

Apartamente

Vânzare apartamente Craiova



Luna 10  
(01.10.25 - 31.10.25)

2.070 €/mp

→ 0,0% (0 €/mp) față de luna 9/2025

↑ +18,4%(321 €/mp) față de luna 10/2024

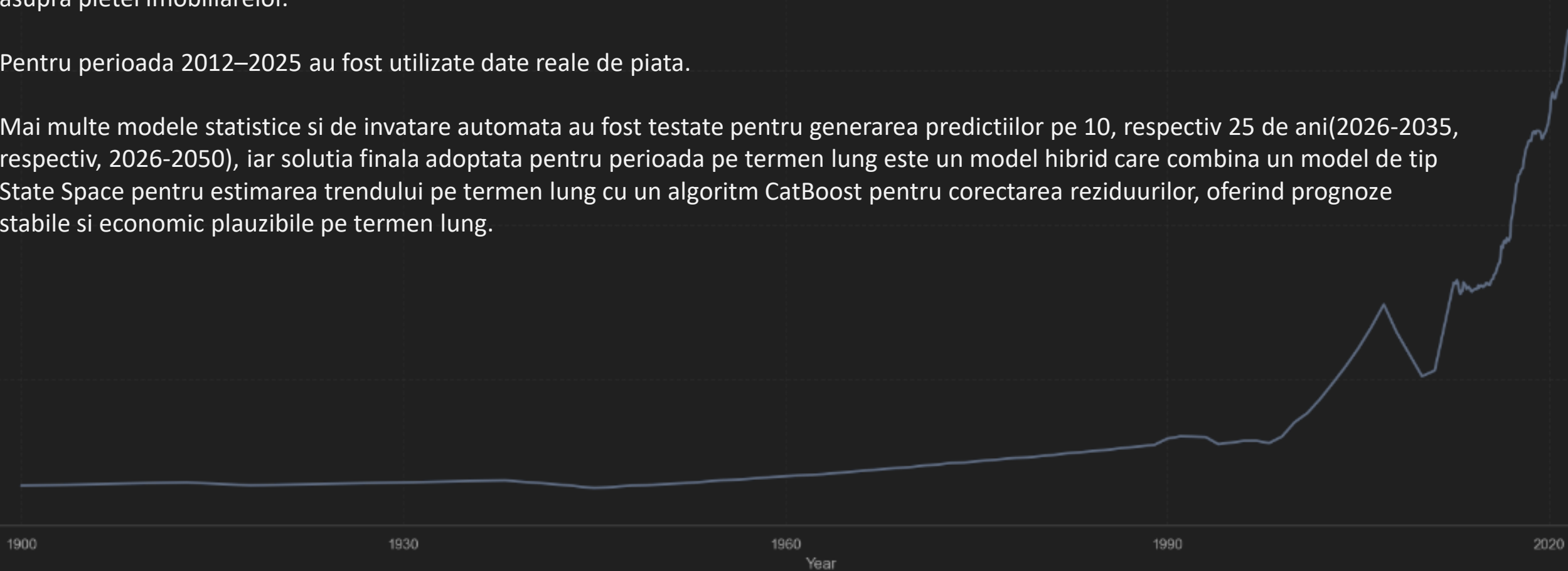
## *O scurta descriere ...*

Acest proiect urmareste analiza si prognoza evolutiei pretului apartamentelor pe metru patrat in municipiul Craiova, pentru perioada 1900–2050.

In lipsa datelor reale pentru intervalul anterior anului 2012, a fost construita o serie istorica sintetica (1900–2011) pe baza unor formule matematice simple, care a fost impartita pe perioade istorice importante(Comunism, WW1, WW2 etc.) si care si-au pus amprenta asupra pietei imobiliarelor.

Pentru perioada 2012–2025 au fost utilizate date reale de piata.

Mai multe modele statistice si de invatare automata au fost testate pentru generarea predictiilor pe 10, respectiv 25 de ani(2026-2035, respectiv, 2026-2050), iar solutia finala adoptata pentru perioada pe termen lung este un model hibrid care combina un model de tip State Space pentru estimarea trendului pe termen lung cu un algoritm CatBoost pentru corectarea reziduurilor, oferind prognoze stabile si economic plauzibile pe termen lung.



## Perioada 2012 - 2025

Ca sa putem face niste predictii avem nevoie de un dataset, iar o parte a acestui dataset trebuie obligatoriu sa aiba niste date reale(daca este posibil toate datele reale) verificate si culese atent.

