1. **Briefly summarise what the author (you) has done.**

In this report, I introduce three spelling correction methods: GED, N-gram and a hybrid method combining the two. I define four evaluation criteria: accuracy, precision, recall and time and compare the relative advantages and disadvantages of different methods by analyzing their performance on the data set.

1. **Indicate what you think the author (you) has done well, and why.**

After reading other reports and reviewing the feedback received, I concluded that my report did well in a few ways:

* 1. The report is generally clear and informative.
  2. The report analyzed a more comprehensive set of spelling correction methods than most. The report’s examination of the original work of Levenshtein and Kondrak, and the discussion of why certain methods were omitted was unusual and insightful.
  3. The potential improvements section was particularly original and insightful for readers.

1. **Indicate what you think could have been improved, and why**

There were several areas in my report that could have been improved:

1. The structure of the report could be improved by reallocating the analysis in section 4 (Methodology and Algorithms) to section 6 (Discussion). At the time of writing the report, I wanted to describe any strengths and shortcomings gained from reading the original work of Levenshtein and Kondrak in section 4 and develop any insights from the specific data set (from UrbanDictionary) in section 6. In retrospect however, it would have been cleaner to refrain from analyzing the spelling correction methods until section 6. There is also a mistake in labelling in section 4 where the Tiered and Consensus models should be 4.3.1 and 4.3.2 instead of 4.4 and 4.5 respectively.
2. The analysis could have been dramatically improved by analyzing the dataset itself and including additional examples from the output. For example, I could have highlighted how only 595 out of the 716 corrected words appeared in the dictionary. I could have also analyzed the 56 words that the GED method was able to correct while N-gram failed and vice versa.
3. I could have supplemented the existing analysis by including statistics on how widespread the identified problems were. For example, I argued that padding penalizes N-gram for shorter strings. I could have included statistics on how many short strings were present in the dataset and how often N-gram resulted in the incorrect prediction.