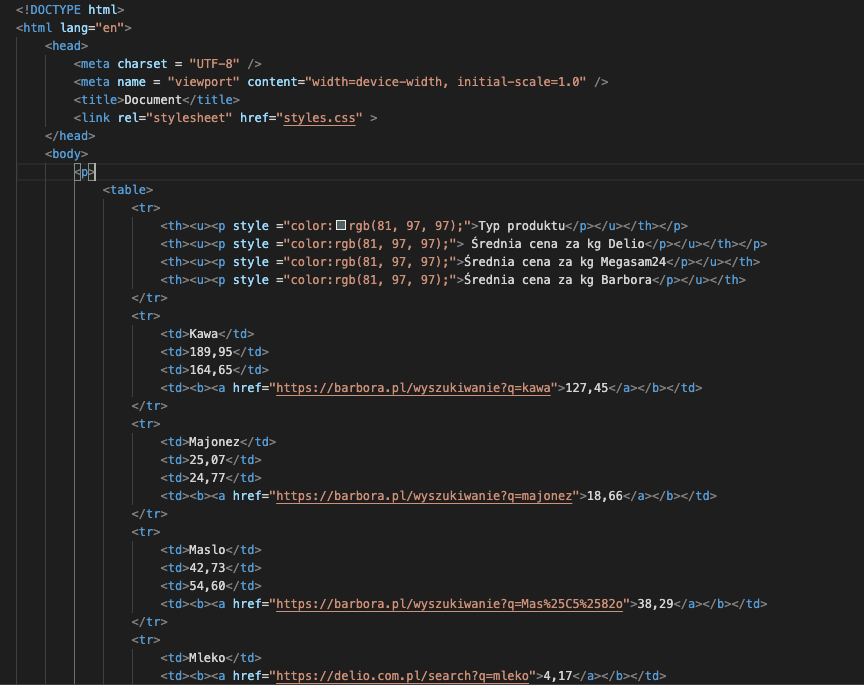




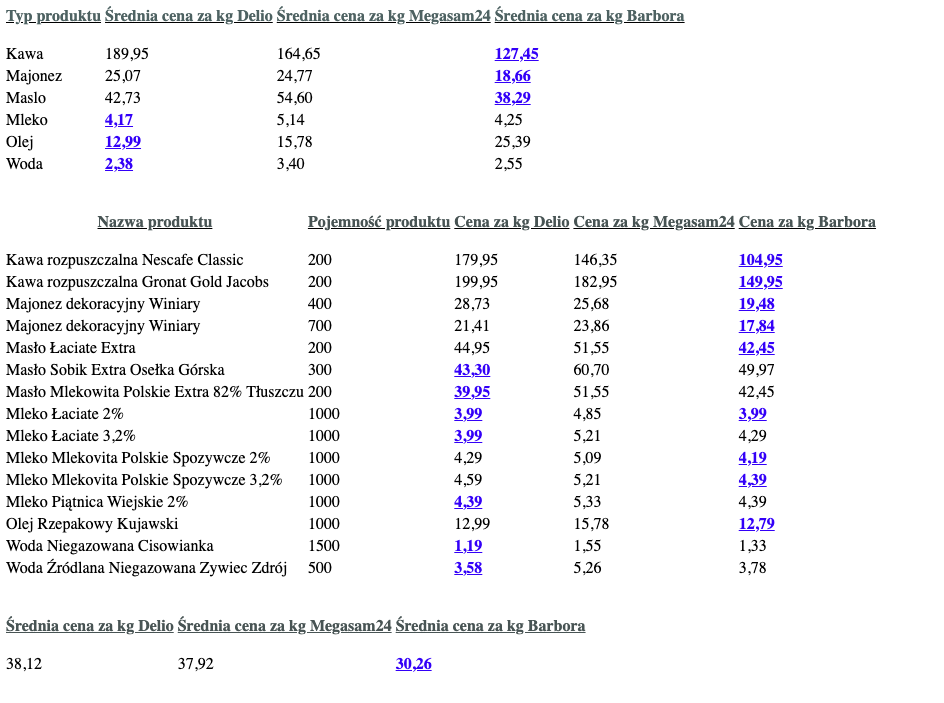
1. Następnie został napisany kod w języku programowania Python ze za pomocą programu Visual Studio Code, który umożliwia przeniesienie pliku na stronę internetową korzystając z Flaska. Poniższy kod umożliwia połączenie się również z plikiem html.



Powyższy kod umożliwia połączenie się z plikiem html, który zawiera najważniejsze informacje o cenach przedstawionych w kilku tabelach.



W kodzie html zostały umieszczone również hyperlinki, pogrubienia i rozmiary wyświetlanych danych. Po wejściu na odpowiednią stronę inernetową korzystając z flaska i kodu html potencjalny użytkownik jest przenoszony na stronę z najważniejszymi wnioskami ze zescrapowanych wówczas danych.



Strona zawiera wykaz cen, produktów i ich pojemności. Użytkownik może przenieść się poprzez kliknięcie w cenę do najtańszej oferty i zakupić produkt.

Z zaprezentowanych danych można łatwo wywnioskować, że sklep Barbora posiada większość najtańszych produktów spożywczych, a Megasam24 jest zdecydowanie najdroższym sklepem.

**Zakończenie:**

Projekt miał na celu porównanie cen produktów spożywczych z wybranych sklepów internetowych jak również przedstawienie najważniejszych wniosków z przekierowaniem do najkorzystniejszych ofert.