Projekt zaliczeniowy

Proces ETL

***Grupa projektowa:...***

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Imię | Nazwisko | Numer albumu | Grupa dziekańska | Wkład w prace nad projektem[[1]](#footnote-0) | Udział procentowy |
| Oleksandra | Hriaznova | 195568 | WZISS2-1111 | Wygląd aplikacji (UI), opracowanie dokumentacji | 30% |
| Karol | Tarłowski | 195029 | WZISS2-1111 | Tworzenie bazy danych, wygląd aplikacji, opracowanie dokumentacji | 35% |
| Grzegorz | Więcek | 195239 | WZISS2-1111 | Opracowanie głównej funkcjonalności aplikacji oraz komunikacji z bazą danych | 35% |

|  |
| --- |
| \_\_/70 pkt |

Spis treści

[**Dokumentacja techniczna**](#_ndo9kno8f933) **3**

[Elementy aplikacji](#_4e48jz4larlu) 3

[Wymagane środowisko](#_whjbsfx18hg1) 3

[Instrukcje instalacji aplikacji](#_lp5xknln4h4w) 4

[Model danych](#_gop4znm96eo) 4

[Opis kodu](#_swtd3znh8bpn) 5

[**Instrukcja obsługi**](#_x7m3ijte2ppg) **8**

[Instrukcja uruchomienia aplikacji](#_1rqdeln4d7gr) 8

[Opis funkcjonalności aplikacji](#_l493o0m7re6j) 9

[Opis scenariuszy użycia aplikacji](#_uk68eq4owegf) 9

[Opis menu i widoków okna aplikacji](#_6zcdhm5n6wl6) 11

# Dokumentacja techniczna

## **Elementy aplikacji**

Aplikacja składa się z trzech elementów:

1. Warstwa front-end.

Warstwa stanowi zasadniczy element aplikacji, udostępnia funkcjonalności przeznaczone dla użytkowników, przedstawia wygląd końcowy aplikacji i umożliwia komunikację z programem.

Użyte technologie:

* HTML5
* AngularJS 1.6.9 (<https://angularjs.org/>)
* CSS3
* Bootstrap 4.1.3 (<https://getbootstrap.com/docs/4.1/getting-started/download/>)

1. Warstwa back-end.

Warstwa ta opisuje całą logikę działania aplikacji, pobierania oraz wysyłania danych, komunikację ze stroną internetową [otodom.pl](http://otodom.pl).

Użyte technologie:

* PHP 7.3 (<http://php.net/downloads.php>)

1. Warstwa dostępu do danych.

Warstwa gromadząca dane pobierane ze strony do bazy danych, umożliwia przechowywanie, przetwarzanie oraz udostępnianie danych.

Użyte technologie:

* MySQL

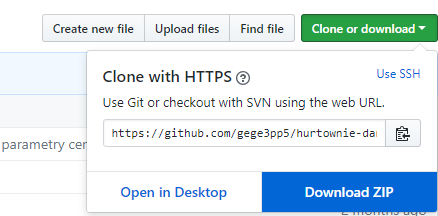
## **Wymagane środowisko**

W celu zapewnienia poprawności działania aplikacji należy spełniać poniższe wymagania:

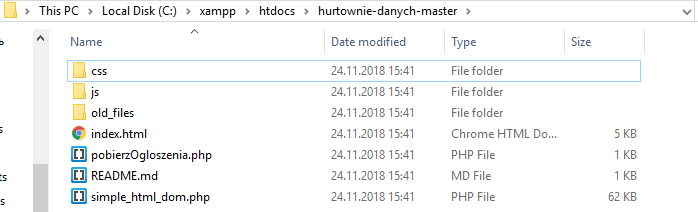
* komputer klasy PC;
* system operacyjny z rodziny Microsoft Windows XP lub wyższe;
* przeglądarka internetowa (np. Google Chrome, Mozilla Firefox lub inna);
* zintegrowany pakiet, składający się z serwera Apache, bazy danych MySQL i interpreterów dla skryptów napisanych w PHP (np. xampp (<https://www.apachefriends.org/index.html>), mamp (<https://www.mamp.info/en/>) )

## **Instrukcje instalacji aplikacji**

Do uruchomienia aplikacji należy pobrać wszystkie pliki w formacie ZIP z repozytorium <https://github.com/gege3pp5/hurtownie-danych>



rozpakować je na uruchomionym serwerze XAMPP w folderze xampp/htdocs/hurtownie-danych (MAMP lub inny).



