

Data Mining

Contents

What is Data Mining?	1
Data mining main components	1
What is the main purpose of data mining?	2
Main data mining techinques	2
Sample Project	2
The Ames Housing Data	2
Brief	2
Introduction to the Ames Housing Price Data Set	2

What is Data Mining?

- Data mining involves exploring and analyzing large blocks of information to glean meaningful patterns and trends.

Data mining main components

- Data mining tools include powerful statistical, mathematical, and analytics capabilities whose primary purpose is to **sift through large sets of data to identify trends, patterns, and relationships to support informed decision-making and planning.**

What is the main purpose of data mining?

- Data mining is the process of finding anomalies, patterns and correlations within large data sets to predict outcomes.
- Using a broad range of techniques, you can use this information to increase revenues, cut costs, improve customer relationships, reduce risks and more.

Main data mining techniques

In recent data mining projects, various major data mining techniques have been developed and used, including **association, classification, clustering, prediction, sequential patterns, and regression**.

Sample Project

The Ames Housing Data

Brief

- In this data mining project, a housing dataset is used which includes all the prices of the different houses. In this project, the dataset for prediction of price is added along with other attributes. Depending on the level of sophistication, you can follow a predictive model with simple techniques such as regressions or other machine learning algorithms that are suitable for the problem. The application of this project is in the real estate companies. This project utilizes algorithms and techniques for price predictions of the houses based on different housing attributes.
- You can carry out exploratory data analysis using R, Google Data Studio, Power BI, Tableau or Excel.
- You can as well extend the solution by fitting an appropriate ML model using R.

Introduction to the Ames Housing Price Data Set

The data set contains information on 2,930 properties in Ames, Iowa, including columns related to:

- house characteristics (bedrooms, garage, fireplace, pool, porch, etc.)
- location (neighborhood)
- lot information (zoning, shape, size, etc.)

- ratings of condition and quality
- sale price

The dataset is available as a csv file. [Click here to see the dataset documentation.](#)