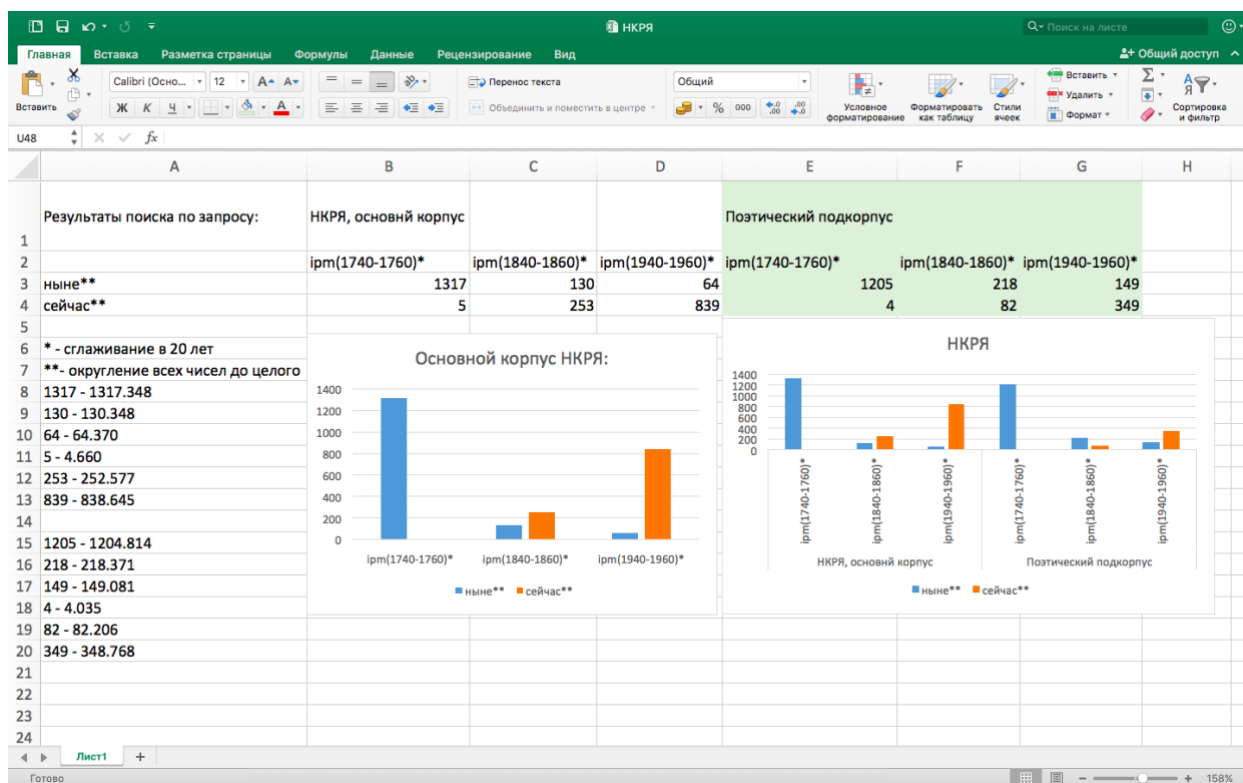


РАБОТА С НАЦИОНАЛЬНЫМ КОРПУСОМ РУССКОГО ЯЗЫКА

Задание 1. Целью моего задания было сравнить употребление таких слов, как «ныне» и «сейчас» в различные периоды, а именно 18, 19 и 20 века. Сравнить их нужно было в основном и поэтическом подкорпусах. Вот что у меня получилось:



Для того, чтобы получить эти результаты я брала период в середине века за 20 лет с сглаживанием в 20 лет, чтобы получить средний результат. Затем, записывала эти результаты и округляла их до целого числа, и уже последние вставляла в таблицу, чтобы было возможно построить диаграмму и увидеть все более наглядно. Исходя из результатов, мы можем сказать, что, как и ожидалось, в XVII веке наиболее популярной оказалась словоформа «ныне» как и в основном, так и в поэтическом подкорпусах. А употребление словоформы «сейчас» настолько мало, что значение не виднеется ни на одной из диаграмм. А вот в XIX веке наблюдаются уже существенные различия между основным и поэтическим подкорпусами, а именно: в основном корпусе словоформа «сейчас» резко приобретает популярность и преобладает над «ныне», а в поэтическом подкорпусе мы видим обратную картину, несмотря на то, что словоформа «сейчас» приобретает популярность и виднеется на диаграмме, словоформа «ныне» также становится более популярной и также преобладает над другой словоформой, но уже не с таким преимуществом. В XX веке «сейчас» уже занимает главенствующую роль, как в основном, так и в поэтическом подкорпусах, однако в поэтическом оно не настолько популярно, как в основном, словоформа «ныне» очень медленно, не хотя, теряет свою популярность в поэтическом подкорпусе, это наглядно можно увидеть как и в таблице, так и на диаграмме.

