



Analiza și Prognoza Seriilor de Timp

Capitolul 10: Recapitulare Completă

Analiză Completă cu Date Reale



Cuprins

- 1 Fluxul Complet de Analiză
- 2 Studiu de Caz 1: Analiza Volatilității Bitcoin
- 3 Studiu de Caz 2: Analiza Ciclului Petelor Solare
- 4 Studiu de Caz 3: Șomajul SUA cu Ruptură Structurală
- 5 Studiu de Caz 4: Analiza Multivariată VAR
- 6 Selecția Modelului: Ghid Practic
- 7 Rezumat și Concluzii Cheie

Prezentare Generală: Metode Studiate

Metode Clasice

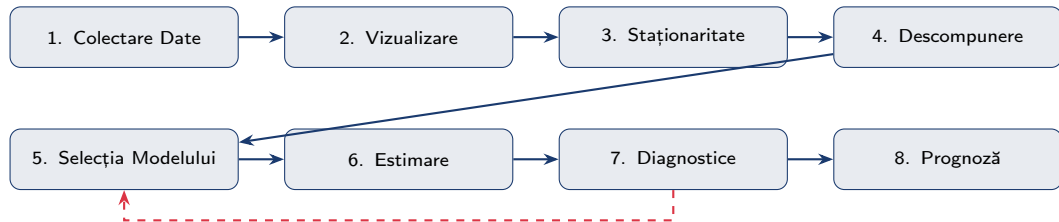
- Cap 1: Fundamentele Seriilor de Timp
- Cap 2: Modele ARMA
- Cap 3: Modele ARIMA
- Cap 4: Modele SARIMA
- Cap 5: Modele GARCH

Metode Avansate

- Cap 6: VAR & Cauzalitate Granger
- Cap 7: Cointegrare & VECM
- Cap 8: Extensii Moderne
- Cap 9: Prophet & TBATS

Astăzi: Aplicăm TOATE pe Date Reale!

Fluxul Complet de Analiză



Principiu Cheie

Diagnosticile pot necesita revenirea la selecția modelului (proces iterativ)

Seturi de Date Reale pentru Acest Capitol

Bitcoin

- Zilnic 2019-2024
- Clustering volatilitate
- ARIMA + GARCH

Pete Solare

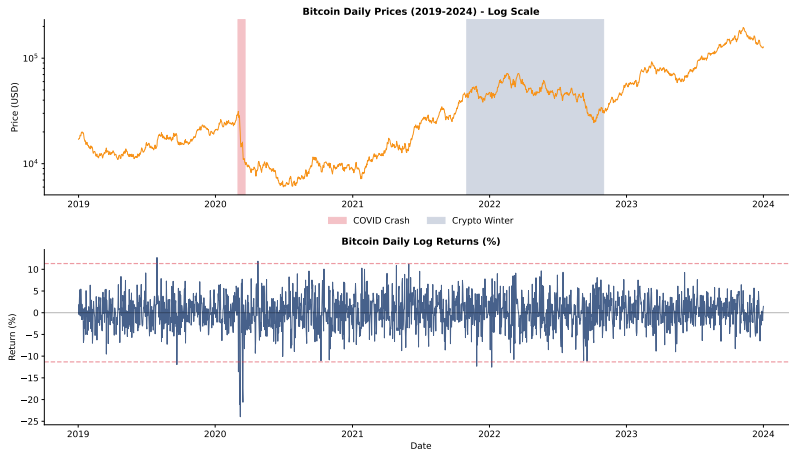
- Anual 1900-2023
- Ciclu de 11 ani
- Termeni Fourier

- Lunar 2010-2023
- Șocul COVID-19
- Prophet

Economic VAR

- Trimestrial 2000-2023
- PIB, Inflație, etc.
- VAR Multivariat

Bitcoin: Prezentare Generală a Datelor



- **Date:** Prețuri zilnice Bitcoin și randamente logaritmice (2019-2024)
- **Evenimente cheie:** Crahul COVID, bull run 2021, iarna crypto 2022

Pasul 1: Testarea Staționarității

Testul Augmented Dickey-Fuller

- H_0 : Rădăcină unitară (nestaționară)
- H_1 : Staționară

Rezultate pentru Bitcoin:

