## Distribution of topics in training dataset

