



Отчёт

Анализ погоды с помощью Weka

Студенты:
Сытюк И.В.,
Кисляковский И.О.
группа Р3420
Преподаватель:
Жукова Н.А.

Санкт-Петербург, 2015

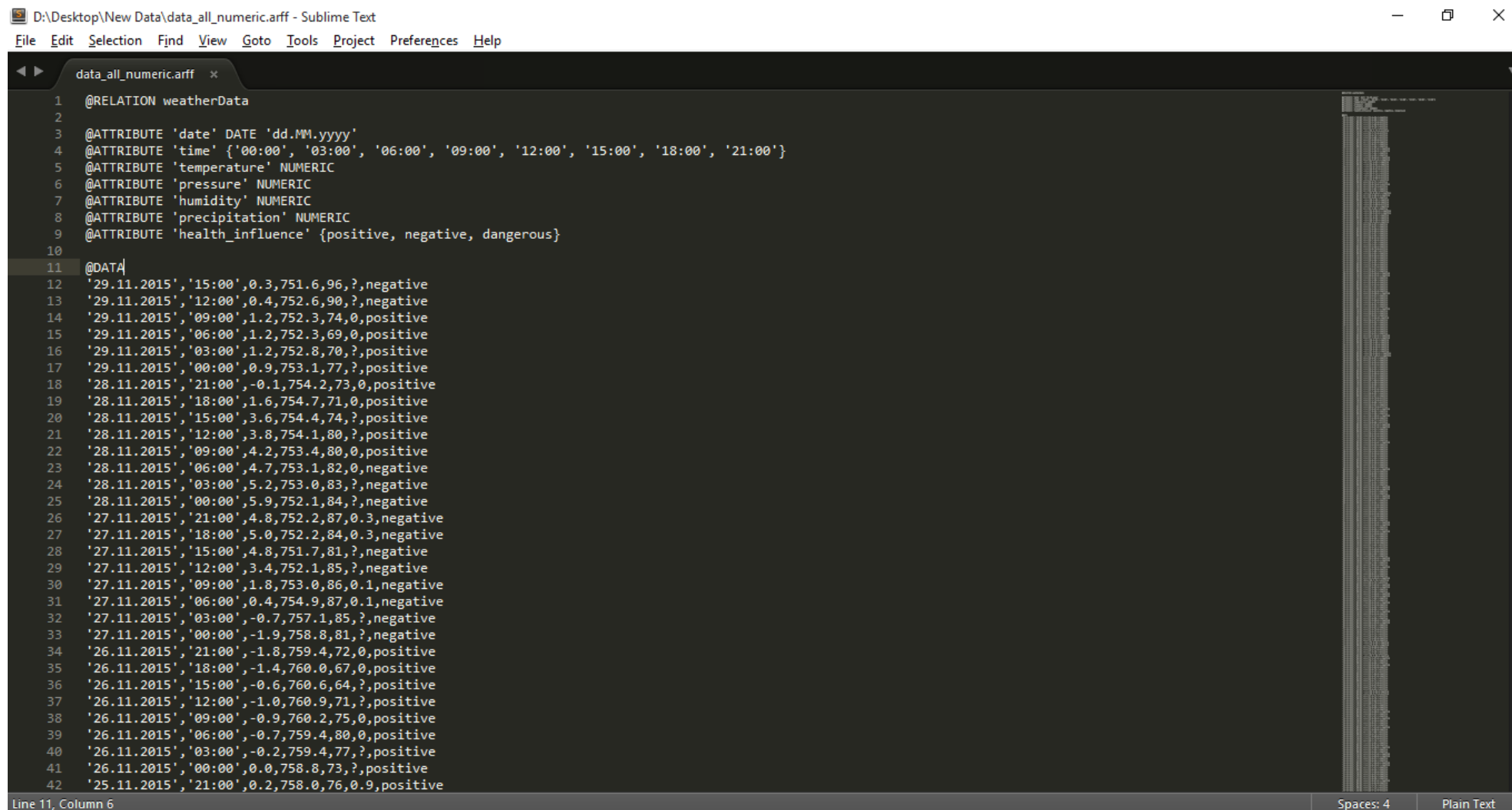
Цель

- Выполнить анализ данных о погоде в Санкт-Петербурге за последние 10 лет

Задачи

- Найти зависимости между атрибутами в наборе данных
 - Выполнить кластеризацию данных и описать полученный результат
 - Выполнить поиск ассоциативных правил
- Выполнить классификацию в соответствии с данными о влиянии погоды на здоровье человека и проанализировать полученные результаты

