***Задача№4***

Задача о нуклеотидной последовательности

Вся наследственная информация в живых организмах хранится внутри молекулы ДНК. Эта молекула состоит из последовательности четырех "букв" — ﻿A, T, G и C.

Напишите функцию most\_significant, которая получает на вход dataframe с произвольным количеством переменных, где каждая переменная это нуклеотидная последовательность.

Функция должна возвращать вектор с ﻿названием переменной (**или переменных**), в которой был получен минимальный p - уровень значимости при проверке гипотезы о равномерном распределении нуклеотидов при помощи критерия хи - квадрат.

**Пример** для прояснения сути задачи:

V1 V2 V3

1 A A C

2 G G A

3 C C C

4 T T A

5 G T T

6 T A G

В этом примере три последовательности V1 , V2, V3.

Для каждой переменной мы можем проверить нулевую гипотезу о том, что все нуклеотиды (A, T, G, C) встречаются равновероятно внутри этой последовательности. Однако, возможно, что в некоторых последовательностях распределение частоты встречаемости каждого нуклеотида отличается от равномерного.

Функция должна возвращать вектор с ﻿названием переменной (**или переменных**), в которой был получен минимальный p - уровень значимости при проверке гипотезы о равномерном распределении нуклеотидов при помощи критерия хи - квадрат.

