

Analyzing Roll Call Votes from the US Senate

Jordan Poles

November 20, 2015

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

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

require(ggplot2)

## Loading required package: ggplot2

require(grid)

## Loading required package: grid

require(reshape2)

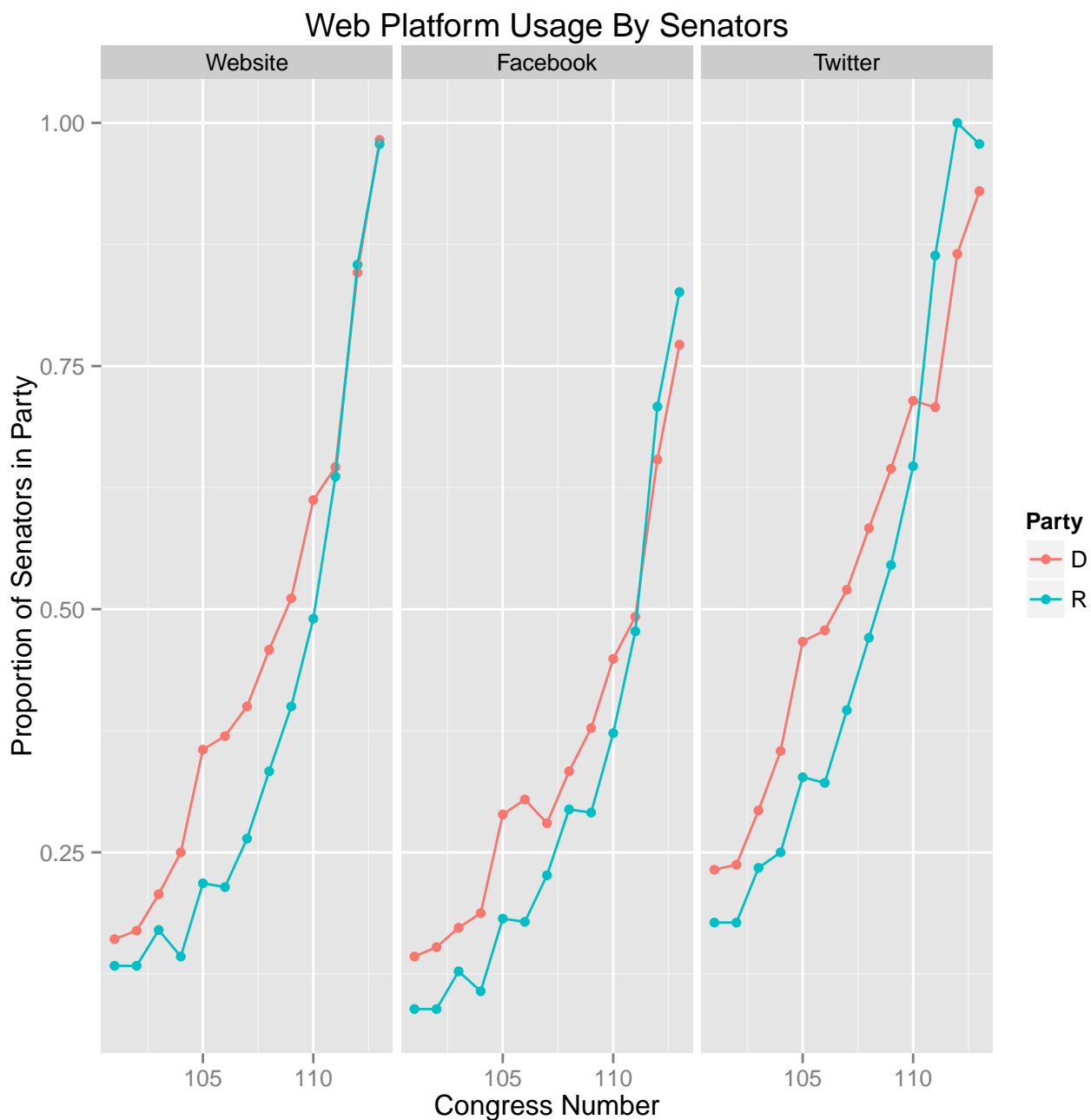
## Loading required package: reshape2

source("../config.R")
source("../voteAnalysis.R")
```

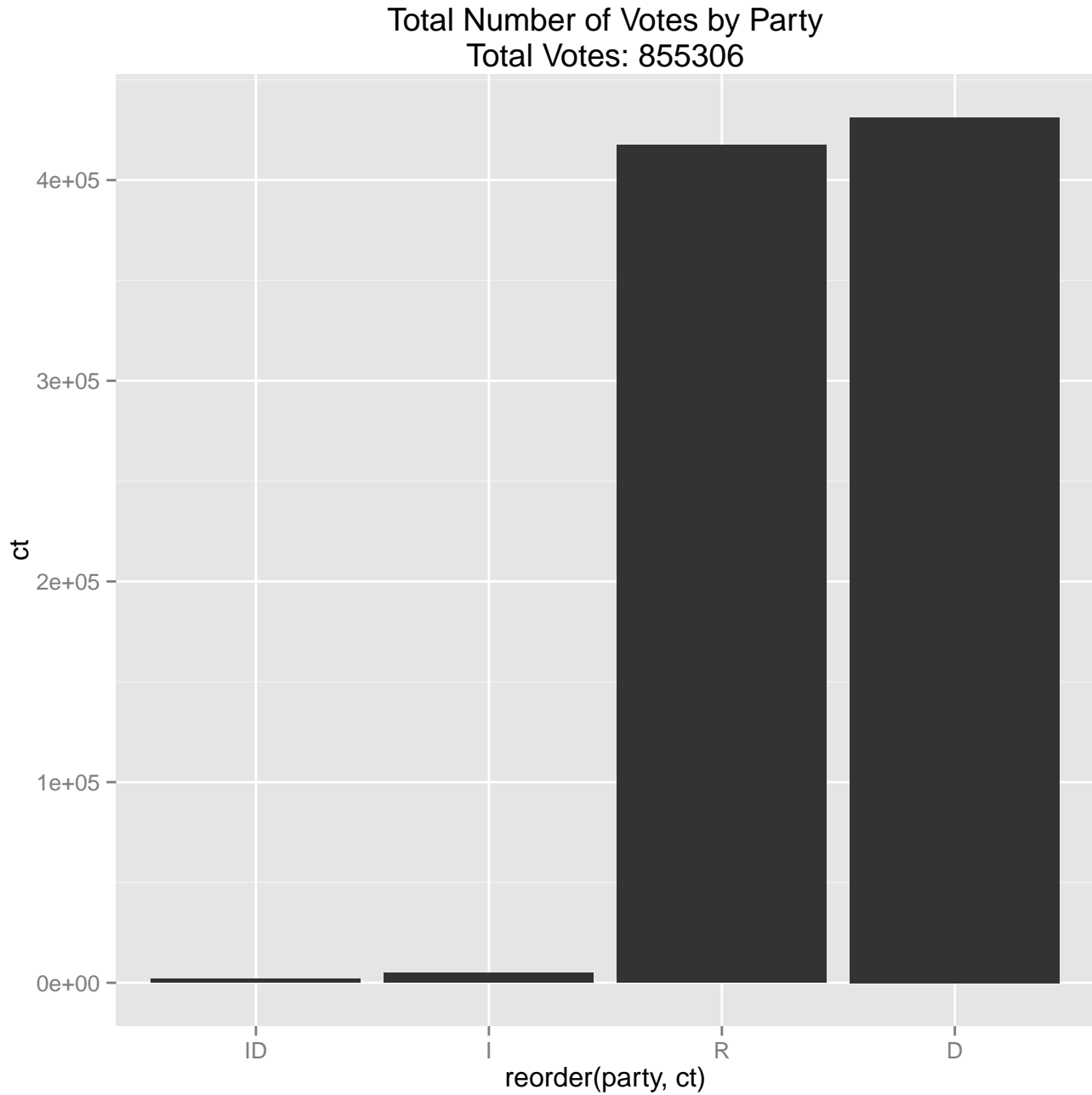
```

