

Google WebLight Transformations Analysis

Ammar Tahir 20100212

Muhammad Adil Inam 20100180

Two Types of Analysis

- External Analysis of the Objects fetched
、
- Internal Analysis of the underlying HTML

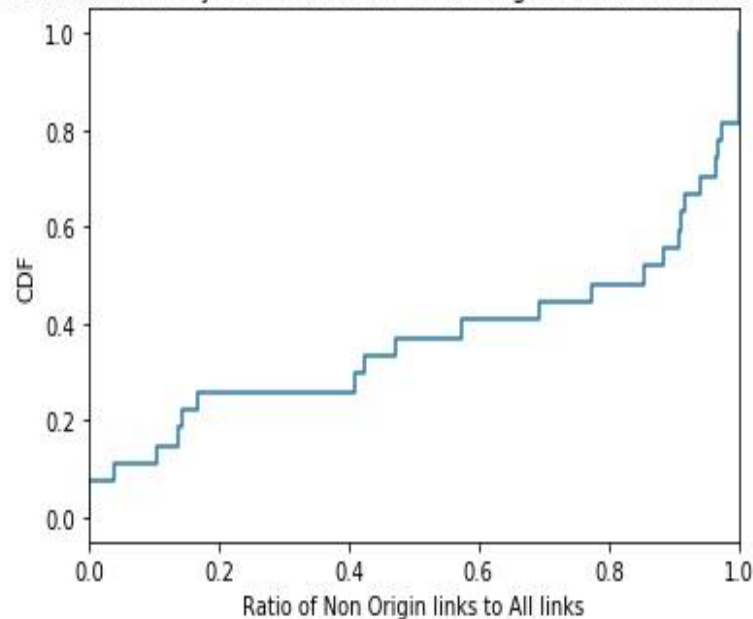
External Analysis

- Comparison of origin and non-origin requests and servers in the transformed and non-transformed pages.
- Comparison of the ratio of different kinds external objects (JS , CSS , HTML, Image) fetched in the transformed and non-transformed pages.

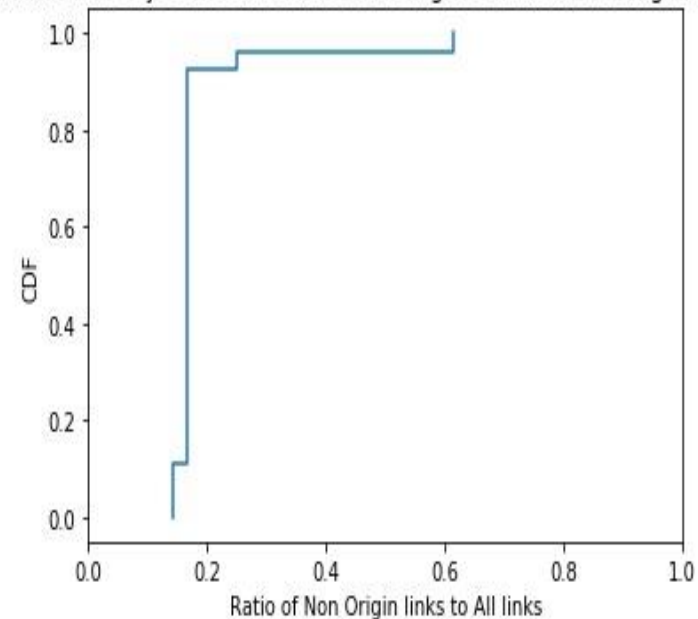
Non-Origin Requests Comparison

- The ratio of non-origin to total requests is significantly larger in non-transformed pages as compared to transformed pages.

Normalised Number of Objects fetched from Non Origin Servers in Normal Mobile pages

