

Podstawy Przetwarzania Szeregów Czasowych w Pythonie

SZYMON MOLIŃSKI, CODE ME

Tydzień
Programisty
2021

Czym są szeregi czasowe?

Jak opisujemy szereg czasowy?

Projekt #1: Rozkład szeregu

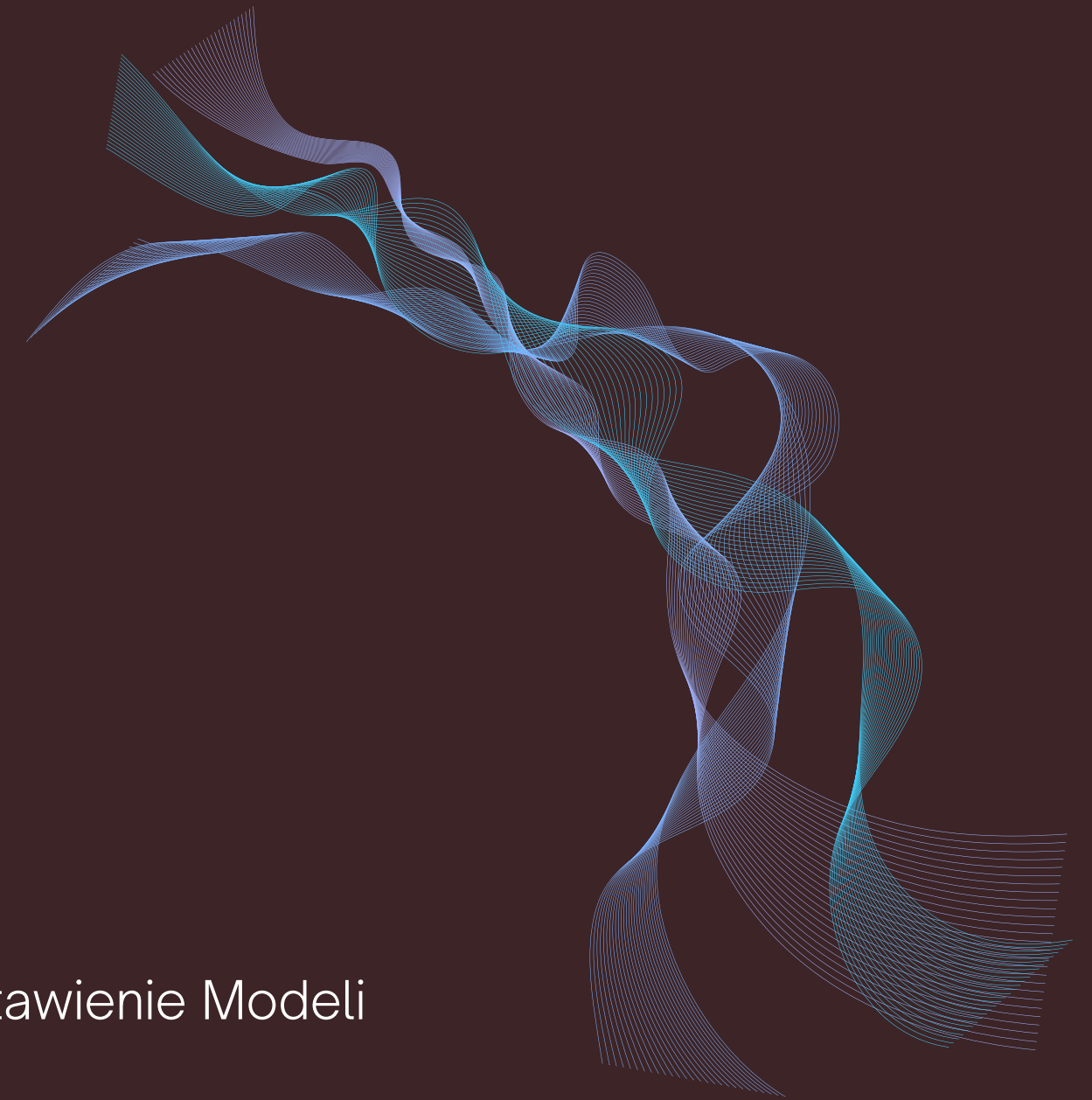
Projekt #2: Losowanie

Projekt #3: Powtórzenie (naiwny model)

