

V tomto projektu budete analyzovat data ze stránky ``www.sreality.cz``. Budete tedy získávat aktuální nabídku bytů ve Zlíně.

Projekt se dělí na 2 části:

1. WebScraping

- pomocí knihovny ``BeautifulSoup`` získajte všechny aktuální inzerce bytů v daném městě.
- Získáte **tyto** informace: ``URL odkaz bytu, rozměry bytu, kompozice bytu, cena bytu, lokace bytu (Ulice+město)``

2. Datová analýza - ETL proces

- V první fázi provedte processing dat tak, aby obsahovala příslušné datové typy, nenulové hodnoty atp. Ulice i město bude v samostatných sloupcích.
- Zprocesovaná a vyčištěná data vyexportujte do `` .csv `` souboru s názvem ``{vašeJméno}_zdrojová_data.csv``
- Druhou část začněte tak, že si do dataframe nahrajete vámi vyexportovaný `` .csv `` soubor s vyscrapovanými daty.

3. Datová analýza - analýza dat

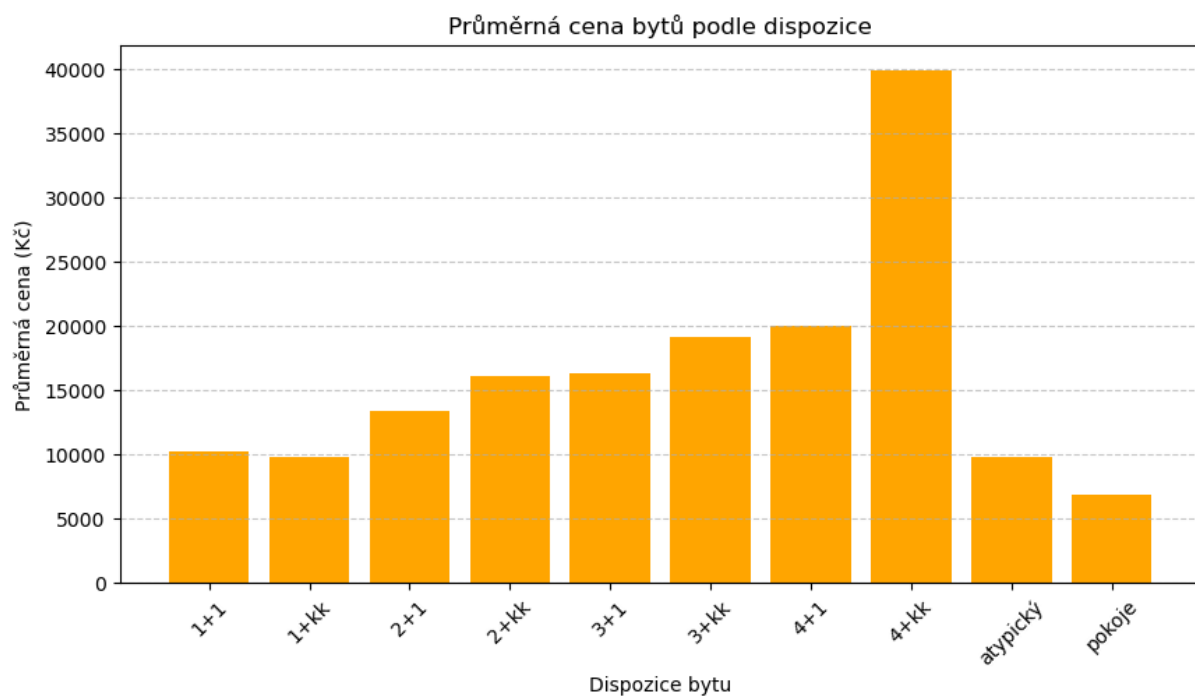
- Ve druhé fázi provedte analýzu dat a odpovězte na tyto dotazy:

- Jaká je průměrná cena bytů?

Průměrná cena bytů je 13.591,64 Kč

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

	dispozice	cena
0	1+1	10228.57
1	1+kk	9788.89
2	2+1	13335.19
3	2+kk	16107.18
4	3+1	16290.62
5	3+kk	19125.00
6	4+1	20000.00
7	4+kk	39900.00
8	atypický	9750.00
9	pokoje	6889.78



