

1)

На графике А наблюдаются стабильно повторяющиеся колебания в относительно маленьких значениях. Не наблюдается зависимости продаж от дня недели или месяца. Соответственно, ошибка по горизонтальной оси не так сильно важна для этой функции. Лучше всего подойдет MAE, которая по сути считает среднее арифметическое всех ошибок. Т.к. ряд не имеет ярко выраженных выбросов, среднее ошибки будет не сильно колебаться.

Теоретически, RMSE и MAPE тут тоже можно использовать, но они обе предназначены для менее стабильных графиков, которые больше меняются по значению.

На графике В - заметная зависимость продаж от дня месяца/недели: можно наблюдать ярко выраженные локальные максимумы, которые чередуются с почти линейными участками. Тут не подойдет MAE, т.к. среднее будет сильно колебаться по значению. Можно использовать RMSE, чтобы отслеживать ошибку в расстоянии. Это отлично подойдет чтобы контролировать предсказание даты максимума функции. MAPE также хорошо подойдет, ведь она показывает процентную ошибку от реального значения => будет лучше для графиков с значительными изменениями в значениях

Политика выбора функции ошибки для графиков В и С одинакова, т.к. характер самих функции почти идентичен.

2)

График похож на график А из прошлого пункта этого же задания. Он имеет стабильные колебания, амплитуда которых не зависит от даты.

Дисперсия характеризует отклонение от средней величины => колебания графика. Тут они относительно небольшие, поэтому выбранная функция ошибки не подходит. Учитывая характер графика, можно использовать обычную линейную регрессию - она будет давать лучшие результаты, чем дисперсия, благодаря его (графика) стабильности и маленьким колебаниям.

3)

Иногда прогнозы суммы отличаются от суммы прогнозов для каждой категории. Но так как графики каждой категории сильно отличаются, более корректные значения стоит ожидать от суммы предсказаний первых трех моделей. Тогда можно написать код, который проверяет, одинаковы ли предсказание суммы и сумма предсказаний, и в противном случае, приравнивает первое ко второму. Это, если точность всех моделей одинакова.

Во втором случае сравним “доверие” к предсказаниям каждой модели. Сравним среднее арифметическое точности первых 3 моделей и 4, впишем в данные результат, у которого лучшая точность. (Либо сумма предсказаний первых 3, либо предсказание 4)