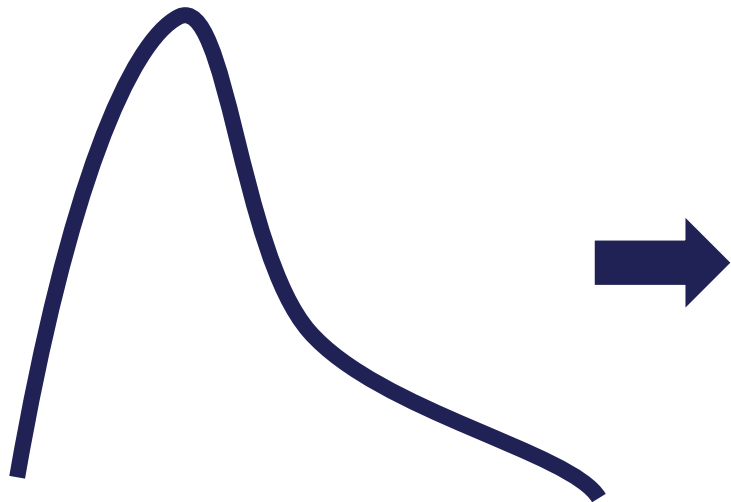




Discretisation

Discretisation



Skewed

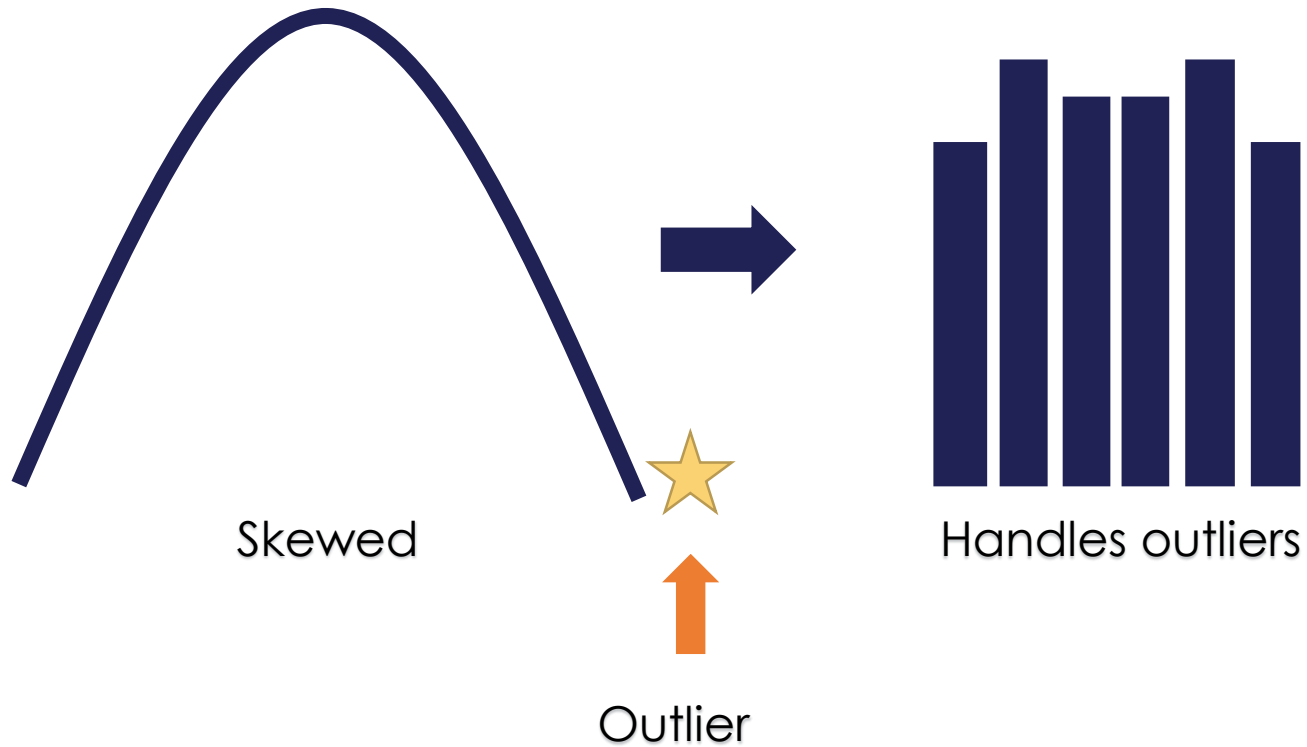


Improved value spread

Discretisation is the process of transforming continuous variables into discrete variables by creating a set of contiguous intervals that span the range of the variable's values.

