**Практичне завдання**

Виконав: Колдаєв Дмитро, група КА-03мп

Прогнозування фінансових процесів сучасними методами

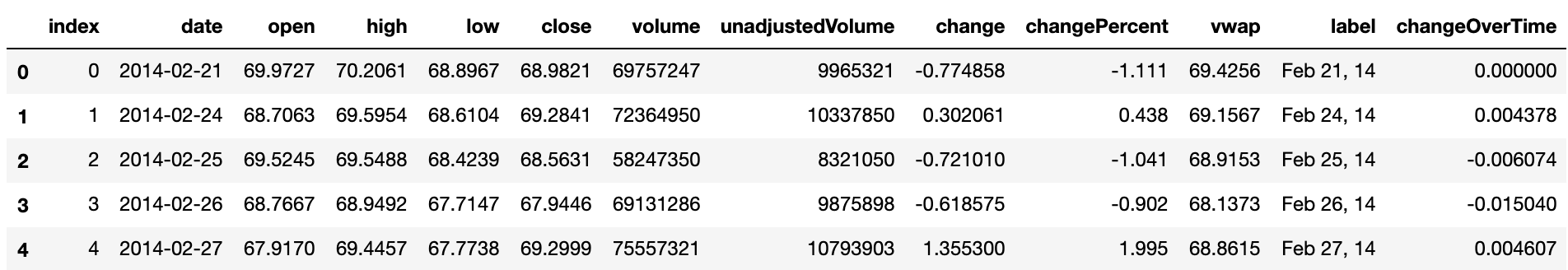
Мета: *навчитися застосовувати сучасні методи прогнозування для*

*аналізу та моделювання реальних фінансових процесів, що характеризуються високою волатильністю і потребують побудови моделей з урахуванням їх гетероскедастичності та сезонності.*

**Опис вхідних даних**

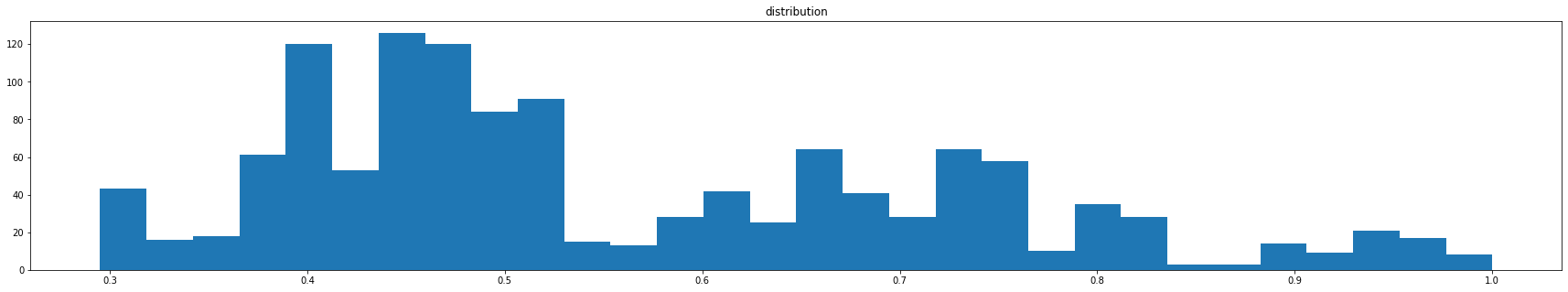
У якості вхідних даних були використані акції компанії Apple за період з 2014-02-21 по 2019-02-20 взяті на платфомі Kaggle

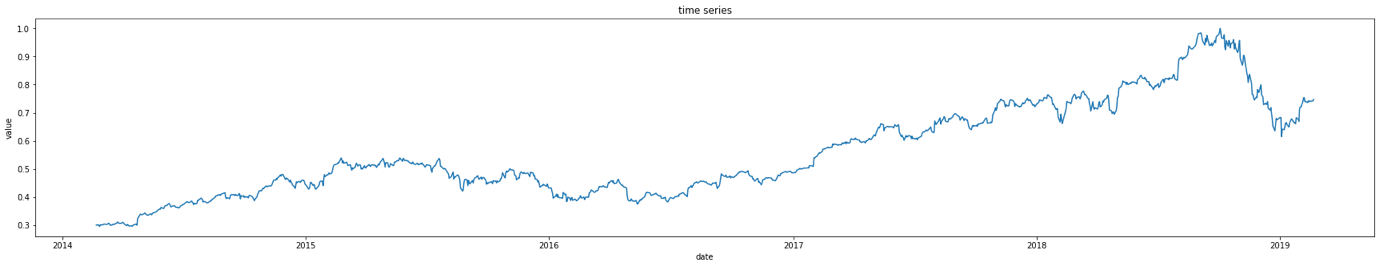
https://www.kaggle.com/timoboz/stock-data-dow-jones?select=AAPL.csv



