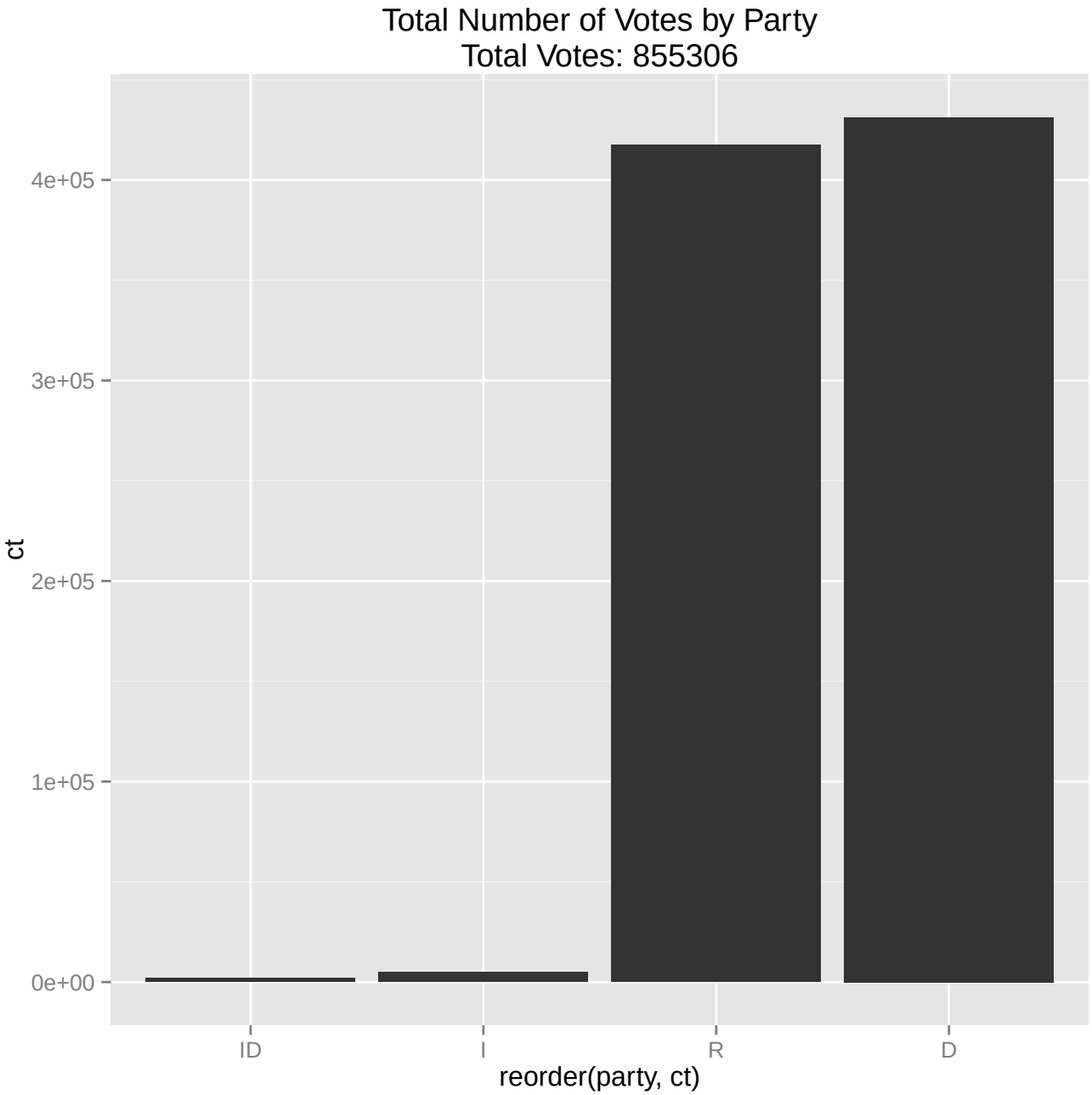
Analyzing Roll Call Votes from the US Senate

Jordan Poles

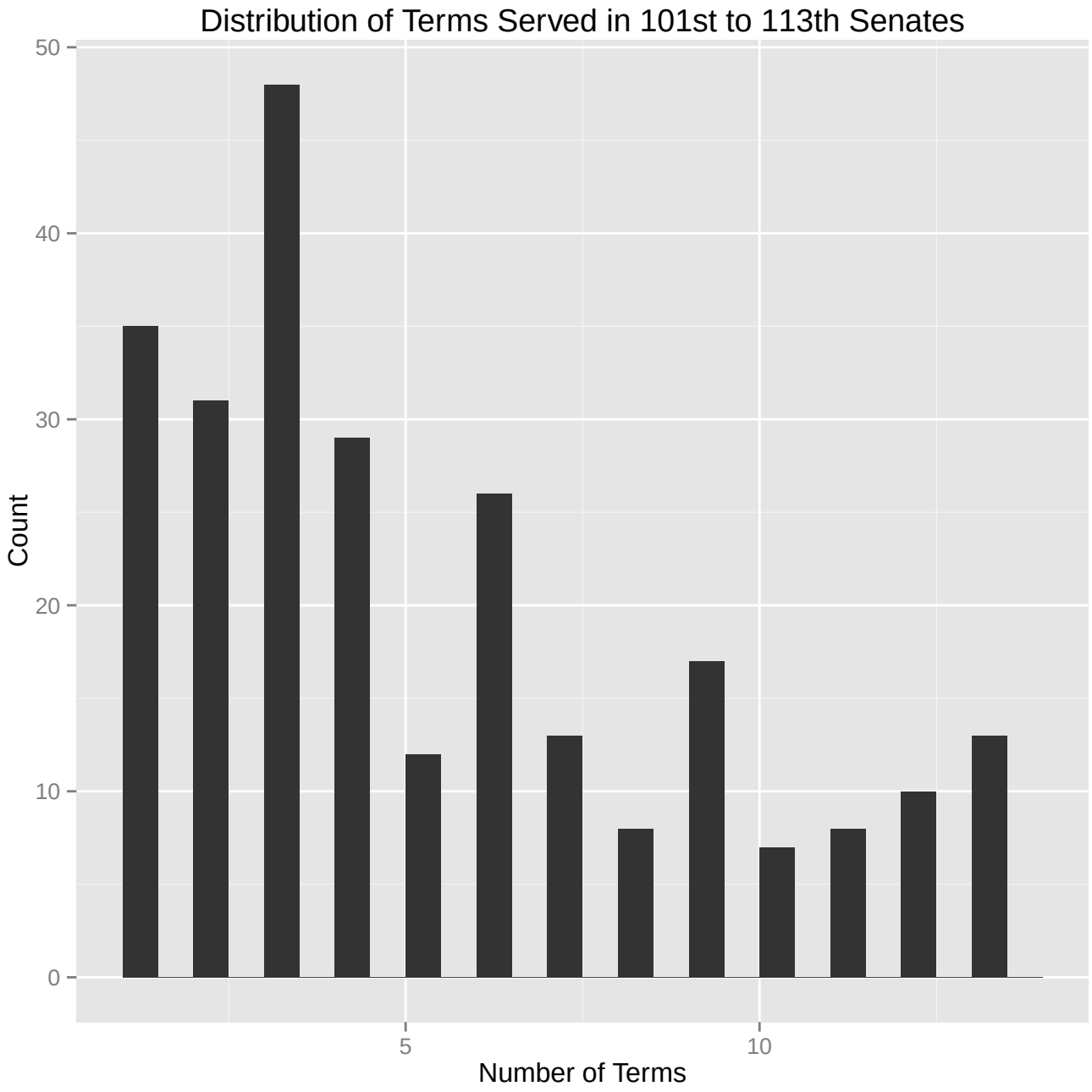
November 19, 2015

```
#Setting up libraries and imports
  require(RSQLite)

## Loading required package:  RSQLite
## Loading required package:  DBI

  require(ggplot2)

## Loading required package:  ggplot2

  library(grid);
  source("../config.R")
  source("../voteAnalysis.R")
```

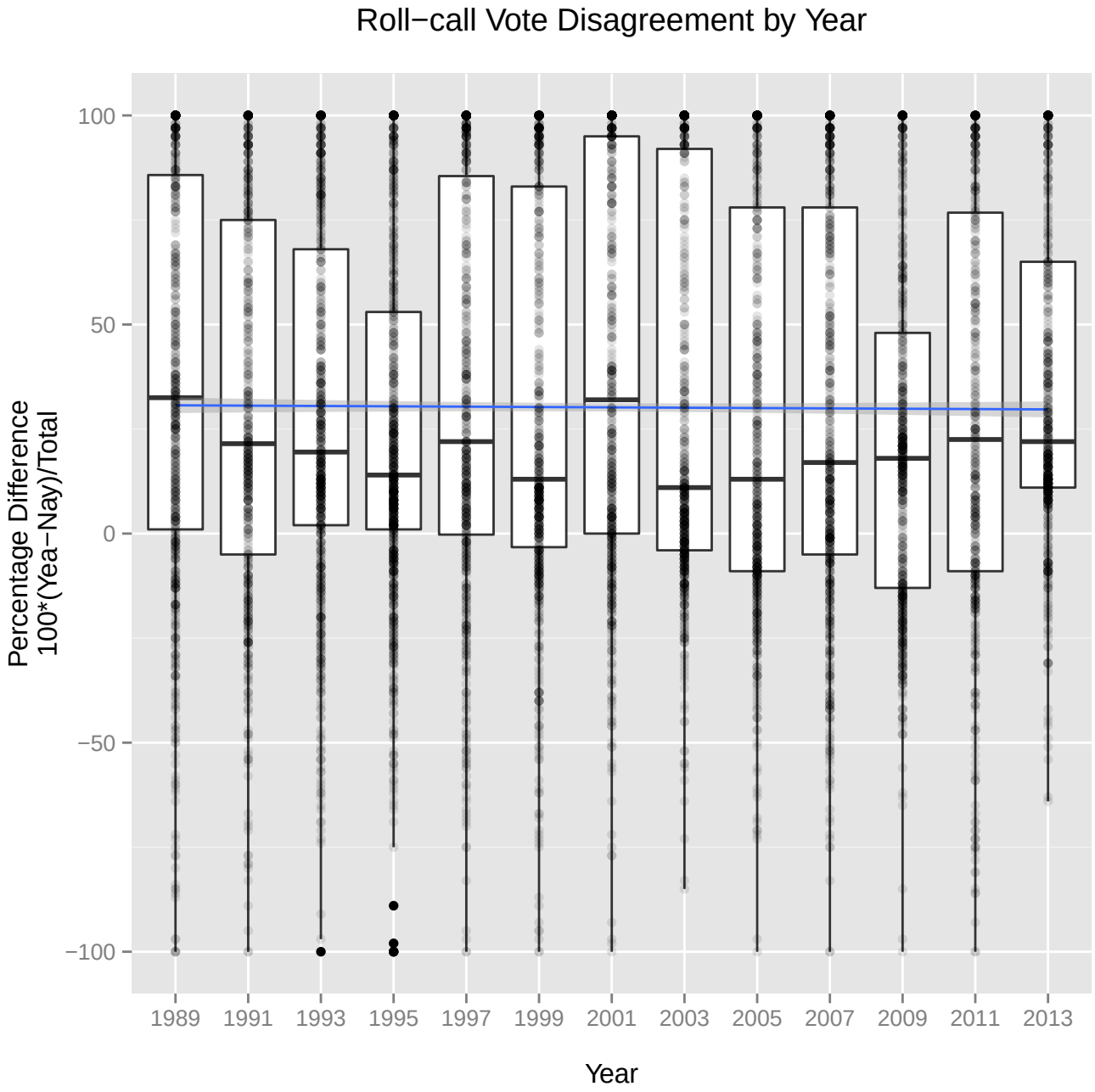
```
query = "SELECT party, count(*) as ct FROM votes GROUP BY party"
partyVoteTotals = queryDB(query, "data.sqlite")
ggplot(partyVoteTotals, aes(x = reorder(party, ct), y = ct)) + geom_bar(stat = "identity") +
    ggtitle(sprintf("Total Number of Votes by Party\nTotal Votes: %s", sum(partyVoteTotals$ct)))
```



