

Выводы:

- **Все региональные выступления** существенно короче по количеству слов.
- **Все региональные выступления** имеют долю уникальных слов >30%, выступление Путина – 21%.
- **Доля** разных частей речи примерно одинакова во всех стенограммах.
- **Во всех стенограммах** помимо очевидных словосочетаний вроде 'прошлый год' и 'последний год' также встречались словосочетания: *'детский сад', 'новые технологии', 'социальная сфера' и 'повышение качества'*.
- **Послание Путина** отличается упоминанием различных стран: 15 раз упоминались США и Америка, 1 раз - Крым, 4 раза - Европа, 85 раз – Россия и 5 раз - Советский Союз. В нем часто упоминается стратегия, стратегические планы делаются акценты на вооружение. Количество упоминаний этих тем сильно перевешивает, например, упоминания уровня жизни и экономического роста.
- **Рязанское послание** отличают акценты на инвестициях, строительстве и производстве. Пять раз упоминаются 'группы компаний', четыре раза – 'бережливое производство', три раза – 'инвестиционный проект'.
Упоминание 2018ого года встречается в 7.25 раза чаще, чем в послании Путина. Слова: 'участие', 'объект', 'район', 'муниципальный', 'дом' – встречаются более чем в 5 раз чаще, чем у Путина.
Слова: 'цифровой', 'являться', 'производство' – встречаются более, чем в 7 раз чаще, чем в тульском послании.
В 29 и 12 раз соответственно чаще встречается упоминание 2018 и 2017 года, чем в тюменском послании. Слова: 'завод' и 'современный' – встречаются более, чем в 7 раз чаще. Слово 'цифровой' встречается в 8 раз чаще в этом послании, чем в ямальском.
- **Тульское послание** отличают упоминания 'Национального проекта' - 4 раза, 'Указа президента' - 4 раза, а также 'Пожилых людей' - 3 раза и 'Местных дорог' - 3 раза. Слова: 'искать', 'оценка', 'бюджет' - встречаются более чем в 7 раз чаще, чем у Путина. Слова: 'бюджет', 'мера', 'финансирование' – встречаются более, чем в 10 раз чаще, чем в рязанском послании.
В 16 раз чаще встречается упоминание 2018, чем в тюменском послании. Слова: 'рубль' и 'работа', 'ребенок', 'ремонт' – встречаются более, чем в 10 раз чаще.
Слова: 'медицинский', 'ремонт', 'городской' - встречаются в более, чем в 7 раз чаще в этом послании, чем в ямальском.
- **Тюменское послание** отличается от остальных упоминаниями 'Качества жизни' - 6 раз, 'Национального проекта' - 5 раз и 'Производительности труда' – тоже 5 раз. Слова: 'округ', 'справедливый' - встречаются более чем в 5 раз чаще, чем у Путина. Слово 'экономика' встречается в 13 раз чаще, чем в рязанском послании. Обращение к 'друзьям' происходит в 7 раз чаще, чем в тульском послании.

Слова: 'сделать', 'должный', 'медицинский', 'национальный' - встречаются в более, чем в 7 раз чаще в этом послании, чем в ямальском.

- В стенограмме **ямальского обращения** встречается несколько интересных уникальных словосочетаний: 'Северный морской путь' - 6 раз, 'Арктическая зона' - 6 раз, 'Автономный округ' - 6 раз.
Также упоминается: 'Аварийное жилье' - 5 раз, 'Коренные народы' - 5 раз, 'Муниципальное образование' - 4 раза.
Слова 'благодарить', 'глава', 'депутат', 'арктический', 'округ' встречаются более чем в 10 раз чаще, чем у Путина.
Частое употребление слов 'сохранение', 'семья' отличают это послание от рязанского.
От тульского – 'северный', 'благодарить' и 'оставаться'.
А от тюменского – 'ребенок' и 'жилье'.

Идеи:

Инфографика о продолжительности выступления и количеству слов в нем. Как соотносятся эти величины.

