Існують два методи прогнозування цін:

* Фундаментальний аналіз — коли оцінюється інформація, яка стосується більше компаніі, ніж її акцій. Рішення стосовно тих чи інших дій приймаються на основі оінки діяльності в минулому.
* Технічний аналіз — розглядається поведінка ціни і як результат виявляються її основні паттерни (використовуючи аналіз часових рядів).

Технічний аналіз використовується для розробки короткострокових стратегій, у той час як фундамендальний допомогає виявити стратегію на роки вперед.

Часові ряди

Часовий ряд (англ. time series) — це ряд точок даних[en], проіндексованих (або перелічених, або відкладених на графіку) в хронологічному порядку. Найчастіше часовий ряд є послідовністю, взятою на рівновіддалених точках в часі, які йдуть одна за одною. Таким чином, він є послідовністю даних дискретного часу[en]. Прикладами часових рядів є висоти океанських припливів, кількості сонячних плям, та щоденне середньозважене значення індексу ПФТС (біржовий індекс, який розраховується щодня за результатами торгів ПФТС\* на основі середньозваженої ціни за угодами) на момент закриття торгів.

*\*Перша Фондова Торгова Система*

Часові ряди дуже часто представляють за допомогою лінійних діаграм. Часові ряди використовуються в статистиці, обробці сигналів, розпізнаванні образів, економетриці, фінансовій математиці, прогнозуванні погоди, розумному транспорті та передбаченні траєкторій, передбаченні землетрусів, електроенцефалографії, автоматичному керуванні, астрономії, технологіях зв'язку, а також значною мірою в будь-якій області прикладної науки та інженерії, яка включає часові вимірювання.

Дані часових рядів мають природний часовий порядок. Це робить аналіз часових рядів відмінним від поперечних досліджень, у яких не існує природного порядку спостережень (наприклад, пояснення заробітної платні людей через посилання на їхні рівні освіти, де дані осіб можуть вводитися у будь-якому порядку). Аналіз часових рядів відрізняється також і від аналізу просторових даних, де спостереження зазвичай відносяться до географічних розташувань (наприклад, підрахунок цін на будинки за розташуванням, а також за власними характеристиками цих будинків).

З огляду на високий попит на електроенергію дуже важливою є проблема прогнозу ціни та споживання електроенергії. Тому такі **питання** як:

Яку ціну очікувати завтра, наступного тижня, наступного року?

Яке споживання електроенергії очікувати завтра, наступного тижня, наступного року?

Є найбільш актуальними.

Існує безліч різних **підходів** для прогнозування часових рядів, таких як Авторегресія, Авторегресійна інтегрована ковзна середня(ARIMA), Просте експонентне згладжування (SES)і т.д.

Автоматичні прогнози ARIMA схильні до великих помилок при зміні тренду поблизу періоду відсікання і не можуть вловити сезонність. Прогнози експоненційного згладжування та наївного сезонного прогнозу вловлюють тижневу сезонність, але втрачають довгострокову сезонність. Усі методи надмірно реагують наприкінці річного періоду, оскільки вони не дуже точно моделюють річну сезонність.

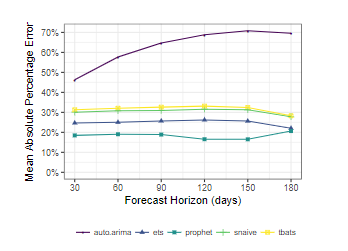
Ми використовуємо модель часового ряду з трьома основними компонентами моделі: тренд, сезонність та свята. Вони об'єднані у наступне рівняння



Тут g(t) - функція тренду, яка моделює неперіодичні зміни у значенні часового ряду, s(t) представляє періодичні зміни (наприклад, тижневу та річну сезонність), аh(t) представляє наслідки свят, що відбуваються за потенційно нерегулярним графіком протягом одного або кілька днів. помилки е --- представляє будь-які зміни, які не враховуються моделлю