## **Model danych**

Dane są gromadzone za pomocą relacyjnej bazy danych MySQL. Tworzona jest baza danych **hurtownieDB**, w której tworzy się tabela **ads** z polami:

* Id - numer ID ogłoszenia;
* Name - nazwa ogłoszenia;
* Price - cena mieszkania/domu;
* NrOfRooms - liczba pokoi;
* Size - metraż mieszkania/domu;
* LandSize - metraż działki (w przypadku domów);
* BuildingType - typ zabudowy;
* ContractType - typ umowy.

## **Opis kodu**

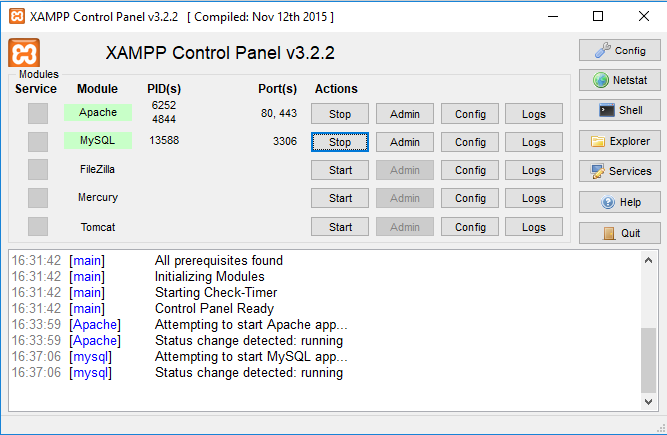
|  |  |
| --- | --- |
| **Nazwa** | **Opis** |
| **hurtownie.js** | |
| poAngielskiemu | funkcja pozwalająca na zmianę wszystkich polskich znaków na angielskie |
| allRegionsName | wybór regionu - Cała Polska |
| regions | nazwy dostępnych województw |
| maxNrOfAdds | maksymalna dostępna liczba pobrań ogłoszeń |
| defaultAddsPerPage | liczba |
| hurtownie |  |
| hurtownie.filter(‘spacedNumber’) | funkcja dodająca spacje do tabeli wynikowej |
| hurtownie.filter('adPrice') | funkcja dodająca cenę do tabeli wynikowej |
| hurtownie.filter('nrOfRooms') | funkcja dodająca liczbę pokoi do tabeli wynikowej |
| hurtownie.filter('size') | funkcja dodająca metraż do tabeli wynikowej |
| hurtownie.filter('landSize') | funkcja dodająca metraż działki do tabeli wynikowej |
| hurtownie.controller("mainCtrl") | funkcja pozwalająca przejść do strony głównej |
| hurtownie.controller("homeCtrl") | funkcja pozwalająca przejść do zakładki menu “Home” |
| hurtownie.controller("etlCtrl") | funkcja pozwalająca przejść do zakładki menu “Proces ETL” |
| .searchParams | metoda opisująca podstawowe parametry |
| c.extract | funkcja, która zbiera dane od użytkownika, wykonuje żądanie php, dzieląc na części składowe - totalNrOfAddsElem; totalNrOfAdds; nrOfAddsToDownload; nrOfPages. |
| extractNextPage() | funkcja, która w przypadku wyszukiwania kilku stron kontynuuje działanie programu przechodząc do kolejnych stron |
| c.transform | funkcja pozwalająca na transformacje danych |
| c.load | funkcja ładująca zebrane dane do bazy danych |
| c.performETL | funkcja wykonująca wszystkie etapy ETL |
| c.buildBaseUrl | funkcja tworząca URL w zależności od parametrów wprowadzonych do formularza |
| c.extractAdsTo | funkcja przygotowująca dane do bazy danych |
| hurtownie.controller("dbCtrl") | funkcja pozwalająca przejść do zakładki menu “Operacje na bazie danych” |
| c.getAds | funkcja wykorzystująca funkcję z pliku getAds.php, pozwalająca przeszukać bazę danych i wybrać rekordy podpadające pod wybrane kryteria |
| c.getCSV | funkcja pozwalająca wyeksportować bazę danych do pliku .csv |
| **Database.php** | |
| Database | klasa służąca zintegrowaniu metod pracy z bazą danych |
| \_\_construct() | konstruktor pozwalający się podłączyć się do bazy danych “hurtownieDB” |
| \_\_destruct() | destruktor pozwalający automatycznie zamknąć połączenie |
| execute\_statement | funkcja wykonująca wyrażenie |
| execute\_prepared\_statement | funkcja wykonująca przygotowane wyrażenie |
| fetch\_prepared\_statement | funkcja ładująca do bazy przygotowane wyrażenie |
| prepare\_statement | funkcja przygotowująca wyrażenie |
| **getAds.php** | |
| getAds() | funkcja wybierająca rekordy z bazy danych wykorzystująca obiekt Database() |
| **load.php** | |
| load() | funkcja zapisująca rekordy do bazy danych, wykorzystująca obiekt Database() |
| **getCSV.php** | |
| getCSV() | funkcja pozwalająca na eksport rekordów z bazy danych do pliku .csv, wykorzystująca obiekt Database() |
| **createDB.php** | |
| createDB() | funkcja nawiązująca połączenie z serwerem MySQL, która następnie tworzy bazę danych hurtownieDB wraz z tabelą o nazwie “ads” |

