1. What is the problem?
    1) What are the significant metrics that can characterize the complexity of the Web sites, which will affect the client's experience?
    2) Which metrics are most critical for predicting page render and load times.

2. What are the contributions in the paper?
    1) Former studies mostly focus on characterizing the Web graph, network footprint, or the rate of content changing, instead of analyzing the Web sites themselves.
    2) Help Web providers to reduce the complexity and customize for different client platforms.

3. What are the conclusions in the paper and how do the authors support their conclusions (for example using measurement study, experimentation, analysis, modeling, etc).
    1) The studies show that popularity and Web traffic are not as significant as categories, non-origin sources, and the number of servers to the complexity.
    2) The paper use correlation analysis, including analyzing the correlation between metrics and complexities in order to find good indicators of metrics.
    3) The paper use regression analysis to analyze the complexity of different metrics.

4. What did you like about the paper (or what is the one new/cool thing you learned from the paper)
    1) In terms of data sources, this paper focused on the landing pages after analyzing the complexity between non-landing pages and landing pages, which can avoid extra meaningless studies.

5. What did you not like about the paper? Is there something you don't agree with within the paper? (or how can you improve the paper)
    1) In terms of analyzing the optimizations for improving performance, researchers used many tools to improve Web performance and emulate the effect in the given dataset. However, the results are not ideal and the advantage is not prominent. For me, I will try to explore more optimizations or develop a tool ourselves to prove it can indeed improve page load times.

Bonus: How do you think you can use the techniques described in the paper to solve another problem in a different domain?
    1) This paper uses LASSO instead of simple linear regression which is more robust and uses a k x 2 cross-validation technique to avoid overfitting. I can use this technique to analyze top k features in data analysis of many other domains such as NLP.