У якості прогнозованої величини використовувався показник “сlose” - показник ціни на акції на кінець дня.

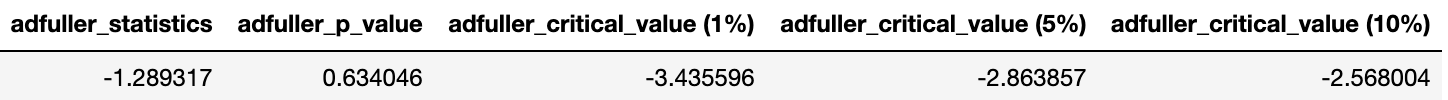
**Експлораторний аналіз**





Для побудови моделей було проведено лінійну інтерполяцію на вихідному наборі даних з метою позбавлення від пропущених значень в цільовій змінній.

**Перевірка стаціонарності**



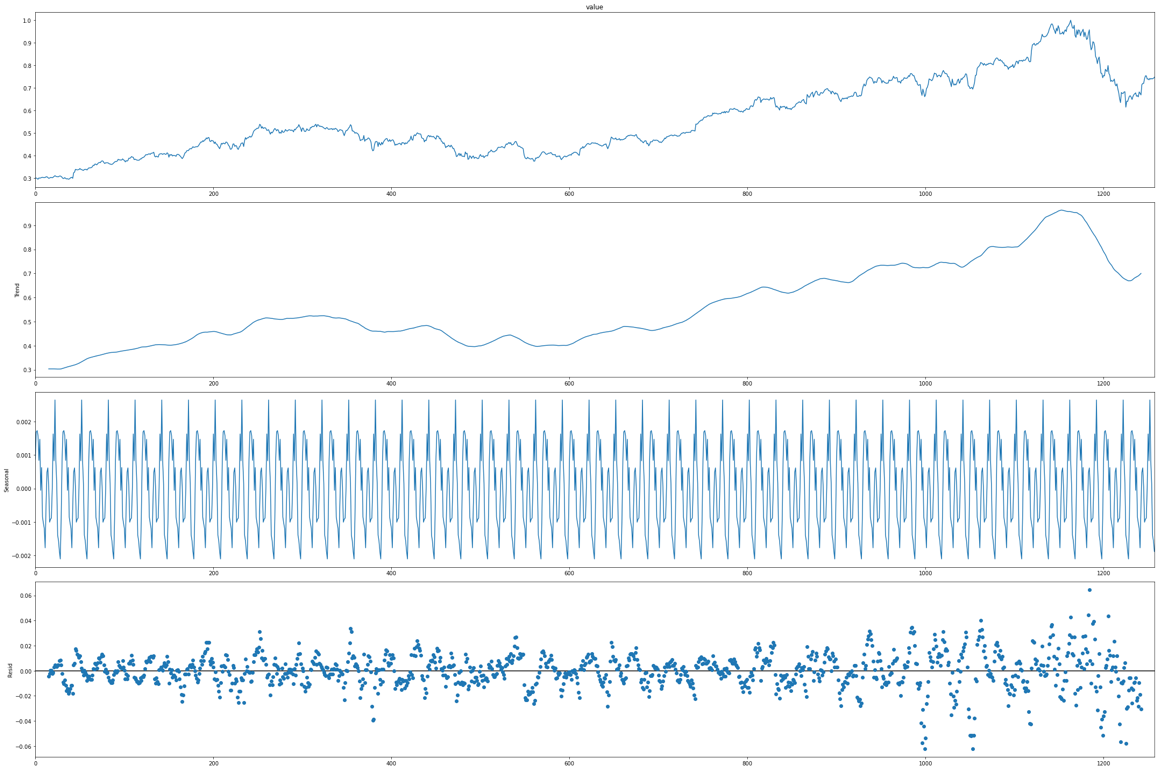


Тести на стаціонарність Дікі-Фюллера та KPSS показуть відсутність стаціонарноті у данному ряді