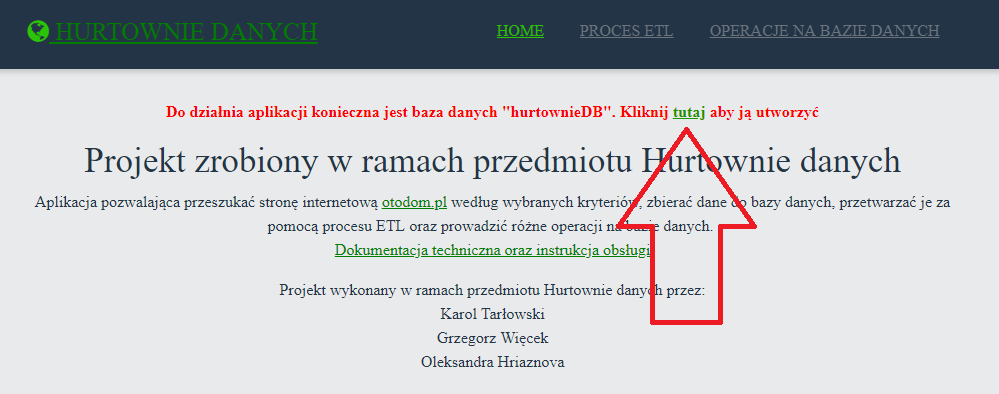
# 

# Instrukcja obsługi

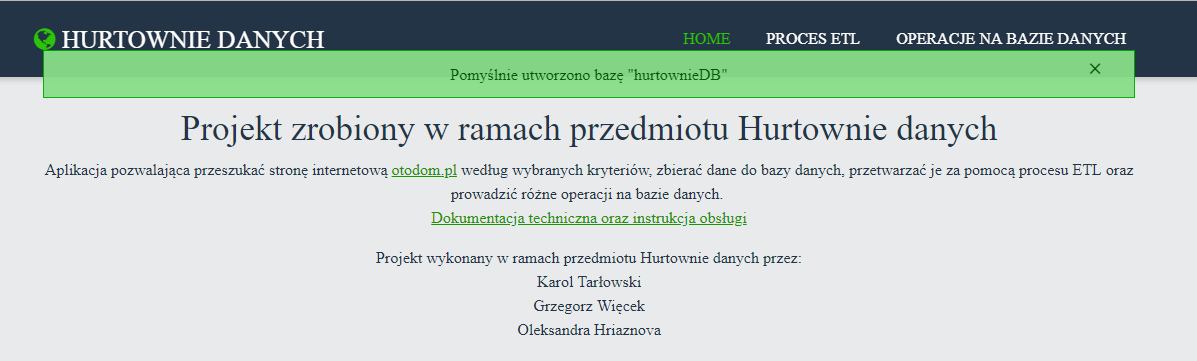
## **Instrukcja uruchomienia aplikacji**

Po rozpakowaniu folderu ZIP na serwerze XAMPP należy przez panel sterowania włączyć Apache oraz MySQL.