Serie	Statistică ADF	p-value
Prețuri	-0.87	0.79
Randamente Log	-42.1	< 0.001

⇒ Prețuri: nestaționare (random walk)

⇒ Randamente: staționare

Testul KPSS (Confirmare)

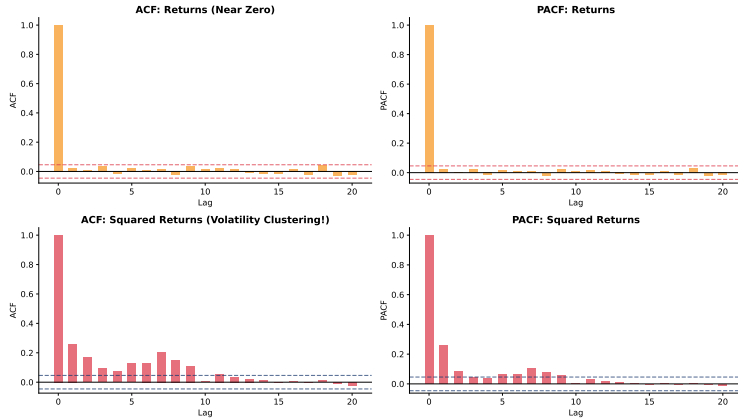
- H_0 : Staționară
- H_1 : Rădăcină unitară

Prețuri: KPSS = 5.83**

Randamente: KPSS = 0.12

Ambele teste confirmă: folosim randamente log!

Pasul 2: Analiza ACF/PACF a Randamentelor



- **Randamente:** Aproape zgomot alb (dependență liniară slabă)
- **Randamente pătratic:** Persistență puternică \Rightarrow clustering volatilitate
- **Implicație:** Modelul GARCH este esențial pentru Bitcoin!

Pasul 3: Model ARIMA pentru Randamente

Selecția Modelului folosind AIC/BIC:

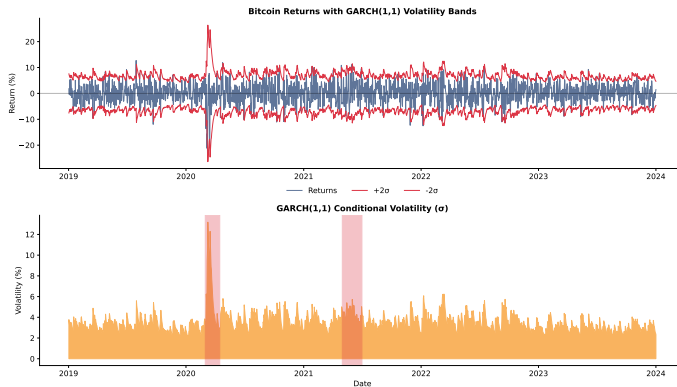
Model	AIC	BIC
ARIMA(0,0,0)	9524	9530
ARIMA(1,0,0)	9522	9534
ARIMA(0,0,1)	9523	9535
ARIMA(1,0,1)	9520	9538

Cel mai bun: ARIMA(1,0,1) dar îmbunătățire marginală

Observație Cheie

Randamentele crypto sunt notoriu imprevizibile.
“Alpha” este în înțelegerea **dinamicii volatilității**, nu în predicția direcției!

Pasul 4: Model GARCH pentru Volatilitate



Modelul GARCH(1,1):

$$\sigma_t^2 = \omega + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2$$

- Bitcoin arată $\alpha + \beta \approx 0.95$ (persistență ridicată)
- Perioadele COVID și Mai 2021 arată vârfuri masive de volatilitate

Bitcoin: Rezumatul Abordării



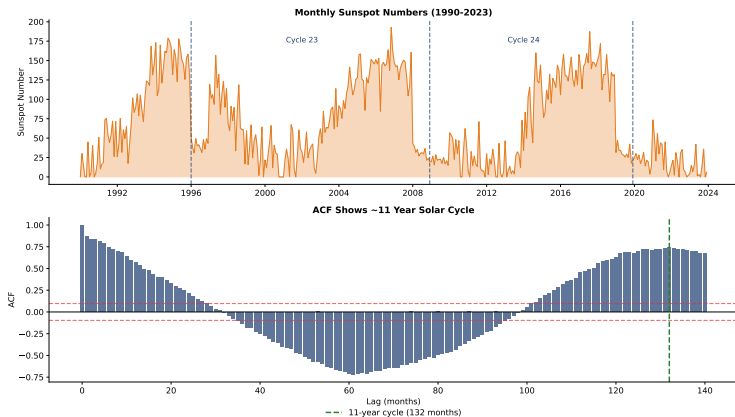
Constatări Cheie:

- Randamentele sunt aproape imprevizibile
- Clustering extrem de volatilitate
- GARCH captează dinamica riscului

Utilizare Practică:

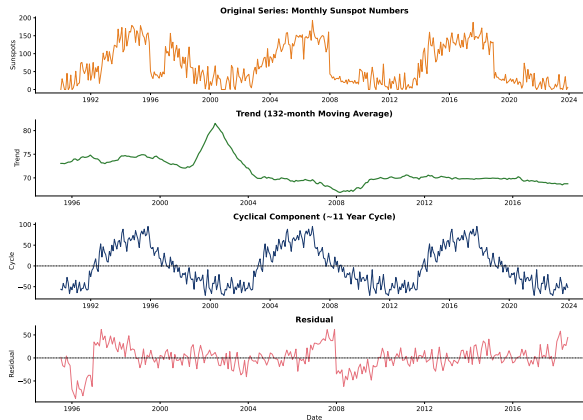
- Managementul riscului (VaR, CVaR)
- Dimensionarea pozițiilor
- Strategii de trading pe volatilitate

Pete Solare: Un Set de Date Clasic cu Ciclu Lung



- **Date:** Numere lunare de pete solare, 1990-2023
- **Caracteristici:** Celebrul ciclu solar de ~ 11 ani (132 luni)

Pasul 1: Analiza Descompunerii



- **Trend:** Media pe termen lung a activității solare
- **Ciclu:** Ciclul solar de 11 ani (ciclul Schwabe)
- **Provocare:** Perioadă sezonieră foarte lungă ($m = 132$)

Pasul 2: Gestionarea Sezonalității Lungi

Provocarea:

- SARIMA standard cu $m = 132$ necesită estimarea multor parametri
- Diferențierea sezonieră la lag 132 pierde 11 ani de date!

Opțiunea 1: Termeni Fourier

- Adăugăm regresori sinus/cosinus
- Perioadă = 132 luni
- Mai puțini parametri decât SARIMA complet

$$\sum_{k=1}^K \left[a_k \sin \left(\frac{2\pi kt}{132} \right) + b_k \cos \left(\frac{2\pi kt}{132} \right) \right]$$

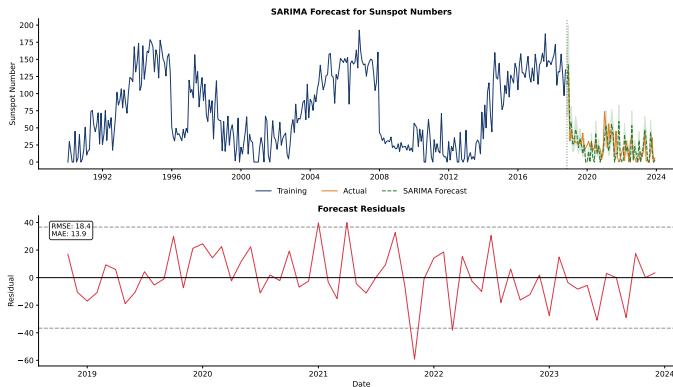
Opțiunea 2: Model AR

- AR de ordin înalt captează ciclul
- AR(12) sau AR(24) adesea suficient
- Simplu și eficient

Rezultat Clasic

Petele solare sunt bine modelate de modele AR(9) sau AR(12) (Yule, 1927)

Pasul 3: Prognoza SARIMA



Model: AR(12) cu termeni Fourier pentru ciclul de 11 ani

- Captează comportamentul cvasi-periodic
- Incertitudinea prognozei crește semnificativ cu orizontul

