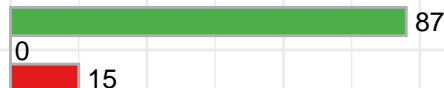


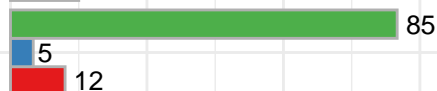
## Distribution of answers to questions about reproducibility of datasets

Question number

Q7.1.1 Does the article mention where the dataset is available



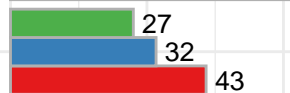
Q7.1.2 Does the article contain information about the number of examples in the training, testing, or development sets?



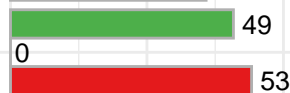
Q7.1.3 Does the article contain information about the number of examples in training, testing, or development sets per category?



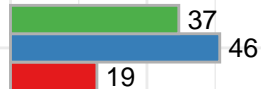
Q7.1.4 Does the article contain information about the dimension of examples in training, testing, or development sets?



Q7.1.5 Does the article mention the version of the dataset explicitly, or can we guess it based on the description of the dataset?



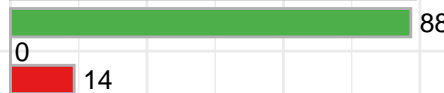
Q7.1.6 Does the article use and explain its pre-processing methods?



Q7.1.7 Is the code for data pre-processing, or a pre-processed copy of the dataset is published in a publicly accessible repository?



Q7.1.8 Does the article use datasets other than 20Newsgroups? If yes, what types of dataset were used?



answer



Yes



Partly



No

0 30 60 90

Count of articles