**Závěrečný projekt – modul Python**

V tomto projektu jsme analyzovali data ze stránky www.sreality.cz pro aktuální nabídku pronájmu bytů ve Zlíně.

Projekt se dělí na 2 části:

**1. WebScraping** - pomocí knihovny BeautifulSoup získejte všechny aktuální inzerce bytů v daném městě.

- Získáte informace: URL odkaz bytu, rozměry bytu, kompozice bytu, cena bytu, lokace bytu (Ulice + město)

**2. Datová analýza - ETL proces**

- V první fázi proveďte 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 proveďte analýzu dat a odpovězte na tyto dotazy:

1. Jaká je průměrná cena bytů ?
2. Jaká je průměrná cena bytů pro každou kompozici (1+1, 2+1, atp.)? Zobrazte i na grafu.
3. Jaká je průměrná velikost bytu pro každou kompozici?
4. Existuje ulice, kde je vyšší koncentrace dražších bytů?
5. Jaký typ kompozice je v daném městě nejčastěji inzerován? Který to je? Proč myslíte, že tomu tak je?
6. 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)?
7. Zjistěte minimální a maximální inzerce pro každou kompozici bytu. Která kompozice má největší rozptyl mezi minimální a maximální inzerovanou cenou?

**Komentář k analýze dat a zjištěným výsledkům analýzy**

1. ***Jaká je průměrná cena bytů ?***

Průměrná cena všech nabízených bytů je: 14216,70 Kč/měsíc.

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

Rozměry bytu Průměrná cena

0 1+1 10228.57

1 1+kk 11200.00

2 2+1 13346.55

3 2+kk 17051.90

4 3+1 16450.00

5 3+kk 19125.00

6 4+1 20000.00

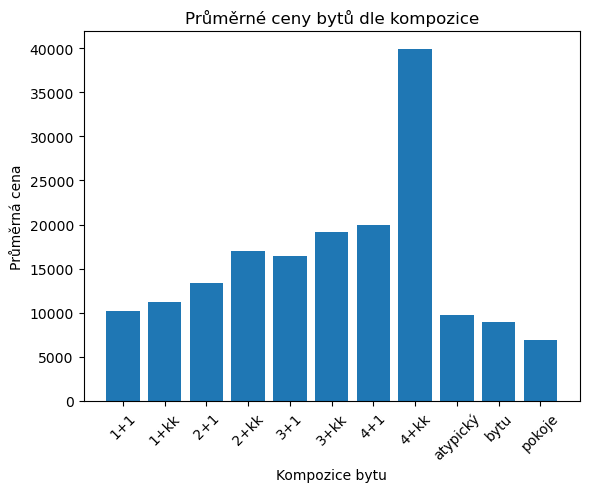
7 4+kk 39900.00

8 atypický 9750.00

9 bytu 9000.00

10 pokoje 6889.78

Zde uvádím výsledky zjištění promítnuté do grafu.



1. ***Jaká je průměrná velikost bytu pro každou kompozici?***

Kompozice bytu Průměrná velikost

0 1+1 35.07

1 1+kk 29.12

2 2+1 56.55

3 2+kk 61.31

4 3+1 78.59

5 3+kk 82.88

6 4+1 81.00

7 4+kk 100.00

8 atypický 27.00

9 bytu NaN

10 pokoje 37.67

1. ***Existuje ulice, kde je vyšší koncentrace dražších bytů?***

**Lokace Počet výskytů inzerce**

Bartošova, Zlín 1

Chelčického, Zlín - Malenovice 1

Dlouhá, Zlín 1

Lešetín I, Zlín 1

Luhačovice 6

Na Honech I, Zlín 1

Na Honech II, Zlín 1

Nádražní, Luhačovice 1

Přístav, Zlín - Příluky 1

Slezská, Zlín 2

Slunečná, Zlín 1

Smetanova, Zlín 2

Zabrání, Zlín - Malenovice 1

třída Svobody, Zlín - Malenovice 1

třída Tomáše Bati, Zlín 1

Uvažovala jsem jako hranici „dražšího“ pronájmu částku 18.900,-Kč/měsíc. S touto hodnotou bylo provedeny výpočty.

Nejvyšší počet nabízených „dražších“ bytů nad touto hranicí má lokace **Luhačovice.**

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

**Kompozice bytu /Počet inzercí**

2+kk 42

2+1 29

1+kk 24

3+1 17

1+1 14

pokoje 9

3+kk 8

4+1 2

atypický 2

4+kk 1

bytu 1

Podle výše uvedené tabulky je patrné zjištění, že nejvíce jsou mezi analyzovanými inzeráty nabízeny byty s **kompozicí 2+kk** nebo **2+1** . Je to přibližně polovina všech nabízených bytů. Na otázku „Proč je tento typ bytu inzerován nejčastěji?“ nám vstupní data neumožnila najít přesnou odpověď. Lze se jen domnívat, že jde o typ bytů, po kterých je největší poptávka, nebo snad že za tento typ bytu lze dostat vyšší nájemné.

1. ***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)?***

**Cena Výměra Kompozice bytu Lokace**

39900 100.0 4+kk Luhačovice

29000 95.0 3+kk třída Tomáše Bati, Zlín

25000 75.0 3+kk Na Honech I, Zlín

25000 96.0 2+kk Lešetín I, Zlín

25000 111.0 2+kk Smetanova, Zlín

23000 120.0 2+kk Nádražní, Luhačovice

22000 85.0 2+kk Luhačovice

22000 85.0 2+kk Luhačovice

22000 85.0 2+kk Luhačovice

22000 85.0 2+kk Luhačovice

22000 85.0 2+kk Luhačovice

21000 83.0 4+1 Slunečná, Zlín

20690 64.0 2+kk třída Svobody, Zlín – Malenovice

Při analýze dat bylo zjištěno, že je ve zkoumané nabídce je **celkem 13 nabízených bytů**, které stojí více než 20.000,-Kč/měsíc. Mezi nimi se nachází i **9 bytů s kompozicí 2+kk**. To možná potvrzuje hypotézu, že o byty 2+kk je největší zájem, pokud vycházíme ze zjištění, že mezi cenou za 2+kk a 3+kk je minimální nebo žádný rozdíl.

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

**Kompozice min max rozdíl**

**bytu**

2+kk 9000 25000 16000

3+kk 13000 29000 16000

1+1 4200 13500 9300

1+kk 6500 14900 8400

2+1 9000 16500 7500

3+1 11900 19000 7100

pokoje 3708 10500 6792

atypický 7500 12000 4500

4+1 19000 21000 2000

4+kk 39900 39900 0

bytu 9000 9000 0

Největší rozdíl mezi minimální a maximální nabízenou cenou byl zjištěn u bytů **2+kk a 3+kk**, kde **rozdíl ceny za pronájem činí 16.000,-Kč/ měsíc**.