Pearson Correlation Coefficient for Word Counts (10K-GNAD) REGEX **PYTHON** NLTK_BASE WORD_PUNKT NLTK_TWEET NLTK_NIST 0.99 SPACY **GENSIM** 0.99 NLTK_BASE_GER SPACY_GER GENSIM REGEX SPACY NLTK_BASE SPACY_GER **NORD_PUNKT** NLTK_TWEET NLTK_NIST K_BASE_GER

1.00 0.75 0.50 0.25 0.00 -0.25-0.50 -0.75 -1.00