Part-A Observations

❏ Primary Observations

- For this part, we took 5 websites and followed all the hyperlinks on them and recorded the time for the corresponding pages to load (an average of 5 attempts), the number of deadlinks and the number of links that work. After that we calculated the website score using the following formulae:

  - col1: Average link load time
  - col2: Number of dead links
  - col3: Number of working links
  - A: normalized Average Link Load Time (Mobile Network)
  - $A = (val(col1) - min(col1)) / (max(col1) - min(col1))$
  - B: Fraction of dead links
  - $B = val(col2) / (val(col2) + val(col3))$
  - Website score = $(A + B) / 2$

1. https://www.usa.gov/ :
   - The average link load time is 1141.05 ms
   - The number of dead links are 0
   - The number of working links are 561
   - The overall website score is 0

2. https://www.isro.gov.in/
   - The average link load time is 1166.2745 ms
   - The number of dead links are 1
   - The number of working links are 102
   - The overall website score is 0.01127

3. https://www.bits-pilani.ac.in/
   - The average link load time is 1379.2668 ms
   - The number of dead links are 1
   - The number of working links are 512
   - The overall website score is 0.06157

4. https://nrega.nic.in/netnrega/home.aspx
   - The average link load time is 1951.8496 ms
   - The number of dead links are 12

- The number of working links are 125
- The overall website score is 0.25005

5. https://medium.com/
  - The average link load time is 3106.6025 ms
  - The number of dead links are 0
  - The number of working links are 80
  - The overall website score is 0.5

❏ Some other observations:

- Website with highest score (worst website in terms of performance): https://medium.com/ (score: 0.5)
- Website with least score (best website in terms of performance): https://www.usa.gov/ (score: 0)
- Website with highest average loading time: https://medium.com/ (time: 3106.6025 ms)
- Website with least average loading time: https://www.usa.gov/ (time: 1141.05 ms)
- Website with highest number of dead links: https://nrega.nic.in/netnrega/home.aspx (12 dead links = 8.75 % dead links)
- Websites with no dead links: https://medium.com/ and https://www.usa.gov/
- Websites with none or almost none dead links indicates they are updated at regular intervals and managed if a link is broken or dead.
- Performance of a website is inversely proportional to its score as the score is the average of normalized link load time and fraction of dead links.

❏ Checking if a Link is Dead or not

- We obtain the performance log of the link where we capture the status code sent by the server.
- For every status code code 4xx & 5xx, we consider that link to be dead.
- If the hostname of the link is unable to resolve, we consider that link to be dead. It generates the exception server.gaierror and it gets handled in the except block.
- Any link which is taking too long to respond and throws a TimeoutException is handled and is considered dead.

❏ Calculating Link Load Time

- We are using Navigation Timing JavaScript APIs to measure the link load time. Though there are other libraries available, Navigation Timing is more reliable & accurate.
- **navigationStart** – This attribute returns the time spent after the user agent completes unloading the previous page/document. If there was no document prior to loading the new page, navigationStart returns the same value as fetchStart.
- **domComplete** – This attribute returns the time just before the current document/page readiness is set to 'complete'.
- The differences calculate the total Link Loading Time.
- For the calculation of Average Link Load Time, we are only taking the working links of the websites.