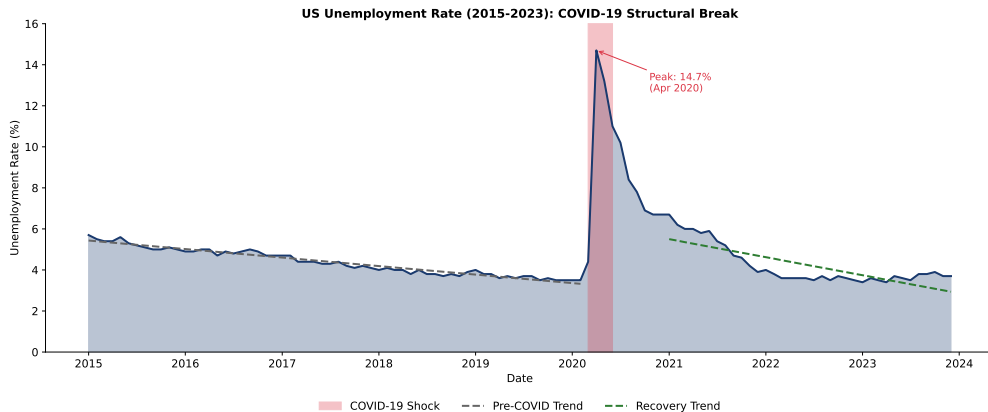
Model	RMSE	MAE	Note
AR(12)	28.4	22.1	Simplu, interpretabil
ARIMA(2,0,2) + Fourier	26.8	20.5	Captură bună a ciclului
TBATS	25.2	19.8	Detecrie automată a ciclului
Prophet	29.1	23.4	Mai puțin potrivit pentru cicluri lungi

Lecție Cheie

Pentru perioade sezoniere foarte lungi, considerați:

- Termeni de regresie Fourier
- TBATS (selecție automată a ciclului)
- Modele AR de ordin înalt

Șomajul SUA: Șocul COVID-19



- **Date:** Rata Șomajului SUA, lunar, 2015-2023 (BLS)
- **Șoc:** De la 3.5% la 14.7% într-o singură lună (Aprilie 2020)!

Opțiunea 1: Trunchierea Datelor

- Folosim doar date post-COVID
- Pro: Curate, fără rupturi
- Contra: Pierdem tipare istorice

Opțiunea 2: Variabile Dummy

- Adăugăm indicator COVID
- Pro: Folosește toate datele
- Contra: Complex în ARIMA

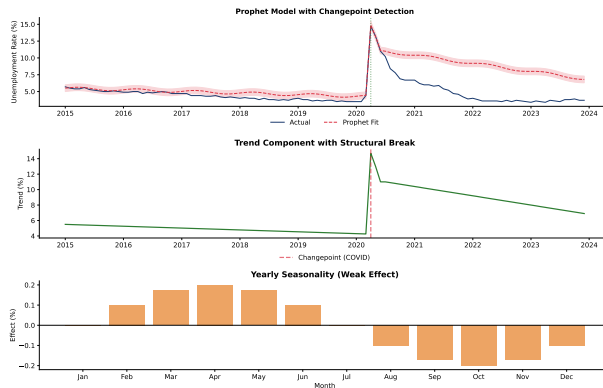
Opțiunea 3: Prophet cu Changepoints

- Detecție automată
- Pro: Gestionează rupturile natural
- Contra: Poate necesita tuning

Recomandare

Pentru date afectate de COVID, detecția changepoint-urilor Prophet sau modelele de schimbare de regim funcționează cel mai bine.

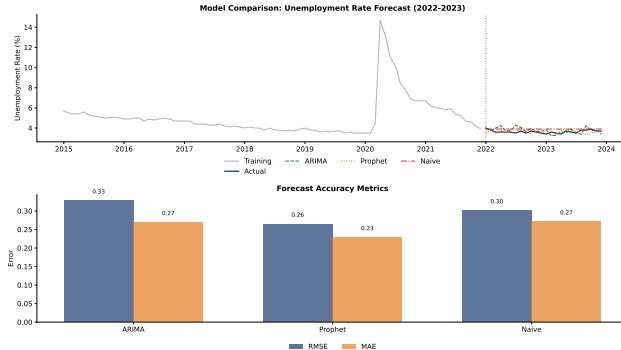
Prophet pentru Șomaj



Configurare Prophet:

- `changepoint_prior_scale = 0.5` (flexibil pentru șocul COVID)
- Changepoint automat în Aprilie 2020
- Captează tiparul de recuperare în V

Comparația Modelelor pe Șomaj



Lecție Cheie

Când datele au rupturi structurale extreme:

- ARIMA tradițional poate eșua sau necesită analiză de intervenție
- Flexibilitatea Prophet cu changepoints captează schimbările de regim
- Considerați modele de schimbare de regim (Markov-switching)

De Ce Analiză Multivariată?

