WOJSKOWA AKADEMIA TECHNICZNA

im. Jarosława Dąbrowskiego

WYDZIAŁ CYBERNETYKI



**Sprawozdanie do projektu z przedmiotu:  
Analiza i wizualizacja danych**

# 

**Prowadzący: Dr Jarosław Olejniczak**

**Student:**

Kita Tomasz Nr grupy: I8J1N1  
Nr indeksu: 64674

## Założenia:

Projekt zakładał budowę aplikacji typu dashboard z użyciem języka R i pakietu Shiny. Do wykonania analiz potrzebny był zbiór danych. **W tym przypadku wzięto do analiz rynek sprzedaży nieruchomości w Warszawie obejmujący stan na 18.06.2021.**

Do pozyskania danych napisano aplikację w języku Python do scrapingu danych z wykorzystaniem biblioteki Beautifufsoup (bs4), Requests. Dane były pobrane z popularnego serwisu z ogłoszeniami nieruchomości **gratka.pl** i dotyczyły rynku warszawskiego, konkretnie sprzedaży nieruchomości.

W trakcie pobierania dane były zapisywane do utworzonej na potrzeby projektu bazy danych SQLite.

Stworzona baza danych jest częścią projektu. Zbiór danych do „obróbki” w aplikacji R, z pakietem Shiny jest więc bezpośrednio pobierany z bazy danych. Do pobierania danych z bazy został użyty pakiet RSQLite. Dodatkowo do prezentacji wykorzystano pakiety DT (do prezentacji tabeli), wspomniany pakiet Shiny, bibliotekę shinydashboard, a do prezentacji także ggplot2.

## Szczegółowy opis prac w projekcie

Dane pobrane z serwisu gratka za pomocą crawlera pobierały informację o każdej nieruchomości wystawionej na sprzedaż a konkretnie:

* dzielnicę,
* metraż,
* cenę nieruchomości.

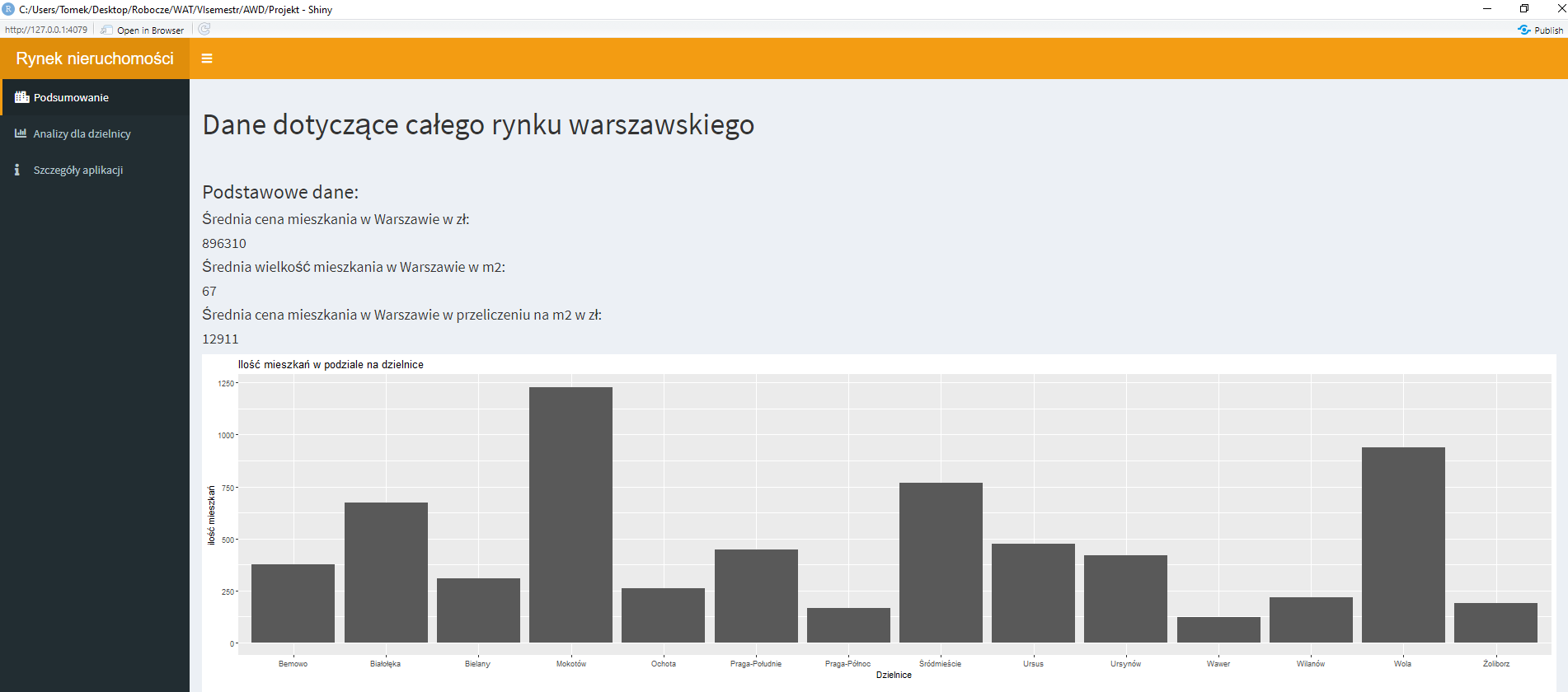
W pakiecie R dokonano przyporządkowania uzyskanych wyników do listy dzielnic. Warto dodać, że pozyskany z internetu zbiór zawierał niepoprawne dane np. użytkownik zamiast dzielnicy podawał osiedle lub ulicę, a zamiast ceny wpisywał cena na telefon. Takie dane zostały pominięte. Warto dodać, że były to bardzo nieliczne przypadki.

