

ЭЛЕКТРОННЫЕ ТАБЛИЦЫ: ПРИМЕНЯЕМ С ПОЛЬЗОЙ

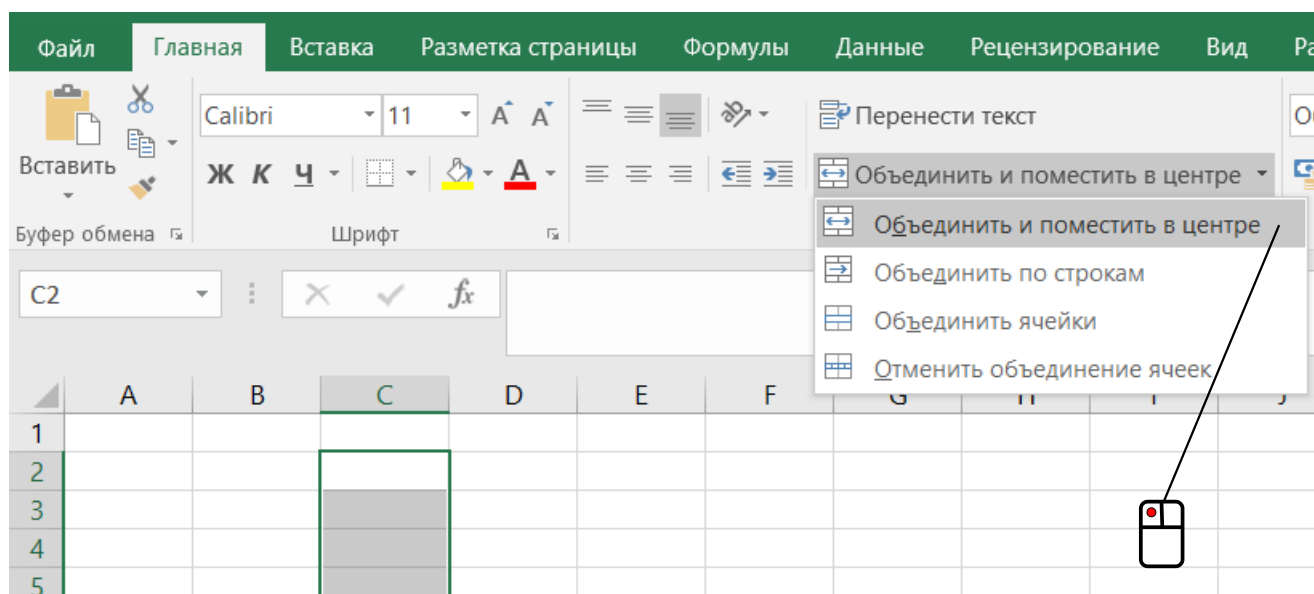
Ищем аллитерации в стихах

Для версий:
Microsoft Office 2016
LibreOffice 6.4

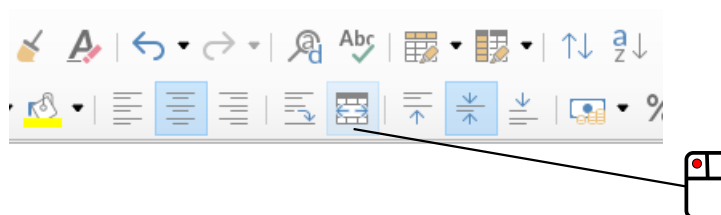
Аллитерация – это скопление одинаковых или схожих по звучанию согласных в каком-либо тексте. В поэзии её нередко используют в качестве приёма, создающего дополнительную звуковую выразительность стиха, кроме того аллитерации встречаются в скороговорках и логопедических упражнениях. Ниже предлагается описание создания инструмента для анализа частот встречаемости букв в текстах для выявления этих самых аллитераций.