websiteCt = queryDB("SELECT party as Party, congressNumber, count(*) as webct FROM members WHERE URL!='' AND P
twitterCt = queryDB("SELECT party as Party, congressNumber, count(*) as twitterct FROM members WHERE twitter_a
fbCt = queryDB("SELECT party as Party, congressNumber, count(*) as fbct FROM members WHERE facebook_account!='
totalCt = queryDB("SELECT party as Party, congressNumber, count(*) as totalct FROM members WHERE Party!='I' GP
mediaCt = merge(websiteCt, twitterCt, by=c("Party", "congressNumber"))
mediaCt = merge(mediaCt, fbCt, by=c("Party", "congressNumber"))
mediaCt = merge(mediaCt, totalCt, by=c("Party", "congressNumber"))
mediaCt$Website = mediaCt$webct/mediaCt$totalct
mediaCt$Twitter = mediaCt$twitterct/mediaCt$totalct
mediaCt$Facebook = mediaCt$fbct/mediaCt$totalct
mediaPlotData = melt(mediaCt, measure.vars = c("Website", "Facebook", "Twitter"), id.vars = c("Party", "congre
ggplot(mediaPlotData, aes(x=congressNumber, y=value, color=Party))+geom_point()+geom_line()+facet_wrap(~variab

```



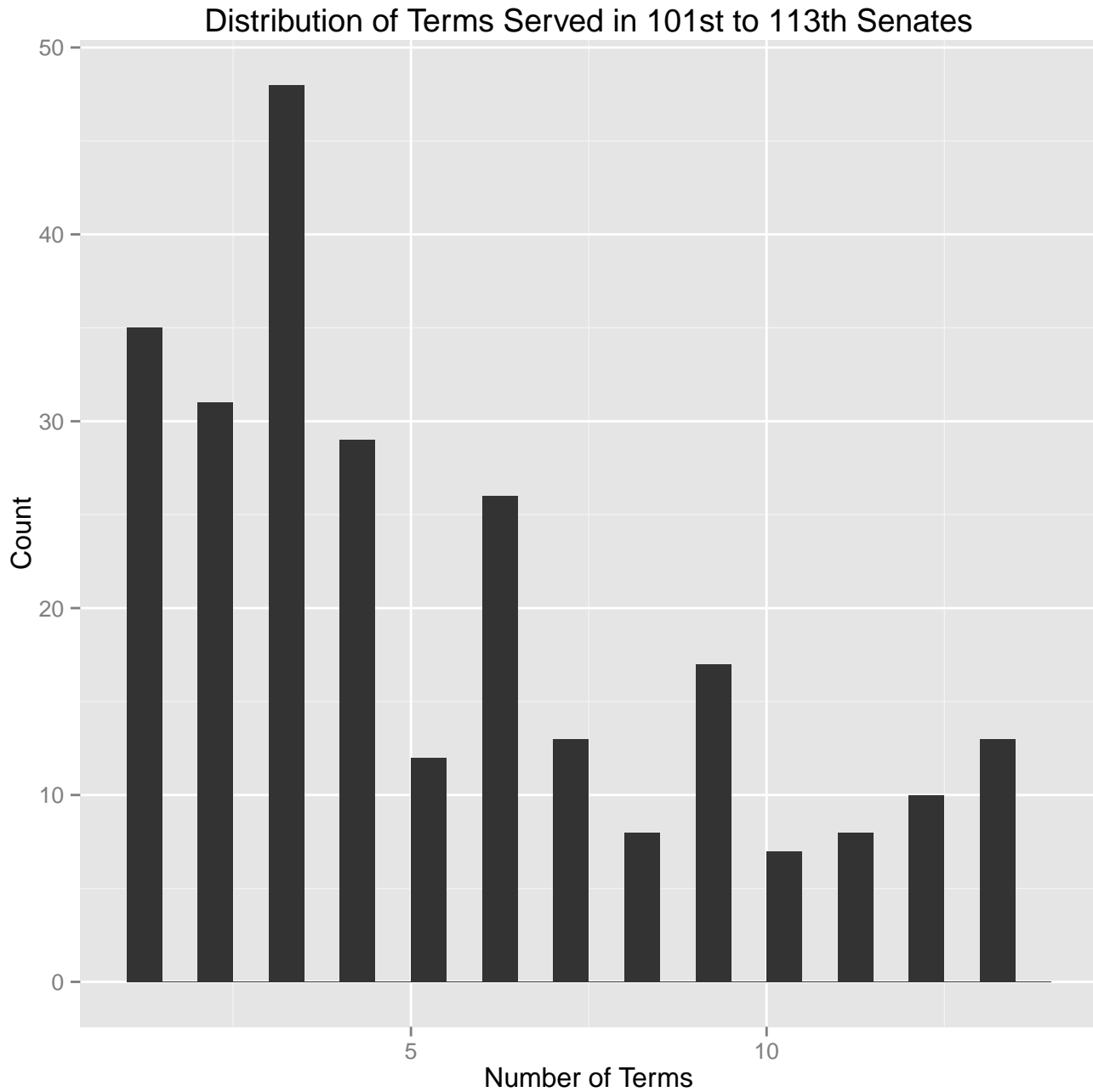
```
query = "SELECT party, count(*) as ct FROM votes GROUP BY party"
partyVoteTotals = queryDB(query, "data.sqlite")
ggplot(partyVoteTotals, aes(x = reorder(party, ct), y = ct)) + geom_bar(stat = "identity") +
  ggtitle(sprintf("Total Number of Votes by Party\nTotal Votes: %s", sum(partyVoteTotals$ct)))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



```

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")

