Детекция сарказма

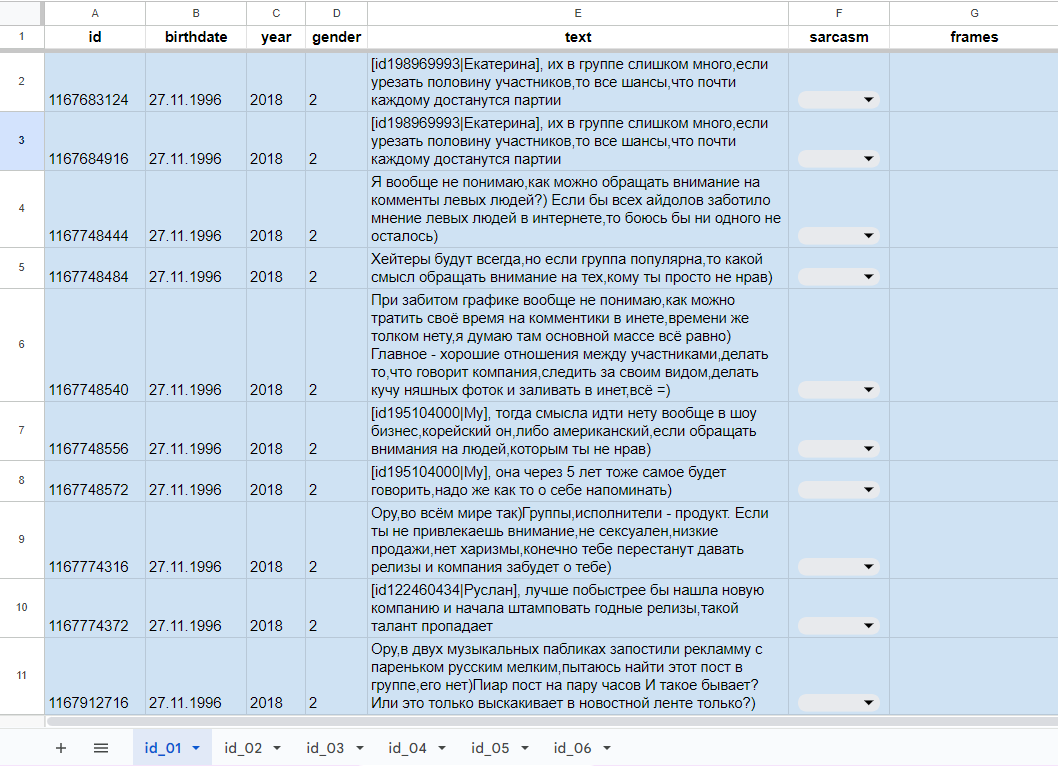
**Сарказм** — это особый вид высказывания, целью которого является выражение насмешки, презрения или оскорбления в сторону собеседника, объекта речи или описываемой ситуации и смысл которого прямо противоречит его форме.

Именно ввиду неоднозначности и противоречивости саркастических высказываний детекция сарказма может быть непростой задачей как для людей, так и для нейронных сетей!

Примеры саркастических выражений:

* О да, как я люблю, когда меня игнорируют!
* Было очень весело отметить МОЙ день рождения в компании тебя и твоего парня, просто праздник мечты)))
* Хорошо, что я вам не нравлюсь. Не у всех есть хороший вкус.
* Мне жаль, что я обидел тебя, когда назвал тупым. Я думал, что ты это уже знал.
* Вы так интересно рассказывали, что я всего два раза заснул и выпил пять кружек кофе.

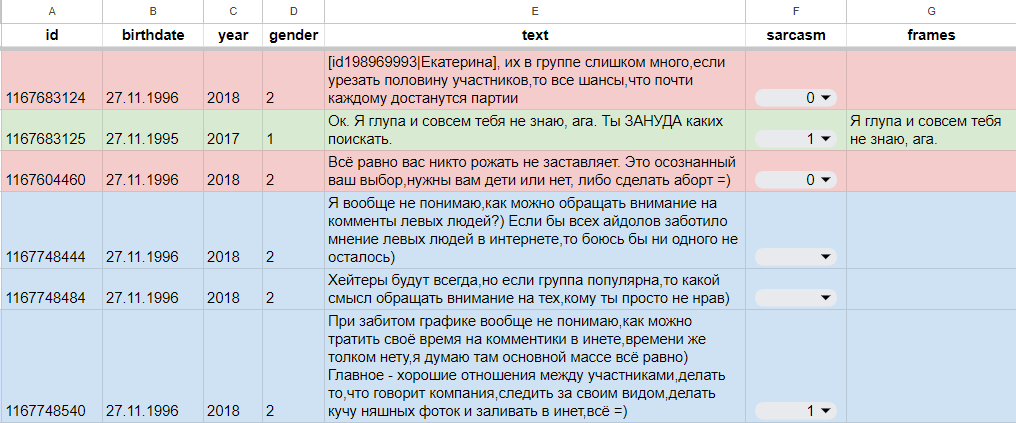
Работа с таблицей



Разметчик работает только на листе, обозначенном своим id-номером, и вносит информацию только в столбцы *sarcasm* и *frames*.