Создайте новую книгу и выделите на листе диапазон ячеек “C2:C26”:

На вкладке «Главная» ленты, в секции «Выравнивание» выберите команду «Объединить и поместить в центре»:



На панели инструментов «Форматирование» нажмите кнопку команды «Объединить и центрировать ячейки»:



После этого действия выделенный диапазон станет выглядеть как одна большая ячейка – установите в ней горизонтальное выравнивание текста по левому краю, а вертикальное – по верхнему (*Пособие*, с. 11). Стоит отметить, что к содержимому такой большой ячейки можно обращаться, используя адрес “C2”.

Далее, воспользовавшись автозаполнением (*Пособие*, с. 40), проставьте в ячейках столбца “А” числа от 1 до 3000:

	A	B
2993	2993	
2994	2994	
2995	2995	
2996	2996	
2997	2997	
2998	2998	
2999	2999	
3000	3000	
3001		
3002		
3003		

Затем необходимо заполнить и отформатировать лист примерно следующим образом:

	A	B	C	D	E	F
1	1		Анализируемый текст:	Буква	Количество	Относительная частота встречаемости
2	2			а		
3	3			б		
4	4			в		
5	5			г		
6	6			д		
7	7			е		
8	8			ё		
9	9			ж		
10	10			з		
11	11			и		
12	12			й		
13	13			к		
14	14			л		
15	15			м		
16	16			н		
17	17			о		
18	18			п		
19	19			р		
20	20			с		
21	21			т		
22	22			у		
23	23			ф		
24	24			х		
25	25			ц		
26	26			ч		
27	27			ш		
28	28		щ			
29	29		ъ			
30	30		ы			
31	31		ь			
32	32		э			
33	33		ю			
34	34		я			
35	35					
36	36			Всего:		
37	37					

Нетрудно видеть, что “С2” предназначена для размещения в ней анализируемого текста – введите в неё какую-нибудь фразу, например: «Тихо шифером шурша, крыша едет не спеша». Так как скорее всего анализируемый текст будет в дальнейшем вставляться в ячейку из буфера обмена, то рекомендую делать это, предварительно установив текстовый курсор в строке формул (строке ввода) – дело в том, что длинное стихотворение может содержать много

символов перевода строки и при обычной вставке такой текст вставится не только в “С2”, а ещё и расположенные под ней ячейки (“С27”, “С28” и т. д.).

Из сказанного становится ясно, что текст перед анализом лучше предварительно подготовить. В ячейку “С29” введите следующую формулу:

```
=СТРОЧН (ПЕЧСИМВ (СЖПРОБЕЛЫ (С2) ) )
```

```
=LOWER (CLEAN (TRIM (С2) ) )
```

Работает формула так. Функция

```
СЖПРОБЕЛЫ ( )
```

```
TRIM ( )
```

убирает лишние пробелы, заменяя несколько подряд идущих одним. После этого функция

```
ПЕЧСИМВ ( )
```

```
CLEAN ( )
```

убирает из текста специальные непечатаемые знаки, в том числе – символы перевода строки. Заканчивается обработка функцией

```
СТРОЧН ( )
```

```
LOWER ( )
```

– она делает все заглавные буквы строчными.

В “С28” введите вот такую формулу:

```
= "Подготовленный текст (число знаков " & ДЛСТР (С29) & " ) : "
```

```
= "Подготовленный текст (число знаков " & LEN (С29) & " ) : "
```

Использованная в ней функция

```
ДЛСТР ( )
```

```
LEN ( )
```

возвращает число символов в тексте (в данном случае – «очищенном» от лишних знаков). Обратите также внимание на особенность рассматриваемой формулы – она формирует строку, в которую внедрено число, присоединяемое к текстовым фрагментам при помощи символов & (амперсанд), сами текстовые фрагменты в формуле берутся в кавычки.

