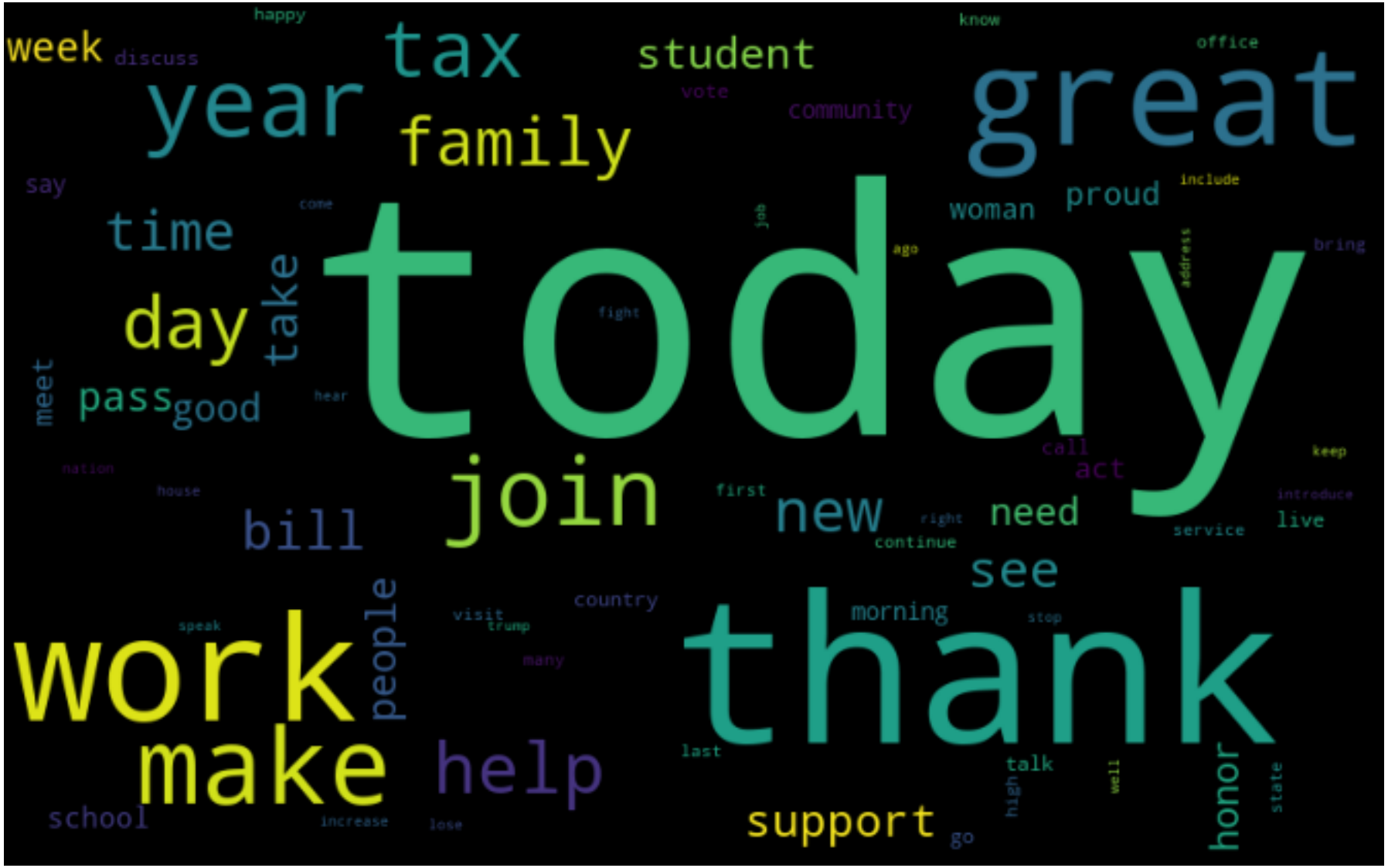Words removed through TF-IDF from individual tweets proportional by their size.