

data source: https://www.rankedchoicevoting.org/data_clearinghouse
(https://www.rankedchoicevoting.org/data_clearinghouse)

RCV definition: [https://ballotpedia.org/Ranked-choice_voting_\(RCV\)](https://ballotpedia.org/Ranked-choice_voting_(RCV)) ([https://ballotpedia.org/Ranked-choice_voting_\(RCV\)](https://ballotpedia.org/Ranked-choice_voting_(RCV)))

Additional analysis:

- <http://archive3.fairvote.org/press/san-leandro-facts/> (<http://archive3.fairvote.org/press/san-leandro-facts/>)
- <https://laurendo.wordpress.com/2010/11/24/running-the-numbers/>
(<https://laurendo.wordpress.com/2010/11/24/running-the-numbers/>)
- <http://www.acgov.org/rov/rcv/results/index.htm> (<http://www.acgov.org/rov/rcv/results/index.htm>)

Objective for this notebook: separate the elections into the following categories:

1. Leading candidate in the first round has greater than 50% first choice votes
2. Leading candidate in the first round has between 45-50% first choice votes
3. Leading candidate in the first round has less than 45% of first choice votes

```
In [1]: import glob
import pandas
print('pandas', pandas.__version__)

pandas 0.23.4
```

data gathering: download all folders from drive manually

all the data: https://drive.google.com/drive/folders/1DJzlrTaDW3GSGJTkPTGAIpAMbozFG_pm
(https://drive.google.com/drive/folders/1DJzlrTaDW3GSGJTkPTGAIpAMbozFG_pm)

Then download all content as a zip. Size is 1.5 GB. Of this, Sante Fe is 1.4GB

I started with just "Alameda County, CA (Berkeley, Oakland, San Leandro)" which is 18MB as a .zip

https://drive.google.com/drive/folders/1u_airJzoLC2PMYMHcF2KYJEKxxKBi5H7
(https://drive.google.com/drive/folders/1u_airJzoLC2PMYMHcF2KYJEKxxKBi5H7)

```
!mkdir voting_data !mkdir voting_data/Alameda !unzip voting_data/Alameda/drive-download-20190724T221439Z-001.zip
```

parse ballotimage files

```
In [2]: list_of_files = glob.glob('voting_data/Alameda/Alameda (Oakland, San Leandro, Berkeley) 2010/ballot_image_*')
print('number of election results to parse:',len(list_of_files))
```

number of election results to parse: 17

```
In [3]: list_of_files[0]
```

```
Out[3]: 'voting_data/Alameda/Alameda (Oakland, San Leandro, Berkeley) 2010/ballot_image_Member, City Council, District 4 - Oakland_Nov 2010.txt'
```

```
In [4]: def sort_elections_into_bins(reslts,ballot):
    #print(ballot)
    df = pandas.read_fwf(ballot,
                        header=None,
                        widths=[7,9,7,3,7,3,7,1,1])
    df.columns=['contest_id','pref_voter_id',
                'serial_number','tally_type_id',
                'precinct_id','vote_rank',
                'candidate_id','over_vote','under_vote']
    #print(df.shape)
    #print(df['candidate_id'].unique())
    df_cand = df[df['candidate_id']!=0] # drop rows where no candidate is specified
    series_of_candidates_and_first_choice_count = df_cand[df_cand['vote_rank']==1].groupby('candidate_id')['vote_rank'].count()
    number_of_first_choice_votes = series_of_candidates_and_first_choice_count.sum()
    #print('number of first choice votes:',number_of_first_choice_votes)
    if (series_of_candidates_and_first_choice_count > number_of_first_choice_votes*0.5).any():
        #print('Leading candidate in the first round has greater than 50% first choice votes')
        reslts['leading candidate in first round has more than 50% of first choice votes'].append(ballot)
    elif ((series_of_candidates_and_first_choice_count <= number_of_first_choice_votes*0.5).any() and
          (series_of_candidates_and_first_choice_count >= number_of_first_choice_votes*0.45).any()):
        #print('Leading candidate in the first round has between 45-50% first choice votes')
        reslts['leading candidate in first round vote has between 50% and 45% of first choice votes'].append(ballot)
    elif (series_of_candidates_and_first_choice_count < number_of_first_choice_votes*0.45).any():
        #print('Leading candidate in the first round has less than 45% of first choice votes')
        reslts['leading candidate in first round vote has less than 45% of first choice votes'].append(ballot)
    else:
        raise Exception('should not reach this condition')
    return reslts
```

```
In [5]: results={'leading candidate in first round has more than 50% of first choice votes':[],
                'leading candidate in first round vote has between 50% and 45% of first choice votes':[],
                'leading candidate in first round vote has less than 45% of first choice votes':[]}

for ballot in list_of_files:
    results = sort_elections_into_bins(results,ballot)
```

```
In [6]: for k,v in reslts.items():
        print(k,':')
        for election in v:
            print('    ',election.replace('voting_data/','').replace('.txt',
            ''))
```

