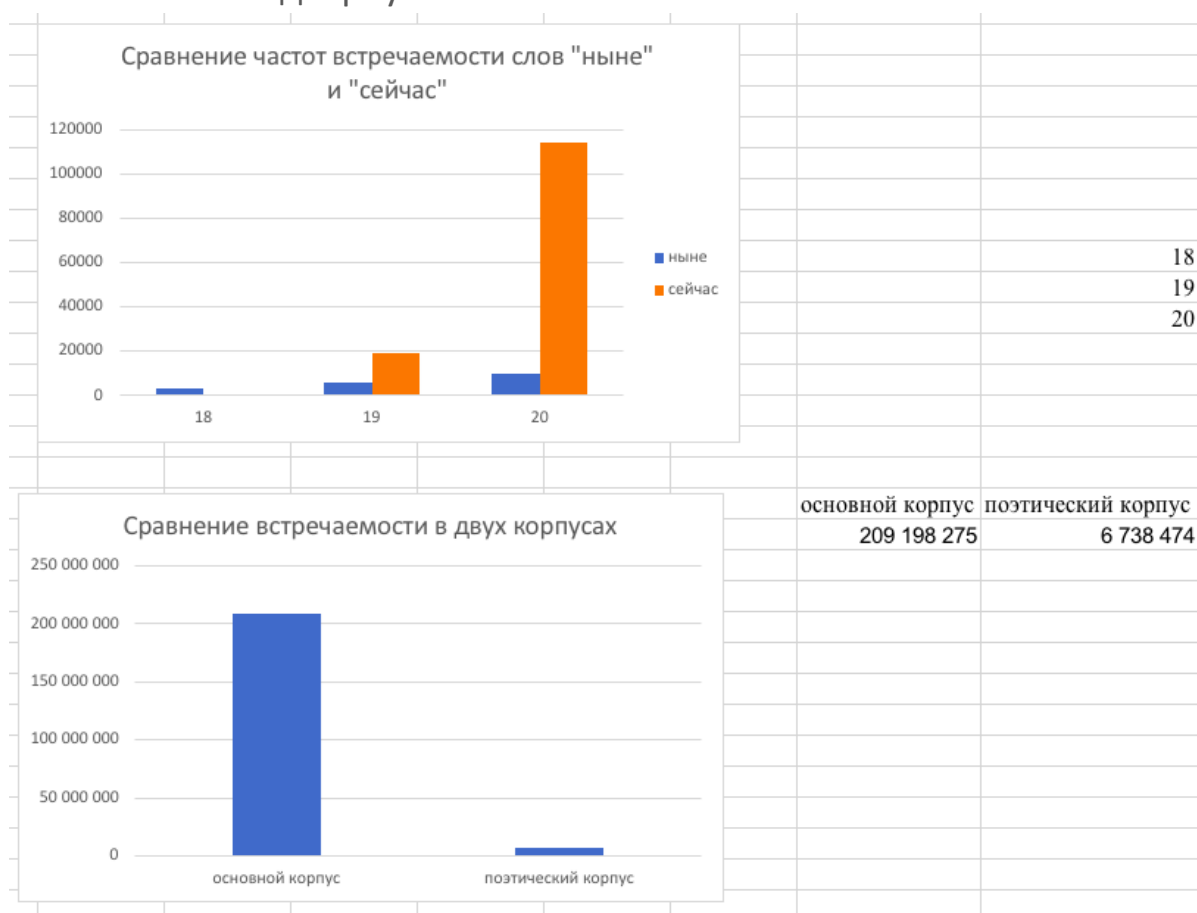


Отчет (работа с корпусами)

- Первым шагом я сравнила частоту употреблений таких слов, как «ныне» и «сейчас» в разное время (18 – 20 века). На основе данных Национального корпуса русского языка, я построила столбчатые диаграммы, которые позволили проследить частотность употребления данных слов в основном и поэтическом подкорпусах.



- Исходя из полученных данных, мы можем сделать вывод: частотность встречаемость слов «ныне» и «сейчас» гораздо больше в основном корпусе, нежели в поэтическом. Основной корпус (209 198 275) и поэтический корпус (6 738 474).
- Второе задание заключалось в поиске слов из данных нам диалогов и выяснении того, употреблялись ли эти слова в годы правления Екатерины Второй (1762 – 1796).

- По результатам исследования слов из первого диалога, можно сделать вывод, что данные диалоги не могли встретиться в таком виде во времена правления Екатерины Второй.

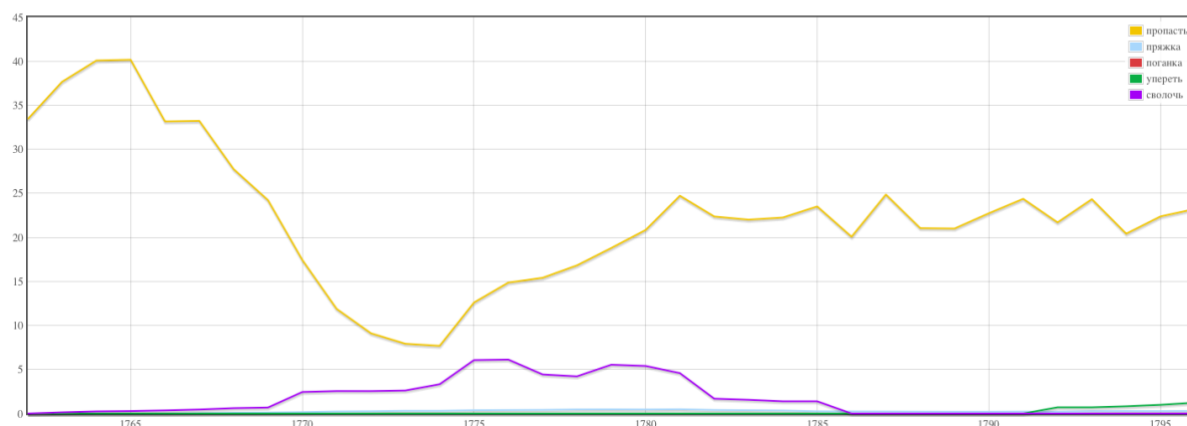
1) График употребления слов из первого диалога:

Распределение по годам (частота на миллион словоформ)

Слова или сочетания слов, через запятую:

Годы с 1762 по 1796 со сглаживанием 3

[построить](#)



Если построить таблицу, то мы поймем, что слово «поганка», например, начало употребляться только с 1839 года. Поэтому диалог не мог существовать. Примеры:

<u>поганка</u>	
<u>2015</u>	1
<u>2012</u>	1
<u>2011</u>	1
<u>2010</u>	1
<u>2008</u>	1
<u>2007</u>	1
<u>2005</u>	2
<u>2004</u>	2
<u>2003</u>	6

2001	1
2000	2
1998	1
1997	3
1996	1
1995	2
1994-2003	1
1992	4
1991	3
1990-2002	2
1987	2
1986	1
1984	1
1982-1985	1
1975-1987	2
1972	1
1971	2
1970-1992	1
1969-1970	1
1969	2
1967	6
1966	2

1965-1994	3
1965-1982	3
1964	1
1956	1
1948	1
1942-1952	1
1938-1947	2
1930-1950	1
1928-1940	5
1928-1931	9
1928	1
1924-1932	1
1912	1
1911	1
1909	2
1903	1
1899	2
1898	1
1894	1
1892	3
1869-1872	2
1866	1

1857-1863	1
1847	1
1839	1

По результатам проверки слов из второго диалога, удалось выяснить, что слова из данного текста начали использоваться только с 1834 года. Следовательно, второй диалог также не может существовать.

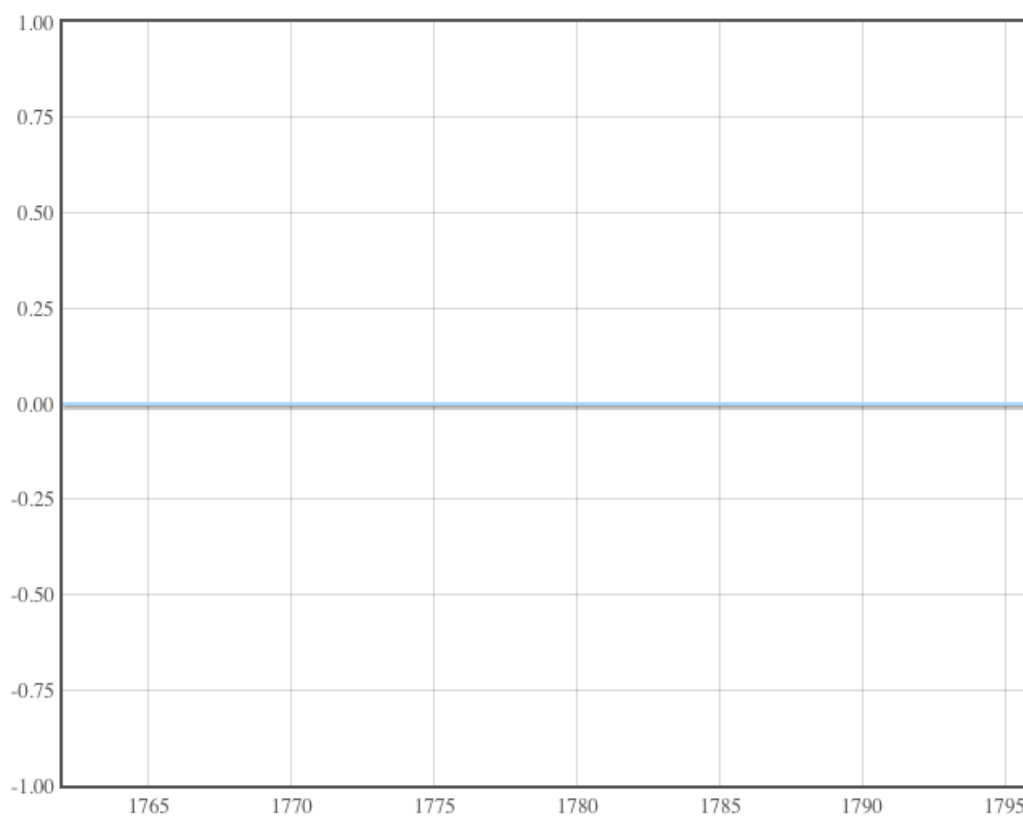
2) График употребления слов из второго диалога.

Распределение по годам (частота на миллион словоформ)

Слова или сочетания слов, через запятую:

Годы с по со сглаживанием

[построить](#)

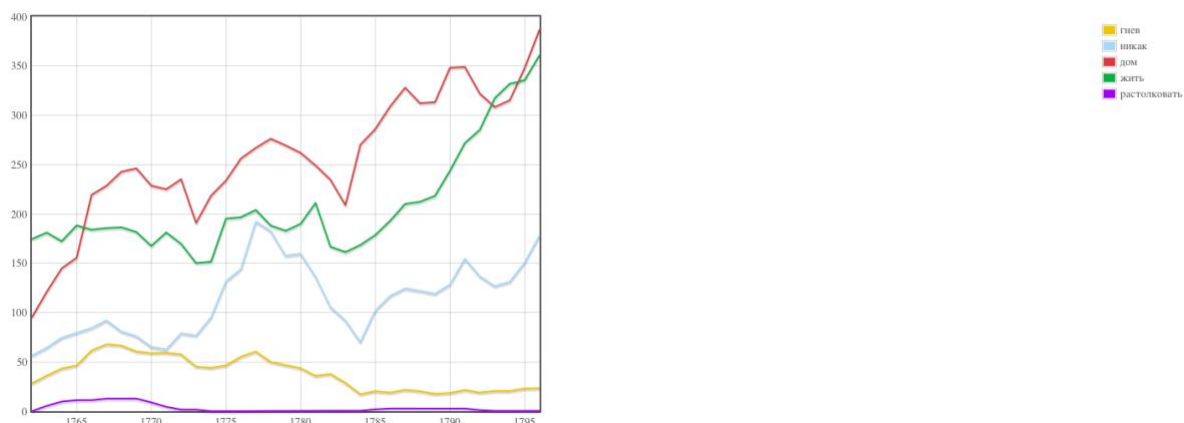


3) Заключительный анализ слов третьего диалога оказался успешен. Мне удалось определить слова, которые использовались в этот период:

Распределение по годам (частота на миллион словоформ)

Слова или сочетания слов, через запятую:

Годы с по со сглаживанием



Вывод: я закончила поиск слов из данных диалогов. Результат показал, что третий диалог мог существовать во времена Екатерины Второй.

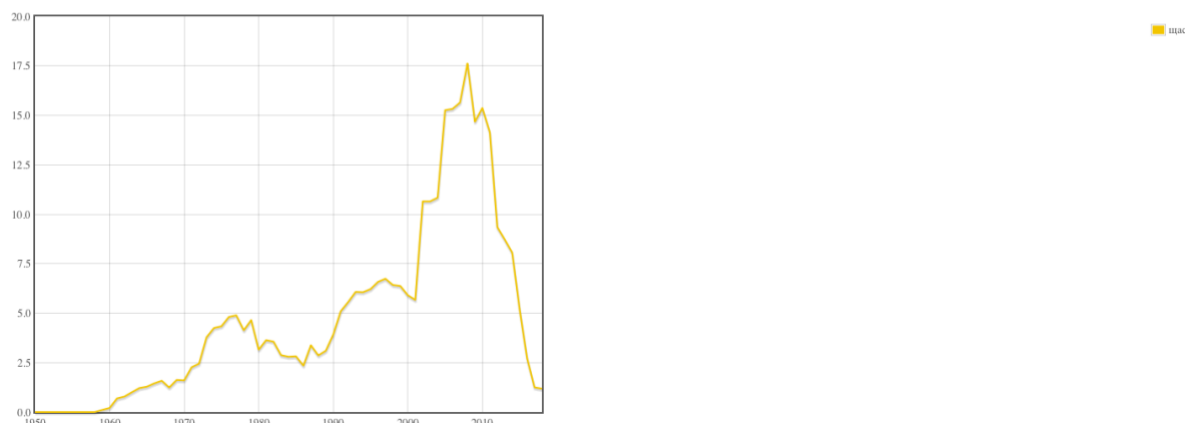
Бонус.

1. Для начала я посмотрела динамику употребления слова «щас» в основном корпусе с 1950 года по наши дни и построила следующий график:

Распределение по годам (частота на миллион словоформ)

Слова или сочетания слов, через запятую:

Годы с по со сглаживанием



Исходя из графика, можно увидеть, что пик употребления слова «щас» приходится на 2011 – 2015 года:

Всего:	
2015	845230.023810
2014	1316853.571429
2013	1798549.506349
2012	2351268.275397
2011	1716153.745094

2. Далее я сравнила употребление слова «щас» в текстах разных жанров. И получила следующие данные:

Распределение по годам (частота на миллион словоформ)

щас

График построен относительно заданного подкорпуса.

Годы от 1950 до 2018 со сглаживанием 3 построить



Жанр [свернуть](#)

Значение	Найдено документов	Найдено словоформ
нежанровая проза	122	279 (54.17%)
детектив, боевик	18	84 (16.31%)
юмор и сатира	16	41 (7.96%)
драматургия	14	36 (6.99%)
фантастика	15	36 (6.99%)
фантастика приключения	2	14 (2.72%)
детская	7	8 (1.55%)
нежанровая проза любовная история	1	4 (0.78%)
документальная проза	2	3 (0.58%)
фантастика юмор и сатира	2	2 (0.39%)
любовная история	1	2 (0.39%)
фантастика детская	2	2 (0.39%)
историческая проза	2	2 (0.39%)
фантастика детектив, боевик	1	1 (0.19%)
приключения	1	1 (0.19%)

Если изучить статистику, то можно понять, что слово «щас» чаще всего употреблялось в нежанровой прозе (документов – 122, словоформ – 279), детективах и боевиках (документов – 18, словоформ – 84), юмор и сатира (документов – 16, словоформ – 41).

Тип текста [свернуть](#)

Значение	Найдено документов	Найдено словоформ
роман	77	239 (46.41%)
повесть	45	111 (21.55%)
рассказ	46	75 (14.56%)
пьеса	14	36 (6.99%)
миниатюра	3	27 (5.24%)
анекдот	11	12 (2.33%)
повесть цикл	1	6 (1.17%)
рассказ цикл	3	3 (0.58%)
миниатюра анекдот	2	2 (0.39%)
записки	1	1 (0.19%)
очерк цикл	1	1 (0.19%)
повесть рассказ	1	1 (0.19%)
сказка	1	1 (0.19%)

Далее я просмотрела статистику типов текстов, в которых употреблялось слово «щас». С помощью этого я выяснила, что данное слово встречалось в основном в романах (документы – 77, словоформы – 239), повестях (документы – 45, словоформы – 111) и рассказах (документы 46, словоформы – 75).

Тематика текста [свернуть](#)

Значение	Найдено документов	Найдено словоформ
(нет)	203	510 (99.03%)
политика и общественная жизнь	2	4 (0.78%)
армия и вооруженные конфликты	1	1 (0.19%)

Результат, связанный с тематикой текста показал, что слово «щас» встретилось в текстах без значения (документов – 203, словоформ – 510).