Данные



```
D:\Desktop\New Data\data_all_numeric.arff - Sublime Text
File Edit Selection Find View Goto Tools Project Preferences Help

data_all_numeric.arff x
1 @RELATION weatherData
2
3 @ATTRIBUTE 'date' DATE 'dd.MM.yyyy'
4 @ATTRIBUTE 'time' {'00:00', '03:00', '06:00', '09:00', '12:00', '15:00', '18:00', '21:00'}
5 @ATTRIBUTE 'temperature' NUMERIC
6 @ATTRIBUTE 'pressure' NUMERIC
7 @ATTRIBUTE 'humidity' NUMERIC
8 @ATTRIBUTE 'precipitation' NUMERIC
9 @ATTRIBUTE 'health_influence' {positive, negative, dangerous}
10
11 @DATA
12 '29.11.2015','15:00',0.3,751.6,96,?,negative
13 '29.11.2015','12:00',0.4,752.6,90,?,negative
14 '29.11.2015','09:00',1.2,752.3,74,0,positive
15 '29.11.2015','06:00',1.2,752.3,69,0,positive
16 '29.11.2015','03:00',1.2,752.8,70,?,positive
17 '29.11.2015','00:00',0.9,753.1,77,?,positive
18 '28.11.2015','21:00',-0.1,754.2,73,0,positive
19 '28.11.2015','18:00',1.6,754.7,71,0,positive
20 '28.11.2015','15:00',3.6,754.4,74,?,positive
21 '28.11.2015','12:00',3.8,754.1,80,?,positive
22 '28.11.2015','09:00',4.2,753.4,80,0,positive
23 '28.11.2015','06:00',4.7,753.1,82,0,negative
24 '28.11.2015','03:00',5.2,753.0,83,?,negative
25 '28.11.2015','00:00',5.9,752.1,84,?,negative
26 '27.11.2015','21:00',4.8,752.2,87,0.3,negative
27 '27.11.2015','18:00',5.0,752.2,84,0.3,negative
28 '27.11.2015','15:00',4.8,751.7,81,?,negative
29 '27.11.2015','12:00',3.4,752.1,85,?,negative
30 '27.11.2015','09:00',1.8,753.0,86,0.1,negative
31 '27.11.2015','06:00',0.4,754.9,87,0.1,negative
32 '27.11.2015','03:00',-0.7,757.1,85,?,negative
33 '27.11.2015','00:00',-1.9,758.8,81,?,negative
34 '26.11.2015','21:00',-1.8,759.4,72,0,positive
35 '26.11.2015','18:00',-1.4,760.0,67,0,positive
36 '26.11.2015','15:00',-0.6,760.6,64,?,positive
37 '26.11.2015','12:00',-1.0,760.9,71,?,positive
38 '26.11.2015','09:00',-0.9,760.2,75,0,positive
39 '26.11.2015','06:00',-0.7,759.4,80,0,positive
40 '26.11.2015','03:00',-0.2,759.4,77,?,positive
41 '26.11.2015','00:00',0.0,758.8,73,?,positive
42 '25.11.2015','21:00',0.2,758.0,76,0.9,positive
```

Line 11, Column 6 Spaces: 4 Plain Text

Кластеризация

- Clusterer - SimpleKMeans

Cluster centroids:

		Cluster#			
Attribute	Full Data	0	1	2	3
	(29054)	(8351)	(7549)	(6569)	(6585)
=====					
temperature	6.701	13.5077	1.7119	14.9865	-4.4094
pressure	759.5137	759.166	751.0944	761.6832	767.4281
humidity	76.7951	81.6027	88.0359	51.3795	82.6989

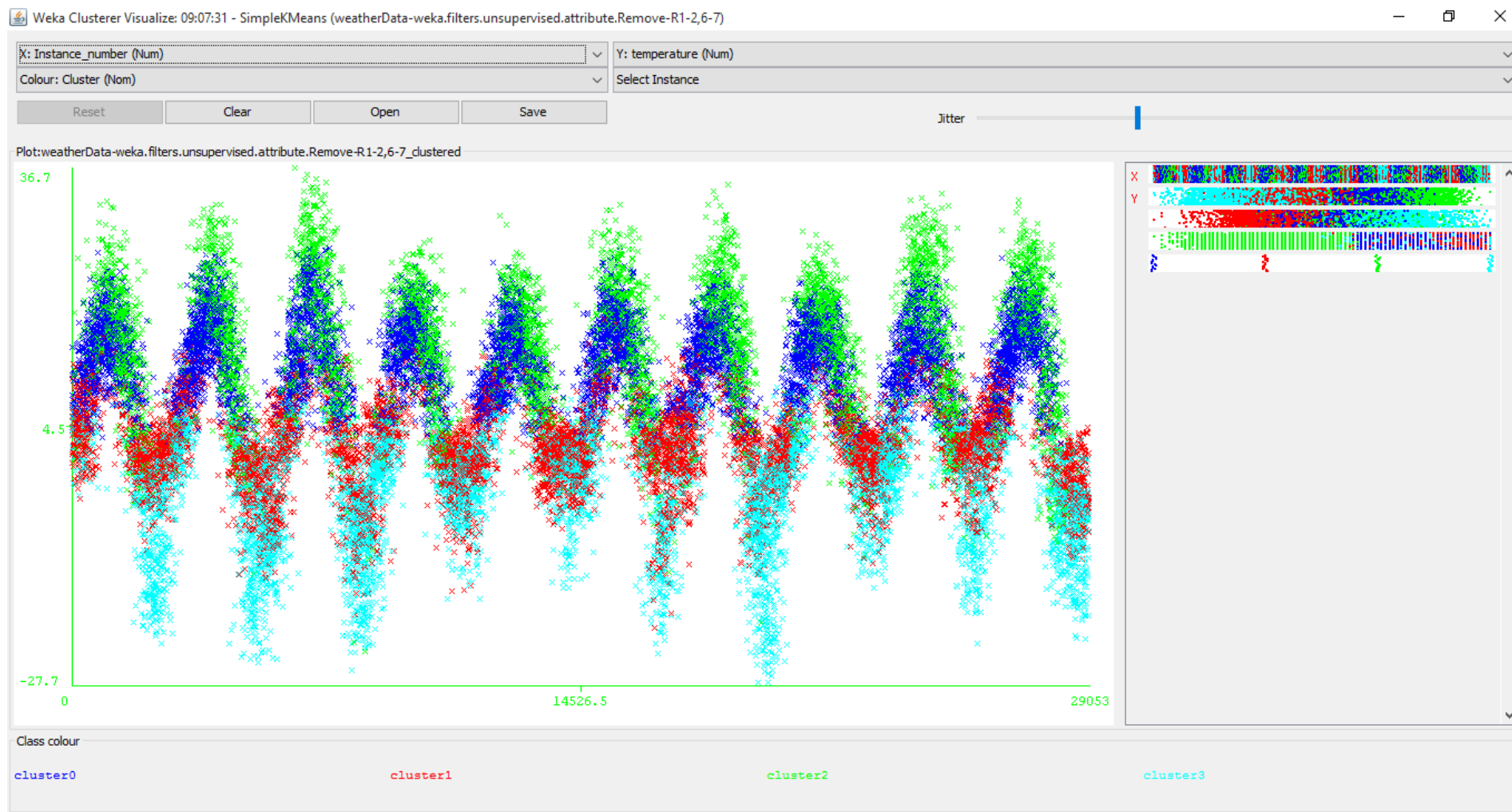
Time taken to build model (full training data) : 1.98 seconds

