

# Временные ряды 2

...

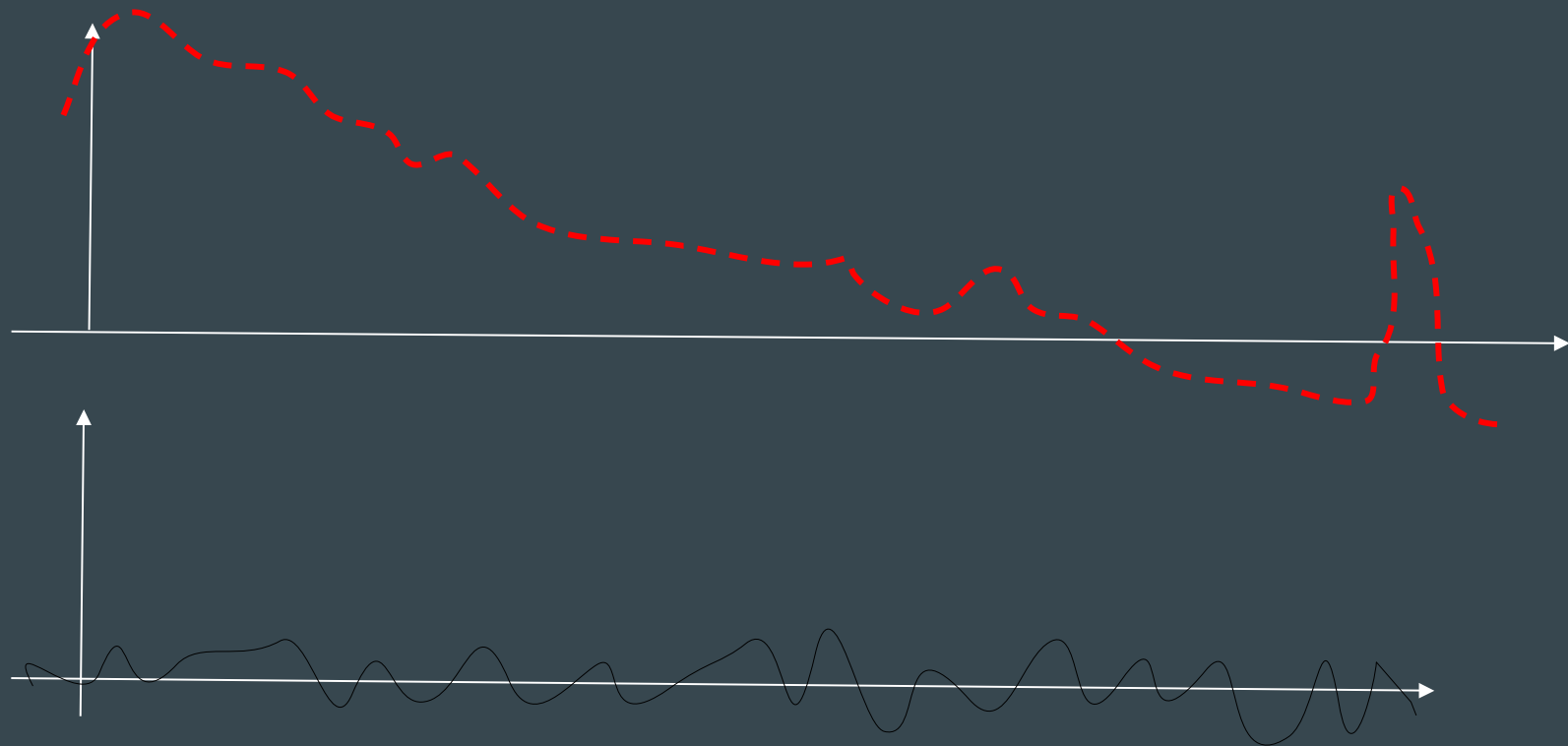
Построение признаков

2019

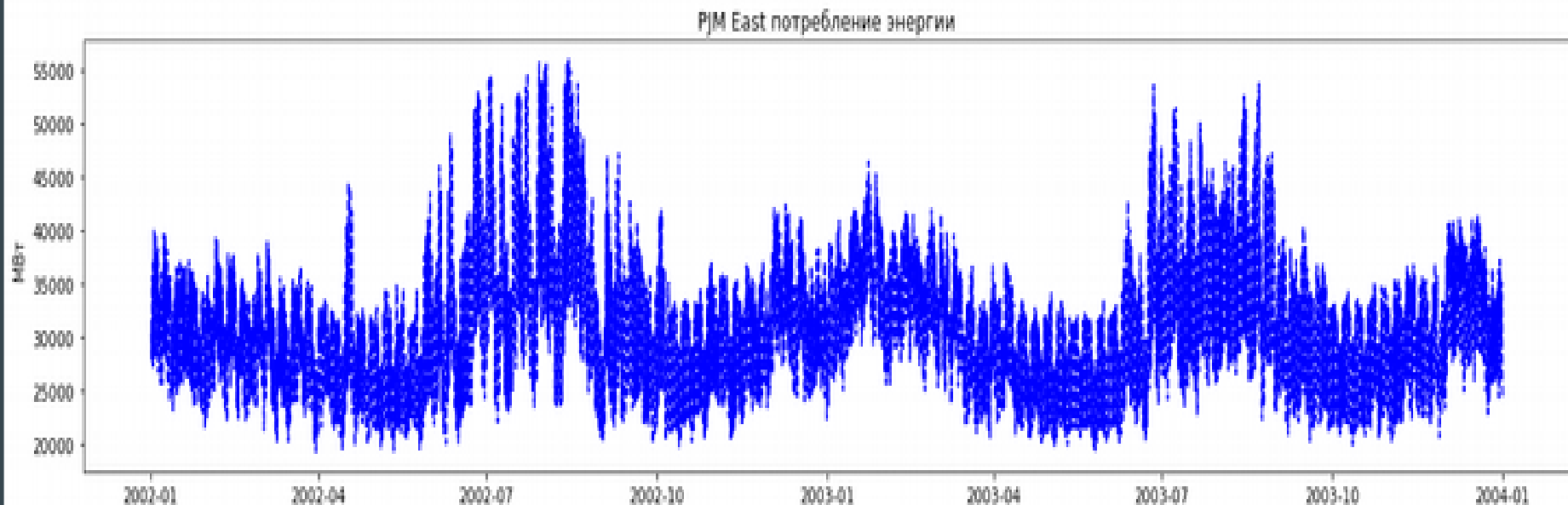
# План

1. ДЗ1
2. Теоретическая часть
  - а. Понятие стационарного ряда
  - б. Тренд , цикл , Сезон
  - с. Интегральный ряд
  - д. Модели временного ряда
  - е. Сглаживание
  - ф. тест ADF
3. Практическая часть - построение признаков от значений ряда
  - а. как запустить анализ стационарности
  - б. как запустить сглаживание
  - с. как собрать модель
  - д. ДЗ 2

