

## ДОДАТОК А

### Фрагменти коду

```
import numpy as np, pandas as pd
from statsmodels.graphics.tsaplots import plot_acf, plot_pacf
import matplotlib.pyplot as plt
plt.rcParams.update({'figure.figsize':(9,7), 'figure.dpi':120})

Import data
df = pd.read_csv('onlineretail.csv', names=['value'], header=0)
fig, axes = plt.subplots(3, 2, sharex=True)
axes[0, 0].plot(df.value); axes[0, 0].set_title('Original Series')
plot_acf(df.value, ax=axes[0, 1])

Differencing
axes[1, 0].plot(df.value.diff()); axes[1, 0].set_title('1st Order Differencing')
plot_acf(df.value.diff().dropna(), ax=axes[1, 1])
axes[2, 0].plot(df.value.diff().diff()); axes[2, 0].set_title('2nd Order Differencing')
plot_acf(df.value.diff().diff().dropna(), ax=axes[2, 1])

plt.show()

Find order
plt.rcParams.update({'figure.figsize':(9,3), 'figure.dpi':120})

fig, axes = plt.subplots(1, 2, sharex=True)
axes[0].plot(df.value.diff()); axes[0].set_title('1st Differencing')
axes[1].set(ylim=(0,5))
plot_pacf(df.value.diff().dropna(), ax=axes[1])

plt.show()
```

Find order

```
fig, axes = plt.subplots(1, 2, sharex=True)
axes[0].plot(df.value.diff()); axes[0].set_title('1st Differencing')
axes[1].set(ylim=(0,1.2))
plot_acf(df.value.diff().dropna(), ax=axes[1])
```

plt.show()

Build model

```
model = ARIMA(df.value, order=(1,1,2))
model_fit = model.fit(dispatch=0)
print(model_fit.summary())

model = pm.auto_arima(df.value, start_p=1, start_q=1,
                      test='adf',      # use adftest to find optimal 'd'
                      max_p=3, max_q=3, # maximum p and q
                      m=1,             # frequency of series
                      d=None,          # let model determine 'd'
                      seasonal=False,  # No Seasonality
                      start_P=0,
                      D=0,
                      trace=True,
                      error_action='ignore',
                      suppress_warnings=True,
                      stepwise=True)
```

```
print(model.summary())
```

```
#> Fit ARIMA: order=(1, 2, 1); AIC=525.586, BIC=535.926, Fit time=0.060 seconds
#> Fit ARIMA: order=(0, 2, 0); AIC=533.474, BIC=538.644, Fit time=0.005 seconds
#> Fit ARIMA: order=(1, 2, 0); AIC=532.437, BIC=540.192, Fit time=0.035 seconds
#> Fit ARIMA: order=(0, 2, 1); AIC=525.893, BIC=533.648, Fit time=0.040 seconds
```

#> Fit ARIMA: order=(2, 2, 1); AIC=515.248, BIC=528.173, Fit time=0.105 seconds  
#> Fit ARIMA: order=(2, 2, 0); AIC=513.459, BIC=523.798, Fit time=0.063 seconds  
#> Fit ARIMA: order=(3, 2, 1); AIC=512.552, BIC=528.062, Fit time=0.272 seconds  
#> Fit ARIMA: order=(3, 2, 0); AIC=515.284, BIC=528.209, Fit time=0.042 seconds  
#> Fit ARIMA: order=(3, 2, 2); AIC=514.514, BIC=532.609, Fit time=0.234 seconds

ДОДАТОК Б  
Слайди презентації

# Аналіз замовлень в інтернет-магазині з застосуванням СТАТИСТИЧНИХ МЕТОДІВ

СТУДЕНТ ІПЗМ-17-1: КУШНАРЕНКО О.О.  
КЕРІВНИК РОБОТИ: ВЛАСЕНКО Л.А.

## Зміст презентації

- ▶ Передмова
- ▶ Цілі
- ▶ Методи
- ▶ Результат
- ▶ Висновки

# Електронна комерція



## Електронна комерція

У 2019 в інтернеті вже було створено замовлень на 3.5\$

К 2040 року очікується 95% закупівель через інтернет

80% людей у розвинених країнах робили хоча б 1 замовлення в інтернеті

Майже 20% закупівель сьогодні

## Цілі бізнесу

- ▶ Заробити більше грошей
  1. Збільшувати об'єми продажів
  2. Масштабувати бізнес
  3. Поширити круг споживачів
  4. Провести персоналізацію
- ▶ Зменшувати проміжні витрати
  1. Зберігання на складі
  2. Нестача товарів

## Цілі роботи

- ▶ Створити систему якаб могла прогнозувати інформацію для бізнесу
- ▶ Зручно вбудовуватися в готові рішення
  1. Створити бібліотку
  2. Створити API для клієнтів

## Як прогнозування може допомогти досягненню цілей

- ▶ Дозволяє зрозуміти, у яких регіонах користувачу потребуватимуть більше товарів
- ▶ Розуміти, які товари частіше повертають
- ▶ Скоротити витрати на зберігання непотрібних товарів

## Джерело даних

- ▶ Замолення інтернет магазину
- ▶ Для розробки - <https://archive.ics.uci.edu>

Дані містять 500 000 замовлень в Британському інтернет магазині

Атрибути:

- назва товару
- кількість
- ціна

## Способи прогнозування даних

- ▶ Регрессионные модели и методы
- ▶ Авторегрессионные модели и методы
- ▶ Модели и методы экспоненциального сглаживания
- ▶ Нейросетевые модели и методы

## Авторегрессійні методи

- ▶ При відносній простоті ці методи мають велику кількість способів застосування.
- ▶ Достатньо універсальні для різних сфер

## Інструменти розробки

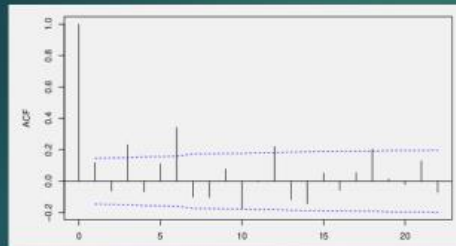
- ▶ Модель авторегресії — ковзного середнього
- ▶ Мова Python с бібліотеками



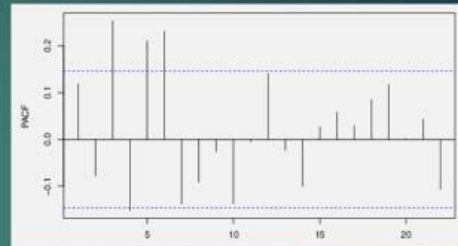
## ARMA

- ▶ Для використання моделі ARMA ( $p, q$ ), необхідно знайти відповідні  $p$  та  $q$ .
- ▶ Поєднує в собі модель авторегресії (autoregressive, AR) та модель ковзного середнього (moving average, MA)
- ▶ Є однією з найпопулярніших моделей

## Підбір параметрів



Автокореляційна функція



Приватна автокореляційна функція

## Результати роботи





ДЯКУЮ ЗА УВАГУ