Projekt #4: Regresja Wielomianowa

Projekt #5: Holt-Winters

Zadanie Domowe: Zestawienie Modeli

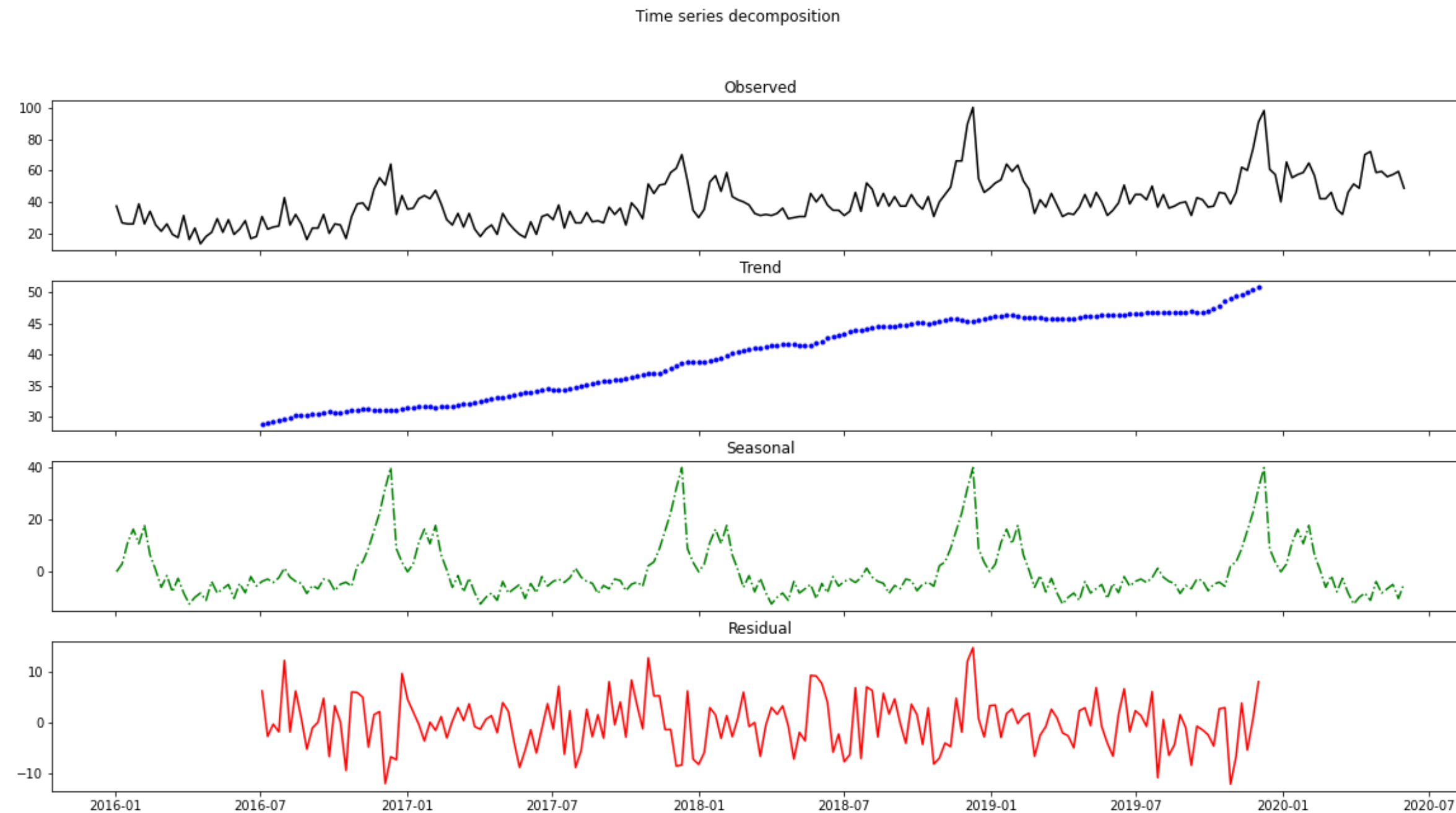


Program warsztatów

Tydzień
Programisty
2021

Czym jest szereg czasowy?





