Data Mining vs Machine Learning vs Statistics vs AI

Below are excerpts from StackExchange:
https://stats.stackexchange.com/questions/5026/what-is-the-difference-between-data-mining-statistics-machine-learning-and-ai
Google: "data mining versus statistics versus machine learning versus AI"

Artificial intelligence is fairly distinct from the rest. AI is the study of how to program a computer to behave and perform a task as an intelligent agent (say, a person) would. This does not have to involve learning or induction at all.

Machine learning is a large area within AI. Machine learning involves the study of algorithms that can extract information automatically (i.e., without on-line human guidance) and learn. A computer program is said to learn some task from experience if its performance at the task improves with experience. Some of these procedures include ideas derived directly from, or inspired by, classical statistics, but they don't have to be. Similarly to AI, machine learning is very broad and can include almost everything, so long as there is some inductive component to it.

Data mining is an area that has taken much of its inspiration and techniques from machine learning (and some, also, from statistics), but is put to different ends. Data mining is carried out by a person, in a specific situation, on a particular data set, with a goal in mind. Typically, this person wants to leverage the power of the various pattern recognition techniques that have been developed in machine learning. Quite often, the data set is massive, complicated, and/or may have special problems (such as there are more variables than observations). Usually, the goal is either to discover / generate some preliminary insights in an area where there really was little knowledge beforehand, or to be able to predict future observations accurately. Moreover, data mining procedures could be either 'unsupervised' (we don't know the answer--discovery) or 'supervised' (we know the answer--prediction). Note that the goal is generally not to develop a more sophisticated understanding of the underlying data generating process.

Statistics is a sub-topic within mathematics. It is largely the intersection of what we know about probability and what we know about optimization. It is mostly understood as more practical and applied in character than other, more rarefied areas of mathematics. As such (and notably in contrast to data mining above), it is mostly employed towards better understanding some particular data generating process. Thus, it usually starts with a formally specified *model*, and from this are derived procedures to accurately extract that model from noisy instances (i.e., estimation--by optimizing some loss function) and to be able to distinguish it from other possibilities (i.e., inferences based on known properties of sampling distributions). The prototypical statistical technique is regression.