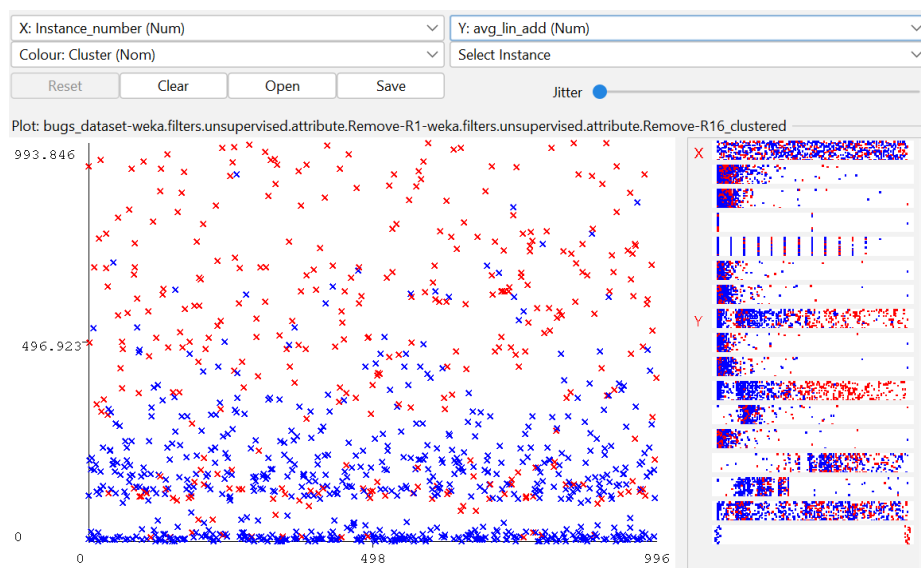
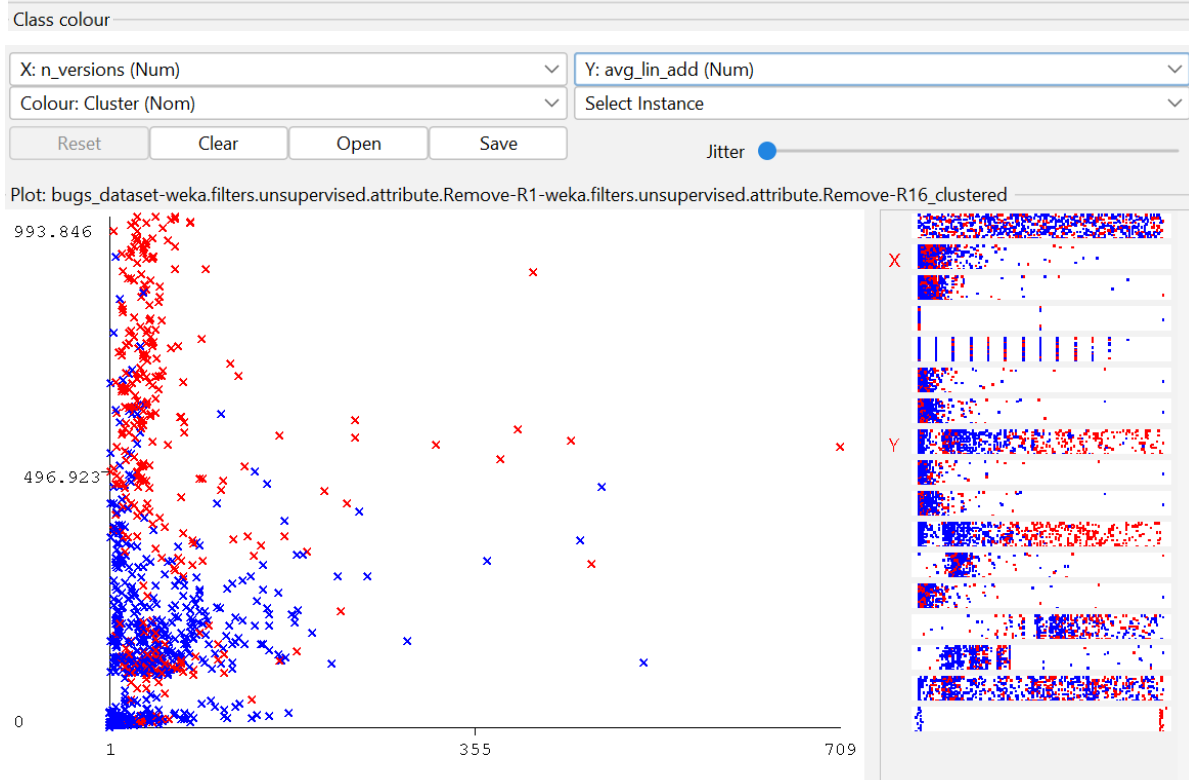