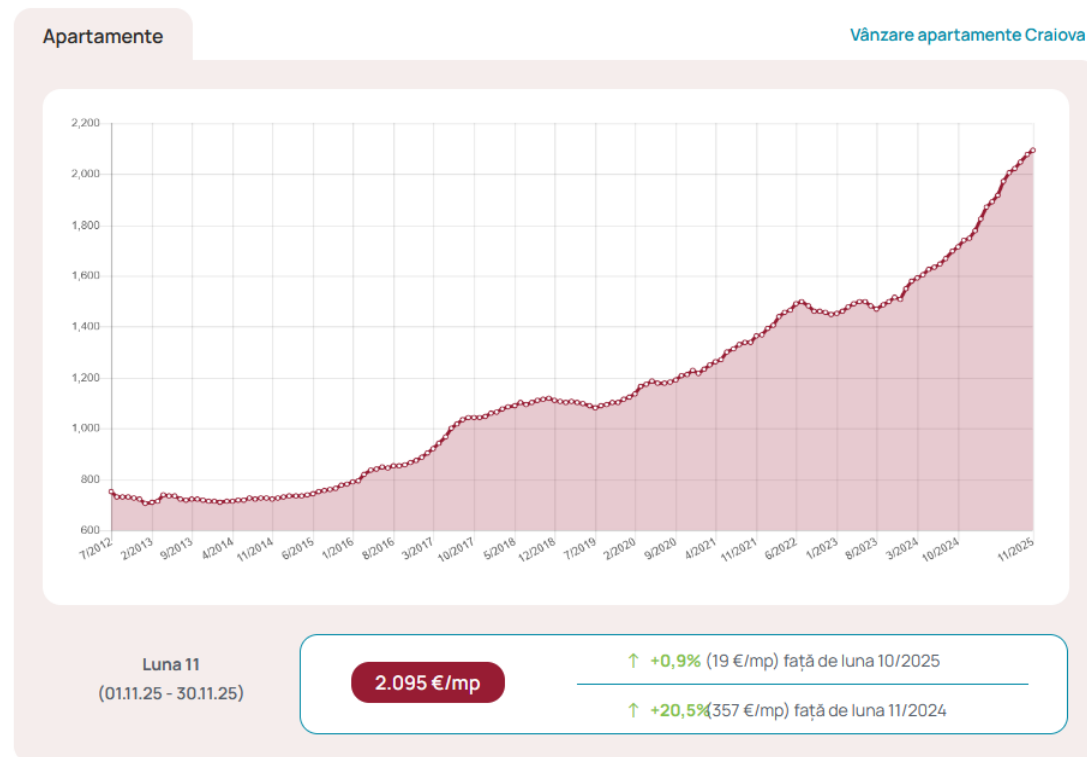
*Cum culegem si mai ales cum alegem datele noastre?*

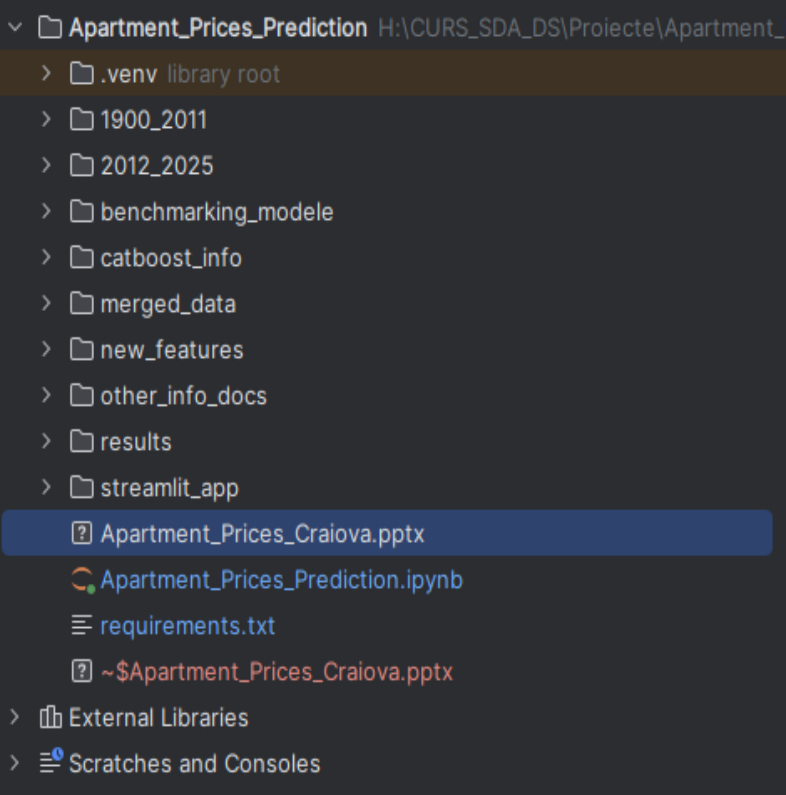
Simplu. Din surse sigure, postate pe site-uri de specialitate. Dar, aici intervine o problema:

Pentru Craiova, din pacate, date multe si corecte puse la atentia publicului larg nu exista, decat, in mare parte **platite**

De aici si adevarata provocare legata de proiectul de fata: *Cautarea si gasirea unor date din perioade istorice de mult apuse pentru predictiile noastre.*

Pentru intervalul 2012 – 2025 site-ul **imobiliare.ro** pune la dispozitia oamenilor curiosi grafice cu pretul pe metrul patrat pentru fiecare oras mare al tarii noastre. Asa ca, putem sa tragem o privire atenta asupra graficului pentru **Craiova**, de mai sus. Ulterior, am extras datele cu ajutorul unui **web scrapper** scris in **Python** si le-am salvat sub forma de fisier **CSV**.





Impartirea pe foldere a proiectului

Refacand conexiunea cu graficul pretului pe metrul patrat al apartamentelor, **dupa preluarea** datelor in CSV, datele au fost si **prelucrate**.

*De ce am prelucrat datele daca erau cele reale?*

*Pentru ca preturile la ora actuala sunt calculate in euro si le-am convertit in RON ca sa putem lucra cat mai bine cu ele.*

```
EUR_RON_AVG = {
    2012: 4.4577,
    2013: 4.4189,
    2014: 4.4441,
    2015: 4.4457,
    2016: 4.4899,
    2017: 4.5697,
    2018: 4.6542,
    2019: 4.7464,
    2020: 4.8364,
    2021: 4.9212,
    2022: 4.9307,
    2023: 4.9482,
    2024: 4.9753,
    2025: 5.0382,
}

def convert_euro_to_lei( 1 usage  @andrei-necsulea *
    input_csv="craiova_apartment_prices_2012_2025_euro.csv",
    output_csv="craiova_apartment_prices_2012_2025_lei.csv",
):
```



**craiova\_index\_page.html** - este pagina HTML descarcata de pe site pe care am aplicat web scrapper-ul si am extras CSV-ul cu datele exacte

**currency\_converter.py** – scriptul cu ajutorul caruia am convertit preturile din EUR in RON, avand grija sa pastream rata de conversie anuala( anul 2012 are o alta rata de conversie fata de 2025 spre exemplu)

## Perioada 1900 - 2011

Ideea de baza : Prezicerea viitorului si a trecutului pe baza prezentului( 2012 – 2025)

Totusi, daca nu am gasit date, asta nu inseamna ca studiul se incheie. DataScience-ul inseamna adaptabilitate, asa ca urmatoarul pas a fost sa generez datele. Cum? Cu ajutorul a 8 generari (incercari) pe care am sa le prezint pe rand (de la datele din prezent sa ajungem la datele din trecut):