В столбец *sarcasm* вносится информация о том, содержит ли текст в строке саркастическое выражение (1 – содержит, 0 – не содержит). Столбец *frames* заполняется, только если значение столбца *sarcasm* равно 1. В столбце *frames* нужно указать конкретное место, которое можно обозначить как сарказм (обязательно **копируйте** информацию из столбца *text*, а не переписывайте её!).

Например:



Все несделанные строки в таблице окрашены в голубой цвет.

Сделанные и не содержащие сарказма автоматически окрашиваются в красный.

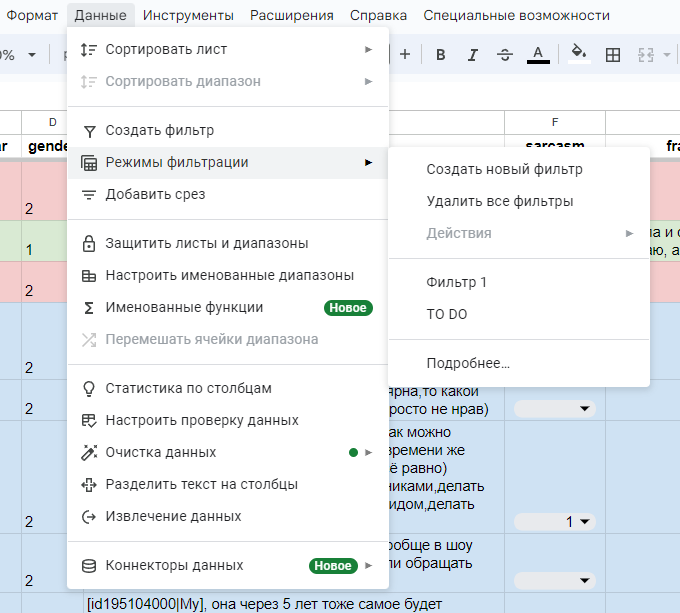
Сделанные и содержащие сарказм – в зелёный.

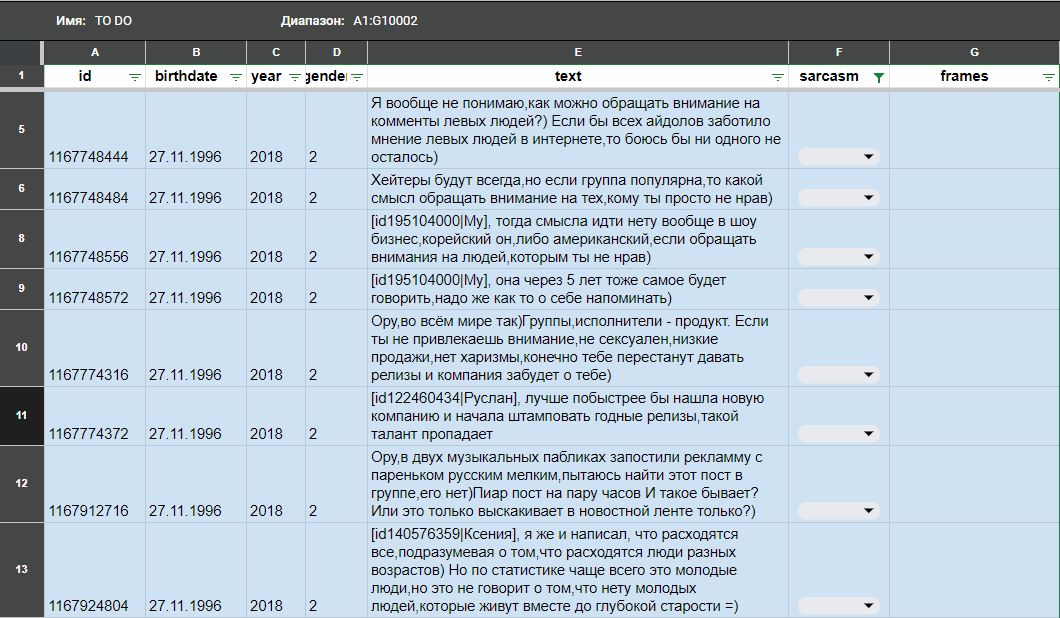
Строка, у которой в столбце *sarcasm* стоит значение *1*, при этом в столбце *frames* отсутствует какое-либо значение, считается несделанной и остаётся выделенной голубым цветом.

Строки, в которых весь текст состоит на иностранном языке или почти полностью дублируется с другим текстом (например, спам), следует отнести к дефектным и не размечать. В столбце defect, можно помечать забракованные вами строки. Для этого нужно указать значение в ячейке defect, равное «1». Для незабракованных текстов ставить «0» в этом столбце не обязательно, можно оставить его пустым. чтобы помечать строчки для удаления. Такие строки становятся жёлтыми:



Посмотреть несделанные строки можно при помощи режима фильтрации *TO DO* (Данные / Режимы фильтрации / TO DO):





Либо пролистав до выделенных голубым ячеек :)