1 Анализ временных рядов

1.1 Теория

Анализ временных рядов ─ совокупность математико-статистических методов анализа, предназначенных для выявления структуры временных рядов и для их прогнозирования. Под структурой временного ряда понимается наличие тренда, сезонной и циклической компоненты, случайной ошибки. Временной ряд может являться суммой вышеперечисленных компонент (аддитивная модель), или произведением (мультипликативная модель). Под прогнозированием понимается предсказание будущих значений временного ряда по настоящим и прошлым значениям

Одним из хороших способов выделения структуры временного ряда является использование процедуры Хольта-Винтерса, которое содержит уравнение прогноза и три уравнения сглаживания – показанного в формуле 1 (тут ссылка на носко). Процедура учитывает сезонность и тренд.

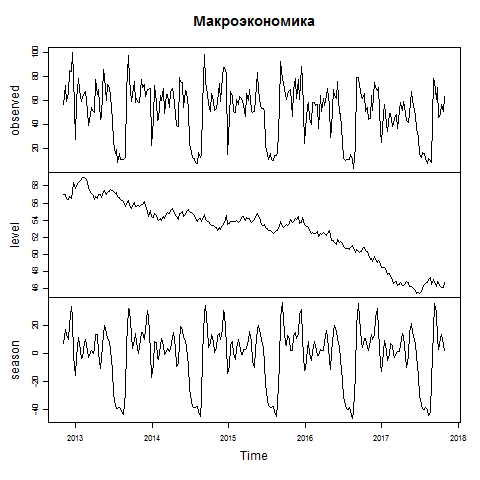
(формула 1.)

Развитием процедуры Хольта-Винтерса является процедура TBATS (Trigonometric Seasonal, Box-Cox Transformation, ARMA residuals, Trend and Seasonality) (тут ссылка на статью). Данная процедура реализована в языке R, и она включает или исключает уравнения сглаживания, исходя из значения AIC, использует ошибки вида ARMA, и выбирает по AIC использовать или нет трансформацию Бокса-Кокса. Сезонность в модели представлена тригонометрически.

1.2 Анализ временных рядов запросов предметов

Для анализа предложенных рядов использовалась процедура TBATS.

Для временного ряда запросов "Макроэкономика" была получена модель без тренда – Рисунок 1. Показатель «level » определяется для периода i как среднее арифметическое всех предыдущих значений, и, при отсутствии ярко выраженного тренда, может использоваться для прогнозирования. (тут ссылка на носко)



(Рисунок 1.)

Сезонная составляющая повторяется каждый год, и были выделены следующие закономерности:

1) Падение числа запросов в период летних каникул

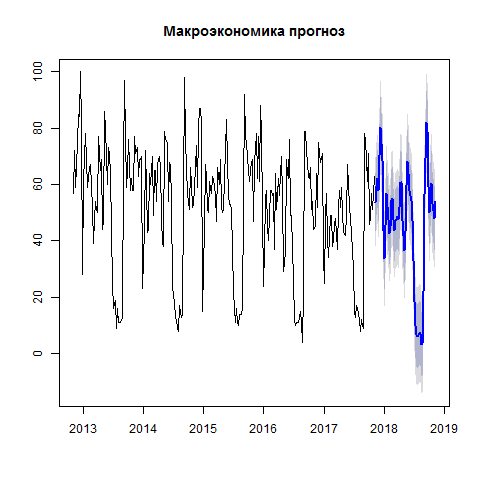
2) Падение числа запросов в праздничные недели Нового года, 1 и 9 мая

