In [64]:
```
# import pandas as pd
# shopperData=pd.read_csv("D:\\intership\\New folder\\task-7-shopperData\\TeePub
import pandas as pd
data = pd.read_csv("D:\intership\cognoriseinfotech\TASK 7 SHOPPER SENTIMENTS\\Te
data.head()
```

Out[64]:

	reviewer_id	store_location	latitude	longitude	date	month	year	
0	0.0	US	37.090240	-95.712891	2023	6	2015 00:00:00	Great hel
1	1.0	US	37.090240	-95.712891	2023	6	2024 00:00:00	I ordered si hadï
2	2.0	US	37.090240	-95.712891	2023	6	2017 00:00:00	These guy customerï
3	3.0	US	37.090240	-95.712891	2023	6	2024 00:00:00	(
4	4.0	CA	56.130366	-106.346771	2023	6	2023 00:00:00	My orde timelyï

In [4]:
```
data.isna().sum()
```

Out[4]:
```
reviewer_id           1
store_location        0
latitude              0
longitude             0
date                  0
month                 0
year                  0
title                12
review            30503
review-label          0
dtype: int64
```

In [14]:
```
data['title'].fillna("No Title", inplace=True)
data['title'].isna().sum()
```

Out[14]:
```
0
```

In [15]:
```
data['review'].fillna("No Review", inplace=True)
data['review'].isna().sum()
```

Out[15]:
```
0
```

In [16]:
```
data['reviewer_id'].fillna("0", inplace=True)
data['reviewer_id'].isna().sum()
```

```
C:\Users\Admin\AppData\Local\Temp\ipykernel_15732\1642739307.py:1: FutureWarning:
Setting an item of incompatible dtype is deprecated and will raise in a future er
ror of pandas. Value '0' has dtype incompatible with float64, please explicitly c
ast to a compatible dtype first.
  data['reviewer_id'].fillna("0", inplace=True)
```

Out[16]:
```
0
```

In [17]:
```
data.isna().sum()
```

Out[17]:
```
reviewer_id       0
store_location    0
latitude          0
longitude         0
date              0
month             0
year              0
title             0
review            0
review-label      0
dtype: int64
```

In [27]:
```
import string
from nltk.corpus import stopwords
from nltk.tokenize import word_tokenize
from nltk.stem import WordNetLemmatizer
import nltk

# Download NLTK data files (only need to run once)
nltk.download('punkt')
nltk.download('stopwords')
nltk.download('wordnet')

# Initialize stopwords and lemmatizer
stop_words = set(stopwords.words('english'))
```

```
lemmatizer = WordNetLemmatizer()

def preprocess_text(text):
    # Convert to lowercase
    text = text.lower()
    # Remove punctuation
    text = text.translate(str.maketrans('', '', string.punctuation))
    # Tokenize
    tokens = word_tokenize(text)
    # Remove stopwords and lemmatize
    tokens = [lemmatizer.lemmatize(word) for word in tokens if word not in stop_
    return ' '.join(tokens)

# Apply preprocessing to the review column
data['cleaned_review'] = data['review'].apply(preprocess_text)

# Display the first few rows of the cleaned reviews
data[['review', 'cleaned_review']].head()
```

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

Out[27]:

	review	cleaned_review
0	I had an order that was lost in transit. When ...	order lost transit called help customer servic...
1	I ordered the wrong size tee and had difficult...	ordered wrong size tee difficulty returning we...
2	These guys offer the best customer service in ...	guy offer best customer service retail product...
3	Looked for an obscure phrase on a shirt. Teepu...	looked obscure phrase shirt teepublic process ...
4	My order arrived in a good timely fashion & th...	order arrived good timely fashion item receive...