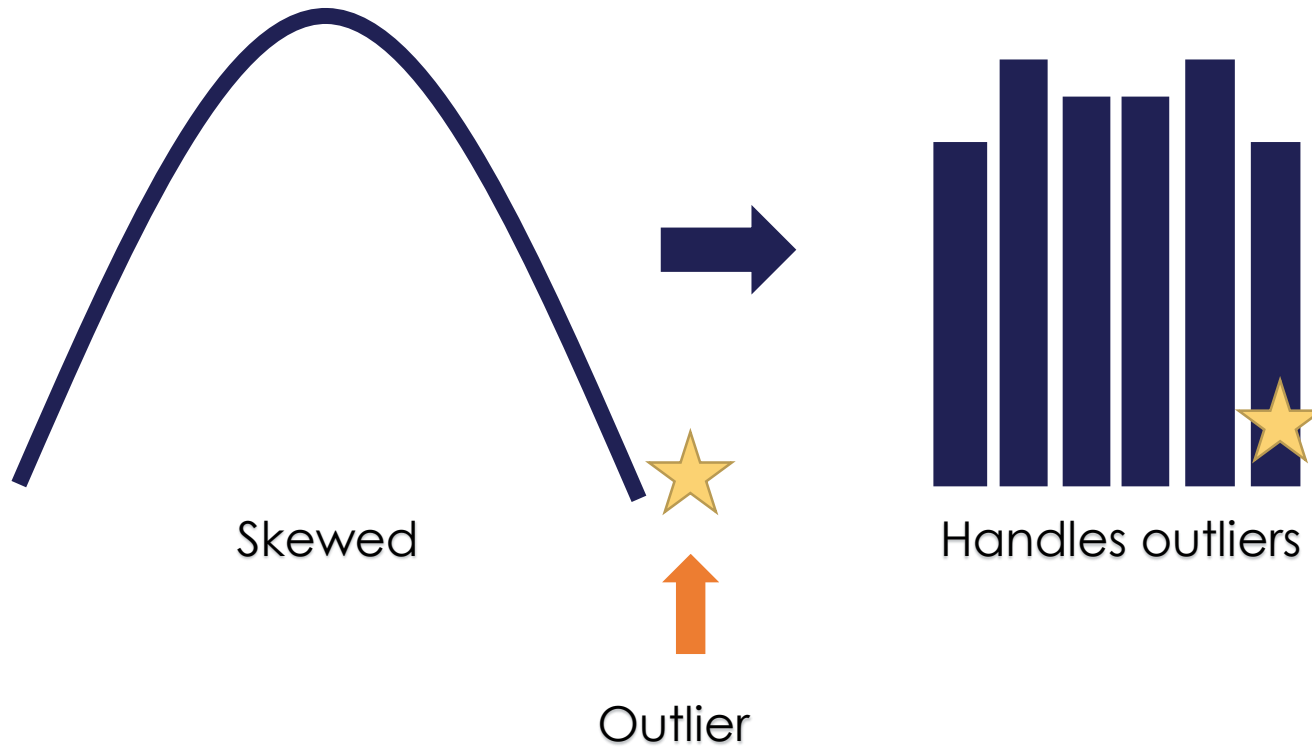
Discretisation is also called **binning**, where bin is an alternative name for interval.

Discretisation



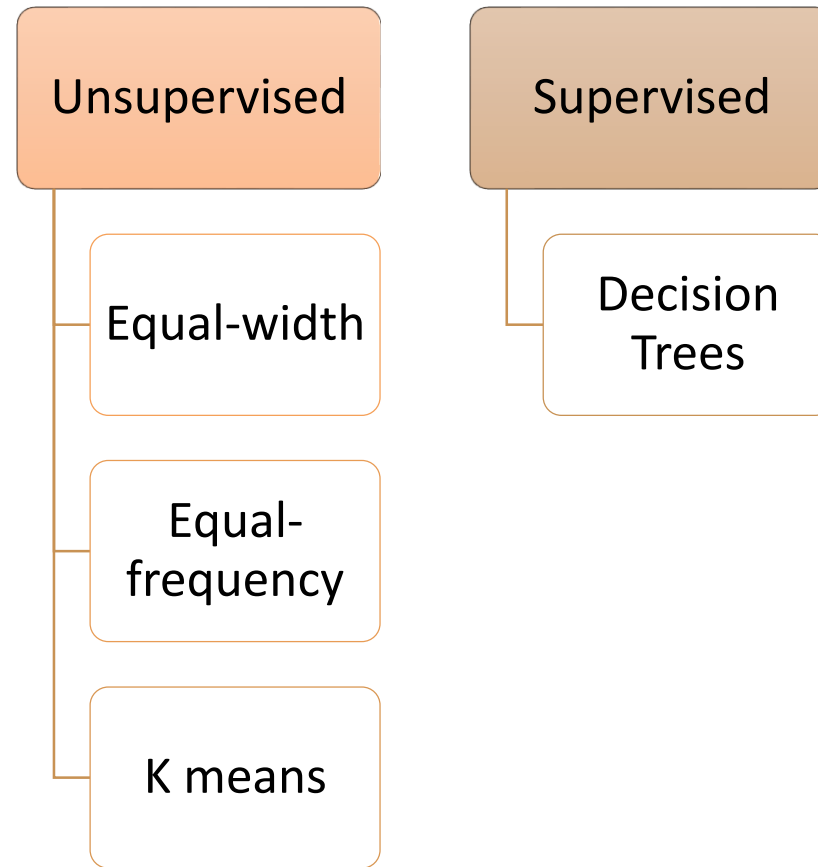
Outliers are placed into the lower or upper intervals, together with the remaining inlier values at the ends of the distribution.

Discretisation



Outliers are placed into the lower or upper intervals, together with the remaining inlier values at the ends of the distribution.

Discretisation approaches



Accompanying Jupyter Notebook



- Accompanying Jupyter Notebook
- How to perform discretisation:
 - Pandas
 - Scikit-learn
 - Feature-engine
- Effect of discretisation on variable distributions

THANK YOU

www.trainindata.com