NOTE###

Uncomment the cell below to run 'newsapi''transformer','torch' installation

```
!pip install newsapi-python transformers torch
#print("Necessary libraries installed.")
```

```
Requirement already satisfied: newsapi-python in c:\users\lenovo\anaconda3\lib\site-packages (0.2.7)
Requirement already satisfied: transformers in c:\users\lenovo\anaconda3\lib\site-packages (4.57.6)
Requirement already satisfied: torch in c:\users\lenovo\anaconda3\lib\site-packages (2.9.1)
Requirement already satisfied: requests<3.0.0 in c:\users\lenovo\anaconda3\lib\site-packages (from newsapi-python) (2.32.3)
Requirement already satisfied: charset-normalizer<4,>=2 in c:\users\lenovo\anaconda3\lib\site-packages (from requests<3.0.0->new
Requirement already satisfied: idna<4,>=2.5 in c:\users\lenovo\anaconda3\lib\site-packages (from requests<3.0.0->newsapi-python)
Requirement already satisfied: urllib3<3,>=1.21.1 in c:\users\lenovo\anaconda3\lib\site-packages (from requests<3.0.0->newsapi-p
Requirement already satisfied: certifi>=2017.4.17 in c:\users\lenovo\anaconda3\lib\site-packages (from requests<3.0.0->newsapi-p
Requirement already satisfied: filelock in c:\users\lenovo\anaconda3\lib\site-packages (from transformers) (3.17.0)
Requirement already satisfied: huggingface-hub<1.0,>=0.34.0 in c:\users\lenovo\anaconda3\lib\site-packages (from transformers) (
Requirement already satisfied: numpy>=1.17 in c:\users\lenovo\anaconda3\lib\site-packages (from transformers) (2.3.5)
Requirement already satisfied: packaging>=20.0 in c:\users\lenovo\anaconda3\lib\site-packages (from transformers) (24.2)
Requirement already satisfied: pyyaml>=5.1 in c:\users\lenovo\anaconda3\lib\site-packages (from transformers) (6.0.2)
Requirement already satisfied: regex!=2019.12.17 in c:\users\lenovo\anaconda3\lib\site-packages (from transformers) (2024.11.6)
Requirement already satisfied: tokenizers<=0.23.0,>=0.22.0 in c:\users\lenovo\anaconda3\lib\site-packages (from transformers) (0
Requirement already satisfied: safetensors>=0.4.3 in c:\users\lenovo\anaconda3\lib\site-packages (from transformers) (0.7.0)
Requirement already satisfied: tqdm>=4.27 in c:\users\lenovo\anaconda3\lib\site-packages (from transformers) (4.67.1)
Requirement already satisfied: fsspec>=2023.5.0 in c:\users\lenovo\anaconda3\lib\site-packages (from huggingface-hub<1.0,>=0.34.
Requirement already satisfied: typing-extensions>=3.7.4.3 in c:\users\lenovo\anaconda3\lib\site-packages (from huggingface-hub<1
Requirement already satisfied: sympy>=1.13.3 in c:\users\lenovo\anaconda3\lib\site-packages (from torch) (1.13.3)
Requirement already satisfied: networkx>=2.5.1 in c:\users\lenovo\anaconda3\lib\site-packages (from torch) (3.4.2)
Requirement already satisfied: jinja2 in c:\users\lenovo\anaconda3\lib\site-packages (from torch) (3.1.6)
Requirement already satisfied: setuptools in c:\users\lenovo\anaconda3\lib\site-packages (from torch) (80.9.0)
Requirement already satisfied: mpmath<1.4,>=1.1.0 in c:\users\lenovo\anaconda3\lib\site-packages (from sympy>=1.13.3->torch) (1.