```

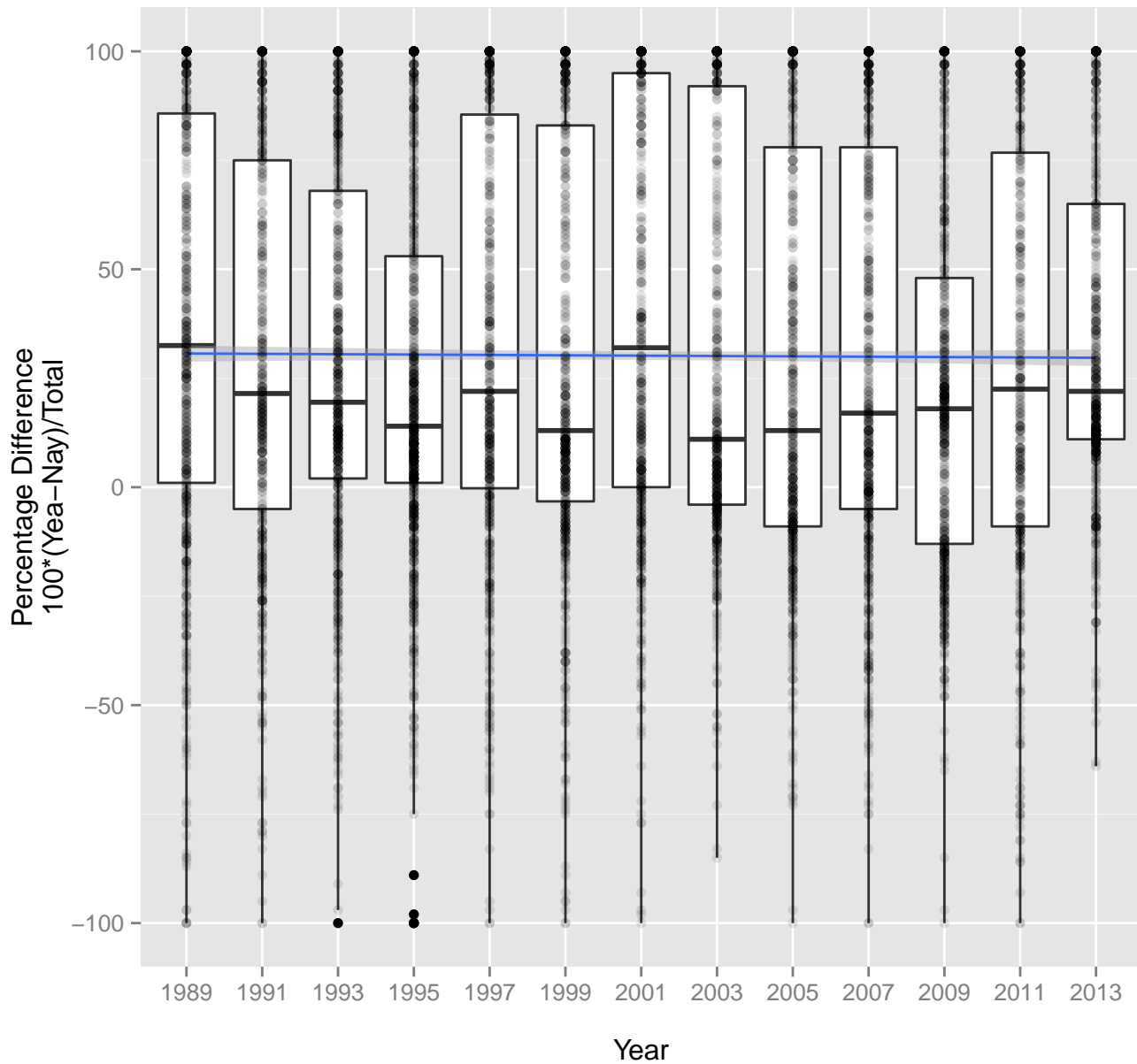


```

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")

