Диснеевские принцессы

Соня Акулова, Яна Лабенская, Кошкина Ксения

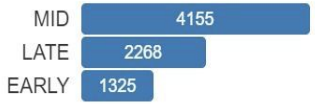
Общее

Disney_Period

Categorical

HIGH_CORRELATION

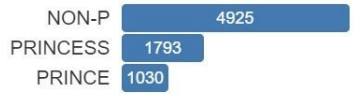
Distinct count	3
Unique (%)	< 0.1%
Missing	0
Missing (%)	0.0%
Memory size	60.5 KiB



Speaker_Status

Categorical

Distinct count	3
Unique (%)	< 0.1%
Missing	0
Missing (%)	0.0%
Memory size	60.5 KiB



Movie

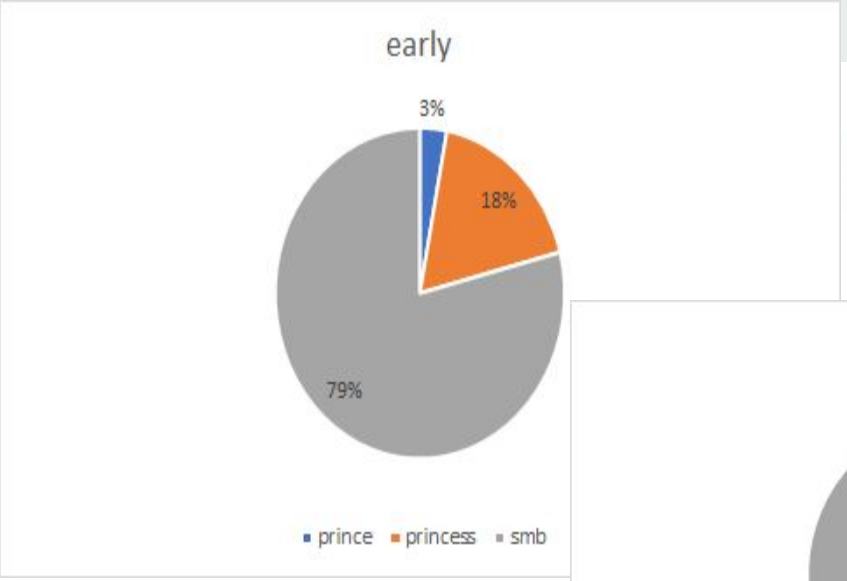
Categorical

HIGH CORRELATION

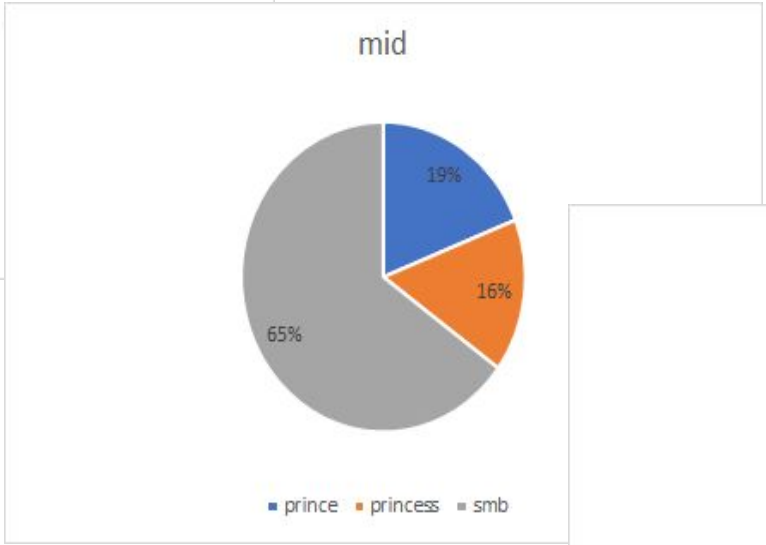
Distinct count	13
Unique (%)	0.2%
Missing	0
Missing (%)	0.0%
Memory size	60.5 KiB

