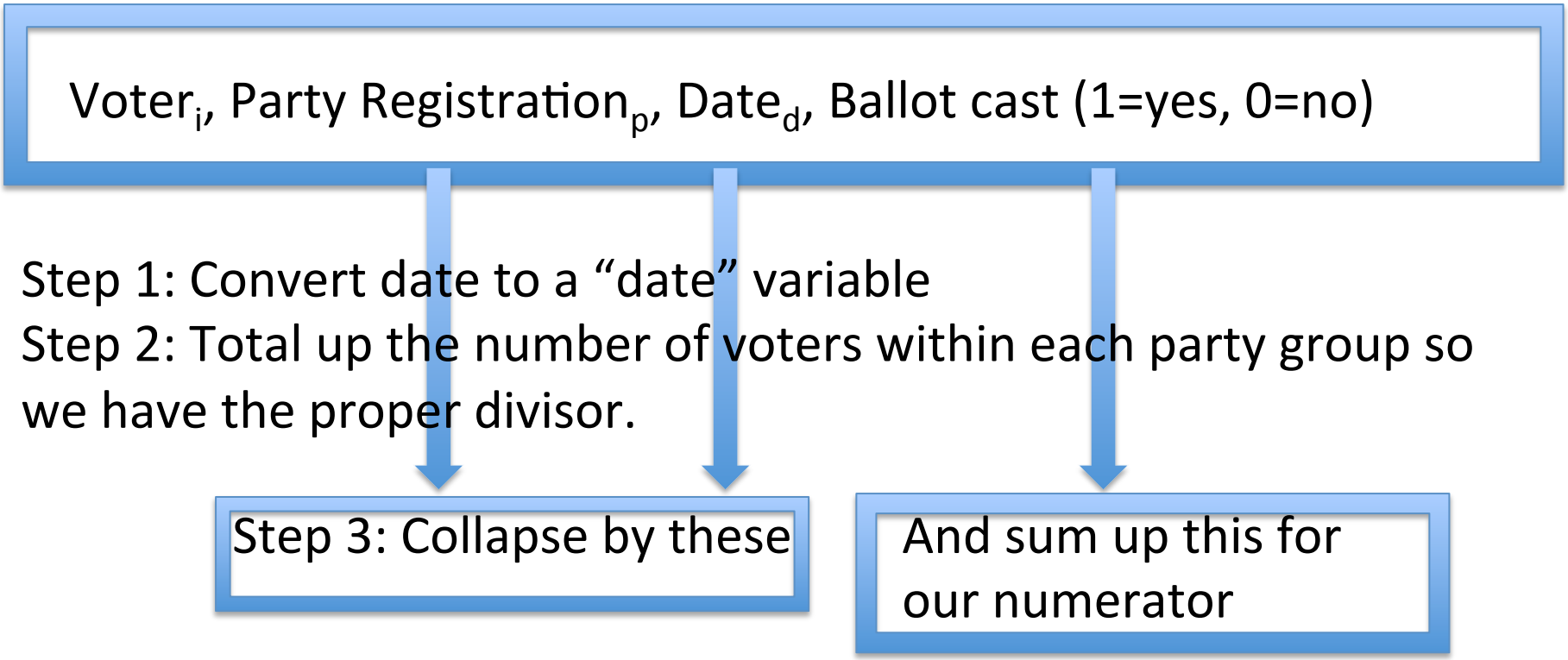
Collapsing the NC Voter File

Political Science 311

Computer Lab 5

Starting and finishing

- Starting:
  - Unit of analysis is the individual voter, the voter's partisanship, whether or not the voter cast an accepted ballot, and the date the ballot was cast.
- Finishing:
  - We want to plot by day the number of accepted ballots and what percent of registered voters have cast a ballot
- Data:
  - ftp://www.app.sboe.state.nc.us/enrs/absentee11xx06xx2012.zip