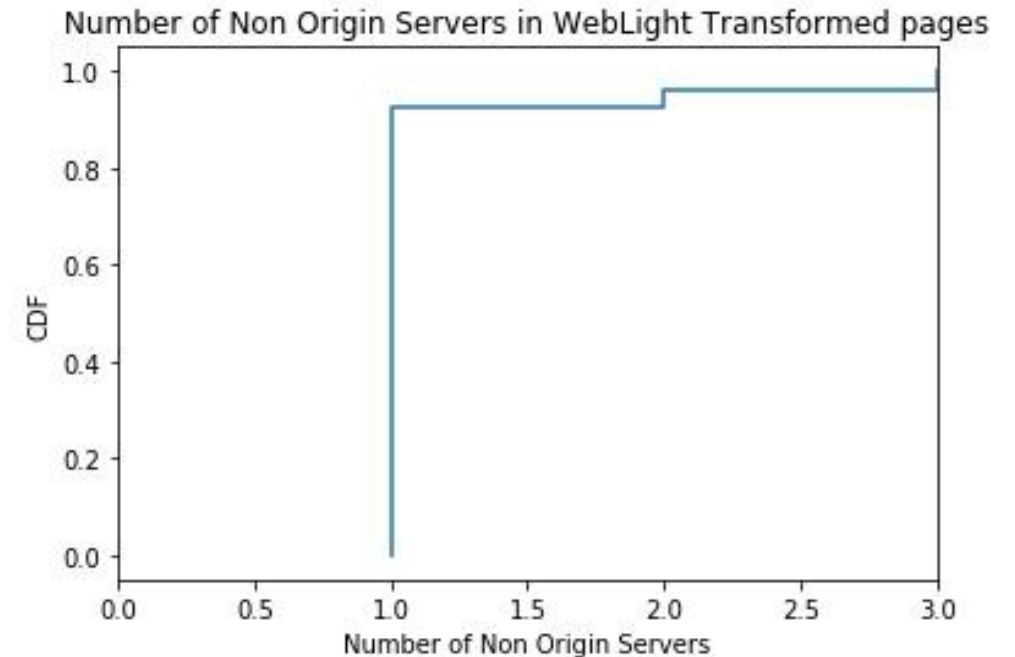
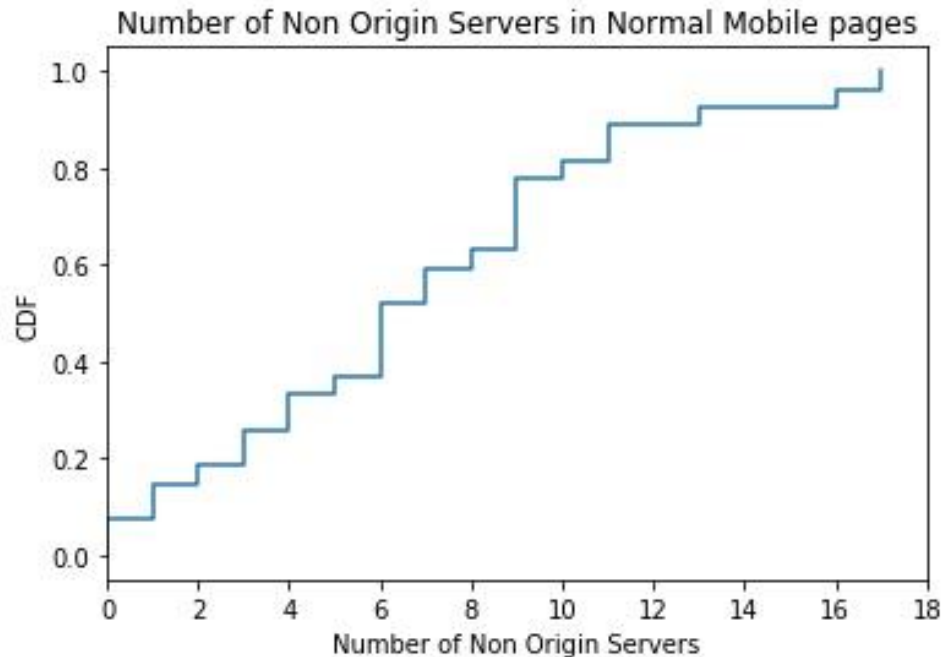


Normalised Number of Objects fetched from Non Origin Servers in WebLight Transformed pages



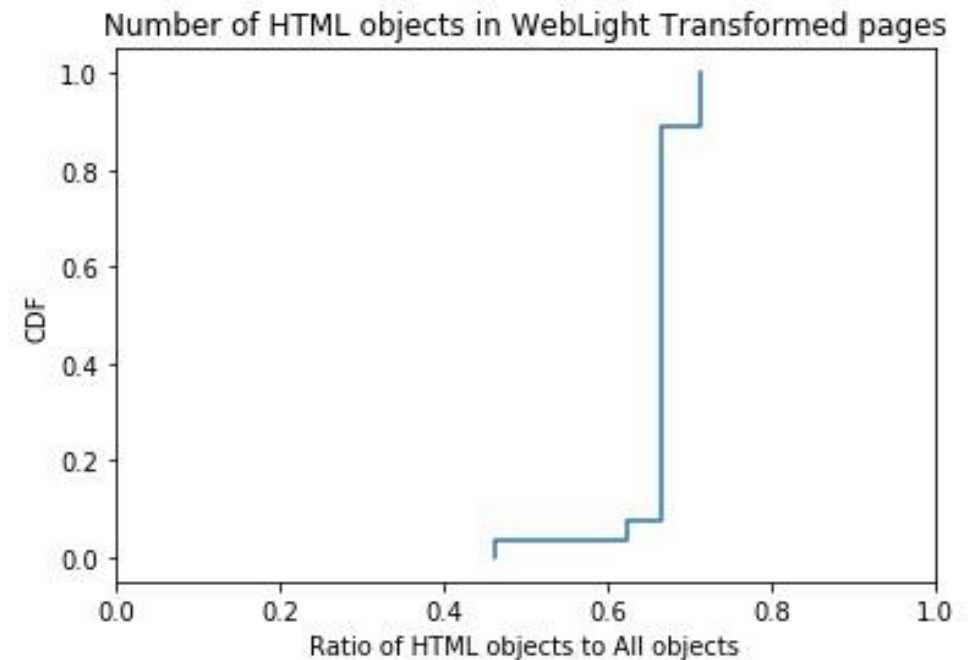
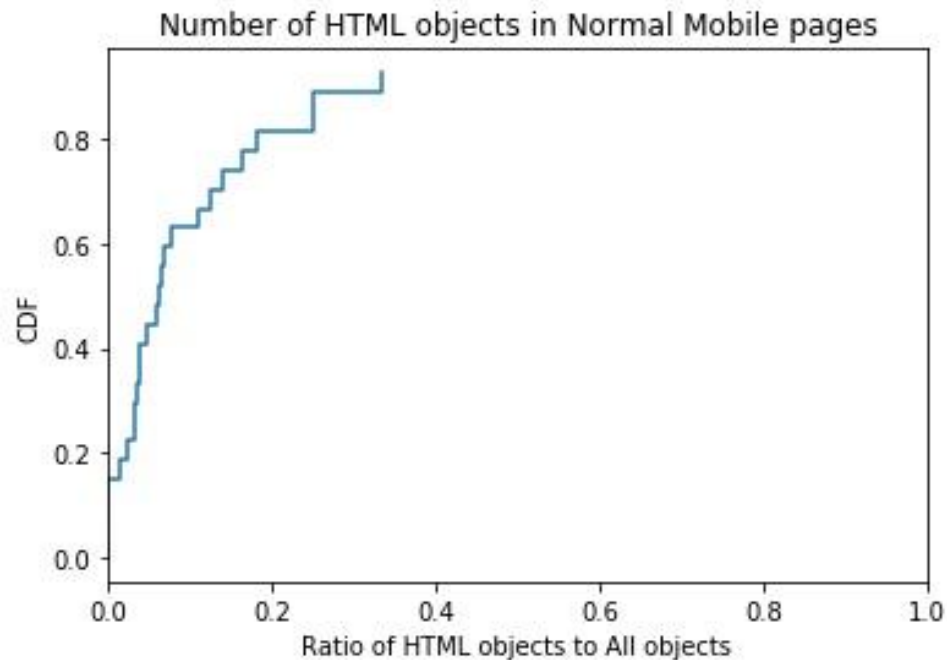
Non-Origin Servers Comparison