In [26]:
```
pip install  geopandas
```

```
Collecting geopandas
  Downloading geopandas-0.14.4-py3-none-any.whl.metadata (1.5 kB)
Collecting fiona>=1.8.21 (from geopandas)
  Downloading fiona-1.9.6-cp311-cp311-win_amd64.whl.metadata (51 kB)
     ---------------------------------------- 0.0/51.5 kB ? eta -:--:--
     -------------- ---------------------- 20.5/51.5 kB 330.3 kB/s eta 0:00:01
     ---------------------------------------- 51.5/51.5 kB 529.4 kB/s eta 0:00:00
Requirement already satisfied: numpy>=1.22 in c:\users\admin\anaconda3\lib\site-p
ackages (from geopandas) (1.26.4)
Requirement already satisfied: packaging in c:\users\admin\anaconda3\lib\site-pac
kages (from geopandas) (23.1)
Requirement already satisfied: pandas>=1.4.0 in c:\users\admin\anaconda3\lib\site
-packages (from geopandas) (2.1.4)
Collecting pyproj>=3.3.0 (from geopandas)
  Downloading pyproj-3.6.1-cp311-cp311-win_amd64.whl.metadata (31 kB)
Collecting shapely>=1.8.0 (from geopandas)
  Downloading shapely-2.0.4-cp311-cp311-win_amd64.whl.metadata (7.2 kB)
Requirement already satisfied: attrs>=19.2.0 in c:\users\admin\anaconda3\lib\site
-packages (from fiona>=1.8.21->geopandas) (23.1.0)
Requirement already satisfied: certifi in c:\users\admin\anaconda3\lib\site-packa
ges (from fiona>=1.8.21->geopandas) (2024.2.2)
Requirement already satisfied: click~=8.0 in c:\users\admin\anaconda3\lib\site-pa
ckages (from fiona>=1.8.21->geopandas) (8.1.7)
Collecting click-plugins>=1.0 (from fiona>=1.8.21->geopandas)
  Downloading click_plugins-1.1.1-py2.py3-none-any.whl.metadata (6.4 kB)
Collecting cligj>=0.5 (from fiona>=1.8.21->geopandas)
  Downloading cligj-0.7.2-py3-none-any.whl.metadata (5.0 kB)
Requirement already satisfied: six in c:\users\admin\anaconda3\lib\site-packages
(from fiona>=1.8.21->geopandas) (1.16.0)
Requirement already satisfied: python-dateutil>=2.8.2 in c:\users\admin\anaconda3
\lib\site-packages (from pandas>=1.4.0->geopandas) (2.8.2)
Requirement already satisfied: pytz>=2020.1 in c:\users\admin\anaconda3\lib\site-
packages (from pandas>=1.4.0->geopandas) (2023.3.post1)
Requirement already satisfied: tzdata>=2022.1 in c:\users\admin\anaconda3\lib\sit
e-packages (from pandas>=1.4.0->geopandas) (2023.3)
Requirement already satisfied: colorama in c:\users\admin\anaconda3\lib\site-pack
ages (from click~=8.0->fiona>=1.8.21->geopandas) (0.4.6)
Downloading geopandas-0.14.4-py3-none-any.whl (1.1 MB)
   ---------------------------------------- 0.0/1.1 MB ? eta -:--:--
   ------- -------------------------------- 0.2/1.1 MB 5.9 MB/s eta 0:00:01
   ------------- -------------------------- 0.4/1.1 MB 5.8 MB/s eta 0:00:01
   ---------------------------------- ----- 1.0/1.1 MB 7.6 MB/s eta 0:00:01
   ---------------------------------------- 1.1/1.1 MB 7.0 MB/s eta 0:00:00
Downloading fiona-1.9.6-cp311-cp311-win_amd64.whl (22.9 MB)
   ---------------------------------------- 0.0/22.9 MB ? eta -:--:--
   - -------------------------------------- 0.7/22.9 MB 21.5 MB/s eta 0:00:02
   - -------------------------------------- 1.1/22.9 MB 13.8 MB/s eta 0:00:02
   -- ------------------------------------- 1.6/22.9 MB 12.6 MB/s eta 0:00:02
   --- ------------------------------------ 1.9/22.9 MB 10.0 MB/s eta 0:00:03
   ---- ----------------------------------- 2.5/22.9 MB 11.2 MB/s eta 0:00:02
   ----- ---------------------------------- 3.0/22.9 MB 11.2 MB/s eta 0:00:02
   ----- ---------------------------------- 3.3/22.9 MB 10.7 MB/s eta 0:00:02
   ------ --------------------------------- 3.9/22.9 MB 10.9 MB/s eta 0:00:02
   ------- -------------------------------- 4.4/22.9 MB 10.8 MB/s eta 0:00:02
   -------- ------------------------------- 4.9/22.9 MB 10.4 MB/s eta 0:00:02
   --------- ------------------------------ 5.3/22.9 MB 10.3 MB/s eta 0:00:02
   --------- ------------------------------ 5.7/22.9 MB 10.4 MB/s eta 0:00:02
   ---------- ----------------------------- 6.2/22.9 MB 10.4 MB/s eta 0:00:02
   ----------- ---------------------------- 6.5/22.9 MB 10.1 MB/s eta 0:00:02
   ----------- ---------------------------- 6.8/22.9 MB 10.1 MB/s eta 0:00:02
```

```
   ------------ --------------------------- 7.2/22.9 MB 9.8 MB/s eta 0:00:02
   ------------ --------------------------- 7.8/22.9 MB 9.9 MB/s eta 0:00:02
   ------------- -------------------------- 8.2/22.9 MB 10.0 MB/s eta 0:00:02
   -------------- ------------------------- 8.6/22.9 MB 10.0 MB/s eta 0:00:02
   -------------- ------------------------- 9.0/22.9 MB 10.1 MB/s eta 0:00:02
   --------------- ------------------------ 9.4/22.9 MB 10.0 MB/s eta 0:00:02
   ---------------- ----------------------- 9.8/22.9 MB 10.0 MB/s eta 0:00:02
   ---------------- ----------------------- 10.3/22.9 MB 9.8 MB/s eta 0:00:02
   ----------------- ---------------------- 10.7/22.9 MB 9.6 MB/s eta 0:00:02
   ------------------ --------------------- 11.1/22.9 MB 9.6 MB/s eta 0:00:02
   ------------------ --------------------- 11.4/22.9 MB 9.5 MB/s eta 0:00:02
   ------------------- -------------------- 11.8/22.9 MB 9.5 MB/s eta 0:00:02
   -------------------- ------------------- 12.1/22.9 MB 9.4 MB/s eta 0:00:02
   -------------------- ------------------- 12.4/22.9 MB 9.2 MB/s eta 0:00:02
   --------------------- ------------------ 12.7/22.9 MB 9.4 MB/s eta 0:00:02
   --------------------- ------------------ 13.0/22.9 MB 9.1 MB/s eta 0:00:02
   ---------------------- ----------------- 13.2/22.9 MB 8.8 MB/s eta 0:00:02
   ---------------------- ----------------- 13.6/22.9 MB 8.8 MB/s eta 0:00:02
   ----------------------- ---------------- 14.0/22.9 MB 8.7 MB/s eta 0:00:02
   ----------------------- ---------------- 14.4/22.9 MB 8.8 MB/s eta 0:00:01
   ------------------------ --------------- 14.8/22.9 MB 8.7 MB/s eta 0:00:01
   ------------------------ --------------- 15.2/22.9 MB 8.6 MB/s eta 0:00:01
   ------------------------- -------------- 15.6/22.9 MB 8.6 MB/s eta 0:00:01
   ------------------------- -------------- 15.9/22.9 MB 8.7 MB/s eta 0:00:01
   -------------------------- ------------- 16.3/22.9 MB 8.7 MB/s eta 0:00:01
   -------------------------- ------------- 16.6/22.9 MB 8.7 MB/s eta 0:00:01
   --------------------------- ------------ 17.1/22.9 MB 8.7 MB/s eta 0:00:01
   --------------------------- ------------ 17.4/22.9 MB 8.8 MB/s eta 0:00:01
   ---------------------------- ----------- 17.9/22.9 MB 8.8 MB/s eta 0:00:01
   ---------------------------- ----------- 18.5/22.9 MB 8.7 MB/s eta 0:00:01
   ----------------------------- ---------- 18.9/22.9 MB 8.7 MB/s eta 0:00:01
   ----------------------------- ---------- 19.3/22.9 MB 8.7 MB/s eta 0:00:01
   ------------------------------ --------- 19.7/22.9 MB 8.7 MB/s eta 0:00:01
   ------------------------------ --------- 20.1/22.9 MB 8.8 MB/s eta 0:00:01
   ------------------------------- -------- 20.5/22.9 MB 9.0 MB/s eta 0:00:01
   ------------------------------- -------- 20.9/22.9 MB 8.8 MB/s eta 0:00:01
   -------------------------------- ------- 21.2/22.9 MB 8.8 MB/s eta 0:00:01
   -------------------------------- ------- 21.6/22.9 MB 8.8 MB/s eta 0:00:01
   --------------------------------- ------ 22.1/22.9 MB 9.0 MB/s eta 0:00:01
   --------------------------------- ------ 22.6/22.9 MB 9.0 MB/s eta 0:00:01
   --------------------------------- ------ 22.9/22.9 MB 9.1 MB/s eta 0:00:01
   --------------------------------- ------ 22.9/22.9 MB 9.1 MB/s eta 0:00:01
   --------------------------------- ------ 22.9/22.9 MB 9.1 MB/s eta 0:00:01
   --------------------------------- ------ 22.9/22.9 MB 8.0 MB/s eta 0:00:00
Downloading pyproj-3.6.1-cp311-cp311-win_amd64.whl (6.1 MB)
   ---------------------------------------- 0.0/6.1 MB ? eta -:--:--
   --- ------------------------------------ 0.5/6.1 MB 15.5 MB/s eta 0:00:01
   ------ --------------------------------- 0.9/6.1 MB 11.7 MB/s eta 0:00:01
   -------- ------------------------------- 1.3/6.1 MB 10.4 MB/s eta 0:00:01
   ----------- ---------------------------- 1.8/6.1 MB 11.5 MB/s eta 0:00:01
   ---------- ----------------------------- 2.3/6.1 MB 11.2 MB/s eta 0:00:01
   ---------- ----------------------------- 2.6/6.1 MB 10.4 MB/s eta 0:00:01
   ------------ --------------------------- 3.0/6.1 MB 10.0 MB/s eta 0:00:01
   ------------- -------------------------- 3.5/6.1 MB 10.0 MB/s eta 0:00:01
   ------------- -------------------------- 4.0/6.1 MB 10.2 MB/s eta 0:00:01
   --------------- ------------------------ 4.4/6.1 MB 10.1 MB/s eta 0:00:01
   ---------------- ----------------------- 4.9/6.1 MB 10.2 MB/s eta 0:00:01
   ---------------- ----------------------- 5.3/6.1 MB 10.0 MB/s eta 0:00:01
   ----------------- ---------------------- 5.7/6.1 MB 9.8 MB/s eta 0:00:01
   ---------------------------------------  6.1/6.1 MB 9.7 MB/s eta 0:00:01
```

```
   ---------------------------------------- 6.1/6.1 MB 9.1 MB/s eta 0:00:00
Downloading shapely-2.0.4-cp311-cp311-win_amd64.whl (1.4 MB)
   ---------------------------------------- 0.0/1.4 MB ? eta -:--:--
   ------- -------------------------------- 0.3/1.4 MB 5.9 MB/s eta 0:00:01
   ------------------ --------------------- 0.7/1.4 MB 8.8 MB/s eta 0:00:01
   ------------------------------ -------- 1.1/1.4 MB 7.9 MB/s eta 0:00:01
   ---------------------------------------- 1.4/1.4 MB 8.3 MB/s eta 0:00:00
Downloading click_plugins-1.1.1-py2.py3-none-any.whl (7.5 kB)
Downloading cligj-0.7.2-py3-none-any.whl (7.1 kB)
Installing collected packages: shapely, pyproj, cligj, click-plugins, fiona, geop
andas
Successfully installed click-plugins-1.1.1 cligj-0.7.2 fiona-1.9.6 geopandas-0.1
4.4 pyproj-3.6.1 shapely-2.0.4
Note: you may need to restart the kernel to use updated packages.
```

In [28]:

```
import geopandas as gpd

# Convert the dataframe to a GeoDataFrame
gdf = gpd.GeoDataFrame(data, geometry=gpd.points_from_xy(data.longitude, data.la

# Verify the GeoDataFrame
gdf.head()
```

Out[28]:

	reviewer_id	store_location	latitude	longitude	date	month	year	
0	0.0	US	37.090240	-95.712891	2023	6	2015 00:00:00	Great hel
1	1.0	US	37.090240	-95.712891	2023	6	2024 00:00:00	I ordered si hadï
2	2.0	US	37.090240	-95.712891	2023	6	2017 00:00:00	These guy customerï
3	3.0	US	37.090240	-95.712891	2023	6	2024 00:00:00	
4	4.0	CA	56.130366	-106.346771	2023	6	2023 00:00:00	My orde timelyï

In [33]:

```
from sklearn.feature_extraction.text import TfidfVectorizer

# Initialize the TF-IDF vectorizer
vectorizer = TfidfVectorizer()

# Fit and transform the cleaned review text
tfidf_matrix = vectorizer.fit_transform(data['cleaned_review'])
print(tfidf_matrix)
# Display the shape of the TF-IDF matrix
print(tfidf_matrix.shape)
```

```
  (0, 68550)    0.10822834414008975
  (0, 100003)   0.15698893199336114
  (0, 100132)   0.7068607939800713
  (0, 44734)    0.1284136153022964
  (0, 9359)     0.16115268143071207
  (0, 71162)    0.13046834684243114
  (0, 62069)    0.16923917230957372
  (0, 54726)    0.2643863618010808
  (0, 36445)    0.10915477858954839
  (0, 70009)    0.2211741851869513
  (0, 74024)    0.09025129024357859
  (0, 20937)    0.09964953878975828
  (0, 39311)    0.17150800346374082
  (0, 14180)    0.21823511239924742
  (0, 88693)    0.2285302317873252
  (0, 48763)    0.17546140807131583
  (0, 57842)    0.25884782427467457
  (1, 28964)    0.15075279975790737
  (1, 63167)    0.3476347779716208
  (1, 39782)    0.18731408618279524
  (1, 75437)    0.07468179658961509
  (1, 73842)    0.16476012122894193
  (1, 43650)    0.16812866412220415
  (1, 70296)    0.24238828546980723
  (1, 41792)    0.20517287495753816
  :     :
  (278098, 36088)       0.19259558784674996
  (278098, 71162)       0.28144817117136917
  (278099, 57659)       0.30384514626962644
  (278099, 85438)       0.28744585561288105
  (278099, 72765)       0.25058143894462476
  (278099, 5594)        0.23125754403355886
  (278099, 96062)       0.15951663076991182
  (278099, 30672)       0.13675951640679748
  (278099, 21991)       0.20818988755851103
  (278099, 15002)       0.16336793631627103
  (278099, 52123)       0.1557712406113355
  (278099, 7348)        0.17779198999112208
  (278099, 91739)       0.14924390282658798
  (278099, 14992)       0.16393654199505034
  (278099, 12298)       0.15013680786238665
  (278099, 48527)       0.09074261864867765
  (278099, 59251)       0.1315491545179948
  (278099, 64016)       0.0946364167121712
  (278099, 99833)       0.10278199502259579
  (278099, 78153)       0.11521501825782585
  (278099, 94583)       0.09635113799287348
  (278099, 38537)       0.09403079782445133
  (278099, 37133)       0.057687792419871986
  (278099, 75437)       0.191949640950626
  (278099, 100132)      0.5919496843056041
(278100, 100224)
```

In [40]:
```
l=['date','month', 'year', 'title']
data[l]
```

Out[40]:

	date	month	year	title
0	2023	6	2015 00:00:00	Great help with lost order
1	2023	6	2024 00:00:00	I ordered the wrong size tee and hadï¿½ï¿½ï¿½
2	2023	6	2017 00:00:00	These guys offer the best customerï¿½ï¿½ï¿½
3	2023	6	2024 00:00:00	Good Stuff
4	2023	6	2023 00:00:00	My order arrived in a good timelyï¿½ï¿½ï¿½
...	...	...	...	...
278095	2018	4	2027 00:00:00	Highly recommend!
278096	2018	4	2027 00:00:00	Great quality
278097	2018	4	2027 00:00:00	Dudes rock.
278098	2018	4	2027 00:00:00	Shipping was fast the T-shirt was justï¿½ï¿½ï¿½
278099	2018	4	2027 00:00:00	Not great quality