=== Model and evaluation on training set ===

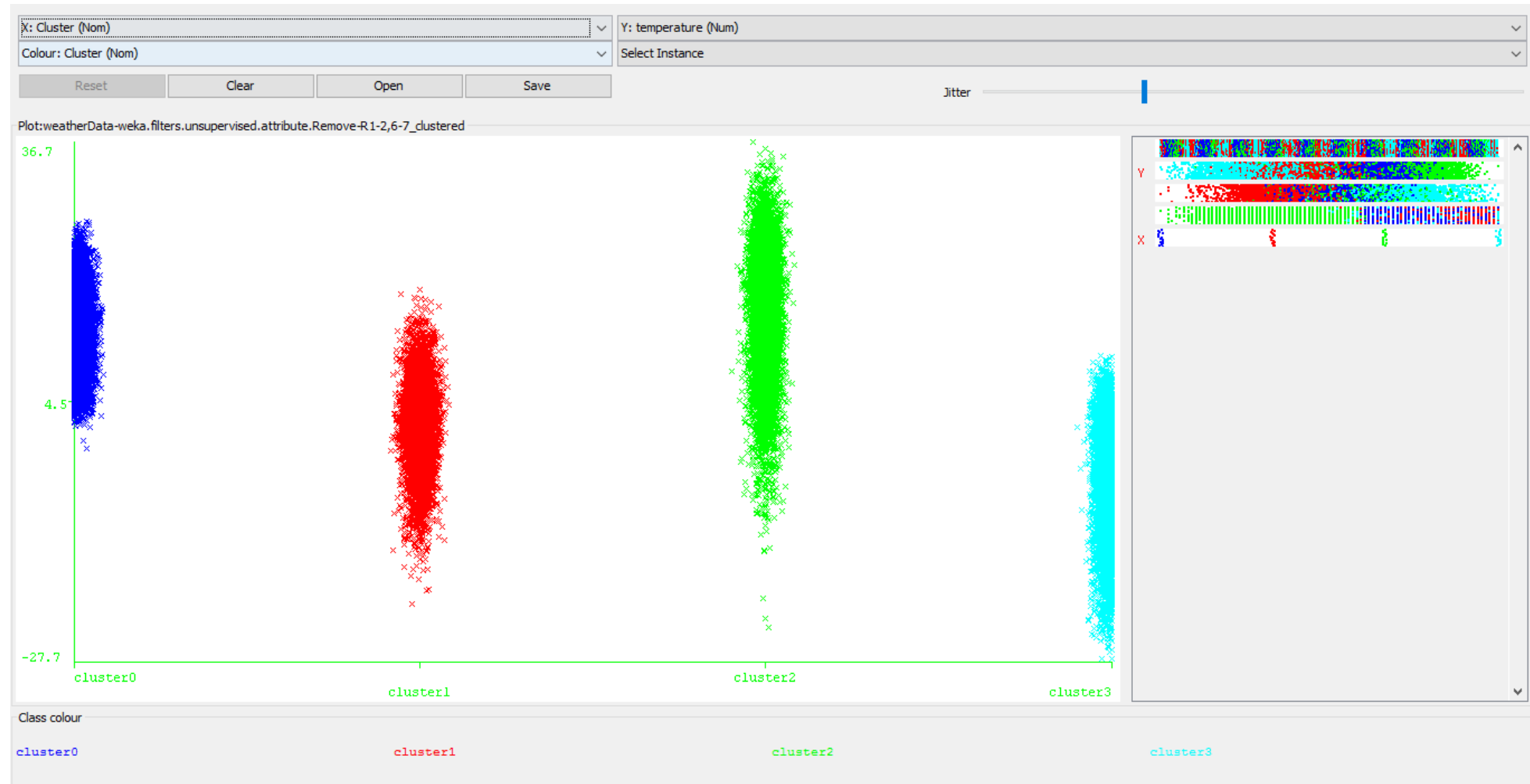
Clustered Instances

0	8351 (29%)
1	7549 (26%)
2	6569 (23%)
3	6585 (23%)