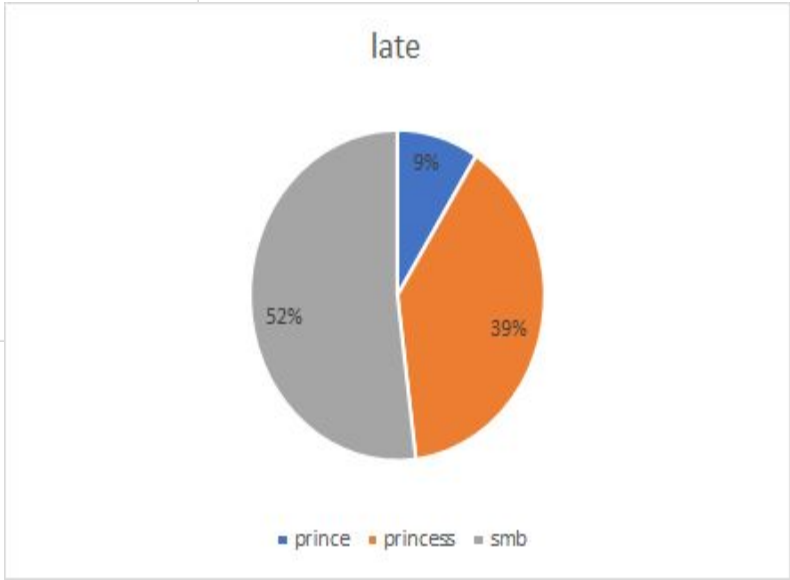
Количество реплик по группам персонажей и периодам (в процентах)

epoch	prince	princess	smb
early	[0.027924528301886794]	[0.17735849056603772]	[0.7947169811320755]
mid	[0.19181708784596871]	[0.16197352587244285]	[0.6462093862815884]
late	[0.08641975308641975]	[0.39021164021164023]	[0.5233686067019401]


early
3%
18%
79%
prince
princess
smb

prince
princess
smb


mid
19%
16%
65%
prince
princess
smb

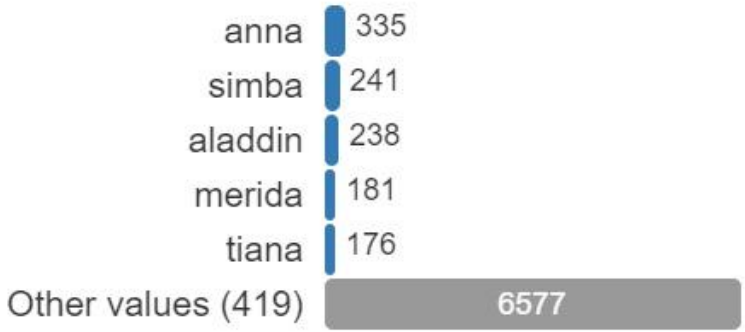

late
9%
39%
52%
prince
princess
smb

Speaker

Categorical

HIGH CARDINALITY

Distinct count	424
Unique (%)	5.5%
Missing	0
Missing (%)	0.0%
Memory size	60.5 KiB



Принцессы и период

Всего 14 принцесс, 3 из которых относятся к периоду EARLY, 6 - MID, 5 - LATE

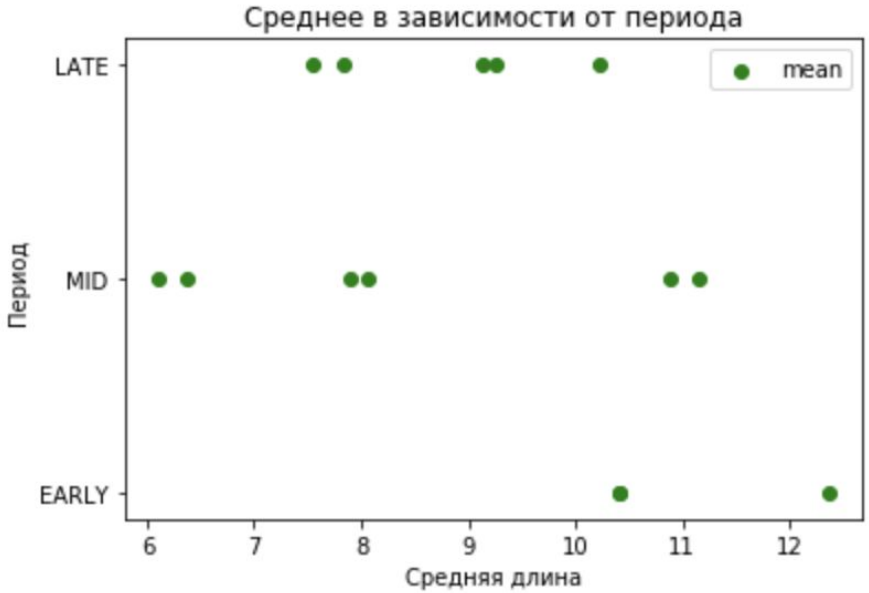
	PERIOD	STATUS	PRINCESS
0	EARLY	PRINCESS	snow white
1	EARLY	PRINCESS	cinderella
2	EARLY	PRINCESS	briar rose
3	MID	PRINCESS	ariel
4	MID	PRINCESS	belle
5	MID	PRINCESS	jasmine
6	MID	PRINCESS	nala
7	MID	PRINCESS	pocahontas
8	MID	PRINCESS	mulan
9	LATE	PRINCESS	tiana
10	LATE	PRINCESS	rapunzel
11	LATE	PRINCESS	merida
12	LATE	PRINCESS	elsa
13	LATE	PRINCESS	anna

Среднее количество слов принцесс в зависимости от периода

	PERIOD	MEAN_LEN
0	EARLY	11.140426
1	MID	8.435364
2	LATE	8.597740

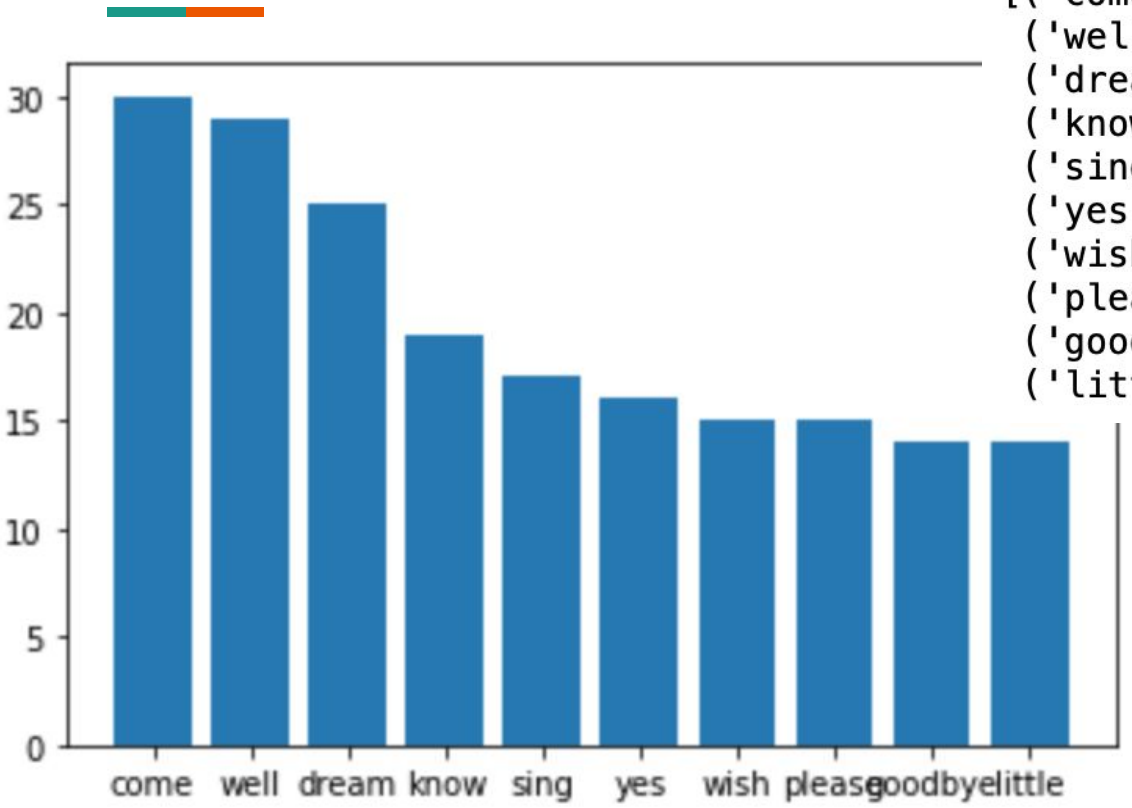
Среднее количество слов каждой принцессы

	PERIOD	PRINCESS	MEAN_LEN
0	EARLY	Snow white	12.363636
1	EARLY	Cinderella	10.410714
2	EARLY	Briar Rose	10.400000
3	MID	Ariel	11.148649
4	MID	Belle	7.893333
5	MID	Jasmine	6.111111
6	MID	Nala	6.364865
7	MID	Pocahontas	10.887324
8	MID	Mulan	8.068966
9	LATE	Tiana	10.232955
10	LATE	Rapunzel	9.250000
11	LATE	Merida	9.139665
12	LATE	Elsa	7.828829
13	LATE	Anna	7.540299



Количество слов по годам
Количество слов
0
50
100
150
200
250
300
350
1940
1950
1960
1970
1980
1990
2000
2010
Год

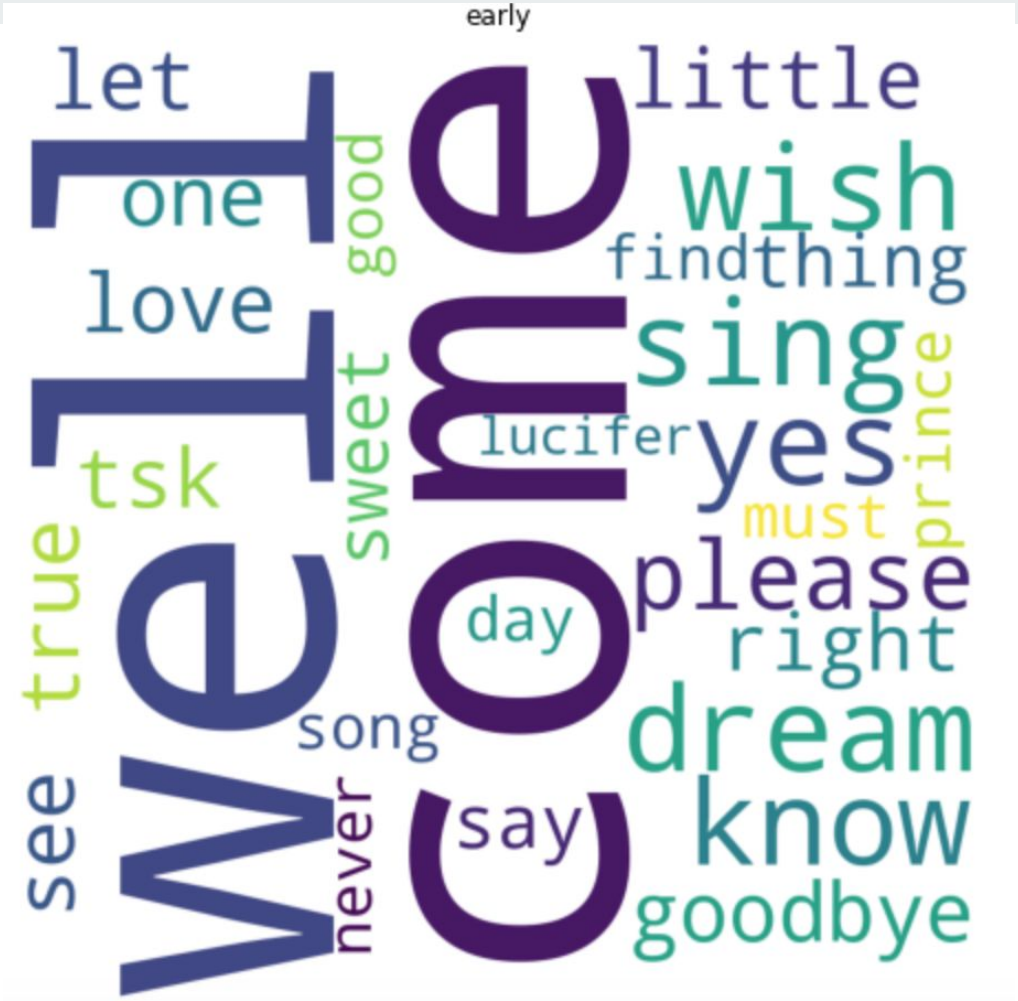
Частотные слова




```
1 early_cnt.most_common(10)
```

```
[('come', 30),
 ('well', 29),
 ('dream', 25),
 ('know', 19),
 ('sing', 17),
 ('yes', 16),
 ('wish', 15),
 ('please', 15),
 ('goodbye', 14),
 ('little', 14)]
```

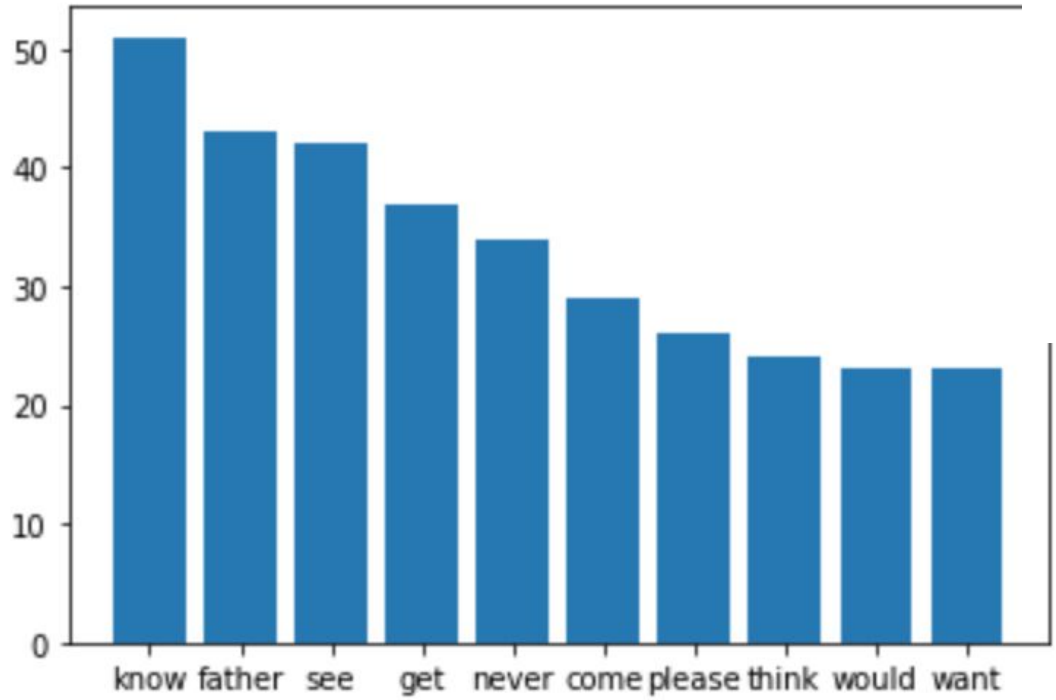
EARLY


early
welcome
let
little
one
wish
good
findthing
love
sing
sweet
lucifer
yes
prince
tsk
must
true
please
day
right
dream
song
see
say
know
never
goodbye

Частотные слова

```
1 mid_cnt.most_common(10)
```

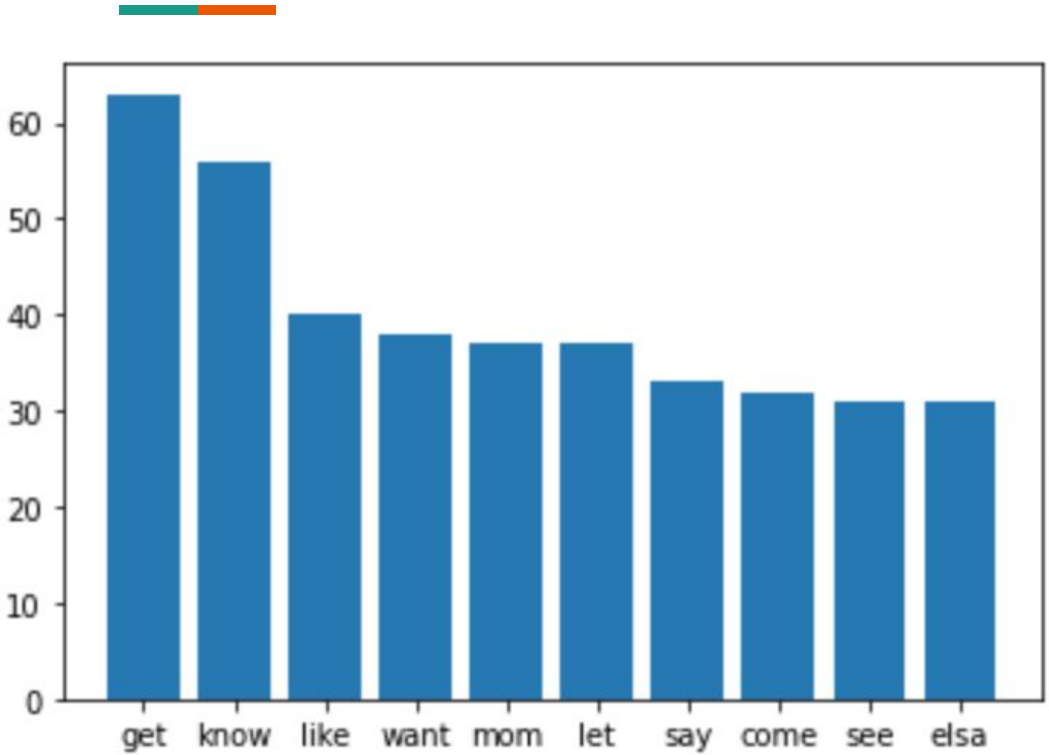
```
[('know', 51),
 ('father', 43),
 ('see', 42),
 ('get', 37),
 ('never', 34),
 ('come', 29),
 ('please', 26),
 ('think', 24),
 ('would', 23),
 ('want', 23)]
```



MIDDLE

middle
come
see
sorry
around
really
right
simba
father
help
let
never
tell
mean
gaston
well
papa
back
look
want
love
know
one
say
please
think
something

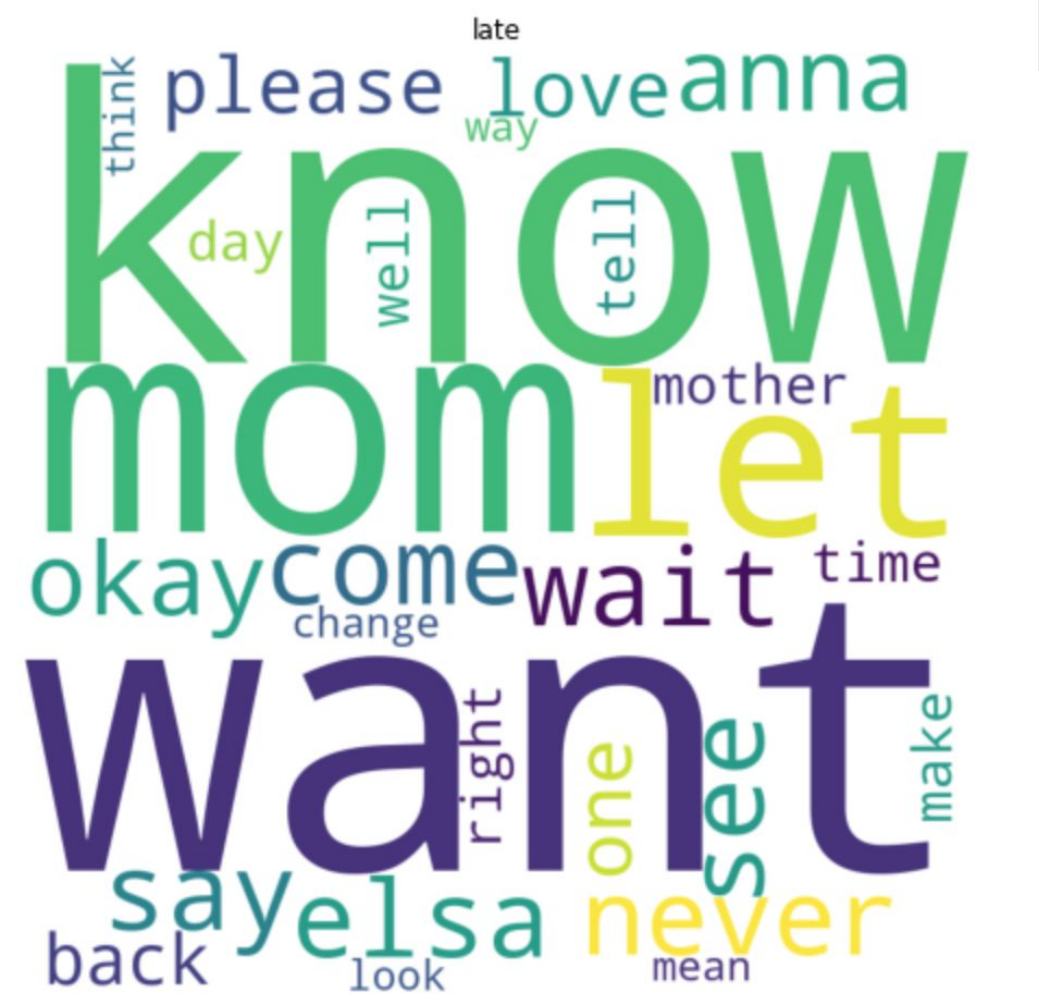
Частотные слова




```
1 late_cnt.most_common(10)
```

```
[('get', 63),
 ('know', 56),
 ('like', 40),
 ('want', 38),
 ('mom', 37),
 ('let', 37),
 ('say', 33),
 ('come', 32),
 ('see', 31),
 ('elsa', 31)]
```

LATE


late
know
want
mom
let
please
love
anna
think
way
day
well
tell
mother
okay
come
wait
time
change
right
see
make
one
say
elsa
never
back
look
mean