

Analýz dat z webu www.sreality.cz

V tomto projektu budeme analyzovat data z webu www.sreality.cz. Obdržíme aktuální nabídku bytů ve Zlíně na 26.01.2025.

Projekt se dělí na 3 části:

1. WebScraping

- Pomocí knihovny `BeautifulSoup` získáme všechny aktuální inzerce bytů v daném městě.
- Získáme informace: `URL odkaz bytu, rozměry bytu, kompozice bytu, cena bytu, lokace bytu (Ulice+město)`

2. Datová analýza - ETL proces

- V první fázi provedeme processing dat tak, aby obsahovala příslušné datové typy, nenulové hodnoty atp. Ulice i město bude v samostatných sloupcích.
- Zprocesovaná a vyčištěná data vyexportujeme do `.csv` souboru.

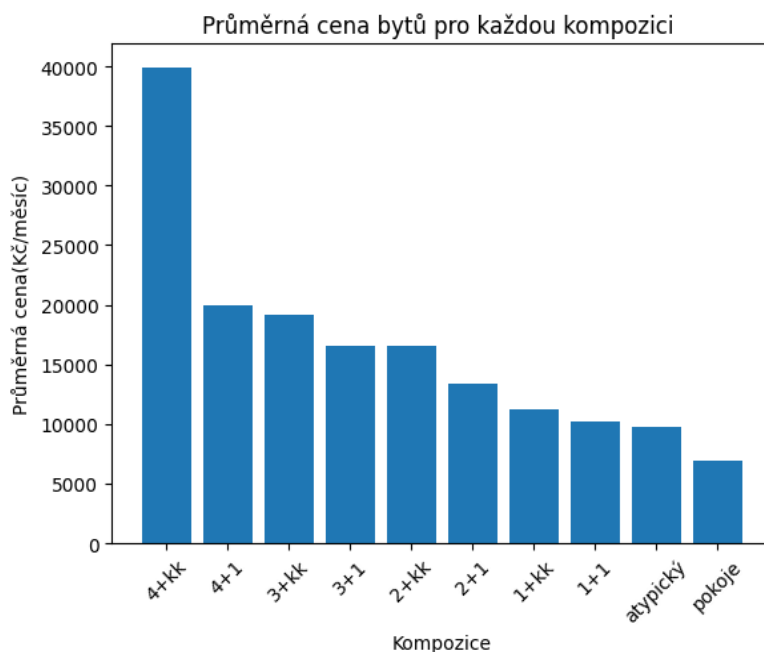
3. Datová analýza - analýza dat

Provedeme analýzu dat a odpovíme na tyto dotazy:

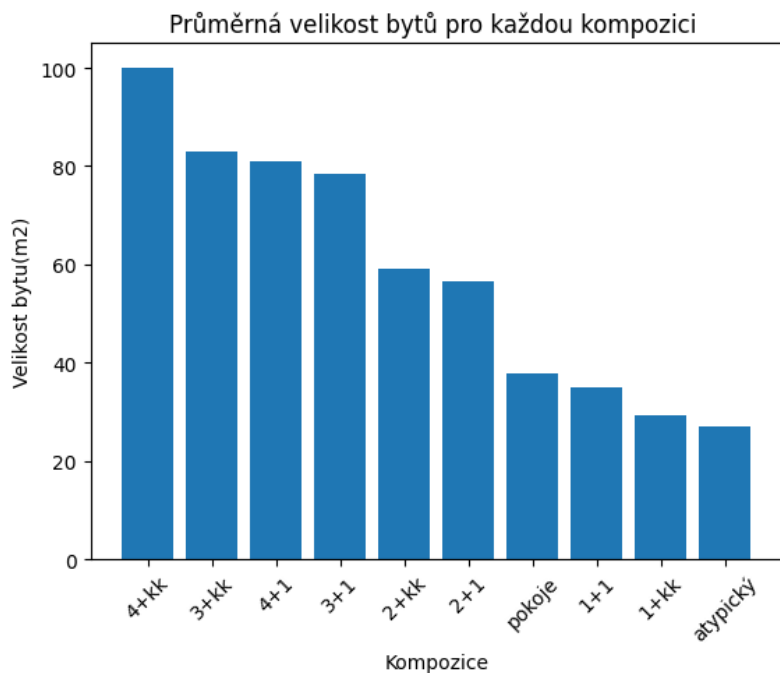
- **Jaká je průměrná cena bytů?**

Průměrná cena bytů je: **14130 Kč** za měsíc

- **Jaká je průměrná cena bytů pro každou kompozici (1+1, 2+1, atp.)? Zobrazte i na grafu.**



- Jaká je průměrná velikost bytu pro každou kompozici?