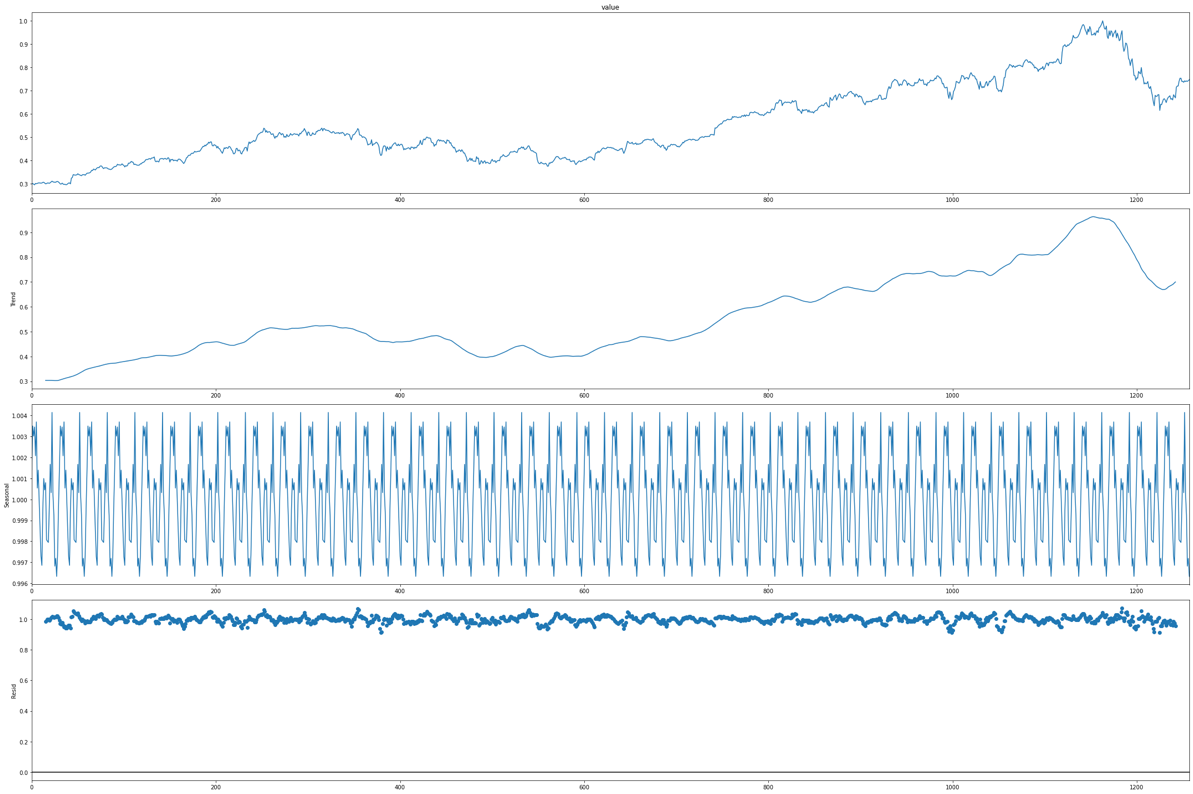
**Декомпозиція часового ряду**

Побудуємо декомпозицію часового ряду за допомогою адитивної та мультиплікативної моделей з періодом 30

