Name: Shahnwaz Khan Roll: AC-1272 Practical 1

In [ ]:
```
import pandas as pd
```

Read the text file

i) Read the data from the file “people.txt”.

In [ ]:
```
data=pd.read_csv('people.txt', delimiter=' ')
display(data)
```

	Age	agegroup	height	status	yearsmarried
0	21	adult	6.0	single	-1
1	2	child	3.0	married	0
2	18	adult	5.7	marri	20
3	221	elderly	5.0	widowed	2
4	34	child	-7.0	married	3

ii) Create a ruleset E that contain rules to check for the following conditions:

1. The age should be in the range 0-150.
2. The age should be greater than yearsmarried.
3. The status should be married or single or widowed.
4. If age is less than 18 the agegroup should be child, if age is between 18 and 65 the agegroup should be adult, if age is more than 65 the agegroup should be elderly.

In [ ]:
```
import ruleset
```

In [ ]:
```
# importing rules
rules=[]
rules.append(ruleset.check_age(data))
rules.append(ruleset.check_agemarried(data))
rules.append(ruleset.check_status(data))
rules.append(ruleset.check_ageGroup(data))
```

In [ ]:
```
print(rules)
```

```
[(1, 'Ruleset 1.'), (0, 'Ruleset 2.'), (1, 'Ruleset 3.'), (1, 'Ruleset 4.')]
```

(iii) Check whether ruleset E is violated by the data in the file people.txt.

In [ ]:
```
flag=False
for violations in rules:
    if violations[0]!=0: # 0 indicates that no row is invalid
        flag=True
if flag:
    print("Ruleset has been violated.")
else:
    print("No ruleset has been violated.")
```

```
Ruleset has been violated.
```

(iV) Summarize the results obtained in part (iii)

In [ ]:
```
print("Violations count in each ruleset.")
for violations in rules:
    print(violations[1]," : ",violations[0])
```

```
Violations count in each ruleset.
Ruleset 1.  :  1
Ruleset 2.  :  0
Ruleset 3.  :  1
Ruleset 4.  :  1
```

(V) Visualize the results obtained in part (iii)

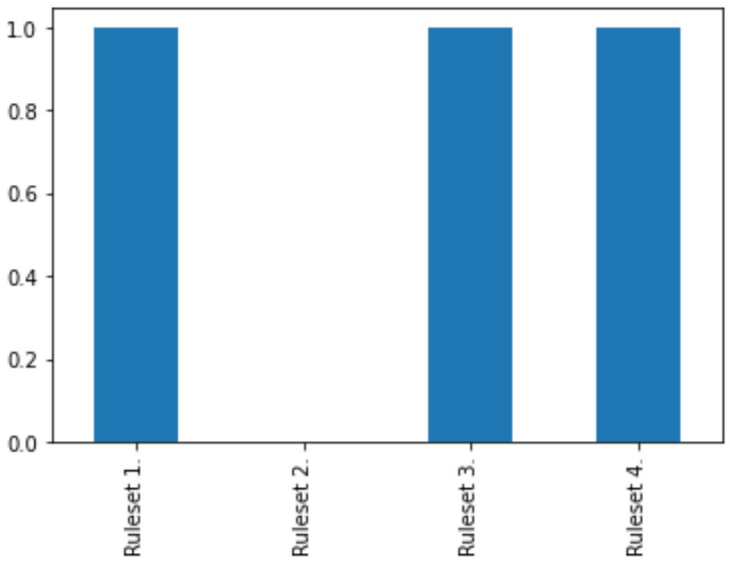
In [ ]:
```
df=pd.Series()
for i in rules:
    df[i[1]]=i[0]
```

In [ ]:
```
display(df)
```

```
Ruleset 1.    1
Ruleset 2.    0
Ruleset 3.    1
Ruleset 4.    1
dtype: int64
```

In [ ]:
```
df.plot.bar()
```

Out[ ]:
```
<AxesSubplot:>
```



```
#Ruletset

import pandas as pd
import numpy as np
def check_age(df):
    count=0
    for i in df.Age.values:
        if(i not in range(0,150)):
            count+=1
    return count,"Ruleset 1."

# age should be greater than yearsmarried
def check_agemarried(df):
   count=0
   for i in range(len(df)):
    if(df['Age'][i]<df['yearsmarried'][i]):
        count+=1
    return count,"Ruleset 2."

# the status should be married or single or widowed
def check_status(df):
    count=0
    if(np.unique(df.status.values) !=
['married','single','widowed']):
        count+=1
    return count,"Ruleset 3."

# If age is less than 18 the agegroup should be child,
if age is between 18 and 65 the agegroup should be
# adult, if age is more than 65 the agegroup should be
elderly.

def check_ageGroup(df):
    count=0
```

```
    for i in range(len(df)):
        if df['Age'][i]<18 and
df['agegroup'][i]!='child':
         count+=1
        elif df['Age'][i]>18 and df['Age'][i] <=65 and
df['agegroup'][i]!='adult':
         count+=1
        elif df['Age'][i]>65 and
df['agegroup'][i]!='elderly':
         count+=1

    return count,"Ruleset 4."
```