Анализ аплодисментов. Их продолжительность. Найти самый долгие и изучить их особенности.

В моем анализе присутствуют пункты о словах и словосочетаниях, отличающих одни послания от других. Это в некотором роде информация, аналогичная той, что есть в исследовании яндекса о рэпе (отличительные слова жанров). Будет куда нагляднее, если к имеющимся стенограммам добавить записи выступлений прошлых лет, других политиков.

Взять аналогичные послания других лет и сравнить изменения: новые слова в топе используемых, продолжительность, количество слов, индекс удобочитаемости, аплодисменты. Можно добавить к сравнению и другие выступления: лидеров других стран, знаковые выступления глав компаний и комиков.

Сухая статистика:

Путин:

- Количество слов: 11252
Уникальных слов: 2350, 21% от всех
Индекс удобочитаемости: 29.1
- 521 прилагательное, 22.2% от всех слов
909 существительных, 38.7% от всех слов
575 глаголов, 24.5% от всех слов
145 наречий, 6.2% от всех слов
- 15 раз упоминались США и Америка, 1 раз Крым, 4 раза Европа и 85 раз Россия, 5 раз Советский Союз
- Популярные в речи словосочетания:

‘Уважаемый коллега’ - 16 раз, ‘Шесть лет’ - 12 раз, ‘Наша страна’ - 9 раз, ‘Стратегическое оружие’ - 9 раз, ‘Ракетный комплекс’ - 7 раз, ‘Качество жизни’ - 6 раз, ‘Экономический рост’ - 6 раз, ‘Вооруженная сила’ - 6 раз, ‘Советский союз’ - 5 раз, ‘Крылатая ракета’ - 5 раз

- Самые популярные слова и количество их повторений:

это - 166, год - 97, весь - 95, наш - 91, Россия - 85, который - 73, новый - 69, нужно - 68, должный - 61, система - 56, страна - 55, человек - 52, развитие - 47, свой - 43, современный - 35, оружие - 35, сегодня - 33, обеспечивать - 31, сделать - 31, мир - 29, возможность - 29, миллион - 28, важный - 27, уровень - 27, сказать - 27, задача - 26, создавать - 26, рост - 26, стратегический - 26, ядерный - 26, процент - 25, гражданин - 24, также - 24, жизнь - 24, позволять - 23, сила - 23, мочь - 22, становиться - 22, комплекс - 22, время - 21

- Самые популярные прилагательные:

новый - 69, должный - 61, современный - 35, важный - 27, стратегический - 26, ядерный - 26, российский - 21, большой - 20, уважаемый - 19, высокий - 19, экономический - 17, военный - 15, социальный - 14, технологический - 14, прошлый - 12

- Самые популярные существительные:

год - 97, Россия - 85, система - 56, страна - 55, человек - 52, развитие - 47, оружие - 35, мир - 29, возможность - 29, миллион - 28, уровень - 27, задача - 26, рост - 26, процент - 25, гражданин - 24

- Самые популярные глаголы:

обеспечивать - 31, сделать - 31, мочь - 31, сказать - 27, создавать - 26, позволять - 23, становиться - 22, хотеть - 21, говорить - 20, работать - 19, идти - 18, иметь - 17, решать - 14, знать - 14, получать - 13

- Самые популярные наречия:

сегодня - 33, также - 24, очень - 20, практически - 14, вместе - 11, далеко - 10

Рязань:

- Количество слов: 3687

Уникальных слов: 1153, 31% от всех

Индекс удобочитаемости: 27.8

- 274 прилагательное, 23.8% от всех слов

493 существительных, 42.8% от всех слов

239 глаголов, 20.7% от всех слов

51 наречий, 4.4% от всех слов

- Популярные в речи словосочетания:

Россия и Российская Федерация упоминались 13 раз.

