

Добрый день!

Данные задания необходимо сделать в **одном** Jupyter Python Notebook на платформе [COLAB](#) и предоставить результаты работы ссылкой, приложив в текстовый файл на платформу.

Для дальнейшей работы нам понадобится датасет, который можно скачать [здесь](#).

Описание датасета находится ниже. Обратите внимание что это оперативная информация по игрокам (игроки дублируются) сквозь сезоны баскетбольной лиги.

№№	Столбец	Описание
1	index	Номер строки в файле
2	player_name	Имя игрока
3	team_abbreviation	Краткое сокращение команды
4	age	Возраст игрока
5	player_height	Рост игрока
6	player_weight	Вес игрока
7	college	Учебное заведение игрока
8	country	Страна игрока
9	draft_year	Год найма (драфта)
10	draft_round	Раунд найма (драфта)
11	draft_number	Номер игрока в найме (драфте)
12	gp	Количество игр игрока за сезон
13	pts	Среднее число очков за игру
14	reb	Среднее число подборов за игру
15	ast	Среднее число голевых пасов за игру
16	net_rating	Разница очков команды за 100 владений, пока игрок находится на площадке
17	oreb_pct	Процент доступных подборов в нападении, которые игрок захватил, находясь на площадке
18	dreb_pct	Процент доступных подборов в защите, которые игрок захватил, находясь на площадке
19	usg_pct	Процент командных игр, использованных игроком, пока он находился на площадке
20	ts_pct	Мера эффективности броска игрока, учитывающая штрафные броски, 2- и 3-очковые броски.
21	ast_pct	Процент бросков с игры товарищей по команде, которым игрок ассистировал, находясь на площадке
22	season	Сезон

- 1) Предварительно загрузите данные в любую SQL реляционную базу данных и затем используйте эти данные с загрузкой в соответствующий dataframe.
- 2) Затем, используя
 - Линейную регрессию
 - Логистическую регрессию
 - SVM

укажите зависимость роста (`player_height`) от веса (`player_weight`) для уникальных игроков всех команд по всем сезонам. Постройте визуализацию используя библиотеки Matplotlib / Seaborn.

По каждому из методов выведите оценку score модели.

- 3) Используя алгоритм Древа Принятия Решения попробуйте определить **название команды (`team_abbreviation`)** как целевое значение функции (Y) за сезон 2010-11 от следующих атрибутов (X_i)
 - Среднее **pts** всех игроков команды за сезон
 - Среднее **reb** всех игроков команды за сезон
 - Среднее **ast** всех игроков команды за сезон
 - Среднее **player_height** всех игроков команды за сезон
 - Среднее **player_width** всех игроков команды за сезон
- 4) Используя ваш личный выбор атрибутов на предложенном датасете постройте классификацию данных по любому (любым) атрибутам-признакам (в том числе можете построить новый вычисляемый атрибут-признак)