

## Analysis of GOP Donor Movements After Rubio and Bush Drop-outs

```
In [1]: import pandas as pd
import datetime as dt
```

### Load and Find Candidate Committees

```
In [2]: cn_headers = open("../data/cn_header_file.csv").read().strip().split(',')
```

```
In [3]: campaigns_to_committees = pd.read_csv("../data/cn.txt", sep="|", names=cn_headers)
```

```
In [4]: CAND_NAMES = [
    "KASICH, JOHN R",
    "TRUMP, DONALD J",
    "RUBIO, MARCO",
    "BUSH, JEB",
    'CRUZ, RAFAEL EDWARD "TED"'
]
```

```
In [5]: COMMITTEE_IDS = campaigns_to_committees[
    (campaigns_to_committees["CAND_OFFICE"] == "P") &
    (campaigns_to_committees["CAND_PTY_AFFILIATION"] == "REP") &
    (campaigns_to_committees["CAND_ELECTION_YR"] == 2016) &
    (campaigns_to_committees["CAND_NAME"].isin(CAND_NAMES))
].set_index("CAND_NAME")["CAND_PCC"].to_dict()
COMMITTEE_IDS
```

```
Out[5]: {'BUSH, JEB': 'C00579458',
         'CRUZ, RAFAEL EDWARD "TED"': 'C00574624',
         'KASICH, JOHN R': 'C00581876',
         'RUBIO, MARCO': 'C00458844',
         'TRUMP, DONALD J': 'C00580100'}
```

## Load and Clean Individual Donations

BuzzFeed News downloaded the "Contributions by Individuals" master file from the [FEC's website](http://www.fec.gov/finance/disclosure/ftpdet.shtml) (<http://www.fec.gov/finance/disclosure/ftpdet.shtml>) on April 25, 2016. It contains information about every single itemized individual donation for every single 2016 campaign for contributors who have given more than \$200 to a committee during this election cycle. You can download a copy of that data [here](https://archive.org/details/fec-contributions-master-file-2016-04-25) (<https://archive.org/details/fec-contributions-master-file-2016-04-25>).

First we select only the donations to the five Republican campaign committees we're interested in analyzing.

Then, we simplify each donor's name (removing suffixes and middle names, which committees have different approaches to reporting) and ZIP code (to the first five digits).

FEC campaign filings do not assign any unique identifiers to donors. So finally, to be able to distinguish and track individual donors, the code below assigns each contribution a `donor_uid` based on the donor's first name, last name, and ZIP code. This approach could result in an undercount of donors if, for instance, there are two people named John Smith in a same ZIP code — but should result in good approximations for the analyses below.

## Load All Itemized Individual Donations

```
In [6]: ind_headers = open("../data/indiv_header_file.csv").read().strip().split(',')
```

```
In [7]: dtypes = {  
    "NAME": str,  
    "ZIP_CODE": str,  
    "TRANSACTION_DT": str,  
    "FILE_NUM": str,  
    "MEMO_CD": str,  
    "MEMO_TXT": str  
}
```

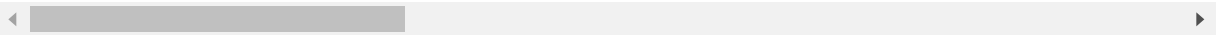
```
In [8]: donors = pd.read_csv(  
    "../data/itcont.txt",  
    sep="|",  
    names=ind_headers,  
    dtype=dtypes  
)
```

In [9]: donors.head()

Out[9]:

	CMTE_ID	AMNDT_IND	RPT_TP	TRANSACTION_PGI	IMAGE_NUM	TRANSACTION_TP	ENTI
0	C00004606	N	M4	P	15951124869	15	
1	C00004606	N	M4	P	15951124869	15	
2	C00004606	N	M4	P	15951124869	15	
3	C00452383	N	M4	P	15951124897	15	
4	C00452383	N	M4	P	15951124897	15	

5 rows × 21 columns



```
In [10]: gop_primary_donors = donors[
    donors["CMTE_ID"].isin(COMMITTEE_IDS.values()) &
    (donors["TRANSACTION_PGI"] == "P")
].copy()
```