В ячейке “В1” укажите ещё одну формулу:

```
=ПСТР ($С$29 ; А1 ; 1)
```

```
=MID ($С$29 ; А1 ; 1)
```

Использованная в ней функция осуществляет вырезание фрагмента из текста. У неё три аргумента: первый – адрес ячейки, в которой содержится обрабатываемый текст, второй – номер символа (позиция знака), с которого надо начинать вырезание, третий аргумент – количество вырезаемых знаков. Таким образом в данном случае формула возвращает первый символ строки

из ячейки “С29”. При помощи маркера заполнения (*Пособие*, с. 25) откопируйте формулу из “В1” вплоть до “В3000”, в результате в диапазоне “В1:В3000” будет отображаться подготовленный анализируемый текст, разделённый на отдельные знаки. Дополнительно следует заметить, что если в используемой формуле второй аргумент функции (начальная позиция для вырезания текста) по своей величине превышает число знаков в обрабатываемом тексте, то сама функция в этом случае возвращает пустую строку "".

В ячейку “Е2” введите формулу такого вида:

$$=СЧЁТЕСЛИ(\$B\$1:\$B\$3000;D2)$$
$$=COUNTIF(\$B\$1:\$B\$3000;D2)$$

Использованная в ней функция (*Пособие*, с. 84) подсчитывает число ячеек в диапазоне “В1:В3000”, содержащих букву «а» – откопируйте формулу вниз до ячейки “Е34” и в диапазоне “Е2:Е34” отобразятся количества каждой буквы алфавита в анализируемом тексте. Для завершения вычислений полученный массив чисел следует пронормировать (*Пособие*, с. 76), определив значения относительных частот встречаемости букв. Для этого в “Е36” вводится формула

$$=СУММ(E2:E34)$$
$$=SUM(E2:E34)$$

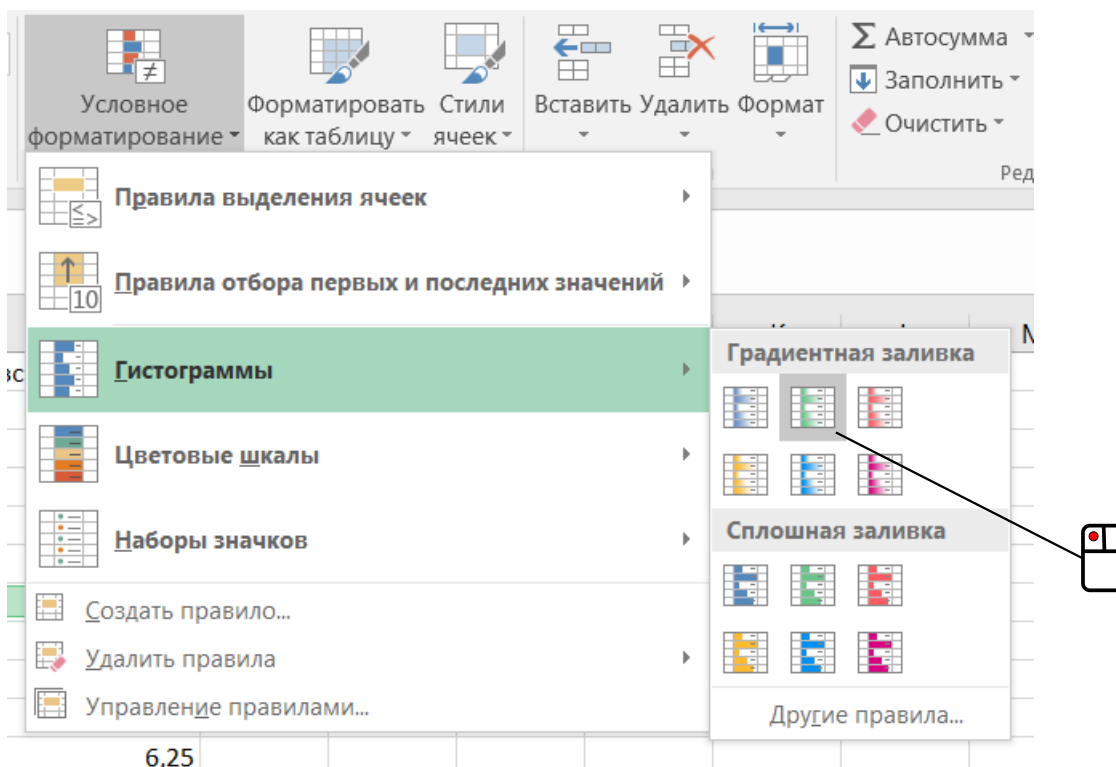
для подсчёта общего количества букв в анализируемом тексте, а в “F2” – формула:

$$=E2/ΣE$36*100$$

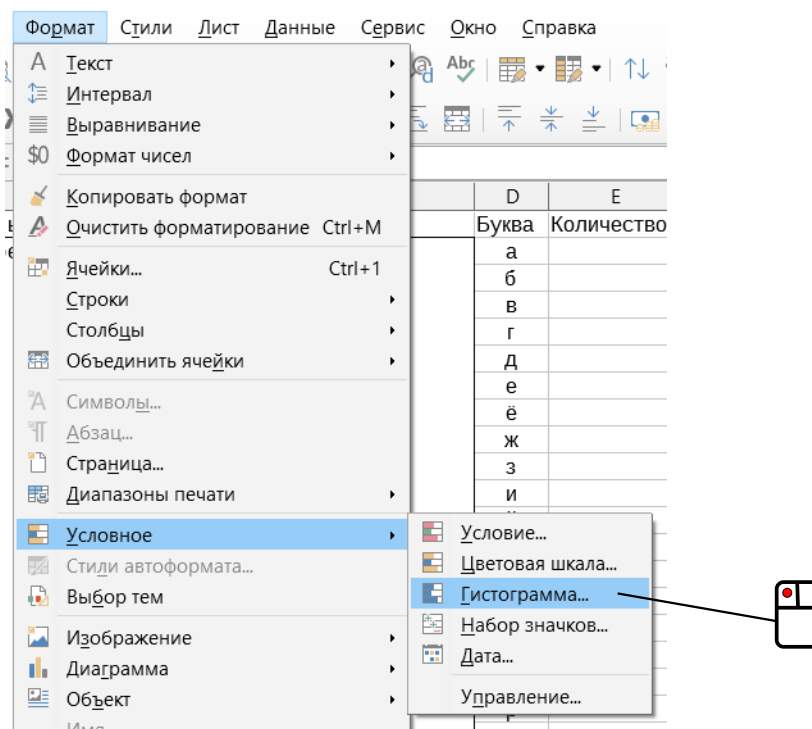
Последнюю формулу также следует откопировать вниз, до “F34”, получив нормированные значения (относительные частоты).

Расчитанные величины лучше как-нибудь визуализировать. В данном случае предлагается использовать условное форматирование ячеек – выделите диапазон “F2:F34”.