```

Roll-call Vote Disagreement by Year



```

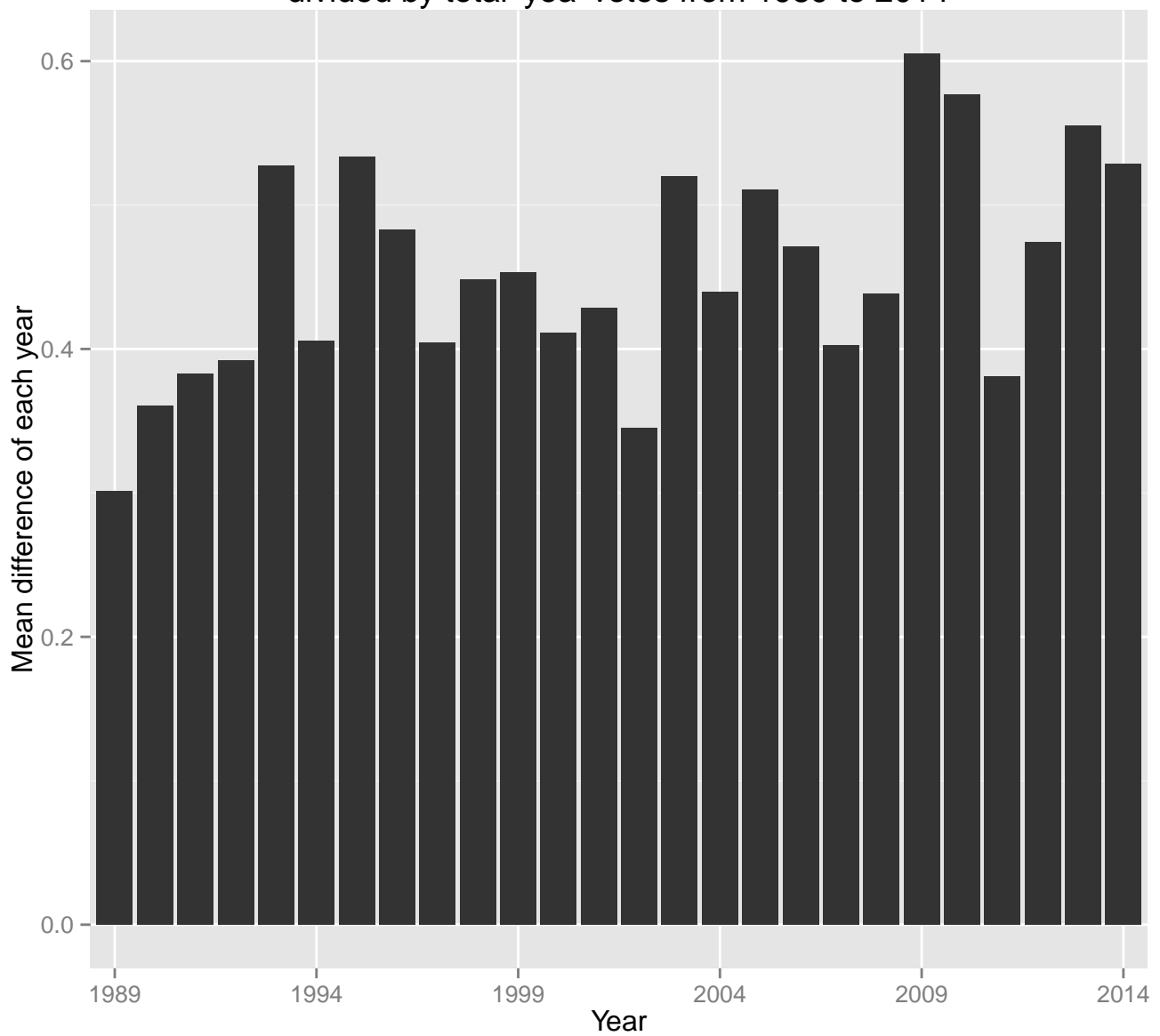
query = "select r.year as year, r.voteNumber as voteNumber,
      abs(r.c - d.c) * 1.0 / (r.c + d.c) as diff
from (select voteNumber, year, count(*) as c
      from votes
      where vote == 'Yea' and party == 'R' group by year, voteNumber)