Składowe szeregu czasowego

Tydzień
Programisty
2021

Projekt #1: rozkład szeregu czasowego



Wczytaj plik `ts_data.csv`

NARZĘDZIA:

Pakiet `pandas`:

- Funkcja `read_csv()` i parametry `parse_dates`, `index_col`



Rozłóż szereg czasowy na czynniki

NARZĘDZIA:

Pakiet `statsmodels`:

- Funkcja `seasonal_decompose()`



Zwizualizuj wyniki

NARZĘDZIA:

Pakiet `matplotlib.pyplot`:

- Funkcja `subplots()`



Projekt #2: Losowanie



Napisz funkcję która losuje n wartości z przedziału 0-100

NARZĘDZIA:

Pakiet `numpy.random`



Napisz funkcję obliczającą RMSE między dwoma wektorami

NARZĘDZIA:

Pakiet `numpy` i narzędzia Data Science



Wydziel zbiór treningowy (90%) i testowy (10%)

NARZĘDZIA:

Pakiet `pandas`



Oblicz ŚREDNIE RMSE między każdym krokiem w zbiorze testowym a wylosowaną wartością

NARZĘDZIA:

Pakiet `pandas` | `numpy`



Stwórz model, który uznaje, że wartość przewidywana powinna być równa ostatniej wartości znanej

NARZĘDZIA:

Pakiet `numpy` | `pandas`



Oblicz RMSE dla modelu naiwnego dla kroku w przyszłość równego 1.

NARZĘDZIA:

Pakiet `numpy` | `pandas` i narzędzia Data Science

Projekt #3: Naiwny Model



Korzystając z narzędzi scikit-learn i statsmodels zbuduj model predykcyjny oparty na regresji wielomianowej od poziomu 1 do poziomu 4, bazujący na 5 ostatnich rekordach ze zbioru treningowego wykrywający 5 pierwszych wartości zbioru testowego.

PROCES:

1. Wykorzystaj funkcję `poly_model()` stworzoną w części teoretycznej.
2. Napisz funkcję, która tworzy obiekt `PolynomialFeatures` zadanego poziomu i modeluje przebieg szeregu przy wykorzystaniu klasy `OLS`. Opieraj się na przykładach z zajęć. Funkcja powinna zwracać tylko przewidziane wartości.
3. Stwórz ramkę danych z predykcjami dla każdego wielomianu.
4. Oblicz RMSE dla każdego wielomianu.

Projekt #4: regresja wielomianowa





Korzystając z narzędzi statsmodels zbuduj model predykcyjny oparty na modelu Holt-Winters

PROCES:

1. Z pakietu `statsmodels.tsa.holtwinters` zaimportuj `ExponentialSmoothing`
2. Stwórz funkcję do prognozowania opartą na modelu Holt-Winters, z założeniami, że trend jest typu addytywnego, a liczba rekordów per sezon wynosi 54.
3. Przetestuj RMSE.

Projekt #5: Holt-Winters (Wygładzanie Wykładnicze)



Zestawienie przewidywań kluczem do sukcesu



(Zadanie domowe)

Korzystając z pakietu pandas stwórz trzy ramki danych z zestawieniem wyników dla 5 kroków naprzód względem każdego modelu.

Zwizualizuj przebiegi. Wyciągnij średnią i odchylenie standardowe dla każdej daty i stwórz trzy nowe przebiegi: $[\text{średnia} - 2 \cdot \text{odchylenie}]$ | $[\text{średnia}]$ | $[\text{średnia} + 2 \cdot \text{odchylenie}]$. Zwizualizuj te trzy przebiegi.

Dziękuję!

MAIL

szymon.molinski@codeme.pl

MATERIAŁY

https://github.com/szymon-datalions/conferences-and-workshops/tree/main/2021/09/2021_09_TP_CODEME