‘Президент РФ’ - 7 раз, ‘Группа компаний’ - 5 раз, ‘Областной центр’ - 5 раз, ‘Развитие региона’ - 4 раза, ‘Бережливое производство’ - 4 раза, ‘Профессиональная команда’ - 3 раза, ‘Социальная сфера’ - 3 раза, ‘Инвестиционный проект’ - 3 раза, ‘Президент РФ Путин’ - 4 раза

- Самые популярные слова и количество их повторений:

год - 59, регион - 36, рязанский - 30, область - 30, проект - 29, 2018 - 29, это - 28, новый - 25, район - 21, необходимо - 21, человек - 20, наш - 19, развитие - 18, который - 18, работа - 17, место - 16, создание - 16, строительство - 16, поддержка - 16, должный - 15, весь - 14, число - 14, производство - 14, рязань - 14, современный - 13, 2017 - 12, компания - 12, центр - 12, направление - 11, задача - 11, объект - 11, позволять - 11, областной - 11, сделать - 10, социальный - 10, становиться - 10, общественный - 10, качество - 10, реализация - 10, федеральный - 10

- Самые популярные прилагательные:

рязанский - 30, новый - 25, должный - 15, современный - 13, областной - 11, социальный - 10, общественный - 10, федеральный - 10, хороший - 8, цифровой - 8, важный - 7, общий - 7, инвестиционный - 7, основной - 6, перспективный - 6, бережливый - 6

- Самые популярные существительные:

год - 59, регион - 36, область - 30, проект - 29, район - 21, человек - 20, развитие - 18, работа - 17, место - 16, создание - 16, строительство - 16, поддержка - 16, число - 14, производство - 14, компания - 12,

- Самые популярные глаголы:

позволять - 11, сделать - 10, становиться - 10, обеспечивать - 8, построить - 8, являться - 8, увеличивать - 7, продолжать - 7, создавать - 7, получать - 7, начинать - 6, хотеть - 5, принимать - 5, ставить - 5, оставаться - 5, определять - 4

- Самые популярные наречия:

наиболее - 7, также - 6, максимально - 5, мало - 5, социально - 4, активно - 4

- Сравнение с остальными посланиями:

Упоминание 2018 года встречается в 7.25 раза чаще, чем в послании Путина. Слова: ‘участие’, ‘объект’, ‘район’, ‘муниципальный’, ‘дом’ – встречаются более чем в 5 раз чаще, чем у Путина.

Слова: ‘цифровой’, ‘являться’, ‘производство’ – встречаются более, чем в 7 раз чаще, чем в тульском послании.

В 29 и 12 раз соответственно чаще встречается упоминание 2018 и 2017 года, чем в тюменском послании. Слова: ‘завод’ и ‘современный’ – встречаются более, чем в 7 раз чаще.

Слово ‘цифровой’ встречается в 8 раз чаще в этом послании, чем в ямальском.

Тула

- Количество слов: 4414

Уникальных слов: 1343, 30% от всех

Индекс удобочитаемости: 30.9

- 312 прилагательное, 23.2% от всех слов

554 существительных, 41.2% от всех слов

288 глаголов, 21.4% от всех слов

68 наречий, 5% от всех слов

- Популярные в речи словосочетания:

9 раз упоминалась Россия, Тульская область – 15 раз

‘В первую очередь’ - 5 раз, ‘Национальный проект’ - 4 раза, ‘Указ президента’ - 4 раза, ‘Детский сад’ - 3 раза, ‘Пожилой человек’ - 3 раза, ‘Местная дорога’ - 3 раза

- Самые популярные слова и количество их повторений:

год - 79, это - 40, область - 35, наш - 35, работа - 34, проект - 32, 2019 - 30, бюджет - 29, весь - 28, новый - 28, рубль - 28, который - 23, развитие - 22, регион - 22, тульский - 21, региональный - 19, становиться - 19, гражданин - 18, млн - 17, 2018 - 16, социальный - 16, число - 16, задача - 15, человек - 14, система - 14, Тула - 14, житель - 13, результат - 13, программа - 13, центр - 13, свой - 13, должный - 13, проблема - 13, правительство - 12, рост - 12, уровень - 12, давать - 12, первый - 12, президент - 11, качество - 11

