

Презентация
к дипломной работе

на тему:

**«Анализ влияния цены акций на объем продаж компании COCO COLA (KO) NYSE -
NYSE Delayed Price за период с 19 января 1962 года по 19 декабря 2021 года.»**

Студент: Тарасова Т.А.

Наставник: Шестакова Е.А.

Введение: Для проведения анализа взята статистика набор данных для цены акций COCO COLA (KO) NYSE - NYSE Delayed Price, валюта в долларах США за период с 19 января 1962 года по 19 декабря 2021 года.

Цель: Проведение исследования данных и построение прогноза влияния цены акции (при открытии, закрытии, самой высокой и самой низкой) на объем продаж акций компании COCO COLA (KO) NYSE - NYSE Delayed Price.

Для проведения анализа в дипломной работе был использован Python и при помощи различных библиотек (полный перечень библиотек прописан в части 1 дипломной работы) были построены модели:

1. **PROPHET;**
2. **Экспоненциальное сглаживание" (Exponential smoothing)**

А) До построения моделей необходимо изучить статистические данные (dataset), убедиться в их корректности и возможности использовать, а именно:

- Корректность загрузки данных;
- Последовательность даты;
- Проверка на наличие пропусков в датасете;
- Проверка формата данных, числовые данные должны быть в формате число двойной точности (dtype: float64)
- Фильтрация данных

В) Проведение разведочного анализа данных EDA (exploratory data analysis), который состоит из:

- Вывода статистики по необходимым показателям;
- Расчета основных статистических метрик;
- Построения графического отображения метрик.

Выводы:

Промежуточный вывод: Видим уровень слабой корреляции между ценой продажи акции в течение одного дня. Строим гипотезу: будет ли в будущем расти объем продаж с учетом снижения цены акции Coca-cola.

На основании этого можно построить гипотезу: Будет ли в будущем расти объем продаж акций Coca-cola с вязи с уменьшением цены на акцию.

С) После того, как данные изучены и обработаны, можно приступать к построению моделей:

- Для всех моделей необходимо определить размер выборки и задать тестовую выборку (test) размером 1 год;
- Построить декомпозицию временного ряда (ETS композицию).

Вывод: На основании ETS можно наблюдать цикличную сезонность продажи акций.

2-я модель PROPHET:

- Необходимо подготовить данные для модели, сбросить индекс и переименовать столбцы;
- Подгоняем модель под наши данные:

Промежуточные выводы:

Алгоритм проигнорировал, выбирая параметры

Disabling yearly seasonality. Run prophet with yearly_seasonality=True to override this.

Недельную сезонность (для недельной сезонности необходимо иметь данные минимум за месяц, чтобы суметь использовать ее в модели)

Дневную сезонность (дневная сезонность может использоваться в случае, если данные собираются по часам/минутам, в нашем случае данные представлены дням: открытие, закрытие, сглаживание закрытия, высокая, низкая).

Метод обнаружил годовую сезонность (для годовой сезонности нам нужно иметь данные минимум за 2 года, чтобы суметь использовать ее в модели) INFO:fbprophet:Disabling daily seasonality. Run prophet with `daily_seasonality=True` to override this.

- Строим датафрейм на 12 месяцев;
- Предсказываем значения модели с доверительным интервалом в 95%;
- Выделяем основные поля для прогнозирования:

