

WEB INTELLIGENCE NETWORK
CONFERENCE 2025
GDANSK, POLAND



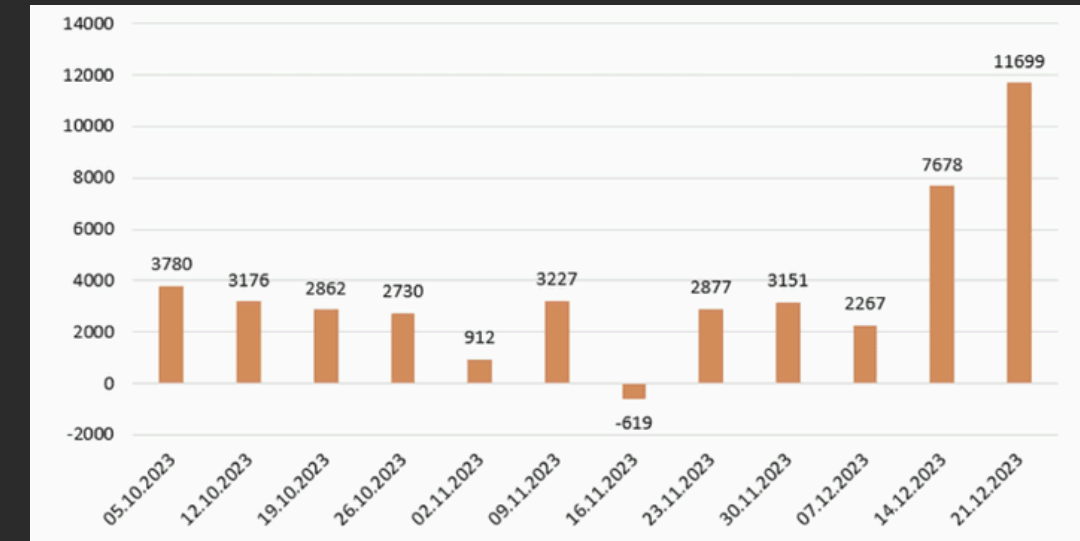
ANALYSING HOUSING IN TRICITY METROPOLITAN AREA IN POLAND