- Самые популярные прилагательные:

новый - 28, тульский - 21, региональный - 19, социальный - 16, должный - 13, федеральный - 9, детский - 9, медицинский - 9, большой - 8, муниципальный - 8, областной - 7, городской - 7, единый - 6, полный - 6, национальный - 6, государственный - 6

- Самые популярные существительные:

год - 79, область - 35, работа - 34, проект - 32, бюджет - 29, рубль - 28, развитие - 22, регион - 22, гражданин - 18, млн - 17, число - 16, задача - 15, человек - 14, система - 14, житель - 13

- Самые популярные глаголы:

становиться - 19, давать - 12, начинать - 10, создавать - 10, расти - 9, использовать - 8, принимать - 7, объем - 7, искать - 7, привлекать - 7, предусматривать - 6, хотеть - 6, выполнять - 6, развивать - 6, работать - 6, мочь - 6

- Самые популярные наречия:

сегодня - 9, активно - 6, менее - 4, достойно - 3, вдвое – 3

- Сравнение с остальными посланиями:

Слова: ‘искать’, ‘оценка’, ‘бюджет’ - встречаются более чем в 7 раз чаще, чем у Путина.

Слова: ‘бюджет’, ‘мера’, ‘финансирование’ – встречаются более, чем в 10 раз чаще, чем в рязанском послании.

В 16 раз чаще встречается упоминание 2018, чем в тюменском послании. Слова: 'рубль' и 'работа', 'ребенок', 'ремонт' – встречаются более, чем в 10 раз чаще.

Слова: 'медицинский', 'ремонт', 'городской' - встречаются в более, чем в 7 раз чаще в этом послании, чем в ямальском.

Тюмень

- Количество слов: 3314

Уникальных слов: 1177, 35% от всех

Индекс удобочитаемости: 30.2

- 279 прилагательных, 23.7% от всех слов
486 существительных, 41.3% от всех слов
264 глагола, 22.4% от всех слов
84 наречия 7% от всех слов
- Популярные в речи словосочетания:

11 раз упоминалась Россия, Тюменская область - 16 раз.

'Качество жизни' - 6 раз, 'Национальный проект' - 5 раз, 'Производительность труда' - 5 раз.

- Самые популярные слова и количество их повторений:

весь - 32, наш - 32, тюменский - 27, новый - 24, область - 24, это - 24, который - 23, должный - 22, год - 19, проект - 18, становится - 17, развитие - 17, друг - 14, экономика - 13, каждый - 12, человек - 12, социальный - 12, гражданин - 10, Россия - 10, свой - 10, качество - 10, жизнь - 10, задача - 10, общество - 9, уровень - 9, — - 9, равенство - 9, сделать - 8, возможность - 8, поддержка - 8, продолжать - 8, национальный - 7, говорить - 7, рост - 7, создавать - 7, общественный - 7, регион - 7, труд - 7, инициатива - 7, медицинский - 7, внимание - 7, главный - 7,

- Самые популярные прилагательные:

тюменский - 27, новый - 24, должный - 22, социальный - 12, национальный - 7, общественный - 7, медицинский - 7, главный - 7, региональный - 6, хороший - 6, крупный - 6, близкий - 6, справедливый - 6, равный - 6, малый - 6, технологический - 5

- Самые популярные существительные:

область - 24, год - 19, проект - 18, развитие - 17, друг - 14, экономика - 13, человек - 12, гражданин - 10, Россия - 10, качество - 10, жизнь - 10, задача - 10, общество - 9, уровень - 9, равенство - 9

- Самые популярные глаголы:

становиться - 17, сделать - 8, продолжать - 8, говорить - 7, создавать - 7, работать - 6, обеспечивать - 6, давать - 5, думать - 5, оставаться - 5, видеть - 4, являться - 4, развивать - 4, показывать - 3, ставить - 3, находиться - 3, объем - 3, капитал - 3, запускать - 3, построить - 3

- Самые популярные наречия:

прежде - 5, научно - 4, мало - 3, активно - 3, действительно – 3

- Сравнение с остальными посланиями:

Слова: 'округ', 'справедливый' - встречаются более чем в 5 раз чаще, чем у Путина.