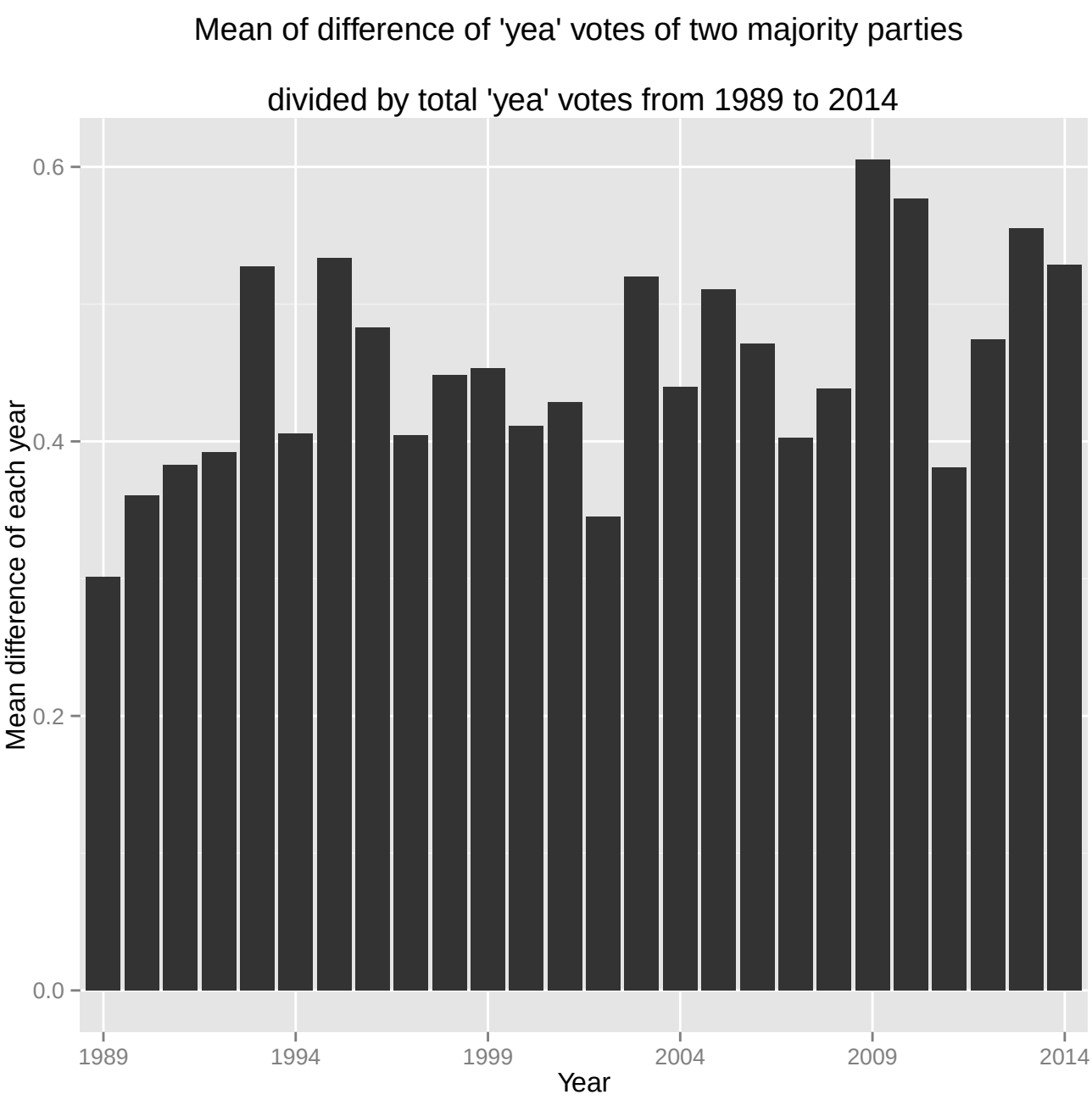
```
query = "SELECT id, first_name, last_name, party, seniority, count(*) AS ct FROM members GROUP BY id ORDER BY ct
senatorTotals = queryDB(query, "data.sqlite")
ggplot(senatorTotals, aes(x = ct)) + xlim(1, 14) + geom_bar(binwidth = 0.5) +
    xlab("Number of Terms") + ylab("Count") + ggtitle("Distribution of Terms Served in 101st to 113th Senates")
```



```
query = "SELECT yeas, nays, (yeas+nays) as total, (100*(yeas-nays)/(yeas+nays)) as voteDiff, congressNumber, ses
rollCallStats = queryDB(query, "data.sqlite")
rollCallStats$year = apply(rollCallStats, 1, function(x) {
    congressToYear(x["congressNumber"], x["sessionNumber"])
})
ggplot(rollCallStats, aes(x = as.factor(year), y = voteDiff)) + geom_boxplot() +
