Ćwiczenia 7

1. (S) Określ ile dni minęło pomiędzy pierwszym stycznia roku 1901, a 31 grudnia roku 2000.

2. Zbiór danych `wages` z pakietu `TSA` zawiera informacje na temat średniej płacy (w dolarach) pracowników przemysłu tekstylnego w USA od lipca 1981 do czerwca 1987.

   - Dokonaj wizualizacji szeregu czasowego.
   - Dopasuj liniowy oraz kwadratowy trend do tych danych. Dodaj oba do wykresu.
   - Wyznacz wartości AIC oraz BIC dla obu modeli. Który model należy wybrać sugerując się tym kryteriami.
   - Zwizualizuj standaryzowane residua obu modeli. Sprawdź normalność reszt.
   - Który model jest poprawny?

3. (S) W zbiorze danych `female.txt` znajdują się dane na temat liczby bezrobotnych kobiet (w tysiącach) w wieku od 16 do 19 lat od lipca 1961 roku do grudnia 1985 roku w USA. Wczytaj dane i przekształć je do szeregu czasowego. Narysuj dane wraz z dopasowaną średnią ruchomą rzędu 17 (dopasowanie kolorem czerwonym). Czy występuje sezonowość badanego zjawiska?

4. W zbiorze danych `unemp.txt` znajdują się dane na temat liczby bezrobotnych robotników w budownictwie w Niemczech od lipca 1975 roku do września 1979 roku. Wczytaj dane i przekształć je do szeregu czasowego. Narysuj dane wraz z dopasowaną średnią ruchomą rzędu 12 (kolorem czerwonym). Czy występuje sezonowość badanego zjawiska? Wykonaj wygładzanie wykładnicze Holta-Wintersa i dodaj otrzymany model do rysunku (kolorem niebieskim).

5. Zidentyfikuj następujące procesy:

   - $y_t = y_{t-1} - 0{,}25y_{t-2} + e_t - 0{,}1e_{t-1}$,
   - $y_t = 2y_{t-1} + y_{t-2} + e_t$,
   - $y_t = 0{,}5y_{t-1} - 0{,}5y_{t-2} + e_t - 0{,}5e_{t-1} + 0{,}25e_{t-2}$.

6. **(S)** Zbadaj stacjonarność poniższych procesów:

   - $y_t = \frac{3}{2}y_{t-1} - \frac{1}{2}y_{t-2} + \varepsilon_t$,
   - $y_t = \frac{5}{6}y_{t-1} - \frac{1}{6}y_{t-2} + \varepsilon_t$,

- $y_t = \frac{2}{3}y_{t-1} - \frac{5}{3}y_{t-2} + \varepsilon_t - \frac{3}{2}\varepsilon_{t-1}$.

7. Zbiór danych `robot` z pakietu `TSA` zawiera informacje o odległościach (w calach) robota od punktu docelowego po wykonaniu określonych czynności.
   - Bazując na wykresie stwierdź czy proces jest stacjonarny.
   - Dopasuj do tych danych modele AR(1) oraz ARIMA$(0, 1, 1)$. Który z nich jest lepszy?
   - Dokonaj prognozy 5 przyszłych wartości za pomocą lepszego modelu.
8. **(S)** Zbiór danych `gold` z pakietu `TSA` zawiera informacje na temat cen złota dla 252 dni w roku 2005.
   - Zwizualizuj proces.
   - Zwizualizuj różnice logarytmów tych danych.
   - Zaproponuj odpowiedni model ARIMA pasujący do przekształconych danych.
9. **(*)** *Ruchoma średnia wykładnicza.* Niech $y_1, y_2, \ldots, y_n$ będzie szeregiem czasowym o długości n. Ruchomą średnią wykładniczą tego szeregu definiujemy wybierając wpierw początkową wartość m_0 oraz współczynnik wygładzania δ. Następnie wyznaczamy $m_1, m_2, \ldots, m_n$ rekurencyjnie w następujący sposób:

$$e_t = y_t - m_{t-1},$$
$$m_t = m_{t-1} + (1 - \delta)e_t,$$

   dla $t = 1, 2, \ldots, n$. Napisz funkcję, która dla zadanego szeregu czasowego wyznacza tę średnią ruchomą oraz rysuje szereg wraz z wygładzeniem. Następnie wykorzystaj ją do zbioru `LakeHuron` (za m_0 przyjmij poziom wody na początku badanego okresu oraz niech $\delta = 0{,}9$).