Code 5

```
import numpy as np
import pandas as pd
import os
import sys
from collections import defaultdict
#C:\Users\Ajinkya\Desktop\sample.csv
data_df = pd.read_csv('C:/Users/nilan/Desktop/sample.csv')
#print(data_df)
d = defaultdict(list)
d1 = []
from collections import Counter
for index,rows in data_df.iterrows():
if (rows['hired'] == 1):
d1.append(rows['tasker_id'])
k = Counter(d1)
print(k)
for x in k:
key = x
value = k[key]
if(value == 1):
print key,
print(k.most_common(1))
```

Code 6

```
import numpy as np
import pandas as pd
import os
import sys
from collections import defaultdict
data_df = pd.read_csv('C:/Users/nilan/Desktop/sample.csv')
d = defaultdict(list)
d4 = defaultdict(list)
d1 = []
d2 = []
from collections import Counter
for index,rows in data_df.iterrows():
if (rows['hired'] == 1):
d1.append(rows['tasker_id'])
elif (rows['hired'] == 0):
d2.append(rows['tasker_id'])
k = Counter(d1)
l =[]
for x in k:
key = x
```

```
l.append(key)
value = k[key]
d[key].append(value)
if(value == 1):
print key,
print(k.most_common(1))
k1 = Counter(d2)
for x in k1:
key = x
value = k1[key]
d4[key].append(value)
for key,value in d.iteritems():
if key in d4:
s = d[key] + d4[key]
sum1 = sum(s)
p = d[key]
sum2 = sum(p)
k = (float(sum2)/float(sum1))
print(k)
```

Code 8

```
import numpy as np
import pandas as pd
import os
import sys
from collections import defaultdict
#C:\Users\nilan\Desktop\sample.csv
df = pd.read_csv('C:/Users/nilan/Desktop/sample.csv')
#from collections import Counter
Hired = df.ix[(df['hired']==1)]
df1 = pd.DataFrame(Hired, columns = ['category', 'hired', 'position'])
data_df=df1.groupby(['category']).mean()
print(data_df)
```

Code 9

```
import numpy as np
import pandas as pd
import os
import sys
from collections import defaultdict
#C:\Users\nilan\Desktop\sample.csv
df = pd.read_csv('C:/Users/nilan/Desktop/sample.csv')
#from collections import Counter
Hired = df.ix[(df['hired']==1)]
df1 = pd.DataFrame(Hired, columns = ['category', 'num_completed_tasks',
'hourly_rate'])
```

```
data_df=df1.groupby(['category']).mean()
print(data_df)
```