      as r
join
(select voteNumber, year, count(*) as c
      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

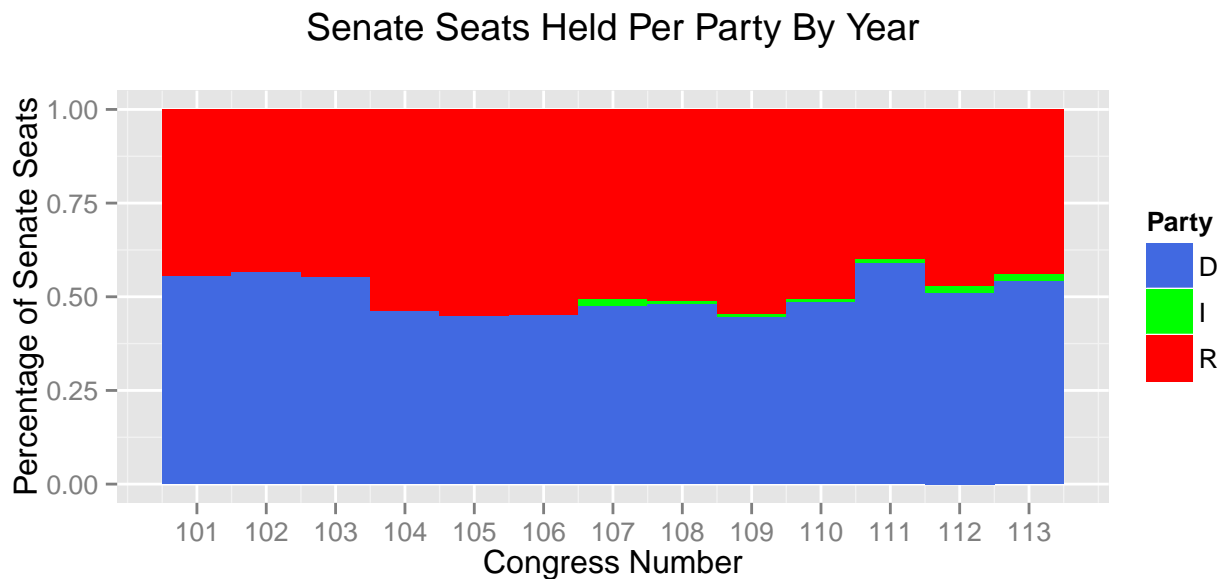
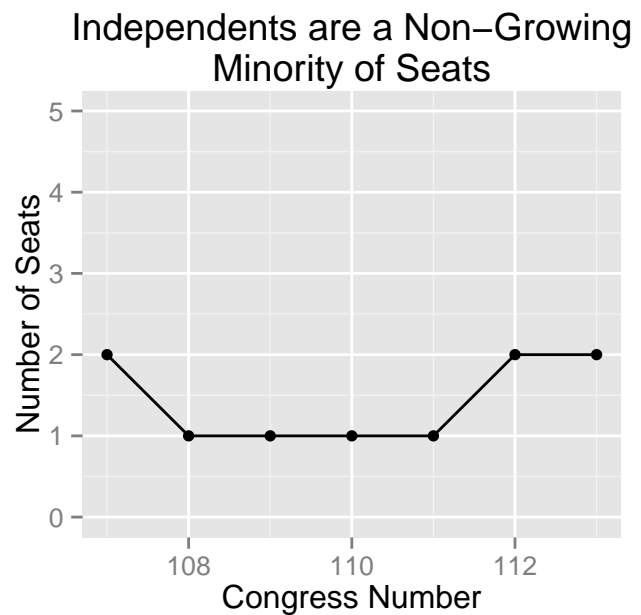
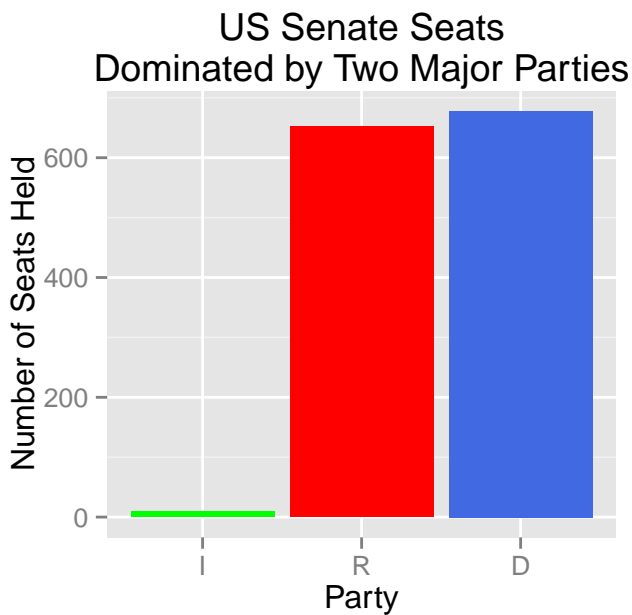
divided by total 'yea' votes from 1989 to 2014



