```python
import pandas as pd
celebs = pd.read_csv("./data/famousbday.txt",sep=':',header=None,
names=['name','occup','bday'])
celeb_mbti = pd.read_csv("./data/myer-briggs.txt",header=None,sep=':',\
names=['mbti','name'])
df = pd.merge(celeb_mbti,celebs)
print df[:4]
print df.shape
```

```
    mbti             name                     occup        bday
0   ENFJ  Abraham Maslow  American psychologist    1/4/1908
1   ENFJ     Ben Stiller          American actor  30/11/1965
2   ENFJ       Bob Saget          American actor   17/5/1956
3   ENFJ  Brenda Vaccaro        American actress  18/11/1939
(537, 4)
```

```python
from datetime import datetime
def f(s):
    try:
        return datetime.strptime(s, '%d/%m/%Y').date().strftime('%Y%m%d')
    except:
        return None
print f('18/11/1939')
print f('18/11/1839')
df['bday2'] = df['bday'].apply(f)
print df[:4]
df.to_csv('/tmp/out.csv',sep=';')
```

```
19391118
None
    mbti             name                     occup        bday      bday2
0   ENFJ  Abraham Maslow  American psychologist    1/4/1908   19080401
1   ENFJ     Ben Stiller          American actor  30/11/1965   19651130
2   ENFJ       Bob Saget          American actor   17/5/1956   19560517
3   ENFJ  Brenda Vaccaro        American actress  18/11/1939   19391118
```

```python
import os
cols = []
lewi = os.listdir('./doc/details/lewi')
lewi = map(lambda x: 'lewi'+x.replace('.html',''),lewi)
cols += lewi
millman = os.listdir('./doc/details/millman')
millman = map(lambda x: 'mill'+x.replace('.html',''),millman)
millman.remove('millnineyearcycle')
cols += millman
chinese = os.listdir('./doc/details/chinese')
chinese = map(lambda x: x.replace('.html',''),chinese)
cols += chinese
spiller = os.listdir('./doc/details/spiller')
spiller = map(lambda x: x.replace('.html',''),spiller)
cols += spiller
for x in cols: df[x] = np.nan
print df.ix[7]
```

```
mbti                                                ENFJ
```

```
name                               Martin Luther King Jr
occup         American civil rights leader, recipient of the...
bday                                         15/1/1929
bday2                                         19290115
lewi143                                            NaN
lewi53                                             NaN
lewi154                                            NaN
lewi64                                             NaN
lewi160                                            NaN
lewi55                                             NaN
lewi91                                             NaN
lewi140                                            NaN
lewi242                                            NaN
lewi155                                            NaN
...
Ox              NaN
Rabbit          NaN
Horse           NaN
Libra           NaN
Scorpio         NaN
Aries           NaN
Cancer          NaN
Pisces          NaN
Gemini          NaN
Virgo           NaN
Aquarius        NaN
Sagittarius     NaN
Taurus          NaN
Leo             NaN
Capricorn       NaN
Name: 7, Length: 354, dtype: object
```

```python
df2 = df[pd.isnull(df['bday2']) == False]
def enrich(x):
    res = mindmeld.calculate(x['bday2'])
    for lew in res['lewi']: x['lewi'+str(lew)] = 1
    if res['spiller']: x[res['spiller']] = 1
    if res['chinese']: x[res['chinese']] = 1
    return x

df3 = df2.apply(enrich, axis=1)
df3.to_csv('/tmp/out3.csv',sep=';')
```