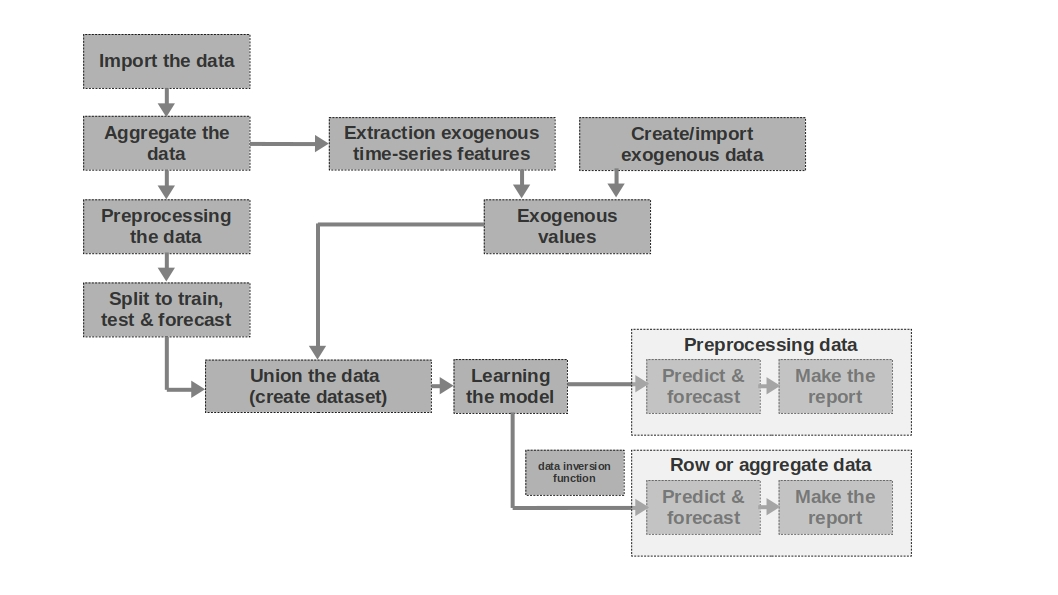
По суті, ми розглядаємо проблему прогнозування як вправу з підгонкою кривих(це метод побудови кривої або функції шляхом визначення тенденції в даних, яка має найбільшу кореляцію з реальними даними. В ідеалі, вона буде відображати тенденцію в даних і дозволить нам робити прогнози про те, як поводитиметься ряд даних у майбутньому., що за своєю суттю відрізняється від моделей часових рядів, які явно враховують структуру тимчасової залежності в даних). Хоча ми відмовляємося від деяких важливих переваг використання генеративної моделі, такої як ARIMA, таке формулювання забезпечує низку практичних переваг:

1. Гнучкість: Ми можемо легко врахувати сезонність за допомогою кількох періодів та далі робити різні припущення про тенденції.
2. На відміну від моделей ARIMA, виміри не обов'язково повинні бути рівномірно рознесені, і нам не потрібно інтерполювати недостатні значення
3. Модель прогнозування має параметри, що легко інтерпретуються, які можуть бути змінені аналітиком, щоб накласти припущення на прогноз. Більше того, аналітики зазвичай мають досвід роботи з регресією і легко можуть розширити модель, включивши в неї нові компоненти.
4. Підгонка відбувається дуже швидко, що дозволяє аналітику інтерактивно досліджувати багато специфікацій моделі



Прогноз зроблений prophet має значно меншу помилку передбачення, ніж інші автоматизовані методи прогнозування.

Моделювання часових рядів:



**Import the data**

Імпорт даних. На основі «сирих» даних моделюються перші 1-2 прогнози.

**Aggregate the data**

Якщо дані представлені з нерівними часовими інтервалами, їх необхідно агрегувати. Зазвичай беруть середнє значення в інтервалі (але за певних умов беруть max/min).

**Preprocessing the data**

Зазвичай під цим пунктом розуміють предобробку даних, після якої ряд стає стаціонарним.

**Split to train, test & forecast**

Часові ряди розбиваються на періоди навчання, тестування та прогнозування.

**Extraction exogenous time-series features**

Виділення зовнішніх додаткових чинників з ряду (вихідні, ремонтні роботи і т.д.) .

**Create/import exogenous data**

Додати зовнішні дані, які описують тимчасові події, що можуть сильно змінити значення ряду (введення санкцій, початок воєнних дій, природні катастрофи).

**Exogenous values**

Об’єднання зовнішніх даних.

**Union the data (create dataset)**

Об’єднання усіх зовнішніх даних. Іншими словами — готуємо датасет, що буде навчати модель та формувати прогноз.

**Learning the model**

Навчання моделі.

**Preprocessing data: predict & forecast**

Якщо була виконана предоброка даних, то якість моделі оцінюється спочатку на отриманих даних, а потім вже на «сирих».

**Row data: predict & forecast**

Заключний етап. Отримаємо результати.

**ДЛЯ ПОБУДОВИ ПРОГНОЗУ СПОЧАТКУ Приводимо до потрібного формату:**

Перший стовпець повинен мати ім'я "ds" і містити дату та час. Другий стовпець повинен мати ім'я «y» та містити спостереження.Це означає, що ми змінюємо імена стовпців у наборі даних. Також потрібно, щоб перший стовпець був перетворений на об'єкти дати та часу, якщо вони ще не були такими.

За допомогою допоміжної функції Prophet.make\_future\_dataframe створюємо dataframe, який містить усі історичні часові точки та ще 20, для яких ми хотіли побудувати прогноз.

Для того, щоб побудувати прогноз, викликаємо у моделі функцію predict

Вивід підсумовує те, що сталося в процесі навчання моделі, зокрема, процеси оптимізації, які виконувались

**У вигляді прогнозу Prophet поверне DataFrame із великою кількістю стовпців. Найцікавіші:**

ds - тимчасова мітка для прогнозованого значення;

yhat - прогнозоване значення;

yhat\_lower - нижня межа прогнозу;

yhat\_upper – найвища межа прогнозу.

З графіку видно, що У моделі є навички та прогноз, який виглядає розумним.

**Для оцінки якості передбачень** підрахували MAPE (mean absolute percentage error) – це середня абсолютна помилка нашого прогнозу.

Та абсолютну помилку MAE - mean absolute error, щоб розуміти, наскільки помиляється модель в абсолютних величинах. Бачимо, що модель працює непогано