**Gen 1** - bazata pe **formule derivate din perioade istorice impactante**(WW1, WW2, etc.)

**Gen 2** - “backcasting” bazat pe **State Space model**, practic o predictie inversa;

**Gen 3** - “backcasting” bazat pe modelul matematic al **Spline-ului**;

**Gen 4** - “backcasting” bazat pe modelul **State Space Logistic**;

**Gen 5** - “backcasting” bazat pe modelul **Norm State Space Logistic** ;

**Gen 6** - bazata pe **formule derivate din perioade istorice impactante**(WW1, WW2, etc.)

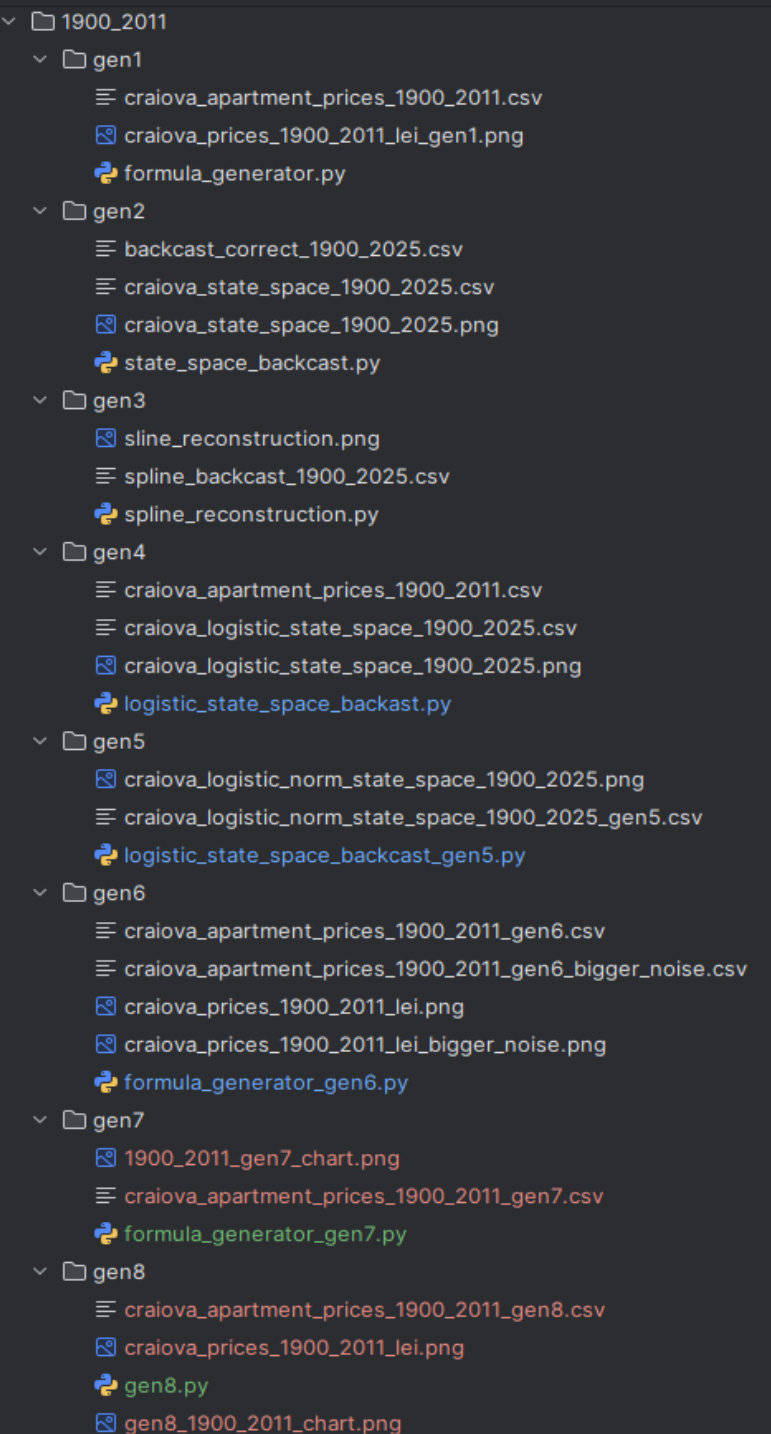
**Gen 7** - bazata pe **formule derivate din perioade istorice impactante**(WW1, WW2, etc.)

**Gen 8** - bazata pe **formule derivate din perioade istorice impactante**(WW1, WW2, etc.)

**Gen 1, Gen 6, Gen 7 si Gen 8** sunt toate incercarile bazate pe formule. DE ce? Pentru ca intre **Gen1** si **Gen6** s-a incercat generarea datelor cu ajutorul modelelor si s-a constatat ca la Gen1 nu au fost calibrate formulele. Initial, am continuat “punerea impreuna a datelor” cu datele de la Gen6, dar se constata ca **Gen 6** pleaca de la o presupunere necalculata, acelasi lucru intamplandu-se si la **Gen7**.

**Gen8 este corecta deoarece nu se mai pleaca de la presupunerea pretului pe metrul patrat la 1900, ci se incearca calcularea efectiva prin aducerea in discutie a mediei salariale de la 1900 si corelarea mediei cu devalorizarea leului (ROL), dinainte de 2005.**

Pentru aceasta perioada nu exista date oficiale, mai ales fiind aproape imposibil sa se fii tinut oficial evidenta inainte de 2000.



## Despre Generarea no. 8 ...

Seria istorica a pretului pe metru patrat este generata pornind de la o estimare economica de baza pentru anul 1900. Salariul mediu lunar din 1900 este estimat prin ajustarea inversa a salariului mediu lunar din anul 2025, folosind o rata anuala constanta de crestere ( $g = 2\%$ ) pe intreaga perioada 1900–2025.

Pretul initial pe metru patrat este calculat ca raport intre valoarea totala a salariilor obtinute intr-un numar mediu de ani necesari pentru achizitia unei locuinte si suprafata medie a unui apartament. Aceasta formula reflecta ideea economica conform careia pretul locuintei este proportional cu capacitatea de cumparare a populatiei.

Evolutia pretului este apoi simulata anual prin aplicarea unor rate de crestere sau scadere diferite, specifice principalelor perioade istorice (razboaie mondiale, comunism, tranzitie, boom imobiliar, criza), la care se adauga un termen aleator pentru a introduce volatilitate realista. Rezultatul este o serie temporala continua, coerenta economic, care reproduce tendintele majore ale pietei imobiliare in lipsa datelor reale.

```
# 2008-2010 (criza imobiliara)
for year in range(2008, 2011):
    current_price = current_price * (1 + np.random.uniform(-0.15, -0.10)) + np.random.normal(loc=0, scale=5)
    add(year, current_price)
```

Pentru anii 2008–2010, pretul pe metru patrat este actualizat anual folosind urmatoarea relatie: **pret\_curent = pret\_anterior  $\times$  (1 + r) + e**

```
def generate_historical_series(seed: int | None = 42) -> pd.DataFrame:
    """usage new"""
    if seed is not None:
        np.random.seed(seed)

    prices = []

    #salarii_lunare_1900 = [60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145, 150, 155, 160,
    #salariu_mediu_lunar_1900 = statistics.mean(salarii_lunare_1900)

    def adjust_salary_to_modern():
        """new"""
        salariu_mediu_2025 = 5387
        period = 2025-1900
        g = 0.02
        modern_salary = salariu_mediu_2025 / ((1.0 + g) ** period)
        return modern_salary

    salariu_mediu_lunar_1900 = adjust_salary_to_modern()

    suprafata_medie_apartament = 50
    cati_ani_de_plata = 2

    current_price = (salariu_mediu_lunar_1900 * 12 * cati_ani_de_plata)/suprafata_medie_apartament

    print(current_price)

    def add(y, p):
        """new"""
        prices.append({
            "date": date(y, month=1, day=1),
            "price_per_sqm": float(p)
        })

    # 1900-1913 (crestere lenta)
    for year in range(1900, 1914):
        current_price = current_price * 1.005 + np.random.normal(loc=0, scale=0.2)
        add(year, current_price)
```

unde:

- **r** este o rata negativa aleatoare, generata uniform in intervalul  $[-0.15, -0.10]$ , care modeleaza scaderea semnificativa a preturilor specifica perioadei de criza;
- **e** este un termen de zgomot aleator, extras dintr-o distributie normala cu medie zero si deviatie standard 5, care introduce fluctuatii anuale realiste.

Aceasta formula descrie o scadere abrupta a pretului, dar cu variatii de la un an la altul, reflectand instabilitatea si volatilitatea pietei imobiliare in timpul crizei financiare.



merged\_data directory

Practic, aici “lipim” rezultatele generate din 1900 - 2011, cu datele reale din 2012 – 2025.

merged\_data

corectare\_apartment\_prices\_csv.py

corectare\_apartment\_prices\_csv\_gen8.py

craiova\_apartment\_prices\_1900\_2025.csv

craiova\_apartment\_prices\_1900\_2025\_gen8.csv

craiova\_apartment\_prices\_1900\_2025\_year\_only.csv

craiova\_apartment\_prices\_1900\_2025\_yearly\_mean.csv

craiova\_apartment\_prices\_1900\_2025\_yearly\_mean\_gen8.csv

craiova\_prices\_1900\_2025.png

craiova\_prices\_1900\_2025\_gen8.png

full.csv

full\_gen8.csv

merge\_data\_gen6.py

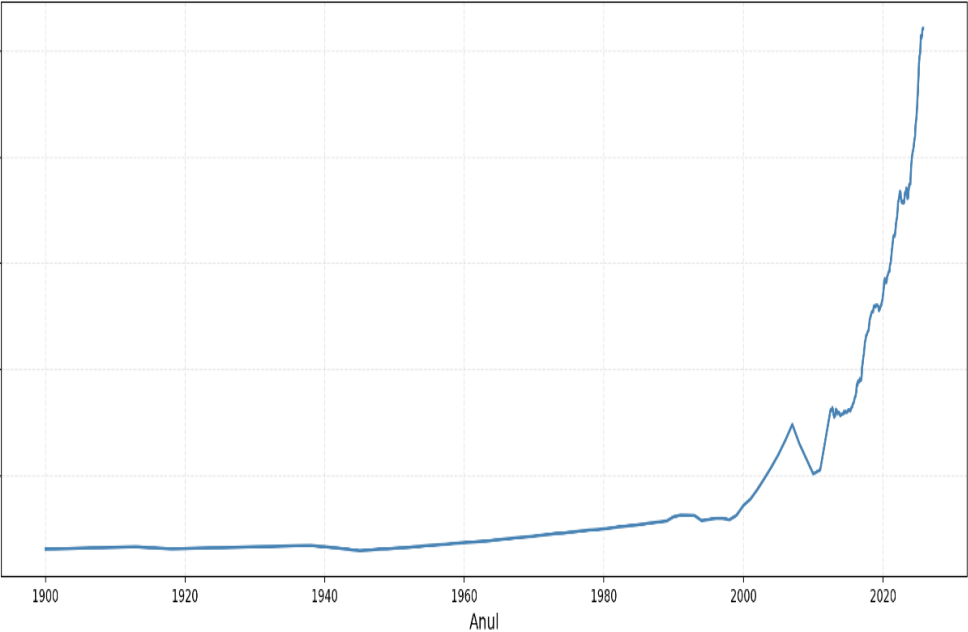
merge\_data\_gen8.py

new\_features\_added.py

new\_features\_added\_gen8.py

Evolutia pretului mediu pe metrul patrat (RON/sqm)

Craiova, 1900-2025

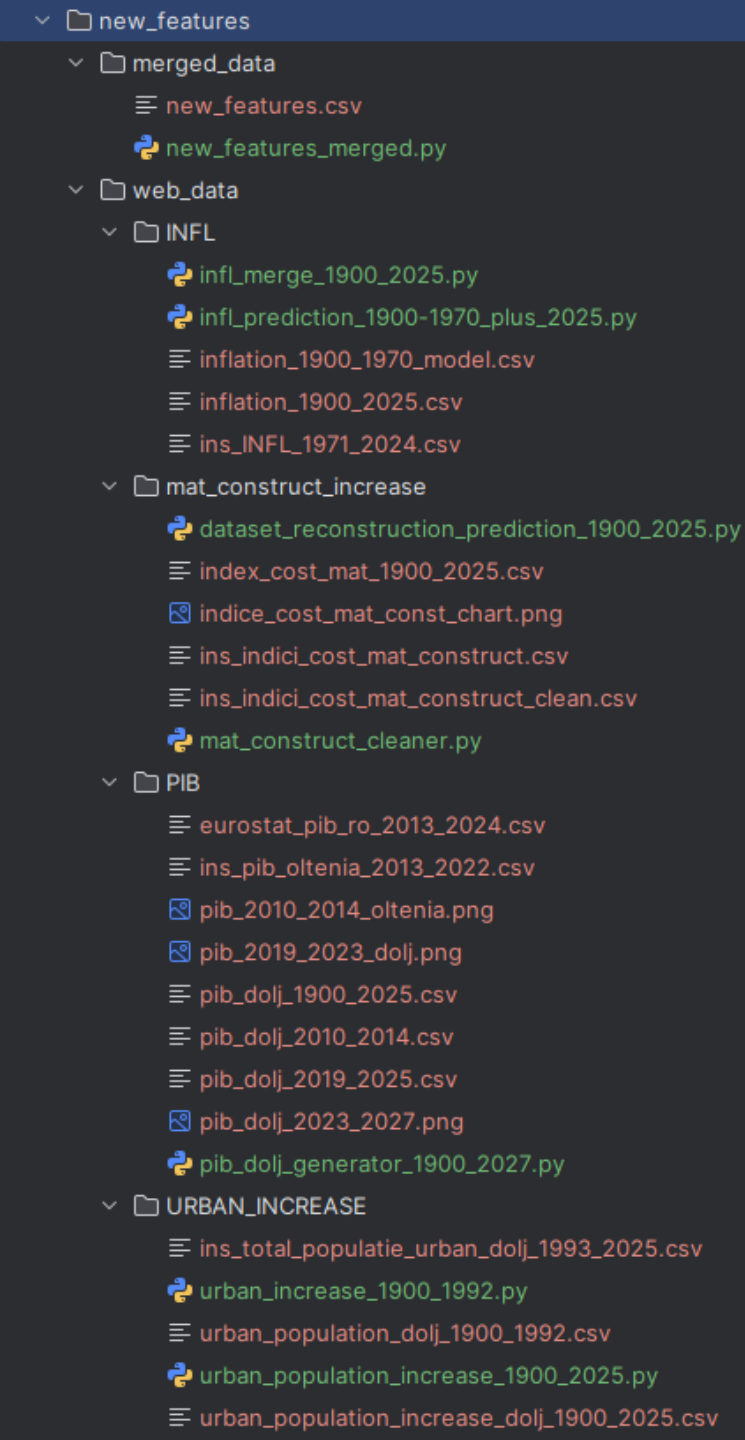


craiova\_apartment\_prices\_1900\_2025\_yearly\_mean\_gen8.csv

	An	price_per_sqm
1	1900	626.16
2	1901	629.21
3	1902	632.73
4	1903	636.76
5	1904	639.81
6	1905	642.88
7	1906	647.0
8	1907	650.67
9	1908	653.66
10	1909	657.23
11	1910	660.26
12	1911	663.29
13	1912	666.74
14	1913	668.98
15	1914	660.32
16	1915	653.07
17	1916	645.38
18	1917	639.29
19	1918	631.85
20	1919	634.25
21	1920	638.21
22	1921	641.28
23	1922	644.53
24	1923	646.97
25	1924	649.91

craiova\_apartment\_prices\_1900\_2025\_yearly\_mean\_gen8.csv

	An	price_per_sqm
101	2000	1448.67
102	2001	1567.42
103	2002	1749.42
104	2003	1955.95
105	2004	2170.25
106	2005	2405.86
107	2006	2673.6
108	2007	2973.11
109	2008	2615.96
110	2009	2329.88
111	2010	2043.36
112	2011	2118.08
113	2012	3237.03
114	2013	3179.4
115	2014	3191.97
116	2015	3305.01
117	2016	3722.13
118	2017	4474.12
119	2018	5075.79
120	2019	5208.78
121	2020	5692.44
122	2021	6397.56
123	2022	7171.29
124	2023	7350.96
125	2024	8344.82
126	2025	10017.45



### **new\_features:**

Acest director contine seturile de date auxiliare si scripturile folosite pentru reconstructia, curatarea si unificarea indicatorilor macroeconomici utilizati ca variabile explicative in modelele de predictie a preturilor imobiliare.

### **merged\_data:**

Contine fisierele rezultate in urma imbinarilor finale ale tuturor indicatorilor economici utilizati in proiect. Aici sunt centralizate datele privind inflatia, costurile de constructie, dinamica urbana si PIB-ul, aliniate temporal si pregatite pentru antrenarea modelelor de predictie.

### **web\_data/INFL:**

Include date istorice si reconstruite privind inflatia, pentru perioada 1900–2025. Sunt utilizate atat date reale (post-1970), cat si serii estimate pentru intervalele fara informatii oficiale, impreuna cu scripturi de predictie si validare a evolutiei inflatiei.

### **web\_data/mat\_construct\_increase:**

Contine date si scripturi pentru reconstructia indicelui costurilor de constructie pe termen lung, inclusiv curatarea seriilor INS si generarea unei evolutii continue pentru perioada 1900–2025.

