The increasing availability of public datasets offers an inexperienced opportunity to conduct data-driven studies. Metric Multi-dimensional Scaling aims to find a low-dimensional embedding of the data, preserving the pairwise dissimilarities amongst the data points in the original space. Along with the visualizability, this dimensionality reduction plays a pivotal role in analyzing and discovering the hidden structures in the data. This work introduces Sparse Kernel-based Least Squares Multi-dimensional Scaling, a least-squares multi-dimensional scaling approach for exploratory data analysis and, when desirable, data visualization. Via the use of sparsity-promoting regularizers, the technique is capable of embedding data on a, typically, lower-dimensional manifold by naturally inferring the embedding dimension from the data itself. In the process, key training samples are identified, whose participation in the embedding map’s kernel expansion is most influential. As we will show, such influence may be given interesting interpretations in the context of the data at hand. Furthermore, given appropriate positive-definite kernel functions, its kernel-based nature allows for such embeddings of data even with non-numerical features. The resulting multi-kernel learning, non-convex framework can be effectively trained via a block coordinate descent approach, which alternates between an accelerated proximal average method-based iterative majorization for learning the kernel expansion coefficients and a simple quadratic program, which deduces the multiple-kernel learning coefficients. Experimental results showcase potential uses of the proposed framework on artificial data as well as real-world datasets, that underline the merits of our embedding framework. Our method discovers genuine hidden structures in the data that, in the case of network data, matches the results of the well-known Multi-level Modularity Optimization community structure detection algorithm.