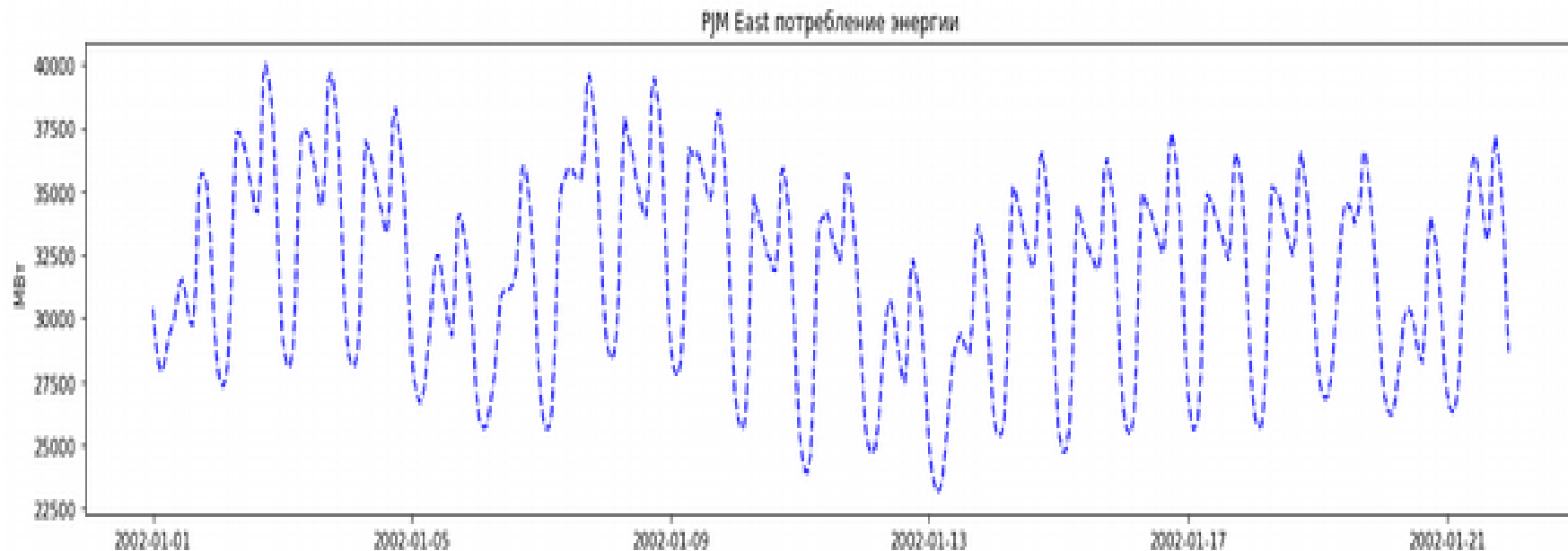
# Стационарность



# Потребление энергии



# Потребление энергии



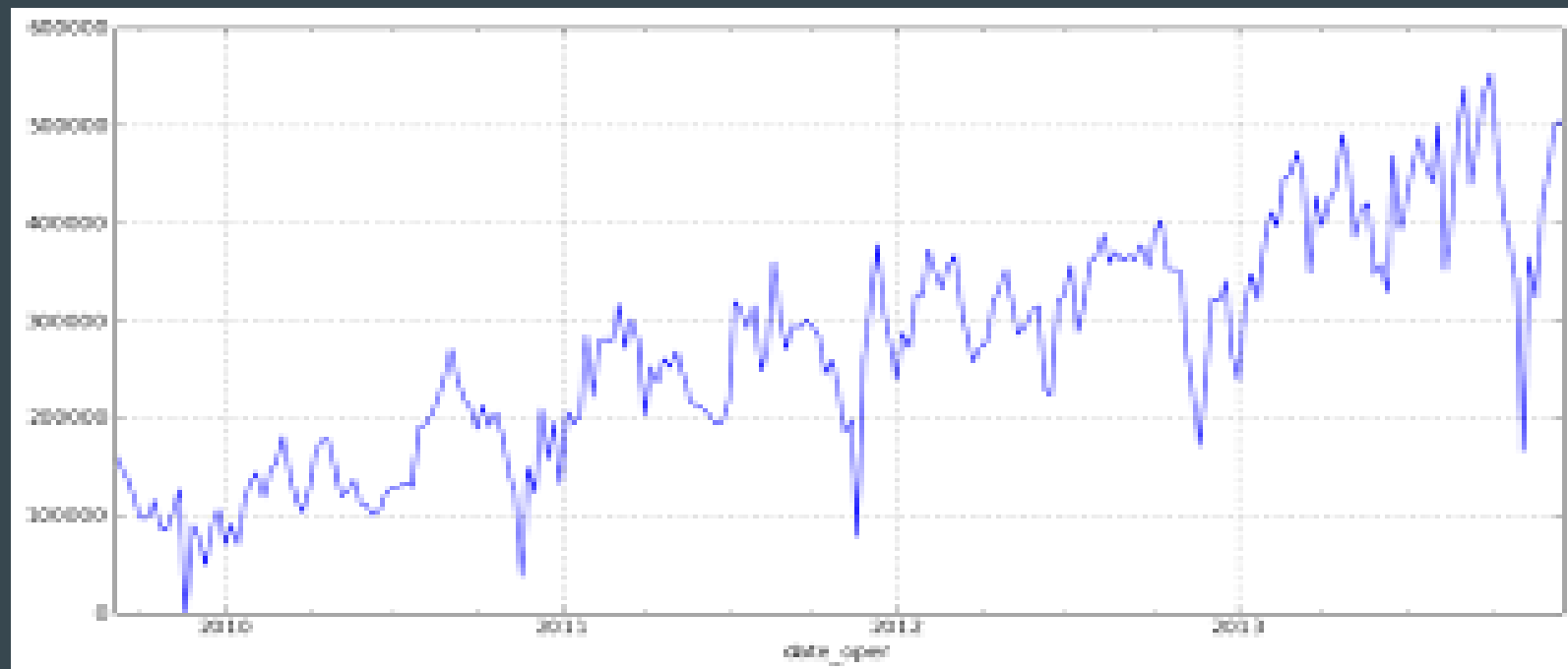
