```
# coding: utf-8

# In[1]:


import numpy as np
import pandas as pd


# In[2]:


df = pd.read_csv('C:/Users/rusha/Documents/Business Analytics
program/ny.csv.csv',index_col=False,low_memory=False)


# In[3]:


parties = {'Clinton, Hillary Rodham':'Democrat', 'Sanders,
Bernard':'Democrat', 'Trump, Donald J.':'Republican',

       "O'Malley, Martin Joseph":'Democrat', "Cruz, Rafael Edward
'Ted'":'Republican',

       'Walker, Scott':'Republican', 'Bush, Jeb':'Republican', 'Rubio,
Marco':'Republican', 'Kasich, John R.':'Republican',

       'Christie, Christopher J.':'Republican', 'Stein, Jill':'Green',
'Johnson, Gary':'Libertarian',

       'Graham, Lindsey O.':'Republican', 'Webb, James Henry
Jr.':'Democrat',

       'Carson, Benjamin S.':'Republican', 'Paul, Rand':'Republican',
'Fiorina, Carly':'Republican',

       'Santorum, Richard J.':'Republican', 'Jindal, Bobby':'Republican',
'Huckabee, Mike':'Republican',

       'Pataki, George E.':'Republican', 'Gilmore, James S
III':'Republican', 'Lessig, Lawrence':'Democrat',

       'Perry, James R. (Rick)':'Republican', 'McMullin,
Evan':'Independent'}


# This is an example of markdown.

# In[4]:


df['party']=df.cand_nm.map(parties)
```

```
# Question 2 : convert the contb_receipt_dt column into an actual data
object.

# In[7]:



df = pd.DataFrame(df) #converted into a Data Frame


# In[9]:


df['contb_receipt_dt'] = pd.to_datetime(df['contb_receipt_dt']) #convert
it to Date Time


# Question 3 : Using group by, show the number of donations given to each
party

# In[11]:


df3 = df.groupby(['contb_receipt_dt','party'])['tran_id'].count() #
Groups the Parties, Date & calculates the total number of Donations
df3


# Question 4 : Using group by, show the number of donations given to each
party, by date

# In[12]:


df3 = df.groupby(['contb_receipt_dt','party'])['tran_id'].count() #
Groups the Parties, Date & calculates the total number of Donations
df3


# Question 5 : Using group by, show the total amount of donations given
to each party

# In[13]:


pd.options.display.float_format = '{:,.2f}'.format
df['contb_receipt_amt'] = df['contb_receipt_amt'].astype(object)
df4 = df.groupby('party')['contb_receipt_amt'].sum()
df4


# In[14]:
```

```
print (df.party)


# Question 6 : Using group by, show the total amount of donations given
to each party, by date

# In[15]:


df5 = df.groupby(['contb_receipt_dt','party'])['contb_receipt_amt'].sum()
df5.head()


# Question 7 : Which occupations donated the top 5 most money?

# In[16]:


df6 = df.groupby('contbr_occupation')['contb_receipt_amt'].sum()
df6 = pd.DataFrame(df6)
df7 = df6.sort_values(['contb_receipt_amt'],ascending=False)
df7
df7.head()


# Question 8 : Which occupations donated the least 5 amount of money?

# In[26]:


df11 = df7.contb_receipt_amt[df7['contb_receipt_amt'] > 0]
df11.tail()


# Question 9 : Which employer's employees gave the most money, give the
top 5.

# In[18]:


df8= df.groupby('contbr_employer')['contb_receipt_amt'].sum()
df8 = pd.DataFrame(df6)
df9 = df8.sort_values(['contb_receipt_amt'],ascending=False)
df9
df9.head()


# Question 10 : For each candidate, what were the top 5 occupations that
donated to their election

# In[37]:
```

```
def top5(group) :
    return
group.groupby('contbr_occupation')['contb_receipt_amt'].sum().sort_values
(ascending = False).head(5)


# In[39]:


df10 = df.groupby('cand_nm').apply(top5)
df10
```