Домашнее задание №1

## Общие замечания.

* Решение каждой задачи помещайте в отдельный каталог вида Task\_NN, например – Task\_01 (см. примеры кода к лекциям).
* Все решения запакуйте в один архив и отправьте по почте тренеру.

## Задача 1.

[CSV-файл](https://en.wikipedia.org/wiki/Comma-separated_values) **data.csv** хранит информацию об отпусках сотрудников некоторой компании. Каждая строка содержит имя и фамилию сотрудника (имя-пробел-фамилия), дату начала отпуска, дату окончания отпуска (даты – в американском формате). Вот фрагмент файла:

Dlsw Fmbnaswsg,1/25/2016,2/3/2016

Iqbidtaw Ndxqjdpsg,6/8/2015,6/19/2015

Lrnywj Fdgyzw,6/8/2015,6/16/2015

Имена сотрудников зашифрованы (расшифровывать их не надо). Полный файл находится в приложении к заданию.

Ваша программа должна считать файл **data.csv**, вычислить для каждого сотрудника суммарную продолжительность всех его отпусков в днях и сохранить эту информацию в файл **out.txt**, расположив сотрудников в порядке убывания продолжительности отпусков. Вот так должен выглядеть файл **out.txt** (фрагмент):

Lrnywj Fdgyzw: 10

Dlsw Fmbnaswsg: 5

Iqbidtaw Ndxqjdpsg: 1

Примечания:

1. Если у двух сотрудников одинаковая продолжительность отпусков, расположите их по алфавитному порядку фамилий.
2. Если при разборе строки из файла **data.csv** обнаруживаются некорректные данные, такая строка записывается в файл **errors.txt** вместе со своим порядковым номером. Например,

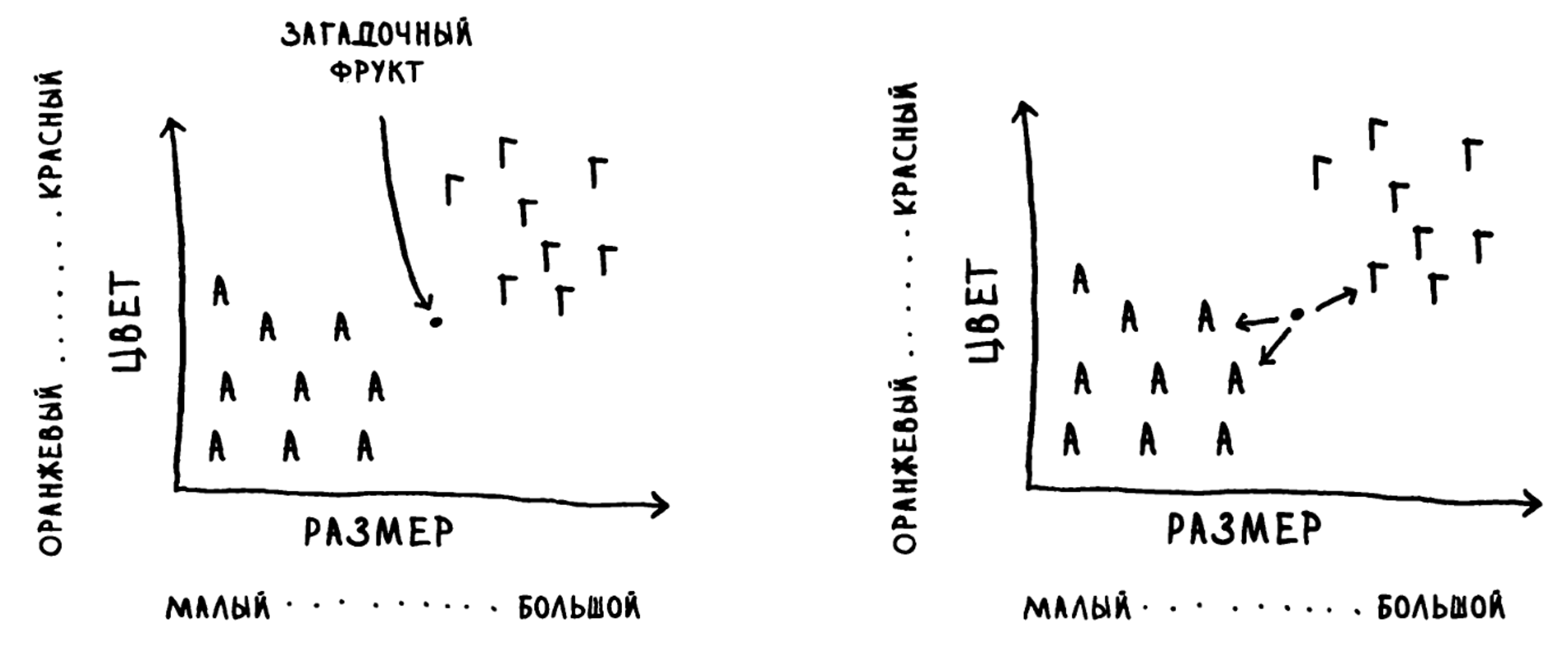
105 Lrnywj XXX Fdgyzw,6/8/2015,6/16/2015

## Задача 2.

*Метод k-ближайших соседей* – один из простейших алгоритмов машинного обучения (<https://en.wikipedia.org/wiki/K-nearest_neighbors_algorithm>).

Суть алгоритма: пусть имеется набор объектов с известными названиями и числовыми характеристиками. Эти характеристики дают возможность описать объект как точку в n-мерном пространстве. Пусть есть объект, у которого заданы характеристики, но нет названия. Чтобы найти название этого объекта, найдём k ближайших к нему известных объектов. То название, которое будет у ближайших соседей повторяться чаще, сделаем названием нашего неизвестного объекта (алгоритм классификации).

Пример: апельсины и грейпфруты, k=3. Неизвестный объект оказался апельсином.



1. Необходимо реализовать алгоритм метода k-ближайших соседей в двумерном пространстве характеристик. Вход алгоритма: набор известных объектов, две характеристики неизвестного объекта и значение k. Каждый объект состоит из строкового названия и двух (для простоты) числовых характеристик. Выход алгоритма: строка с названием неизвестного объекта.
2. Создайте консольное приложение для тестирования алгоритма. При старте приложение читает входные данные из текстового файла **data.txt** (сам файл вам предоставляется). Затем приложение в бесконечном цикле запрашивает у пользователя две характеристики и выводит строку с названием объекта.