Requirement already satisfied: colorama in c:\users\lenovo\anaconda3\lib\site-packages (from tqdm>=4.27->transformers) (0.4.6)
Requirement already satisfied: MarkupSafe>=2.0 in c:\users\lenovo\anaconda3\lib\site-packages (from jinja2->torch) (3.0.2)
```

```
import numpy as np
import pandas as pd
from newsapi import NewsApiClient
import matplotlib.pyplot as plt
from transformers import AutoTokenizer, AutoModelForSequenceClassification, pipeline
import torch
import seaborn as sns
```

```
# IMPORTANT: Replace 'YOUR_NEWS_API_KEY' with your actual News API key.
# It's recommended to store API keys securely, e.g., in environment variables or Colab secrets.
NEWS_API_KEY = 'b341f7e3c7484523947c716f46943cd6'

# Initialize the News API client
newsapi = NewsApiClient(api_key=NEWS_API_KEY)

print("NewsAPI client initialized. Please ensure you've replaced 'YOUR_NEWS_API_KEY' with your actual API key.")
```

```
NewsAPI client initialized. Please ensure you've replaced 'YOUR_NEWS_API_KEY' with your actual API key.
```

```
query = 'Apple OR AAPL stock'
language = 'en'
sort_by = 'relevancy' # or 'publishedAt'
page_size = 100 # Maximum articles per request

try:
    # Fetch articles
    all_articles = newsapi.get_everything(
        q=query,
        language=language,
        sort_by=sort_by,
        page_size=page_size
    )

    articles_data = []
    if all_articles['status'] == 'ok' and all_articles['articles']:
```

```
        for article in all_articles['articles']:
            articles_data.append({
                'source_id': article['source']['id'],
                'source_name': article['source']['name'],
                'author': article['author'],
                'title': article['title'],
                'description': article['description'],
                'url': article['url'],
                'urlToImage': article['urlToImage'],
                'publishedAt': article['publishedAt'],
                'content': article['content']
            })
        df_articles = pd.DataFrame(articles_data)
        print(f"Successfully fetched {len(df_articles)} articles.")
        print("First 5 rows of the articles DataFrame:")
        print(df_articles.head())
    else:
        print("No articles found or API call failed.")
        df_articles = pd.DataFrame()

except Exception as e:
    print(f"An error occurred while fetching articles: {e}")
    df_articles = pd.DataFrame()
```

```
Successfully fetched 90 articles.
First 5 rows of the articles DataFrame:
          source_id       source_name                          author  \
0  business-insider  Business Insider  Dominick Reuter,Robert Scammell
1              None     24/7 Wall St.                  Omor Ibne Ehsan
2  business-insider  Business Insider                      Jordan Hart
3              None         MacRumors                 Hartley Charlton
4              None            Kotaku                 Brittany Vincent

                                           title  \
0  Apple strikes deal with Google's Gemini to pow...
1  Buffett’s $24 Billion Selling Spree: The 6 Sto...
2  Here's how much Apple CEO Tim Cook made last year
3  Apple CEO Tim Cook Buys $3 Million of Nike Shares
4  Apple Quietly Drops the Beats Studio Pro by 49...

                                     description  \
0  Apple and Alphabet have reached a deal that wi...
1  Investors began to notice that Warren Buffett ...
2  Tim Cook's salary was a drop in the bucket of ...
3  Apple CEO Tim Cook disclosed a roughly $3 mill...
4  These might very well become your new favorite...

                                             url  \
0  https://www.businessinsider.com/apple-google-g...
1  https://247wallst.com/investing/2025/12/23/buf...
2  https://www.businessinsider.com/apple-ceo-tim-...
3  https://www.macrumors.com/2025/12/24/tim-cook-...
4  https://kotaku.com/apple-quietly-drops-the-bea...

                                       urlToImage          publishedAt  \
0  https://i.insider.com/6965166604eda4732f2ee3db...  2026-01-12T16:32:48Z
1  https://s.yimg.com/ny/api/res/1.2/4f8G.4_vfcPu...  2025-12-23T14:02:07Z
2  https://i.insider.com/6961245704eda4732f2ec781...  2026-01-09T17:52:23Z
3  https://images.macrumors.com/t/0_RX2LavEjJuQ-0...  2025-12-24T16:41:25Z
4  https://kotaku.com/app/uploads/2025/12/beatstp...  2025-12-25T13:05:26Z

