

```
In [14]: !pip install beautifulsoup4 requests
!pip install fsspec
#Install reqired Libraries
!pip install pandas
!pip install matplotlib
!pip install lxml
!pip install nltk
!pip install wordcloud
```

Requirement already satisfied: beautifulsoup4 in c:\users\grigor\.conda\envs\ads_502\lib\site-packages (4.12.2)

Requirement already satisfied: requests in c:\users\grigor\.conda\envs\ads_502\lib\site-packages (2.29.0)

Requirement already satisfied: soupsieve>1.2 in c:\users\grigor\.conda\envs\ads_502\lib\site-packages (from beautifulsoup4) (2.4.1)

Requirement already satisfied: charset-normalizer<4,>=2 in c:\users\grigor\.conda\envs\ads_502\lib\site-packages (from requests) (2.0.4)

Requirement already satisfied: idna<4,>=2.5 in c:\users\grigor\.conda\envs\ads_502\lib\site-packages (from requests) (3.4)

Requirement already satisfied: urllib3<1.27,>=1.21.1 in c:\users\grigor\.conda\envs\ads_502\lib\site-packages (from requests) (1.26.16)

Requirement already satisfied: certifi>=2017.4.17 in c:\users\grigor\.conda\envs\ads_502\lib\site-packages (from requests) (2023.5.7)

Requirement already satisfied: fsspec in c:\users\grigor\.conda\envs\ads_502\lib\site-packages (2023.4.0)

Requirement already satisfied: pandas in c:\users\grigor\.conda\envs\ads_502\lib\site-packages (1.4.1)

Requirement already satisfied: python-dateutil>=2.8.1 in c:\users\grigor\.conda\envs\ads_502\lib\site-packages (from pandas) (2.8.2)

Requirement already satisfied: pytz>=2020.1 in c:\users\grigor\.conda\envs\ads_502\lib\site-packages (from pandas) (2022.7)

Requirement already satisfied: numpy>=1.18.5 in c:\users\grigor\.conda\envs\ads_502\lib\site-packages (from pandas) (1.21.5)

Requirement already satisfied: six>=1.5 in c:\users\grigor\.conda\envs\ads_502\lib\site-packages (from python-dateutil>=2.8.1->pandas) (1.16.0)

Requirement already satisfied: matplotlib in c:\users\grigor\.conda\envs\ads_502\lib\site-packages (3.7.1)

Requirement already satisfied: contourpy>=1.0.1 in c:\users\grigor\.conda\envs\ads_502\lib\site-packages (from matplotlib) (1.0.5)

Requirement already satisfied: cyclor>=0.10 in c:\users\grigor\.conda\envs\ads_502\lib\site-packages (from matplotlib) (0.11.0)

Requirement already satisfied: fonttools>=4.22.0 in c:\users\grigor\.conda\envs\ads_502\lib\site-packages (from matplotlib) (4.25.0)

Requirement already satisfied: kiwisolver>=1.0.1 in c:\users\grigor\.conda\envs\ads_502\lib\site-packages (from matplotlib) (1.4.4)

Requirement already satisfied: numpy>=1.20 in c:\users\grigor\.conda\envs\ads_502\lib\site-packages (from matplotlib) (1.21.5)

Requirement already satisfied: packaging>=20.0 in c:\users\grigor\.conda\envs\ads_502\lib\site-packages (from matplotlib) (23.0)

Requirement already satisfied: pillow>=6.2.0 in c:\users\grigor\.conda\envs\ads_502\lib\site-packages (from matplotlib) (9.4.0)

Requirement already satisfied: pyparsing>=2.3.1 in c:\users\grigor\.conda\envs\ads_502\lib\site-packages (from matplotlib) (3.0.9)

Requirement already satisfied: python-dateutil>=2.7 in c:\users\grigor\.conda\envs\ads_502\lib\site-packages (from matplotlib) (2.8.2)

Requirement already satisfied: importlib-resources>=3.2.0 in c:\users\grigor\.conda\envs\ads_502\lib\site-packages (from matplotlib) (5.2.0)

Requirement already satisfied: zipp>=3.1.0 in c:\users\grigor\.conda\envs\ads_502\lib\site-packages (from importlib-resources>=3.2.0->matplotlib) (3.11.0)

Requirement already satisfied: six>=1.5 in c:\users\grigor\.conda\envs\ads_502\lib\site-packages (from python-dateutil>=2.7->matplotlib) (1.16.0)

Requirement already satisfied: lxml in c:\users\grigor\.conda\envs\ads_502\lib\site-packages (4.9.2)

Requirement already satisfied: nltk in c:\users\grigor\.conda\envs\ads_502\lib\site-packages (3.8.1)

Requirement already satisfied: click in c:\users\grigor\.conda\envs\ads_502\lib\site-packages (from nltk) (8.0.4)

Requirement already satisfied: joblib in c:\users\grigor\appdata\roaming\python\python38\site-packages (from nltk) (1.2.0)

