**Отчёт по винам.**

Подготовил Михайлишин Виктор Валерьевич.

**Введение.**

На сегодняшний день компании, производящие алкогольную продукцию, сталкиваются с высокой конкуренцией. Из-за этого предварительная оценка того, как вино будет воспринято на рынке экспертами и потребителем очень важна, потому что необходимо создавать положительную репутацию в глаза вышеупомянутых категорий людей. Для этого необходимо соблюдать определённые стандарты, в число которых входит как качество вина, так и соответствие этого качества цене.

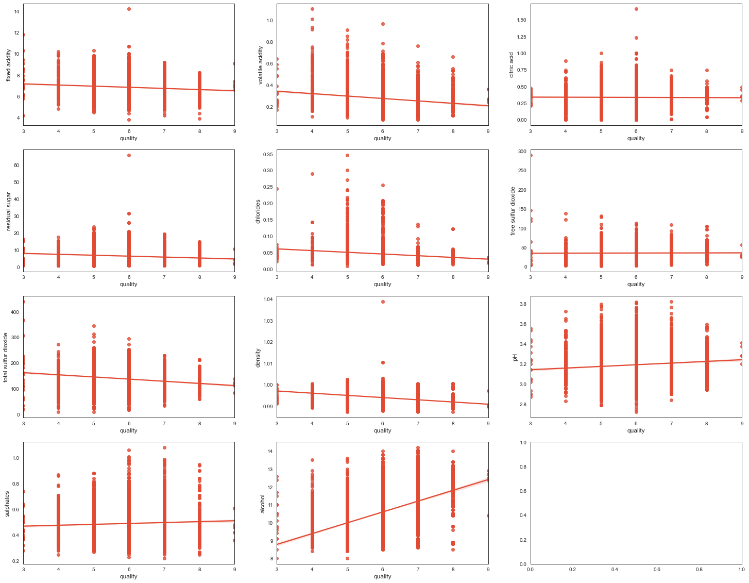
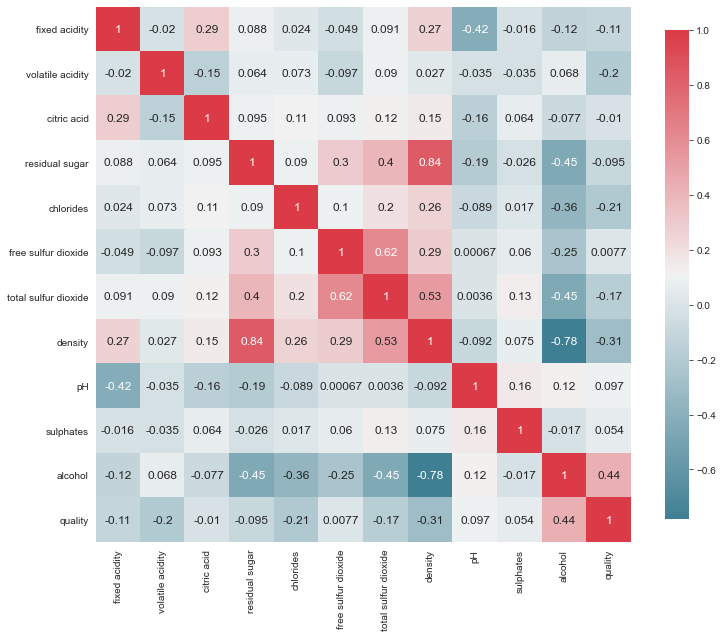
**Задачи исследования и информация о данных.**

В предложенных к изучению данных была предложена таблица с описанием химического анализа разных вин, их характеристика и оценка.

Задачей данного исследования было понять, с какими же критериями наиболее тесно коррелирует качество вина, а также построить модель, которая бы предсказывала качество вина на основании данных химического анализа вина и его типа.

Одной из проблем, с которых пришлось столкнуться – это обработка пропусков, однако ввиду малого их количества было принято отбросить неполные результаты химического исследования.

На основании полученных данных была построена тепловая карта корреляции, чтобы выявить наиболее влияющие на качество вина признаки.

А также построено распределение между признаками и оценкой вина.  

На основании полученных данных удалось выявить следующие корреляции:

- Качество резко повышается при снижении volatile acidity;

- Качество повышается с увеличением citric acid;

- Качество резко повышается с увеличением alcohol.

После чего данные были разделены в случайном порядке на тестовую и тренировочную выборки в соотношении 75:25.

**Выбор и построение модели для определения качества вина.**

Ввиду необходимости проводить регрессию было принято решение выбрать для модели алгоритм стэкинга, состоящего из LinearSVR, RidgeCV и StackingRegressor. После чего модель была обучена и протестирована на имеющихся выборках из данных.

**Оценка**

В ходе проведения предсказаний среднее квадратичное отклонение при 10-ти бальной оценке составило около 0.65. Таким образом данная модель позволяет предсказывать качество вина с высокой точностью, что в будущем позволит применять её при предварительной оценке вина на заводе и последующем грамотном позиционировании его на рынке.