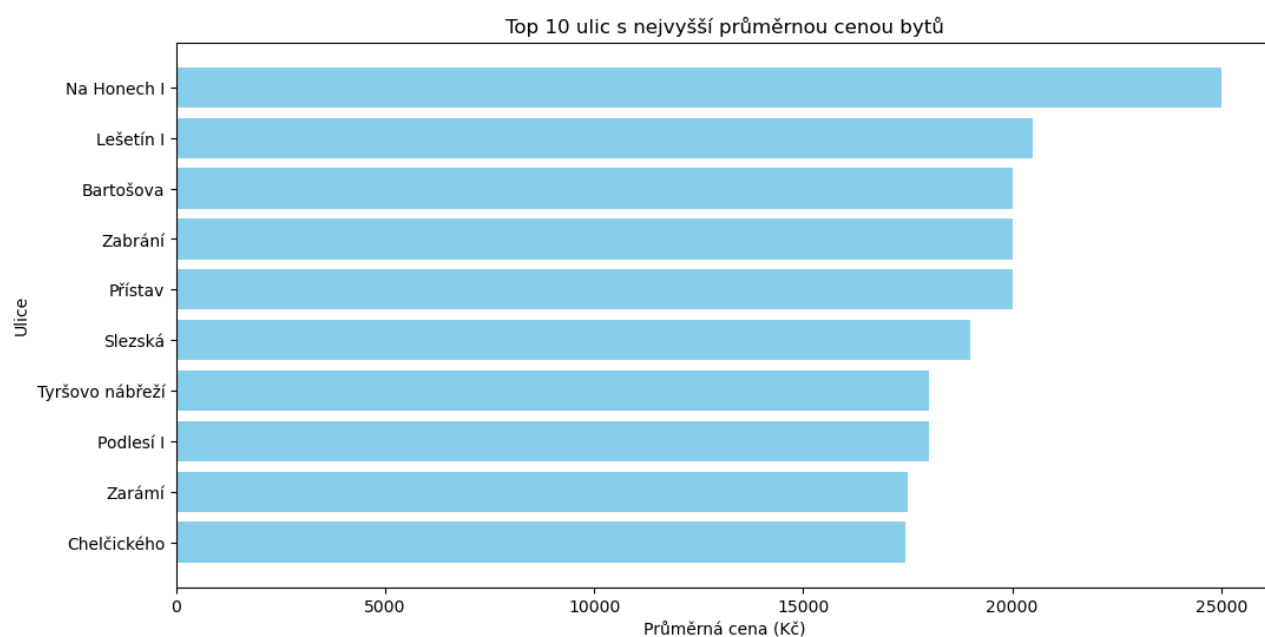
- Jaká je průměrná velikost bytu v m² pro každou dispozici?

	dispozice	prumerne_rozmary
0	1+1	35.07
1	1+kk	29.11
2	2+1	57.41
3	2+kk	59.51
4	3+1	78.94
5	3+kk	82.88
6	4+1	81.00
7	4+kk	100.00
8	atypický	27.00
9	pokoje	37.67

- Existuje ulice, kde je vyšší koncentrace dražších bytů?

Podle výsledků to vypadá, že vyšší koncentrace dražších bytů je v ulicích Na Honech I a Lešetín I.

	ulice	prumerna cena
31	Na Honech I	25000.0
24	Lešetín I	20500.0
0	Bartošova	20000.0
66	Zabrání	20000.0
49	Přístav	20000.0
53	Slezská	19000.0
59	Tyršovo nábreží	18000.0
44	Podlesí I	18000.0
69	Zarámí	17500.0
4	Chelčického	17450.0



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

Z výsledků víme, že 2+kk je nejčastější dispozicí v inzerátech.

	dispozice	pocet_inzeratu
0	2+kk	39
1	1+kk	27
2	2+1	27
3	3+1	16
4	1+1	14
5	pokoje	9
6	3+kk	8
7	atypický	2
8	4+1	2
9	4+kk	1

1. Populární mezi nájemníky a kupujícími

2+kk je často preferovaná velikost pro mladé páry, jednotlivce a investory (např. na pronájem).

Menší byty mají nižší celkovou cenu než větší jednotky, takže jsou dostupnější.

2. Ekonomické faktory

Výstavba 2+kk je nákladově efektivnější pro developery než větší byty.

Menší byty jsou rychleji prodejné než velké dispozice, což znamená vyšší poptávku.

3. Nabídka a poptávka v lokalitě

Pokud ve Zlíně převažují mladí lidé nebo menší domácnosti, mohlo by to vysvětlovat vysoký počet inzerátů na 2+kk.

Jak to můžeme potvrdit daty?

Můžeme se podívat na několik věcí:

1. Průměrná cena za m² pro různé dispozice – jsou 2+kk levnější na m²?
2. Průměrná velikost bytu u různých dispozic – jak velké jsou ve srovnání s ostatními?
3. Lokalita – kde se 2+kk vyskytují nejčastěji?

Možná se soustředí do určitých částí města, např. do novostaveb nebo oblastí s vysokou poptávkou.

- 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)?

Unnamed: 0			url	cena	rozmer	dispozice	lokace	ulice
7	7	https://www.sreality.cz/detail/pronajem/byt/2+...		25000	96	2+kk	Lešetín I, Zlín	Lešetín I
19	19	https://www.sreality.cz/detail/pronajem/byt/2+...		20690	64	2+kk	třída Svobody, Zlín - Malenovice	třída Svobody
113	113	https://www.sreality.cz/detail/pronajem/byt/2+...		25000	111	2+kk	Smetanova, Zlín	Smetanova
127	127	https://www.sreality.cz/detail/pronajem/byt/2+...		22000	85	2+kk	Luhačovice	Luhačovice
135	135	https://www.sreality.cz/detail/pronajem/byt/2+...		23000	120	2+kk	Nádražní, Luhačovice	Nádražní

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

	min	max	rozptyl
dispozice			
1+1	4200	13500	9300
1+kk	0	14900	14900
2+1	9000	16500	7500
2+kk	0	25000	25000
3+1	11900	19000	7100
3+kk	13000	29000	16000
4+1	19000	21000	2000
4+kk	39900	39900	0
atypický	7500	12000	4500
pokoje	3708	10500	6792

Největší rozptyl cen má dispozice 2+kk s rozptylem 25000 Kč.