leading candidate in first round has more than 50% of first choice vote
s :

```
Alameda/Alameda (Oakland, San Leandro, Berkeley) 2010/ballot_image_
Member, City Council, District 2 - Oakland_Nov 2010
Alameda/Alameda (Oakland, San Leandro, Berkeley) 2010/ballot_image_
Member, City Council, District 1 - San Leandro_Nov 2010
Alameda/Alameda (Oakland, San Leandro, Berkeley) 2010/ballot_image_
Member, City Council, District 8 - Berkeley_Nov 2010
Alameda/Alameda (Oakland, San Leandro, Berkeley) 2010/ballot_image_
Member, City Council, District 6 - Oakland_Nov 2010
Alameda/Alameda (Oakland, San Leandro, Berkeley) 2010/ballot_image_
Member, City Council, District 4 - Berkeley_Nov 2010
Alameda/Alameda (Oakland, San Leandro, Berkeley) 2010/ballot_image_
City Auditor - Oakland_Nov 2010
Alameda/Alameda (Oakland, San Leandro, Berkeley) 2010/ballot_image_
Member, City Council, District 5 - San Leandro_Nov 2010
Alameda/Alameda (Oakland, San Leandro, Berkeley) 2010/ballot_image_
City Auditor - Berkeley_Nov 2010
Alameda/Alameda (Oakland, San Leandro, Berkeley) 2010/ballot_image_
School Director, District 6 - Oakland_Nov 2010
Alameda/Alameda (Oakland, San Leandro, Berkeley) 2010/ballot_image_
Member, City Council, District 1 - Berkeley_Nov 2010
Alameda/Alameda (Oakland, San Leandro, Berkeley) 2010/ballot_image_
School Director, District 4 - Oakland_Nov 2010
Alameda/Alameda (Oakland, San Leandro, Berkeley) 2010/ballot_image_
School Director, District 2 - Oakland_Nov 2010
Alameda/Alameda (Oakland, San Leandro, Berkeley) 2010/ballot_image_
Member, City Council, District 3 - San Leandro_Nov 2010
leading candidate in first round vote has between 50% and 45% of first
choice votes :
Alameda/Alameda (Oakland, San Leandro, Berkeley) 2010/ballot_image_
Member, City Council, District 7 - Berkeley_Nov 2010
leading candidate in first round vote has less than 45% of first choice
votes :
Alameda/Alameda (Oakland, San Leandro, Berkeley) 2010/ballot_image_
Member, City Council, District 4 - Oakland_Nov 2010
Alameda/Alameda (Oakland, San Leandro, Berkeley) 2010/ballot_image_
Mayor of San Leandro_Nov 2010
Alameda/Alameda (Oakland, San Leandro, Berkeley) 2010/ballot_image_
Mayor of Oakland_Nov 2010
```