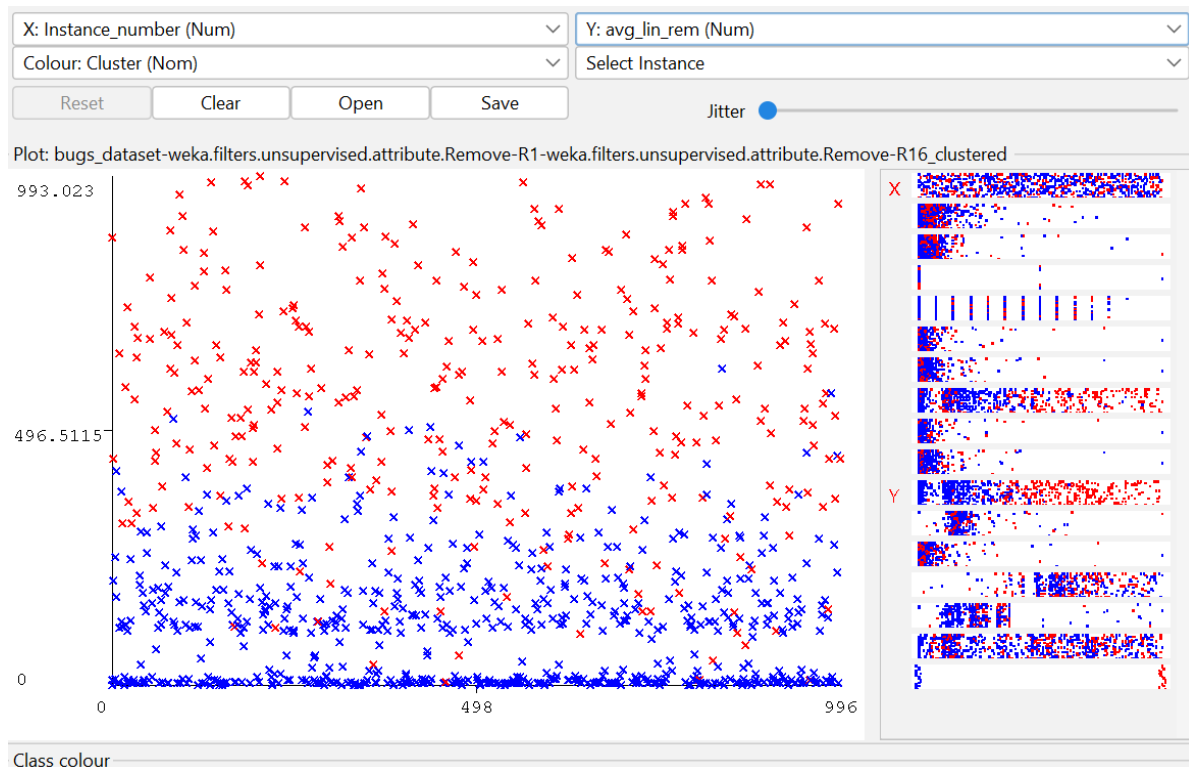


