Getting to know your date: →

Date objects and Attribute types: →

Date sets are made up of date objects. A date object represents an entity - in a sales database, The objects may be customers, store items, and sales; in a Medical database, The objects may be Patients;

in a university date base, The objects may be Students, Professors, and Courses.

Date objects are typically described by attributes. Date objects can also be referred to as samples, examples, instances, date Points, or objects. If The date objects are stored in a date bases, They are date tuples. That is, The row of a date base correspond to the date objects, and The columns correspond to The attributes.

What is an attribute:- An attribute is a date field, representing a characterstic or feature of a date object. The noun attribute, dimension, feature, and Variable are often used interchangeably in The literature. The term dimension is commonly used in datawarehousing. machine learning leterature tend to use the term feature, while statisticians prefer The term variable. Date mining and database Professionals commonly use The term attribute. Attributes describing a customer object / customer date base can Include, customer Id, name and address. The distribution of date involving one attribute (or variable) is called univariate. A bivariate distribution involves Two attributes, and so on.

The type of an attribute is determined by the set of Possible values - nominal, binary, or numeric - the ~~[illegible]~~ attribute can have.

Nominal Attributes (=, ≠) :-

Nominal Means "Relating to names".

The values a nominal attribute are symbols or names of things. Each value represents some kind of category, code, or state, and so nominal attributes are also referred to as categorical. The value do not have any meaningful order.

Ex:- Suppose that hair-color and marital-status are two attributes describing person objects. In our application, Possible values for hair-color are black, brown, blond, red, auburn, gray and white.

The attribute marital-status can take on the value single, married, divorced, and widowed. Both hair color and Marital-status are nominal attributes.

Another example of a nominal attribute is occupation, with the values teacher, dentist, Programmer, farmer, and so on.

Although, we said that the values of a nominal attribute are symbols or "names of things". It is Possible to represents such symbols or "names" with numbers. With hair color, for instances, we can assign a code of 0 for black, 1 for brown, and so on. But on these numbers we can not Perform operations like numerical number (quantitative number).

Since, nominal attribute values do not have any meaningful order about them and are not quantative, it makes no sense to find mean value or Median value

Binary Attributes:-

A binary attribute is a nominal attribute with only Two categories or states: 0 or 1 where 0 typically means that the attribute is absent, and 1 means that it is Present. Binary attributes are referred to as Boolean if ~~they~~ the Two states correspond to True and false.

For Ex:- Given The attribute smoker describing a Patient object, 1 indicates that The Patient smokes, while 0 indicates that The Patient does not. Similarly, suppose that Patient undergoes a medical test that has two Possible outcomes. The attribute medical_test is binary, where a value of 1 means the result of the test for The Patient is Possitive, while 0 means The result of Patient is negative.

A binary attribute is symmetric if both of its states are equally valuable and carry the same weight; that is, There is no Preference on which outcome should be coded as 0 or 1. One such example could be the attribute gender having the states male and female. (Balanced data set)

A binary attribute is asymmetric if both of the states are not equally important, such as The Positive and negative outcomes of a medical test for HIV. By convention, we code the most Important outcome, which is usually the rarest one, by 1 (eg. HIV Positive) and The other by 0 (eg. HIV negative)

(<, >)

ordinal Attributes: →

An ordinal attribute is an attribute with Possible values that have a meaningful order or Ranking among Them, but The magnitude b/w successive values is not known.

for Ex:- suppose that drink-size corresponds to the size of drinks available at a fast-food restaurant. This nominal attribute has three Possible values: small, medium, and large. The value have a meaningful sequence (which corresponds to increasing drink size); however, we cannot tell from The values how much bigger, say, a medium is Than a large. other examples of ordinal attributes include grade (eg. A+, A, A-, B+, and so on) and professional rank (assistant, associate, and Professors).

ordinal attributes are useful for registering subjecti assessments of qualities that cannot be measured objective; Thus ordinal attributes are often used in surveys for rating. In one survey, Participants were asked to Rate how satisfied They were as customers. customer satisfaction had The following ordinal categories: 0: very satisfied
1: somewhat dissatisfied
2: neutral
3: satisfied
4: very satisfied.

ordinal attributes may be obtained from The discretization of numeric quantities by splitting The value range into a finite number of ordered categories.

Numeric Attributes: →

A Numeric attribute is quatitative; that is, it is Measurable quantity, represented in integer or real values. Numeric attributes can be interval-scaled or ratio-scaled.

Interval-scaled attributes: → interval-scaled attributes are measured on a scale of equal-size units. The values of interval-scaled attributes have order and can be Positive, 0, or negative. Thus, in addition to Providing a ranking of values, such attributes allow us to compare and quantify the difference b/w values.

A temperature attribute is interval-scaled. Suppose that we have the outdoor temerature value for a number of different days, where each day is an object. By ordering the values, we obtain a ranking of the objects with respect to temerature. In addition we can quantify the differences b/w values. For Example, a temerature of 20°c is five degrees higher Than a temerature of 15°c. Calender dates are another Example.

For Instance, The years 2002 and 2020 are eight years apart.

* For interval attributes, The differences between values are meaningful, i.e, a unit of Measurement exists. (+, −)

For Ex:- Calender dates, temerature in celsius or fahrenheit.

Ratio-scaled Attributes :→

A ratio-scaled attributes is a numeric attribute with an inherent zero point. That is, if a measurement is ratio-scaled, we can speak of a value as being a multiple (or ratio) of another value. In addition, the values are ordered, and we can also compute the difference b/w values, as well as the mean, median and mode.

* for ratio variables, both differences and ratio are meaningful (*, /)

EX: temperature in kelvin, monetary quantities, counts, age, mass, length, electrical current;

Discrete versus continuous Attributes :→

A discrete attribute has a finite or countably infinite set of values, which may or may not be represented as integers. The attributes hair_color, smoker, medical_test, and drink_size each have numeric values, such as 0 and 1 for binary attributes, or the values 0 to 110 for the attribute age.

An attribute is countably infinite if the set of possible values is infinite but the values can be put in a one-to-one correspondance with natural numbers.

for Example, the attribute customer_id is countably infinite. The number of customers can grow to infinity, but in reality the actual set of values is countable (where the values can be put in one-to-one correspondence with the set of integers). Zip code are another example.

If an attribute is not discrete, it is continuous.

The ~~term~~ numeric attribute can be either integers or real number, but continuous value are only real numbers, where continuous attributes are typically represented as floating-point variables.