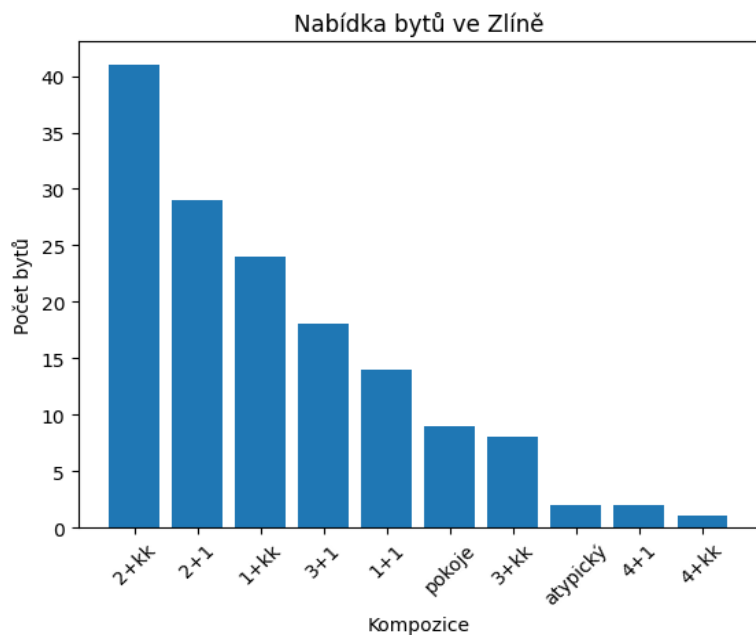
- Existuje ulice, kde je vyšší koncentrace dražších bytů?

Na ulici **Na Honech I** nejdražší **průměrná** cena za byty : 25000 Kč/měsíc.

Nejdražší byt se nachází v ulici Luhačovice, jehož cena je 39900 Kč/měsíc. To je 4+kk byt velikosti 100 m².

- Jaký typ kompozice je v daném městě nejčastěji inzerovan? Který to je?
Proč myslíte, že tomu tak je?

Byt s kompozicí 2+kk je v daném městě nejčastěji inzerovan, protože je to dobrý poměr cena/velikost bytu.



- Existují zde inzerce bytů, které stojí více, než 20.000,-? Pokud ano, jsou v této cenové hladině inzerovány i maximálně dvoupokojové byty (2+1/2+kk)?

Ano, existují zde inzerce bytů, které stojí více, než 20.000 Kč/měsíc. V této cenové hladině inzerovány byty 2+kk(9 bytu), 3+kk (3 byty), 4+1 (1 byt) a 4+kk (1byt)

- Zjistěte minimální a maximální inzerce pro každou kompozici bytu. Která kompozice má největší rozptyl mezi minimální a maximální inzerovanou cenou?

Kompozice **2+kk** a **3+kk** má největší rozptyl mezi minimální a maximální inzerovanou cenou(rozdil je 16000 Kč/měsíc)

...	Kompozici	mean	sum	max	min	Rozdil mezi max a min
0	1+1	10228.571429	143200.0	13500.0	4200.0	9300.0
1	1+kk	11200.000000	268800.0	14900.0	6500.0	8400.0
2	2+1	13346.551724	387050.0	16500.0	9000.0	7500.0
3	2+kk	16562.926829	679080.0	25000.0	9000.0	16000.0
4	3+1	16591.666667	298650.0	19000.0	11900.0	7100.0
5	3+kk	19125.000000	153000.0	29000.0	13000.0	16000.0
6	4+1	20000.000000	40000.0	21000.0	19000.0	2000.0
7	4+kk	39900.000000	39900.0	39900.0	39900.0	0.0
8	atypický	9750.000000	19500.0	12000.0	7500.0	4500.0
9	pokoje	6889.777778	62008.0	10500.0	3708.0	6792.0