**Пример для data science.**

Зависимость между случайными величинами – факт интересный и с теоретической, и с практической точек зрения.

Как правило, во всех учебниках приводится пример линейной зависимости, наличие которой обнаруживается с помощью вычисления линейной корреляции, а затем, скажем, методом наименьших квадратов находится и её явный вид.

Но ведь зависимость может быть достаточно простой, но нелинейной.

Поиск такой зависимости более сложная задача, не для нескольких страничек.

Но, чтобы привлечь внимание к этой теме, привожу простой пример, причём в сочетании теории и практики.

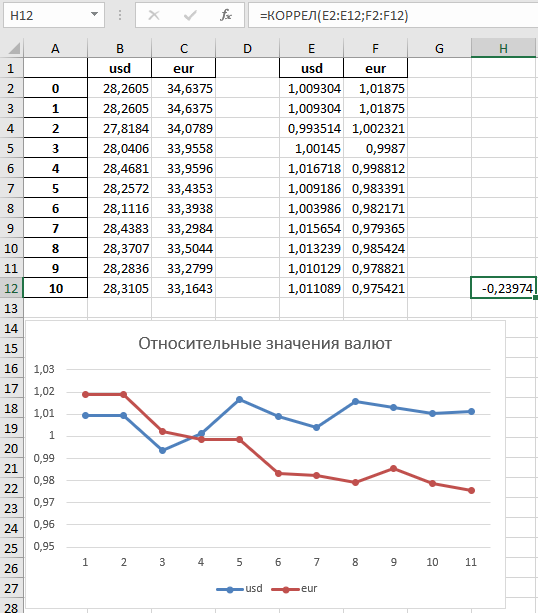
Изображение выглядит как текст

Автоматически созданное описаниеИз базы с курсами валют извлекаю курсы американского доллара и евро.

Изображение выглядит как текст

Автоматически созданное описаниеЧтобы не загромождать пример большими наборами данных, курсы беру раз в 10 дней.

Эта программа создаёт файл curr.xslx с курсами, которые нормирую делением на целую часть курса, а затем изображаю данные в виде графиков. Кроме того, нахожу линейную корреляцию этих данных, которые далее называю относительными значениями валют.

Имею

Никакой корреляции нет, значит, и линейной зависимости – тоже.

Но замечаю, что возрастание одной валюты, как правило, но не строго, соответствует возрастанию и другой.

Изображение выглядит как стол

Автоматически созданное описаниеПоэтому рассматриваю массив разностей относительных значений валют и нахожу корреляцию уже новых данных.

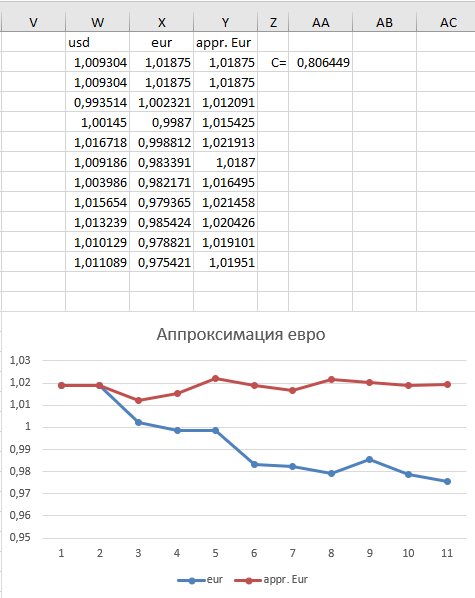
Здесь некоторая корреляция уже имеется.

Поскольку пример модельный, я не пытаюсь улучшить зависимость, а нахожу её.

Изображение выглядит как стол

Автоматически созданное описаниеАппроксимация вполне неплохая, а это значит, что если функции доллара и евро обозначить , соответственно, то

где = 0,00945.

Сравниваю настоящее значение евро и полученное из найденной зависимости.

*Максимальная ошибка аппроксимации – 0,04 от стоимости валюты. Это вполне неплохо.*

***Результат не имеет статистической достоверности, так как на 10 значениях статистика не строится. Напоминаю, это только модельный пример.***

Все вычисления, сделанные в Excel, разумеется, можно реализовать в Piton.