Слово 'экономика' встречается в 13 раз чаще, чем в рязанском послании.

Обращение к 'друзьям' происходит в 7 раз чаще, чем в тульском послании.

Слова: 'сделать', 'должный', 'медицинский', 'национальный' - встречаются в более, чем в 7 раз чаще в этом послании, чем в ямальском.

Ямал

- Количество слов: 5326

Уникальных слов: 1600, 30% от всех

Индекс удобочитаемости: 31.9

- 345 прилагательных, 21.6% от всех слов

695 существительных, 43.4% от всех слов

347 глаголов, 21.7% от всех слов

88 наречий, 5.5% от всех слов

- Популярные в речи словосочетания:

Россия упоминалась 26 раз.

'Дорогие ямальцы' - 7 раз, 'Уважаемые депутаты' - 7 раз, 'Северный морской путь' - 6 раз, 'Арктическая зона' - 6 раз, 'Автономный округ' - 6 раз, 'Глава государства' - 6 раз, 'Аварийное жилье' - 5 раз, 'Коренные народы' - 5 раз, 'Детский сад' - 4 раза, 'Муниципальное образование' - 4 раза

- Самые популярные слова и количество их повторений:

год - 96, это - 48, округ - 37, Ямал - 36, регион - 31, наш - 30, весь - 30, развитие - 26, новый - 26, Россия - 26, работа - 26, проект - 23, становиться - 21, ямалец - 18, программа - 17, строительство - 17, число - 16, социальный - 16, принимать - 15, поддержка - 15, страна - 14, арктический - 14, жилье - 14, территория - 14, житель - 14, который - 14, оставаться - 14, объект - 14, сегодня - 13, рубль - 13, предприятие - 13, следующий - 13, возможность - 13, сфера - 13, свой - 13, задача - 13, ребенок - 13, депутат - 12, внимание - 12, человек - 12, правительство - 12, проводить - 12

- Самые популярные прилагательные:

новый - 26, социальный - 16, арктический - 14, следующий - 13, северный - 11, большой - 11, уважаемый - 10, федеральный - 10, текущий - 9, международный - 9, современный - 9, коренной - 9, морской - 8, крупный - 8, дорогой - 7, последний - 7

- Самые популярные существительные:

год - 96, это - 48, регион - 31, округ - 37, развитие - 26, Россия - 26, работа - 26, проект - 23, ямалец - 18, программа - 17, строительство - 17, число - 16, поддержка - 15, страна - 14, жилье - 14, территория - 14, житель - 14

- Самые популярные глаголы:

Ямал - 36, становиться - 21, принимать - 15, оставаться - 14, проводить - 12, продолжать - 11, благодарить - 10, поручать - 9, вырастать - 8, счет - 8, планировать - 8, составлять - 7, создавать - 7, давать - 7, работать - 7, начинать - 7, вестись - 7, удаваться - 7, построить - 6, являться - 6,

- Самые популярные наречия:

сегодня - 13, свыше - 8, значительно - 5, совместно – 5

- Сравнение с остальными посланиями:

Слова 'благодарить', 'глава', 'депутат', 'арктический', 'округ' встречаются более чем в 10 раз чаще, чем у Путина.

Частое употребление слов 'сохранение', 'семья' отличают это послание от рязанского.

От тульского – 'северный', 'благодарить' и 'оставаться'.

А от тюменского – 'ребенок' и 'жилье'.