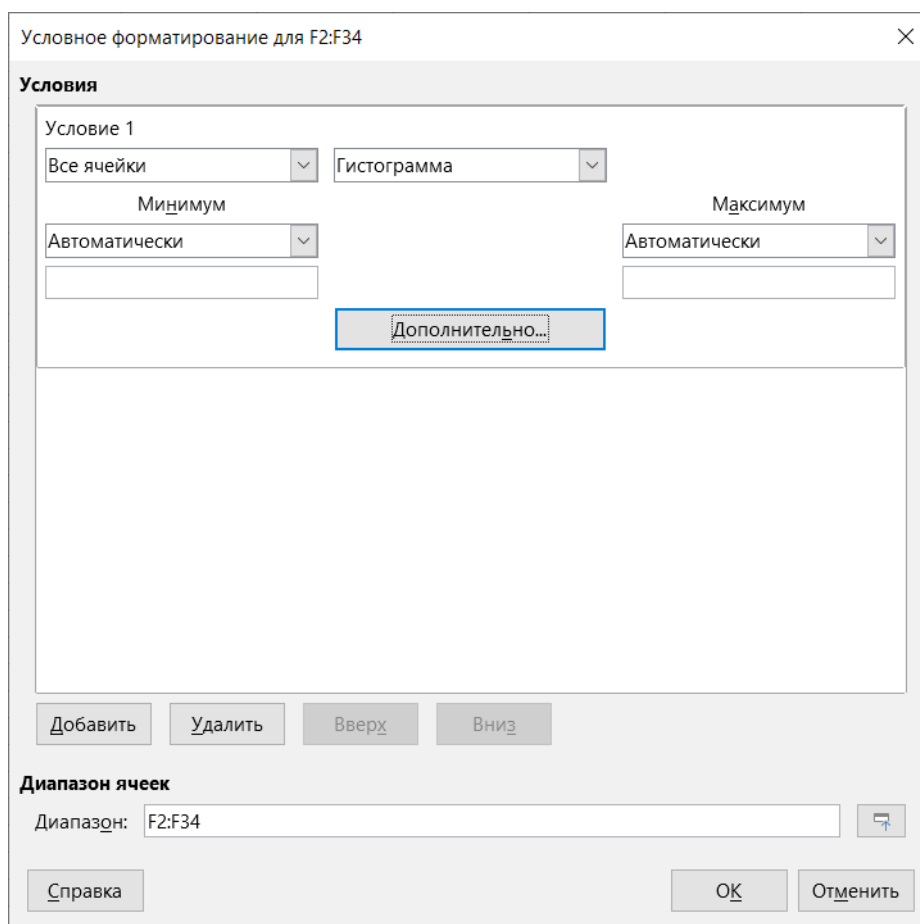
На вкладке ленты «Главная», в секции «Стили» выберите «Условное форматирование» → «Гистограммы» → Градиентная заливка → «Зелёная гистограмма»:



В программном меню выберите «Формат» → «Условное» → «Гистограммы...»:



В появившемся после этого диалоговом окне



можно сразу нажать кнопку «ОК» или (если есть желание) дополнительно настроить параметры гистограммы, вызвав нажатием кнопки «Дополнительно...» ещё одно диалоговое окно:

Гистограмма

Входные значения

Минимум: Автоматически 0

Максимум: Автоматически 0

Цвета

Положительные: Синий

Отрицательные: Красный

Заливка: Градиент

Ось

Положение вертикальной оси: Автоматически

Цвет вертикальной оси: Чёрный

Длина линейки

Минимальная длина (%): 0

Максимальная длина (%): 100

Показать только полосу

ОК Отменить

В результате получится горизонтальная гистограмма, показывающая распределение относительных частот встречаемости букв в анализируемом тексте:

Буква	Количество	Относительная частота встречаемости
а	3	9,375
б	0	0
в	0	0
г	0	0
д	1	3,125
е	5	15,625
ё	0	0
ж	0	0
з	0	0
и	2	6,25
й	0	0
к	1	3,125
л	0	0
м	1	3,125
н	1	3,125
о	2	6,25
п	1	3,125
р	3	9,375
с	1	3,125
т	2	6,25
у	1	3,125
ф	1	3,125
х	1	3,125
ц	0	0
ч	0	0
ш	5	15,625
щ	0	0
ъ	0	0
ы	1	3,125
ь	0	0
э	0	0
ю	0	0
я	0	0

Буква	Количество	Относительная частота встречаемости
а	3	9,375
б	0	0
в	0	0
г	0	0
д	1	3,125
е	5	15,625
ё	0	0
ж	0	0
з	0	0
и	2	6,25
й	0	0
к	1	3,125
л	0	0
м	1	3,125
н	1	3,125
о	2	6,25
п	1	3,125
р	3	9,375
с	1	3,125
т	2	6,25
у	1	3,125
ф	1	3,125
х	1	3,125
ц	0	0
ч	0	0
ш	5	15,625
щ	0	0
ъ	0	0
ы	1	3,125
ь	0	0
э	0	0
ю	0	0
я	0	0

Для поиска аллитераций в стихотворении нужно дополнительно проанализировать какой-либо обычный текст, взяв его в качестве эталона, чтобы полученные распределения частот можно было сравнивать между собой.