Po czym w przeglądarce wprowadzić adres [localhost/hurtownie-danych](http://localhost/hurtownie--danych) i aplikacja zostanie uruchomiona. Przy pierwszym uruchomieniu aplikacji należy utworzyć bazę danych poprzez naciśnięcie linku “tutaj”, który znajduje się w miejscu pokazanym poniżej.  


Po utworzeniu bazy danych zostanie wyświetlony komunikat potwierdzający utworzenie bazę danych oraz zostaną odblokowane wszystkie funkcje aplikacji.





## **Opis funkcjonalności aplikacji**

Aplikacja pozwala na przeprowadzenie całego procesu ETL na raz, jak również przeprowadzenie każdej z jego części osobno: E (extract) - pozyskanie danych ze strony [otodom.pl](http://otodom.pl); T (transform) - przekształcanie danych przed ładowaniem do bazy danych; L (load) - ładowanie danych do bazy danych. Po wykonaniu każdej z części procesu ETL aplikacja wyświetla informacje statystyczne (liczba plików pobranych/przetworzonych). Aplikacja pozwala również na aktualizację danych przechowywanych w bazie danych, ma funkcjonalność pozwalającą wyczyszczać całą bazę danych oraz wyeksportować wszystkie lub wybrane rekordy do pliku .csv.

## **Opis scenariuszy użycia aplikacji**

**Przypadek użycia:** Wyszukiwanie mieszkań

**Użytkownik:** klient.

**Scenariusz główny:**

1. Klient wchodzi na stronę główną.
2. Klika przycisk utworzenia bazy danych.
3. Przechodzi do zakładki “Proces ETL”.
4. Wybiera następujące kryteria w formularzu:
   1. Typ zabudowy: mieszkanie;
   2. Typ umowy: wynajem;
   3. Cena minimalna: null;
   4. Cena maksymalna: 3000 zł;
   5. Powierzchnia minimalna: 30;
   6. Powierzchnia maksymalna: 60;
   7. Województwo: małopolskie.
5. Klient klika przycisk ETL.
6. Aplikacja wyświetla tabelę wyników, zapisuje wszystkie rekordy do bazy danych oraz wyświetla komunikat że rekordy zostały zapisane.

**Scenariusz alternatywny:**

2a. Klient już ma zainstalowaną bazę danych. Pomija krok 2.

5a. Klient klika przycisk E.

6a. Aplikacja wyświetla tabelę wyników. Po czym jest dostępna opcja kliknięcia przycisku “T”.

5b. Klient klika przycisk ETL.

6b. Aplikacja wyświetla tabelę wyników, zapisuje do bazy tylko te rekordy, których jeszcze nie ma w bazie danych oraz wyświetla komunikat że rekordy zostały zapisane.

**Przypadek użycia:** Eksport pliku .csv

**Użytkownik:** klient.

**Scenariusz główny:**

1. Klient wchodzi na stronę główną.
2. Klient już ma zainstalowaną bazę danych z rekordami, które zostały od wcześniejszego przeprowadzenia procesu ETL.
3. Klient przechodzi do zakładki “Operacje na bazie danych”.
4. Klient klika przycisk “Pobierz plik .csv”
5. Klient pobiera plik na swój komputer.

**Scenariusz alternatywny:**

2a. Klient klika przycisk instalujący bazę danych.

3a. Klient przechodzi do zakładki “Proces ETL”

4a. Klient przeprowadza cały proces ETL z wybranymi przez niego parametrami.

5a. Powrót do punktu 3 scenariusza głównego.

4b. Klient wyszukuje rekordy w bazie danych przez formularz dostępny pod zakładką “Operacje na bazie danych”

5b. Klient pobiera plik .csv z wybranymi rekordami bazy danych poprzez kliknięcie przycisku “Pobierz plik .csv”

## **Opis menu i widoków okna aplikacji**

Aplikacja ma górne menu składające się z trzech części: “Home”, “Proces ETL”, “Operacje na bazie danych”.