Таким образом, мы можем сделать вывод, что употребление различных слов в разговорной речи оказывает прямое влияние на употребление слов в поэтической.

Задание 2. Мне были даны несколько диалогов и мне нужно было определить, могли ли они употребляться во времена правления Екатерины II. Мы знаем, что годы правления Екатерины 1762 – 1796. Мы можем взять какие-либо слова или словосочетания из данных диалогов и посмотреть, употреблялись ли они в выше указанный промежуток. Посмотрим на первый диалог:

– *Отдай !*

– *Пиши пропало . Что пряжка , третьего дня эта поганка у меня с груди звезду Александра Невского уперла ! Любит , сволочь , блестящее .*

В нем я выделила такое сочетание, как «пиши пропало» и выяснило, что оно не употреблялось во времена Екатерины II, так как впервые начало употребляться в начале XIX века, следовательно, данный диалог не свойственен эпохе правления Екатерины II. Обратим свое внимание на следующий диалог:

– *Ох , ну просто немыслимо ! Нет , как можно быть таким лопухом ? !*

Из этого диалога я взяла слово «немыслимо» и оно опять же не свойственно 1762-1796 годам, так как впервые начало употребляться ближе к середине XIX века.

И последний диалог, который был нам дан:

– *Я гнева вашего никак не растолкую . Он в доме здесь живет , великая напасть !*

В данном диалоге я обратила внимание на слово «напасть», так как в наши дни оно редко употребляется и ему находят замены такие как «беда», «неприятность» и другие. И, действительно, это слово очень активно употреблялось во времена Екатерины II и именно наибольшая частота употребления этого слова была замечена как раз в XVIII веке.

Задание 3.

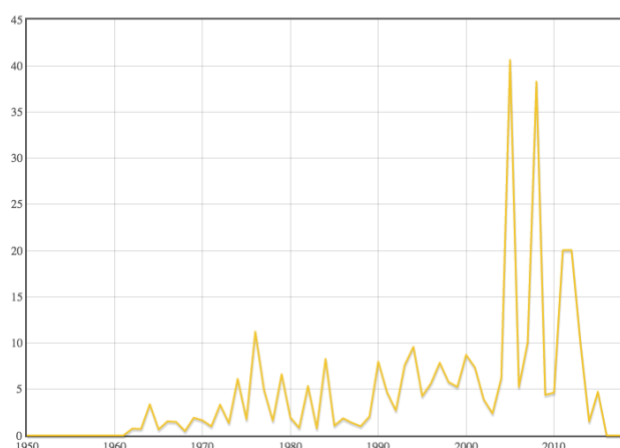
Распределение по годам (частота на миллион словоформ)

Слова или сочетания слов, через запятую: щас

Годы с 1950 по 2018

со сглаживанием 0

построить



Выше Вы видите график словоформы «щас» с 1950 года по наши дни, результаты нам говорят о том, что данная словоформа стала употребляться в 1961, а сейчас уже не употребляется, начиная с 2016, но была особенно популярна с 2003 по 2010-е.

Также мне нужно было выбрать два различных жанра и сравнить в них употребление этой же словоформы. Первым жанром, который я выбрала, оказался детектив, боевик, следующим – любовная история.

График (детектив, боевик):

Распределение по годам (частота на миллион словоформ)

