Consider an instance of Euclidean $k$-means or $k$-medians clustering. We show that the cost of the optimal solution is preserved up to a factor of $(1 + \epsilon)$ under a projection onto a random $O(\log(k/\epsilon)/\epsilon^2)$-dimensional subspace. Further, the cost of every clustering is preserved within $(1 + \epsilon)$. More generally, our result applies to any dimension reduction map satisfying a mild sub-Gaussian-tail condition. Our bound on the dimension is nearly optimal. Additionally, our result applies to Euclidean $k$-clustering with the distances raised to the $p$-th power for any constant $p$.

For $k$-means, our result resolves an open problem posed by Cohen, Elder, Musco, Musco, and Persu (STOC 2015); for $k$-medians, it answers a question raised by Kannan. (Received February 01, 2019)