DMWA Lab -3

Teghdeep Kapoor

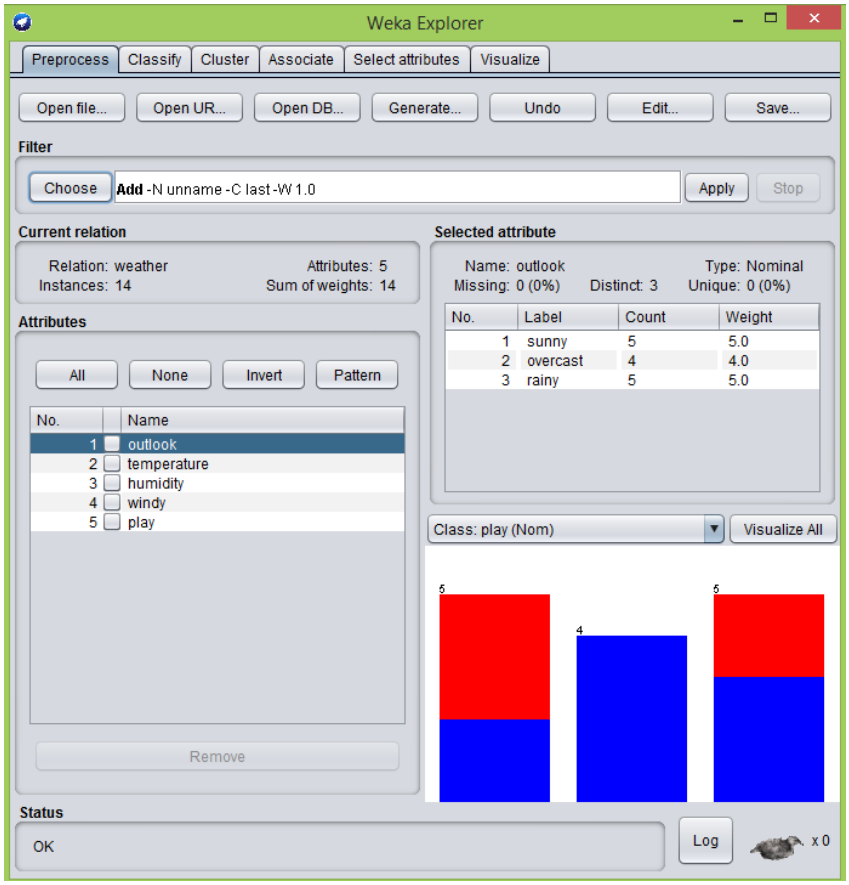
18104050

B12

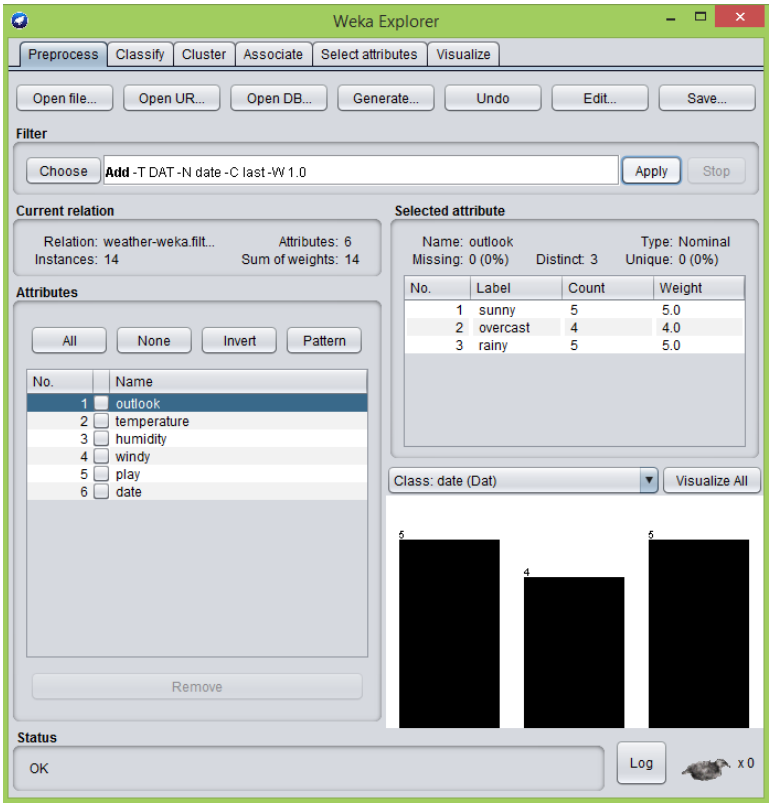
Q1

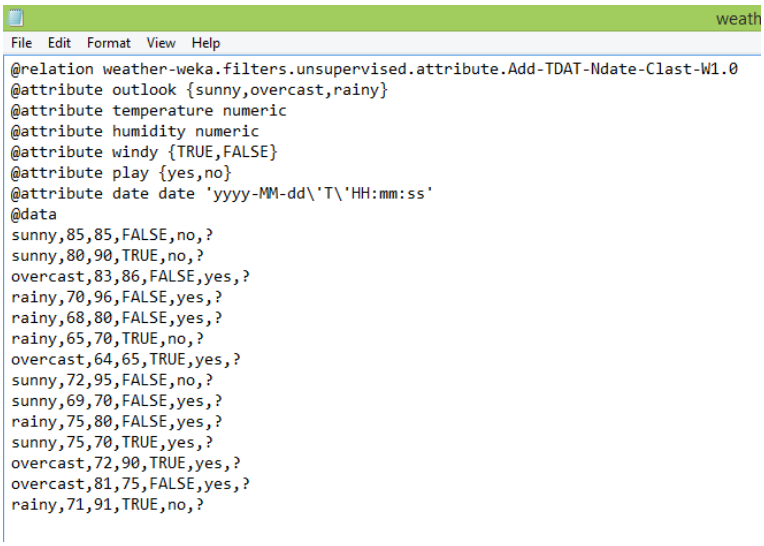
a. Add Attribute(Date)