                                         content
0  Apple and Alphabet have reached a big AI deal....
1  Mark Wilson / Getty Images\r\n<ul><li>Buffett ...
2  Tim Cook made slightly less in 2025 than he di...
3  Apple CEO Tim Cook disclosed a roughly $3 mill...
4  Looking for a new pair of headphones tha sound...
```

```
query = 'Apple OR AAPL stock'
language = 'en'
sort_by = 'relevancy' # or 'publishedAt'
page_size = 100 # Maximum articles per request

try:
    # Fetch articles
    all_articles = newsapi.get_everything(
        q=query,
        language=language,
```

```
        sort_by=sort_by,
        page_size=page_size
    )

    articles_data = []
    if all_articles['status'] == 'ok' and all_articles['articles']:
        for article in all_articles['articles']:
            articles_data.append({
                'source_id': article['source']['id'],
                'source_name': article['source']['name'],
                'author': article['author'],
                'title': article['title'],
                'description': article['description'],
                'url': article['url'],
                'urlToImage': article['urlToImage'],
                'publishedAt': article['publishedAt'],
                'content': article['content']
            })
        df_articles = pd.DataFrame(articles_data)
        print(f"Successfully fetched {len(df_articles)} articles.")
        print("First 5 rows of the articles DataFrame:")
        print(df_articles.head())
    else:
        print("No articles found or API call failed.")
        df_articles = pd.DataFrame()

except Exception as e:
    print(f"An error occurred while fetching articles: {e}")
    df_articles = pd.DataFrame()
```

```
Successfully fetched 90 articles.
First 5 rows of the articles DataFrame:
          source_id       source_name                          author  \
0  business-insider  Business Insider  Dominick Reuter,Robert Scammell
1              None     24/7 Wall St.                  Omor Ibne Ehsan
2  business-insider  Business Insider                      Jordan Hart
3              None         MacRumors                 Hartley Charlton
4              None            Kotaku                 Brittany Vincent

                                          title  \
0  Apple strikes deal with Google's Gemini to pow...
1  Buffett’s $24 Billion Selling Spree: The 6 Sto...
2  Here's how much Apple CEO Tim Cook made last year
3  Apple CEO Tim Cook Buys $3 Million of Nike Shares
4  Apple Quietly Drops the Beats Studio Pro by 49...

                                    description  \
0  Apple and Alphabet have reached a deal that wi...
1  Investors began to notice that Warren Buffett ...
2  Tim Cook's salary was a drop in the bucket of ...
3  Apple CEO Tim Cook disclosed a roughly $3 mill...
4  These might very well become your new favorite...

                                           url  \
0  https://www.businessinsider.com/apple-google-g...
1  https://247wallst.com/investing/2025/12/23/buf...
2  https://www.businessinsider.com/apple-ceo-tim-...
3  https://www.macrumors.com/2025/12/24/tim-cook-...
4  https://kotaku.com/apple-quietly-drops-the-bea...

                                        urlToImage           publishedAt  \
0  https://i.insider.com/6965166604eda4732f2ee3db...  2026-01-12T16:32:48Z
1  https://s.yimg.com/ny/api/res/1.2/4f8G.4_vfcPu...  2025-12-23T14:02:07Z
2  https://i.insider.com/6961245704eda4732f2ec781...  2026-01-09T17:52:23Z
3  https://images.macrumors.com/t/0_RX2LavEjJuQ-0...  2025-12-24T16:41:25Z
4  https://kotaku.com/app/uploads/2025/12/beatstp...  2025-12-25T13:05:26Z

                                         content
0  Apple and Alphabet have reached a big AI deal....
1  Mark Wilson / Getty Images\r\n<ul><li>Buffett ...
2  Tim Cook made slightly less in 2025 than he di...
3  Apple CEO Tim Cook disclosed a roughly $3 mill...
4  Looking for a new pair of headphones tha sound...
```

```
columns_to_combine = ['title', 'description', 'content']

# Fill missing values with empty strings in the selected columns
for col in columns_to_combine:
    if col in df_articles.columns:
        df_articles[col] = df_articles[col].fillna('')
```

```
    else:
        print(f"Warning: Column '{col}' not found in df_articles. Skipping.")