- The number of non-origin servers are greater in number for non-transformed pages as compared to transformed pages.



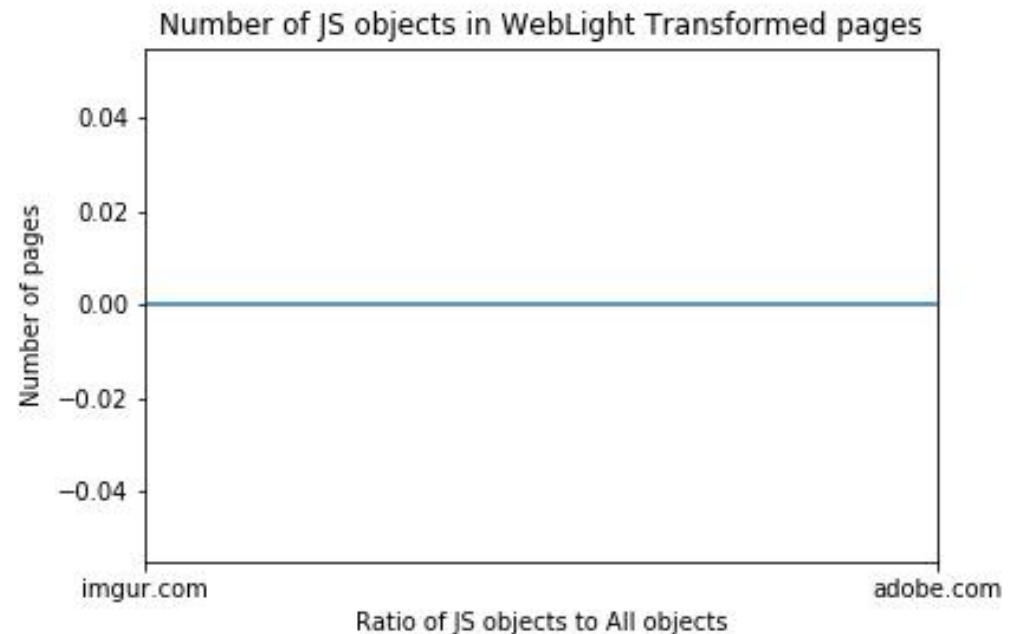
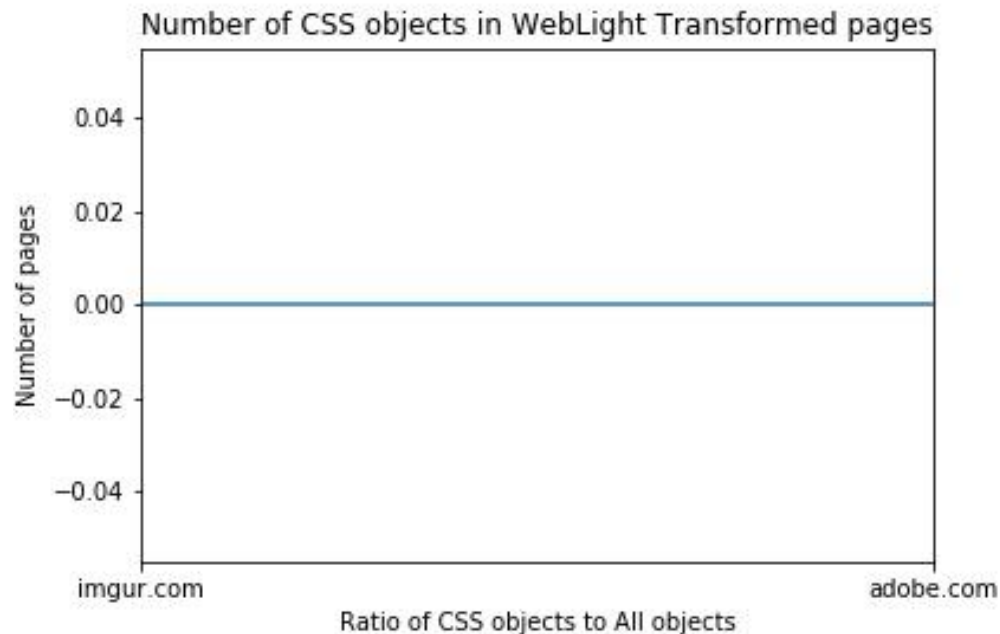
Ratio of HTML Objects

- The ratio of html objects to total objects fetched is significantly higher in transformed pages as most of the external CSS and Js objects are embedded inside the html in transformed pages.



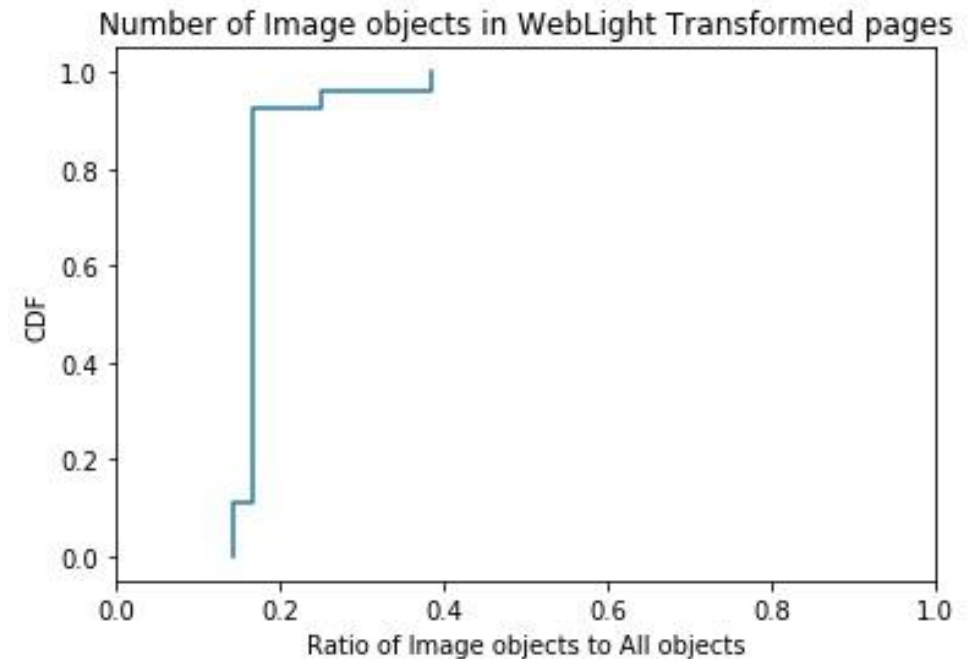
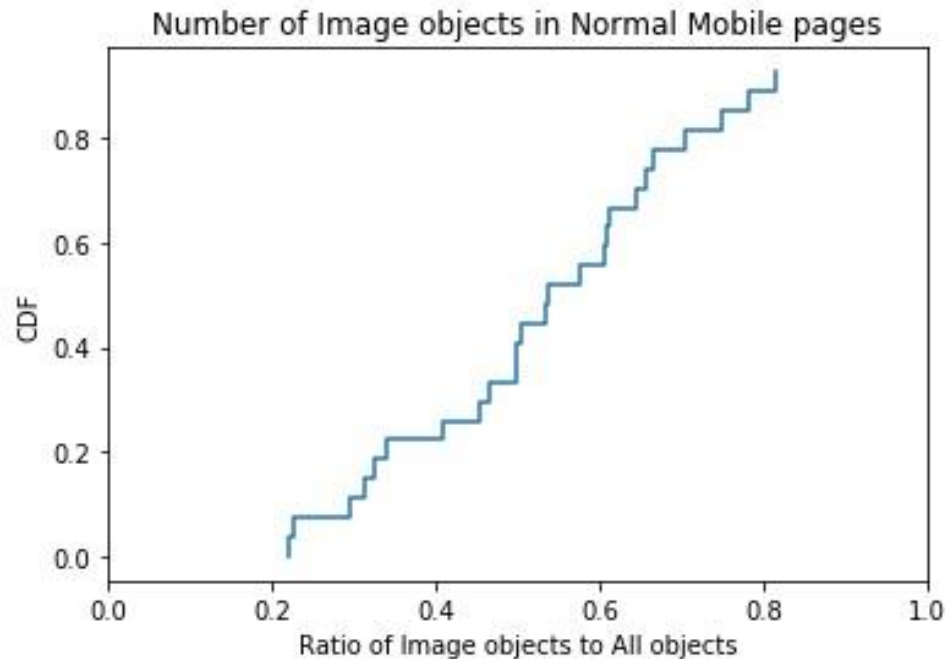
Ratio of Js and CSS Objects

- The transformed pages in the sample dataset made no external CSS or JS object requests.



Ratio of Image Objects

- The ratio of image objects to total objects fetched is significantly lower in transformed pages as most of the images are completely embedded inside the html in the transformed pages.

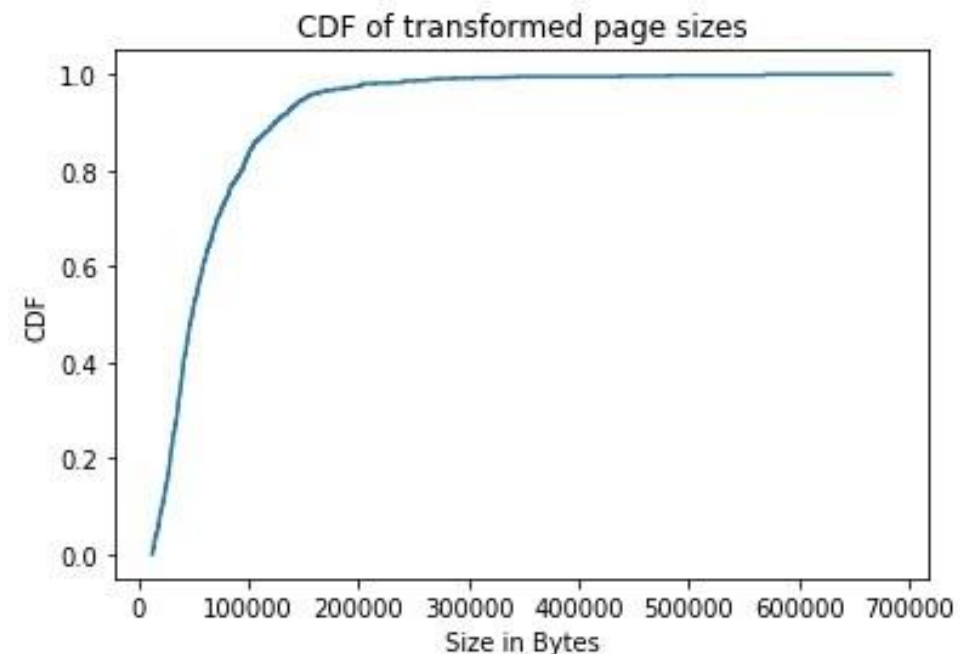
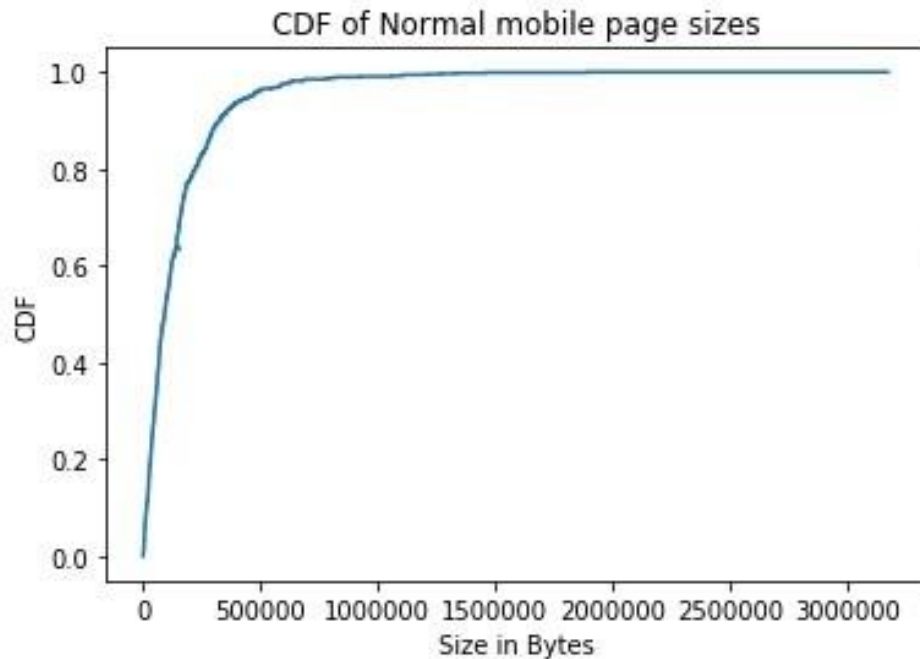


Internal Analysis of the underlying HTML

- Parsing of both the transformed and non-transformed HTML pages
- Comparison of different kind of HTML Tags (Image, Script, Division , Hyperlink) in transformed and non-transformed pages.
- Average Difference of tags computed as:
Avg of [tag] in non-transformed pages - Avg of [tag] in transformed pages

Page Size Comparison

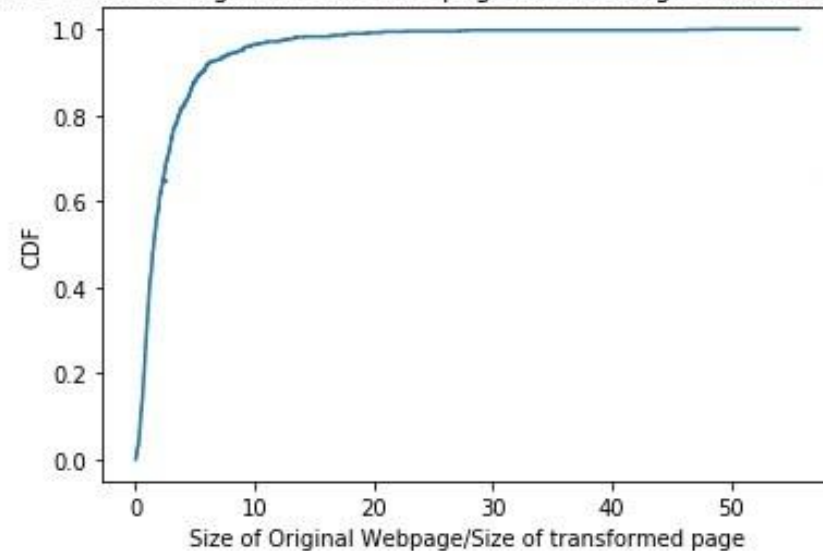
- The CDF plots clearly indicate that the page size of transformed pages is generally lesser than normal pages.



Ratio of Page Sizes

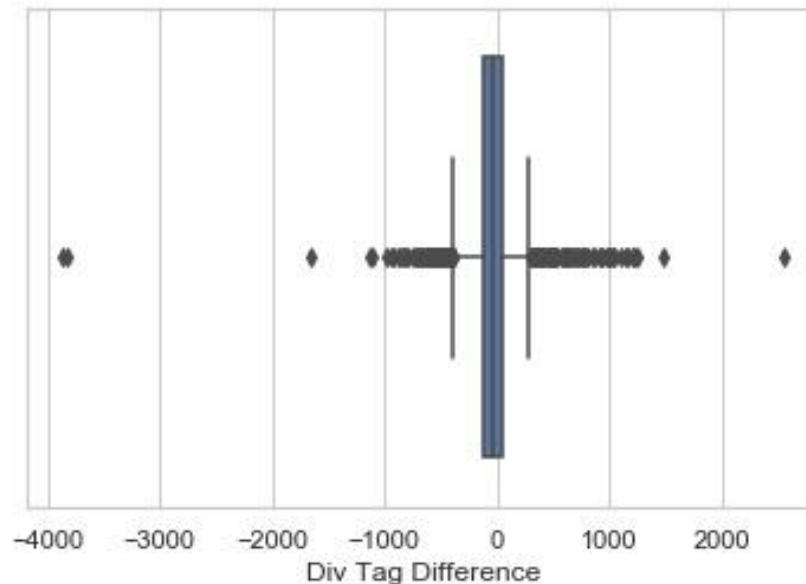
- In the average case, the plot shows that there is a 2 to 3 times decrease in page size for transformed pages.

Ratio of size of Original Mobile web page and WebLight Transformed pages



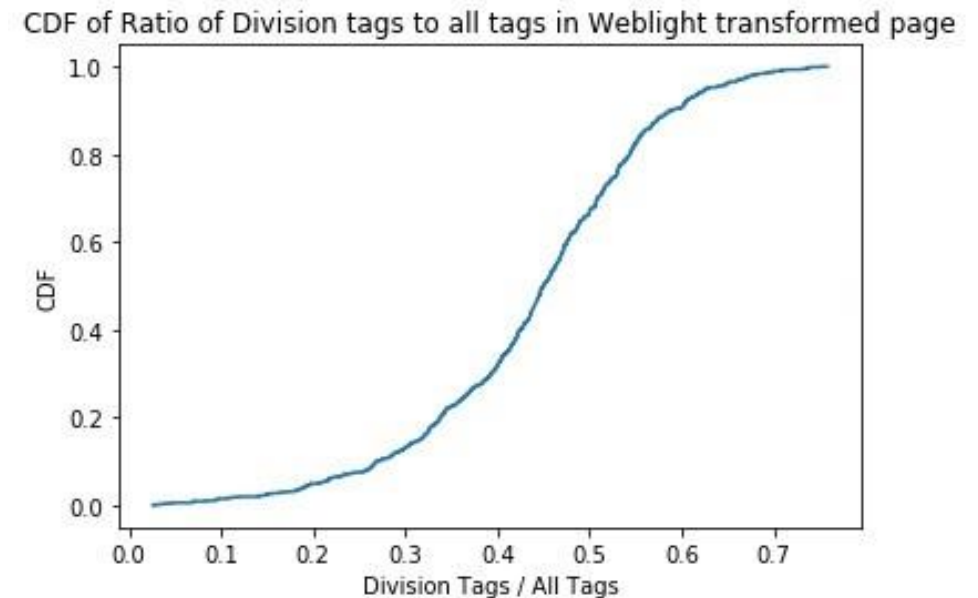
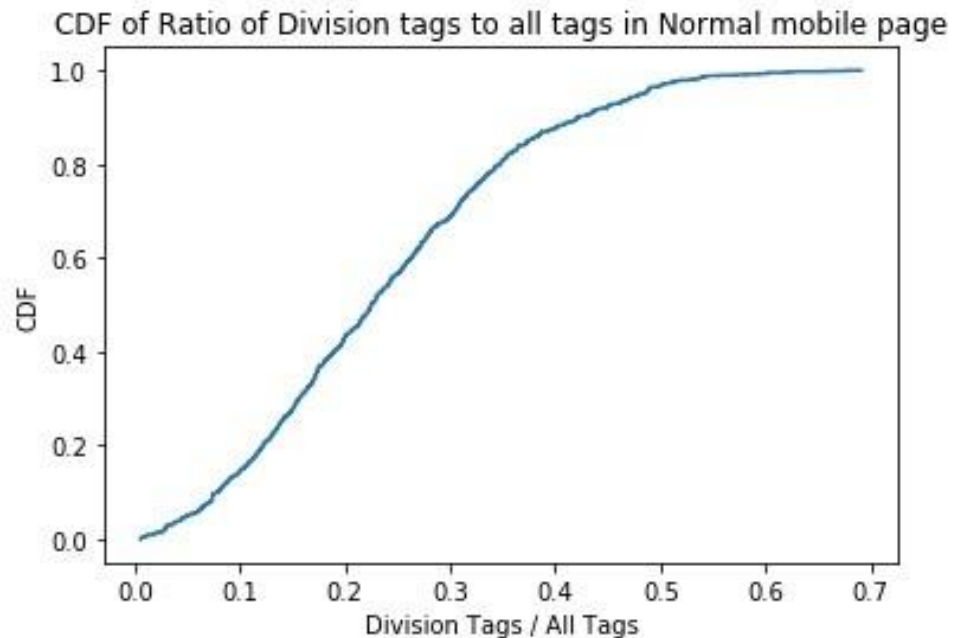
Difference of Division Tags

- The division tags in the average case decrease in transformed pages due to decreased content. However, in some cases, the division tags may increase in transformed pages due to reformatting and embedded content.