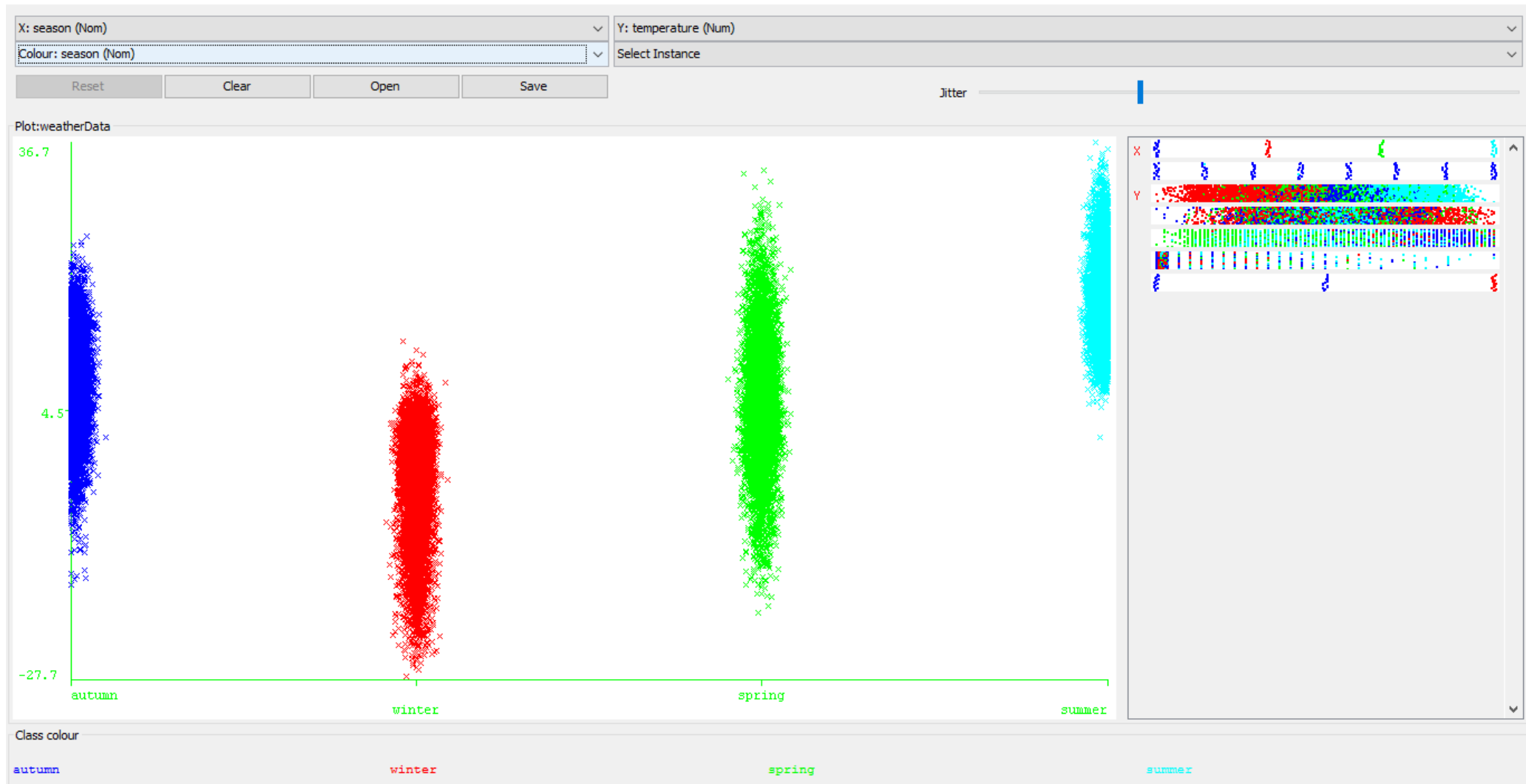
Зависимость температуры от времени



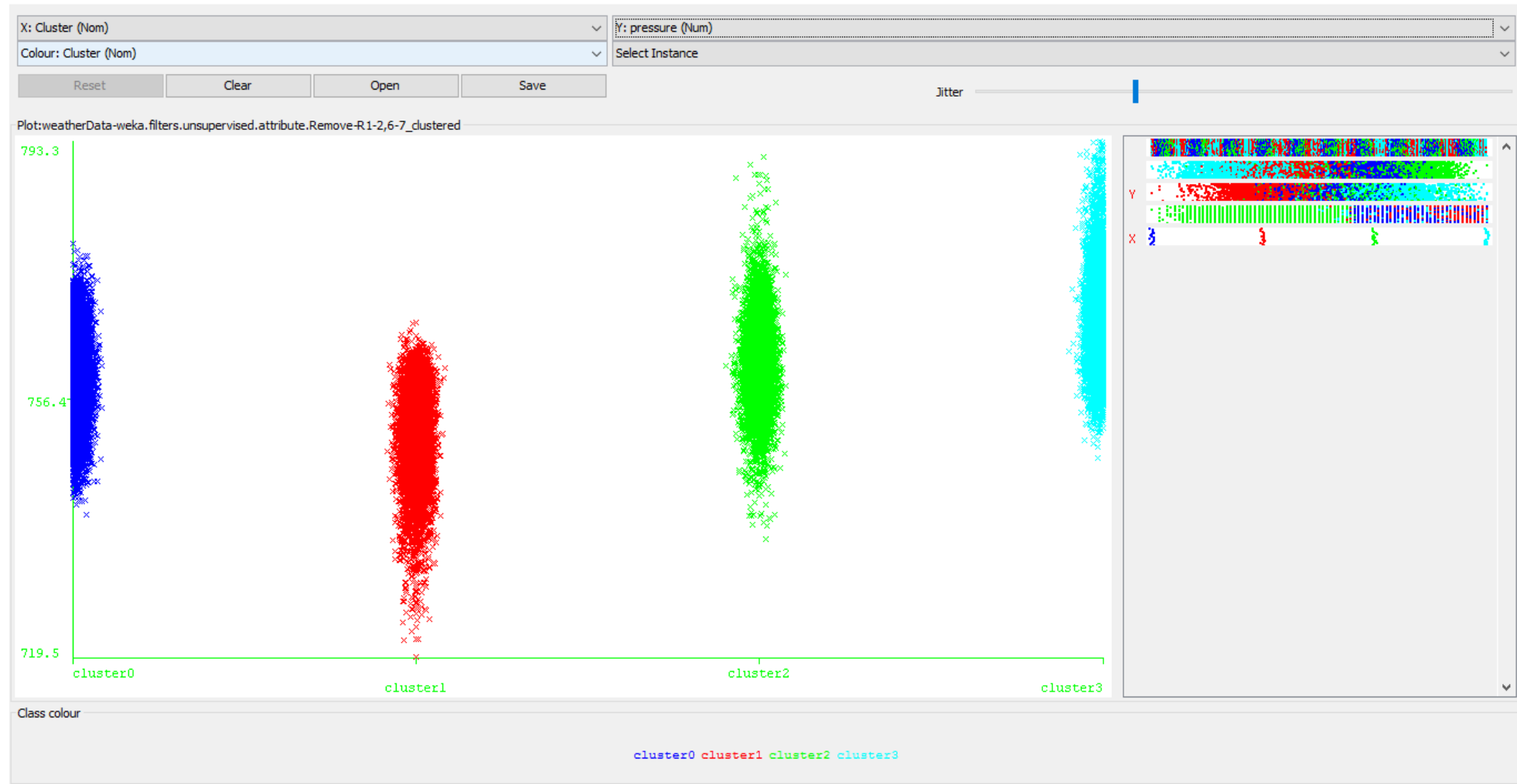
Изменение температуры для каждого кластера