# Concatenate the cleaned text columns into a new 'text_for_sentiment' column
# Use a period as a separator for better readability and sentence separation
text_parts = []
if 'title' in df_articles.columns: text_parts.append(df_articles['title'])
if 'description' in df_articles.columns: text_parts.append(df_articles['description'])
if 'content' in df_articles.columns: text_parts.append(df_articles['content'])

if text_parts:
    df_articles['text_for_sentiment'] = text_parts[0]
    for i in range(1, len(text_parts)):
        df_articles['text_for_sentiment'] = df_articles['text_for_sentiment'] + '. ' + text_parts[i]
else:
    df_articles['text_for_sentiment'] = ''

# Display the first few rows of the DataFrame with the new column
print("DataFrame with 'text_for_sentiment' column:")
print(df_articles[['title', 'description', 'content', 'text_for_sentiment']].head())
```

```
DataFrame with 'text_for_sentiment' column:
                                               title  \
0  Apple strikes deal with Google's Gemini to pow...
1  Buffett’s $24 Billion Selling Spree: The 6 Sto...
2  Here's how much Apple CEO Tim Cook made last year
3  Apple CEO Tim Cook Buys $3 Million of Nike Shares
4  Apple Quietly Drops the Beats Studio Pro by 49...

                                         description  \
0  Apple and Alphabet have reached a deal that wi...
1  Investors began to notice that Warren Buffett ...
2  Tim Cook's salary was a drop in the bucket of ...
3  Apple CEO Tim Cook disclosed a roughly $3 mill...
4  These might very well become your new favorite...

                                           content  \
0  Apple and Alphabet have reached a big AI deal....
1  Mark Wilson / Getty Images\r\n<ul><li>Buffett ...
2  Tim Cook made slightly less in 2025 than he di...
3  Apple CEO Tim Cook disclosed a roughly $3 mill...
4  Looking for a new pair of headphones tha sound...

                                text_for_sentiment
0  Apple strikes deal with Google's Gemini to pow...
1  Buffett’s $24 Billion Selling Spree: The 6 Sto...
2  Here's how much Apple CEO Tim Cook made last y...
3  Apple CEO Tim Cook Buys $3 Million of Nike Sha...
4  Apple Quietly Drops the Beats Studio Pro by 49...
```

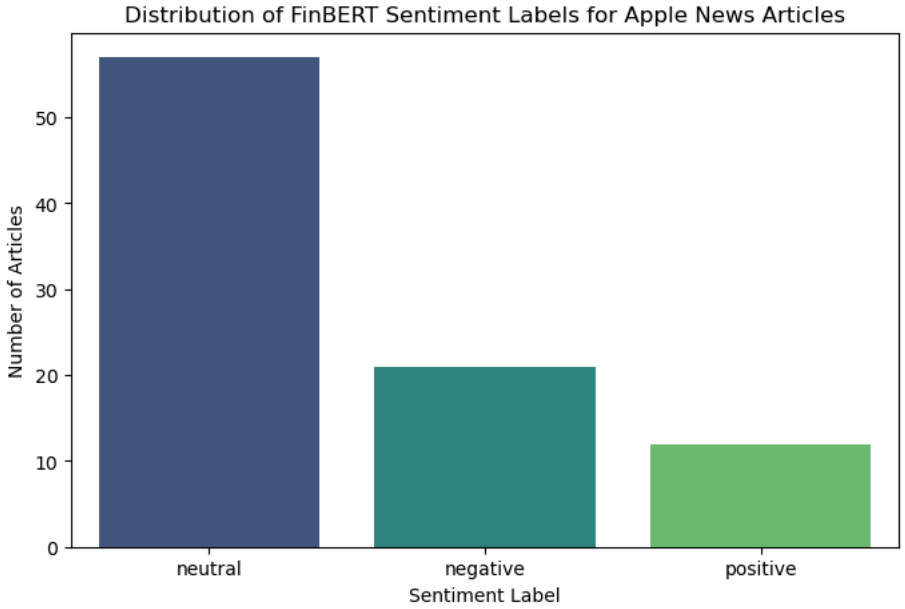
```
# 2. Define the FinBERT model name
model_name = "ProsusAI/finbert"

# 3. Load the tokenizer and model
tokenizer = AutoTokenizer.from_pretrained(model_name)
model = AutoModelForSequenceClassification.from_pretrained(model_name)

# 4. Create a sentiment analysis pipeline
sentiment_pipeline = pipeline("sentiment-analysis", model=model, tokenizer=tokenizer)

# 5. Define a function to get FinBERT sentiment
def get_finbert_sentiment(text):
    if not isinstance(text, str) or not text.strip():
        return 'neutral', 0.0
    try:
        # The FinBERT model outputs a list of dictionaries, e.g., [{'label': 'positive', 'score': 0.99}]
        result = sentiment_pipeline(text[:512]) # Truncate to avoid exceeding model's max input length
        label = result[0]['label']
        score = result[0]['score']
        return label, score
    except Exception as e:
        print(f"Error processing text: {text[:100]}... Error: {e}")
        return 'neutral', 0.0