*Piotr Klopotowski, Eng, Master student
*Olgun Aydin, PhD, Assistant Professor

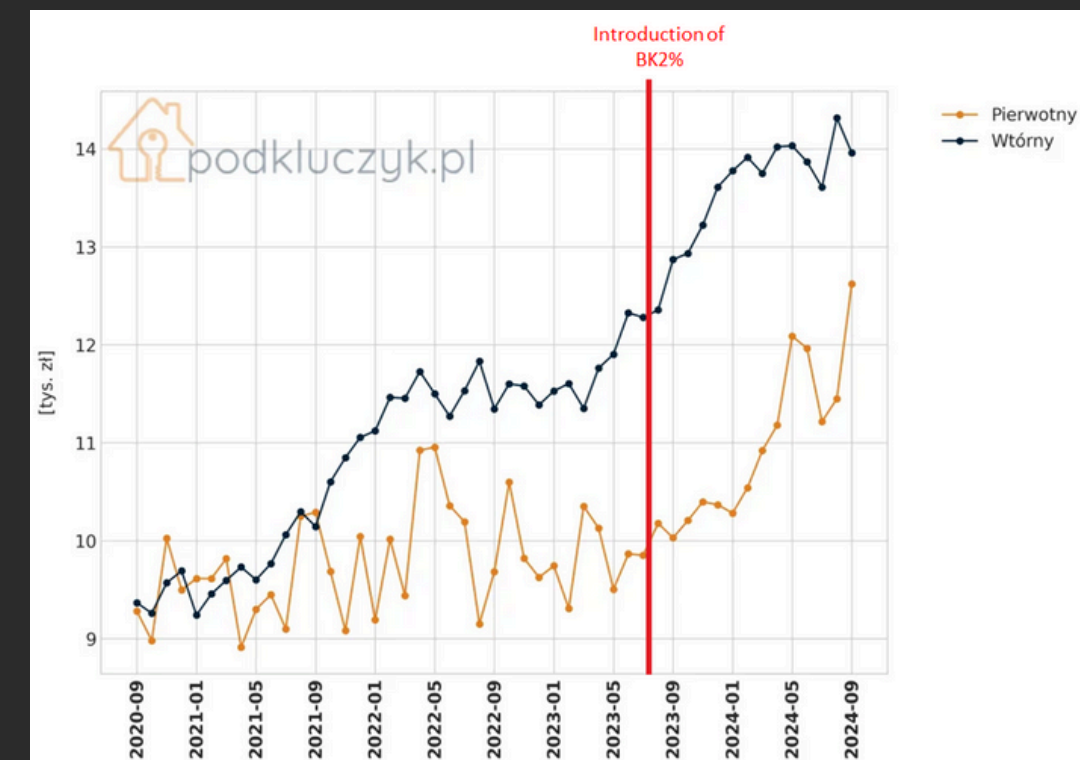
*Gdansk University of Technology

Motivation

- Real Estate price growth in 2023 due to several factors
 - War in Ukraine
 - Bk 2% – a government funded consumer loan program to help people under 45 buy their first flat
- Prices in some districts of the Tricity have risen uncontrollably – even more than the rest of the market



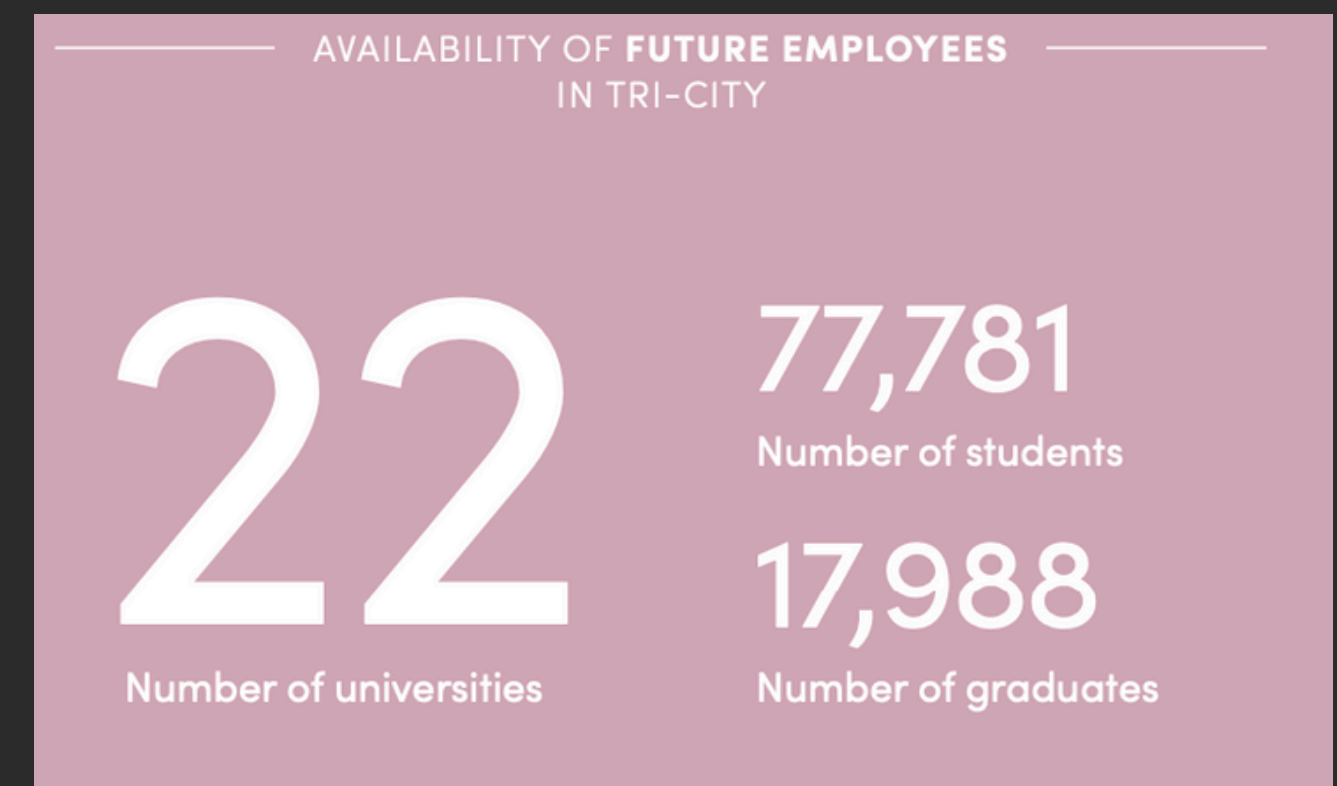
Number of accepted Bk2% requests in Q4 2023 [1]