Step 4: Now our unit of analysis is party group / date. Create a cumulative sum (1% day 1, 1%+3% day 2, 1+3+2 day 3, etc).
Step 5: Divide the cumulative sum / total party registrants = Percent of party registrants who voted, by day.

```
    // Read in the raw data
insheet using "absentee11xx06xx2012.txt"

    // Convert the date of ballot return variable
gen ballot_return_date=date(ballot_rtn_dt,"MDY")
format ballot_return_date %tdNN/DD

    // Check this table for outlier values, probably data entry errors
tab ballot_return_date

    // Set a fixed start date for graphics

replace ballot_return_date=date("08Oct2012","DMY") ///
    if ballot_return_date<date("08Oct2012","DMY")
    // NOTE YOU NEED TO CHANGE THIS AS WE DOWNLOAD NEWER FILES

replace ballot_return_date=. ///
    if ballot_return_date>date("22Oct2012","DMY")
// THIS COMMAND SETS ILLEGAL VALUES TO MISSING


    // Initialize the vote variable (set to zero for all cases)
gen voted=0
    // If ballot has been accepted, then count as a vote
replace voted=1 if ballot_rtn_status=="ACCEPTED"
```

Read in the data

Convert date

Set a start point

Illegal dates

Start vote counter

```
    // The egen commands create, for each record in the dataset, the
    // total number of cases that have the values "DEM", "REP", and "UNA"
    //
    // Change for one stop, create separate measure for all voter, for
    // vbm voters, and for in person voters
    //

egen dems=total(voter_party_code=="DEM")
egen reps=total(voter_party_code=="REP")
egen una=total(voter_party_code=="UNA")
egen vbmdems=total(voter_party_code=="DEM") if ballot_req_delivery=="MAIL"
egen vbmreps=total(voter_party_code=="REP") if ballot_req_delivery=="MAIL"
egen vbmuna=total(voter_party_code=="UNA") if ballot_req_delivery=="MAIL"
egen onedems=total(voter_party_code=="DEM") if ballot_req_delivery=="IN
PERSON"
egen onereps=total(voter_party_code=="REP") if ballot_req_delivery=="IN
PERSON"
egen oneuna=total(voter_party_code=="UNA") if ballot_req_delivery=="IN
PERSON"
```

Use egen to create the divisor (All, VBM, and Early in person, by party. Since all Democrats, for example, will have the same value on the "dems" variables, we only need to grab one value when we collapse

```
// ***** ALL VOTERS *****
//
collapse (max) dems (max) reps (max) una (sum) voted, ///
    by(ballot_return_date voter_party_code)
```

Collapse by date and party

```
// These commands create a cumulative sum of the ballot returns by Party.
//
gen dvotetotal=sum(voted) if voter_party_code=="DEM"
gen rvotetotal=sum(voted) if voter_party_code=="REP"
gen ivotetotal=sum(voted) if voter_party_code=="UNA"
```

Create cumulative sums

```
//  Party totals obtained from here
//  http://www.app.sboe.state.nc.us/webapps/voter_stats/results.aspx?
date=10-20-2012

gen dempct=dvotetotal/2839724
gen reppct=rvotetotal/2036312
gen unapct=ivotetotal/1693275
```

Percent turnout among all registered partisans

```
gen dempct=dvotetotal/vbmdems
gen reppct=rvotetotal/vbmreps
gen unapct=ivotetotal/vbmuna
```

Percent or requested VBM ballots that have been returned, by party

```
twoway (line dempct ballot_return_date ///
               if ballot_return_date>td(01Sept2012)&ballot_return_date<td(23Oct2012)) ///
          (line reppct ballot_return_date ///
               if ballot_return_date>td(01Sept2012)&ballot_return_date<td(23Oct2012)) ///
          (line unapct ballot_return_date ///
     if ballot_return_date>td(01Sept2012)&ballot_return_date<td(23Oct2012)), ///
          ytitle("Proportion Accepted Ballots") xtitle("Ballot Return Date") xtick(#10)
xlabel(#10) ///
          xlabel(, labsize("medium") angle("forty_five")) ///
          title("Proportion Accepted Ballots in NC") ///
          subtitle("Out of All Registered Voters by Return Date") ///
          note("From NC Board of Elections Website") legend(order(1 "Democratic Party"
2 "Republican Party" 3 "Unaffiliated"))

graph export ncabs_proportion_allearly_byparty.png, replace
```

Sample overlay graphic command:

- Three separate twoway line graphs plot the percent voted against date
- Each graph is limited to a legal date range

Graphic options:

- Set ytitle and xtitle
- Set the number of ticks and labels on the x-axis
- Set the graph title and subtitle, add a note (caption)
- Control the legend so that it uses “nice” labels and not the variable labels

Export graphic to a common format.