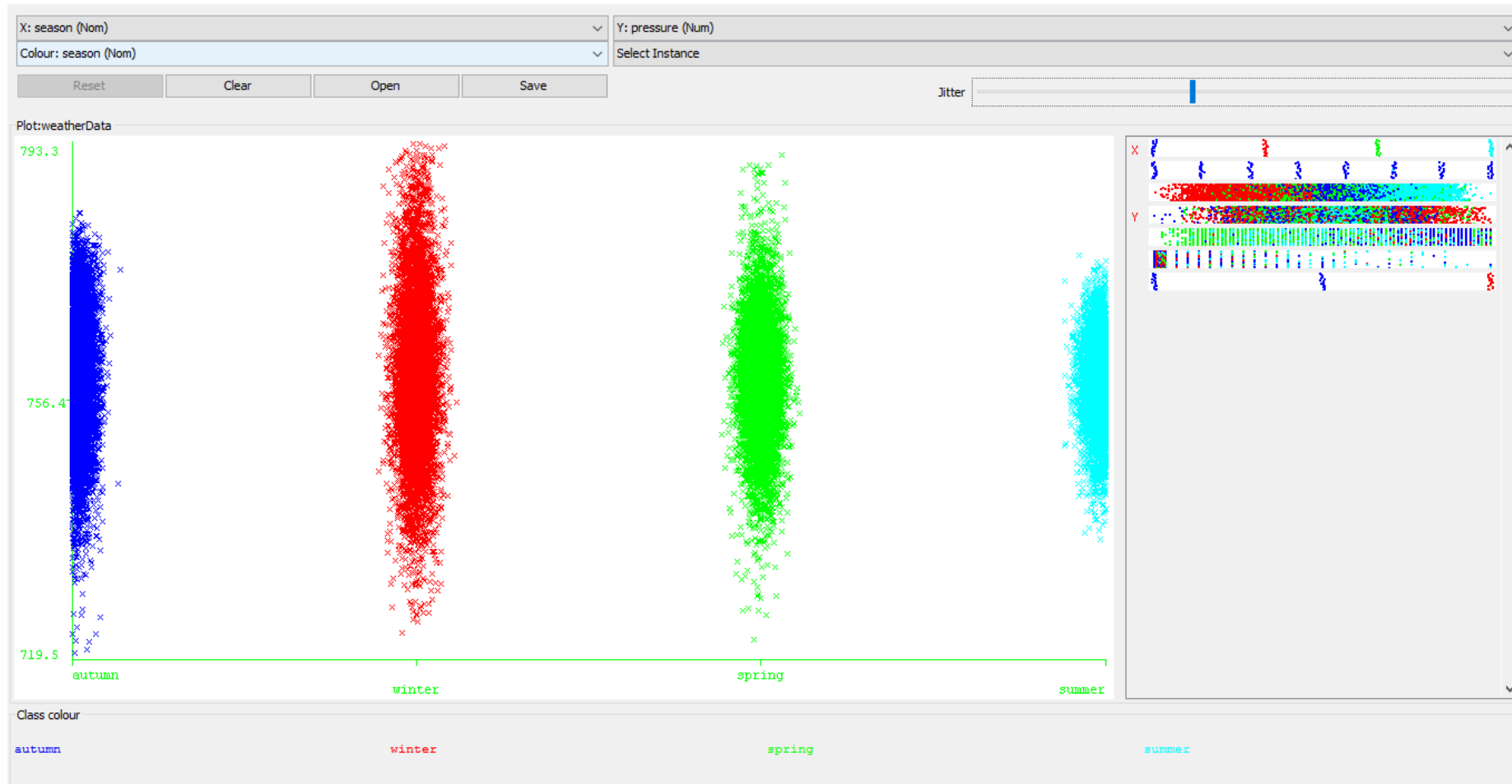
Изменение температуры для каждого сезона