Serie Singulară (Univariată):

- ARIMA, GARCH, Prophet
- O singură variabilă la un moment dat
- Nu captează dinamica între serii

Serii Multiple (Multivariate):

- **VAR**: Vector Autoregression
- **VECM**: Cu cointegrare
- Captează interdependențele

Când să Folosim VAR?

- Indicatori economici (PIB, inflație, șomaj)
- Piețe financiare (acțiuni, obligațiuni, FX)
- Variabile supply chain
- Orice sistem cu bucle de feedback

Modelul VAR: Bazele

Vector Autoregression VAR(p):

Pentru k variabile, VAR(p) este:

$$y_t = c + A_1 y_{t-1} + A_2 y_{t-2} + \cdots + A_p y_{t-p} + \varepsilon_t$$

Exemplu: VAR(1) cu 2 variabile

$$\begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} + \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} y_{1,t-1} \\ y_{2,t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}$$

Cheie: Fiecare variabilă depinde de lag-urile TUTUROR variabilelor

Caracteristici Cheie:

- Captează feedback dinamic
- Toate variabilele sunt endogene
- Permite teste de cauzalitate Granger
- Analiză impulse response

Variabile (Trimestrial, 2000-2023):

- Creștere PIB (YoY %)
- Rata Șomajului (%)
- Inflație (CPI YoY %)
- Rata Fed Funds (%)

Relații Așteptate:

- PIB $\uparrow \Rightarrow$ Șomaj \downarrow (Legea lui Okun)
- PIB $\uparrow \Rightarrow$ Inflație \uparrow (cerere-push)
- Inflație $\uparrow \Rightarrow$ Rata Fed \uparrow (Regula Taylor)

Surse Date (FRED):

- GDPC1: PIB Real
- UNRATE: Șomaj
- CPIAUCSL: Indicele Prețurilor
- FEDFUNDS: Rata Fed Funds

Cauzalitatea Granger

Definiție:

Variabila X *Granger-cauzează* Y dacă valorile trecute ale lui X ajută la predicția lui Y dincolo de ceea ce oferă doar valorile trecute ale lui Y .

Test:

- H_0 : X NU Granger-cauzează Y
- H_1 : X Granger-cauzează Y
- Test F pe restricții de coeficienți

Exemple de Rezultate:

Cauză → Efect	p-value
PIB → Șomaj	0.02*
Șomaj → PIB	0.15
Inflație → Fed	0.01**
Fed → Inflație	0.08

PIB conduce șomajul

Inflația conduce Rata Fed

Atenție

Cauzalitate Granger \neq Cauzalitate reală!

Măsoară cauzalitate *predictivă*, nu structurală.

Funcții de Răspuns la Impuls (IRF)

Ce este IRF?

Arată cum un șoc de o unitate în variabila X afectează toate variabilele în timp.

Analiza Șocului PIB:

- Șoc pozitiv PIB
- \Rightarrow Șomajul scade
- \Rightarrow Inflația crește
- \Rightarrow Fed crește ratele
- Efectele persistă câteva trimestre

Interpretare:

- Arată efectele multiplicator dinamice
- Benzile de încredere arată incertitudinea
- Poate identifica transmisia politicii

Utilizare în Politici

Băncile centrale folosesc IRF pentru a înțelege cum șocurile de politică monetară se propagă în economie.