Before Adding the attribute

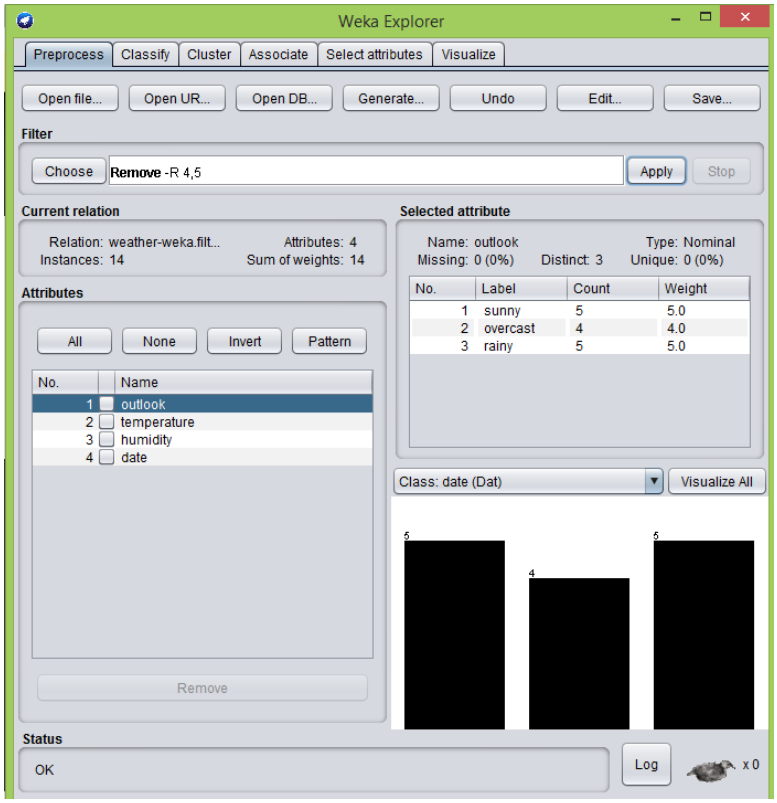


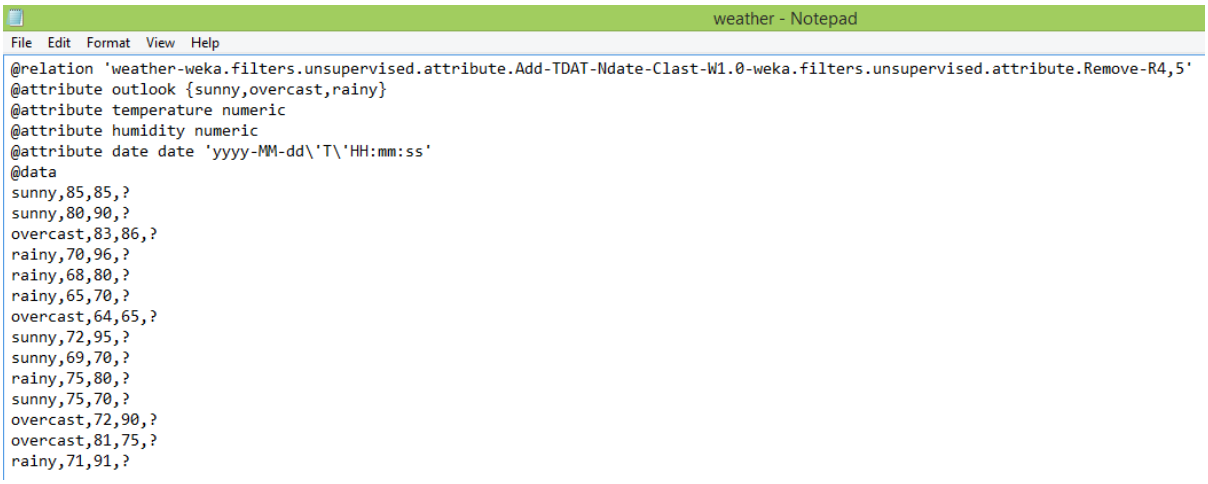