Median price of a m² in Gdańsk over time
The red line indicates start of Bk2% program [2]

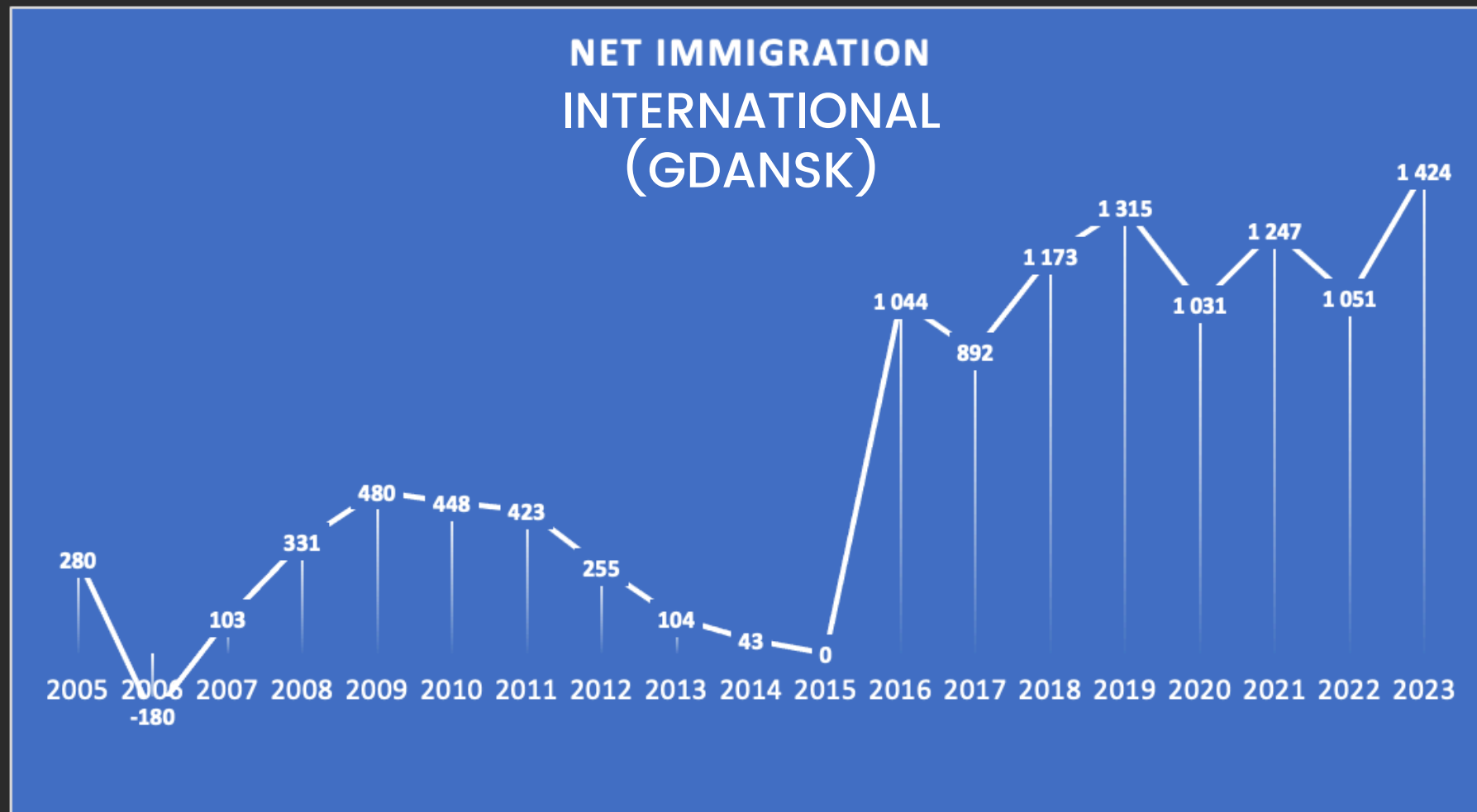
Motivation

- According to Statistics Poland's data, average gross salaries in Tricity is 11% higher than Poland' average.
- Tricity is attractive for investors
- The Tricity is among the four cities in Poland, alongside Warsaw, Kraków, and Wrocław, with a total office stock exceeding one million square meter.
- High quality of life in Tricity