Requirement already satisfied: regex>=2021.8.3 in c:\users\grigor\.conda\envs\ads_502\lib\site-packages (from nltk) (2022.7.9)
Requirement already satisfied: tqdm in c:\users\grigor\.conda\envs\ads_502\lib\site-packages (from nltk) (4.65.0)
Requirement already satisfied: colorama in c:\users\grigor\.conda\envs\ads_502\lib\site-packages (from click->nltk) (0.4.6)
Requirement already satisfied: wordcloud in c:\users\grigor\.conda\envs\ads_502\lib\site-packages (1.9.2)
Requirement already satisfied: numpy>=1.6.1 in c:\users\grigor\.conda\envs\ads_502\lib\site-packages (from wordcloud) (1.21.5)
Requirement already satisfied: pillow in c:\users\grigor\.conda\envs\ads_502\lib\site-packages (from wordcloud) (9.4.0)
Requirement already satisfied: matplotlib in c:\users\grigor\.conda\envs\ads_502\lib\site-packages (from wordcloud) (3.7.1)
Requirement already satisfied: contourpy>=1.0.1 in c:\users\grigor\.conda\envs\ads_502\lib\site-packages (from matplotlib->wordcloud) (1.0.5)
Requirement already satisfied: cyclor>=0.10 in c:\users\grigor\.conda\envs\ads_502\lib\site-packages (from matplotlib->wordcloud) (0.11.0)
Requirement already satisfied: fonttools>=4.22.0 in c:\users\grigor\.conda\envs\ads_502\lib\site-packages (from matplotlib->wordcloud) (4.25.0)
Requirement already satisfied: kiwisolver>=1.0.1 in c:\users\grigor\.conda\envs\ads_502\lib\site-packages (from matplotlib->wordcloud) (1.4.4)
Requirement already satisfied: packaging>=20.0 in c:\users\grigor\.conda\envs\ads_502\lib\site-packages (from matplotlib->wordcloud) (23.0)
Requirement already satisfied: pyparsing>=2.3.1 in c:\users\grigor\.conda\envs\ads_502\lib\site-packages (from matplotlib->wordcloud) (3.0.9)
Requirement already satisfied: python-dateutil>=2.7 in c:\users\grigor\.conda\envs\ads_502\lib\site-packages (from matplotlib->wordcloud) (2.8.2)
Requirement already satisfied: importlib-resources>=3.2.0 in c:\users\grigor\.conda\envs\ads_502\lib\site-packages (from matplotlib->wordcloud) (5.2.0)
Requirement already satisfied: zipp>=3.1.0 in c:\users\grigor\.conda\envs\ads_502\lib\site-packages (from importlib-resources>=3.2.0->matplotlib->wordcloud) (3.11.0)
Requirement already satisfied: six>=1.5 in c:\users\grigor\.conda\envs\ads_502\lib\site-packages (from python-dateutil>=2.7->matplotlib->wordcloud) (1.16.0)

```
In [16]: import requests
         from bs4 import BeautifulSoup
         import pandas as pd
         import matplotlib.pyplot as plt
         import fsspec
```

```
In [17]: #Import required libraries
         import requests
         from bs4 import BeautifulSoup
         import pandas as pd
         import matplotlib.pyplot as plt
         import lxml

         import os
         import re
         import numpy as np
         import string
         import glob
         import nltk
         from wordcloud import WordCloud
         from collections import Counter, defaultdict
         from nltk.corpus import stopwords
         from nltk.tokenize import word_tokenize
         from nltk.stem import PorterStemmer
```

```
from string import punctuation

sw = stopwords.words("english")
```

```
In [18]: Dodgers_url = "https://www.mlbtraderumors.com/los-angeles-dodgers"
response = requests.get(Dodgers_url)
# Parse the HTML content using BeautifulSoup
Beaut_soup = BeautifulSoup(response.content, "html.parser")
articles = Beaut_soup.find_all("article")
```

```
In [19]: # Create a file to save the scraped data
file_path = "C:/Users/Grigor/OneDrive/Desktop/Master_Program/ADS509/Project/dodgers.txt"
with open(file_path, "w", encoding="utf-8") as file:
    # Initialize empty lists to store the extracted values
    titles = []
    authors = []
    dates = []
    contents = []

    # Iterate over the articles and extract the text
    for article in articles:
        # Extract the article title
        title_element = article.find("h2")
        title = title_element.text.strip() if title_element else ""
        titles.append(title)

        # Extract the article author
        author_element = article.find("span", class_="entry-author")
        author = author_element.text.strip() if author_element else ""
        authors.append(author)

        # Extract the article date
        date_element = article.find("time", class_="entry-time")
        date = date_element.text.strip() if date_element else ""
        dates.append(date)

        # Extract the article content
        content_element = article.find("div", class_="entry-content")
        content = content_element.text.strip() if content_element else ""
        contents.append(content)

        # Write the title and content to the file
        file.write("Title: " + title + "\n")
        file.write("Author: " + author + "\n")
        file.write("Date: " + date + "\n")
        file.write("Content: " + content + "\n")
        file.write("---\n")

print("Data saved to", file_path)

# Create a DataFrame from the extracted values
data = {
    "Title": titles,
    "Author": authors,
    "Date": dates,
    "Content": contents
}

df = pd.DataFrame(data)
```

```
df.dropna()
df.head()
```

Data saved to C:/Users/Grigor/OneDrive/Desktop/Master_Program/ADS509/Project/dodgers.txt

Out[19]:

	Title	Author	Date	Content
0	Friedman: Dodgers Increasingly Likely To Targe...	Anthony Franco	June 19, 2023	The Dodgers were off tonight, a chance for a r...
1	Lates On Daniel Hudson, Julio Urias	Nick Deeds	June 18, 2023	That's phenomenal news for San Francisco, as C...
2	Five Bats Improving Their Stock Ahead Of Free ...	Nick Deeds	June 18, 2023	With the midpoint of the 2023 regular season f...
3	Dodgers Notes: Taylor, Peralta, Urias, Hudson	Nick Deeds	June 17, 2023	Dodgers veterans Chris Taylor and David Peralt...
4	Dodgers Select Bryan Hudson	Nick Deeds	June 17, 2023	Ahead of tonight's game against the Giants, th...