Pierce County data

https://www.rankedchoicevoting.org/data_clearinghouse
(https://www.rankedchoicevoting.org/data_clearinghouse)

https://drive.google.com/drive/folders/1DJzlrTaDW3GSGJTkPTGAlpAMbozFG_pm
(https://drive.google.com/drive/folders/1DJzlrTaDW3GSGJTkPTGAlpAMbozFG_pm)

```
In [7]: list_of_files = glob.glob('voting_data/Pierce_County/Pierce County/*')
        len(list_of_files)
```

```
Out[7]: 8
```

```
In [8]: list_of_files
```

```
Out[8]: ['voting_data/Pierce_County/Pierce County/Pierce County Auditor 2009 Ba
llot Image.txt',
        'voting_data/Pierce_County/Pierce County/Pierce County Executive 2008
Master Lookup.txt',
        'voting_data/Pierce_County/Pierce County/Pierce County Assessor - Trea
surer 2008 Ballot Image.txt',
        'voting_data/Pierce_County/Pierce County/Pierce County Council, Distri
ct No. 2 2008 Master Lookup.txt',
        'voting_data/Pierce_County/Pierce County/Pierce County Council, Distri
ct No. 2 2008 Ballot Image.txt',
        'voting_data/Pierce_County/Pierce County/Pierce County Assessor - Trea
surer 2008 Master Lookup.txt',
        'voting_data/Pierce_County/Pierce County/Pierce County Executive 2008
Ballot Image Data.txt',
        'voting_data/Pierce_County/Pierce County/Pierce County Auditor 2009 Ma
ster Lookup.txt']
```

```
In [9]: list_of_ballot_files=[]
for filename in list_of_files:
    if filename.endswith('.txt'):
        with open(filename,'r') as fil:
            file_contents = fil.readlines()
            if len(file_contents[0].strip())==45:
                print(filename)
                list_of_ballot_files.append(filename)
                print(file_contents[1])
```

```
voting_data/Pierce_County/Pierce County/Pierce County Auditor 2009 Ball
ot Image.txt
000071400001543600000010050000002002000044000
```

```
voting_data/Pierce_County/Pierce County/Pierce County Assessor - Treasu
rer 2008 Ballot Image.txt
000019200006315800000010050000002002000000001
```

```
voting_data/Pierce_County/Pierce County/Pierce County Council, District
No. 2 2008 Ballot Image.txt
000019300007697700000010050000063002000013100
```

```
voting_data/Pierce_County/Pierce County/Pierce County Executive 2008 Ba
llot Image Data.txt
000019700006315800000010050000002002000000001
```

```
In [10]: results={'leading candidate in first round has more than 50% of first cho
ice votes':[],
                'leading candidate in first round vote has between 50% and 45% o
f first choice votes':[],
                'leading candidate in first round vote has less than 45% of firs
t choice votes':[]}

for ballot in list_of_ballot_files:
    reslts = sort_elections_into_bins(results,ballot)
```

```
In [11]: for k,v in reslts.items():
          print(k,':')
          for election in v:
              print('    ',election.replace('voting_data/','').replace('.txt',
''))
```

leading candidate in first round has more than 50% of first choice votes :

leading candidate in first round vote has between 50% and 45% of first choice votes :

Pierce_County/Pierce County/Pierce County Auditor 2009 Ballot Image
Pierce_County/Pierce County/Pierce County Council, District No. 2 2008 Ballot Image

leading candidate in first round vote has less than 45% of first choice votes :

Pierce_County/Pierce County/Pierce County Assessor - Treasurer 2008 Ballot Image

Pierce_County/Pierce County/Pierce County Executive 2008 Ballot Image Data

San Francisco

```
In [13]: list_of_files = glob.glob('voting_data/San_Fransisco/San Francisco/**/*')
          len(list_of_files)
```

Out[13]: 46

```
In [14]: list_of_file_extentions=[]
          for filename in list_of_files:
              #print(filename.split('.')[1])
              list_of_file_extentions.append(filename.split('.')[1])
          print(set(list_of_file_extentions))

{'txt', 'pdf'}
```

```
In [15]: for filename in list_of_files:
        if filename.endswith('.txt'):
        #     print(filename)
        if 'ballot' in filename.lower():
            print(filename.split('/')[-1])
            with open(filename,'r') as fil:
                file_contents = fil.readlines()
            print(file_contents[0:2])
            print(len(file_contents[0].strip()))
```


