

Домашнее задание

Тема: Работа с данными Pandas. Dataset по продажам.

Цель задания: Дать возможность закрепить теоретическую часть материала. Рекомендуется использовать помимо лекции сторонние источники информации, как это рекомендует спикер. Очень важно пробовать применять на практике данную информацию, поэтому рекомендуем повторять все действия за спикером максимально внимательно.

Задачи:

1 Перечислите (можно не на память, просто их нужно держать перед глазами и уметь ими пользоваться) статистические характеристики. Повторите за спикером примеры их использования.

Процесс выполнения:

1

Основные статистические характеристики

`amin(a[, axis, out, keepdims, initial])`

Минимальное значение элементов массива. Параметр `axis` позволяет указывать оси, вдоль которых необходим поиск минимальных значений.

`amax(a[, axis, out, keepdims, initial])`

Максимальное значение в массиве. Параметр `axis` позволяет указывать оси, вдоль которых необходим поиск максимальных значений.

`nanmin(a[, axis, out, keepdims])`

Минимальное значение массива или минимальное значение вдоль указанной оси. Элементы с значением `np.nan` игнорируются.

`nanmax(a[, axis, out, keepdims])`

Максимальное значение массива или максимальное значение вдоль указанной оси. Элементы с значением `np.nan` игнорируются.

`numpy.ptp(a[, axis, out, keepdims])`

Возвращает диапазон значений (`[max - min]`) массива или указанной оси массива.

`numpy.percentile(a, q[, axis, out, overwrite_input, interpolation, keepdims])`

Вычисляет `q`-й перцентиль (перцентиль) значений элементов массива или элементов вдоль указанной оси.

`numpy.nanpercentile(a, q[, axis, out, overwrite_input, interpolation, keepdims])`

Вычисление `q`-го перцентиль (перцентиль) значений вдоль указанной оси массива. Элементы с значением `np.nan` игнорируются.

`numpy.quantile(a, q[, axis, out, overwrite_input, interpolation, keepdims])`

Вычисление `q`-го перцентиль (перцентиль) значений вдоль указанной оси массива. Элементы с значением `np.nan` игнорируются.

`numpy.nanquantile(a, q[, axis, out, overwrite_input, interpolation, keepdims])`

Вычисление q-го процентиля (перцентиля) значений вдоль указанной оси массива. Элементы с значением `np.nan` игнорируются.

Средние и отклонения

`median(a[, axis, out, overwrite_input, keepdims])`

Медиана значений элементов массива, расположенных вдоль указанной оси.

`average(a[, axis, weights, returned])`

Средневзвешенное значений элементов массива расположенных вдоль указанной оси.

`mean(a[, axis, dtype, out, keepdims])`

Среднее арифметическое значений элементов массива расположенных вдоль указанной оси.

`std(a[, axis, dtype, out, ddof, keepdims])`

Стандартное отклонение значений элементов массива расположенных вдоль указанной оси.

`var(a[, axis, dtype, out, ddof, keepdims])`

Дисперсия значений элементов массива расположенных вдоль указанной оси.

`nanmedian(a[, axis, out, overwrite_input, ...])`

Медиана значений элементов массива расположенных вдоль указанной оси. Элементы с значением `NaN` игнорируются.

`nanmean(a[, axis, dtype, out, keepdims])`

Среднее арифметическое значений элементов массива расположенных вдоль указанной оси. Элементы с значением `NaN` игнорируются.

`nanstd(a[, axis, dtype, out, ddof, keepdims])`

Стандартное отклонение значений элементов массива расположенных вдоль указанной оси. Элементы с значением `NaN` игнорируются.

`nanvar(a[, axis, dtype, out, ddof, keepdims])`

Дисперсия значений элементов массива расположенных вдоль указанной оси. Элементы с значением `NaN` игнорируются.

Корреляции

`corrcoef(x[, y, rowvar, bias, ddof])`

Коэффициент корреляции Пирсона.

`correlate(a, v[, mode])`

Взаимнокорреляционная функция двух одномерных последовательностей.

`cov(m[, y, rowvar, bias, ddof, fweights, ...])`

Ковариационная матрица.

Гистограммы

`histogram(a[, bins, range, normed, weights, ...])`

Вычисление гистограммы набора данных.

`histogram2d(x, y[, bins, range, normed, weights])`

Вычисление двумерной гистограммы двух наборов данных.

`histogramdd(sample[, bins, range, normed, ...])`

Вычисление N-мерной гистограммы N-го количества наборов данных.

`bincount(x[, weights, minlength])`

Количество вхождений значений в массиве.

`numpy.histogram_bin_edges(a, bins=10, range=None, weights=None)`

Значения для прямоугольников гистограммы.

`digitize(x, bins[, right])`

Вычисление индексов числовых интервалов массива `bins` в которые входит каждое последующее значение элемента массива `x`.

Критерии оценки:

Задание считается выполненным при предоставлении студентом списка статистических характеристик и основных скринов экрана с выполненными действиями.