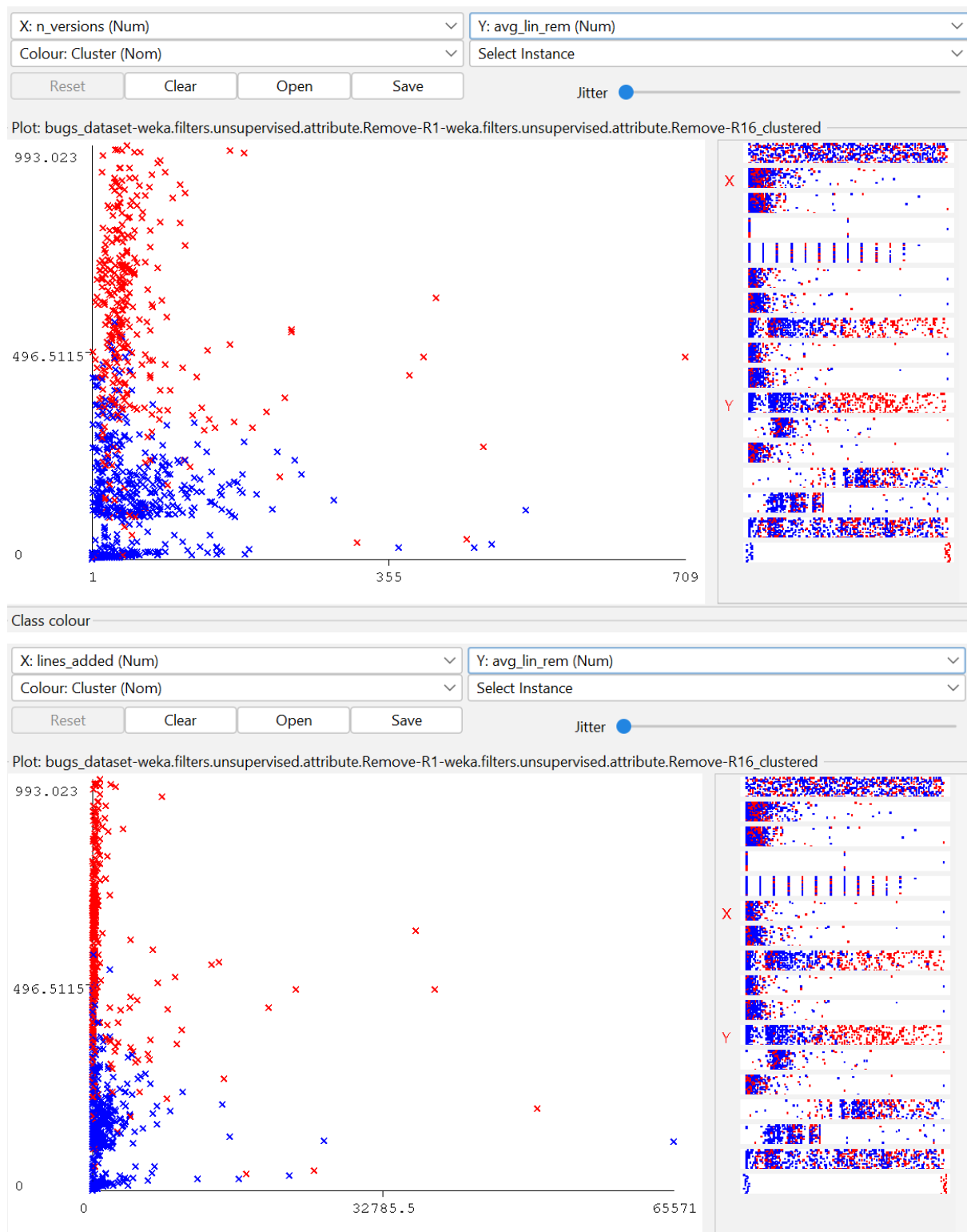
Utilizando o método K-means obtemos as seguintes posições para os centróides:

Final cluster centroids:			
Attribute	Full Data	Cluster#	
		0	1
	(997.0)	(659.0)	(338.0)
=====			
n_versions	45.6189	41.217	54.2012
n_fixes	7.6439	6.8725	9.1479
n_refact	0.01	0.0091	0.0118
n_authors	5.7934	5.3869	6.5858
lines_added	1209.4604	1019.2291	1580.355
max_lin_add	222.2528	209.261	247.5828
avg_lin_add	266.2617	146.9404	498.903
lines_rem	981.2768	820.2003	1295.3284
max_lin_rem	203.2869	189.7511	229.6775
avg_lin_rem	267.2122	113.0976	567.6901
code_churn	228.1836	199.0288	285.0266
max_code_churn	78.2678	72.5857	89.3462
avg_code_churn	134.6316	118.2855	166.5016
age	302.0616	290.3167	324.9607
weighted_age	443.057	409.6414	508.2075

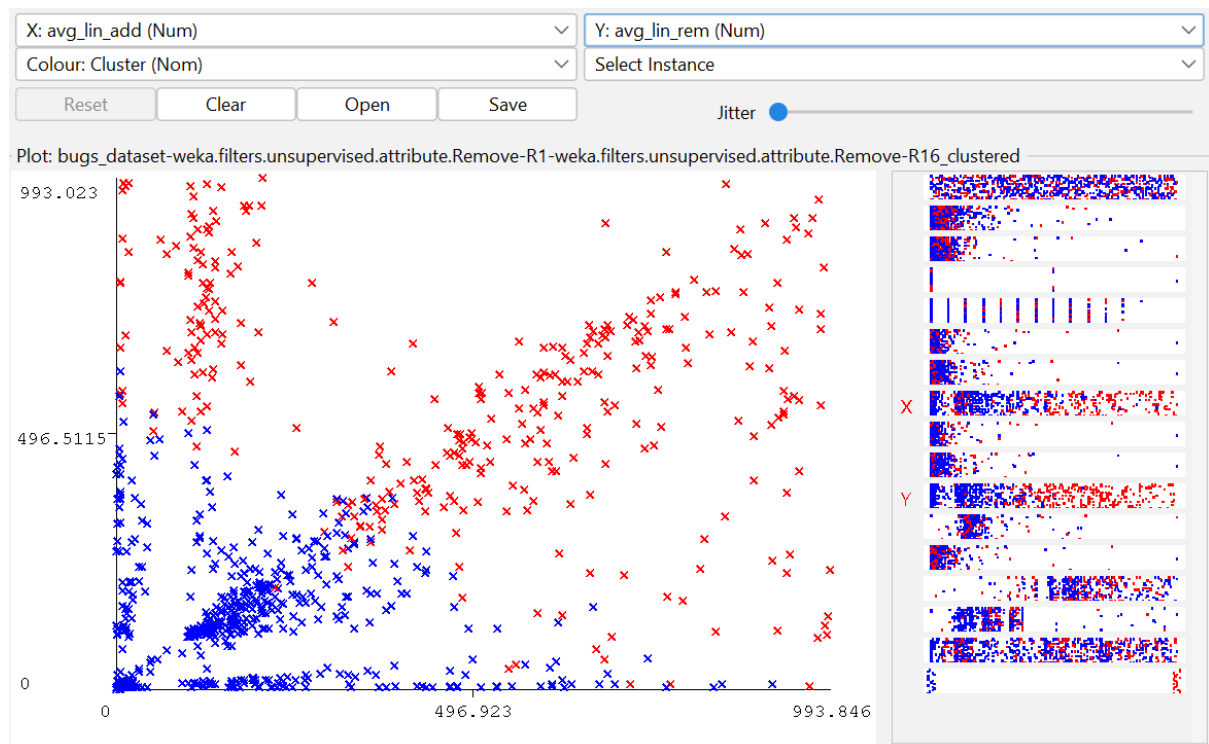
Para tentar definir o método de separação, fica difícil classificar para ver a posição sem a proporção da distância entre os centróides, então partimos para a visualização em que podemos notar resultados mais relevantes com atributos relativos a linhas adicionadas e linhas removidas. Podemos ver essa separação nas imagens a seguir:



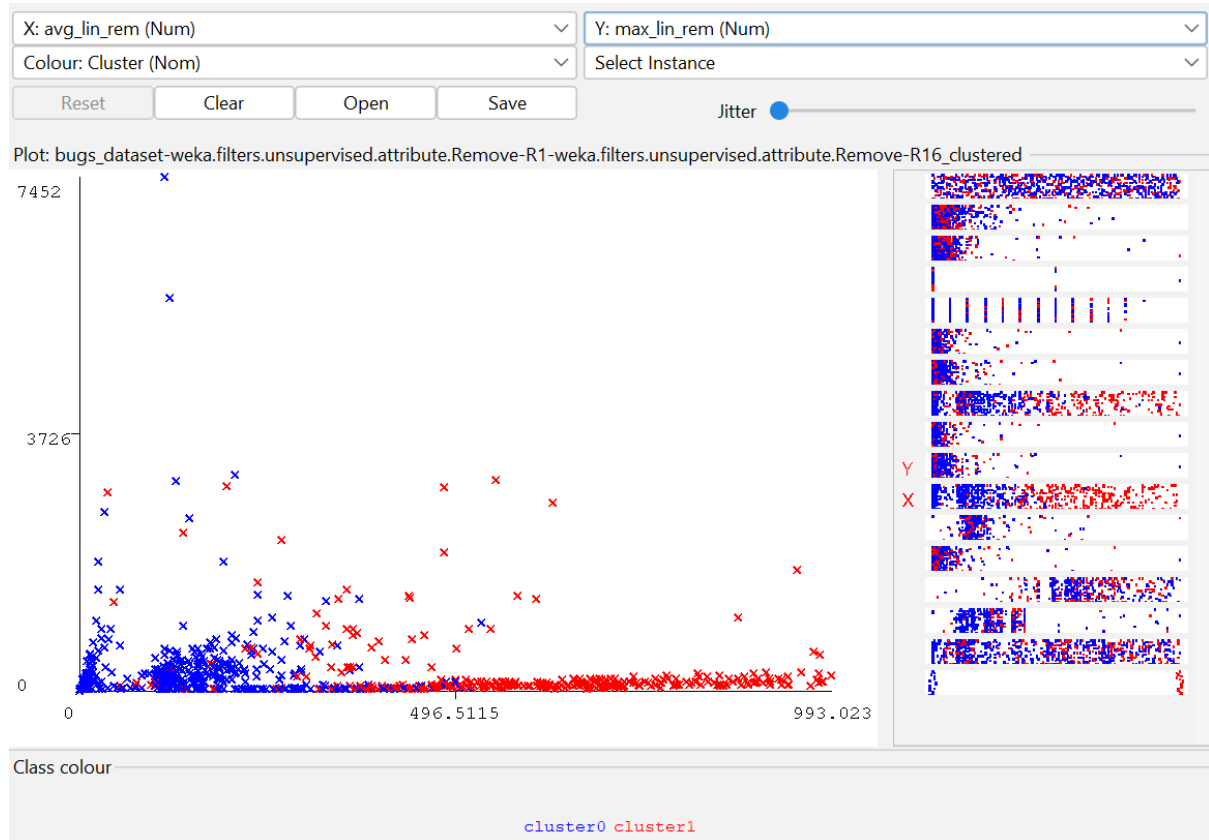




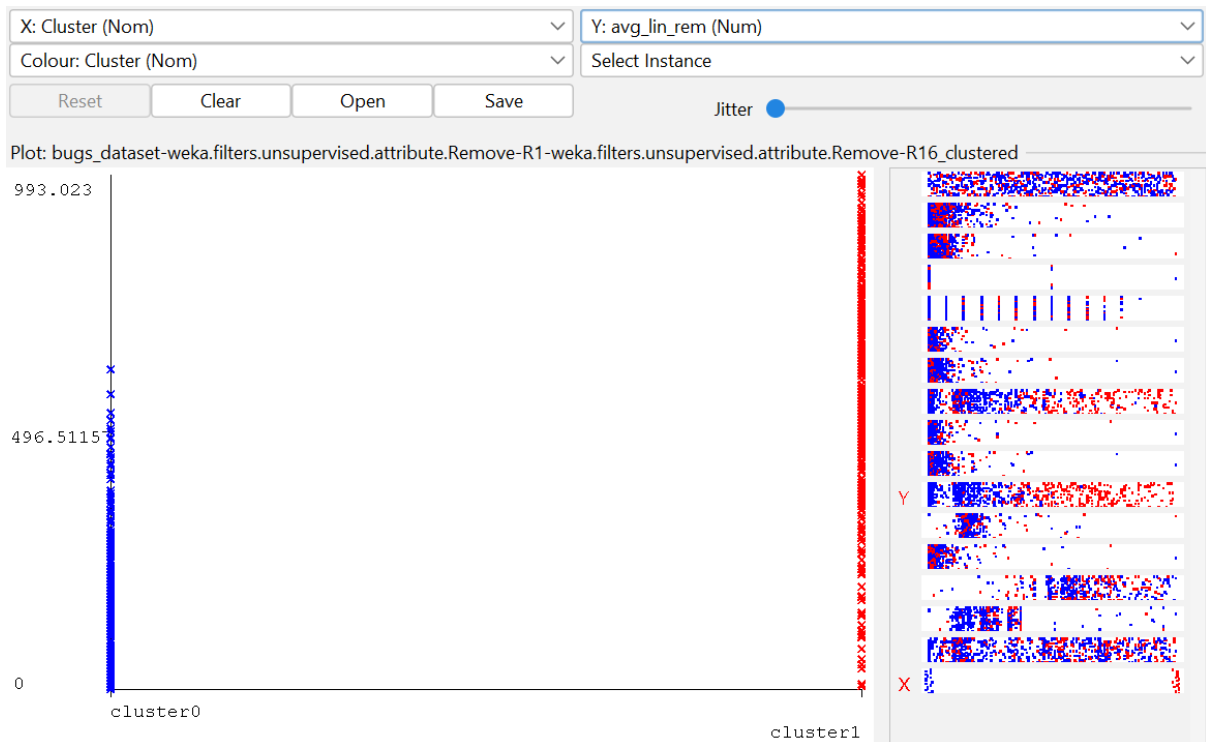
Também podemos visualizar quando comparamos diretamente um com um outro o que indica que a combinação de ambos é um bom preditor da classe:



Alguns atributos também mostram comportamentos interessantes. Por serem de natureza similar, não importa quanto um varia, tendo a separação dada apenas em um eixo:

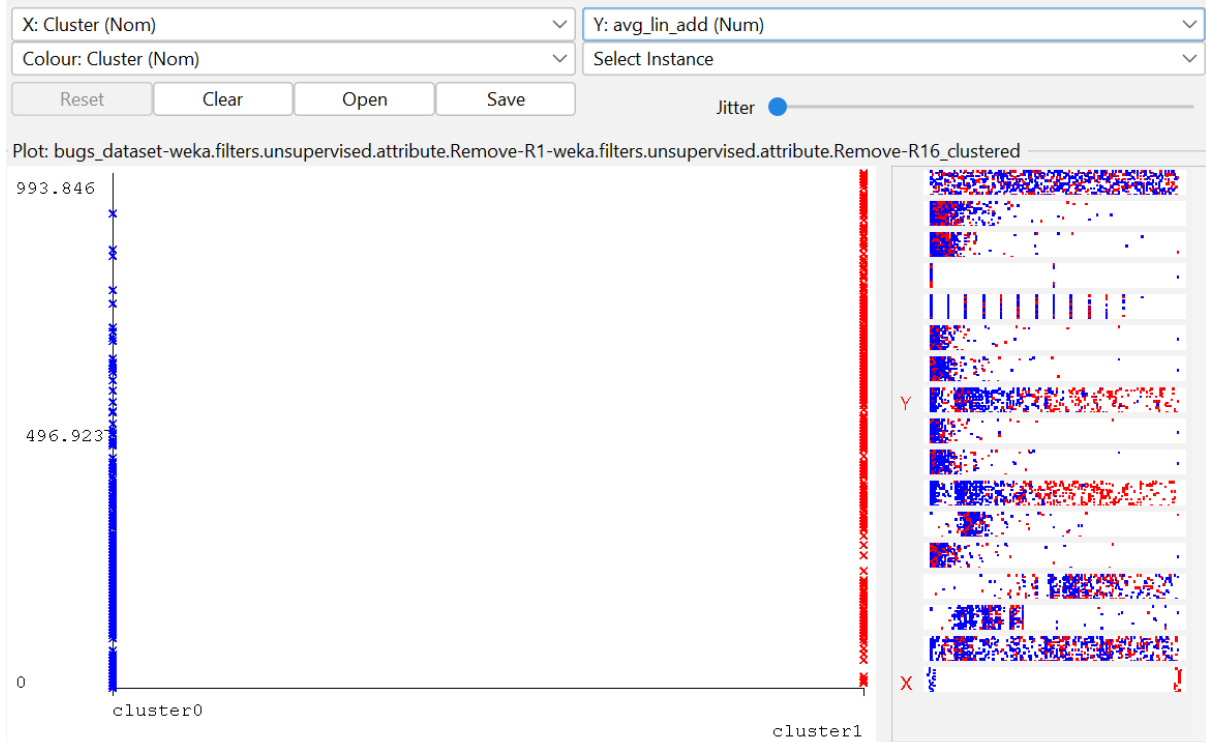


Também podemos comparar com clusters para ver essa correlação com valores altos ou baixos de determinado atributo:



Class colour

cluster0 cluster1



Class colour

cluster0 cluster1