BallotImage-D10.txt
['000000600001706700000040020000274001000014900\n', '000000600001706700
000040020000274002000015700\n']
45
BallotImage-D2.txt
['000000700001712400000090020000331001000012600\n', '000000700001712400
000090020000331002000000001\n']
45
BallotImageListing.txt
['\x1bE\x1b&l2a0o7c067F\x1b(s0p16.66h3b6T\x1b&a00L\n', 'BALLOT IMAGE LI
STING
SAN FRANCISCO
OFFICIAL RESULTS\n']
36
BallotImageSummary.txt
['\x1bE\x1b&l2a0o7c067F\x1b(s0p16.66h3b6T\x1b&a00L\n', 'BALLOT IMAGE EL
ECTION SUMMARY\n']
36
20151119_ballotimage.txt
['000000100002610500000010020000012001000003600\n', '000000100002610500
000010020000012002000003700\n']
45
D10_BallotImage.txt
['000003300000385300000010020000054001000012900\n', '000003300000385300
000010020000054002000012800\n']
45
Sheriff-BallotImage.txt
['000000200004728200000010020000003001000002800\n', '000000200004728200
000010020000003002000002800\n']
45
DA-BallotImage.txt
['000000100004728200000010020000003001000002300\n', '000000100004728200
000010020000003002000002300\n']
45
Nov2004_BallotImage.txt
['pct,1,2,3\n', '0001,,,\n']
9
BallotImage-D8.txt
['000001000001706800000040020000513001000014400\n', '000001000001706800
000040020000513002000014200\n']
45
Mayor-BallotImage.txt
['000000300004728200000010020000003001000003700\n', '000000300004728200
000010020000003002000003700\n']
45
D7-BallotImage.txt
['000002300000759400000010020000244001000011500\n', '000002300000759400
000010020000244002000011800\n']
45
BallotImage-D6.txt
['0000009000017076000000130020000457001000013600\n', '000000900001707600
000130020000457002000012800\n']
45
20161206_ballotimage.txt
['000000900000660300000010020000406001000012800\n', '000000900000660300
000010020000406002000000001\n']
45
20180621_ballotimage.txt

```
['000002000001288600000010020000009001000018100\n', '000002000001288600
000010020000009002000018500\n']
45
D5-BallotImage.txt
['000002200000739700000010020000161001000011400\n', '000002200000739700
000010020000161002000010800\n']
45
CityWide_Ballot_Image.txt
['000000300000835700000030020000066001000007600\n', '000000300000835700
000030020000066002000007700\n']
45
BallotImage_San Fran_Mayor_Nov 2007.txt
['0001\n', '0002\n']
4
BallotDescription_San Fran_2007.txt
[' 15      BALLOTS CAST - TOTAL
\n', ' 20      BALLOTS CAST - TOTAL CARD 1
\n']
27
```

```
In [16]: list_of_ballot_files=[]
         for filename in list_of_files:
             if filename.endswith('.txt'):
                 with open(filename,'r') as fil:
                     file_contents = fil.readlines()
                 if len(file_contents[0].strip())==45:
                     print(filename)
                     list_of_ballot_files.append(filename)
                     print(file_contents[1])
```

voting_data/San_Fransisco/San Francisco/San Fran_Nov 2010_District 10 S
upervisors/BallotImage-D10.txt
000000600001706700000040020000274002000015700

voting_data/San_Fransisco/San Francisco/San Fran_Nov 2010_District 2 Su
pervisors/BallotImage-D2.txt
000000700001712400000090020000331002000000001

voting_data/San_Fransisco/San Francisco/2015 All offices/20151119_ballo
timage.txt
000000100002610500000010020000012002000003700

voting_data/San_Fransisco/San Francisco/San Fran_Nov 2014_District 10 S
upervisors/D10_BallotImage.txt
000003300000385300000010020000054002000012800