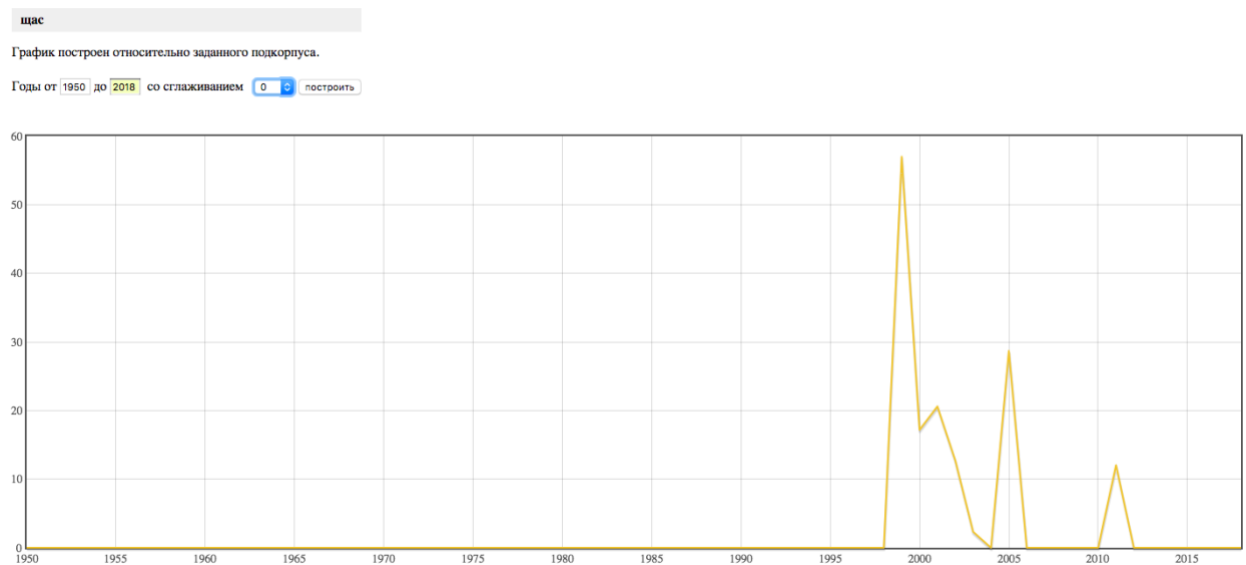
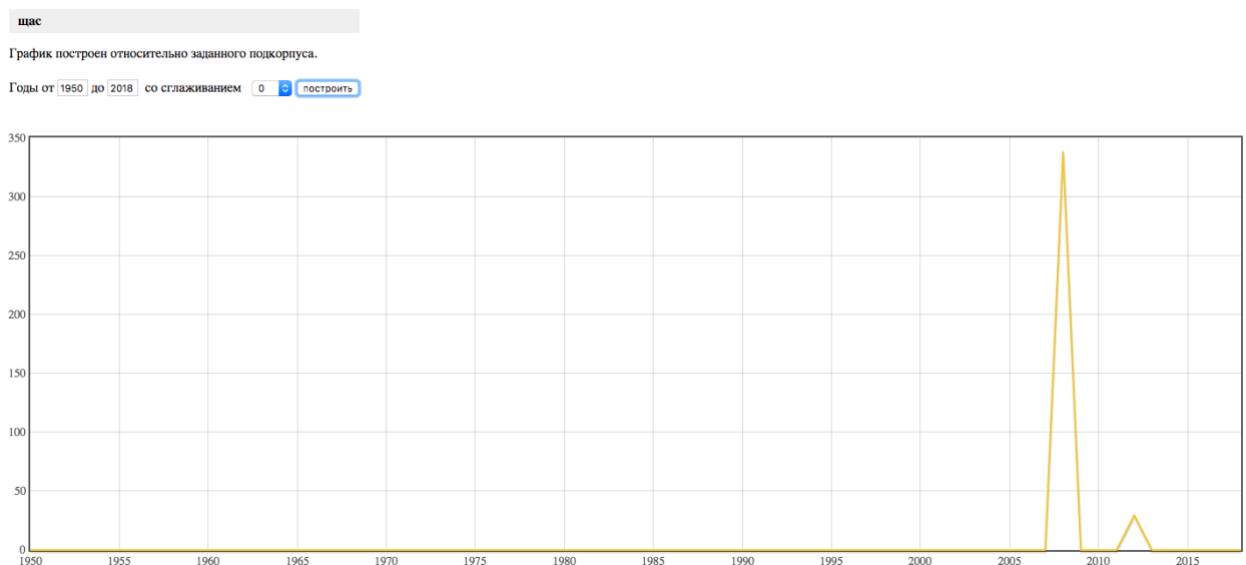


График (любовная история):

Распределение по годам (частота на миллион словоформ)



Исходя из данных результатов мы можем сказать, что словоформа «щас» в детективах и боевиках была более популярна, нежели чем в любовных историях. И она была популярная в различные периоды, например, в детективах и боевиках с 1998-2006 и 2010-2012. А вот в любовных историях с 2007 по 2009 и с 2011 по 2013.

Задание 4. После скачивания выдачи в формате Excel, я отсортировала строки по столбцу F по алфавиту, где были существительные в родительном падеже. Там я выделила слова, которые встречаются более 5 раз, такие как(слова будут в родительном падеже): автора, архитектуры, бизнеса, власти, вражды, имени, искусства, истории, колокола, культуры, литературы, любви, математики, мира, мышления, народа, науки, начала, общения, описания, оригинала, поэзии, программирования, происхождения, религии, романа, современности, страны, судопроизводства, теории, улицы, физики, философии, человека.