## Opis aplikacji

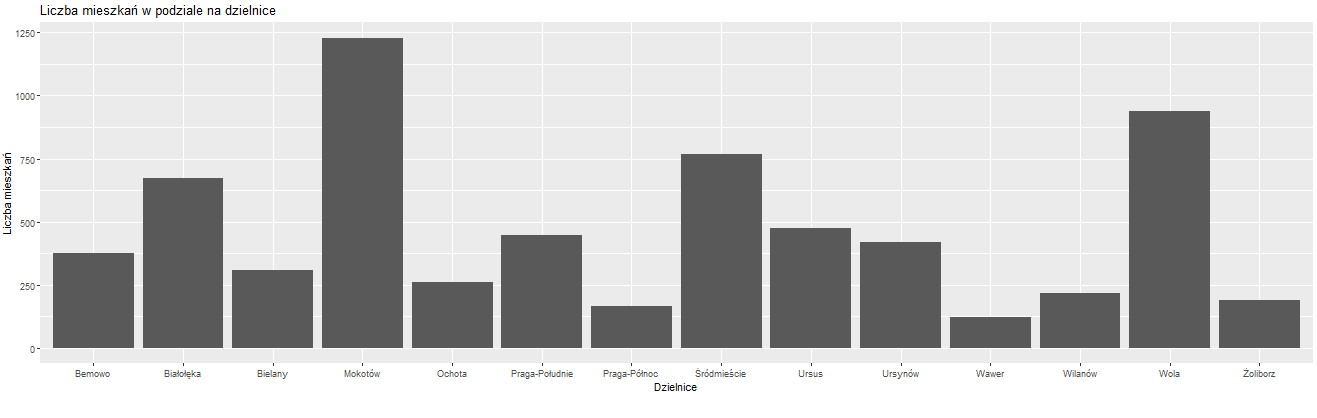
Aplikacja została zbudowana z użyciem biblioteki shinydahsboard. Została pogrupowane na 3 zakładki. 1. Zakładka dotyczy całego warszawskiego rynku nieruchomości. 2. Jest zakładką, gdzie należy dokonać wyboru dzielnicy i na podstawie wyboru pokazuje się wykres oraz szczegółowe informacje dotyczące rynku nieruchomości w danej dzielnicy (tu na podstawie wyboru są na bieżąco obliczane są średnia cena m2, średnia cena mieszkania, średnia . 3. To zakładka informacyjna ze statyczną treścią .