# Модель ряда

Временной Ряд = Явление от времени + случайность

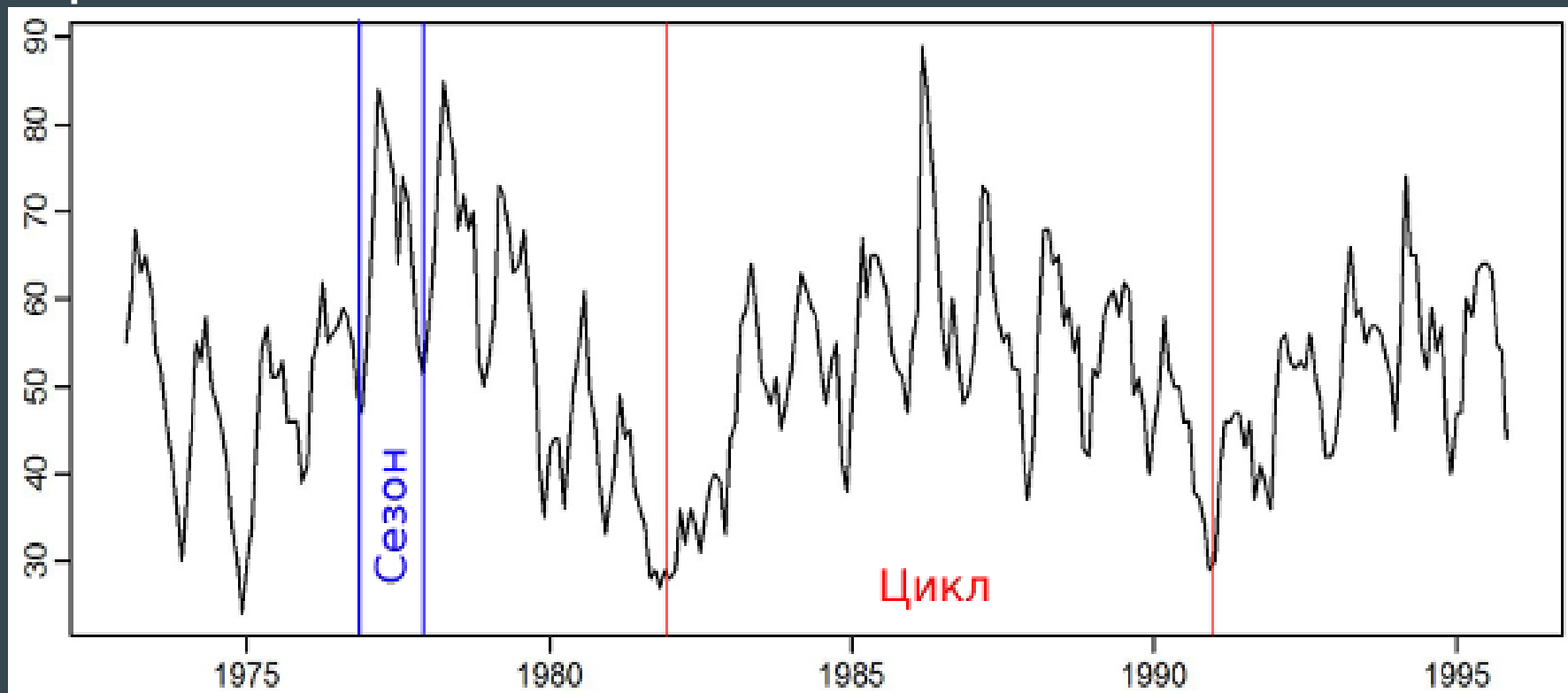
Временные явления ряда:

- долговременные;
- сезонные;
- циклические;
- Случайные.

