

```
1 /opt/anaconda3/envs/python/bin/python /Users/  
Federica/Documents/PhD/chronic_conditions/  
reddit_LDAtomotopy.py  
2 /opt/anaconda3/envs/python/lib/python3.9/site-  
packages/past/builtins/misc.py:45:  
DeprecationWarning: the imp module is deprecated in  
favour of importlib; see the module's  
documentation for alternative uses  
3 from imp import reload  
4 Data Cleaning...  
5 Number of Documents: 368443  
6 Vocabulary Size: 36460  
7 2021-10-25 14:32:10.487980_____0:12:43.  
874381  
8  
9 Performing Topic Modeling...  
10 Num docs:368443  
11 Vocabulary Size: 36460  
12 Removed Top words: HD Ph observed Endocrinologists  
brine  
13 Iteration: 0 Log-likelihood: -7.976490896311141  
14 Iteration: 10 Log-likelihood: -7.8676948211466105  
15 Iteration: 20 Log-likelihood: -7.642125015877292  
16 Iteration: 30 Log-likelihood: -7.525287112473097  
17 Iteration: 40 Log-likelihood: -7.469008664024656  
18 Iteration: 50 Log-likelihood: -7.441453506067228  
19 Iteration: 60 Log-likelihood: -7.424287091859234  
20 Iteration: 70 Log-likelihood: -7.414824535018156  
21 Iteration: 80 Log-likelihood: -7.40806875134714  
22 Iteration: 90 Log-likelihood: -7.402389829548536  
23 2021-10-25 14:32:35.917832____Topic modeling 5, 5  
time:____0:00:25.429745  
24  
25 Performing Topic Modeling...  
26 Num docs:368443  
27 Vocabulary Size: 36460  
28 Removed Top words: 6x6 3x3 bymy OBYGN Huntington  
HD Ph observed Endocrinologists brine  
29 Iteration: 0 Log-likelihood: -8.082023503885432  
30 Iteration: 10 Log-likelihood: -7.929213401718376  
31 Iteration: 20 Log-likelihood: -7.74834615423005
```

```

32 Iteration: 30    Log-likelihood: -7.638606894261593
33 Iteration: 40    Log-likelihood: -7.570407925368479
34 Iteration: 50    Log-likelihood: -7.536143615031014
35 Iteration: 60    Log-likelihood: -7.5196241051751675
36 Iteration: 70    Log-likelihood: -7.509803602588022
37 Iteration: 80    Log-likelihood: -7.503591046289951
38 Iteration: 90    Log-likelihood: -7.49880768959023
39 2021-10-25 14:32:57.375741____Topic modeling 5, 10
   time:____0:00:21.457824
40
41 Performing Topic Modeling...
42 Num docs:368443
43 Vocabulary Size: 36460
44 Removed Top words:  HD Ph observed Endocrinologists
   brine
45 Iteration: 0      Log-likelihood: -8.49658106040562
46 Iteration: 10     Log-likelihood: -8.113131603221873
47 Iteration: 20     Log-likelihood: -7.815727970677854
48 Iteration: 30     Log-likelihood: -7.689528641319329
49 Iteration: 40     Log-likelihood: -7.625122208216227
50 Iteration: 50     Log-likelihood: -7.5890465585033695
51 Iteration: 60     Log-likelihood: -7.566088187226067
52 Iteration: 70     Log-likelihood: -7.546433972047594
53 Iteration: 80     Log-likelihood: -7.531897019256159
54 Iteration: 90     Log-likelihood: -7.523660616062498
55 2021-10-25 14:33:24.658240____Topic modeling 10, 5
   time:____0:00:27.282418
56
57 Performing Topic Modeling...
58 Num docs:368443
59 Vocabulary Size: 36460
60 Removed Top words:  6x6 3x3 bymy OBYGN Huntington
   HD Ph observed Endocrinologists brine
61 Iteration: 0      Log-likelihood: -8.595403712653194
62 Iteration: 10     Log-likelihood: -8.193580162805683
63 Iteration: 20     Log-likelihood: -7.894952482265365
64 Iteration: 30     Log-likelihood: -7.771375872485265
65 Iteration: 40     Log-likelihood: -7.715811416493037
66 Iteration: 50     Log-likelihood: -7.682899392890313
67 Iteration: 60     Log-likelihood: -7.660065702505145
68 Iteration: 70     Log-likelihood: -7.644681975713711

```

```

69 Iteration: 80    Log-likelihood: -7.
   6315363841979185
70 Iteration: 90    Log-likelihood: -7.623993054411781
71 2021-10-25 14:33:52.049868____Topic modeling 10,
   10 time:____0:00:27.391470
72
73 Performing Topic Modeling...
74 Num docs:368443
75 Vocabulary Size: 36460
76 Removed Top words:  HD Ph observed
   Endocrinologists brine
77 Iteration: 0      Log-likelihood: -8.810225918445914
78 Iteration: 10     Log-likelihood: -8.201548660280874
79 Iteration: 20     Log-likelihood: -7.923552187153283
80 Iteration: 30     Log-likelihood: -7.789769000731277
81 Iteration: 40     Log-likelihood: -7.712829302852153
82 Iteration: 50     Log-likelihood: -7.664883685962193
83 Iteration: 60     Log-likelihood: -7.629612622237007
84 Iteration: 70     Log-likelihood: -7.604640899037096
85 Iteration: 80     Log-likelihood: -7.586696246344023
86 Iteration: 90     Log-likelihood: -7.571247945227252
87 2021-10-25 14:34:24.981304____Topic modeling 15, 5
   time:____0:00:32.931347
88
89 Performing Topic Modeling...
90 Num docs:368443
91 Vocabulary Size: 36460
92 Removed Top words:  6x6 3x3 bymy OBYGN Huntington
   HD Ph observed Endocrinologists brine
93 Iteration: 0      Log-likelihood: -8.922704283990685
94 Iteration: 10     Log-likelihood: -8.291697652059154
95 Iteration: 20     Log-likelihood: -7.990861804189091
96 Iteration: 30     Log-likelihood: -7.864067758872466
97 Iteration: 40     Log-likelihood: -7.793471135829386
98 Iteration: 50     Log-likelihood: -7.749570339332396
99 Iteration: 60     Log-likelihood: -7.72103391131209
100 Iteration: 70     Log-likelihood: -7.698486985478237
101 Iteration: 80     Log-likelihood: -7.681812434229997
102 Iteration: 90     Log-likelihood: -7.666883569864263
103 2021-10-25 14:34:57.521431____Topic modeling 15,
   10 time:____0:00:32.539985

```

104

105

106 Process finished with exit code 0

107