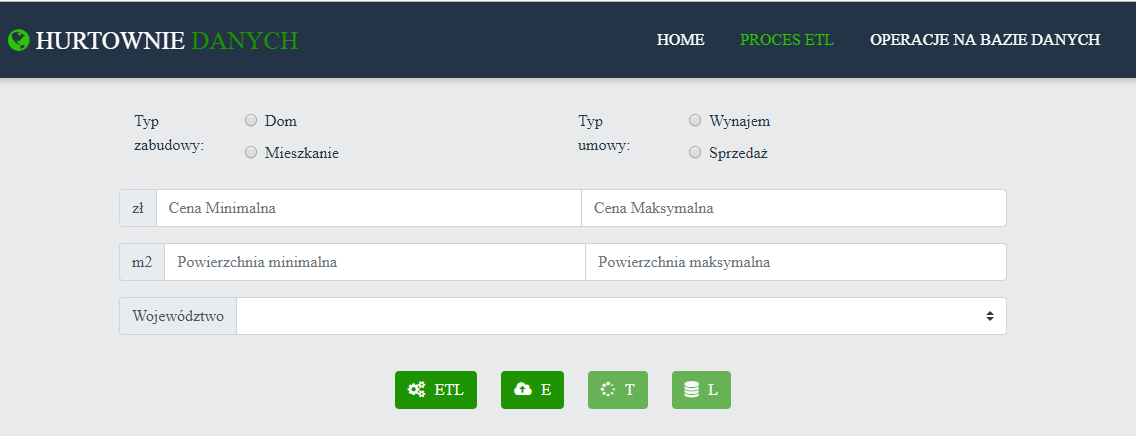
Po uruchomieniu aplikacji zostaną wyświetlone informacje dostępne pod “Home”. Są to informacje ogólne, opisujące czym jest dana aplikacja, kto jest twórcą oraz zawierająca dokumentację techniczną i instrukcje obsługi.



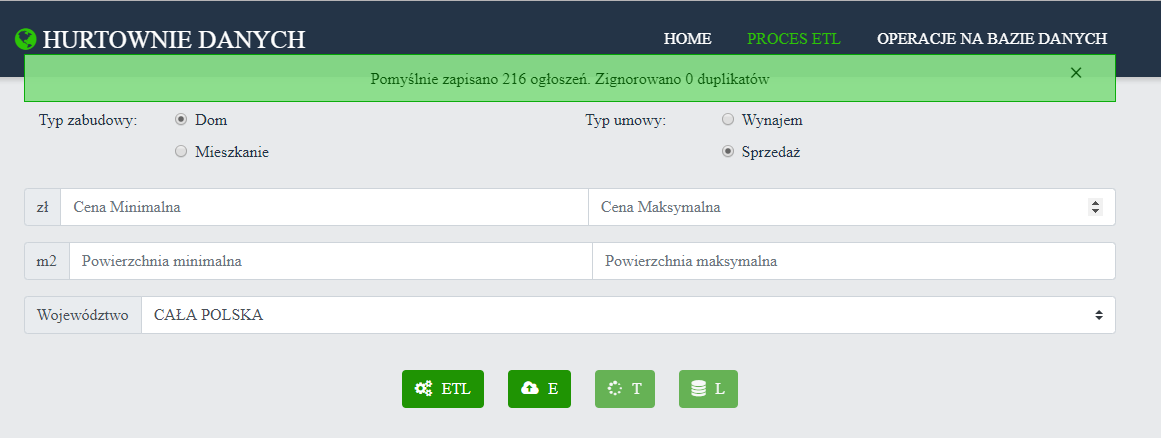
Po przejściu do zakładki “Proces ETL” zostanie wyświetlony formularz umożliwiający wyszukiwanie ogłoszeń na stronie [otodom.pl](http://otodom.pl) oraz przycisk umożliwiający przeprowadzenie całego procesu ETL, jak również jego części osobno.

Dostępne parametry wyszukiwania:

* typ zabudowy (dom/mieszkanie);
* typ umowy (sprzedaż/wynajem);
* cena minimalna oraz maksymalna;
* powierzchnia minimalna oraz maksymalna.