# Тренд

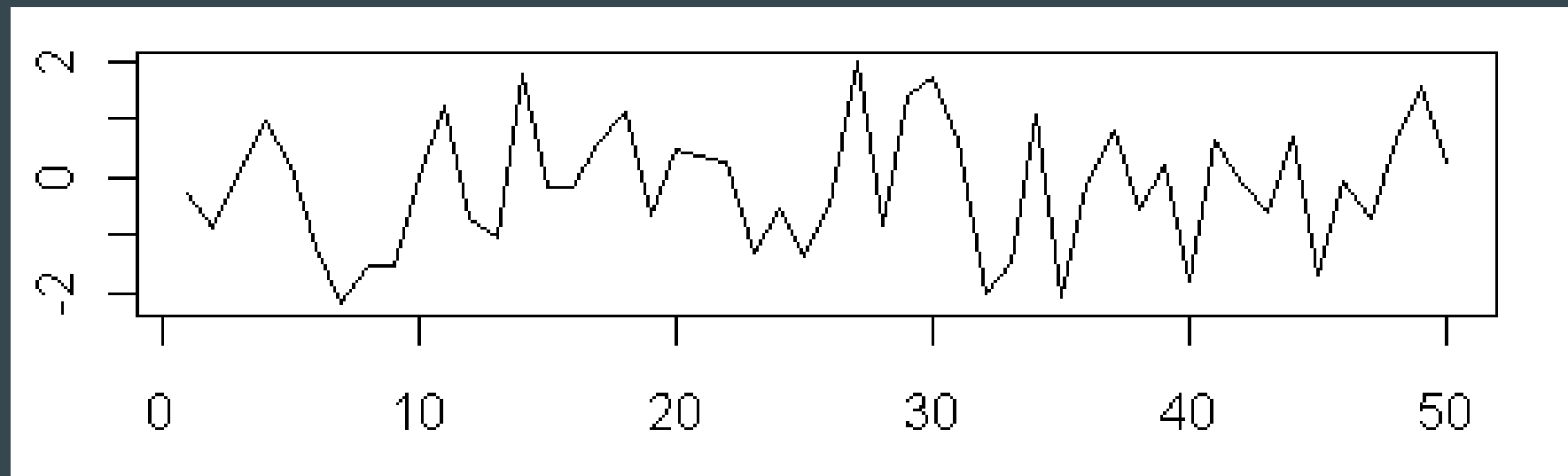


# Цикл

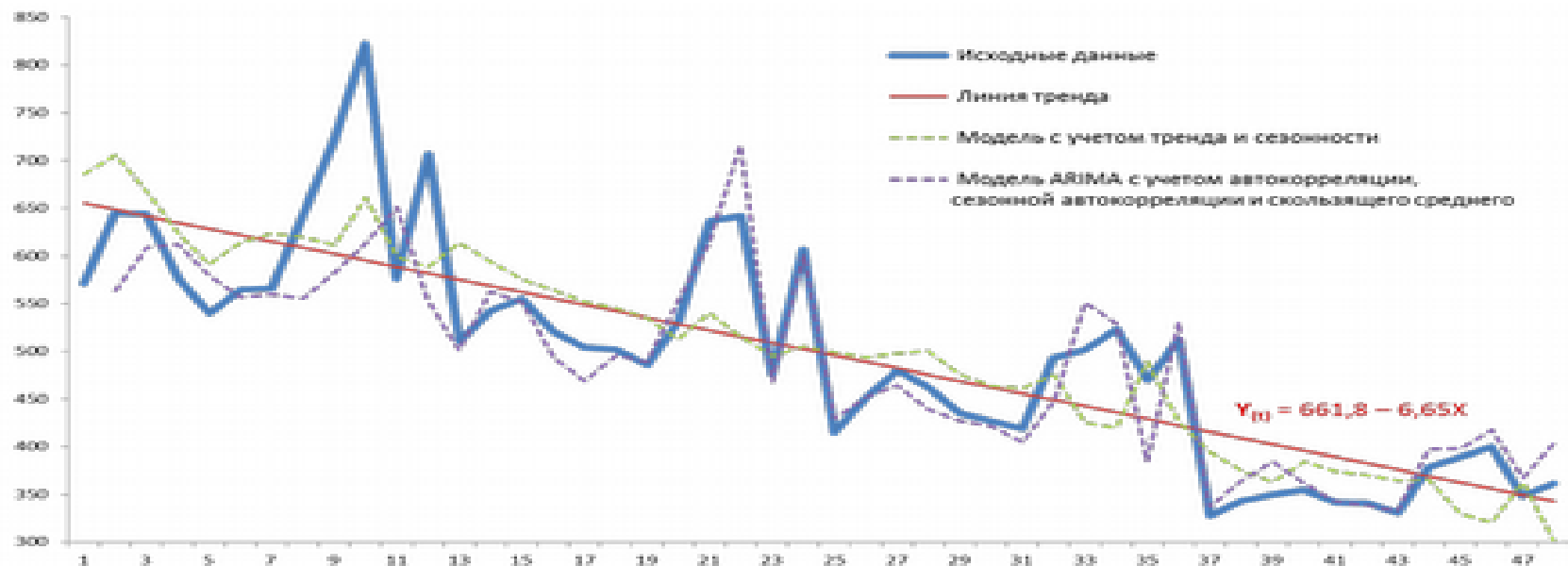




шум



# ВСЕ ВМЕСТЕ



# Модель ряда

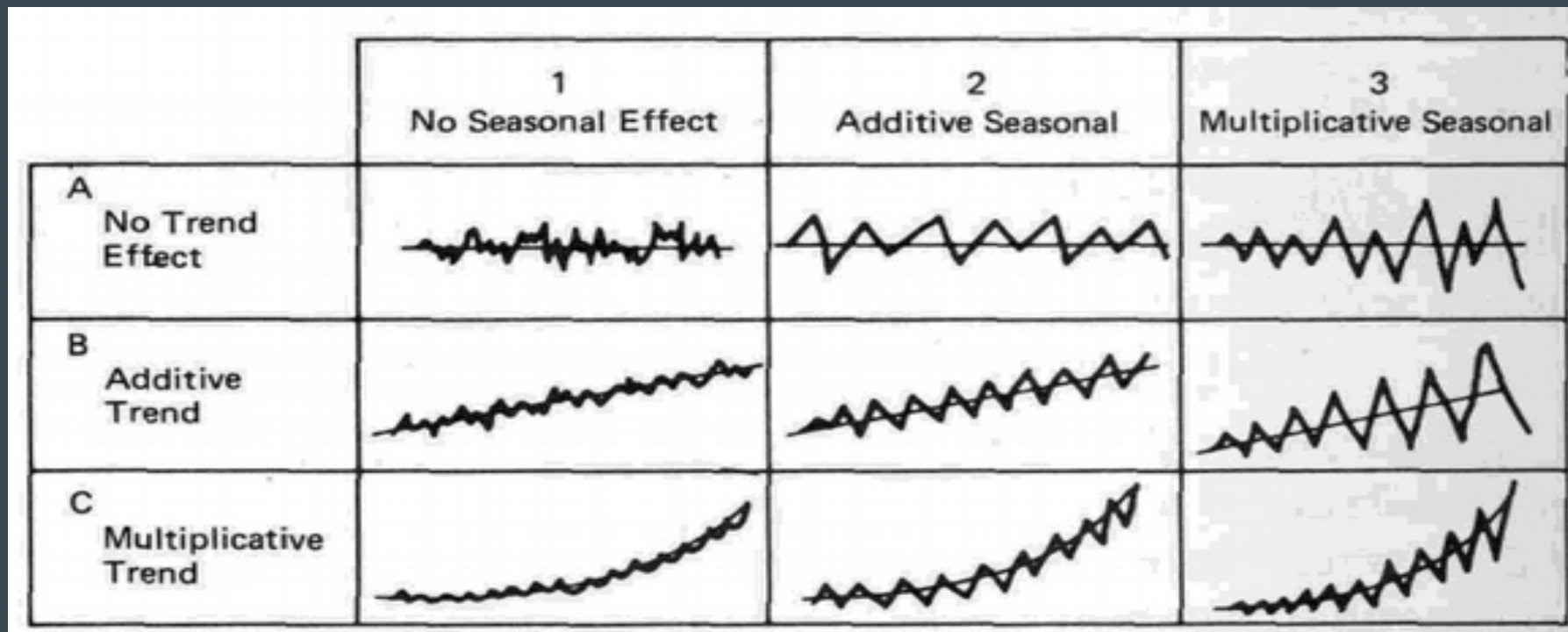
- аддитивная :

$$y(t) = X_{\text{trend}}(t) + X_{\text{cycle}}(t) + e(t)$$

- мультипликативная :

$$y(t) = X_{\text{trend}}(t) X_{\text{cycle}}(t) e(t)$$

# Виды рядов



# Стационарность

Важно:

- Интервал анализа

# Стационарность

Постоянство :

- Мат.ожидания
- Дисперсии
- автоковариация стационарного ряда с лагом  $L$
- коэффициент автокорреляция стационарного ряда с лагом  $L$

# Как отделить случайное

Вычесть не случайное

- Найти сглаженную модель  $y\_случ(i) = x(i) - f(x(i-p:i))$ .
  - размер окна  $p$ ?
  - вид функции  $f()$ ?

Модель авторегрессии  $y(t) = c + y(t-1) * a_1 + y(t-2) * a_2 + \dots + y(t-p) * a_p + e$

- Взять разницы  $y\_случ?(i) = x(i) - x(i-1)$ 
  - сколько раз брать?
  - как узнать, что осталась случайная часть?