```

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", "#4169E1"))
independentDataByYear = queryDB("SELECT party as Party, congressNumber, count(*) as ct FROM members WHERE party='ID'")
b=ggplot(independentDataByYear, aes(x=congressNumber, y=ct))+geom_point()+geom_line()+ggtitle("")+ylim(0,5)+xlab("Congress Number")
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=c("green", "red", "#4169E1"))
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))

```




```

passedQuery <- "select 'P' as res, congressNumber, count(*) as cnt
               from senateRollCalls
               where result == 'Agreed to'
                  or result == 'Confirmed'
                  or result == 'Passed'
               group by congressNumber"
passedResults <- queryDB(passedQuery, "data.sqlite")
failedQuery <- "select 'F' as res, congressNumber, count(*) as cnt
               from senateRollCalls
               where result == 'Rejected'
               group by congressNumber"
failedResults <- queryDB(failedQuery, "data.sqlite")
results <- rbind(passedResults, failedResults)
ggplot(results) + aes(x=congressNumber, y=cnt, fill=res) +
  geom_histogram(position="fill", stat="identity", width=1) +
  scale_fill_manual(name = "res", values = c("red", "green")) +
  xlab("Congress Number\n") +
  scale_x_continuous(breaks=101:113) +
  ylab("\nResult counts of bills") +
  ggtitle("Passed/failed bills count By Year\n")

```

Passed/failed bills count By Year