278100 rows × 4 columns

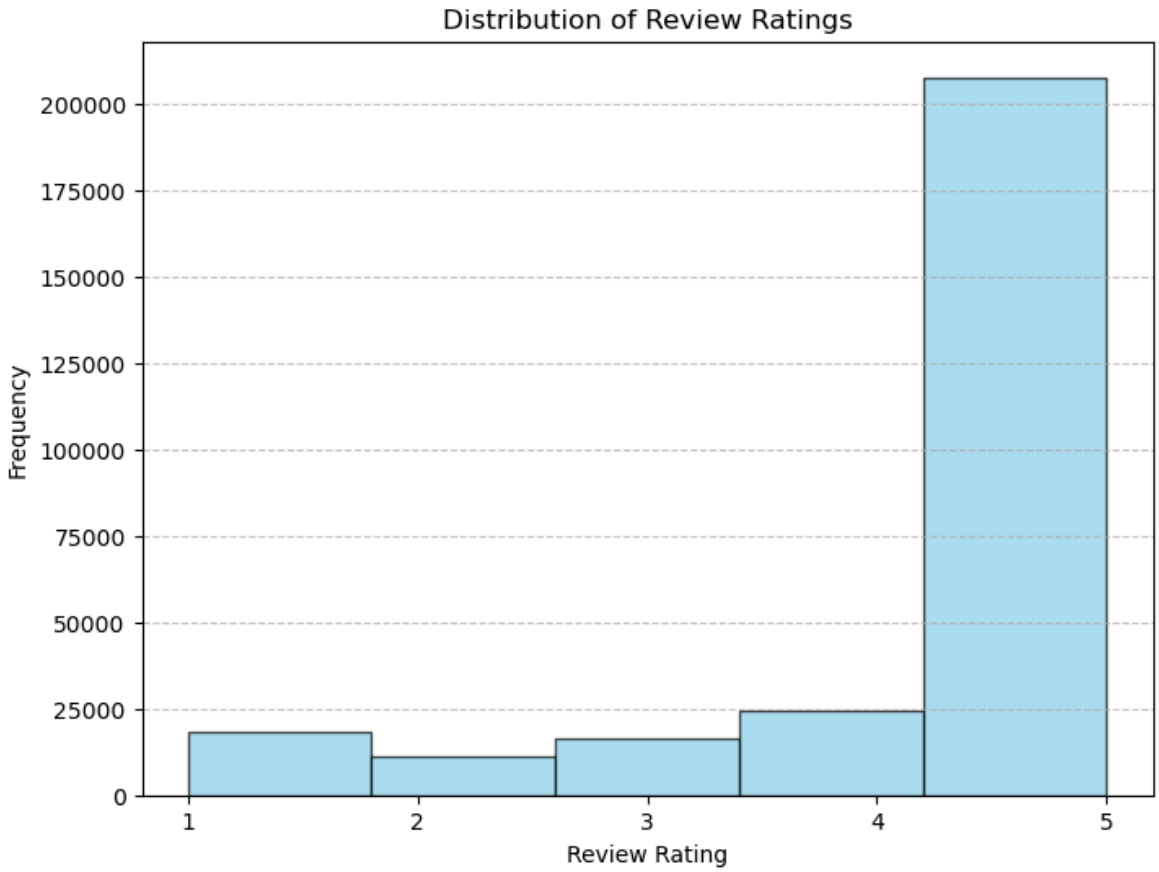
In [39]:
```
data.columns
```

Out[39]:
```
Index(['reviewer_id', 'store_location', 'latitude', 'longitude', 'date',
       'month', 'year', 'title', 'review', 'review-label', 'cleaned_review'],
      dtype='object')
```

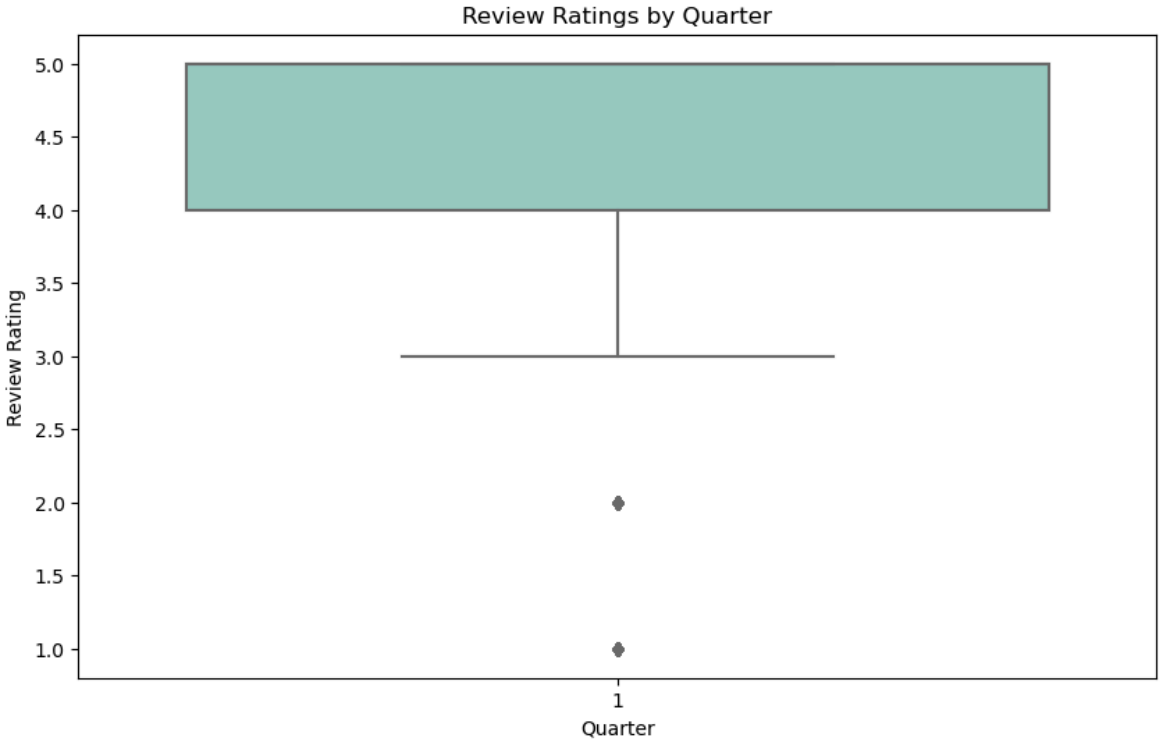
In [46]:
```
# data['date'] = pd.to_datetime(data['date'].astype(str) + '0101', format='%Y%m%
# Extract additional time features
data['day_of_week'] = data['date'].dt.dayofweek
data['quarter'] = data['date'].dt.quarter

# Display the first few rows with the new features
print(data[['date', 'day_of_week', 'quarter']].head())
```

```
        date  day_of_week  quarter
0 2023-01-01            6        1
1 2023-01-01            6        1
2 2023-01-01            6        1
3 2023-01-01            6        1
4 2023-01-01            6        1
```

