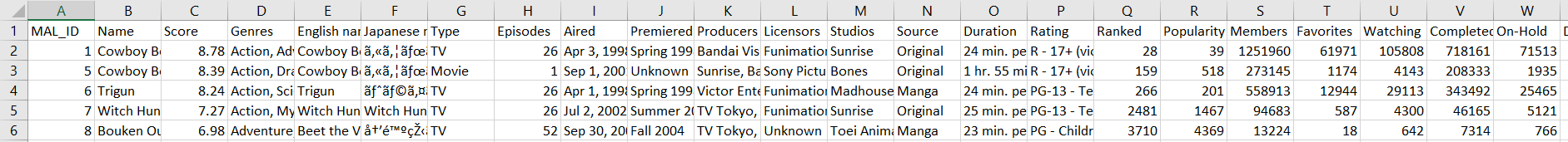
Name – Soumen samanta

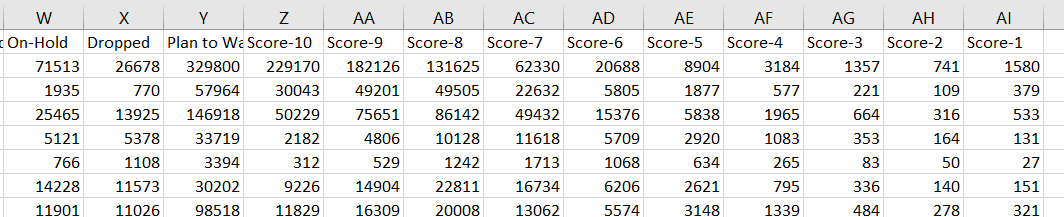
Roll-16010420133

IT B1

Name of data set – anime.csv

<https://www.kaggle.com/hernan4444/anime-recommendation-database-2020?select=anime.csv>





Data attribute of columns: A, C,H ,Q,R,S ,T, U, V, W, X, Y, Z, AA, AB, AC, AD, AE, AF, AG, AH, AI

NUMERIC

Data attribute of columns :B D E F G I J K L M N O P X -CATEGORIAL ATTRIBUTE

POST LAB !!!!

Discrete data can only take particular values. There may potentially be an infinite number of those values, but each is distinct and there's no grey area in between. Discrete data can be numeric -- like numbers of apples -- but it can also be categorical -- like red or blue, or male or female, or good or bad.

Continuous data are not restricted to defined separate values, but can occupy any value over a continuous range. Between any two continuous data values, there may be an infinite number of others. Continuous data are always essentially numeric.

It sometimes makes sense to treat discrete data as continuous and the other way around:

For example, something like height is continuous, but often we don't really care too much about tiny differences and instead group heights into a number of discrete bins -- i.e. only measuring centimetres --.

Conversely, if we're counting large amounts of some discrete entity

-- i.e. grains of rice, or termites, or pennies in the economy -- we may choose not to think of 2,000,006 and 2,000,008 as crucially

different values but instead as nearby points on an approximate

continuum.

It can also sometimes be useful to treat numeric data as categorical, eg: underweight, normal, obese. This is usually just another kind of binning.

It seldom makes sense to consider categorical data as continuous.