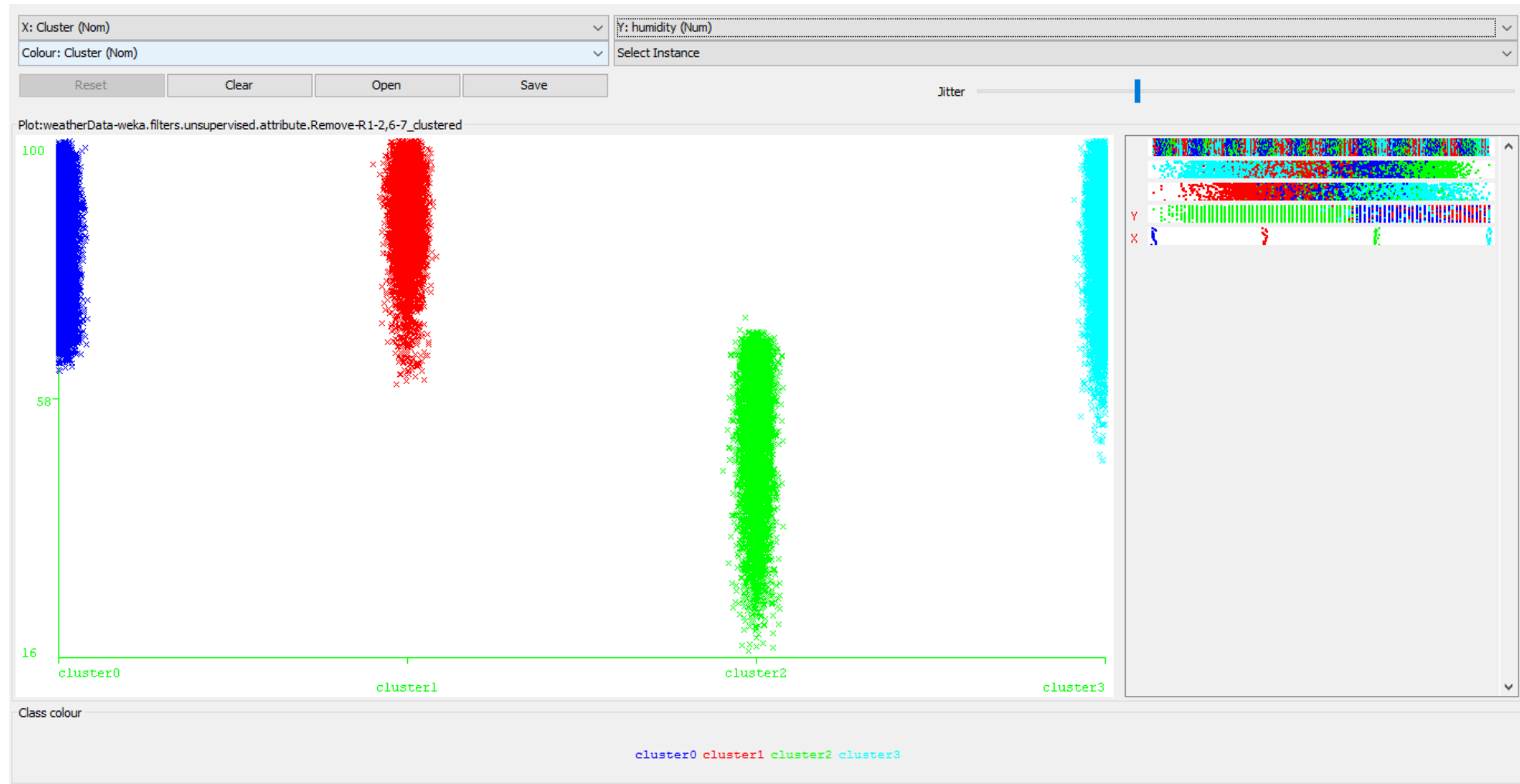
Изменение давления для каждого кластера



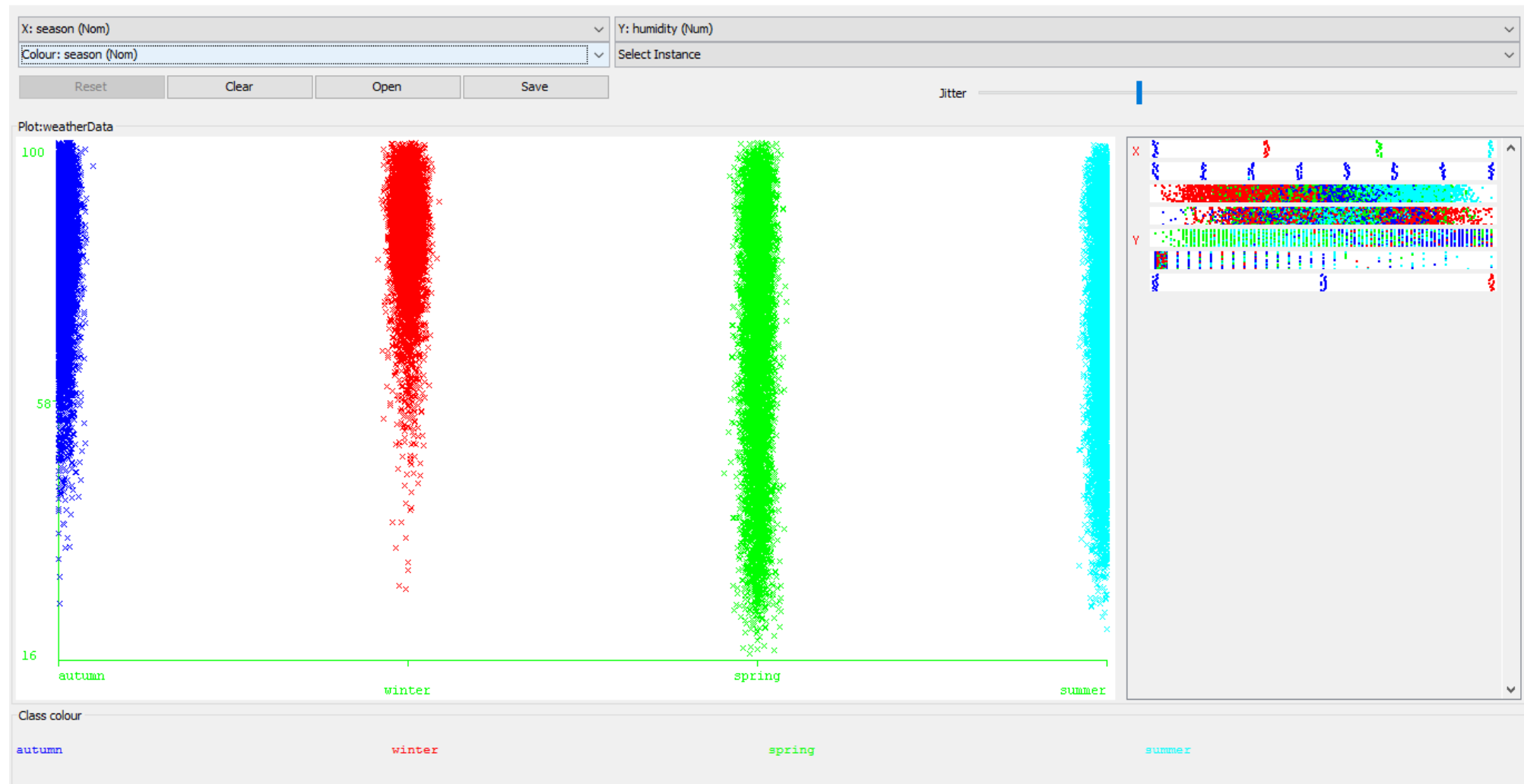
Изменение давления для каждого сезона



Изменение влажности для каждого кластера



Изменение влажности для каждого сезона



Поиск ассоциативных правил

- Associator - Apriori

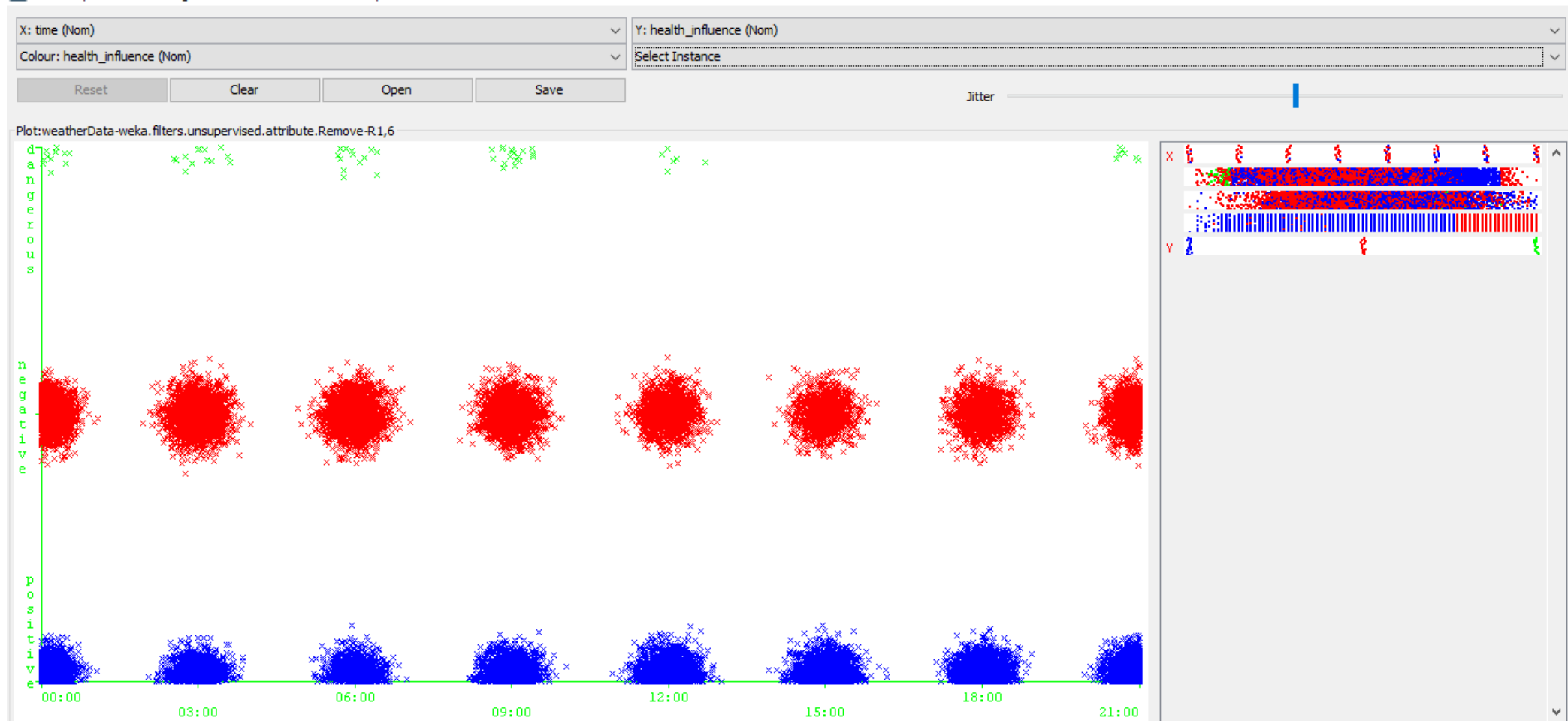
Best rules found:

```
1. temperature=hot humidity=normal 576 ==> dropPressure=no 564    conf:(0.98)
2. temperature=hot pressure=normal humidity=normal 431 ==> dropPressure=no 421    conf:(0.98)
3. temperature=hot 714 ==> dropPressure=no 697    conf:(0.98)
4. temperature=hot pressure=normal 539 ==> dropPressure=no 526    conf:(0.98)
5. pressure=normal humidity=normal 770 ==> dropPressure=no 735    conf:(0.95)
6. pressure=high 598 ==> dropPressure=no 568    conf:(0.95)
7. humidity=normal 1153 ==> dropPressure=no 1089    conf:(0.94)
8. pressure=normal 1607 ==> dropPressure=no 1497    conf:(0.93)
9. pressure=normal humidity=high 833 ==> dropPressure=no 758    conf:(0.91)
10. temperature=mild humidity=normal 544 ==> dropPressure=no 495    conf:(0.91)
```

