Day 1: Let’s Play with Data!

2008 Presidential Election Donations
$1,644,712,232
6X
Earth’s circumference!
10%
U.S. Dept of Energy’s Budget!
Federal Election Commission
(FEC)
2008 Donations Dataset

C00430470,"P80002801",
"McCain, John S",
"WAHLERS, STUART E. MR.",
"A.P.O.", "AE", "091280013",
"U.S. MILITARY","SOLDIER",
250, 31-JUL-08,
"","X",
"TRANSFER FROM MCCAIN VICTORY 2008",
"SA18",377957
Contributions to All Candidates

Summary

What's Included Here?

Summary

- Contributions
  - Individual: $1,334,224,621
  - PAC: $5,178,919
  - Party: -$74,043
  - Candidate: $60,230,846
- Federal Funds: $103,732,646
- Transfers-In: $154,259,577
- Disbursements: $1,603,258,818
- Cash On Hand: $49,237,448

Size of Donations

- $200 and Under: $427,817,410
- $200.01 - $499: $140,622,952
- $500 - $999: $134,573,216
- $1000 - $1999: $206,433,548
- $2000 and Over: $418,956,583

This content is in the public domain, from the Federal Election Commission.
Negative Donations
We provide `defaultdict` a function to call and return if the key doesn't exist. This is nice because we can assume a default value. It is otherwise used as a normal python dictionary:

```python
>>> d = collections.defaultdict(lambda:0)
>>> d['foo']
0
>>> d['bar'] += 1
```

- We parse the dates using the `datetime` module's `strptime` function. The string `%%d-%%b-%%y` is called a date format string.
- In the loop that reads the data, we record the total donation amount for each date.
- Finally, we need to sort the data in `obamdonations` by the date (the key). `sorted(l, key=f)` returns a sorted copy of `l` and calls `f` to extract the key to use for comparison.
- `zip(*pairs)` then unzips the list of pairs into two lists.

You should see something like this:

![Graph](image-url)