Модель скользящего среднего  $y(t) = c + e(t) - e(t-1) * b_1 + e(t-2) * b_2 + \dots + e(t-p) * b_p$

# На работу модели временной ряд влияет:

- как вели себя предшествующие куски модели
- какие ошибки они делали

Можем сделать признаки:

- сглаживания
- разности



# Признаки для формирования модели

- Прошлые значения ряда
- Преобразования по сглаживанию для ряда
- Преобразования - разности

# Признаки из истории

$$x\_NEW(i) = F(x(i-n), x(n-i+1), \dots, x(i))$$

n - порядок

# Признаки из Сглаживания

Скользящие средние

$$x\_NEW(i) = (x(i-n) + x(n-i+1) + \dots + x(i)) * 1/(n + 1).$$

n - порядок

# Признаки из Сглаживания

Экспоненциальное

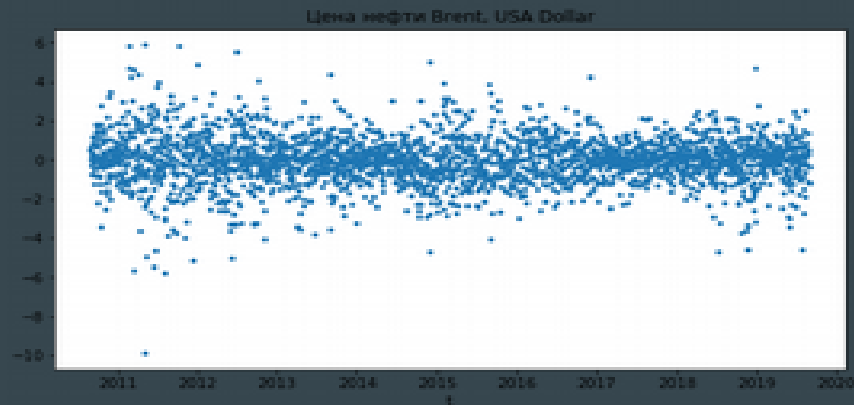
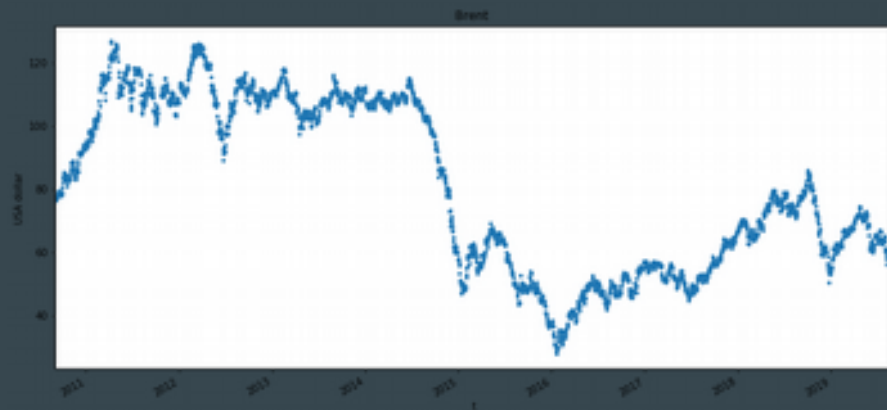
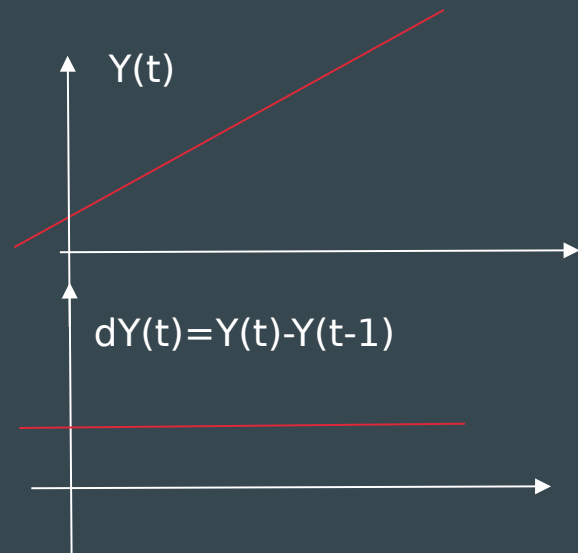
$$x\_NEW(i) = (x(i)*a + (1-a)*(x(n-i+1)* a + (1-a)*( ...+ (1-a)*(x(i-n))))$$

$$x\_NEW(i) = (x\_NEW(i-1)*(1-a) + a*x(i))$$

$a$  - порядок

# Признаки из Первые разности

$$dY(t) = Y(t) - Y(t-1)$$



# Тест Стационарности

- Посмотреть !!!
- Сравнить статистику для разных отрезков
- Тест Дики-Фулера

# Тест Дики-Фуллера

авторегрессия  $Y(t) = b * Y(t-1) + e_t$

авторегрессия для 1-х разностей  $dY(t) = p * Y(t-1) + e_t$

$d$  - оператор разности  $dY(t) = Y(t) - Y(t-1)$

$Y(t) = a_1 Y(t-1) + a_2 Y(t-2) + e_t$

$dY(t) = Y(t) - Y(t-1)$

$dY(t) = (a_1 + a_2 - 1) Y(t-1) - a_2 dY(t-1) + e_t$

проверяем  $a_1 + a_2 - 1 = 0$

# Тест Дики-Фуллера

основная гипотеза  $H_0: b=1$

авторегрессия  $Y(t) = b * Y(t-1) + \epsilon_t$

на следующем этапе:  $H_0: p=0$

авторегрессия для разностей  $dY(t) = p * Y(t-1) + \epsilon_t$  ,  $p=b-1$

$d$  - оператор разности  $dY(t) = Y(t)-Y(t-1)$

$Y(t) = a_1 Y(t-1) + a_2 Y(t-2) + \epsilon_t$

$dY(t) = (a_1 + a_2 - 1) Y(t-1) - a_2 dY(t-1) + \epsilon_t$

проверяем  $a_1 + a_2 - 1 = 0$

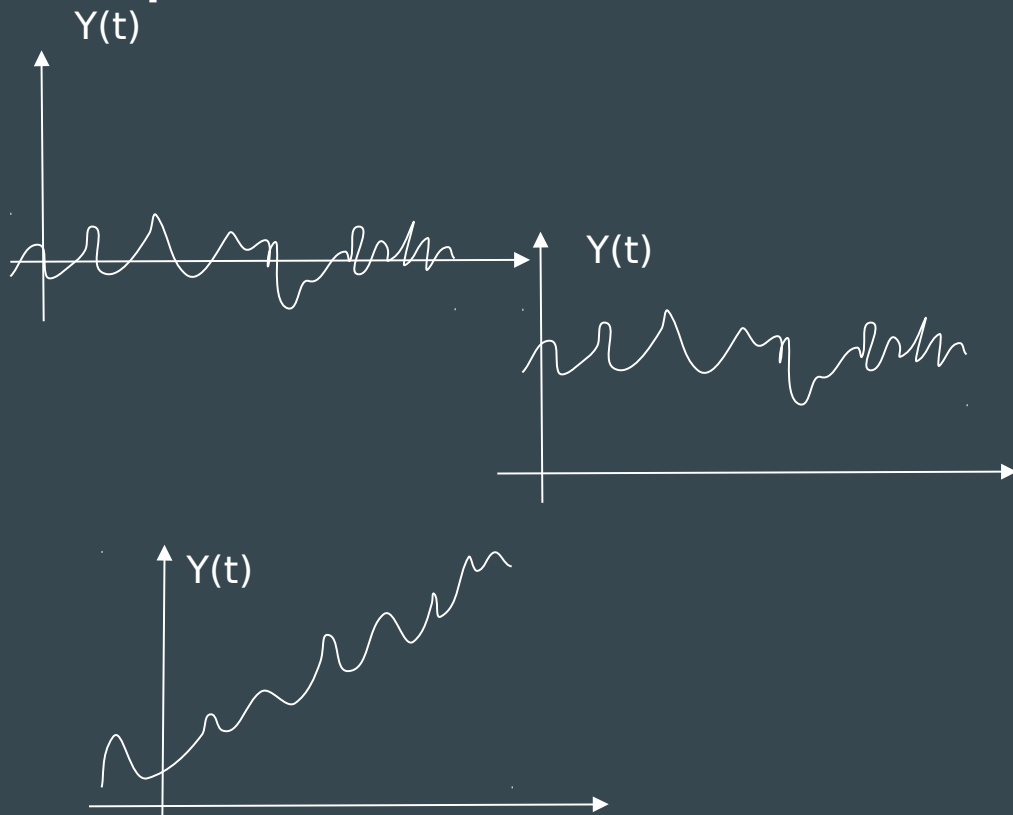


# Тест Дики-Фуллера

$$dY(t) = pY(t-1) + et$$

$$dY(t) = a + pY(t-1) + et$$

$$dY(t) = a + c t + pY(t-1) + et;$$



# Тест Дики-Фуллера

Односторонний

Порядок тестирования:

- ➔ вторые разности ряда :

- ➔ если гипотеза отвергается, то идем дальше

- ➔ не более чем 2 раза интегрируемый,

- ➔ первые разности ряда :

- ➔ если гипотеза отвергается, то идем дальше

- ➔ не более чем 1 раз интегрируемый,

- ➔ исходный ряд :

- ➔ если гипотеза отвергается, то стационарный