

How to compare Training to Test Q & A data

Quality -> How well does the answer document answer the question

1. Compare the Frequency of Overlap of Per Quality vs Query Length
 1. Do longer queries have more good answers and vice versa
2. Query length analysis of user and LPO queries - Query Length
 1. Compare the dataset's query length stats wrt real world user query length stats
3. Query Types
 1. Compare the distribution of types of query intents/questions. Why, When, Who, How, What, Does, Is, Can etc
4. Compare answer/document metadata distribution as function of answer quality