3) Увеличение числа запросов в неделю первого сентября

4) Увеличение числа запросов в неделю перед неделей празднования Нового года

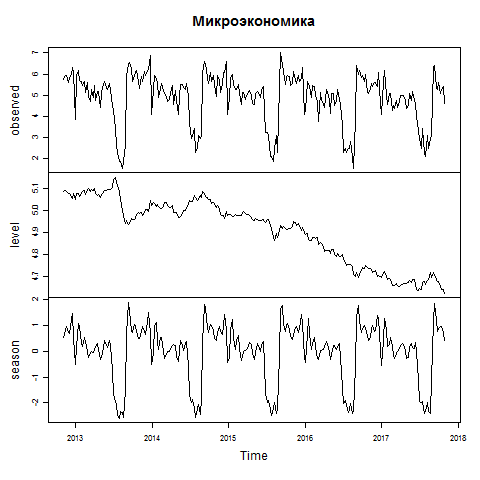
5) Увеличение числа запросов в неделю перед летними каникулами

Исходный ряд не является стационарным, но тест Льюнга-Бокса не отверг на уровне значимости 0,69 нулевую гипотезу о том, что остатки являются случайными данными (белым шумом). Таким образом, есть возможность получить прогноз поисковых запросов на следующий год ─ Рисунок 2.



(Рисунок 2)

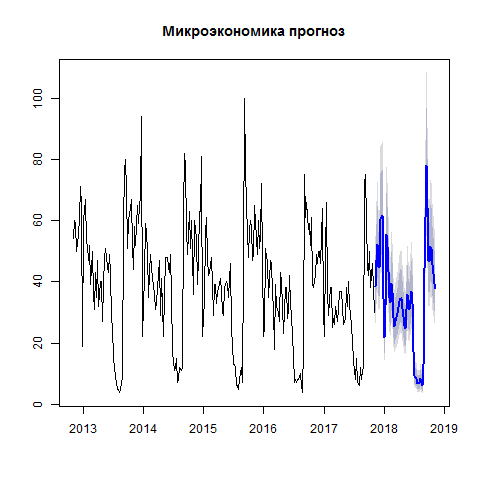
Как видно из прогноза, временной ряд запросов «Макроэкономика» имеет стабильный уровень, без тенденции к росту, что свидетельствует о неизменном уровне интереса к данному предмету.

Для временного ряда запросов «Микроэкономика» также была получена модель без тренда – Рисунок 3.

(Рисунок 3.)

Сезонная компонента повторяется каждый год, с аналогичным ростом количества поисковых запросов после знакомства с предметом, перед экзаменами. Наблюдается падение числа запросов в период летних и зимних каникул, праздников.

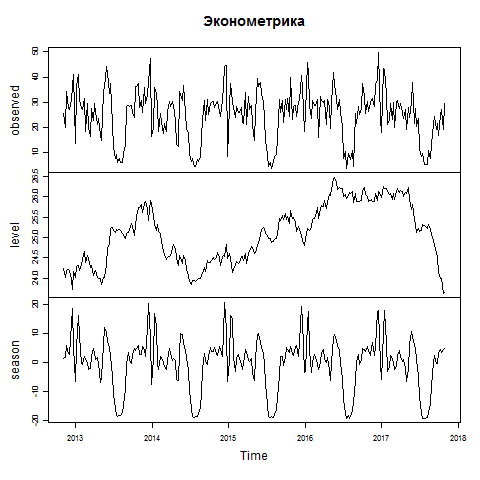
Тест Льюнга-Бокса не отверг на уровне значимости 0,79 нулевую гипотезу о том, что остатки являются случайными данными (белым шумом). Таким образом, есть возможность получить прогноз поисковых запросов на следующий год ─ Рисунок 4.



(Рисунок 4)

Временной ряд запросов «Микроэкономика» схож с временным рядом «Макроэкономика», что объясняется тесным симбиозом предметов и обязательному наличию их обоих в учебной программе. Ряд «Микроэкономика» не имеет долгосрочного временного тренда, что свидетельствует о стабильном уровне интереса к данному предмету.