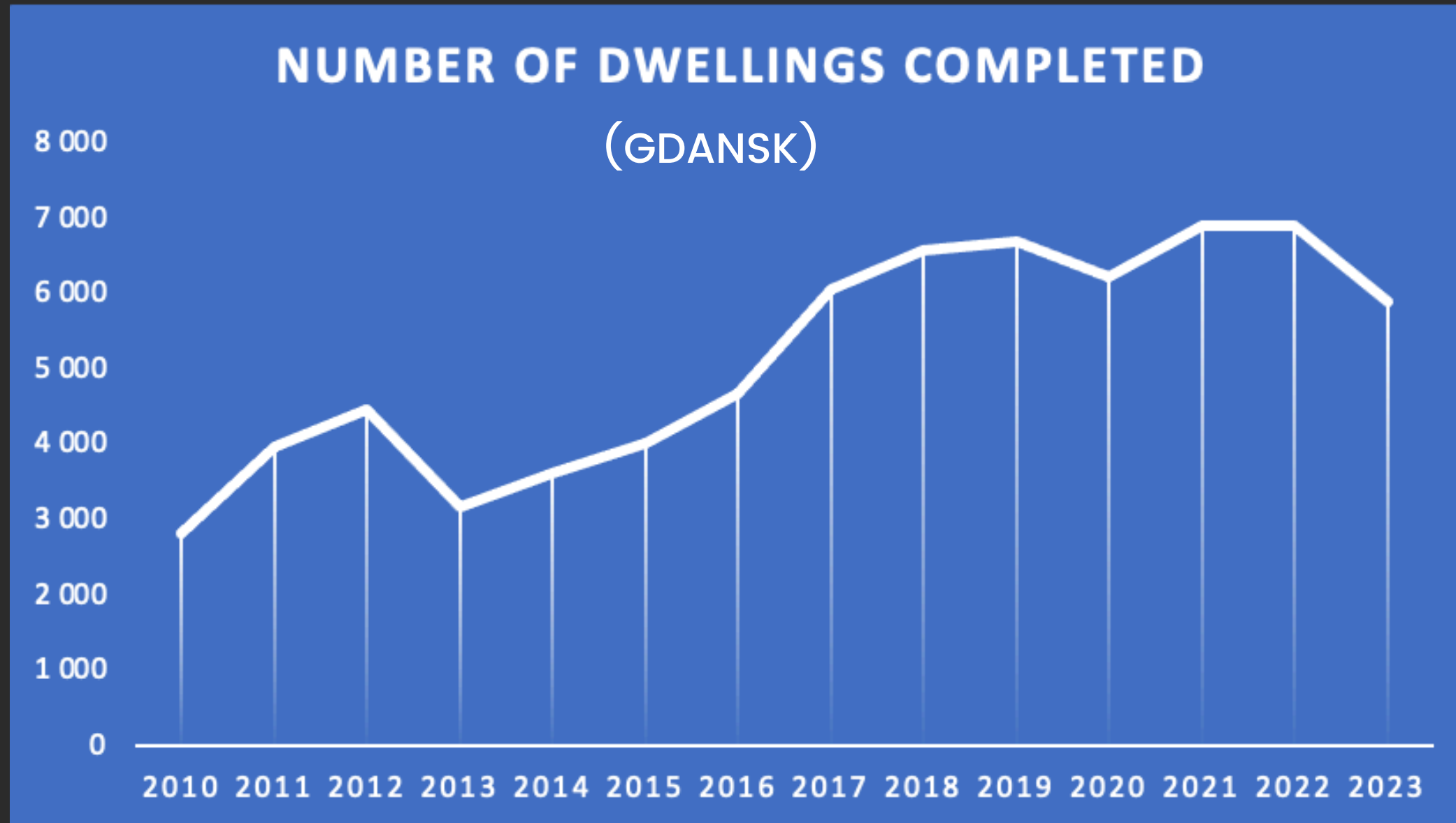
Motivation

2023



Internal Immigration	1 683
International Immigration	1 424
Total	3 107

Motivation

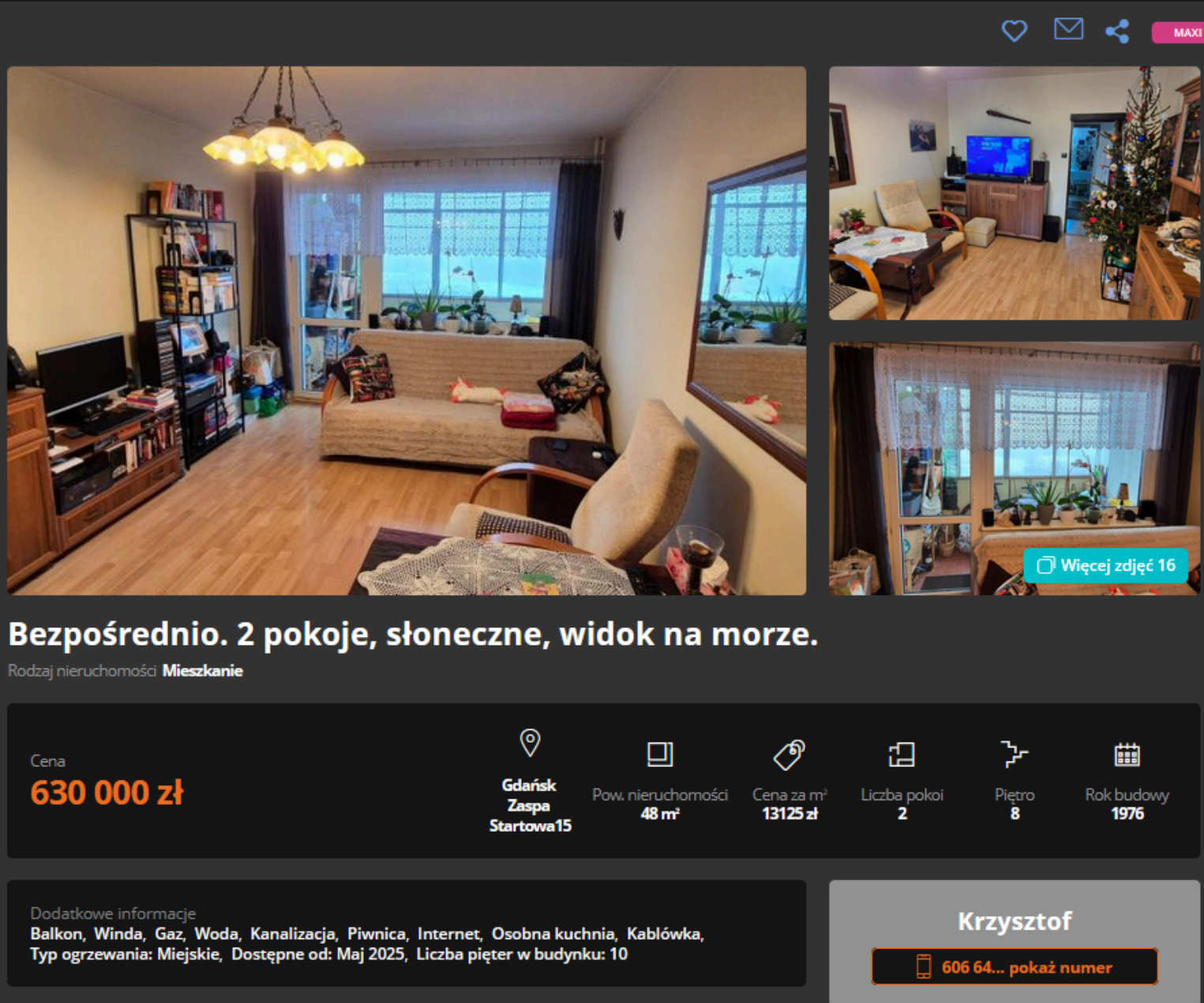


2023
Net Immigration/New Supply

52,89%

Web scraping

- Retrieving data from websites in automatic way
- Key tools: BeautifulSoup, JSoup, Selenium
- Useful, when websites does not provide APIs
- Retrieving data from websites at relatively low cost
- Websites differ in layout and data presentation model, which requires to create distinct scrapers for various sites
- Often they have to be adjusted when website layout changes



Bezpośrednio. 2 pokoje, słoneczne, widok na morze.