## Elementy w zakładce nr 1:

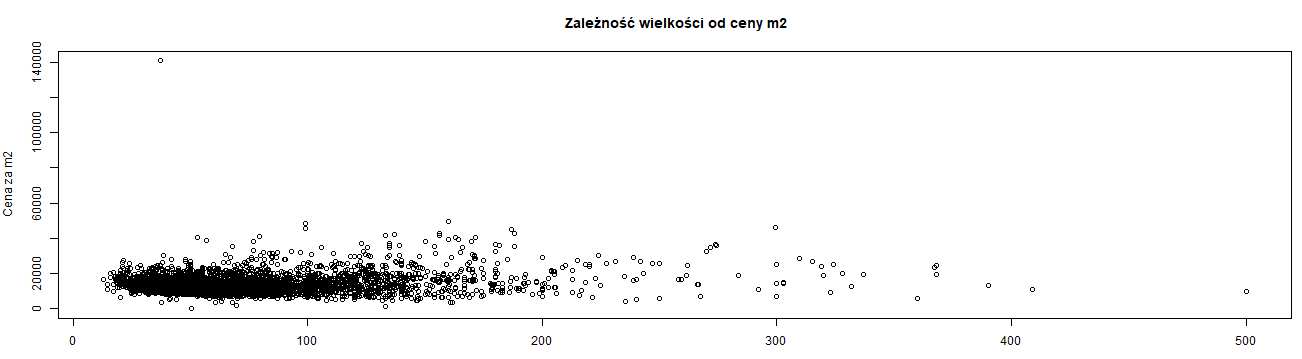
Podstawowe informacje o rynku warszawskim wyliczane są z użyciem funkcji do liczenia średniej. Dowiadujemy się tu, że średnia cena mieszkania w Warszawie to 896310 zł, średnia wielkość mieszkania w Warszawie to 67 m2 a średnia cena m2 to 12911zł.



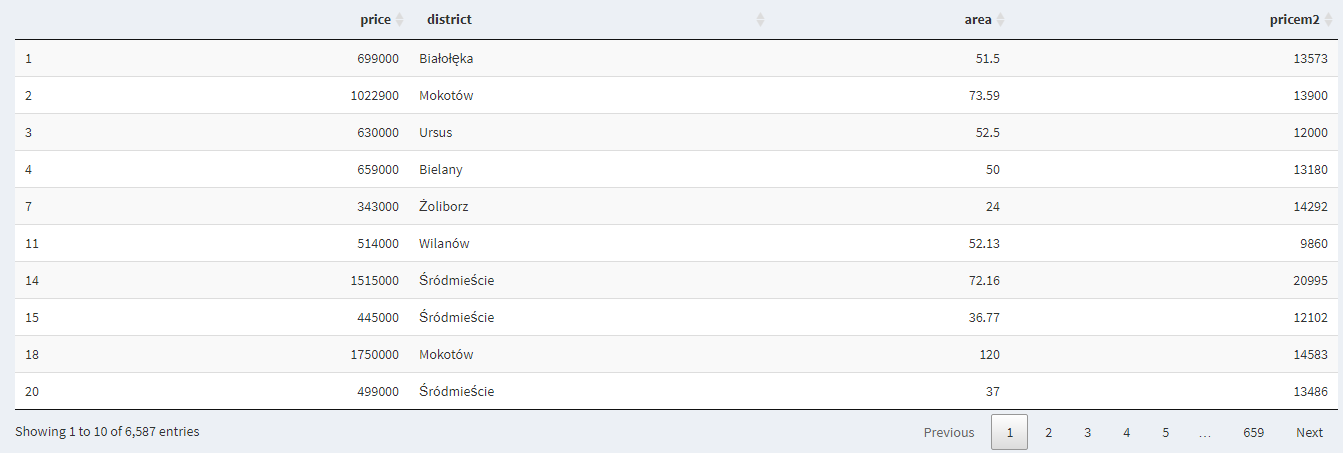
Wykres liczby mieszkań w podziale na dzielnice – wynika z niego, że najwięcej nieruchomości jest do sprzedaży na Mokotowie, Woli i Śródmieściu.



Wykres obrazujący zależność ceny za m2 do wielkości nieruchomości. Z tego wykresu można na pewno wyczytać, że najwięcej nieruchomości ma wielkość do około 150m2 i kosztuje poniżej 20tys zł za m2. Względnie dużo nieruchomości powyżej 150 m2 ma ceny powyżej 20 tys zł za m2.

