

Бриф «Девхаб». Анализ схем севооборота

Фермеры и агрохолдинги стремятся увеличивать доход, повышая урожайность и снижая издержки. Для этого применяется **севооборот** — поочерёдное выращивание различных культур: например, сначала сеют озимую рожь, на следующий год — картофель, потом — яровую пшеницу, а после неё — клевер. Это помогает бороться с сорняками и поддерживать плодородность почвы, обогащая её питательными веществами. Чередую культуры в определённом порядке, фермеры экономят на удобрениях и пестицидах.

Зная стратегии севооборота, которые применяют фермеры, можно предсказать, как будут изменяться характеристики почвы и урожайность. Это позволяет прогнозировать объём произведённых продуктов в масштабе регионов и государств, а также изменение цен на международном рынке сельхозпродукции.

Вам необходимо:

- проанализировать набор данных,
- выявить последовательность смены посевных культур на заданных полях за десять лет,
- предсказать, какую культуру там сеяли в 2020 году.

Набор данных собран из открытой статистики о посевах во Франции за 2010–2019 год. Предсказание необходимо сделать для 100 полей в регионе Haute-de-France, а также в рандомизированной выборке из 100 000 полей для 2019 года (**этой выборки не будет** в представленном наборе данных).

Ваша задача — для каждой точки `Point()` вывести прогнозируемую культуру посева на указанный год в файле `predict.csv`.

Дополнительные сведения

Ограничения схемы севооборота — одну и ту же культуру не сеют два года подряд. Исключения: виноградники, лесные хозяйства и луга, а в некоторых регионах — кукуруза.

Схемы севооборота можно разделить на **три группы**:

- короткую — 2–3 года;
- среднюю — 3–6 лет;
- долгосрочную 6–10 лет.

Короткий цикл применяют для более прибыльных культур — он сильнее истощает почву. Средние и длительные циклы истощают почву меньше. Когда цикл завершается, последовательность культур повторяют.

На предпочтительный выбор посевных культур влияет **специфика климата**: на севере, юге, востоке и западе выращивают разное.

Координаты центра поля приведены в формате WKT в проекции Lambert-93 ([ESPG:2154](#)), единица координат — 1 метр.

Источник данных — [открытая статистика о посевах во Франции](#).

Метрика

Подходящую для решения задачи метрику вам необходимо выбрать самостоятельно.

Данные

- Данные о культивируемом посеве с привязкой к центру поля за 2015–2019 годы. За год собирается около 9.5 млн записей.

- Данные об основной культивируемой группе культур с привязкой к центру участка за 2010–2014 года. За год собирается около 6 млн записей.

Участок — это непрерывный кусок земли, принадлежащий фермеру. На этой земле фермер выращивает одну или несколько культур, по-разному разбивая участок каждый год.

- [Список культур и их принадлежность к группе](#)
- Данные представлены [по ссылке на Kaggle](#).
- [Данные об основной культивируемой группе культур](#) с привязкой к центру участка за 2010–2014 годы.