`ds` — дата прогноза

`yhat` — спрогнозированное значение

`yhat_lower` — нижняя граница доверительного интервала для прогноза

`yhat_upper` — верхняя граница доверительного интервала для прогноза

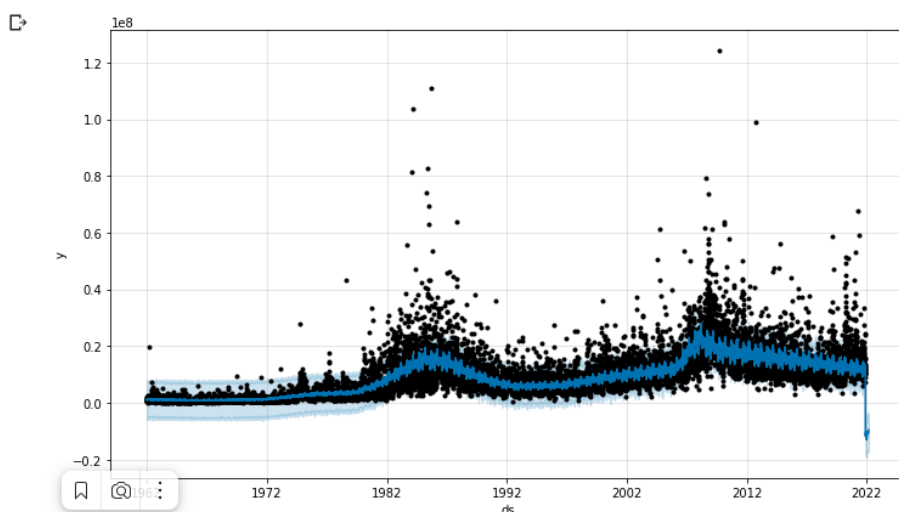
Смотрим прогноз на графике и раскладываем ряд на основные компоненты:

Видим, что тренд объема акций возрастающий и у нас есть годовая сезонность. Видим каким образом изменяются объемы по месяцам.

- Оцениваем качество модели методом MSE, RMSE, MAE, MAPE;
- Делаем прогноз на 1 год вперед;
- Устанавливаем индексы;
- Строим график на будущее.

✓
0 сек.

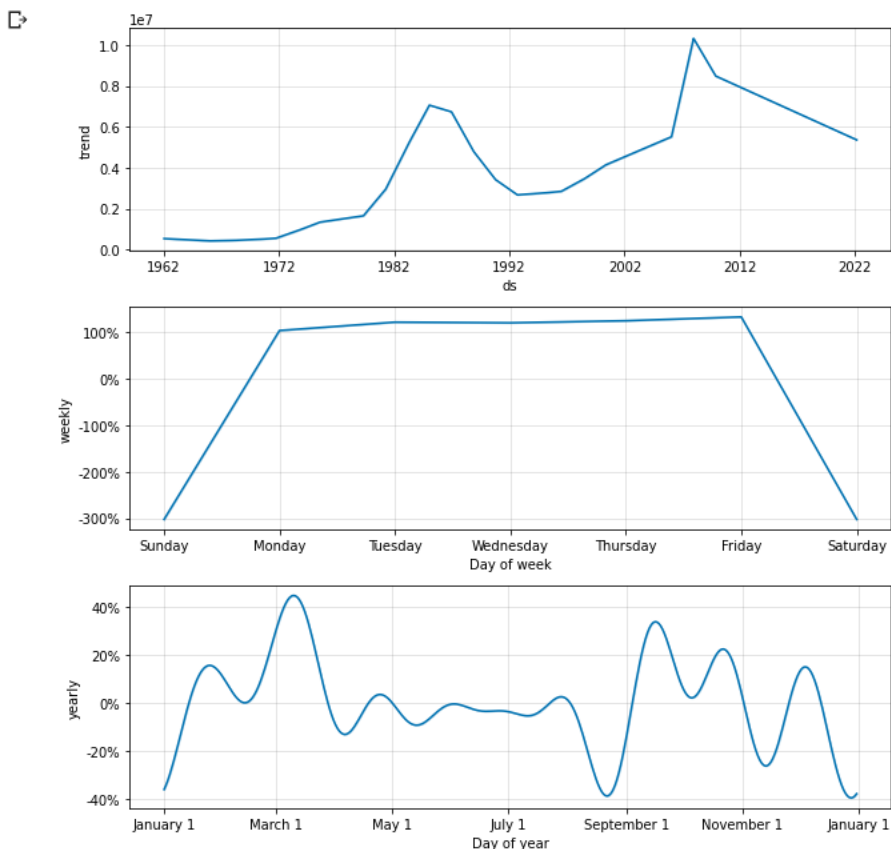
#сравниваем прогноз и тестовую выборку, где черные точки - выборка
`model_prophet.plot(prediction_prophet);`