# 6. Apply the get_finbert_sentiment function to the 'text_for_sentiment' column
# Note: This might take some time depending on the number of articles and their length.
df_articles[['finbert_sentiment_label', 'finbert_sentiment_score']] = df_articles['text_for_sentiment'].apply(lambda x: pd.Seri

# 7. Verify the new columns by displaying the head of df_articles
```

```
print("DataFrame with FinBERT sentiment scores:")
print(df_articles[['title', 'finbert_sentiment_label', 'finbert_sentiment_score']].head())
```

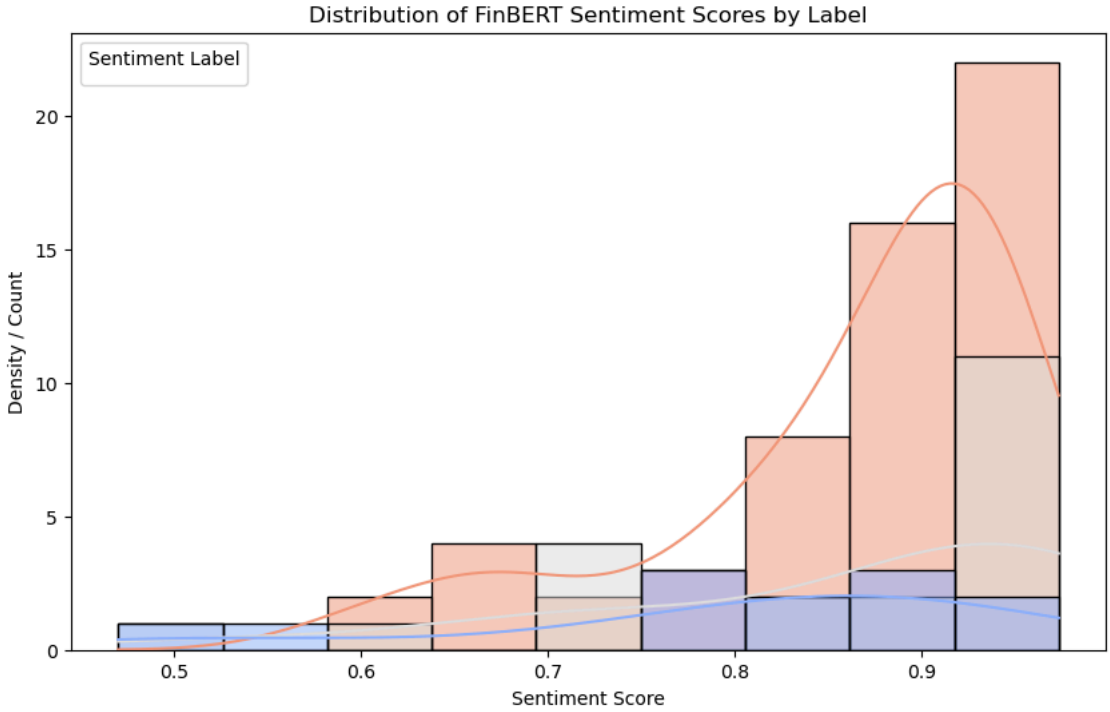
```
Xet Storage is enabled for this repo, but the 'hf_xet' package is not installed. Falling back to regular HTTP download. For bett
pytorch_model.bin:  62%|######2   | 273M/438M [00:00<?, ?B/s]
Device set to use cpu
DataFrame with FinBERT sentiment scores:
                                               title finbert_sentiment_label  \
0  Apple strikes deal with Google's Gemini to pow...                positive
1  Buffett’s $24 Billion Selling Spree: The 6 Sto...                negative
2  Here's how much Apple CEO Tim Cook made last year                negative
3  Apple CEO Tim Cook Buys $3 Million of Nike Shares                 neutral
4  Apple Quietly Drops the Beats Studio Pro by 49...                 neutral

   finbert_sentiment_score
0                 0.884715
1                 0.954643
2                 0.918814
3                 0.923582
4                 0.668575
Xet Storage is enabled for this repo, but the 'hf_xet' package is not installed. Falling back to regular HTTP download. For bett
```

```
# 1. Analyze the distribution of FinBERT sentiment labels
sentiment_counts = df_articles['finbert_sentiment_label'].value_counts()
print("\nSentiment Label Distribution:")
print(sentiment_counts)

# 2. Visualize the sentiment label distribution
plt.figure(figsize=(8, 5))
sns.barplot(x=sentiment_counts.index, y=sentiment_counts.values, palette='viridis')
plt.title('Distribution of FinBERT Sentiment Labels for Apple News Articles')
plt.xlabel('Sentiment Label')
plt.ylabel('Number of Articles')
plt.show()

# 3. Visualize the distribution of sentiment scores
plt.figure(figsize=(10, 6))
sns.histplot(df_articles, x='finbert_sentiment_score', hue='finbert_sentiment_label', kde=True, palette='coolwarm')
plt.title('Distribution of FinBERT Sentiment Scores by Label')
plt.xlabel('Sentiment Score')
plt.ylabel('Density / Count')
plt.legend(title='Sentiment Label')
plt.show()

print("Sentiment analysis distribution and score visualization complete.")
```

```
Sentiment Label Distribution:
finbert_sentiment_label
neutral     57
