

# [Code] Topic Modeling

## Code Snippet

Package Info : <https://bab2min.github.io/tomotopy/v0.6.2/en/>

```
# Package Import
import tomotopy as tp

# LDA Topic Modeling Function
def FindTopic(dataframe, attr=10, iteration=100):

    mdl = tp.LDAModel(k=attr, min_cf=5) # LDA

    # model store
    for text in dataframe["topic_words"].tolist(): # Tokenized Text column
        mdl.add_doc(text)

    # model train
    for i in range(iteration):
        mdl.train()
        print('Iteration: {} \t Log-likelihood: {}'.format(i, mdl.ll_per_word))

    # result
    topic_words = []
    for k in range(mdl.k):
        topic_words.append(mdl.get_topic_words(k, top_n=30)) # frequency top 30 words
```

```

return topic_words

# Do Topic modeling, k=20
topic_words = FindTopic(your_dataframe, 20)

```

## Output Example