Адитивна

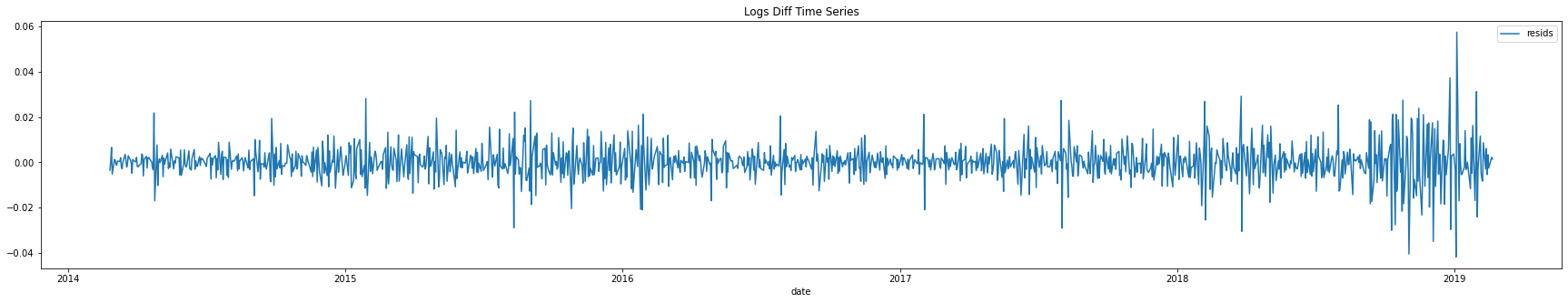


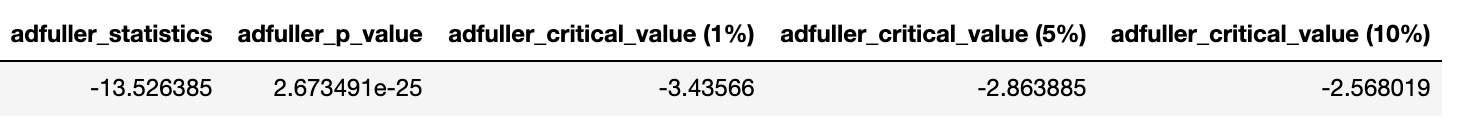
Мультиплікативна



**Побудова моделей**

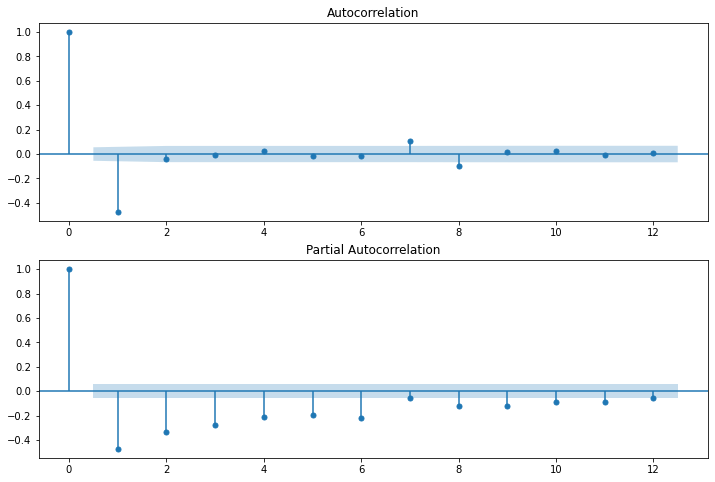
Моделі AR, ARMA будуємо на другій різниці логарифмованого числового ряду адже саме такий ряд виявився стаціонарним



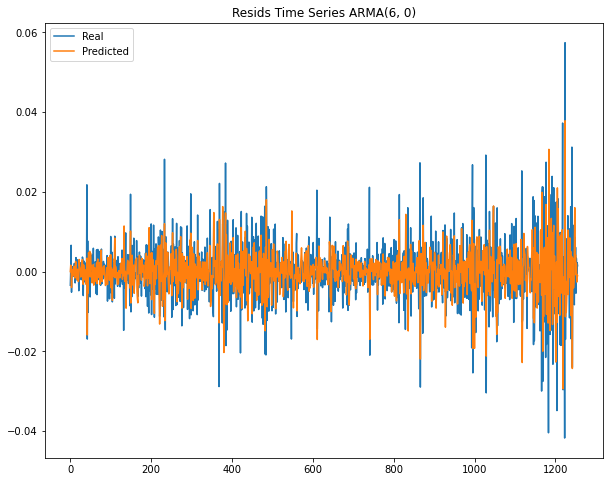
**Модель AR**

Для побудови моделі AR побудуємо АКФ та ЧАКФ нашого ряду

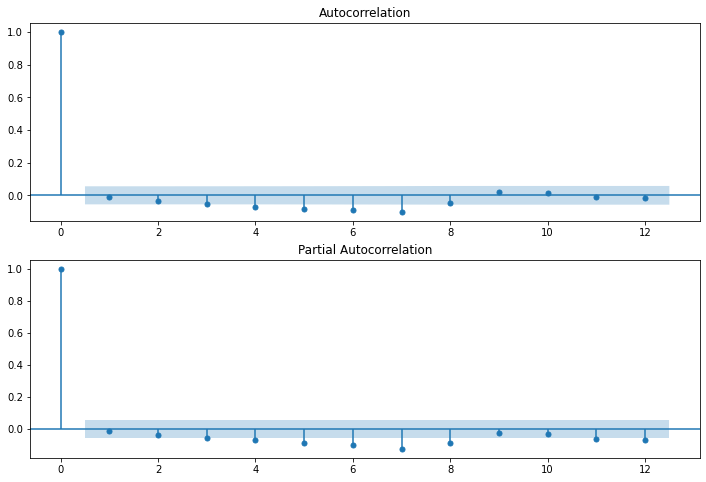


Взявши рівень значущості 0.2 далі будемо будувати моделі з авторегресійним лагом 6

Для обраного лагу побудуємо модель AR



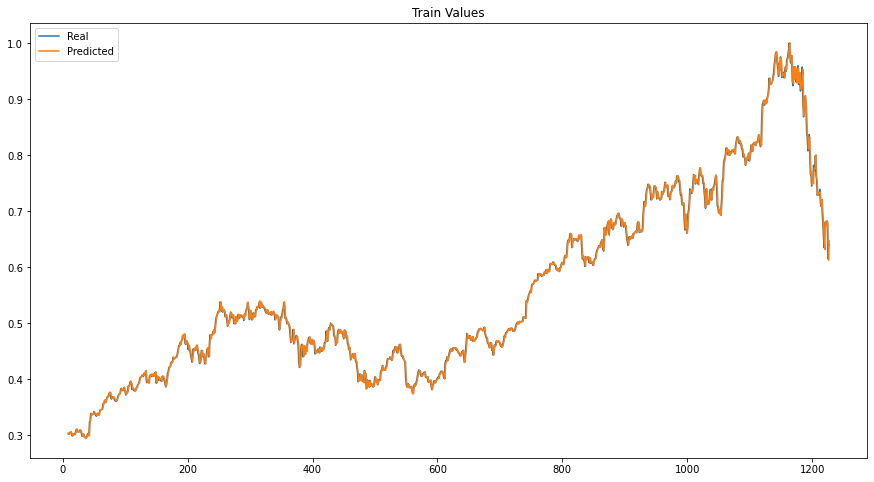
А для її залишків моделі АКФ та ПАКФ



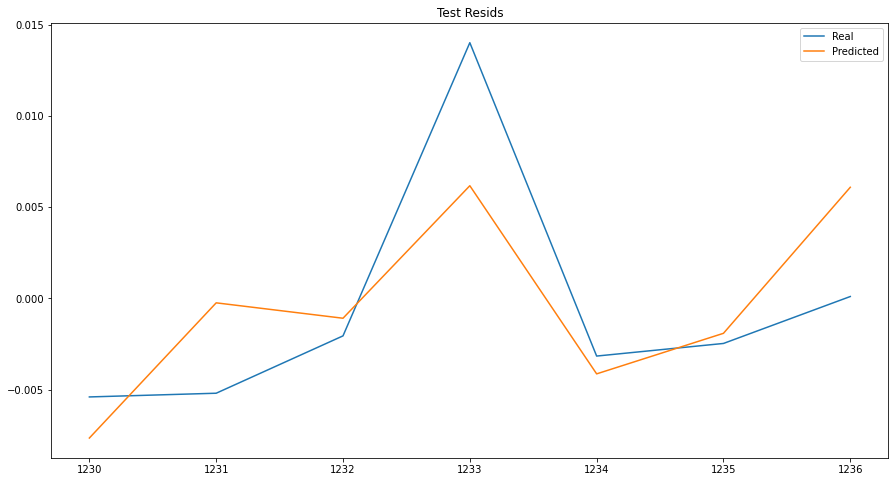
Як бачимо, 7й порядок ковзного середнього є оптимальним

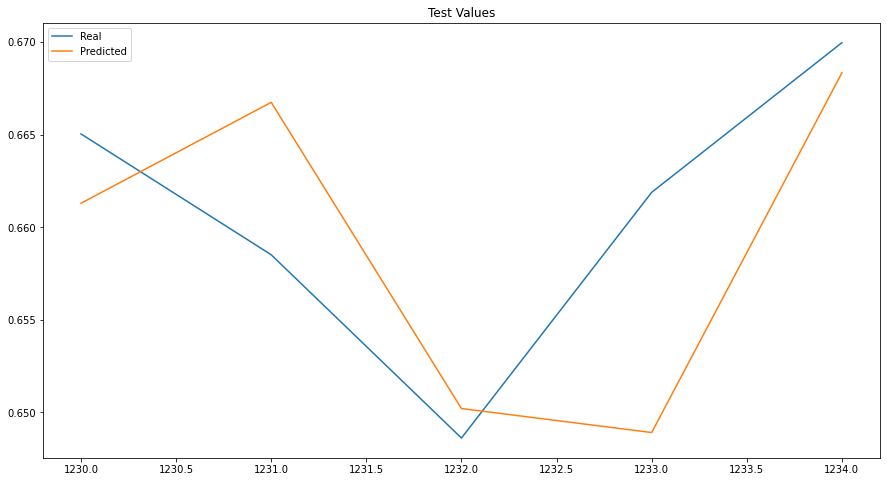
**Модель** **ARMA**

Власне побудова моделі ARMA відбувалась за допомогою кросс-валідації на 7 розбиттях. Під час кросс-валідації моделі валідувались за такими метриками: R2, MAE, MSE, MAPE, RSS, RMSE, а також статистики Дарбіна-Уотсона та інформаційних критеріїв Акайке та Баєса.

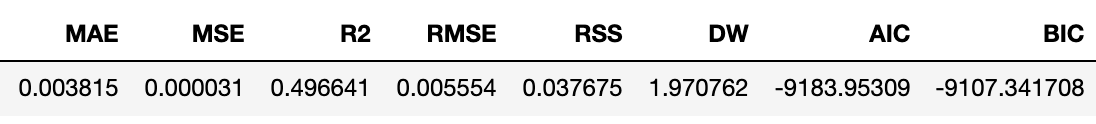


Валідацію будемо проводити шляхом повернення ряду різниці до вихідного

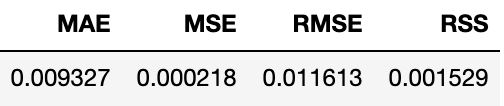




Отже за результами маємо наступні значення метрик на тренувальній вибірці

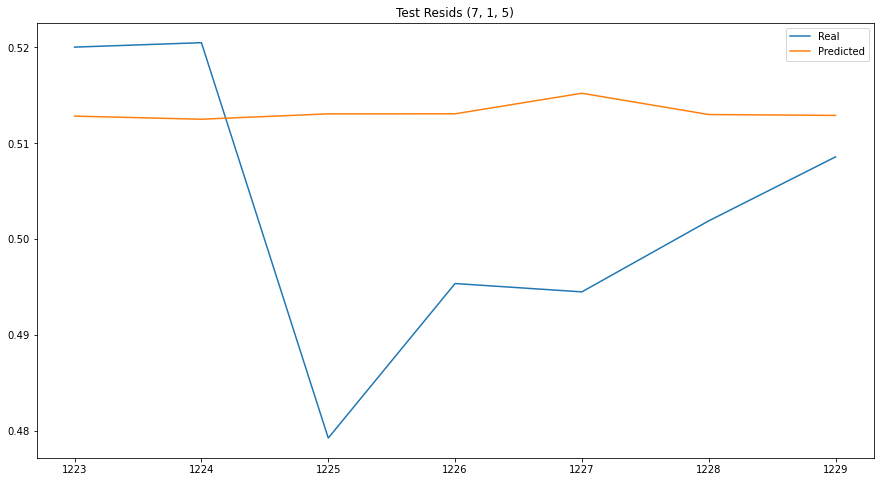


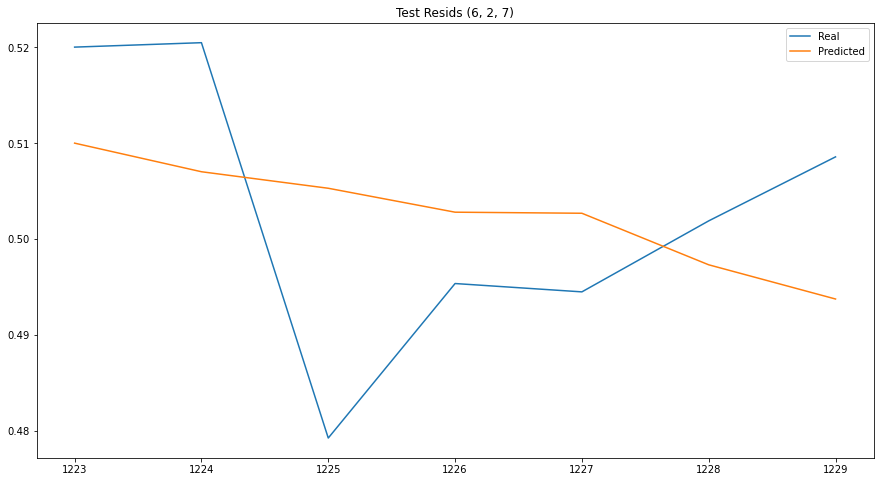
Значення метрик на валідаційній вибірці усереднимо з усіх розбиттів



**Модель ARIMA**

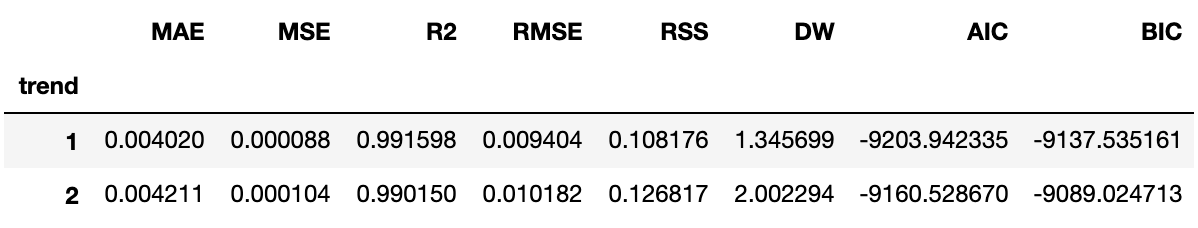
Для моделі ARIMA перевіримо також тренд першого порядку, для якого найкращими параметрами є АР = 7 і КС = 5 та порівняємо результати з обраною моделлю з трендом другого порядку та АР = 6 і КС = 7



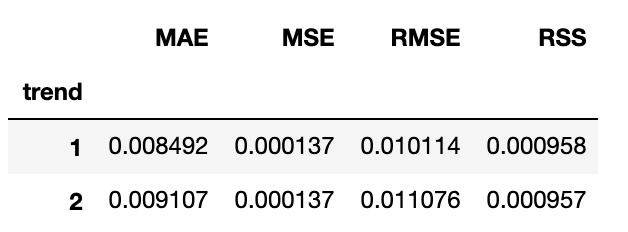


Порівняємо дані моделі на ключових метриках

Тренувальна вибірка



Валідаційна вибірка



Як бачимо обидві моделі показали непогані результати, але в моделі з трендом першого порядку більш виражена автокореляція залишків, тому надалі обираємо модель з другим порядком тренду – ARIMA (6,2,7)

**Гетероскедастичні моделі**

Проведемо тест Бройша-Пагана на гетероскедастичність

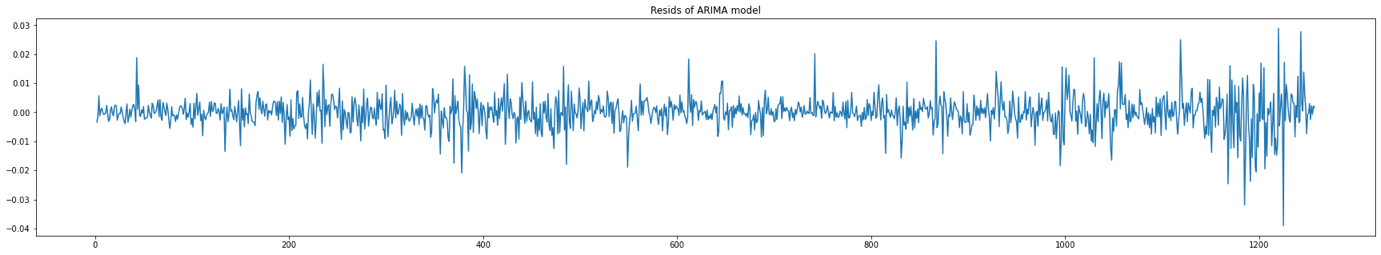
Breusch-Pagan Lagrange Multiplier test for heteroscedasticity

LM Statistic 0.241749

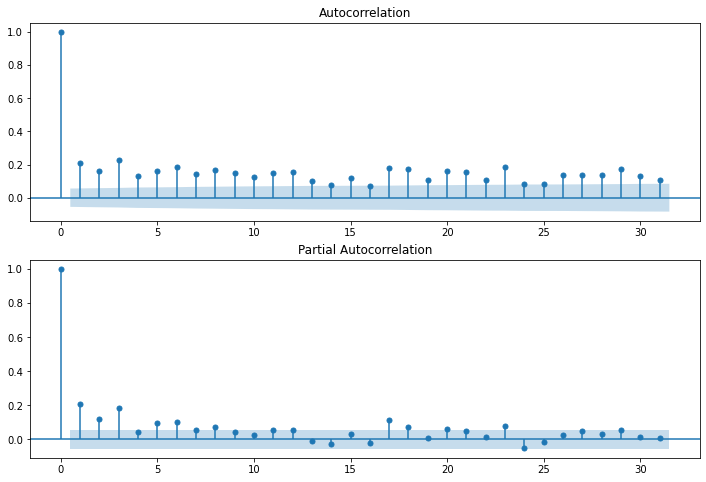
F-Statistic 0.241603

F-Test p-value 0.623137

Отже можемо стверджувати про наявність гетероскедастичності в залишках моделі ARIMA (6,2,7)



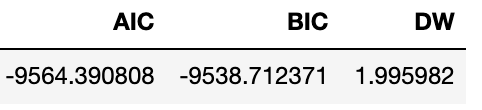
Побудуємо АКФ і ПАКФ для залишків цієї моделі



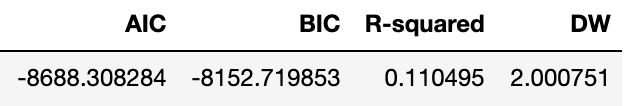
Взявши рівень значущості за 0.2 будемо будувати гетероскедатичні моделі

Приведемо результати пошуку найкращих параметрів для окремих класів гетероскедастичних моделей

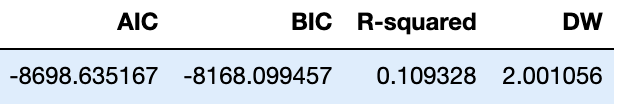
**Модель ARCH (3)**

****

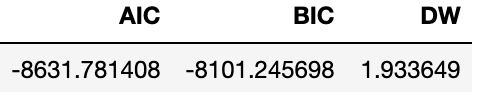
**GARCH (3,1)**

****

**FIGARCH (1,1)**

****

**EGARCH (2, 1)**

****

Також побудуємо прогноз дисперсії ціни на 1 крок за цими моделями

Real value 0.000003

ARCH predict 1.8251373618685284e-05

GARCH predict 2.6009998665357653e-05

FIGARCH predict 3.838774772749344e-05

EGARCH predict 3.57105494926086e-05

**Висновки**

Під виконання практичного завдання було застосовано кілька моделей аналізу часових рядів для дослідження заданого часового ряду, а саме цін акцій компанії Apple проміжок часу у 5 років за допомогою різних моделей, виконано пошук оптимальних параметрів моделей а також побудовано прогноз дисперсії ціни для найкращої моделі за допомогою гетероскедастичних моделей