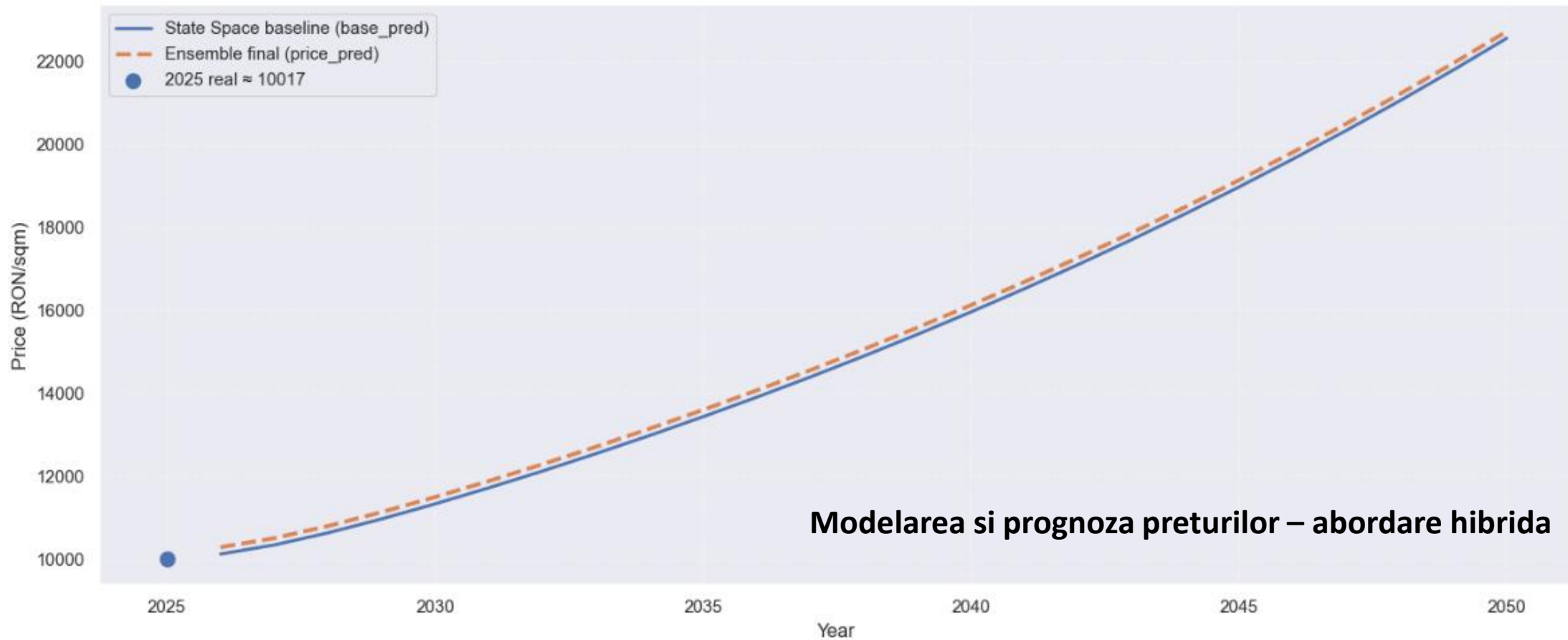
### **web\_data/PIB:**

Include date despre produsul intern brut la nivel regional (judetul Dolj si Oltenia). Sunt combinate surse reale (Eurostat, INS) cu serii generate pentru perioadele lipsa, rezultand o serie continua PIB 1900–2025 utilizata ca indicator al dezvoltarii economice locale.

### **web\_data/URBAN\_INCREASE:**

Contine date privind evolutia populatiei urbane si a gradului de urbanizare in judetul Dolj. Seria este construita prin combinarea datelor INS cu estimari pentru perioadele istorice, reflectand presiunea urbana asupra pietei imobiliare.





### Modelarea si prognoza preturilor – abordare hibrida

Pentru prognoza pretului pe metru patrat a fost adoptata o **abordare hibrida** care **combina un model de tip State Space cu un algoritm de invatare automata CatBoost**. Modelul State Space este utilizat pentru captarea trendului pe termen lung si a continuitatii structurale a seriei temporale, fiind capabil sa modeleze evolutii graduale si schimbari de nivel specifice proceselor economice. Acest model asigura o extrapolare stabila si economic plauzibila in afara intervalului de date observate.

CatBoost este utilizat complementar, fiind antrenat pe reziduurile modelului State Space. Rolul sau este de a surprinde relatiile neliniare si corectiile locale care nu pot fi modelate eficient printr-un model pur structural. Prin aceasta separare a rolurilor, CatBoost ajusteaza fin predictiile fara a distorsiona trendul global.

Utilizarea exclusiva a modelului State Space permite obtinerea unor prognoze stabile si coerente din punct de vedere economic, insa acestea pot subestima dinamica accelerata a pietei in anumite perioade. Pe de alta parte, modelele de invatare automata aplicate individual tind sa ofere rezultate bune pe intervale scurte, dar pot genera extrapolari instabile sau regresii catre valori medii atunci cand sunt utilizate pentru prognoze pe termen lung. Abordarea hibrida propusa valorifica avantajele ambelor tipuri de modele, oferind prognoze robuste, consistente si adaptate evolutiei reale a pietei imobiliare.

# WEBOGRAFIE

<https://www.imobiliare.ro/indicele-imobiliare-ro/craiova>

[https://ec.europa.eu/eurostat/databrowser/view/tec00001\\_custom\\_19266702/default/table](https://ec.europa.eu/eurostat/databrowser/view/tec00001_custom_19266702/default/table)

<https://insse.ro/cms/ro/content/ipc%E2%80%93serie-de-date-anuala>

<https://cnp.ro/economia-regionala-a-romaniei/>

<https://www.primariacraiova.ro/pozearticole/userfiles/files/RAPORT%202015%20final.pdf>

Other additional info : GOOGLE