In [20]: Attendance_url = "https://www.baseball-reference.com/teams/LAD/attend.shtml"

```
# Read the HTML table into a list of DataFrames
tables = pd.read_html(Attendance_url)

# Select the table of interest (index 0 in this case)
table = tables[0]
```

In [21]: Dodgers_Attendance = "C:/Users/Grigor/OneDrive/Desktop/Master_Program/ADS509/Project/Dodgers_Attendance

Out[21]: 'C:/Users/Grigor/OneDrive/Desktop/Master_Program/ADS509/Project/Dodgers_Attendance.csv'

In [22]: path = "C:/Users/Grigor/OneDrive/Desktop/Master_Program/ADS509/Project/"

```
Dodgers_Attendance = pd.DataFrame(pd.read_csv(path + 'Dodgers_Attendance.csv'))
Dodgers_Attendance
```

Out[22]:

	Year	Tm	Lg	W	L	Finish	Playoffs	Attendance	Attend/G	Rank	Est. Payr
0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	N
1	2023.0	Los Angeles Dodgers	NL West	36.0	28.0	2.0	NaN	1,437,473	47,916	1st of 15	\$198,971,6
2	2022.0	Los Angeles Dodgers	NL West	111.0	51.0	1.0	Lost NLDS (3-1)	3,861,408	47,672	1st of 15	\$261,180,1
3	2021.0	Los Angeles Dodgers	NL West	106.0	56.0	2.0	Lost NLCS (4-2)	2,804,693	34,626	1st of 15	\$248,108,3
4	2020.0	Los Angeles Dodgers	NL West	43.0	17.0	1.0	Won WS (4-2)	NaN	NaN	6th of 15	\$216,508,3
...
136	1888.0	Brooklyn Bridegrooms	AA	88.0	52.0	2.0	NaN	NaN	NaN	NaN	N
137	1887.0	Brooklyn Grays	AA	60.0	74.0	6.0	NaN	NaN	NaN	NaN	N
138	1886.0	Brooklyn Grays	AA	76.0	61.0	3.0	NaN	NaN	NaN	NaN	N
139	1885.0	Brooklyn Grays	AA	53.0	59.0	5.0	NaN	NaN	NaN	NaN	N
140	1884.0	Brooklyn Atlantics	AA	40.0	64.0	9.0	NaN	NaN	NaN	NaN	N

141 rows × 14 columns

```

In [23]: Dodgers_05_df = pd.DataFrame(pd.read_csv(path + 'Dodgers_Attendance.csv'))
Dodgers_05_df = Dodgers_05_df[Dodgers_05_df["Year"] > 2004]
Dodgers_05_df

```

Out[23]:

	Year	Tm	Lg	W	L	Finish	Playoffs	Attendance	Attend/G	Rank	Est. Payroll	
1	2023.0	Los Angeles Dodgers	NL West	36.0	28.0	2.0	NaN	1,437,473	47,916	1st of 15	\$198,971,667	1
2	2022.0	Los Angeles Dodgers	NL West	111.0	51.0	1.0	Lost NLDS (3-1)	3,861,408	47,672	1st of 15	\$261,180,131	1
3	2021.0	Los Angeles Dodgers	NL West	106.0	56.0	2.0	Lost NLCS (4-2)	2,804,693	34,626	1st of 15	\$248,108,334	1
4	2020.0	Los Angeles Dodgers	NL West	43.0	17.0	1.0	Won WS (4-2)	NaN	NaN	6th of 15	\$216,508,333	
5	2019.0	Los Angeles Dodgers	NL West	106.0	56.0	1.0	Lost NLDS (3-2)	3,974,309	49,066	1st of 15	\$193,553,333	
6	2018.0	Los Angeles Dodgers	NL West	92.0	71.0	1.0	Lost WS (4-1)	3,857,500	47,043	1st of 15	\$164,703,429	
7	2017.0	Los Angeles Dodgers	NL West	104.0	58.0	1.0	Lost WS (4-3)	3,765,856	46,492	1st of 15	\$201,466,263	
8	2016.0	Los Angeles Dodgers	NL West	91.0	71.0	1.0	Lost NLCS (4-2)	3,703,312	45,720	1st of 15	\$231,342,096	
9	2015.0	Los Angeles Dodgers	NL West	92.0	70.0	1.0	Lost NLDS (3-2)	3,764,815	46,479	1st of 15	\$265,140,429	
10	2014.0	Los Angeles Dodgers	NL West	94.0	68.0	1.0	Lost NLDS (3-1)	3,782,337	46,696	1st of 15	\$233,386,026	
11	2013.0	Los Angeles Dodgers	NL West	92.0	70.0	1.0	Lost NLCS (4-2)	3,743,527	46,216	1st of 15	\$254,161,000	
12	2012.0	Los Angeles Dodgers	NL West	86.0	76.0	2.0	NaN	3,324,246	41,040	3rd of 16	\$177,033,600	
13	2011.0	Los Angeles Dodgers	NL West	82.0	79.0	3.0	NaN	2,935,139	36,236	6th of 16	\$103,785,477	
14	2010.0	Los Angeles Dodgers	NL West	80.0	82.0	4.0	NaN	3,562,320	43,979	2nd of 16	\$95,358,016	
15	2009.0	Los Angeles Dodgers	NL West	95.0	67.0	1.0	Lost NLCS (4-1)	3,761,655	46,440	1st of 16	\$102,117,592	
16	2008.0	Los Angeles	NL West	84.0	78.0	1.0	Lost NLCS	3,730,553	46,056	2nd of 16	\$118,588,536	

	Year	Tm	Lg	W	L	Finish	Playoffs	Attendance	Attend/G	Rank	Est. Payroll	
		Dodgers					(4-1)					
17	2007.0	Los Angeles Dodgers	NL West	82.0	80.0	4.0	NaN	3,857,036	47,618	1st of 16	\$108,454,524	1
18	2006.0	Los Angeles Dodgers	NL West	88.0	74.0	2.0	Lost NLDS (3-0)	3,758,545	46,402	1st of 16	\$98,447,187	1
19	2005.0	Los Angeles Dodgers	NL West	71.0	91.0	4.0	NaN	3,603,646	44,489	1st of 16	\$83,039,000	

```
In [24]: #Identify any noise in the data
RE_SUSPICIOUS = re.compile(r'([<>{}\\[\]\'])')

def impurity(text, min_len=10):
    """returns the share of suspicious characters in a text"""
    if text == None or len(text) < min_len:
        return 0
    else:
        return len(RE_SUSPICIOUS.findall(text))/len(text)
```

```
In [25]: df['Content'].apply
```



```

Out[25]: <bound method Series.apply of 0      The Dodgers were off tonight, a chance for a r...
1      That's phenomenal news for San Francisco, as C...
2      With the midpoint of the 2023 regular season f...
3      Dodgers veterans Chris Taylor and David Peralta...
4      Ahead of tonight's game against the Giants, th...
5      The Dodgers announced a handful of moves befor...
6      The Dodgers are going to select right-hander E...
7      The Dodgers announced that left-hander Adam Ko...
8      Right-hander Alex Reyes won't pitch in 2023 af...
9      The Dodgers have selected the contract of left...
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
Name: Content, dtype: object>

```

```

In [26]: pd.options.display.max_colwidth = 100 ###
# add new column to data frame
df['impurity'] = df['Content'].apply(impurity, min_len=10)

# get the top 3 records
df[['Content', 'impurity']].sort_values(by='impurity', ascending=False).head(3)

```

```

Out[26]:

```

	Content	impurity
6	The Dodgers are going to select right-hander Emmet Sheehan to their roster, reports Juan Toribio...	0.000877
0	The Dodgers were off tonight, a chance for a reset after a disastrous series that saw them swept...	0.000000
30		0.000000

```
In [27]: # Example function to normalize the text in the "Contents" column
def normalize_text(text):
    # Lowercase the text
    text = text.lower()

    # Remove punctuation
    text = text.translate(str.maketrans("", "", string.punctuation))

    # Remove special characters and digits
    text = re.sub(r"[^a-zA-Z]", " ", text)

    # Tokenize the text
    tokens = word_tokenize(text)

    # Remove stopwords
    stop_words = set(stopwords.words("english"))
    tokens = [token for token in tokens if token not in stop_words]

    # Perform stemming
    stemmer = PorterStemmer()
    tokens = [stemmer.stem(token) for token in tokens]

    return tokens

# Apply normalization to the "Contents" column
df["Contents_Normalized"] = df["Content"].apply(normalize_text)

# Print the head of the DataFrame with the normalized contents
df[["Content", "Contents_Normalized"]].head()
```

```
Out[27]:
```

	Content	Contents_Normalized
0	The Dodgers were off tonight, a chance for a reset after a disastrous series that saw them swept...	[dodger, tonight, chanc, reset, disastr, seri, saw, swept, archriv, giant, pull, past, lo, angel...
1	That's phenomenal news for San Francisco, as Cobb has been among the club's most reliable starte...	[phenomen, news, san, francisco, cobb, among, club, reliabl, starter, season, era, fip, inning, ...
2	With the midpoint of the 2023 regular season fast approaching, the 2023-24 free agent class is b...	[midpoint, regular, season, fast, approach, free, agent, class, begin, solidifi, come, class, lo...
3	Dodgers veterans Chris Taylor and David Peralta are both out of the lineup today after exiting y...	[dodger, veteran, chri, taylor, david, peralta, lineup, today, exit, yesterday, game, giant, bil...
4	Ahead of tonight's game against the Giants, the Dodgers have selected the contract of left-hande...	[ahead, tonight, game, giant, dodger, select, contract, lefthand, bryan, hudson, correspond, mov...

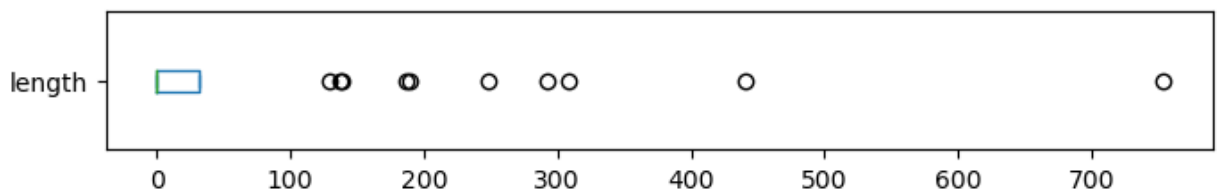
```
In [28]: df['length'] = df['Contents_Normalized'].str.len()
df.head()
```

Out[28]:

	Title	Author	Date	Content	impurity	Contents_Normalized	length
0	Friedman: Dodgers Increasingly Likely To Target Pitching Help At Deadline	Anthony Franco	June 19, 2023	The Dodgers were off tonight, a chance for a reset after a disastrous series that saw them swept...	0.0	[dodger, tonight, chanc, reset, disastr, seri, saw, swept, archiv, giant, pull, past, lo, angel...	441
1	Lates On Daniel Hudson, Julio Urias	Nick Deeds	June 18, 2023	That's phenomenal news for San Francisco, as Cobb has been among the club's most reliable starte...	0.0	[phenomen, news, san, francisco, cobb, among, club, reliabl, starter, season, era, fip, inning, ...	137
2	Five Bats Improving Their Stock Ahead Of Free Agency	Nick Deeds	June 18, 2023	With the midpoint of the 2023 regular season fast approaching, the 2023-24 free agent class is b...	0.0	[midpoint, regular, season, fast, approach, free, agent, class, begin, solidifi, come, class, lo...	753
3	Dodgers Notes: Taylor, Peralta, Urias, Hudson	Nick Deeds	June 17, 2023	Dodgers veterans Chris Taylor and David Peralta are both out of the lineup today after exiting y...	0.0	[dodger, veteran, chri, taylor, david, peralta, lineup, today, exit, yesterday, game, giant, bil...	292
4	Dodgers Select Bryan Hudson	Nick Deeds	June 17, 2023	Ahead of tonight's game against the Giants, the Dodgers have selected the contract of left- hande...	0.0	[ahead, tonight, game, giant, dodger, select, contract, lefthand, bryan, hudson, correspond, mov...	189

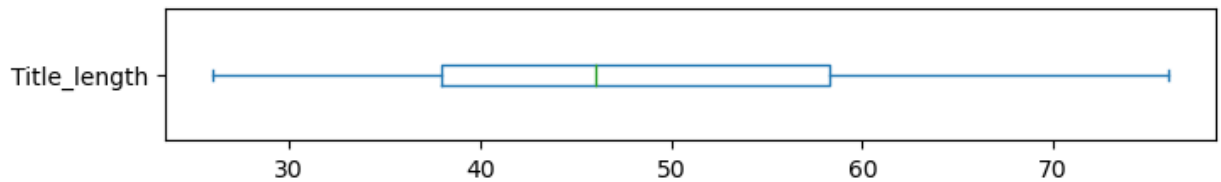
In [29]: `df['length'].plot(kind='box', vert=False, figsize=(8, 1))`

Out[29]: <Axes: >



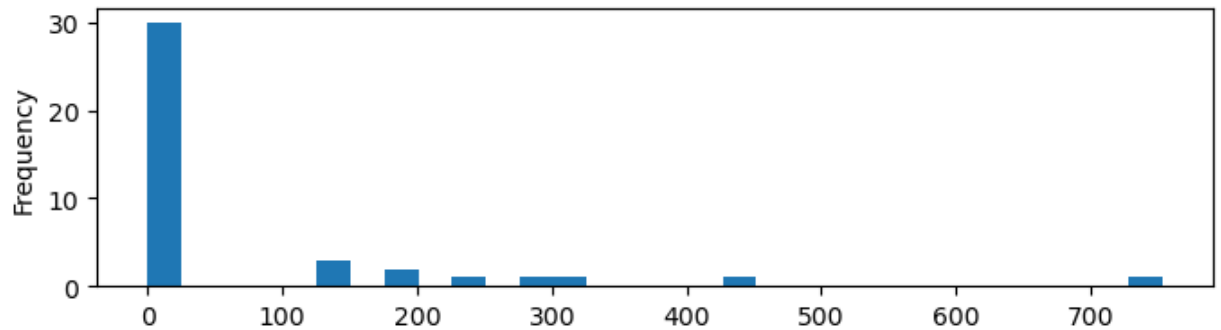
In [30]: `df['Title_length'] = df['Title'].str.len()
df['Title_length'].plot(kind='box', vert=False, figsize=(8, 1))`

Out[30]: <Axes: >



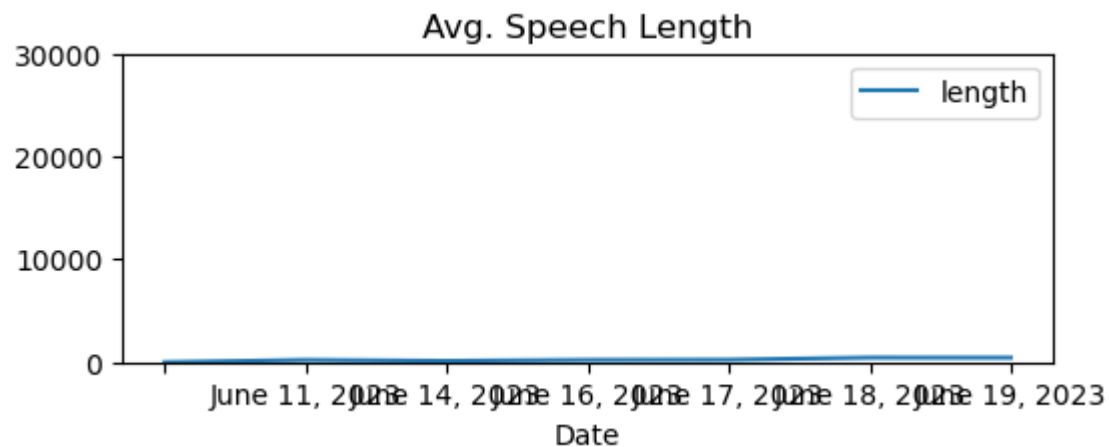
```
In [31]: df['length'].plot(kind='hist', bins=30, figsize=(8,2))
```

```
Out[31]: <Axes: ylabel='Frequency'>
```



```
In [32]: df.groupby('Date').agg({'length': 'mean'}) \
        .plot(title="Avg. Speech Length", ylim=(0,30000), figsize=(6,2))
```

```
Out[32]: <Axes: title={'center': 'Avg. Speech Length'}, xlabel='Date'>
```

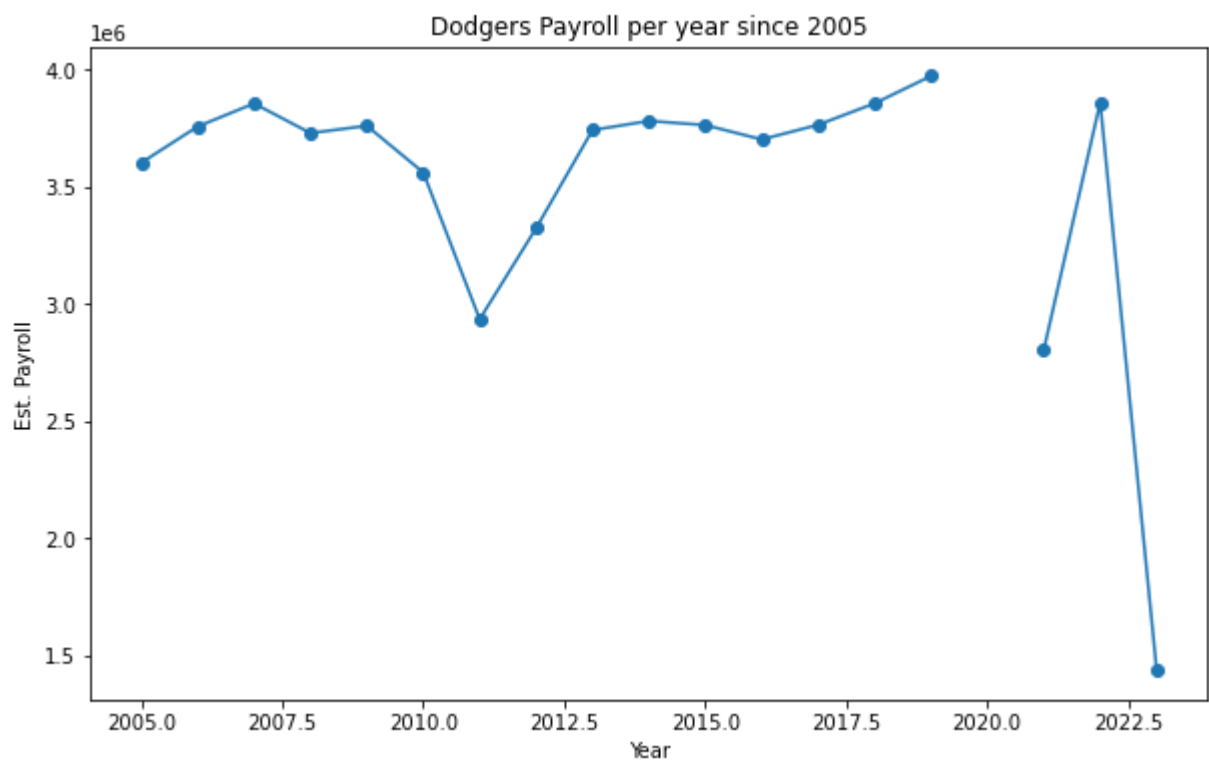


```
In [12]: Dodgers_05_df['Est. Payroll'] = Dodgers_05_df['Est. Payroll'].replace({'\$':''}, regex=True)
Dodgers_05_df['Est. Payroll'] = Dodgers_05_df['Est. Payroll'].replace({'\,':''}, regex=True)
Dodgers_05_df['Attendance'] = Dodgers_05_df['Attendance'].replace({'\,':''}, regex=True)
```

```
In [54]: #plot data
Dodgers_05_df['Attendance'] = Dodgers_05_df['Attendance'].astype(float)
Dodgers_05_df.dtypes
# Plotting
plt.figure(figsize=(10, 6))

# Plotting Attendance over the years
plt.plot(Dodgers_05_df['Year'], Dodgers_05_df['Attendance'], marker='o')
plt.xlabel('Year')
plt.ylabel('Est. Payroll')
plt.title('Dodgers Attendance per year since 2005')
```

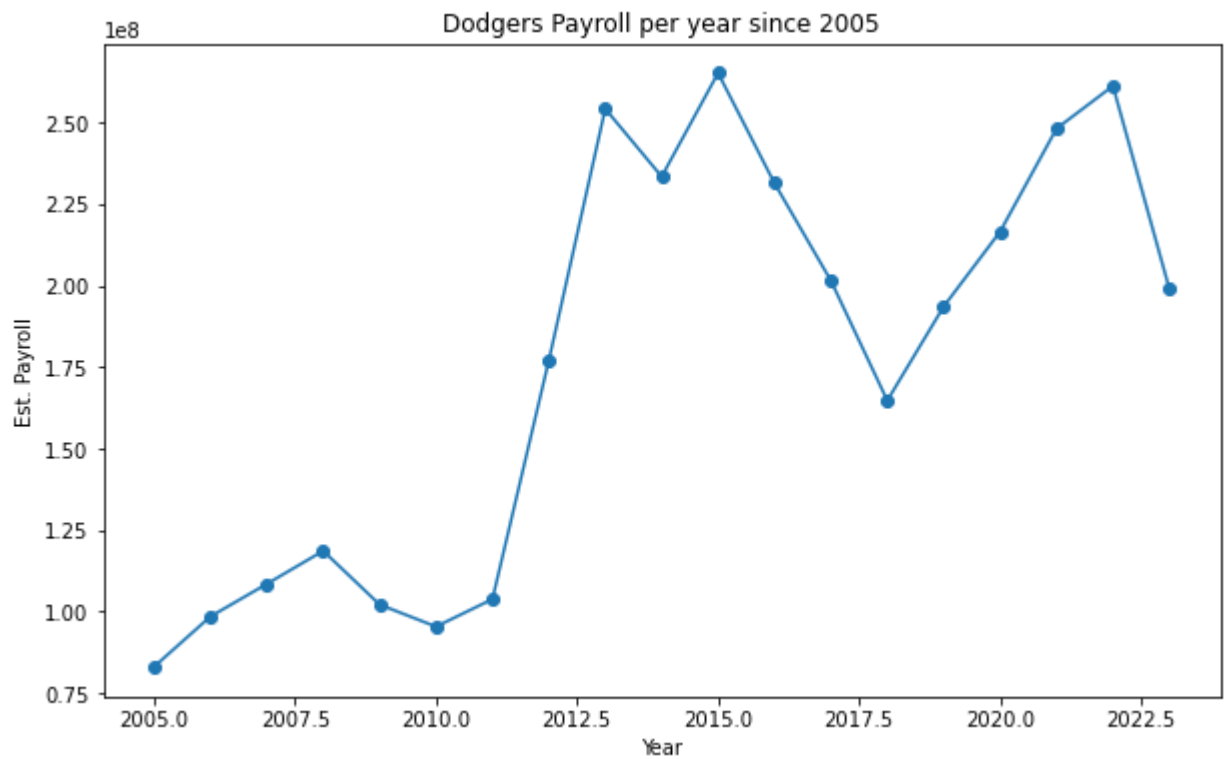
Out[54]: Text(0.5, 1.0, 'Dodgers Payroll per year since 2005')



```
In [50]: #plot data
Dodgers_05_df['Est. Payroll'] = Dodgers_05_df['Est. Payroll'].astype(float)
Dodgers_05_df.dtypes
# Plotting
plt.figure(figsize=(10, 6))

# Plotting Attendance over the years
plt.plot(Dodgers_05_df['Year'], Dodgers_05_df['Est. Payroll'], marker='o')
plt.xlabel('Year')
plt.ylabel('Est. Payroll')
plt.title('Dodgers Payroll per year since 2005')
```

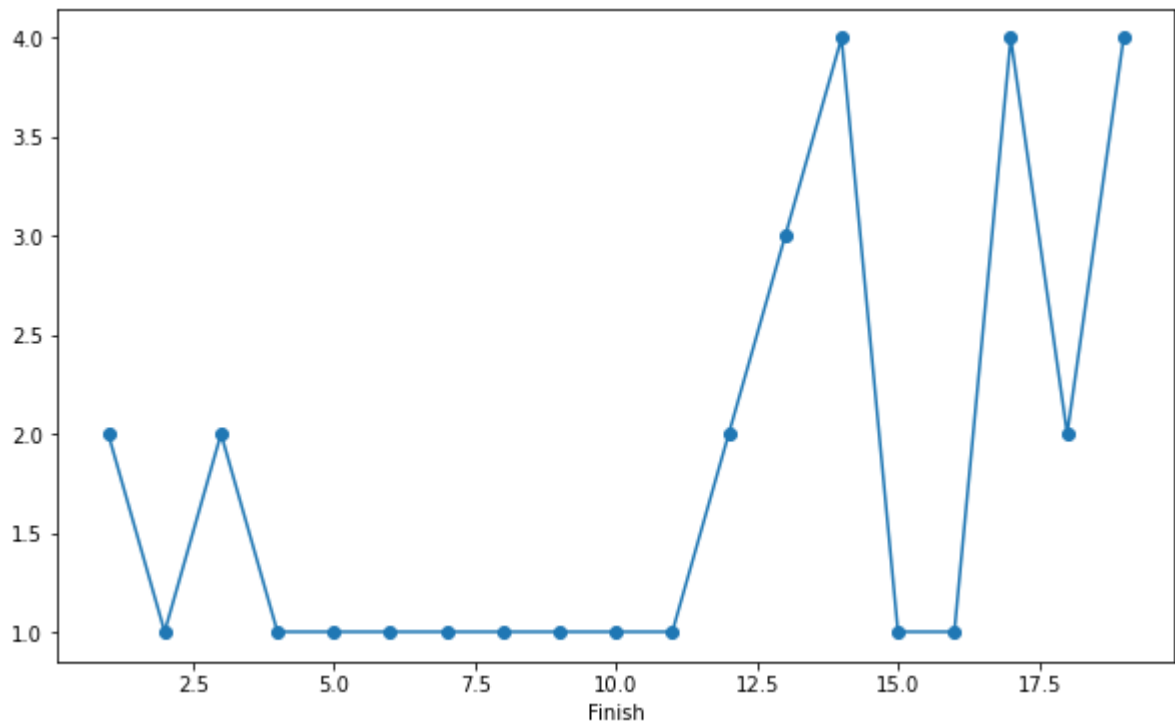
Out[50]: Text(0.5, 1.0, 'Dodgers Payroll per year since 2005')



```
In [45]: plt.figure(figsize=(10, 6))

# Plotting Attendance over the years
plt.plot(Dodgers_05_df['Finish'], marker='o')
plt.xlabel('Finish')
```

```
Out[45]: Text(0.5, 0, 'Finish')
```

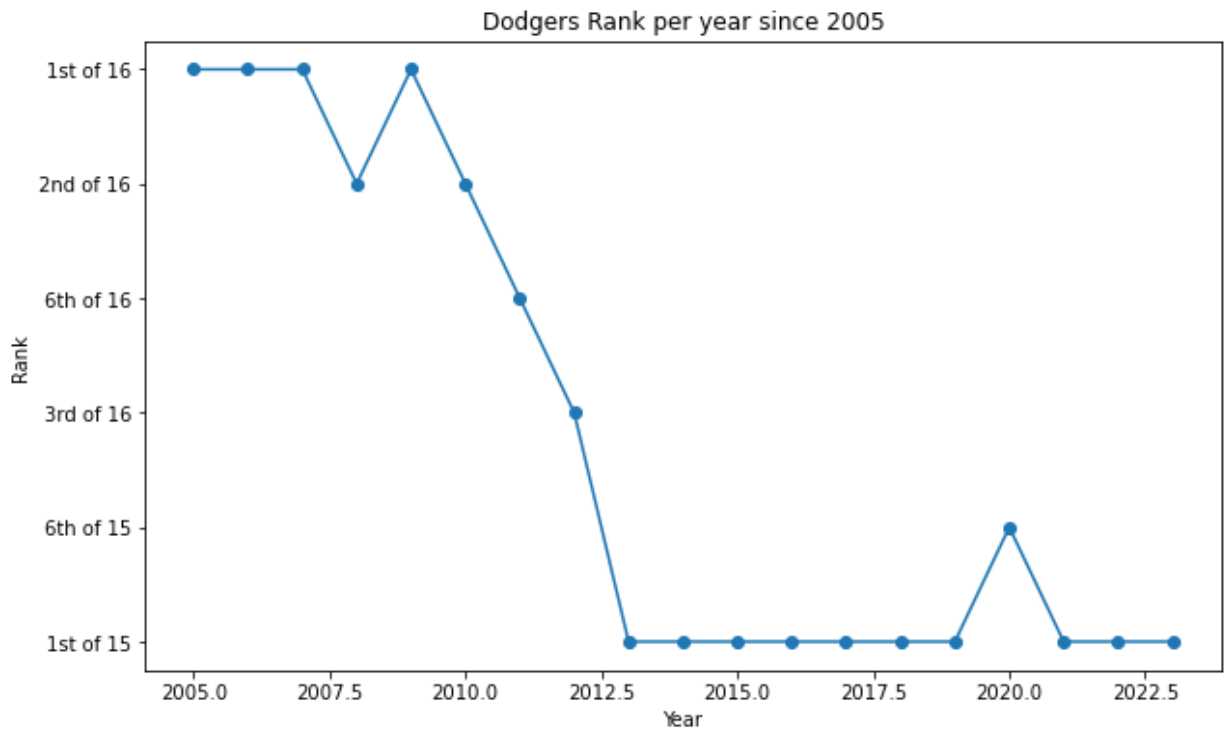


```
In [55]: plt.figure(figsize=(10, 6))

# Plotting Attendance over the years
```

```
plt.plot(Dodgers_05_df['Year'], Dodgers_05_df['Rank'], marker='o')
plt.xlabel('Year')
plt.ylabel('Rank')
plt.title('Dodgers Rank per year since 2005')
```

Out[55]: Text(0.5, 1.0, 'Dodgers Rank per year since 2005')



```
In [59]: corr = Dodgers_05_df["Est. Payroll"].corr(Dodgers_05_df['Finish'])
print(corr)

-0.6477777557448725
```

In []: