Data Analytics Internship Report

*Uses dummy data for the execution.

0. Executing files:

File name: runAll.py

Description: It is a file that allows all three programs to run at once. Separate files are created for easy management of the files and their management. The runAll.py allows to get outputs at once.

1. Engagement Metrics:

File name: analysis.py

Description: This script visualizes user interaction metrics for five webtoon pages, including page views, average time spent on each page, and bounce rates. The visualizations are designed to help identify patterns in user behavior, making it easier to suggest optimization strategies based on how users engage with the content.
This code provides three key visualizations:

- ➔ A bar chart showing the popularity of each page through page views.
- ➔ A line chart showing how long users typically spend on each page.
- ➔ A bar chart highlights each page's bounce rate, indicating user engagement or disengagement.

Libraries: *pandas* is used for data manipulation and creation of the DataFrame.
matplotlib.pyplot and *seaborn* are used to create visualizations.

Suggestions: My suggestions for increasing the time spent on the page are as follows:
Interactive Elements: Incorporate quizzes, polls, or interactive story elements to encourage users to spend more time on the page (ex: “Which Tower of God character are you?”, Asking for their favorite character at the end of the article).
Enhance Content Interaction: Along with a comment section, add a filter/section where fan theories can be posted about the upcoming chapters or the story endings.

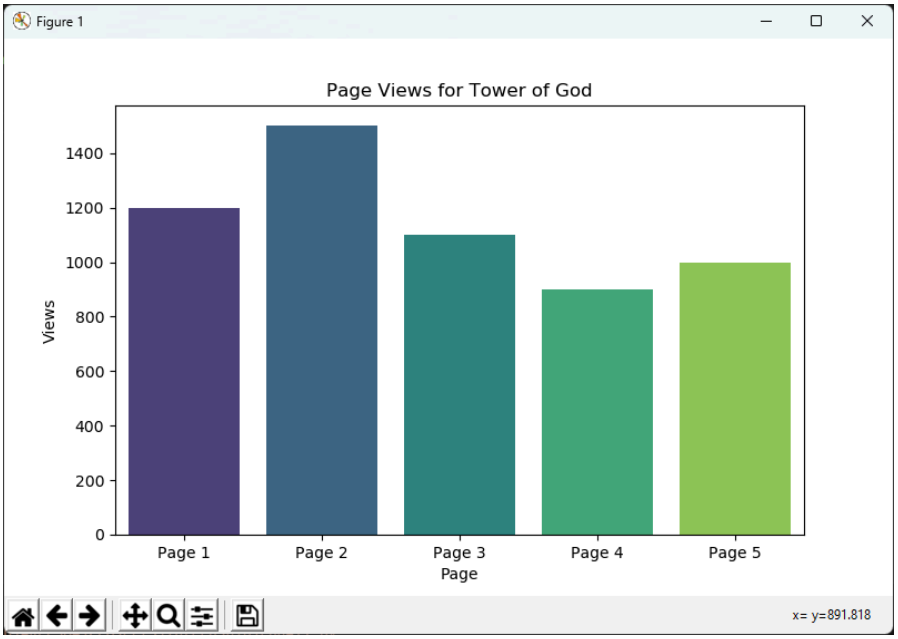


Figure 1
Page Views for Tower of God
Views
0
200
400
600
800
1000
1200
1400
Page 1
Page 2
Page 3
Page 4
Page 5
Page
x= y=891.818

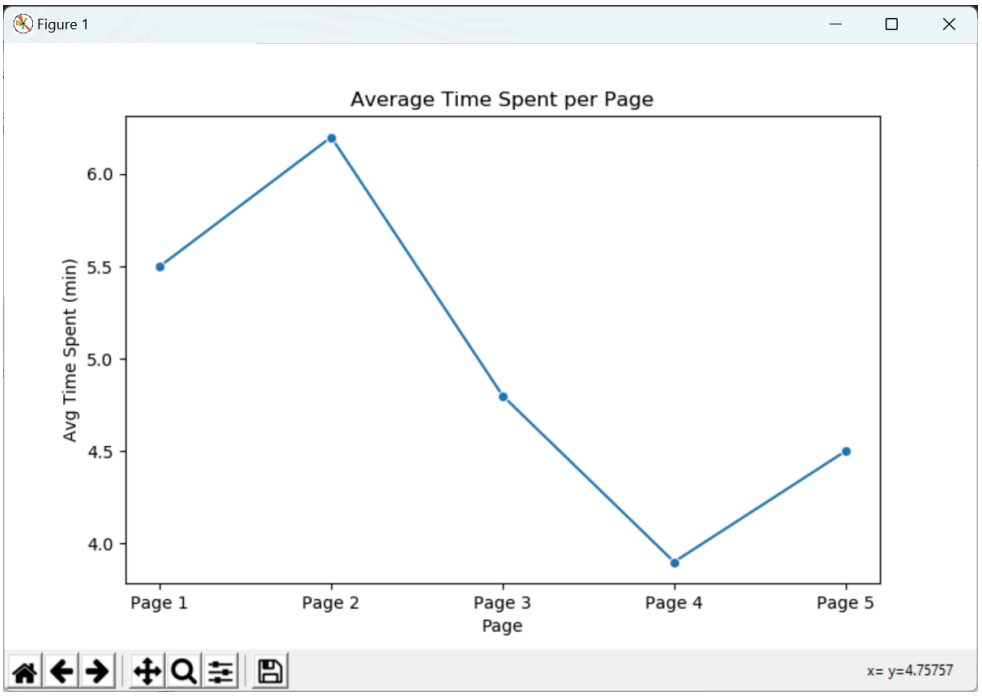
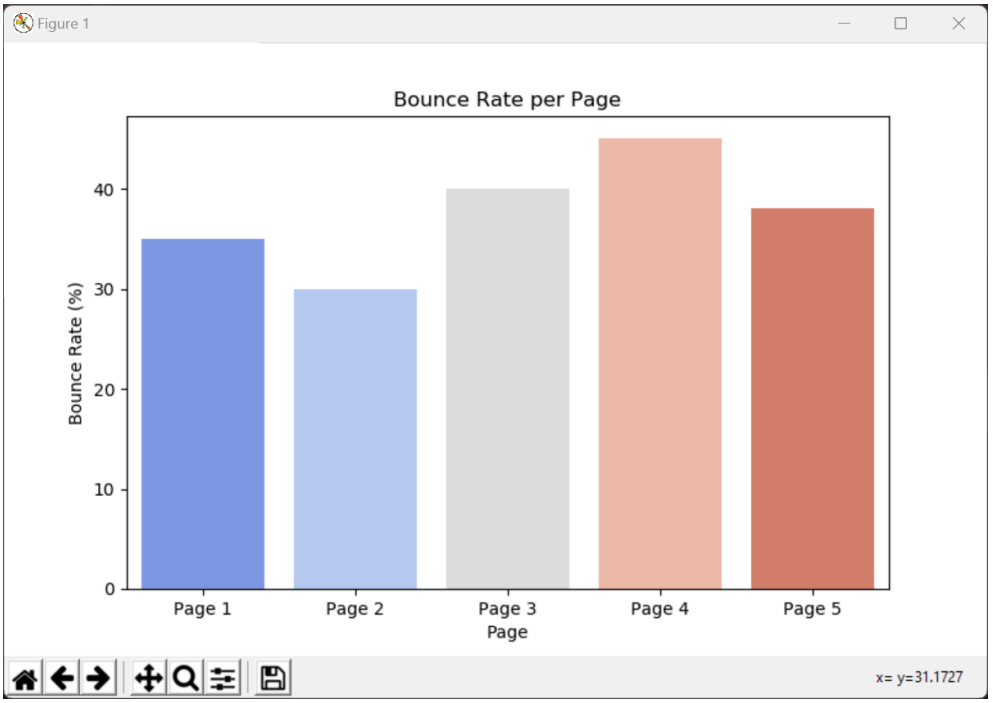


Figure 1
Average Time Spent per Page
Avg Time Spent (min)
4.0
4.5
5.0
5.5
6.0
Page 1
Page 2
Page 3
Page 4
Page 5
Page
x= y=4.75757



2. A/B Testing Strategy

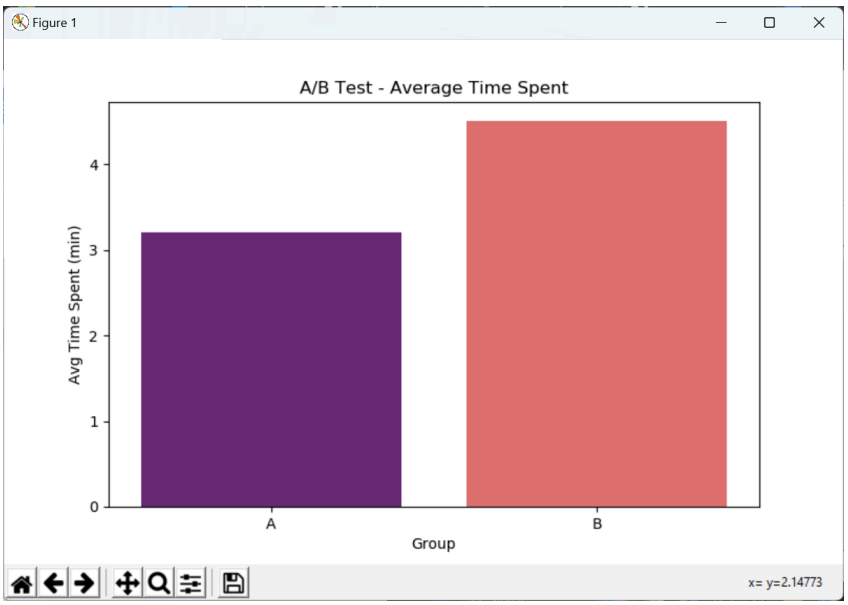
File name: Test.py

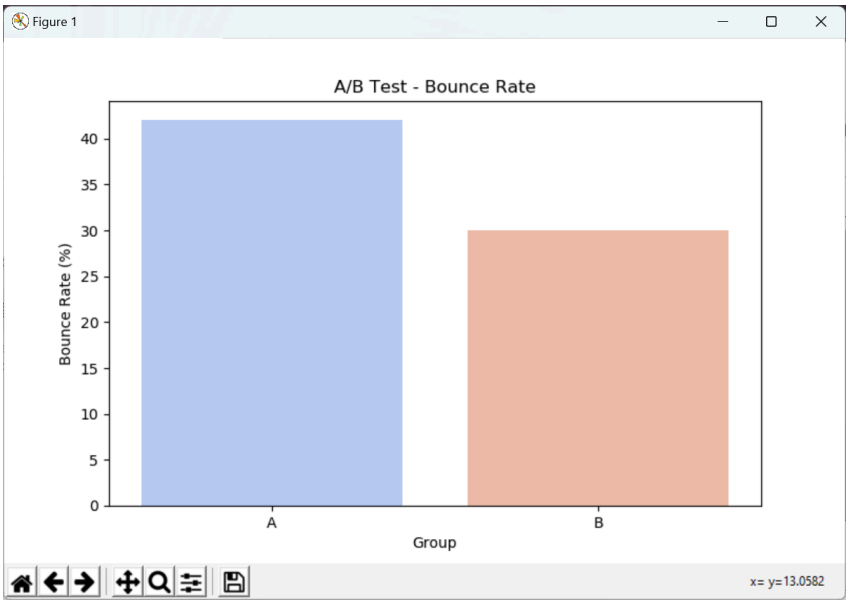
Description: The goal of the A/B test is to identify which headline and visual design drives higher user engagement and lower bounce rates. This code simulates and visualizes the results of an A/B test designed to compare user engagement metrics between two different versions of a webpage.

A/B Test Setup:

Variable to Test: Headline text and background images.
Hypothesis: Changing the headline to a more engaging text and using brighter visuals will improve user engagement.

Using the pandas library, it creates a DataFrame with data for two groups: Group A, which uses an "Old Headline," and Group B, which features a "New Engaging Headline." The dataset includes metrics such as the average time spent on the page and the bounce rate for each group. The seaborn and matplotlib libraries are used to generate bar charts that visually compare the performance of both groups. The first bar chart illustrates the average time users spend in each group, highlighting that Group B has a higher engagement time. The second bar chart compares the bounce rate, showing that Group B has a lower bounce rate than Group A. The visualization helps identify that the new headline in Group B performs better in terms of increasing user time on the page and reducing the bounce rate, suggesting it is the more effective option.





3. User Segmentation

File name: segment.py

Description: This code analyzes and visualizes user engagement metrics based on different user segments using Pandas, Matplotlib, and Seaborn. It creates a DataFrame containing data for various segments, such as age groups (18-24, 25-34, 35-44) and user behavior (returning vs. new visitors), including metrics like average time spent and page views. The first visualization uses a bar plot to display the average time spent by each user segment, helping to identify which groups spend more time on the content. The second bar plot illustrates the page views for each segment, showing which groups generate the most traffic. The

visualizations help highlight differences in user engagement and behavior across different segments, providing insights into how to target each group effectively.