

```

import nltk
from nltk.corpus import stopwords
from nltk.tokenize import word_tokenize
# Download required resources
nltk.download('punkt')
nltk.download('stopwords')

# Sample text
text = """The New Zealand national cricket team, known as the Black
Caps, is a strong force in international cricket.
As of March 2025, they are led by captain Mitchell Santner, a skilled
all-rounder. In the 2025 ICC Champions Trophy,
New Zealand reached the final but lost to India by four wickets.

Key Players:
- Kane Williamson: Experienced batsman and former captain.
- Devon Conway: Reliable left-handed batsman.
- Daryl Mitchell: Vice-captain and all-rounder.
- Rachin Ravindra: Emerging star, scored a century on Champions Trophy
debut.
- Tom Latham: Wicket-keeper and dependable batsman.
- Matt Henry & Kyle Jamieson: Leading pace bowlers.
- Ish Sodhi & Mitchell Santner: Spin bowling strengths.

New Zealand remains a competitive team, blending experienced players
with rising talents,
aiming for continued success on the global stage."""

# Tokenize text
tokens = word_tokenize(text.lower())

# Remove punctuation and stopwords
stop_words = set(stopwords.words('english'))
tokens = [word for word in tokens if word.isalpha() and word not in
stop_words]
print(tokens[0:30])

['new', 'zealand', 'national', 'cricket', 'team', 'known', 'black',
'caps', 'strong', 'force', 'international', 'cricket', 'march', 'led',
'captain', 'mitchell', 'santner', 'skilled', 'icc', 'champions',
'trophy', 'new', 'zealand', 'reached', 'final', 'lost', 'india',
'four', 'wickets', 'key']

[nltk_data] Downloading package punkt to
[nltk_data] C:\Users\kamli\AppData\Roaming\nltk_data...
[nltk_data] Package punkt is already up-to-date!
[nltk_data] Downloading package stopwords to
[nltk_data] C:\Users\kamli\AppData\Roaming\nltk_data...
[nltk_data] Package stopwords is already up-to-date!

