**Тестовое задание**

**Вопросы**

Ниже представлен список вопросов, на которые следует развёрнуто ответить, уделяя особое внимание нюансам задач класса «Marketing mix modeling».

1. Опишите последовательность, суть и нюансы основных этапов проекта по эконометрическому моделированию влияния медиакампаний (offline и Интернет) на целевой показатель (например, продажи). Какие этапы, на Ваш взгляд, являются наиболее ресурсоёмкими, а какие -  наукоёмкими? На каких этапах необходимо активное взаимодействие с клиентом и в чём оно выражается?
2. Каким образом можно сравнивать несколько моделей регрессии? Какие критерии и подходы при этом следует применять?
3. Статистика R^2 линейной регрессионной модели составляет 95%. Что это значит? Можно ли утверждать, что данная модель будет хорошим инструментом прогнозирования зависимой переменной?

**Практическое задание**

В прилагаемом файле «data\_for\_test.xlsx» представлены (на листе «Raw data») временные ряды предикторов и целевого показателя (суммарное количество телефонных звонков возможных покупателей в дилерские центры, реализующие автомобиль определённой марки). Описание факторов дано на листе «Data description». Допускается дополнение данных новыми факторами (в т.ч. производными от имеющихся).

**Основная задача: построить регрессионную модель предсказания зависимой величины с помощью предоставленных факторов. Подзадачи:**

1. провести разведочный анализ данных (описательные статистики, распределения, сводные таблицы, графики), среди графиков желателен график динамики целевой величины с отметками периодов рекламной активности; содержательно описать результаты анализа;
2. построить регрессионную модель, выполнить её валидацию на отдельной подвыборке;
3. оценить качество и адекватность модели, выполнить необходимые тесты выполнения предпосылок используемой модели (если тесты не выполняются, то объяснить, почему, и как с этим бороться);
4. количественно оценить изменение степени влияния рекламной активности на ТВ;
5. построить график динамики целевого показателя, дополненный графиком оценок целевого показателя, полученных с помощью модели;
6. дать содержательную интерпретацию полученным результатам и итоговой модели, сформулировать рекомендации по совершенствованию модели и привести дополнительные факторы, которые можно было бы использовать для улучшения объясняющей силы и прогностической способности модели.

**Результаты работы следует представить в виде текстового отчёта, включающего:**

1. описание задачи и факторов;
2. описание результатов разведочного анализа;
3. описание полученной модели, оценок её качества и адекватности, выявленных проблем;
4. интерпретацию выявленных закономерностей;
5. необходимые таблицы и графики.

**Отдельно следует приложить воспроизводимый программный код на предпочитаемом языке программирования, использованный для решения задачи. Также можно выполнить всё задание в виде единого файла Jupyter Notebook / R Markdown.**

**Срок выполнения задания:** 1 неделя с момента получения задания.