In [48]:
```
# import pandas as pd
import matplotlib.pyplot as plt
# Plot histogram of review ratings
plt.figure(figsize=(8, 6))
plt.hist(data['review-label'], bins=5, color='skyblue', edgecolor='black', alpha
plt.xlabel('Review Rating')
plt.ylabel('Frequency')
plt.title('Distribution of Review Ratings')
plt.xticks(range(1, 6))
plt.grid(axis='y', linestyle='--', alpha=0.7)
plt.show()
```

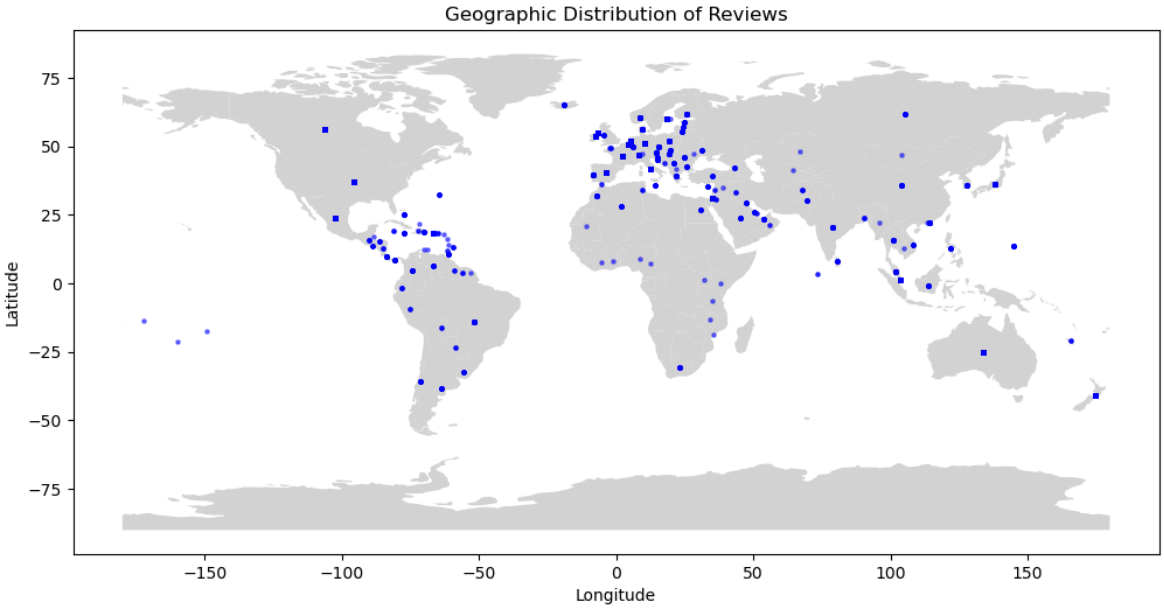


In [50]:
```
import seaborn as sns



# Plot boxplot of review ratings by quarter
plt.figure(figsize=(10, 6))
sns.boxplot(x='quarter', y='review-label', data=data, palette='Set3')
plt.xlabel('Quarter')
plt.ylabel('Review Rating')
plt.title('Review Ratings by Quarter')
plt.show()
```



In [52]:
```
import matplotlib.pyplot as plt
import geopandas as gpd
from shapely.geometry import Point


# Convert to GeoDataFrame
gdf = gpd.GeoDataFrame(data, geometry=gpd.points_from_xy(data.longitude, data.la

# Plotting the map
world = gpd.read_file(gpd.datasets.get_path('naturalearth_lowres'))
fig, ax = plt.subplots(figsize=(12, 8))
world.plot(ax=ax, color='lightgrey')

# Plot points
gdf.plot(ax=ax, markersize=5, color='blue', alpha=0.5)

plt.title('Geographic Distribution of Reviews')
plt.xlabel('Longitude')
plt.ylabel('Latitude')
plt.show()
```

```
C:\Users\Admin\AppData\Local\Temp\ipykernel_15732\3223844924.py:10: FutureWarnin
g: The geopandas.dataset module is deprecated and will be removed in GeoPandas 1.
0. You can get the original 'naturalearth_lowres' data from https://www.naturalea
rthdata.com/downloads/110m-cultural-vectors/.
  world = gpd.read_file(gpd.datasets.get_path('naturalearth_lowres'))
```