After Adding the Attribute



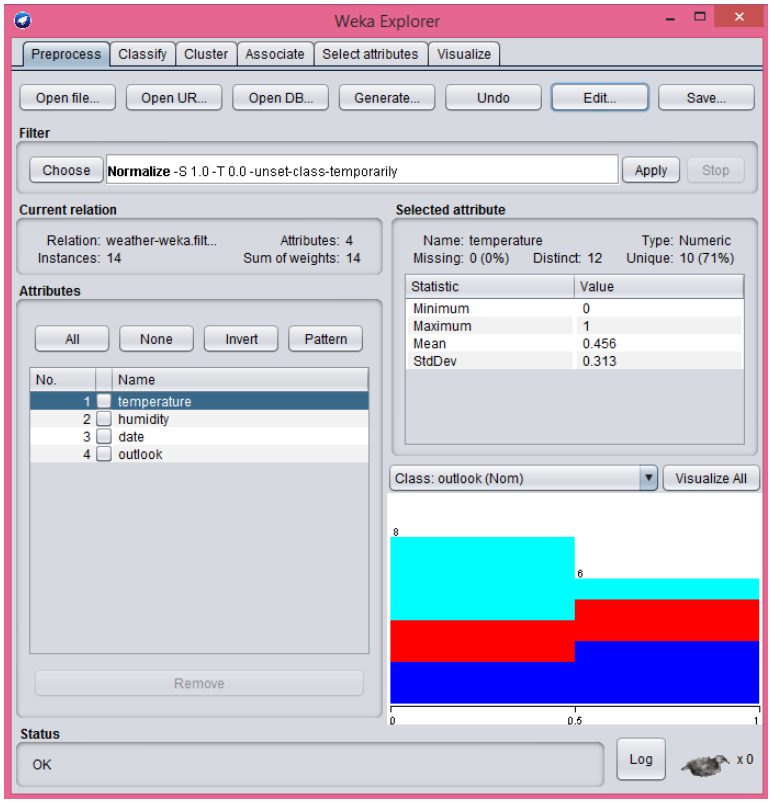


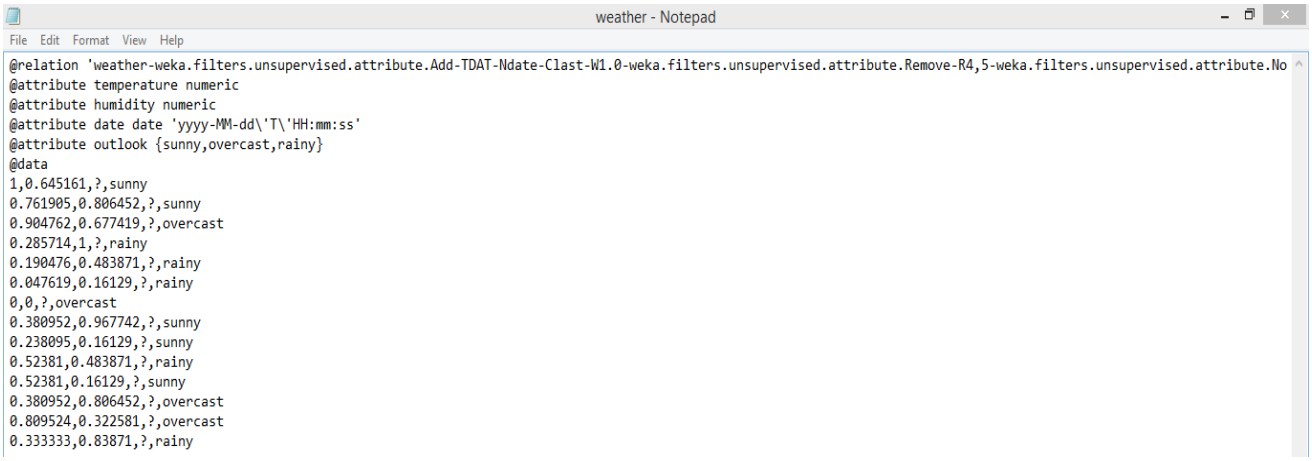
b. Removing the