```
1 import random
```

```
1 # universe
2 U = list(range(1,1000))
3
4 # two similar sets
5 C1 = list(range(1,700))
6 C2 = list(range(100,1000))
7 # Jaccard sim = 60%
8
9 # generate a random permutation of U
10 random.shuffle(U) # new U is a permutation of the previous U
11
12 def minhash(C, U):
13   for x in U:
14     if x in C:
15       return x
16
17
18 m = 100
19 agreements = 0;
20 for i in range(m):
21   random.shuffle(U)
22   minhashC1 = minhash(C1,U)
23   minhashC2 = minhash(C2,U)
24   if minhashC1 == minhashC2:
25     agreements +=1
26   print(minhashC1, minhashC2)
27
28 print('Jaccard similarity estimation', agreements/m)
```