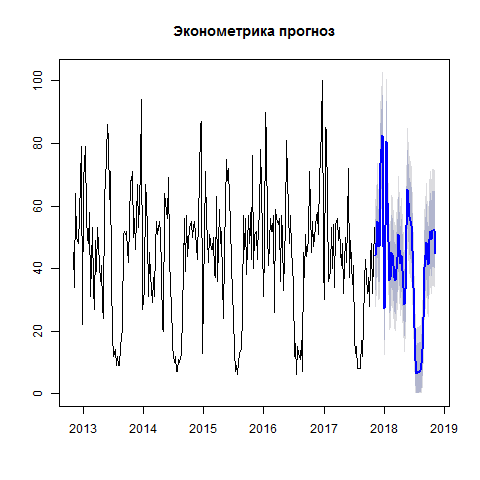
Для временного ряда запросов «Эконометрика» была получена модель без тренда – Рисунок 5.



(Рисунок 5.)

Сезонная компонента повторяется каждый год, и, в отличие от предыдущих предметов, не имеет место увеличение числа запросов в неделю первого сентября – когда происходит знакомство с предметом. Наблюдается падение числа запросов в период каникул, праздников, и рост перед зимними и летними экзаменами.

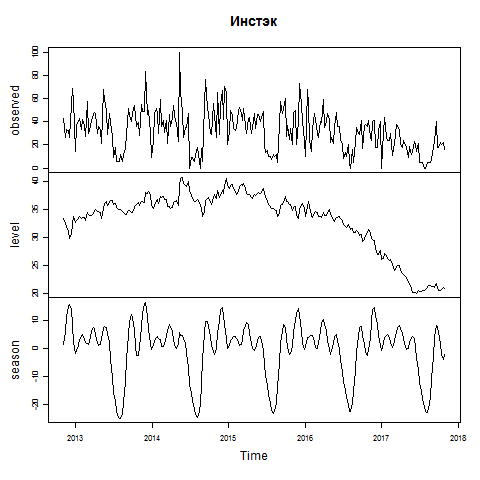
Тест Льюнга-Бокса не отверг на уровне значимости 0,17 нулевую гипотезу о том, что остатки являются случайными данными, поэтому есть возможность получить прогноз поисковых запросов на следующий год ─ Рисунок 6.



(Рисунок 6)

Временной ряд запросов колеблется вокруг значения 40, что меньше чем у предыдущих предметов. Это объясняется меньшей распространенностью данного предмета, зачастую он проходит лишь один семестр. Так как в модели отсутствует долгосрочный тренд, интерес к данному предмету стабилен.

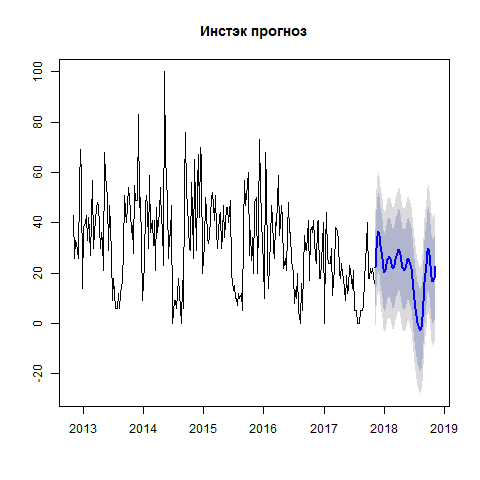
Для временного ряда запросов «Институциональная экономика» была получена модель без тренда – Рисунок 7.



(Рисунок 7.)

Сезонная компонента повторяется каждый год, наблюдается высокий скачок числа запросов после знакомства с предметом в первую неделю учебного года, что характеризует сложность предмета для восприятия. Активность перед экзаменами и снижение активности в период каникул и праздничных недель аналогична предыдущим предметам.

Тест Льюнга-Бокса не отверг на уровне значимости 0,79 нулевую гипотезу о том, что остатки являются случайными данными (белым шумом). Таким образом, есть возможность получить прогноз поисковых запросов на следующий год ─ Рисунок 8.



(Рисунок 8)

Временной ряд колеблется вокруг значения около 30, что является самым маленьким среди предыдущих трех предметов. Это объясняется редкостью курса, отсутствием курса на части экономических специальностей. В выбранной наилучшей модели отсутствует долгосрочный тренд, что характеризует стабильный интерес к предмету, но значение локального уровня снижается последние два года.