Rodzaj nieruchomości: Mieszkanie

Cena: **630 000 zł**

Gdańsk Zaspa Startowa15

Pow. nieruchomości: 48 m²

Cena za m²: 13125 zł

Liczba pokoi: 2

Piętro: 8

Rok budowy: 1976

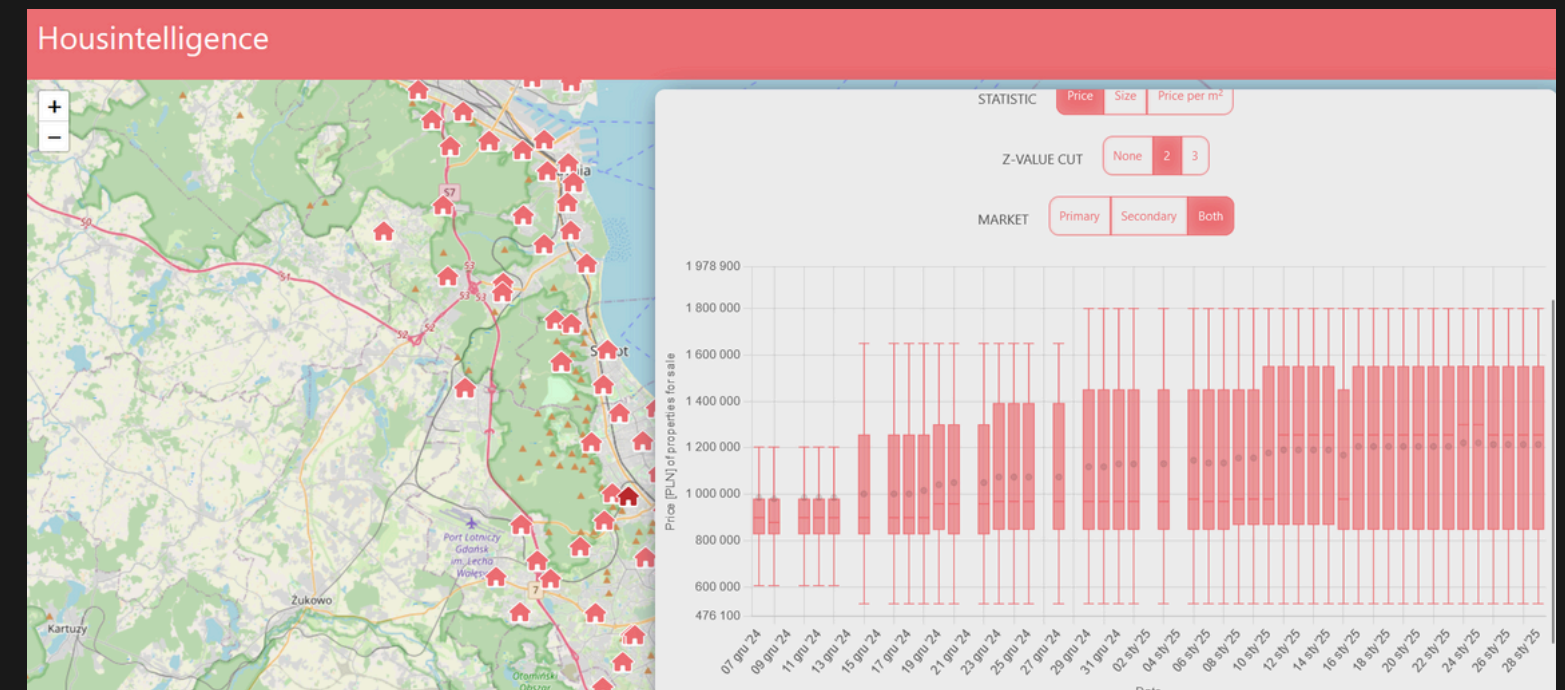
Dodatkowe informacje: Balkon, Winda, Gaz, Woda, Kanalizacja, Piwnica, Internet, Osobna kuchnia, Kabłówka, Typ ogrzewania: Miejskie, Dostępne od: Maj 2025, Liczba pięter w budynku: 10

Krzysztof

606 64... pokaż numer

Application

- Developed as a part of Engineering Thesis on Data Engineering degree and continued during Master Thesis
- Data retrieved from local and national real estate listing websites being scraped every day
- Statistics presented as a boxplot in time series
- Outliers elimination
- Showing graphs for each district of Tricity using various parameters (market, property type, etc.)
- Further development plans: price prediction model, chatbot for searching properties



Key technologies used



Chart.js

An open-source visualization tool allowing to render interactive and lightweight charts inside a browser window.



Selenium

Despite being library primary used for automated testing, this open source tool may be used for web scraping. It may mimic most of the human-user behaviors.