voting_data/San_Fransisco/San Francisco/San Fran Nov 2011 Sheriff/Sheri
ff-BallotImage.txt
000000200004728200000010020000003002000002800

voting_data/San_Fransisco/San Francisco/San Fran Nov 2011 District Atto
rney/DA-BallotImage.txt
000000100004728200000010020000003002000002300

voting_data/San_Fransisco/San Francisco/San Fran_Nov 2010_District 8 Su
pervisors/BallotImage-D8.txt
000001000001706800000040020000513002000014200

voting_data/San_Fransisco/San Francisco/San Fran Nov 2011 Mayor/Mayor-B
allotImage.txt
000000300004728200000010020000003002000003700

voting_data/San_Fransisco/San Francisco/San Fran_Nov 2012_District 7 Su
pervisors/D7-BallotImage.txt
000002300000759400000010020000244002000011800

voting_data/San_Fransisco/San Francisco/San Fran_Nov 2010_District 6 Su
pervisors/BallotImage-D6.txt
000000900001707600000130020000457002000012800

voting_data/San_Fransisco/San Francisco/San Fran_Nov 2016/20161206_ball
otimage.txt
000000900000660300000010020000406002000000001

voting_data/San_Fransisco/San Francisco/San Fran June 2018/20180621_bal
lotimage.txt
000002000001288600000010020000009002000018500

voting_data/San_Fransisco/San Francisco/San Fran_Nov 2012_District 5 Su
pervisors/D5-BallotImage.txt
000002200000739700000010020000161002000010800

voting_data/San_Fransisco/San Francisco/San Fran_Nov 2008_All District
races/CityWide_Ballot_Image.txt
000000300000835700000030020000066002000007700

```
In [17]: results={'leading candidate in first round has more than 50% of first cho
ice votes':[],
                'leading candidate in first round vote has between 50% and 45% o
f first choice votes':[],
                'leading candidate in first round vote has less than 45% of firs
t choice votes':[]}

for ballot in list_of_ballot_files:
    results = sort_elections_into_bins(results,ballot)
```

```
In [18]: for k,v in results.items():
          print(k,':')
          for election in v:
              print('    ',election.replace('voting_data/','').replace('.txt',
''))
```

```
leading candidate in first round has more than 50% of first choice vote
s :
leading candidate in first round vote has between 50% and 45% of first
choice votes :
    San_Fransisco/San Francisco/San Fran_Nov 2014_District 10 Superviso
rs/D10_BallotImage
leading candidate in first round vote has less than 45% of first choice
votes :
    San_Fransisco/San Francisco/San Fran_Nov 2010_District 10 Superviso
rs/BallotImage-D10
    San_Fransisco/San Francisco/San Fran_Nov 2010_District 2 Supervisor
s/BallotImage-D2
    San_Fransisco/San Francisco/2015 All offices/20151119_ballotimage
    San_Fransisco/San Francisco/San Fran Nov 2011 Sheriff/Sheriff-Ballo
tImage
    San_Fransisco/San Francisco/San Fran Nov 2011 District Attorney/DA-
BallotImage
    San_Fransisco/San Francisco/San Fran_Nov 2010_District 8 Supervisor
s/BallotImage-D8
    San_Fransisco/San Francisco/San Fran Nov 2011 Mayor/Mayor-BallotIma
ge
    San_Fransisco/San Francisco/San Fran_Nov 2012_District 7 Supervisor
s/D7-BallotImage
    San_Fransisco/San Francisco/San Fran_Nov 2010_District 6 Supervisor
s/BallotImage-D6
    San_Fransisco/San Francisco/San Fran_Nov 2016/20161206_ballotimage
    San_Fransisco/San Francisco/San Fran June 2018/20180621_ballotimage
    San_Fransisco/San Francisco/San Fran_Nov 2012_District 5 Supervisor
s/D5-BallotImage
    San_Fransisco/San Francisco/San Fran_Nov 2008_All District races/Ci
tyWide_Ballot_Image
```