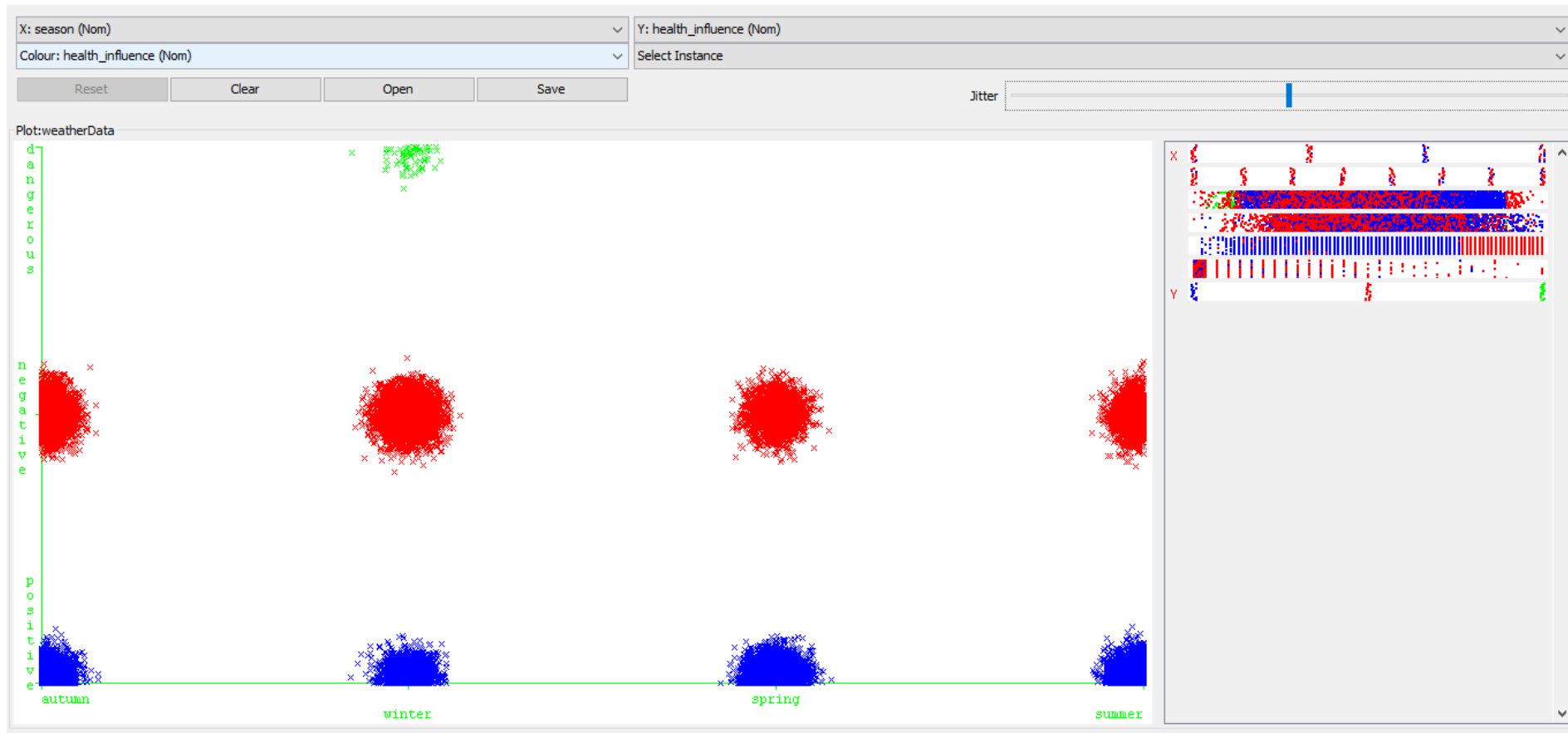
Критерии

	Инфекционные заболевания	Заболевания почек	Бронхиальная астма	Аллергия	Сердечно-сосудистые заболевания	Болезни дыхательных путей	Снижение иммунитета
H >85%	+	+					
H < 10%			+	+			
T < -20	+	+	+				+
T > 30				+	+	+	
P >750			+	+	+		+
P<750					+		

Зависимость влияния погоды на здоровье от времени суток



Зависимость влияния погоды на здоровье от сезона



Количество дней в соответствии с влиянием погоды на здоровье

