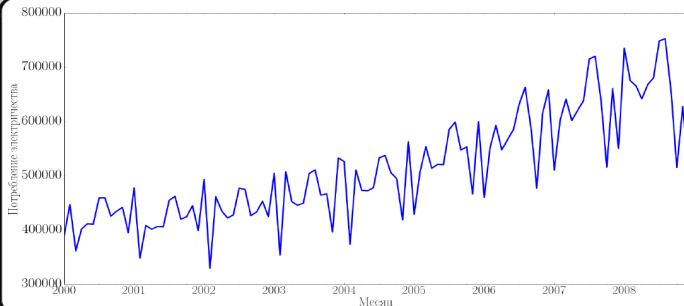


# РЕГРЕССИОННЫЙ ПОДХОД К ПРОГНОЗИРОВАНИЮ

---

# ПРАЗДНИКИ

- Падения соответствуют месяцам, на которые выпадают праздники по исламскому календарю
- (год примерно на 11 дней короче, чем в григорианском)

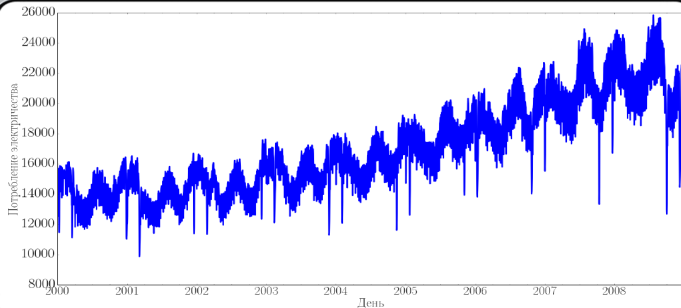


$$y_t = \sum_{j=1}^k \beta_j x_{jt} + z_t$$

$$\begin{aligned} z_t = & \alpha + \phi_1 z_{t-1} + \dots + \phi_p z_{t-p} + \\ & + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} + \\ & + \phi_S z_{t-S} + \dots + \phi_{PS} z_{t-PS} + \\ & + \theta_S \varepsilon_{t-S} + \dots + \theta_{PS} \varepsilon_{t-PS} + \varepsilon_t \end{aligned}$$

# СЛОЖНАЯ СЕЗОННОСТЬ

- › Недельная сезонность
- › Годовая сезонность
- › Праздники по исламскому календарю



## » Сложности:

- ▶ при длинных периодах сезонности в модели SARIMA получается слишком много параметров
- ▶ поведение дневного ряда вряд ли определяется его значением ровно **365** дней назад
- ▶ длина года — **365.25** дней и **52.18** недель

» Решение:

брать в качестве  $S$  период самой короткой сезонности, а сезонность более высоких порядков учитывать регрессией на фурье-гармоники с периодами, например,  $365.25$ ,  $365.25/2$ ,  $365.25/3$  и т.д.

- ▶ гармоника по длинным периодам сезонности
- ▶ индикаторы номера периода в коротких сезонностях
- ▶ индикаторы праздников
- ▶ индикаторы пред- и постпраздничных дней
- ▶ тренды (линейный, квадратичный и т.д.)
- ▶ скользящие средние ряда за предыдущие периоды
- » При хорошем подборе признаков регрессии часто оказывается достаточно

- Пример: дневные продажи товаров в магазинах
- Информация: продажи, остатки, цены, скидки, промо-акции, иерархия товаров, иерархия и расположение торговых точек
- Задача: построить прогнозы продаж всех товаров во всех магазинах



- › Задача: построить прогнозы продаж всех товаров во всех магазинах
- › Проблема: ручной подбор прогнозирующих моделей для каждой пары товар-магазин невозможен
- › Решение: регрессионная модель с хорошо подобранными признаками

- › В ARIMA можно учитывать внешние факторы
- › При хорошем подборе внешних факторов специфические модели временных рядов часто оказываются не нужны