Домашнее задание №6

Загрузите данные при помощи spark.read.csv из приложенного файла

## Воспользуйтесь командами groupBy, max, min и avg, чтобы вывести максимальное минимальное и среднее значние каждого из аттрбутов цветка ('sepal\_length','sepal\_width','petal\_length','petal\_width')

## Выведите по аналогии график petal\_length и petal\_width, так же раскрашенный по типам цветков

## Посмотрите внимательно на получившийся график и сформируйте правило вида ЕСЛИ ЗНАЧЕНИЕ ПАРАМЕТРА A < ЗНАЧЕНИЕ, ТО IRIS ОТНОСИТСЯ К КЛАССУ SETOSA - ИНАЧЕ К КАКОМУ-ТО ДРУГОМУ (НАЗОВЕМ, ЕГО НАПРИМЕР UNKNOWN)

## При помощи команд withColumn, when, otherwise создайте в датафрейме новую колонку с таким предсказанием

## Проверьте качество своего предсказания

Для этого:

1. Подсчитайте количество строк, где Вы предсказали класс Setosa и правильный класс был так же Setosa
2. Подсчитайте количество строк, где Вы предскзали класс Setosa, а правильный класс был какой-то другой (воспользуйтесь оператором "не равно" - !=)