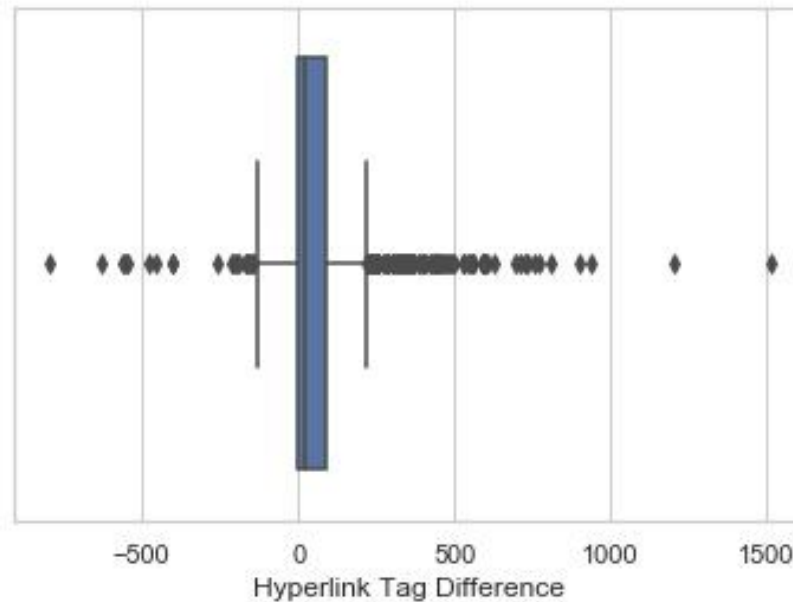
Ratio of Div Tags

- The ratio of div tags to all tags in normal pages is smaller in number as compared to that of transformed pages in the average case. The increased ratio of div tags in transformed pages suggest the removal of other tags in the html document.



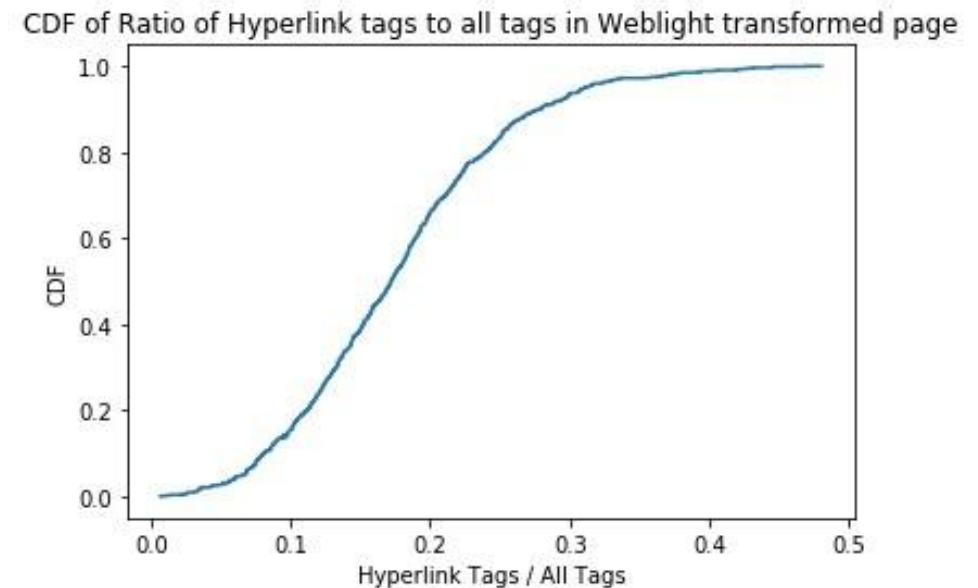
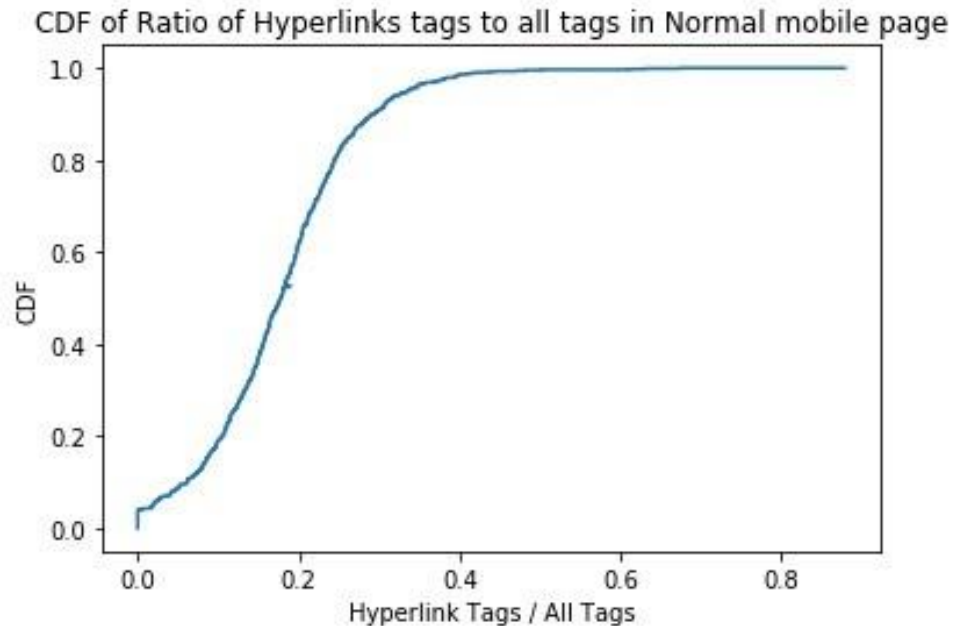
Difference of Hyperlink Tags

- The hyperlink tags in the average case decrease in transformed pages due to decreased content and removal of advertisements and navigation bar. However, there are certain exceptions/outliers.



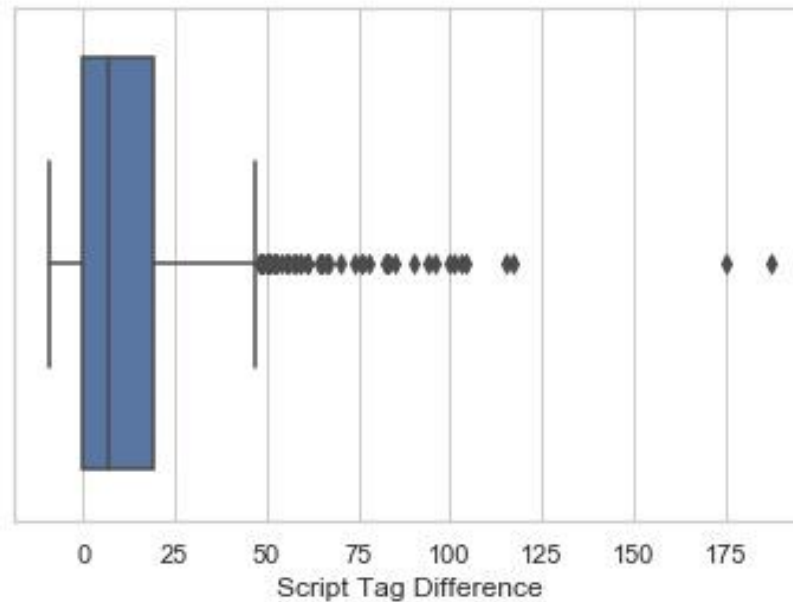
Ratio of Hyperlink Tags

- The ratio of hyperlink tags to total tags in the page approximately displays the same trend for both the transformed and normal pages.



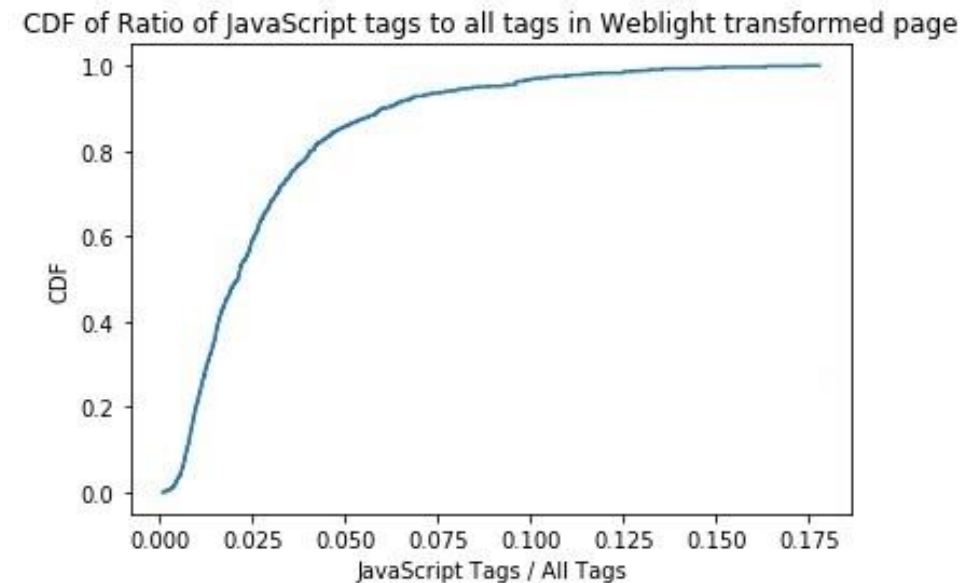
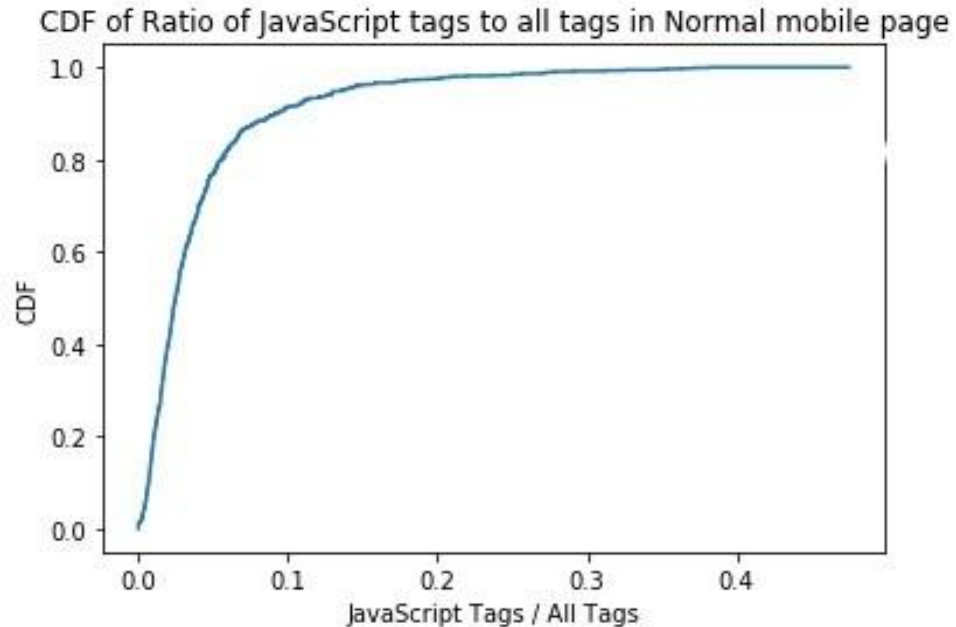
Difference of Script Tags

- The script tags also usually reduce on average in transformed pages, however, the content within the script tags may change.



Ratio of Script Tags

- The ratio of script tags to all tags usually decrease in WebLight transformed pages suggesting the removal of certain JavaScript content in the transformed pages.



Difference of Image Tags

- The image tags slightly decrease in transformed pages on average. However, one possible reason for increased image tags in some transformed pages is that images fetched from hyperlinks etc. are converted to img