```

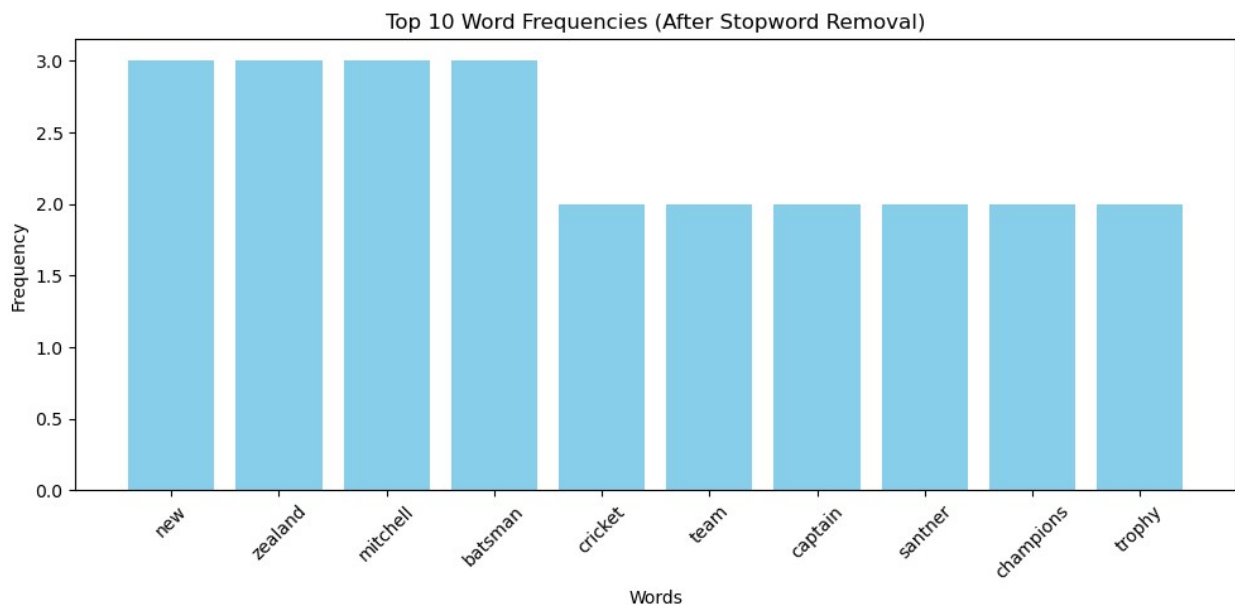
```

from collections import Counter
# Count word frequencies
word_counts = Counter(tokens)
# Select top 10 most common words
most_common_words = word_counts.most_common(10)
words, counts = zip(*most_common_words)
words, counts

(('new',
  'zealand',
  'mitchell',
  'batsman',
  'cricket',
  'team',
  'captain',
  'santner',
  'champions',
  'trophy'),
 (3, 3, 3, 3, 2, 2, 2, 2, 2, 2))

import matplotlib.pyplot as plt
# Plot bar chart
plt.figure(figsize=(10, 5))
plt.bar(words, counts, color='skyblue')
plt.xlabel('Words')
plt.ylabel('Frequency')
plt.title('Top 10 Word Frequencies (After Stopword Removal)')
plt.xticks(rotation=45)
plt.tight_layout()
plt.show()

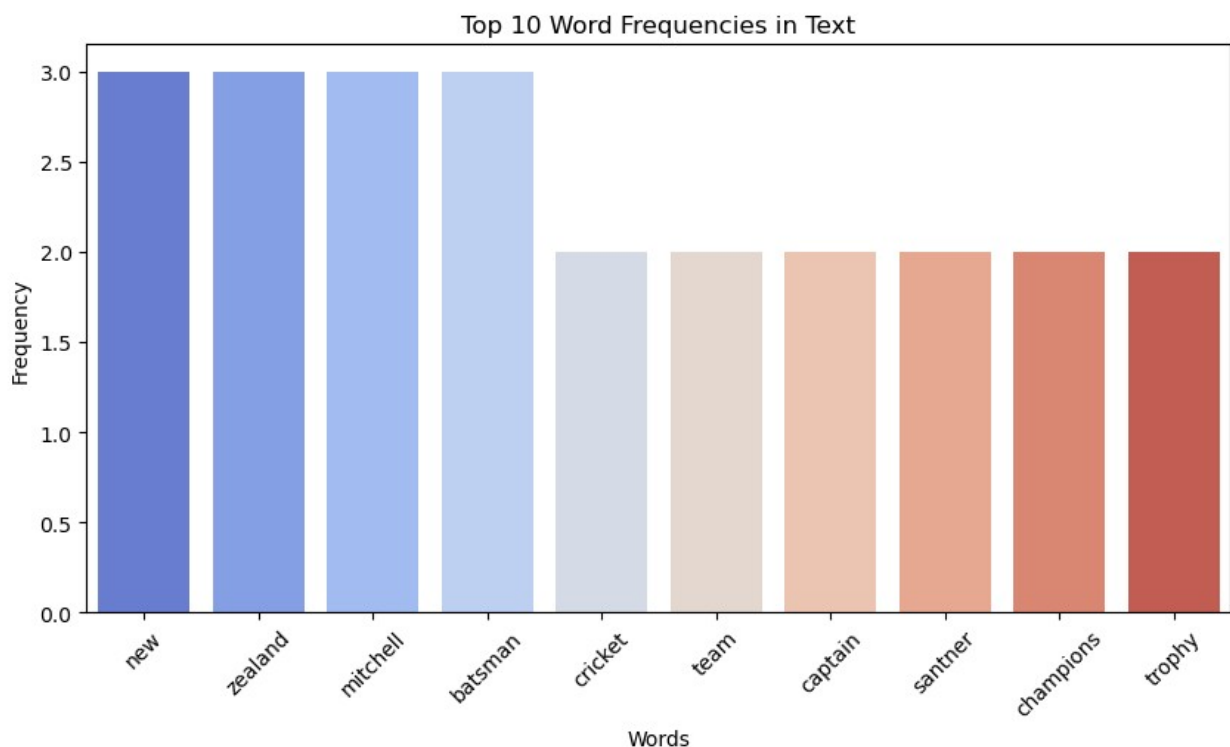
```



```
import seaborn as sns
# Plot with Seaborn
plt.figure(figsize=(10, 5))
sns.barplot(x=list(words), y=list(counts), palette="coolwarm")
plt.xlabel("Words")
plt.ylabel("Frequency")
plt.title("Top 10 Word Frequencies in Text")
plt.xticks(rotation=45)
plt.show()
```