VAR: Selecția Modelului și Diagnostiche

Pasul 1: Selecția Ordinului Lag

- Folosiți criterii informaționale (AIC, BIC)
- BIC tinde să selecteze modele mai simple
- Validare încrucișată pe date reținute

Pasul 2: Verificare Staționaritate

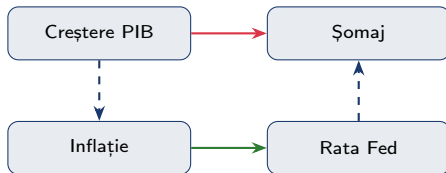
- Toate variabilele trebuie să fie staționare
- Sau folosiți VECM dacă sunt cointegrate
- Testați fiecare variabilă cu ADF

Pasul 3: Diagnostiche

- Autocorelație reziduuri (Portmanteau)
- Normalitate (Jarque-Bera)
- Stabilitate (verificare eigenvalue)

Pasul 4: Evaluare Prognoză

- RMSE pentru fiecare variabilă
- Comparați cu benchmark-uri univariate
- VAR adesea câștigă pentru sisteme interdependente

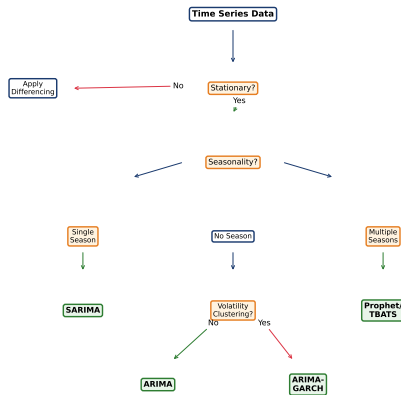


Constatări Cheie:

- PIB Granger-cauzează șomajul
- Inflația Granger-cauzează politica Fed
- VAR captează aceste dinamici

Aplicații Practice:

- Prognoză economică
- Analiză de politici
- Managementul riscului în portofolii



Rezumat Selecție Model

Tip Date		Caracteristici	Model Recomandat	Alternative
Randamente	finan- ciare	Fără trend, clustering vol.	ARIMA-GARCH	EGARCH, GJR
Sezonalitate simplă		Trend + o perioadă sezon.	SARIMA	ETS, Prophet
Cicluri lungi		Pete solare, cicluri business	AR + Fourier, TBATS	Metode spectrale
Rupturi structurale		COVID, schimbări regim	Prophet	ARIMA intervenție
Serii multiple		Interdependențe	VAR, VECM	Modele factoriale

Metrici de Evaluare a Prognozei

Metrici pentru Prognoze Punctuale:

RMSE (Root Mean Square Error):

$$\sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

MAE (Mean Absolute Error):

$$\frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

MAPE (Mean Absolute % Error):

$$\frac{100}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right|$$

Când să Folosim Fiecare:

- **RMSE**: Penalizează erorile mari
- **MAE**: Robust la outlieri
- **MAPE**: Independent de scală

Validare Încrucișată

Folosiți întotdeauna CV pentru serii de timp:

- Fereastră rulantă
- Fereastră expandabilă
- Nu amestecați niciodată!

Înțelegerea Datelor

- Vizualizare mai întâi!
- Testați staționaritatea (ADF, KPSS)
- Identificați tipare sezoniere
- Verificați rupturi structurale

Modele Clasice

- ARIMA: Date nesezoniere
- SARIMA: Sezonalitate simplă
- GARCH: Modelarea volatilității

Abordări Moderne

- Prophet: Interpretabil, gestionează rupturi
- TBATS: Sezonality multiple/lungi
- VAR/VECM: Serii de timp multiple

Cele Mai Bune Practici

- Verificați întotdeauna diagnosticele
- Folosiți validare încrucișată
- Comparați modele multiple
- Cunoașterea domeniului contează!

Recomandări Finale

- ❶ **Începeți Simplu:** Începeți cu vizualizare și statistici de bază
- ❷ **Testați Ipotezele:** Staționaritate, normalitate, independență
- ❸ **Iterați:** Model → Diagnostice → Îmbunătățire
- ❹ **Comparați:** Nu vă bazați niciodată pe un singur model
- ❺ **Validați:** Testarea out-of-sample este esențială
- ❻ **Comunicați:** Vizualizări și interpretări clare

Amintiți-vă

“Toate modelele sunt greșite, dar unele sunt utile.” — George Box

Scopul nu este predicția perfectă, ci perspective utile și prognoze rezonabile.

Întrebări?

Întrebări?

Pași Următori:

- Exersați cu notebook-ul Jupyter
- Aplicați aceste metode pe propriile date
- Comparați diferite modele pe același set de date

Materiale Curs: github.com/danpele/Time-Series-Analysis