negative    21
positive    12
Name: count, dtype: int64
C:\Users\Lenovo\AppData\Local\Temp\ipykernel_22472\3454559281.py:8: FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be removed in v0.14.0. Assign the `x` variable to `hue` and se

  sns.barplot(x=sentiment_counts.index, y=sentiment_counts.values, palette='viridis')
```




```
C:\Users\Lenovo\AppData\Local\Temp\ipykernel_22472\3454559281.py:20: UserWarning: No artists with labels found to put in legend
  plt.legend(title='Sentiment Label')
```




```
Sentiment analysis distribution and score visualization complete.
```

```
print("\nDescriptive statistics for FinBERT Sentiment Scores:")
print(df_articles['finbert_sentiment_score'].describe())

# Display articles with highest positive sentiment
print("\nTop 5 Most Positive Articles:")
positive_articles = df_articles[df_articles['finbert_sentiment_label'] == 'positive'].sort_values(by='finbert_sentiment_score',
if not positive_articles.empty:
```

```
    print(positive_articles[['title', 'finbert_sentiment_label', 'finbert_sentiment_score', 'url']].head())
else:
    print("No positive articles found.")

# Display articles with lowest (most negative) sentiment
print("\nTop 5 Most Negative Articles:")
negative_articles = df_articles[df_articles['finbert_sentiment_label'] == 'negative'].sort_values(by='finbert_sentiment_score',
if not negative_articles.empty:
    # Assuming score is confidence for the assigned label, a higher score for 'negative' means more confidently negative.
    print(negative_articles[['title', 'finbert_sentiment_label', 'finbert_sentiment_score', 'url']].head())
else:
    print("No negative articles found.")

print("\nSummary of key findings complete.")
```

```
Descriptive statistics for FinBERT Sentiment Scores:
count    89.000000
mean      0.843971
std       0.123041
min       0.455196
25%       0.796448
50%       0.889979
75%       0.938135
max       0.973493
Name: finbert_sentiment_score, dtype: float64

Top 5 Most Positive Articles:
                                          title finbert_sentiment_label  \
51  Tesla stock pops as Robotaxi testing with no s...               positive
14  Nike shares are up after Apple CEO Tim Cook bu...               positive
72  Palantir Stock Jumps Amid Contract Wins and AI...               positive
24               Why small hedge funds ruled in 2025               positive
9   Apple spent 2025 setting itself up for the fut...               positive

    finbert_sentiment_score                                          url
51                 0.951494  https://finance.yahoo.com/news/tesla-stock-pop...
