

Описание

В этом задании вам предстоит провести анализ некоторого временного ряда и попробовать предсказать значения для последующих месяцев. В рамках данного задания мы:

- научимся правильно считывать данные и их визуализировать;
- поймём, как определять свойства временных рядов и познакомимся с различными моделями для предсказания значений;
- оценим качество полученных моделей.

Цель **задания 2** в том, чтобы продолжить ваше изучение языка Python, попутно познакомиться с такими библиотеками, как Pandas (работа с данными), statsmodels (работа с различными статистическими моделями, в том числе и временными рядами).

Что вам предстоит сделать:

1. Считать данные из Данные.xlsx. Ответы на тестовой выборке Ответы.xlsx не следует использовать ни в каких экспериментах, кроме финального. Проверить, является ли ряд стационарным в широком смысле. Это можно сделать двумя способами:
 - i. Провести визуальную оценку, отрисовав ряд и скользящую статистику (среднее, стандартное отклонение). Постройте график, на котором будет отображен сам ряд и различные скользящие статистики.
 - ii. Провести тест Дики-Фуллера.

Сделать выводы из полученных результатов. Оценить достоверность статистики. (25 баллов)

2. Разложить временной ряд на тренд, сезонность, остаток в соответствии с аддитивной, мультипликативной моделями. Визуализировать их, оценить стационарность получившихся рядов, сделать выводы. (15 баллов)
3. Проверить является ли временной ряд интегрированным порядка k . Если является, применить к нему модель ARIMA, подобрав необходимые параметры с помощью функции автокорреляции и функции частичной автокорреляции. Выбор параметров обосновать. Отобрать несколько моделей. Предсказать значения для тестовой выборки. Визуализировать их, посчитать r^2 score для каждой из моделей. Произвести отбор наилучшей модели с помощью информационного критерия Акаике. Провести анализ получившихся результатов. (50 баллов)

За все правильно выполненные пункты можно получить 90 баллов.

+10 баллов – соблюдение PEP8

+10 баллов – использование для визуализации библиотек bokeh или seaborn. Надо сделать, чтобы было красиво...).

Сдача

Задание **должно** выполняться индивидуально. **Сдача до 23:59 18 декабря 2022г.** **Задания, отправленные позднее проверяться не будут.** Задания оформлять в виде архива, содержащего рабочий код и другие ресурсы, решающие поставленную задачу, и файл readme с ФИО автора работы, инструкцией по запуску, включая необходимое дополнительное ПО и кратким описанием подхода к решению. Данный архив отправить на почту polonskaya.d@mail.ru и проверить, что отмечено его получение в гугл-лабличке.

Литература и инструменты

[Pandas](#) — библиотека для работы со структурированными данными.

[Statsmodels](#) — библиотека для построения различных статистических моделей.