```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from scipy import stats

# Loading datasets
bookseller_df = pd.read_csv('bookseller.csv')
debugging_df = pd.read_csv('debugging.csv')

bookseller_df.head(), debugging_df.head()
```

```
(   S.No        Date  Number of Books Sold
 0     1  01-01-2020                    90
 1     2  02-01-2020                   100
 2     3  03-01-2020                   100
 3     4  04-01-2020                    97
 4     5  05-01-2020                    93,
    Bug ID  Time Taken to fix the bug
 0   12986                       2.42
 1   12987                       2.03
 2   12988                       2.74
 3   12989                       3.21
 4   12990                       3.40)
```

# Analysis of Bookseller Data

1. **Descriptive Statistics**: We start by analyzing the distribution of `Number of Books Sold`.
2. **Inferential Analysis**: Perform t-tests to check hypotheses related to average sales.

```
# Descriptive statistics for Bookseller data
bookseller_stats = bookseller_df['Number of Books Sold'].describe()
print("Descriptive Statistics for Number of Books Sold:")
print(bookseller_stats)

# Hypothesis Testing - One Sample t-test against a hypothetical mean
(e.g., 95 books sold)
t_stat, p_val = stats.ttest_1samp(bookseller_df['Number of Books
Sold'], 95)
print("\nOne Sample t-test results:")
print(f"T-statistic: {t_stat}, P-value: {p_val}")
```

```
Descriptive Statistics for Number of Books Sold:
count    366.000000
mean      94.961749
std        3.178465
min       90.000000
```

```
25%         92.000000
50%         95.000000
75%         98.000000
max        100.000000
Name: Number of Books Sold, dtype: float64

One Sample t-test results:
T-statistic: -0.23023428813896255, P-value: 0.8180386995785215
```

## Analysis of Debugging Data

1. **Descriptive Statistics**: Analyze the time taken to fix bugs.
2. **Inferential Analysis**: Perform t-tests to evaluate debugging times.

```
# Descriptive statistics for Debugging data
debugging_stats = debugging_df['Time Taken to fix the bug'].describe()
print("Descriptive Statistics for Time Taken to Fix Bugs:")
print(debugging_stats)

# Hypothesis Testing - One Sample t-test against a hypothetical mean
time (e.g., 2.5 hours)
t_stat_debug, p_val_debug = stats.ttest_1samp(debugging_df['Time Taken
to fix the bug'], 2.5)
print("\nOne Sample t-test for debugging times:")
print(f"T-statistic: {t_stat_debug}, P-value: {p_val_debug}")

Descriptive Statistics for Time Taken to Fix Bugs:
count    2098.000000
mean        3.012531
std         1.147148
min         1.010000
25%         2.010000
50%         3.005000
75%         4.030000
max         5.000000
Name: Time Taken to fix the bug, dtype: float64

One Sample t-test for debugging times:
T-statistic: 20.46460352859365, P-value: 5.205283711498065e-85
```