Кроме того, Prophet позволяет также наглядно разложить ряд на основные компоненты — тренд и сезонность:

`plot_components()` — возвращает несколько графиков, среди которых тренд и столько сезонностей, сколько он найдет.

```
# разложим ряд на основные компоненты — тренд и сезонность
model_prophet.plot_components(prediction_prophet);
```



Промежуточные выводы: Видим, наличие возрастающего тренда стоимости акций и годовую сезонность. Видим каким образом изменяются цены по месяцам.

Промежуточные выводы: Видим, наличие возрастающего тренда стоимости акций и годовую сезонность. Видим каким образом изменяются цены по месяцам.

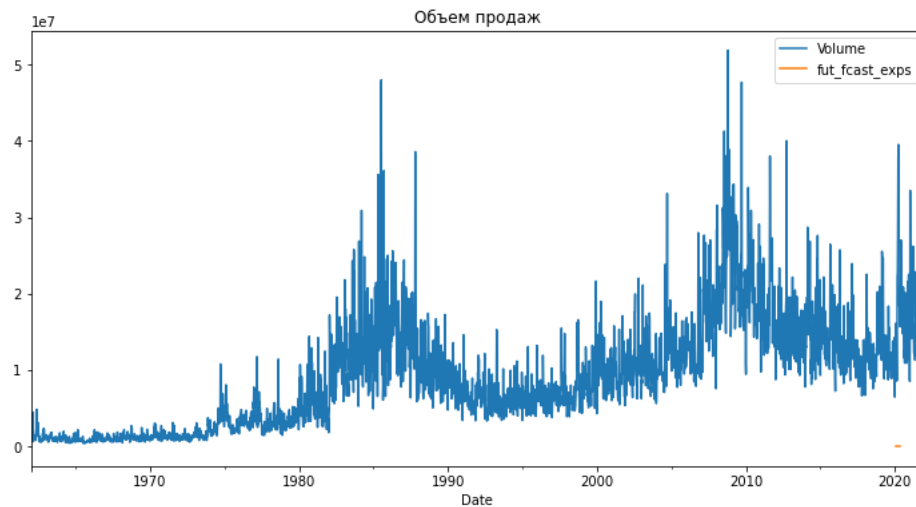
3й метод прогнозирования "Экспоненциальное сглаживание" (Exponential smoothing)

Экспоненциальное сглаживание — метод математического преобразования, используемый при прогнозировании временных рядов.

Метод также известен как метод простого экспоненциального сглаживания, или метод Брауна

- Создаем модель с подобранными параметрами;
- Обучаем модель на обучающей выборке;
- Предсказываем значения, передав модели точку начала и окончания;
- Сравниваем прогноз и тестовую выборку;
- Оцениваем качество модели методом MSE, RMSE, MAE, MAPE;
- Делаем прогноз на год вперед;
- Строим график на будущее.

```
[ ] #строим график на будущее
ax = df1['Volume'].plot(legend=True, figsize=(12,6),title='Объем продаж')
fut_fcast_exps['fut_fcast_exps'].plot(legend=True)
ax.autoscale(axis='x',tight=True)
```



Сравнительный анализ

Выводы по работе метода Экспоненциальное сглаживание: модель показала $MAPE = 20.42\%$

На основании того, что $MAPE > 90\%$ метод Prophet худший для прогнозирования данного временного ряда

Общие выводы:

- Проведен анализ данных с использованием современных методов обработки статистической информации.
- Рассчитаны основные статистические метрики, позволяющие судить о характере исследуемого явления.
- Результаты анализа представленных данных помогли выявить зависимость акции в зависимости от сезона : в летние месяцы наблюдается ровная динамика, в отличие от других месяцев.

Сравнивая 2 модели, лучше всего себя показала модель Экспоненциальное сглаживание.