## Clean Up Donation Data

```
In [11]: def parse_date(date_string):
    if pd.isnull(date_string):
        return None
    else:
        return dt.datetime.strptime(date_string.strip(), "%m%d%Y")
```

```
In [12]: gop_primary_donors["date"] = gop_primary_donors["TRANSACTION_DT"].apply(parse_date)
```

```
In [13]: def extract_last_first(name):
    return " ".join(name.split(" ")[1:2])
```

```
In [14]: gop_primary_donors["last_first"] = gop_primary_donors["NAME"].apply(extract_last_first)
gop_primary_donors["zip_first_five"] = gop_primary_donors["ZIP_CODE"].fillna(
    "").apply(lambda x: x[:5])
```

```
In [15]: def make_uid(row):
    if pd.isnull(row["last_first"]) or (row["zip_first_five"] == ""):
        return None
    else:
        return "|".join([row["last_first"], row["zip_first_five"]])
```

```
In [16]: gop_primary_donors["donor_uid"] = gop_primary_donors.apply(lambda x: make_uid(x), axis=1)
```

## Analyze the Data

The code below uses `donor_uid` to find the donors who made their first donation to a campaign committee after Jeb Bush and Marco Rubio dropped out of the 2016 Republican presidential primary—Feb. 20, 2016 and March 3, 2016 respectively. It then counts how many of those donations were made by donors who had previously given to the Bush or Rubio campaigns.

```
In [17]: # This method aggregates total donations, by contributor and committee,
# before and after a given date
def calculate_movements(since_date):
    grp = gop_primary_donors.groupby([
        "donor_uid",
        gop_primary_donors["date"] > since_date,
        "CMTE_ID",
    ])
    total_contribs = grp["TRANSACTION_AMT"].sum()
    return total_contribs.unstack().unstack().fillna(0)
```

```
In [18]: calculate_movements("2016-02-20").head()
```

```
Out[18]:
```

	CMTE_ID	C00458844	C00574624	C00579458	C00580100	C00581876					
	date	False	True	False	True	False	True	False	True	False	True
	donor_uid										
	AANONSEN, PAUL 20009	0.0	0.0	0.0	0.0	0.0	0.0	323.0	0.0	0.0	0.0
	AARNIO, TERRANCE 97267	0.0	0.0	0.0	0.0	2700.0	0.0	0.0	0.0	0.0	0.0
	AARON, CHARLES 91361	0.0	0.0	250.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	AARON, DAVID 30125	0.0	0.0	0.0	0.0	0.0	0.0	253.0	303.0	0.0	0.0
	AARON, FRED 34240	200.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

```
In [19]: def select_movements(from_candidate, to_candidate, since_date):
    movements = calculate_movements(since_date)
    return movements[
        (movements[(COMMITTEE_IDS[from_candidate], False)] > 0) &
        (movements[(COMMITTEE_IDS[to_candidate], False)] == 0) &
        (movements[(COMMITTEE_IDS[to_candidate], True)] > 0)
    ][(COMMITTEE_IDS[to_candidate], True)]
```

```
In [20]: REMAINING_CANDIDATES = [  
    'CRUZ, RAFAEL EDWARD "TED"',  
    "KASICH, JOHN R",  
    "TRUMP, DONALD J"  
]  
  
DROPOUTS = [  
    ("BUSH, JEB", "2016-02-20"),  
    ("RUBIO, MARCO", "2016-03-15")  
]
```

```
In [21]: for cand_r in REMAINING_CANDIDATES:  
    for cand_drop, date_drop in DROPOUTS:  
        m = select_movements(cand_drop, cand_r, date_drop)  
        print("{0} from {1}\n{2} donors\n${3:,.0f}\n".format(cand_r, cand_drop  
    , len(m), m.sum()))
```

```
CRUZ, RAFAEL EDWARD "TED" from BUSH, JEB  
63 donors  
$99,710
```

```
CRUZ, RAFAEL EDWARD "TED" from RUBIO, MARCO  
120 donors  
$126,650
```

```
KASICH, JOHN R from BUSH, JEB  
131 donors  
$173,850
```

```
KASICH, JOHN R from RUBIO, MARCO  
97 donors  
$92,168
```

```
TRUMP, DONALD J from BUSH, JEB  
9 donors  
$13,817
```

```
TRUMP, DONALD J from RUBIO, MARCO  
6 donors  
$4,504
```