    geom_smooth(aes(group = 1), method = "lm") + geom_point(alpha = 0.1) + ggtitle("Roll-call Vote Disagreement
    xlab("\nYear") + ylab("Percentage Difference\n100*(Yea-Nay)/Total")
```



```
query = "select r.year as year, r.voteNumber as voteNumber,
            abs(r.c - d.c) * 1.0 / (r.c + d.c) as diff
         from (select voteNumber, year, count(*) as c
               from votes
               where vote == 'Yea' and party == 'R' group by year, voteNumber)
               as r
              join
            (select voteNumber, year, count(*) as c
               from votes
               where vote == 'Yea' and party == 'D' group by year, voteNumber)
               as d
            on r.voteNumber == d.voteNumber and r.year == d.year"
yeaDiff = queryDB(query, "data.sqlite")
yeaDiffMean = setNames(aggregate(diff ~ year, yeaDiff, mean), c("year", "mean"))
yeaDiffSd = setNames(aggregate(diff ~ year, yeaDiff, sd), c("year", "std"))
yeaDiffDistribution = merge(yeaDiffMean, yeaDiffSd, by="year")
ggplot(yeaDiffDistribution) +
aes(x = year, y = mean) +
scale_x_discrete(breaks=c(1989,1994,1999,2004,2009,2014)) +
labs(title="Mean of difference of 'yea' votes of two majority parties
     \ndivided by total 'yea' votes from 1989 to 2014") +
xlab("Year") +
ylab("Mean difference of each year") +
geom_bar(stat="identity")
```


Mean of difference of 'yea' votes of two majority parties
divided by total 'yea' votes from 1989 to 2014
Mean difference of each year
0.6
0.4
0.2
0.0
1989
1994
1999
2004
2009
2014
Year

```
partyData = queryDB("SELECT party as Party, count(*) as ct FROM members WHERE party!='ID' GROUP BY party")
a = ggplot(partyData, aes(x=reorder(Party, ct), y=ct))+geom_bar(stat="identity", fill=c("green", "red", "#4169
independentDataByYear = queryDB("SELECT party as Party, congressNumber, count(*) as ct FROM members WHERE part
b=ggplot(independentDataByYear, aes(x=congressNumber, y=ct))+geom_point()+geom_line()+ggtitle("")+ylim(0,5)+xl
partyDataByYear = queryDB("SELECT party as Party, congressNumber, count(*) as ct FROM members WHERE party!='ID
c = ggplot(partyDataByYear, aes(x=congressNumber, y=ct, fill=Party))+scale_fill_manual(name = "Party", values
grid.newpage()
pushViewport(viewport(layout=grid.layout(2, 2)))
print(a, vp = viewport(layout.pos.row=1, layout.pos.col=1))
print(b, vp = viewport(layout.pos.row=1, layout.pos.col=2))
print(c, vp = viewport(layout.pos.row=2, layout.pos.col=1:2))
```