Attribute (Windy and Play)





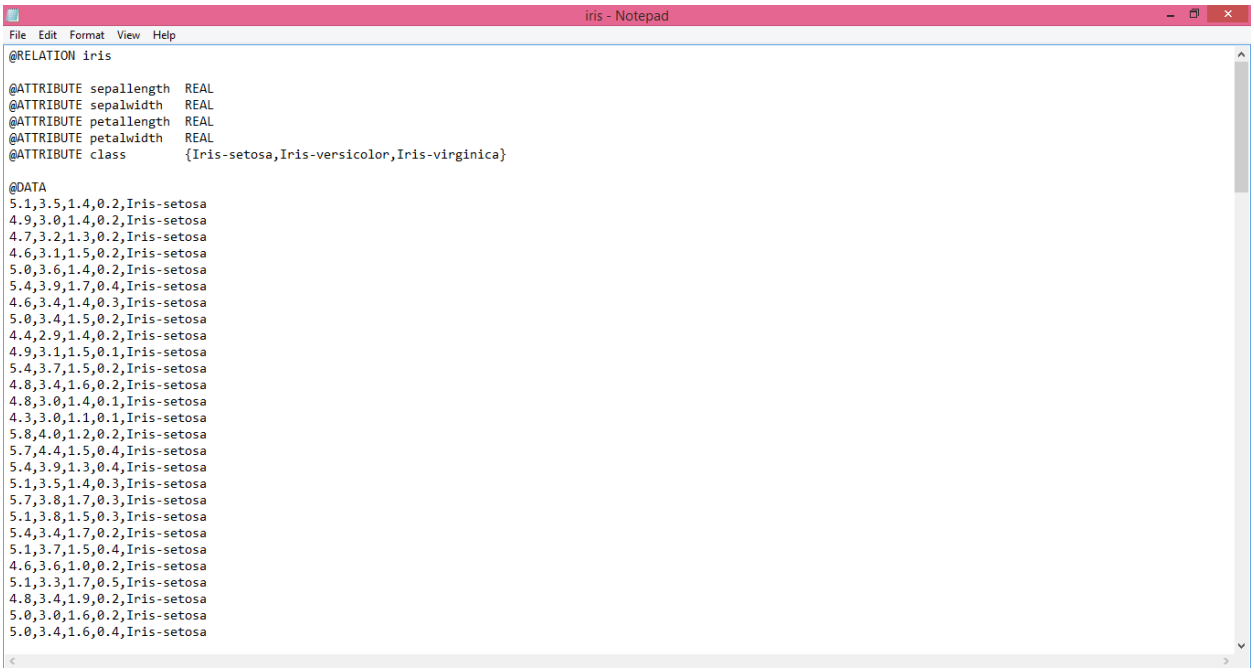
c. Normalizing the Data (Temperature, and Humidity)