У Ломоносова М.В. есть любопытное стихотворение «О сомнительном произношении буквы Г в русском языке». Если честно, то аллитерация там видна, как говорится, невооружённым глазом, просто сам пример весьма показателен:

*«Бугристы берега, благоприятны влаги,
О горы с гроздами, где греет юг ягнят,
О грады, где торги, где мозгокружны браги
И деньги, и гостей, и годы их губят.
Драгие ангелы, пригожие богини,
Бегущие всегда от гадкия гордыни,
Пугливы голуби из мягкого гнезда,
Угодность с негою, огромные чертоги,
Недуги наглые и гнусные остроги,
Богатство, нагота, слуги и господа,
Угрюмы взглядами, игреки, пеги, смуглы,
Багровые глаза, продолговаты, круглы;
И кто горазд гадать и лгать, да не мигать,
Играть, гулять, рыгать и ногти огрызгать,
Ногаи, б_о_лгары, гуроны, геты, гунны,
Тугие головы, о_и_готи чугунны,
Гневливые враги и гладкословный друг,
Толпыги, щеголи, когда вам есть досуг,
От вас совета ждуду, я вам даю на волю:
Скажите, где быть _га_ и где стоять _глаголю_?»**

Относительная частота встречаемости буквы «Г» в данном стихе – 14,53%. Насколько это много? Для сравнения возьмём другой текст Ломоносова, а именно – три первых абзаца седьмой главы его труда «Древняя Российская история»**:

«Самодержавного своего в России государствования Владимир полагает начало, мнимым благочестием по древнему предков многобожию, однако и заблуждением показывает в себе способность к приятию веры в единого истинного Бога. Его повелением поставлен в Киеве перед двором теремным, на высоком холме главный идол Перун, деревянный с серебряною головою и золотым усом. Жертва приносилась - огонь неугасимый. За угашение, небрежением случившееся, жрецы смертной казни предавались. Сей богом грома и молнии почитавшийся Перун был Зевес древних наших предков.

Меньших богов Нестор именует: Хорса, Дажбога, Стрибога, Семаргла, Мокошь, не показав знаменования и приписыванной им от идолопоклонников силы и власти. По Перуне имел Волос первое место, коему покровительство скота приписывалось (рачение о скотопасстве большее, нежели у римлян, нижним божкам оное препоручившим); Погвизд, Похвист или Вихрь - бог ветра, дождя и вёдра, Еол российский; Лада (Венера), Дида и Лель (купидоны), любви и браков покровители, толь усердно от древних предков наших почитались, что оттуда и поныне в любовных простых песнях, особливо на брачных празднествах, упоминаются со многим повторительным восклицанием.

* Текст стихотворения взят с сайта «РусСтих». URL: <https://rustih.ru/mixail-lomonosov-o-somnitelnom-proiznoshenii-bukvy-g-v-rossijskom-yazyke/> (дата обращения: 30.10.2020)

** Для цитирования использован текст, опубликованный на сайте «Библиотекарь.Ру». URL: <http://www.bibliotekar.ru/rus/18.htm> (дата обращения: 30.10.2020)

Купалу, богу плодов земных, соответствующему Цересе и Помоне, праздновали перед началом сенокоса и жатвы в двадцать четвертый день июня. Остатки сего идолопоклонства толь твердо вкоренились, что и поныне почти во всей России ночные игры, особливо скакание около огня, в великом употреблении; и святая Агриппина, которой тогда память празднуется, по древнему идолу проименована от престоноародия Купальницею.»

В приведённом тексте значение частоты для «Г» составляет всего 2,27%. Таким образом, в рассматривавшемся выше стихотворении буква «Г» встречается в 14,53% / 2,27% \approx 6,4 раза чаще обычного.

Строго говоря частотный анализ встречаемости букв пригоден и для других целей – по аналогии с аллитерацией можно исследовать стихи на предмет наличия в них ассонанса (скопления определённых гласных букв), кроме того, такой анализ хорошо подходит для криптографической области – им можно «взламывать» текст, зашифрованный по методу простой замены (то есть когда в зашифрованном тексте каждая буква заменена на какую-то другую).

© Широков Александр, 30.10.2020