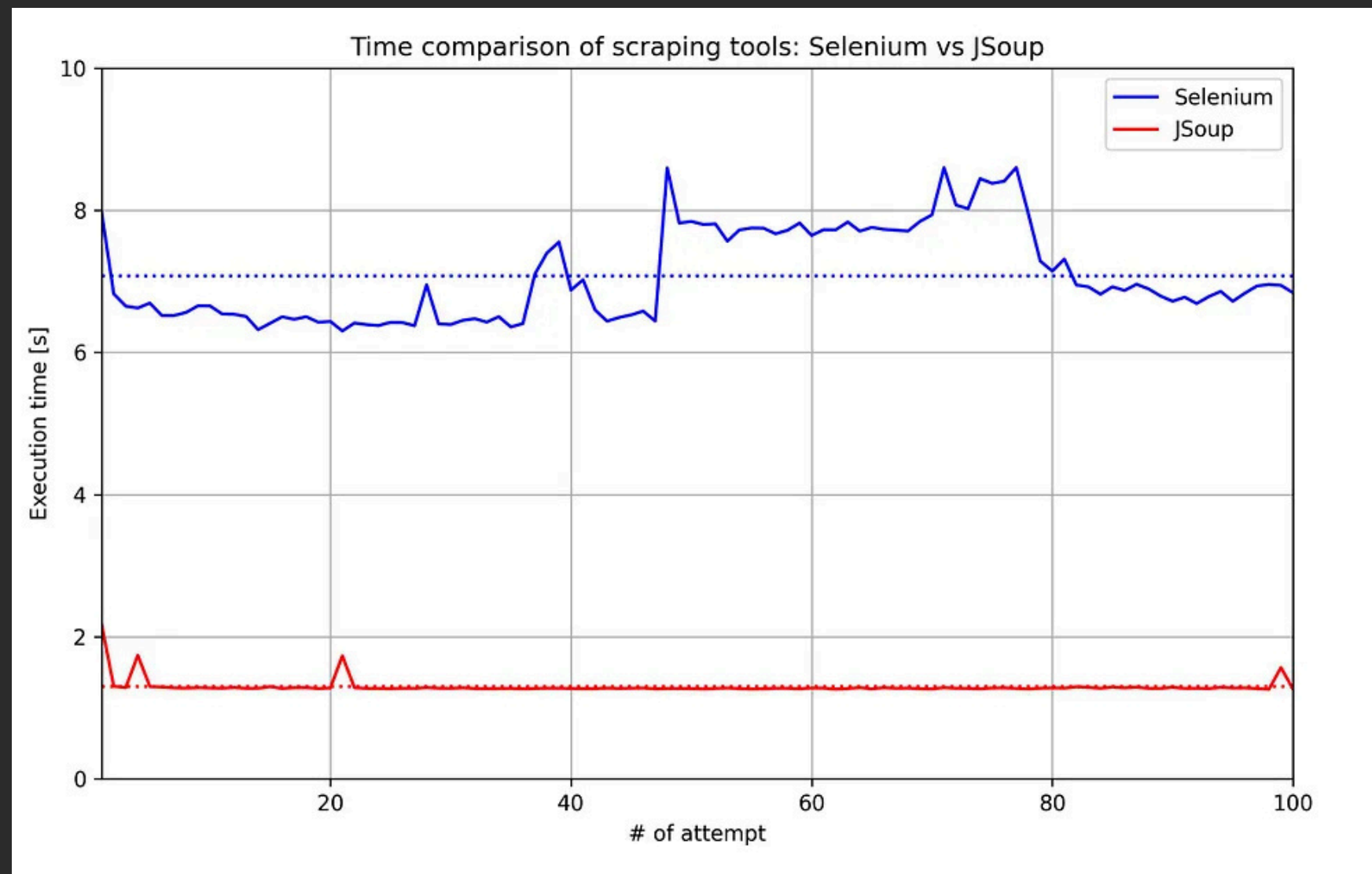


JSoup

Library created solely for the sake of web scraping. Can accomplish scraping tasks much faster than Selenium [3], yet it cannot be used in some specific tasks.