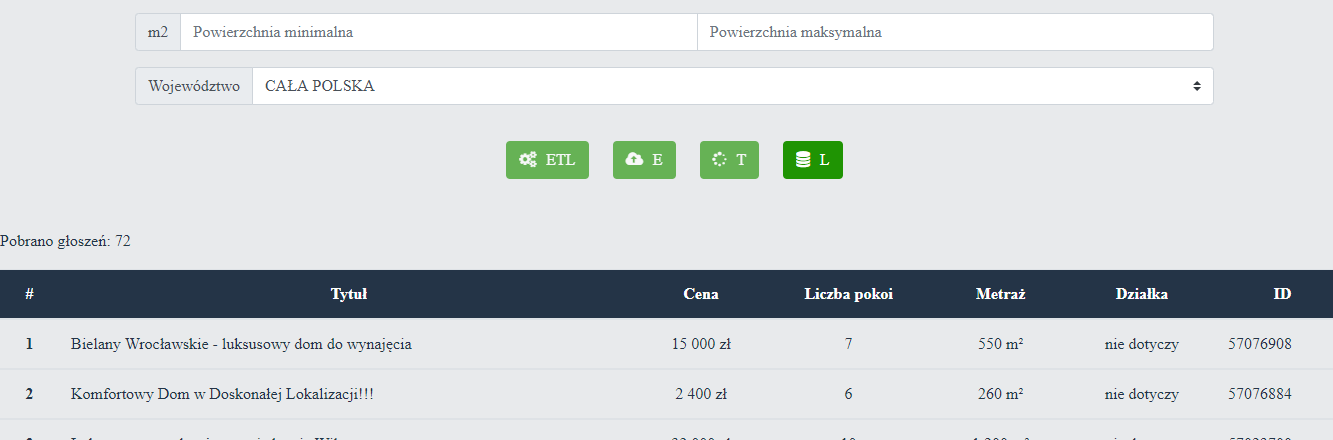
Po wpisaniu parametrów wyszukiwania i kliknięciu na przycisk ETL zostaje wyświetlony komunikat z informacjami statystycznymi.



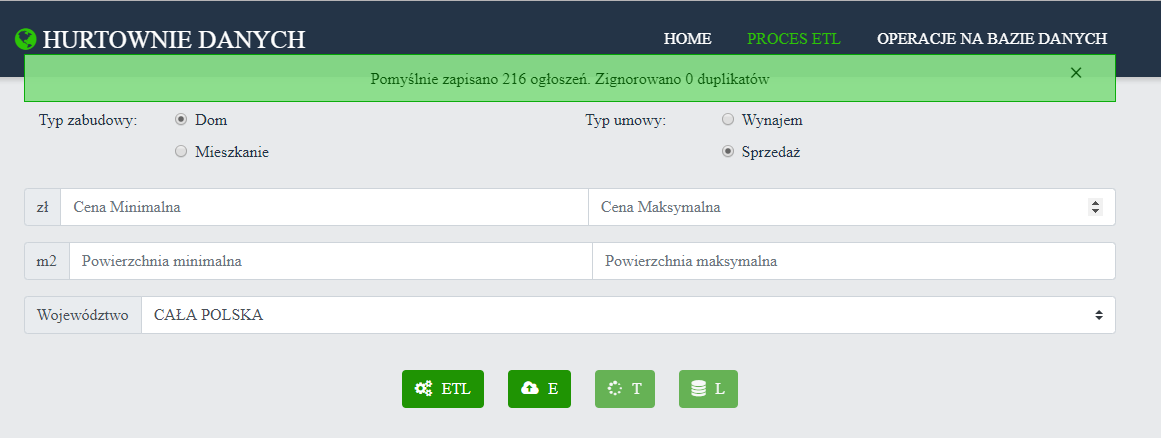
Po kliknięciu na przycisk E zostaje wyświetlona tabela z wynikami oraz zostaje dostępny tylko przycisk T.



