Distribution of the tokenized articles lengths