14                 0.942736  https://www.reutersconnect.com/detail?id=tag:r...
72                 0.922132  https://finance.yahoo.com/news/palantir-stock-...
24                 0.917266  https://www.businessinsider.com/small-hedge-fu...
9                  0.886936  https://finance.yahoo.com/news/apple-spent-202...

Top 5 Most Negative Articles:
                                          title finbert_sentiment_label  \
60  MongoDB, AppLovin, and SoundHound AI Stocks Tr...               negative
70  PepsiCo (PEP) Stock Sinks As Market Gains: Her...               negative
18  Oracle’s AI Push Is Leading to Its Worst Quart...               negative
52  Nvidia Shares Slide After Massive H200 Orders ...               negative
1   Buffett’s $24 Billion Selling Spree: The 6 Sto...               negative

    finbert_sentiment_score                                          url
60                 0.973493  https://finance.yahoo.com/news/mongodb-applovi...
70                 0.970650  https://finance.yahoo.com/news/pepsico-pep-sto...
18                 0.968430  https://gizmodo.com/oracles-ai-push-is-leading...
52                 0.965723  https://finance.yahoo.com/news/nvidia-stock-wo...
1                  0.954643  https://247wallst.com/investing/2025/12/23/buf...

Summary of key findings complete.
```

```
# 1. Convert 'publishedAt' column to datetime objects
df_articles['publishedAt'] = pd.to_datetime(df_articles['publishedAt'], errors='coerce')

# 2. Verify the data type of the 'publishedAt' column
print("Data type of 'publishedAt' column after conversion:")
print(df_articles['publishedAt'].dtype)

# 3. Display the first few rows of the DataFrame with the updated 'publishedAt' column
print("\nFirst few rows with converted 'publishedAt' column:")
print(df_articles[['publishedAt', 'title', 'finbert_sentiment_label']].head())
```

```
Data type of 'publishedAt' column after conversion:
datetime64[ns, UTC]

First few rows with converted 'publishedAt' column:
                publishedAt  \
0 2026-01-12 16:32:48+00:00
1 2025-12-23 14:02:07+00:00
2 2026-01-09 17:52:23+00:00
3 2025-12-24 16:41:25+00:00
```

```
4 2025-12-25 13:05:26+00:00

                                          title finbert_sentiment_label
0  Apple strikes deal with Google's Gemini to pow...                positive
1  Buffett’s $24 Billion Selling Spree: The 6 Sto...                negative
2  Here's how much Apple CEO Tim Cook made last year                negative
3  Apple CEO Tim Cook Buys $3 Million of Nike Shares                 neutral
4  Apple Quietly Drops the Beats Studio Pro by 49...                 neutral
```

```
# 1. Create a new column for the date only
df_articles['published_date'] = df_articles['publishedAt'].dt.date

# 2. Group by 'published_date' and calculate the average 'finbert_sentiment_score'
df_sentiment_trend = df_articles.groupby('published_date')['finbert_sentiment_score'].mean().reset_index()

# Sort by date to ensure correct plotting order
df_sentiment_trend = df_sentiment_trend.sort_values(by='published_date')

# 3. Plot the daily average sentiment score over time
plt.figure(figsize=(12, 6))
sns.lineplot(x='published_date', y='finbert_sentiment_score', data=df_sentiment_trend)

# 4. Add appropriate labels and title
plt.title('Daily Average Sentiment Trend for Apple News Articles')
plt.xlabel('Date')
plt.ylabel('Average Sentiment Score')
plt.grid(True)
plt.xticks(rotation=45) # Rotate x-axis labels for better readability
plt.tight_layout() # Adjust layout to prevent labels from overlapping

# 5. Display the plot
plt.show()

print("Sentiment trend analysis and visualization complete.")
```