

# Анализ временных рядов

## Распознавание дат и формирование новых индексов

```
In data = pd.read_csv('filename.csv',  
                      index_col=[0], # список номеров столбцов или названий столбцов  
                      parse_dates=[0]) # список номеров столбцов или названий столбцов
```

## Проверка, что индекс монотонный

```
In print(data.index.is_monotonic)
```

## Ресемплинг

```
In # ресемплинг – среднее по каждому часу  
data.resample('1H').mean()  
  
# ресемплинг – максимум по каждому двум неделям  
data.resample('2W').max()
```

## Скользящее среднее с окном размером 7

```
In data.rolling(7).mean()
```

## Разложение временного ряда на тренд, сезонность и остаток

```
In from statsmodels.tsa.seasonal import seasonal_decompose  
  
decomposed = seasonal_decompose(data)  
  
decomposed.trend      # тренд  
decomposed.seasonal   # сезонная составляющая  
decomposed.resid      # остаток декомпозиции
```

## Сдвиг на один шаг с заполнением нулевого значения

```
In print(data.shift(fill_value=0))
```

# Словарь

## **Временные ряды** (англ. *time series*)

это последовательности чисел на оси времени. Интервал между значениями ряда постоянный

## **Ресемплирование** или **ресемплинг** (англ. *resample*)

это изменение интервала со значениями ряда. Его выполняют в два этапа:

1. Выбирают новую длину интервала. Причём значения из текущего интервала группируются
2. В каждой группе вычисляется агрегированное значение ряда. Это может быть медиана, среднее, максимум или минимум

## **Скользящее среднее** (англ. *rolling mean*)

или **движущееся среднее** (англ. *moving average*)

метод сглаживания временных рядов. Его суть заключается в поиске значений, которые меньше всего подвержены колебаниям, то есть средних арифметических

## **Тренд** (англ. *trend*)

плавное изменение среднего значения ряда без повторяющихся закономерностей

## **Сезонность** (англ. *seasonality*)

циклично повторяющиеся закономерности во временном ряду

## **Стохастический процессом** (англ. *stochastic process*)

это случайная величина, у которой со временем меняется её распределение

Стохастический процесс **стационарный** (англ. *stationary stochastic process*), если его распределение со временем не меняется. Если распределение меняется, то процесс называется **нестационарным** (англ. *nonstationary stochastic process*)

## **Разности временного ряда** (англ. *time series difference*)

это набор разностей между соседними элементами временного ряда, т. е. из каждого значения вычитается предыдущее