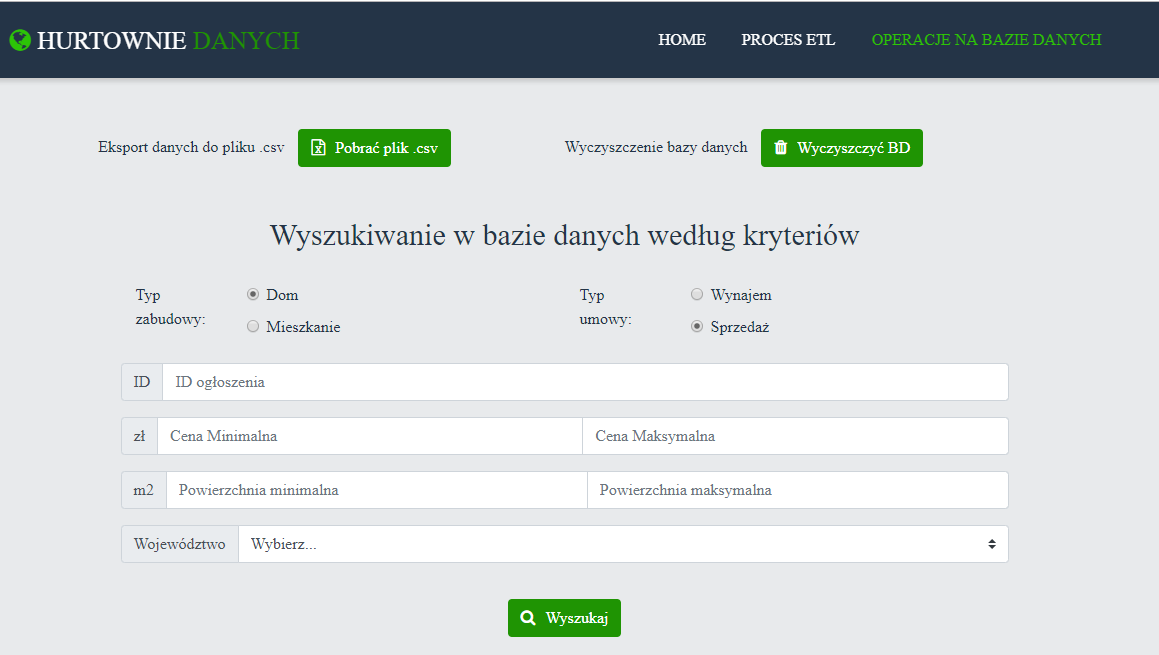
Po kliknięciu na przycisk T zostaje te same wyniki, co w poprzednim kroku oraz zostaje dostępny tylko przycisk L.



Po kliknięciu na przycisk L zostaje wyświetlony komunikat o zapisaniu rekordów do bazy danych oraz znów są dostępny przyciski ETL i E.

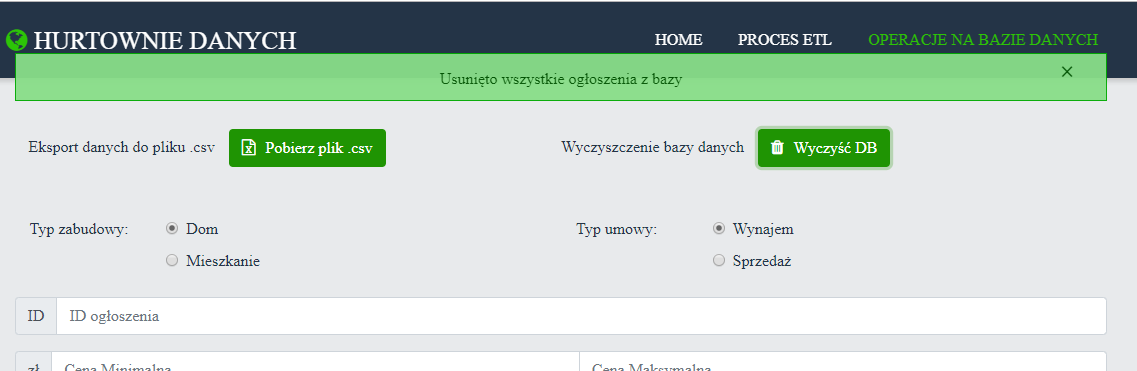


Po przejściu do zakładki “Operacje na bazie danych” wyświetlone zostają funkcjonalności pozwalające operować na bazie danych.



Przycisk “Pobrać plik .csv” pozwala wyeksportować do pliku .csv wszystkie rekordy zawarte w bazie danych.

Przycisk “Wyczyść BD” pozwala usunąć całą zawartość bazy danych.



Formularz zawarty na stronie pozwala przeszukać rekordy zawarte w bazie danych według wymienionych kryteriów (id, cena, metraż, województwo) i wyeksportować je do pliku .csv.

1. proszę wymienić konkretne zadania [↑](#footnote-ref-0)