C:\Users\kamli\anaconda3\Lib\site-packages\seaborn_oldcore.py:1765: FutureWarning:

unique with argument that is not not a Series, Index, ExtensionArray, or np.ndarray is deprecated and will raise in a future version.



```
import plotly.express as px
# Plot with Plotly
df = {"Words": [w[0] for w in most_common_words], "Counts": [w[1] for w in most_common_words]}
fig = px.bar(df, x="Words", y="Counts", title="Top 10 Word Frequencies in Text", color="Counts")
fig.show()
```

```

{"config":{"plotlyServerURL":"https://plot.ly"},"data":
[{"alignmentgroup":"True","hovertemplate":"Words=%{x}<br>Counts=%
{marker.color}<extra></extra>","legendgroup":"","marker":{"color":
[3,3,3,3,2,2,2,2,2,2],"coloraxis":"coloraxis","pattern":
{"shape":""},"name":"","offsetgroup":"","orientation":"v","showlegend
":false,"textposition":"auto","type":"bar","x":
["new","zealand","mitchell","batsman","cricket","team","captain","sant
ner","champions","trophy"],"xaxis":"x","y":
[3,3,3,3,2,2,2,2,2,2],"yaxis":"y"}],"layout":
{"barmode":"relative","coloraxis":{"colorbar":{"title":
{"text":"Counts"},"colorscale":[[0,"#0d0887"],
[0.1111111111111111,"#46039f"],[0.2222222222222222,"#7201a8"],
[0.3333333333333333,"#9c179e"],[0.4444444444444444,"#bd3786"],
[0.5555555555555556,"#d8576b"],[0.6666666666666666,"#ed7953"],
[0.7777777777777778,"#fb9f3a"],[0.8888888888888888,"#fdca26"],
[1,"#f0f921"]]},"legend":{"tracegroupgap":0},"template":{"data":
{"bar":[{"error_x":{"color":"#2a3f5f"},"error_y":
{"color":"#2a3f5f"},"marker":{"line":
{"color":"#E5ECF6","width":0.5},"pattern":
{"fillmode":"overlay","size":10,"solidity":0.2},"type":"bar"}],"barpo
lar":[{"marker":{"line":{"color":"#E5ECF6","width":0.5},"pattern":
{"fillmode":"overlay","size":10,"solidity":0.2},"type":"barpolar"}],"
carpet":[{"aaxis":
{"endlinecolor":"#2a3f5f","gridcolor":"white","linecolor":"white","min
orgridcolor":"white","startlinecolor":"#2a3f5f"},"baxis":
{"endlinecolor":"#2a3f5f","gridcolor":"white","linecolor":"white","min
orgridcolor":"white","startlinecolor":"#2a3f5f"},"type":"carpet"}],"ch
oropleth":[{"colorbar":
{"outlinewidth":0,"ticks":"","type":"choropleth"}],"contour":
[{"colorbar":{"outlinewidth":0,"ticks":"","colorscale":
[[0,"#0d0887"],[0.1111111111111111,"#46039f"],
[0.2222222222222222,"#7201a8"],[0.3333333333333333,"#9c179e"],
[0.4444444444444444,"#bd3786"],[0.5555555555555556,"#d8576b"],
[0.6666666666666666,"#ed7953"],[0.7777777777777778,"#fb9f3a"],
[0.8888888888888888,"#fdca26"],
[1,"#f0f921"]]},"type":"contour"}],"contourcarpet":[{"colorbar":
{"outlinewidth":0,"ticks":"","type":"contourcarpet"}],"heatmap":
[{"colorbar":{"outlinewidth":0,"ticks":"","colorscale":
[[0,"#0d0887"],[0.1111111111111111,"#46039f"],
[0.2222222222222222,"#7201a8"],[0.3333333333333333,"#9c179e"],
[0.4444444444444444,"#bd3786"],[0.5555555555555556,"#d8576b"],
[0.6666666666666666,"#ed7953"],[0.7777777777777778,"#fb9f3a"],
[0.8888888888888888,"#fdca26"],
[1,"#f0f921"]]},"type":"heatmap"}],"heatmapgl":[{"colorbar":
{"outlinewidth":0,"ticks":"","colorscale":[[0,"#0d0887"],
[0.1111111111111111,"#46039f"],[0.2222222222222222,"#7201a8"],
[0.3333333333333333,"#9c179e"],[0.4444444444444444,"#bd3786"],
[0.5555555555555556,"#d8576b"],[0.6666666666666666,"#ed7953"],
[0.7777777777777778,"#fb9f3a"],[0.8888888888888888,"#fdca26"],
[1,"#f0f921"]]},"type":"heatmapgl"}],"histogram":[{"marker":{"pattern":

```

```

{"fillmode":"overlay","size":10,"solidity":0.2},"type":"histogram"}],
"histogram2d":[{"colorbar":{"linewidth":0,"ticks":"","","colorscale":
[[0,"#0d0887"],[0.1111111111111111,"#46039f"],
[0.2222222222222222,"#7201a8"],[0.3333333333333333,"#9c179e"],
[0.4444444444444444,"#bd3786"],[0.5555555555555556,"#d8576b"],
[0.6666666666666666,"#ed7953"],[0.7777777777777778,"#fb9f3a"],
[0.8888888888888888,"#fdca26"],
[1,"#f0f921"]],"type":"histogram2d"}],"histogram2dcontour":
{"colorbar":{"linewidth":0,"ticks":"","","colorscale":
[[0,"#0d0887"],[0.1111111111111111,"#46039f"],
[0.2222222222222222,"#7201a8"],[0.3333333333333333,"#9c179e"],
[0.4444444444444444,"#bd3786"],[0.5555555555555556,"#d8576b"],
[0.6666666666666666,"#ed7953"],[0.7777777777777778,"#fb9f3a"],
[0.8888888888888888,"#fdca26"],
[1,"#f0f921"]],"type":"histogram2dcontour"}],"mesh3d":[{"colorbar":
{"linewidth":0,"ticks":"","","type":"mesh3d"}],"parcoords":[{"line":
{"colorbar":{"linewidth":0,"ticks":"","","type":"parcoords"}],"pie":
[{"automargin":true,"type":"pie"}],"scatter":[{"fillpattern":
{"fillmode":"overlay","size":10,"solidity":0.2},"type":"scatter"}],"scatter3d":[{"line":{"colorbar":{"linewidth":0,"ticks":"","","marker":
{"colorbar":
{"linewidth":0,"ticks":"","","type":"scatter3d"}],"scattercarpet":
[{"marker":{"colorbar":
{"linewidth":0,"ticks":"","","type":"scattercarpet"}],"scattergeo":
[{"marker":{"colorbar":
{"linewidth":0,"ticks":"","","type":"scattergeo"}],"scattergl":
[{"marker":{"colorbar":
{"linewidth":0,"ticks":"","","type":"scattergl"}],"scattermapbox":
[{"marker":{"colorbar":
{"linewidth":0,"ticks":"","","type":"scattermapbox"}],"scatterpolar":
[{"marker":{"colorbar":
{"linewidth":0,"ticks":"","","type":"scatterpolar"}],"scatterpolargl":
[{"marker":{"colorbar":
{"linewidth":0,"ticks":"","","type":"scatterpolargl"}],"scatterternary":
[{"marker":{"colorbar":
{"linewidth":0,"ticks":"","","type":"scatterternary"}],"surface":
[{"colorbar":{"linewidth":0,"ticks":"","","colorscale":
[[0,"#0d0887"],[0.1111111111111111,"#46039f"],
[0.2222222222222222,"#7201a8"],[0.3333333333333333,"#9c179e"],
[0.4444444444444444,"#bd3786"],[0.5555555555555556,"#d8576b"],
[0.6666666666666666,"#ed7953"],[0.7777777777777778,"#fb9f3a"],
[0.8888888888888888,"#fdca26"],
[1,"#f0f921"]],"type":"surface"}],"table":[{"cells":{"fill":
{"color":"#EBF0F8"},"line":{"color":"white"},"header":{"fill":
{"color":"#C8D4E3"},"line":
{"color":"white"},"type":"table"}],"layout":{"annotationdefaults":
{"arrowcolor":"#2a3f5f","arrowhead":0,"arrowwidth":1},"autotypenumbers":
"strict","coloraxis":{"colorbar":
{"linewidth":0,"ticks":"","","colorscale":{"diverging":
[[0,"#8e0152"],[0.1,"#c51b7d"],[0.2,"#de77ae"],[0.3,"#f1b6da"],

```

```

[0.4,"#fde0ef"],[0.5,"#f7f7f7"],[0.6,"#e6f5d0"],[0.7,"#b8e186"],
[0.8,"#7fb341"],[0.9,"#4d9221"],[1,"#276419"]],"sequential":
[[0,"#0d0887"],[0.1111111111111111,"#46039f"],
[0.2222222222222222,"#7201a8"],[0.3333333333333333,"#9c179e"],
[0.4444444444444444,"#bd3786"],[0.5555555555555556,"#d8576b"],
[0.6666666666666666,"#ed7953"],[0.7777777777777778,"#fb9f3a"],
[0.8888888888888888,"#fdca26"],[1,"#f0f921"]],"sequentialminus":
[[0,"#0d0887"],[0.1111111111111111,"#46039f"],
[0.2222222222222222,"#7201a8"],[0.3333333333333333,"#9c179e"],
[0.4444444444444444,"#bd3786"],[0.5555555555555556,"#d8576b"],
[0.6666666666666666,"#ed7953"],[0.7777777777777778,"#fb9f3a"],
[0.8888888888888888,"#fdca26"],[1,"#f0f921"]]],"colorway":
["#636efa","#EF553B","#00cc96","#ab63fa","#FFA15A","#19d3f3","#FF6692",
"#B6E880","#FF97FF","#FECB52"],"font":{"color":"#2a3f5f"},"geo":
{"bgcolor":"white","lakecolor":"white","landcolor":"#E5ECF6","showlake
s":true,"showland":true,"subunitcolor":"white"},"hoverlabel":
{"align":"left"},"hovermode":"closest","mapbox":
{"style":"light"},"paper_bgcolor":"white","plot_bgcolor":"#E5ECF6","po
lar":{"angularaxis":
{"gridcolor":"white","linecolor":"white","ticks":""},"bgcolor":"#E5ECF
6","radialaxis":
{"gridcolor":"white","linecolor":"white","ticks":""}}},"scene":
{"xaxis":
{"backgroundcolor":"#E5ECF6","gridcolor":"white","gridwidth":2,"lineco
lor":"white","showbackground":true,"ticks":"","zerolinecolor":"white"}
,"yaxis":
{"backgroundcolor":"#E5ECF6","gridcolor":"white","gridwidth":2,"lineco
lor":"white","showbackground":true,"ticks":"","zerolinecolor":"white"}
,"zaxis":
{"backgroundcolor":"#E5ECF6","gridcolor":"white","gridwidth":2,"lineco
lor":"white","showbackground":true,"ticks":"","zerolinecolor":"white"}
},"shapedefaults":{"line":{"color":"#2a3f5f"}}},"ternary":{"aaxis":
{"gridcolor":"white","linecolor":"white","ticks":""},"baxis":
{"gridcolor":"white","linecolor":"white","ticks":""},"bgcolor":"#E5ECF
6","caxis":
{"gridcolor":"white","linecolor":"white","ticks":""}}},"title":
{"x":5.0e-2},"xaxis":
{"automargin":true,"gridcolor":"white","linecolor":"white","ticks":"","
"title":
{"standoff":15},"zerolinecolor":"white","zerolinewidth":2},"yaxis":
{"automargin":true,"gridcolor":"white","linecolor":"white","ticks":"","
"title":
{"standoff":15},"zerolinecolor":"white","zerolinewidth":2}}},"title":
{"text":"Top 10 Word Frequencies in Text"},"xaxis":
{"anchor":"y","domain":[0,1],"title":{"text":"Words"},"yaxis":
{"anchor":"x","domain":[0,1],"title":{"text":"Counts"}}}]

```

```

from wordcloud import WordCloud
# Generate Word Cloud
wordcloud = WordCloud(width=800, height=400,

```



```

background_color='white').generate_from_frequencies(word_counts)
# Display Word Cloud
plt.figure(figsize=(10, 5))
plt.imshow(wordcloud, interpolation="bilinear")
plt.axis("off")
plt.title("Word Cloud Visualization")
plt.show()

```