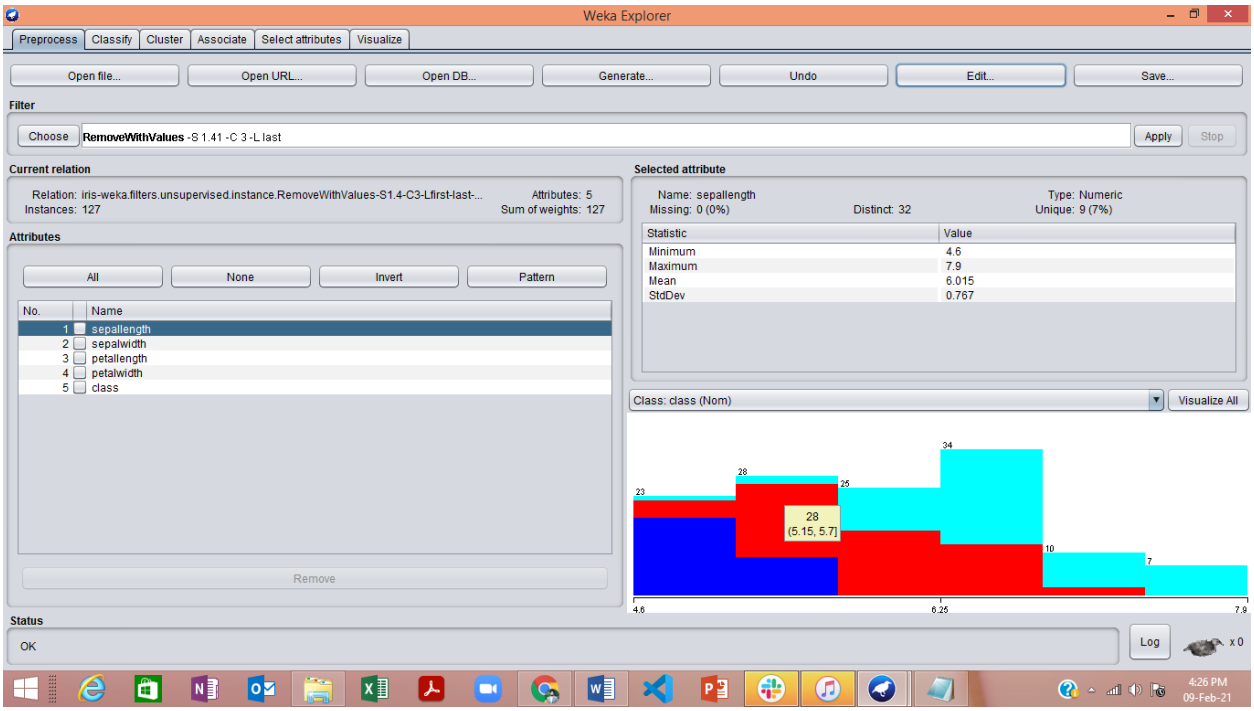


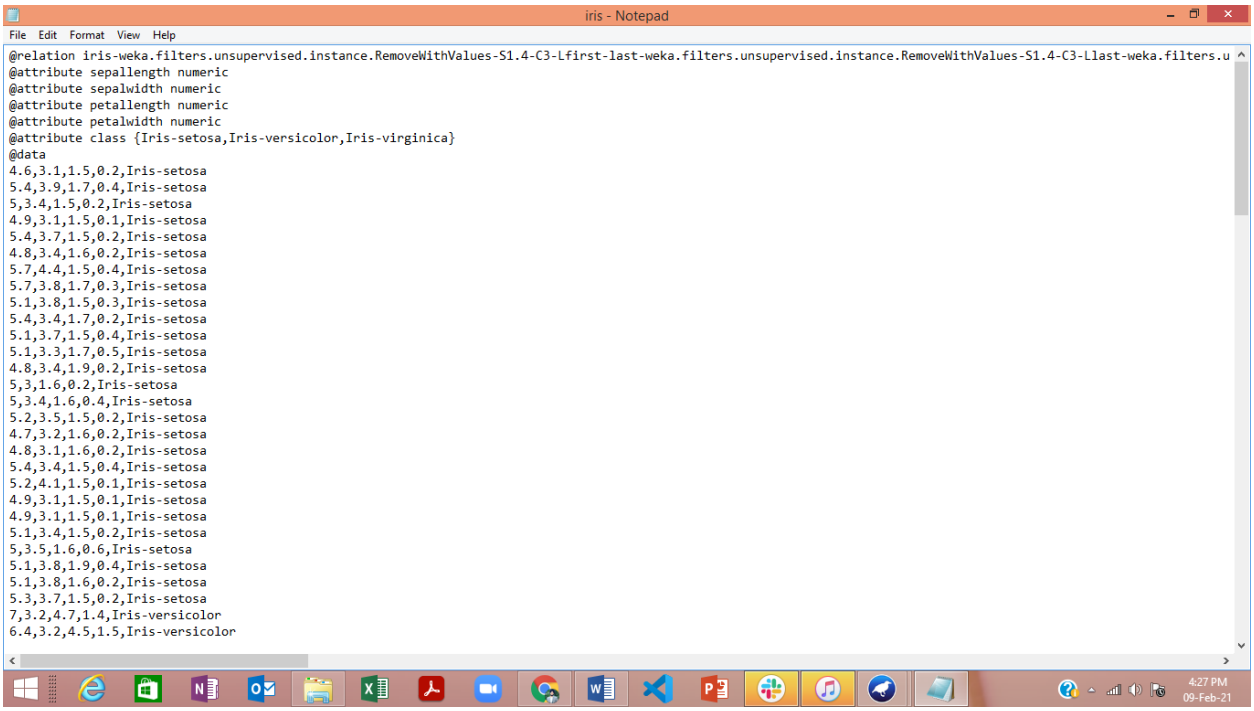
Q2

a. Downloading the IRIS dataset from online website

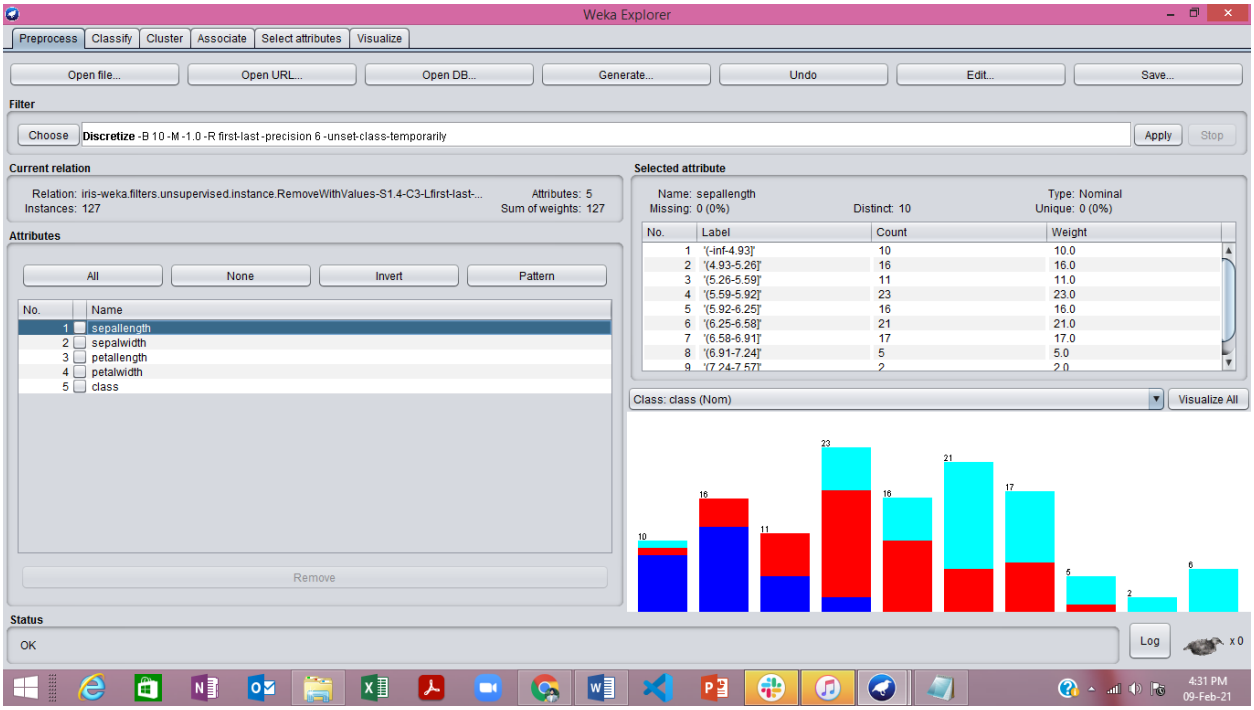


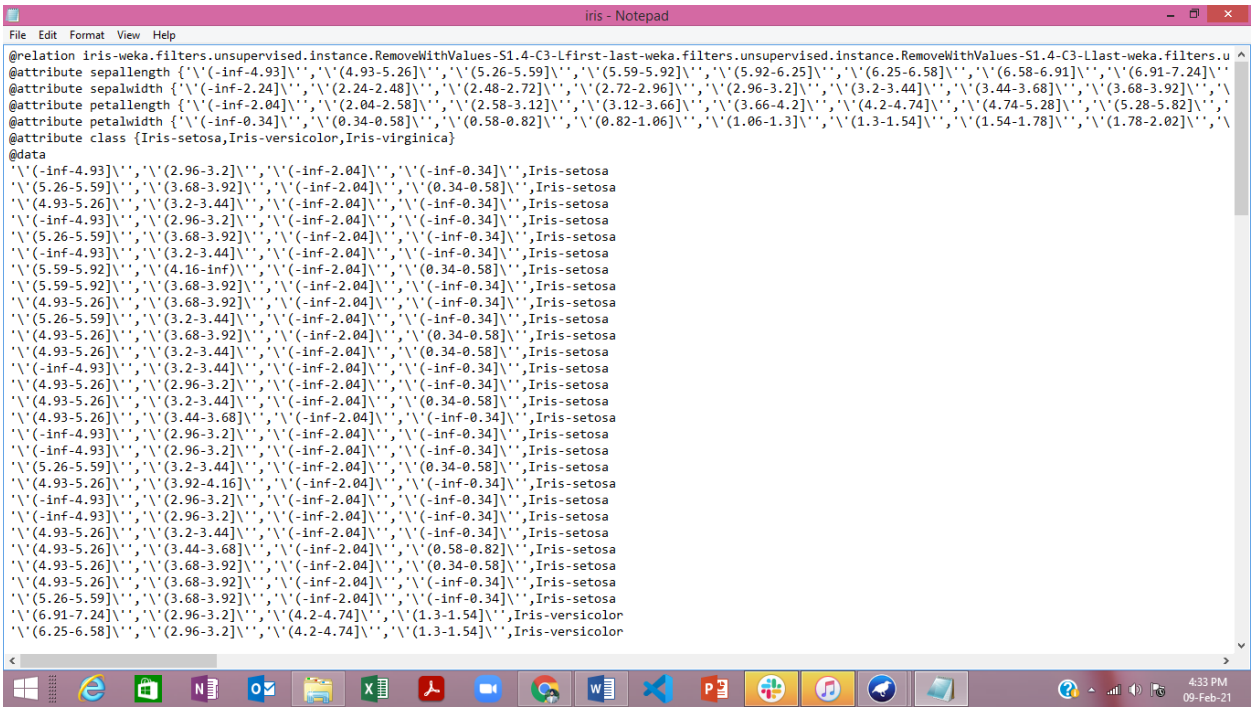
b. Removing the Values in the dataset where petalWidth = 1.4





c. Applying Various Discretization Filters on the Dataset





Q3

a. The automatic generation and smoothing for numeric data based on equal-frequency Partitioning rule.

```
a = len(arr1)
n = int(a / m)
for i in range(0, m):
arr = []
for j in range(i * n, (i + 1) * n):
if j >= a:
break
arr = arr + [arr1[j]]
print(arr)
```

b. The automatic generation of co-relation coefficient between two random variables.

```
def correlationCoefficient(X, Y, n) :
sum_X = 0
sum_Y = 0
sum_XY = 0
squareSum_X = 0
squareSum_Y = 0

i = 0
while i < n :
sum_X = sum_X + X[i]
sum_Y = sum_Y + Y[i]
sum_XY = sum_XY + X[i] * Y[i]
squareSum_X = squareSum_X + X[i] * X[i]
squareSum_Y = squareSum_Y + Y[i] * Y[i]

i = i + 1

corr = (float)(n * sum_XY - sum_X * sum_Y)/
(float)(math.sqrt((n * squareSum_X -
sum_X * sum_X)* (n * squareSum_Y -
sum_Y * sum_Y)))
return corr
```