Scraping performance comparison

JSoup vs Selenium

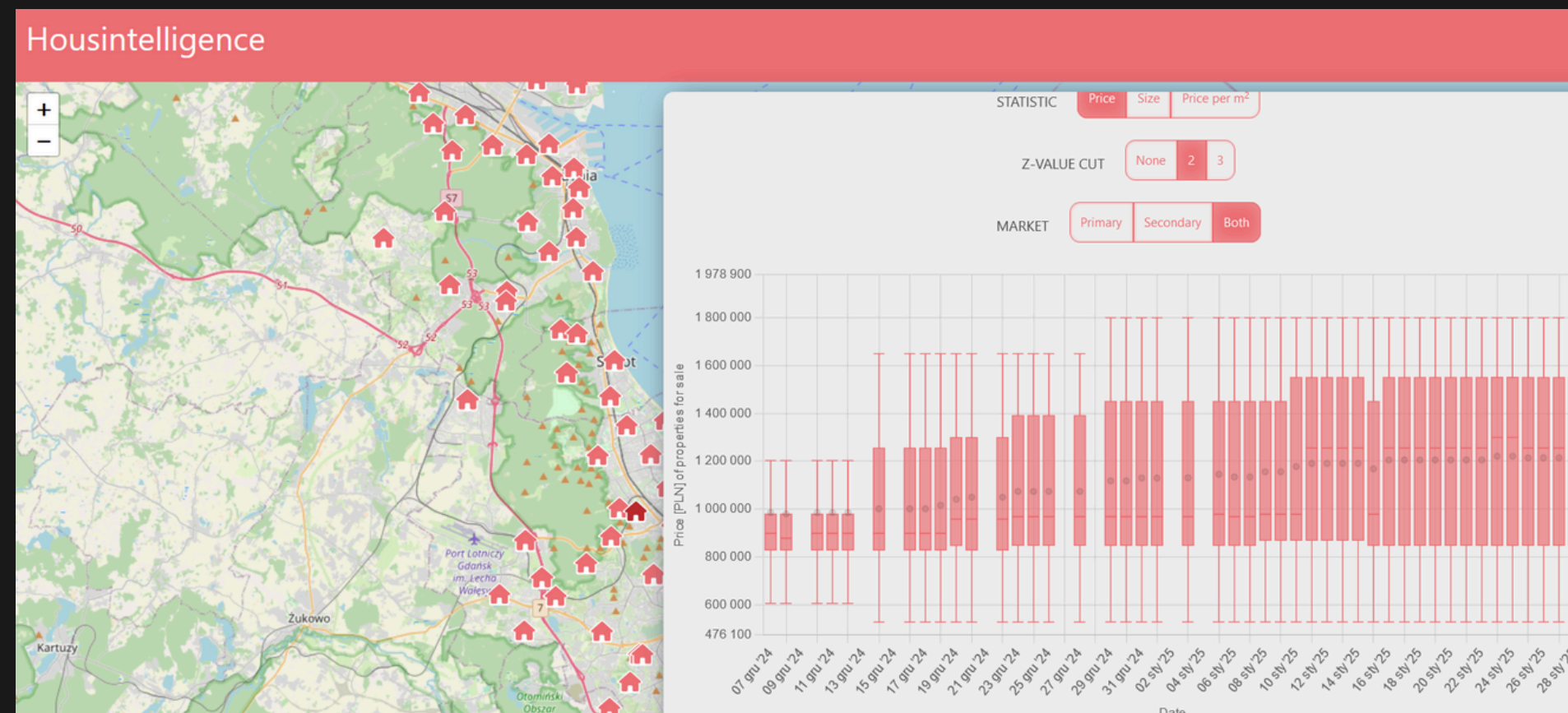


- Selenium needs to render the whole site before scraping
- JSoup is approx. 4 times faster than Selenium

Read the full article (with code examples) on Medium:



Application showcase





Conclusion

- Web scraping is a very efficient and low-cost tool that allows retrieval of huge amounts of data for further processing.
- During 2 months of scraper operation, more than 30 thousand unique advertisements were scraped along with more than 652 thousand daily price observations connected to it.
- The chart changes in the application reflect the changes of the market

Thank you for your attention!

Piotr Kłopotowski, Eng



<https://linkedin.com/in/piotr-kopotowski>



<https://medium.com/@kopotowski>



<https://github.com/kopotowski>

Olgun Aydin, PhD



<https://linkedin.com/in/olgun-aydin>



<https://github.com/olgnaydn>

Bibliography

- 1 <https://bank.pl/ceny-mieszkan-2023-za-duzo-o-trzy-wstrzasy-na-ryнку-nieruchomosci/>
- 2 <https://podkluczyk.pl/raport-ceny-mieszkan-za-metr-w-gdansku-w-poszczegolnych-dzielnicach-we-wrzesniu-2024/>
- 3 <https://medium.com/@klopotowski/scraping-using-jsoup-and-medium-performance-comparison-3a7544091893>
- 4 <https://content.knightfrank.com/research/2300/documents/en/tricity-city-attractiveness-and-office-market-q1-2024-11273.pdf>
- 5 <https://www.paih.gov.pl/en/wp-content/uploads/sites/2/2024/08/Investment-potential-of-Tri-City-2023.pdf>