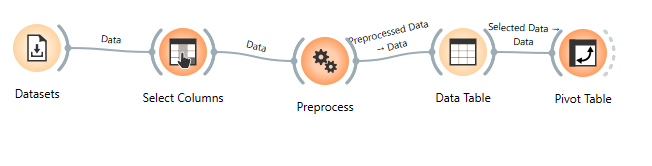
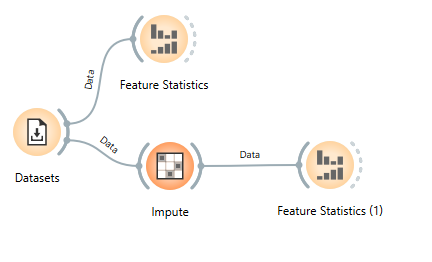
Exp -1 DATA PRE-PROCESSING AND DATA CUBE

In select columns , move all features to ignored  
then choose age,education,income shit back to featured

dataset: adult

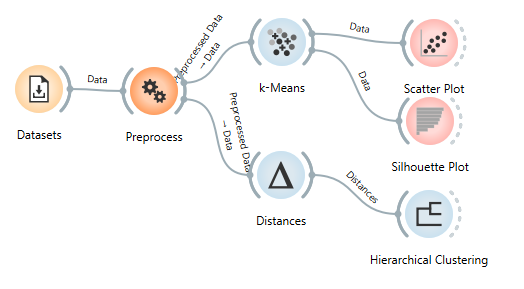
Exp – 2 DATA CLEANING



Dataset: baker’s yeast

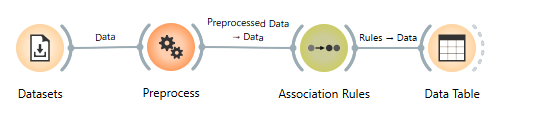
Exp – 3 EXPLORATORY ANALYSIS

k-means & mst:  
in preprocess , choose first option



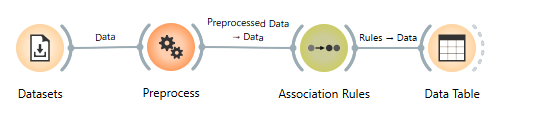
dataset: iris

Exp – 4 ASSOCIATION ANALYSIS  
  
go to options , add ons , associate (double click) , scroll to bottom and choose association rules ,   
preprocess and choose discretize option , choose rules – data (keep any number as rules)



Dataset: market basket

Exp – 5 HYPOTHESIS GENERATION  
same process as exp 4 , keep association rules as 30 , 50  
download associate if not shown from options – add ons



Dataset: market basket

Exp – 6 TRANSFORMATION TECHNIQUES

Haar:

Code:

import numpy as np

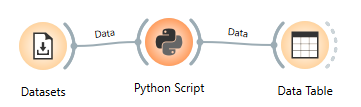
from Orange.data import Table, Domain, ContinuousVariable

data = np.array(in\_data.X)

haar = [[r.mean(), r.std(), r.max()-r.min(), np.sqrt((r\*\*2).sum())] for r in data]

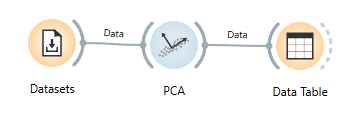
out\_data = Table(Domain([ContinuousVariable(f"Haar\_{i+1}") for i in range(4)]), np.array(haar))

print("Haar transform done")



PCA:

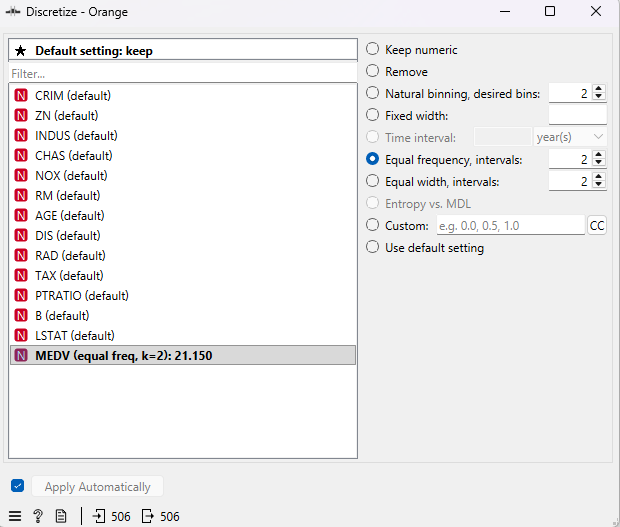
Inside PCA choose 5 in components if question has 5 dimensions

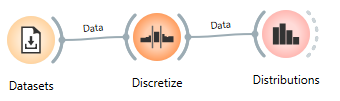


Dataset: Housing

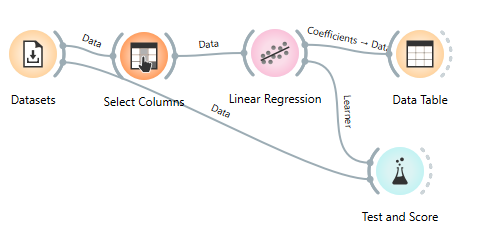
Exp – 7 DATA VISUALIZATION

Binning:



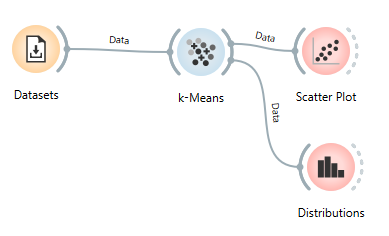


Linear regression:



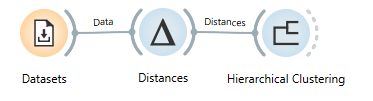
Dataset: Housing

Exp – 8 CLUSTERS ASSESSMENT



Dataset: iris

Exp - 9 HIERARCHICAL CLUSTERING

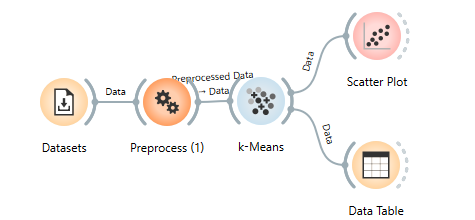


Dataset: iris

Exp – 10 SCALABILITY ALGORITHMS

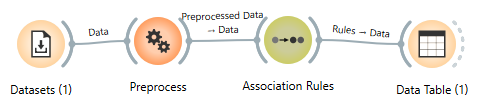
Scalable Clustering (K-Means)

K means choose random initializitation



Dataset:iris

Scalable Apriori Algorithm (Association Rules)



Dataset – market basket