

```
import numpy

#1
f=open ("Data1.txt","r")


a=16-1
b=17-1
feature_a=[]
feature_b=[]
# feature31 is for later use
feature31=[]
summ=[]


for l in (f):
    l = l.split(';')


    if l[a]=="yes":
        l[a]=1
    else:
        l[a]=0


    feature_a.append(l[a])

    if l[b]=="yes":
        l[b]=1
    else:
        l[b]=0
    feature_b.append(l[b])


feature31.append(int((l[31].strip(''))))
```

#2

```
def EuDistance(i,j):  
    d_a=feature_a[i]-feature_a[j]  
    d_b=feature_b[i]-feature_b[j]  
    return(d_a**2+d_b**2)**(1/2)
```

#3

```
dInGrades=[]  
dInBetterG=[]  
dBetween=[]
```

```
for i in range (len(feature_a)):  
    for j in range (len(feature_a)):  
        if i==j:  
            continue  
        elif feature31[j]>10 and feature31[i]>10:  
            dInBetterG.append(EuDistance(i,j))  
        elif (feature31[j]>10 and feature31[i]<=10) or (feature31[j]<=10 and feature31[i]>10):  
            dBetween.append(EuDistance(i,j))  
        else:  
            dInGrades.append(EuDistance(i,j))
```

#4

```
def averageList(L):  
    return numpy.mean(L)  
  
print ("mean distance between students with a grade higher than 10 is",  
        averageList(dInBetterG))  
  
print("mean distance between other students is",averageList(dInGrades))  
print("mean distance between mixed students is",averageList(dBetween))
```

#5

set a=18-1 and b=21-1 in the code above

#here we assumed that the above code was executed, so feature31 is already constructed

```
f=open ("Data1.txt","r")
```

```
a=18-1
```

```
b=21-1
```

```
feature_a=[]
```

```
feature_b=[]
```

```
for l in (f):
```

```
    l = l.split(';')
```

```
    if l[a]=="yes":
```

```
        l[a]=1
```

```
    else:
```

```
        l[a]=0
```

```
feature_a.append(l[a])
```

```
    if l[b]=="yes":
```

```
        l[b]=1
```

```
    else:
```

```
        l[b]=0
```

```
feature_b.append(l[b])
```

```
def EuDistance(i,j):
```

```
    d_a=feature_a[i]-feature_a[j]
```

```
    d_b=feature_b[i]-feature_b[j]
```

```

return(d_a**2+d_b**2)**(1/2)

dInGrades=[]
dInBetterG=[]
dBetween=[]

for i in range (len(feature_a)):
    for j in range (len(feature_a)):
        if i==j:
            continue
        elif feature31[j]>10 and feature31[i]>10:
            dInBetterG.append(EuDistance(i,j))
        elif (feature31[j]>10 and feature31[i]<=10) or (feature31[j]<=10 and feature31[i]>10):
            dBetween.append(EuDistance(i,j))
        else:
            dInGrades.append(EuDistance(i,j))

def averageList(L):
    return numpy.mean(L)

print ("mean distance between students with a grade higher than 10 is",
        averageList(dInBetterG))

print("mean distance between other students is",averageList(dInGrades))
print("mean distance between mixed students is",averageList(dBetween))

#6

a=14-1
b=24-1

feature_a=[]
feature_b=[]

f=open ("Data1.txt","r")

for l in (f):

```

```
l = l.split(';')
```

```
if int(l[a])>1:
```

```
    l[a]=1
```

```
else:
```

```
    l[a]=0
```

```
feature_a.append(l[a])
```

```
if l[b]=="1" or l[b]=="2":
```

```
    l[b]=0
```

```
else:
```

```
    l[b]=1
```

```
feature_b.append(l[b])
```

```
dInGrades=[]
```

```
dInBetterG=[]
```

```
dBetween=[]
```

```
for i in range (len(feature_a)):
```

```
    for j in range (len(feature_a)):
```

```
        if i==j:
```

```
            continue
```

```
        elif feature31[j]>10 and feature31[i]>10:
```

```
            dInBetterG.append(EuDistance(i,j))
```

```
        elif (feature31[j]>10 and feature31[i]<=10) or (feature31[j]<=10 and feature31[i]>10):
```

```
            dBetween.append(EuDistance(i,j))
```

```
        else:
```

```
            dInGrades.append(EuDistance(i,j))
```

```
def averageList(L):  
    return numpy.mean(L)  
  
print ("mean distance between students who got more than 10 is",  
       averageList(dInBetterG))  
  
print("mean distance between other students is",averageList(dInGrades))  
print("mean distance between mixed students is",averageList(dBetween))
```