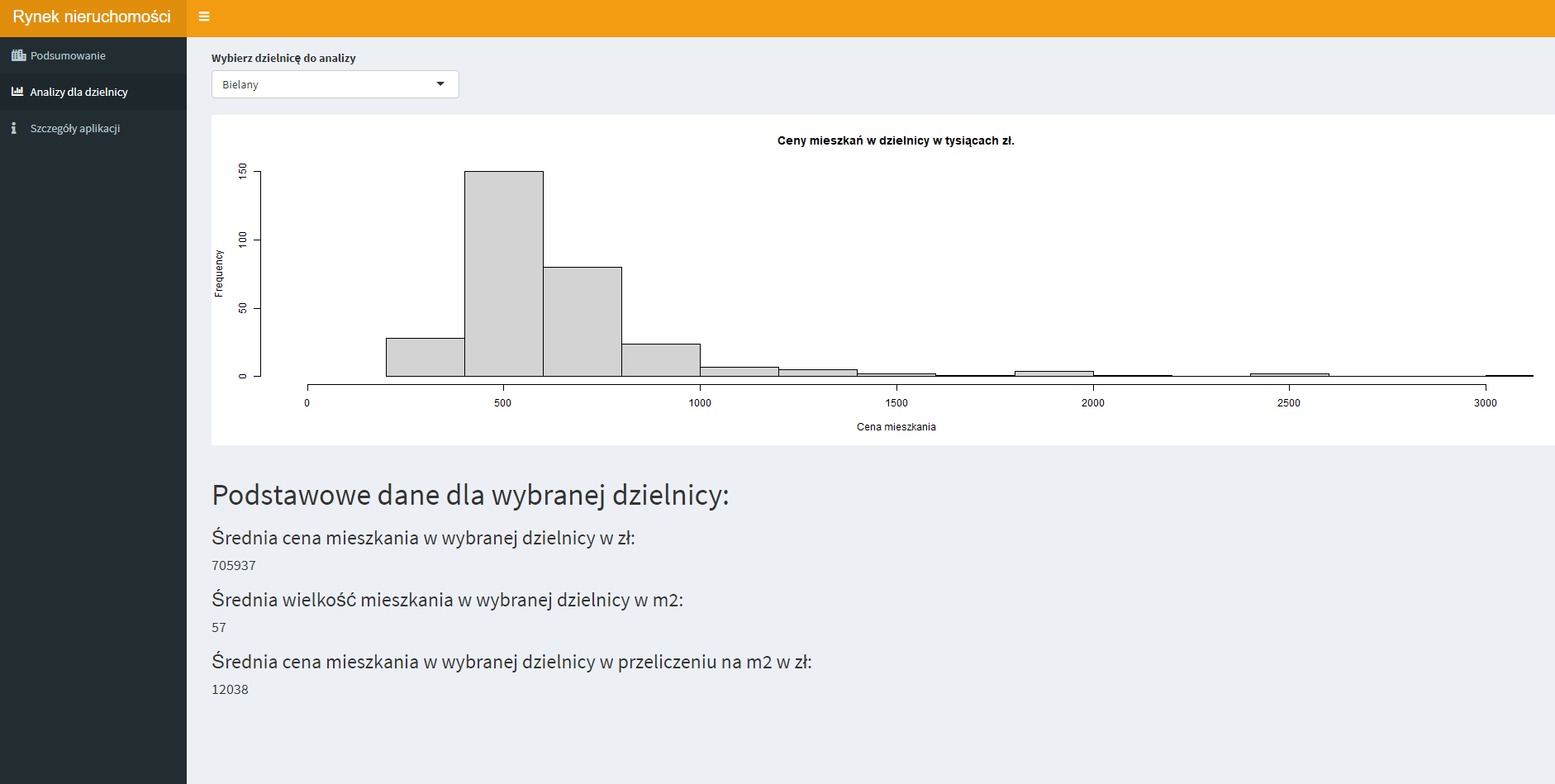


W aplikacji znajduje się też interaktywna tabela, gdzie można obejrzeć wszystkie dane, dokonać sortowania po wybranej kolumnie.



## Elementy w zakładce nr 2

Kolejna zakładka jest interaktywna. Użytkownik musi dokonać wyboru dzielnicy z rozwijanej listy. Na podstawie wyboru generowany jest wykres ze Średnią ceną mieszkania oraz poniżej podstawowe informacje o dzielnicy.



## Elementy w zakładce nr 3

Ostatnia zakładka to zakładka ze szczegółami dotyczącymi aplikacji. W niej znajdziemy statyczne informacje o użytych pakietach.

Obraz zawierający tekst

Opis wygenerowany automatycznie