In [53]:
```
pip install plotly geopandas
```

```
Requirement already satisfied: plotly in c:\users\admin\anaconda3\lib\site-packag
es (5.9.0)
Requirement already satisfied: geopandas in c:\users\admin\anaconda3\lib\site-pac
kages (0.14.4)
Requirement already satisfied: tenacity>=6.2.0 in c:\users\admin\anaconda3\lib\si
te-packages (from plotly) (8.2.2)
Requirement already satisfied: fiona>=1.8.21 in c:\users\admin\anaconda3\lib\site
-packages (from geopandas) (1.9.6)
Requirement already satisfied: numpy>=1.22 in c:\users\admin\anaconda3\lib\site-p
ackages (from geopandas) (1.26.4)
Requirement already satisfied: packaging in c:\users\admin\anaconda3\lib\site-pac
kages (from geopandas) (23.1)
Requirement already satisfied: pandas>=1.4.0 in c:\users\admin\anaconda3\lib\site
-packages (from geopandas) (2.1.4)
Requirement already satisfied: pyproj>=3.3.0 in c:\users\admin\anaconda3\lib\site
-packages (from geopandas) (3.6.1)
Requirement already satisfied: shapely>=1.8.0 in c:\users\admin\anaconda3\lib\sit
e-packages (from geopandas) (2.0.4)
Requirement already satisfied: attrs>=19.2.0 in c:\users\admin\anaconda3\lib\site
-packages (from fiona>=1.8.21->geopandas) (23.1.0)
Requirement already satisfied: certifi in c:\users\admin\anaconda3\lib\site-packa
ges (from fiona>=1.8.21->geopandas) (2024.2.2)
Requirement already satisfied: click~=8.0 in c:\users\admin\anaconda3\lib\site-pa
ckages (from fiona>=1.8.21->geopandas) (8.1.7)
Requirement already satisfied: click-plugins>=1.0 in c:\users\admin\anaconda3\lib
\site-packages (from fiona>=1.8.21->geopandas) (1.1.1)
Requirement already satisfied: cligj>=0.5 in c:\users\admin\anaconda3\lib\site-pa
ckages (from fiona>=1.8.21->geopandas) (0.7.2)
Requirement already satisfied: six in c:\users\admin\anaconda3\lib\site-packages
(from fiona>=1.8.21->geopandas) (1.16.0)
Requirement already satisfied: python-dateutil>=2.8.2 in c:\users\admin\anaconda3
\lib\site-packages (from pandas>=1.4.0->geopandas) (2.8.2)
Requirement already satisfied: pytz>=2020.1 in c:\users\admin\anaconda3\lib\site-
packages (from pandas>=1.4.0->geopandas) (2023.3.post1)
Requirement already satisfied: tzdata>=2022.1 in c:\users\admin\anaconda3\lib\sit
e-packages (from pandas>=1.4.0->geopandas) (2023.3)
Requirement already satisfied: colorama in c:\users\admin\anaconda3\lib\site-pack
ages (from click~=8.0->fiona>=1.8.21->geopandas) (0.4.6)
Note: you may need to restart the kernel to use updated packages.
```

In [57]:
```
import pandas as pd
import geopandas as gpd
import plotly.express as px

# Load the dataset
# data = pd.read_csv('shopper_sentiments.csv')

# Convert to GeoDataFrame
gdf = gpd.GeoDataFrame(data, geometry=gpd.points_from_xy(data.longitude, data.la

# Load world map data
world = gpd.read_file(gpd.datasets.get_path('naturalearth_lowres'))

# Check column names to ensure compatibility
print("Columns in 'world' GeoDataFrame:", world.columns)
print("Columns in 'gdf' GeoDataFrame:", gdf.columns)
```

```
Columns in 'world' GeoDataFrame: Index(['pop_est', 'continent', 'name', 'iso_a3',
'gdp_md_est', 'geometry'], dtype='object')
Columns in 'gdf' GeoDataFrame: Index(['reviewer_id', 'store_location', 'latitud
e', 'longitude', 'date',
       'month', 'year', 'title', 'review', 'review-label', 'cleaned_review',
       'day_of_week', 'quarter', 'geometry'],
      dtype='object')
```

```
C:\Users\Admin\AppData\Local\Temp\ipykernel_15732\3801850089.py:12: FutureWarnin
g: The geopandas.dataset module is deprecated and will be removed in GeoPandas 1.
0. You can get the original 'naturalearth_lowres' data from https://www.naturalea
rthdata.com/downloads/110m-cultural-vectors/.
  world = gpd.read_file(gpd.datasets.get_path('naturalearth_lowres'))
```

In [ ]:

In [ ]:

In [ ]: