Analyzing Roll Call Votes from the US Senate

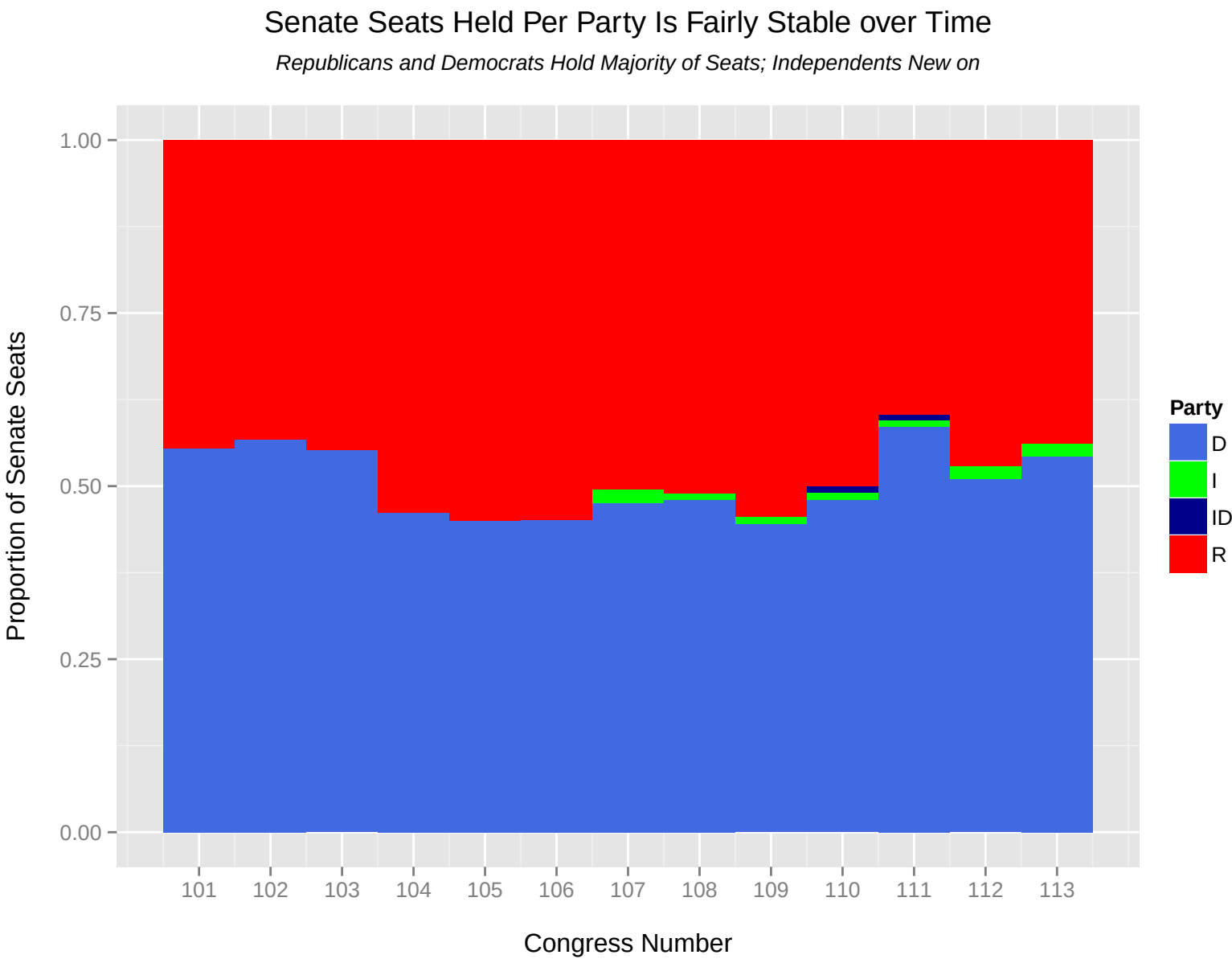
Jordan Poles

November 24, 2015

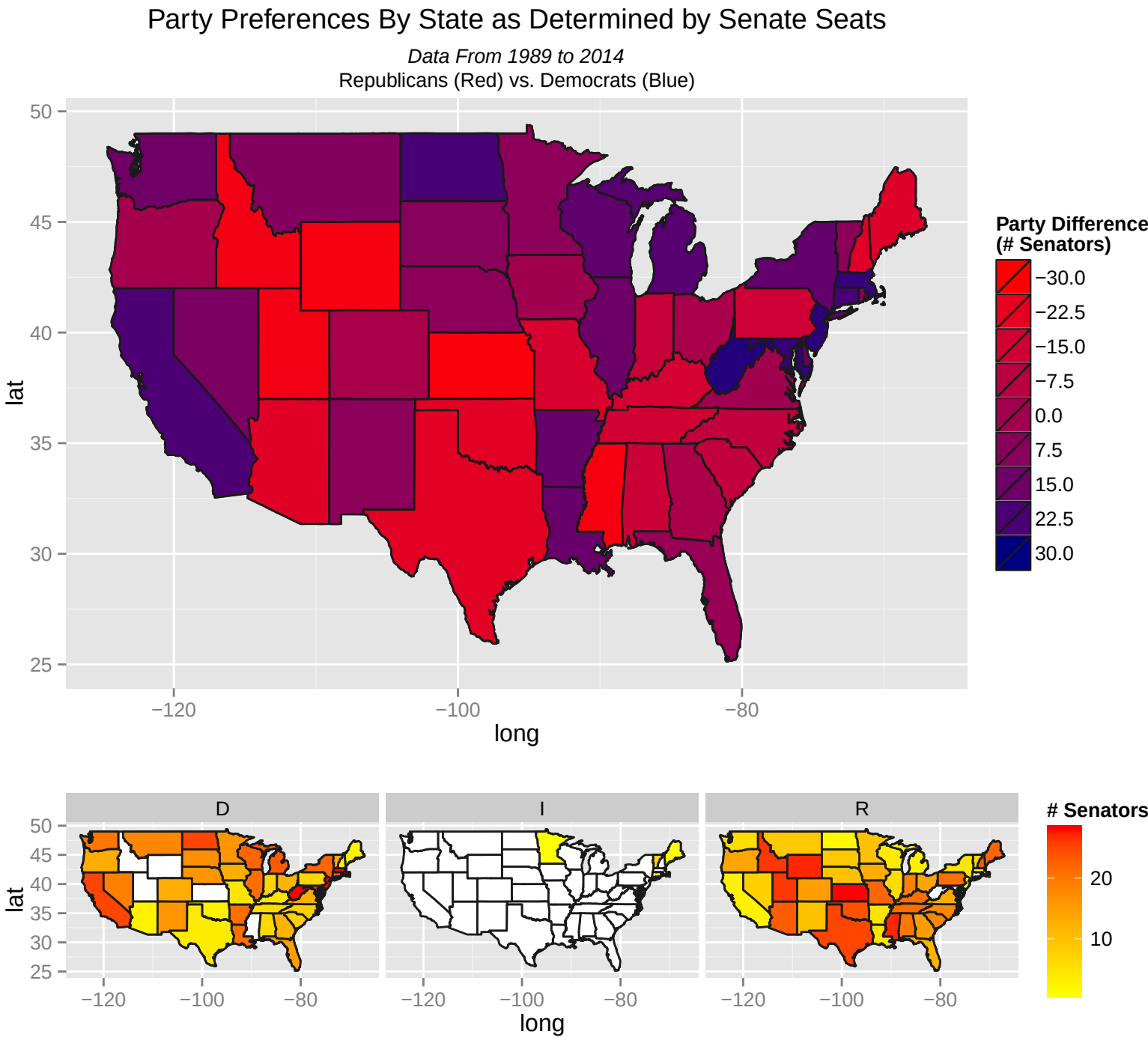
```
## Loading required package:  RSQLite
## Loading required package:  DBI
## Loading required package:  ggplot2
## Loading required package:  grid
## Loading required package:  reshape2
```

Examining The Members of Senate

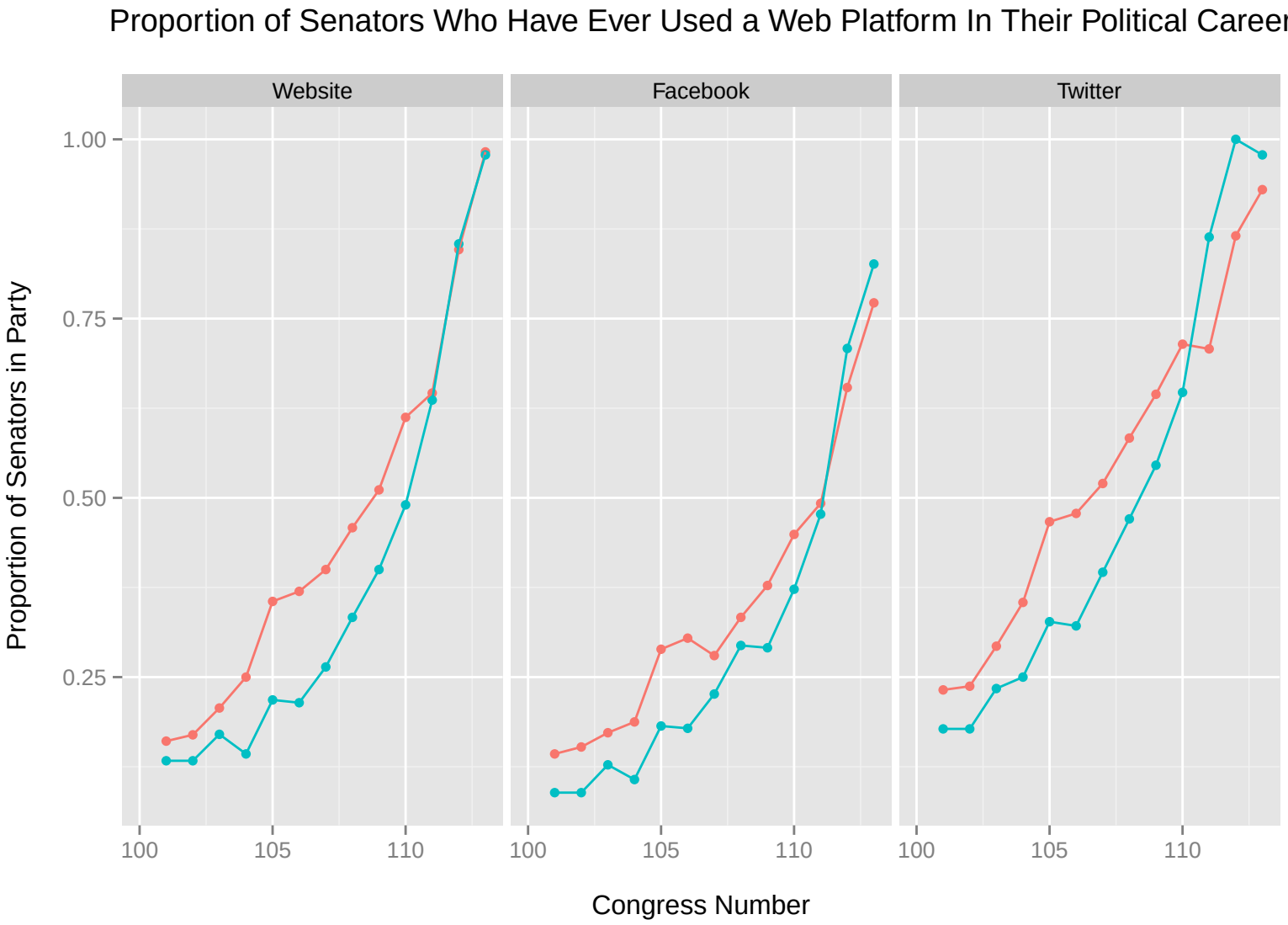
```
partyDataByYear = queryDB(paste("SELECT party as Party, congressNumber, count(*) as ct",
    "FROM members", "GROUP BY party, congressNumber"))
ggplot(partyDataByYear, aes(x = congressNumber, y = ct, fill = Party)) + scale_fill_manual(name = "Party",
    values = c("#4169E1", "green", "dark blue", "#FF0000")) + geom_histogram(position = "fill",
    stat = "identity", width = 1) + xlab("\nCongress Number") + scale_x_continuous(breaks = 101:113) +
    ylab("Proportion of Senate Seats\n") + ggtitle(expression(atop("Senate Seats Held Per Party Is Fairly Stable ov
    atop(italic("Republicans and Democrats Hold Majority of Seats; Independents New on"),
        ""))))
```



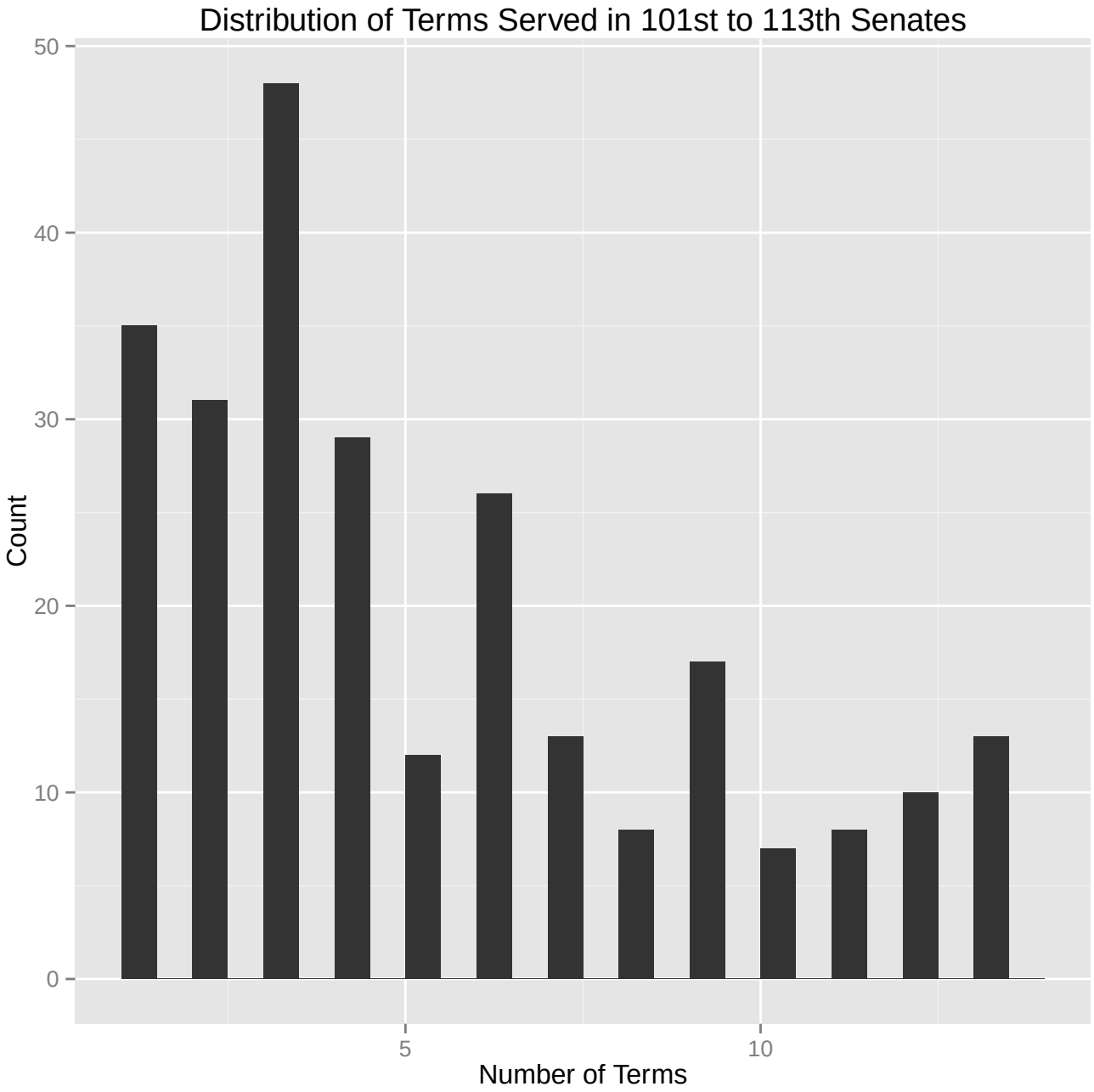
```
stateParty = queryDB(paste("SELECT state as stateAbrev, party, count(*) as ct",
    "FROM members", "WHERE party IN ('D', 'R') GROUP BY party, state"))
statePartyWide = dcast(stateParty, stateAbrev ~ party, value.var = "ct")
statePartyWide[is.na(statePartyWide)] = 0
statePartyWide$diff = statePartyWide$D - statePartyWide$R
statePartyWide$state = apply(statePartyWide, 1, FUN = function(x) {
    stateAbrevToFull(x["stateAbrev"])
})
stateMap = map_data("state")
ggplot(statePartyWide) + geom_map(data = stateMap, map = stateMap, aes(x = long,
    y = lat, map_id = region), fill = "#ffffff", color = "grey10") + geom_map(data = statePartyWide,
    map = stateMap, aes(fill = diff, map_id = state), color = "grey10") + scale_fill_gradient(name = "Party Differe
    low = "red", high = "navy blue", guide = "legend", limits = c(-30, 30),
    breaks = seq(-30, 30, length.out = 9)) + ggtitle(expression(atop("Party Preferences By State as Determined by S
    atop(italic("Data From 1989 to 2014"), "Republicans (Red) vs. Democrats (Blue)"))))
```



```
websiteCt = queryDB("SELECT party as Party, congressNumber, count(*) as webct FROM members WHERE URL!='' AND Party
twitterCt = queryDB("SELECT party as Party, congressNumber, count(*) as twitterct FROM members WHERE twitter_accou
fbCt = queryDB("SELECT party as Party, congressNumber, count(*) as fbct FROM members WHERE facebook_account!='' AN
totalCt = queryDB("SELECT party as Party, congressNumber, count(*) as totalct FROM members WHERE Party!='I' GROUP
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", "congressNumber"))
ggplot(mediaPlotData, aes(x = congressNumber, y = value, color = Party)) + geom_point() +
    geom_line() + facet_wrap(~variable) + xlim(100, 113) + xlab("\nCongress Number") +
    ylab("Proportion of Senators in Party\n") + ggtitle("Proportion of Senators Who Have Ever Used a Web Platform
```

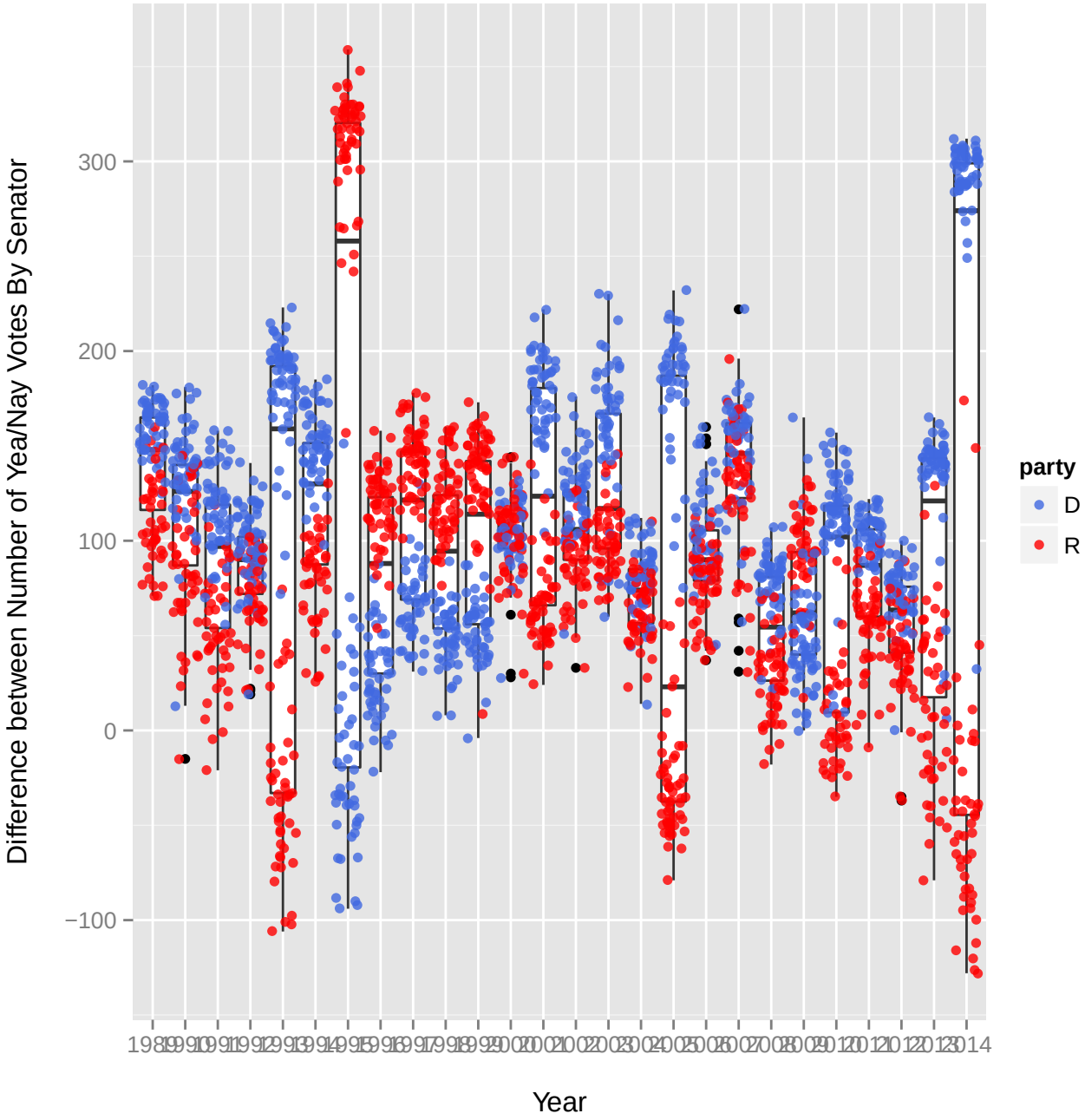


```
query = "SELECT id, first_name, last_name, party, seniority, count(*) AS ct FROM members GROUP BY id ORDER BY ct DE
senatorTotals = queryDB(query, "data.sqlite")
ggplot(senatorTotals, aes(x = ct)) + xlim(1, 14) + geom_bar(binwidth = 0.5) +
    xlab("Number of Terms") + ylab("Count") + ggtitle("Distribution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



```
  votedata = queryDB("SELECT id, party, vote, count(*) as ct, year FROM votes WHERE party in ('D', 'R') GROUP BY id
  votedatawide = dcast(votedata, id+year+party~vote, value.var="ct")
  votedatawide$diff = (votedatawide$Yea-votedatawide$Nay)
  ggplot(votedatawide, aes(x=year, y=diff))+geom_boxplot()+geom_jitter(aes(color=party), alpha=.8)+scale_color_manu
```

```
## Warning:  Removed 3 rows containing non-finite values (stat_boxplot).
## Warning:  Removed 3 rows containing missing values (geom_point).
```



```r
query = "SELECT party, yeas, nays, (yeas+nays) as total, (100*(yeas-nays)/(yeas+nays)) as voteDiff, congressNumber
rollCallStats = queryDB(query, "data.sqlite")
```

```
## Error in sqliteSendQuery(conn, statement):  error in statement:  no such column:  party
```

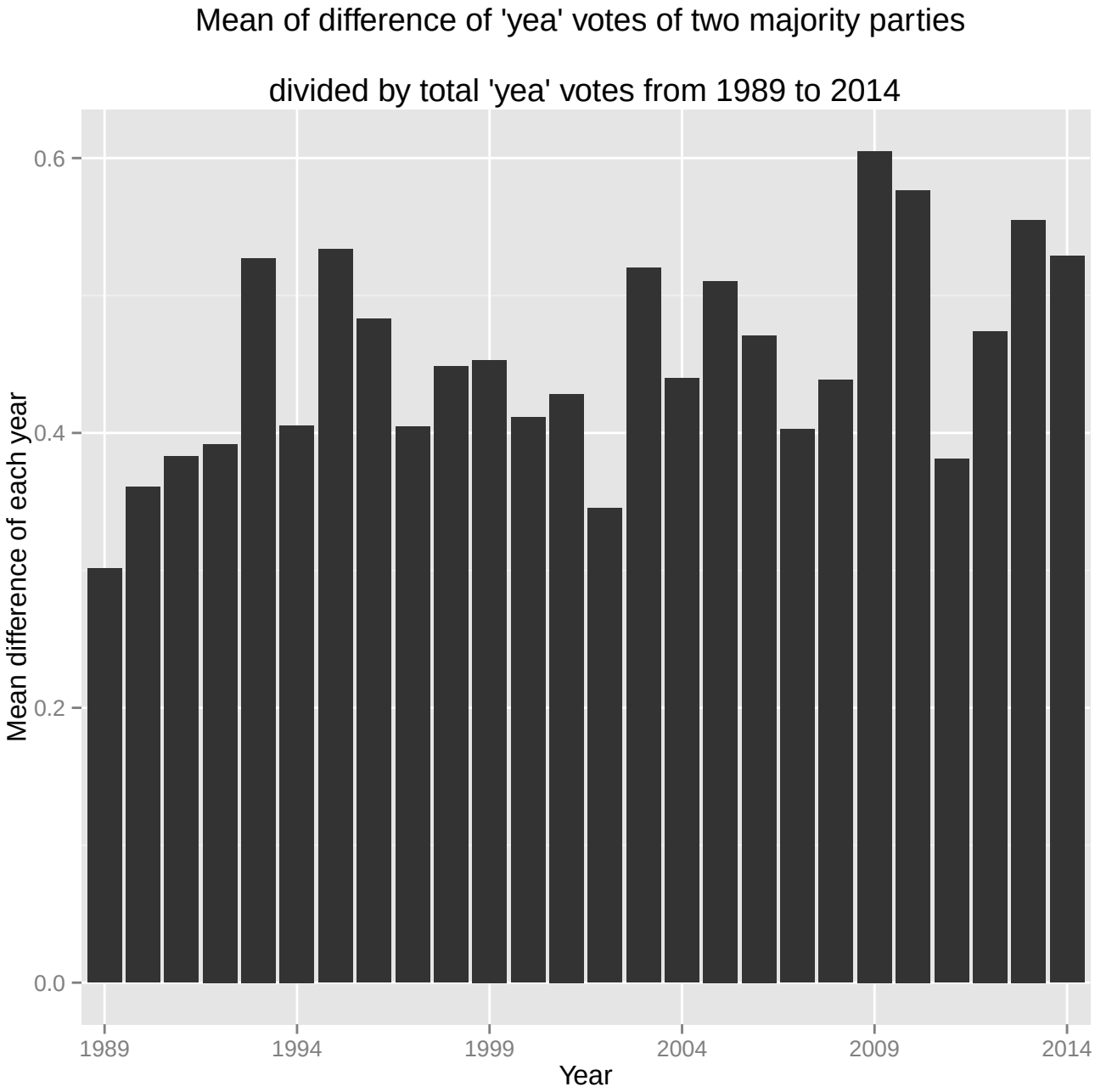
```r
rollCallStats$year = apply(rollCallStats, 1, function(x) {
    congressToYear(x["congressNumber"], x["sessionNumber"])
})
```

```
## Error in apply(rollCallStats, 1, function(x) {:  object 'rollCallStats' not found
```

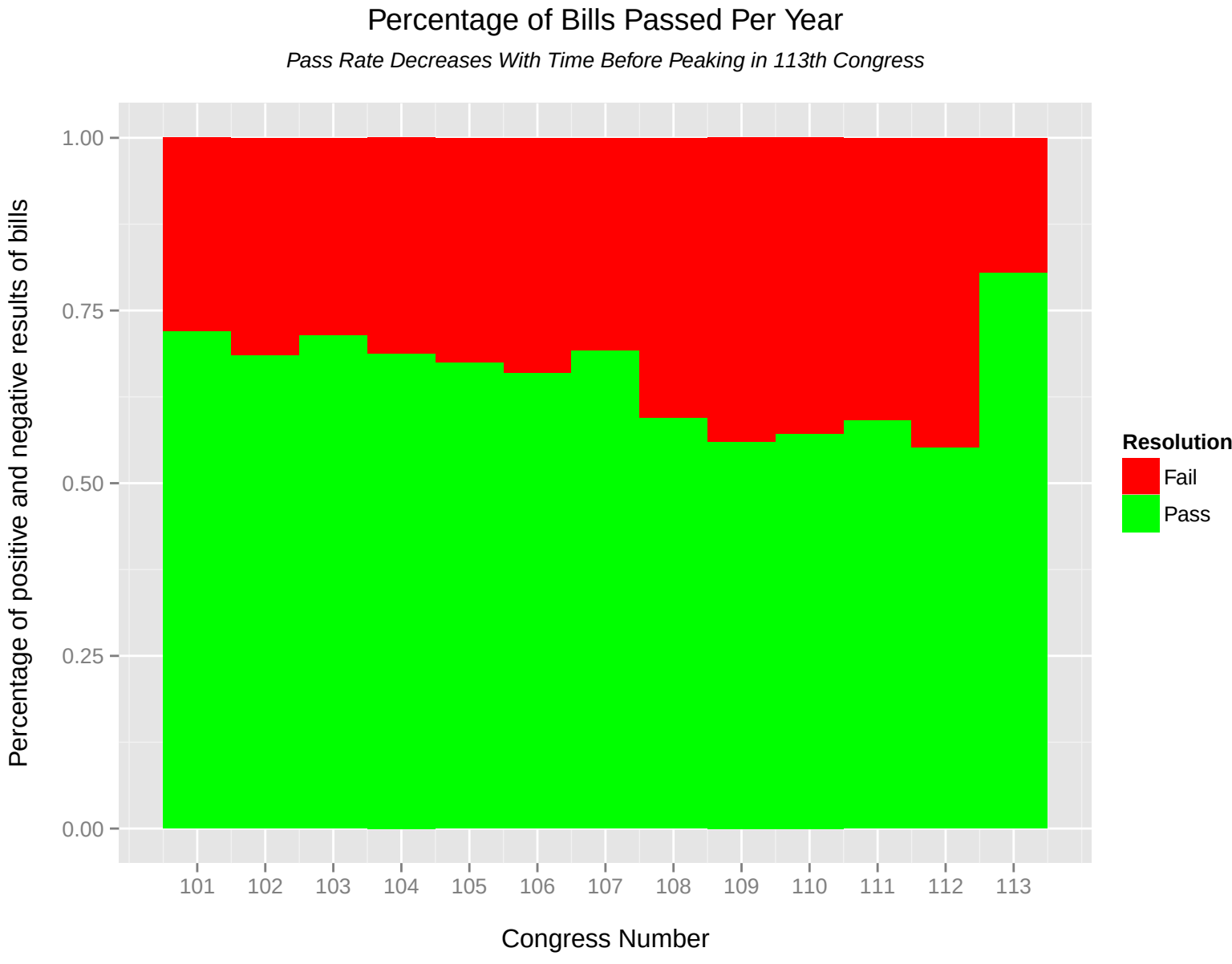
```r
ggplot(rollCallStats, aes(x = as.factor(year), y = voteDiff)) + geom_boxplot() +
    geom_smooth(aes(group = 1), method = "lm") + geom_point(alpha = 0.1) + ggtitle("Roll-call Vote Disagreement by
    xlab("\nYear") + ylab("Percentage Difference\n100*(Yea-Nay)/Total")
```

```
## Error in ggplot(rollCallStats, aes(x = as.factor(year), y = voteDiff)):  object 'rollCallStats' not found
```

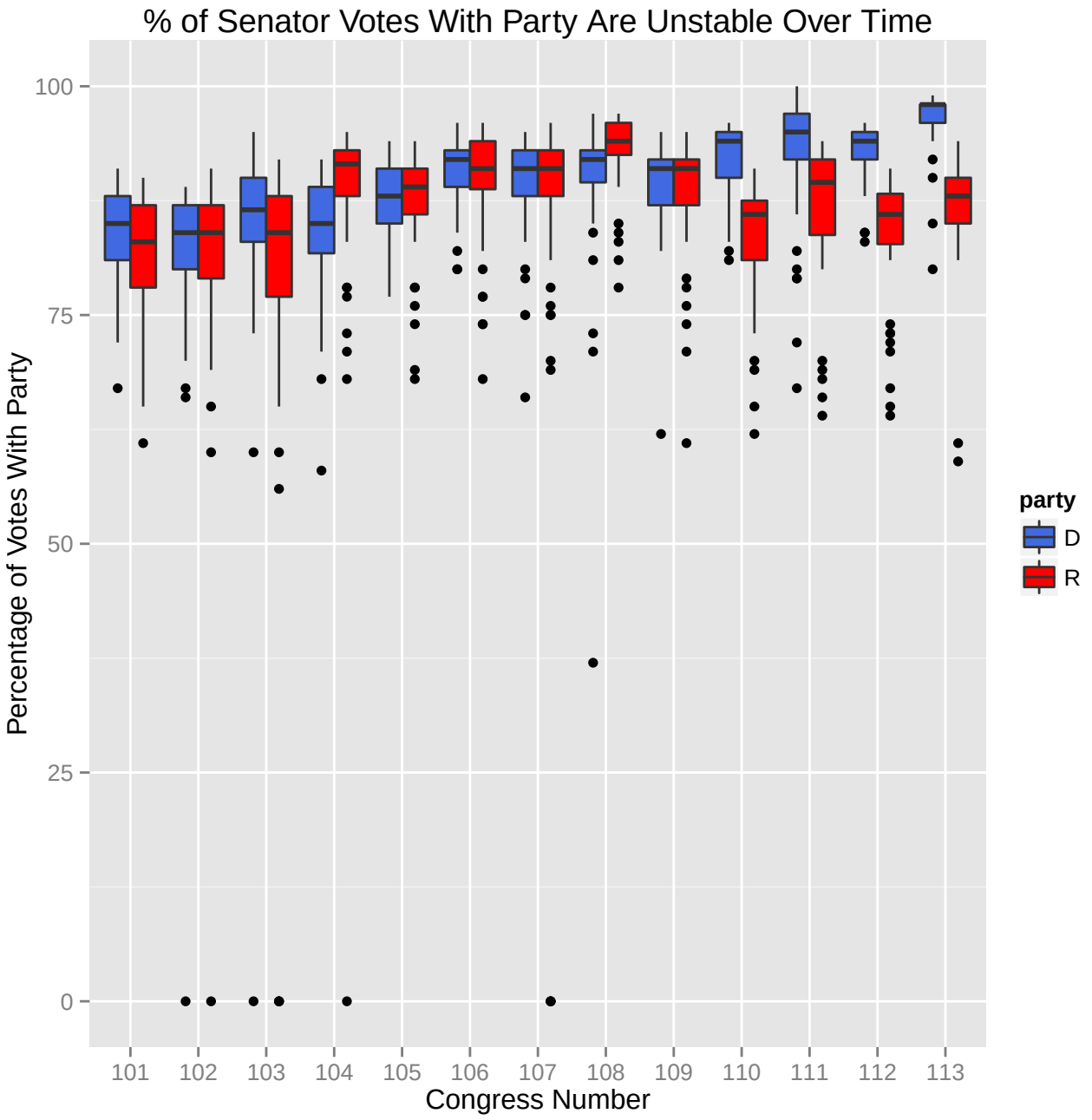
```
query = "select r.year as year, r.voteNumber as voteNumber,
            abs(r.c - d.c) * 1.0 / (r.c + d.c) as diff
         from (select voteNumber, year, count(*) as c
               from votes
               where vote == 'Yea' and party == 'R' group by year, voteNumber)
               as r
              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
0.0
0.2
0.4
0.6
1989
1994
1999
2004
2009
2014
Year

```
passedQuery <- "select 'Pass' 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 'Fail' 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 = "Resolution", values = c("red", "green")) +
  xlab("\nCongress Number") +
  scale_x_continuous(breaks=101:113) +
  ylab("Percentage of positive and negative results of bills\n") +
  ggtitle(expression(atop("Percentage of Bills Passed Per Year", atop(italic("Pass Rate Decreases With Time Befo
```

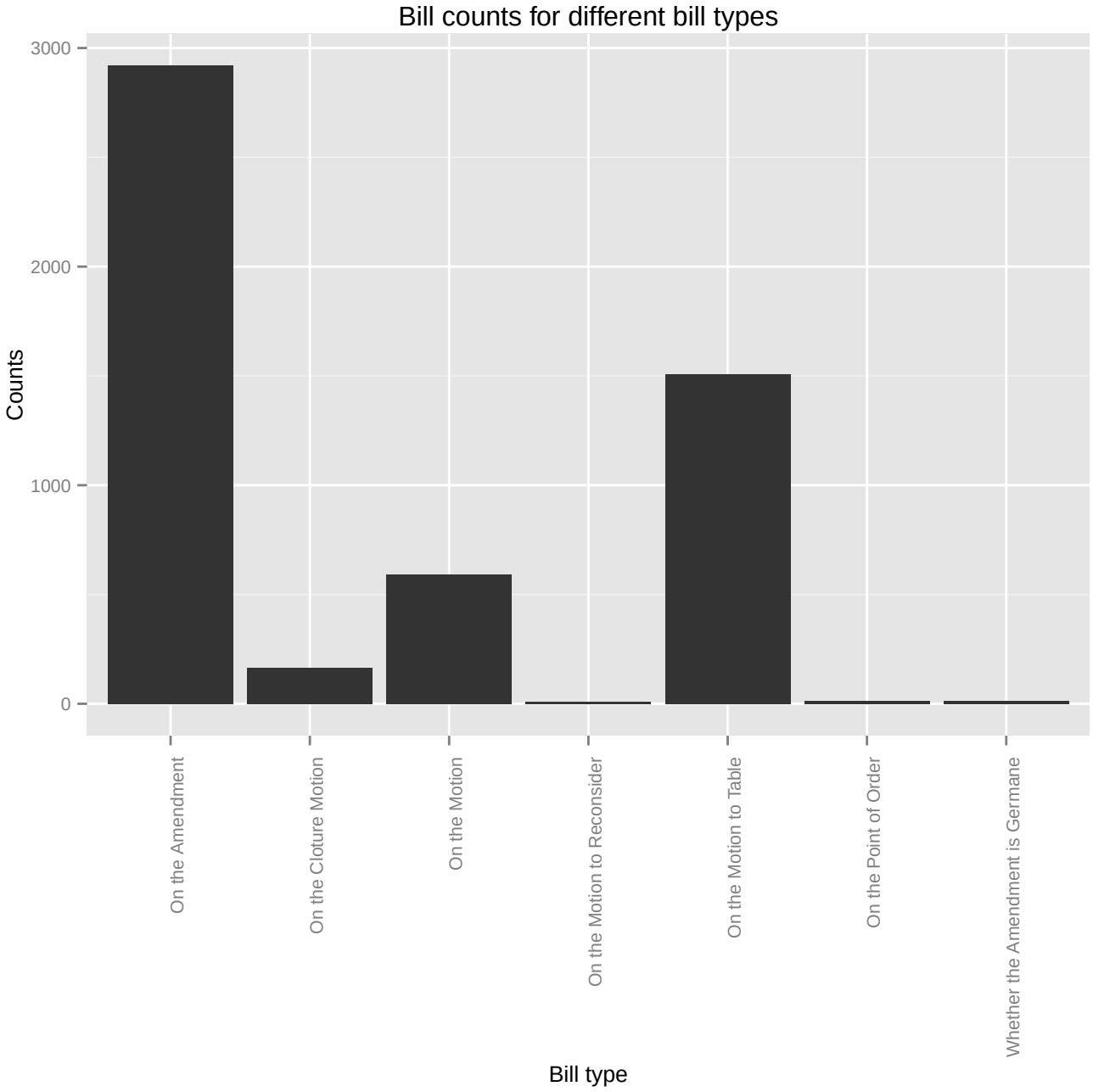


```
memberPct = queryDB("SELECT id, party, missed_votes_pct as missed, votes_with_party_pct as withParty, next_electi
memberPct$withParty = as.integer(as.character(memberPct$withParty))
memberPct$congressNumber = as.factor(memberPct$congressNumber)
ggplot(memberPct, aes(x=congressNumber, y=withParty, fill=party))+geom_boxplot()+xlab("Congress Number")+ylab("Pe
```

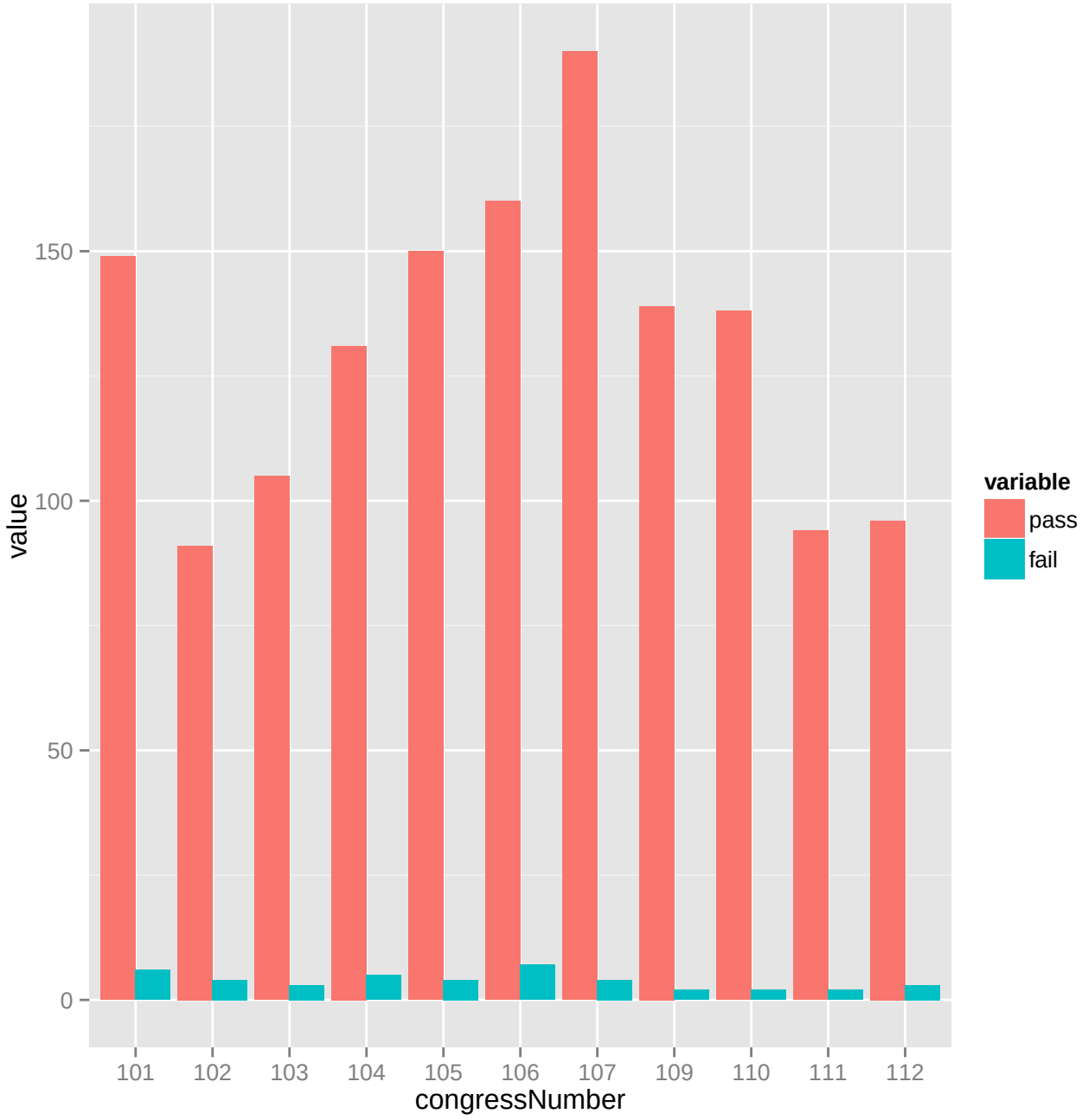



```
query <- "select * from
          (select questionType, count(*) as cnt
           from senateRollCalls
           where questionType not null
           group by questionType)
         where cnt > 10"
type <- queryDB(query, "data.sqlite")
ggplot(type) +
aes(x = questionType, y = cnt) +
labs(title="Bill counts for different bill types") +
xlab("Bill type") +
ylab("Counts") +
geom_bar(stat="identity") +
```

```
theme(text = element_text(size=10),
    axis.text.x = element_text(angle=90, hjust = 1))
```



```
library(reshape2)
query <- "select pass, fail, p.congressNumber
          from
          (select congressNumber, count(*) as pass
           from senateRollCalls
           where  CAST(nays as INTEGER) <= 5
           group by congressNumber) as p
          join
          (select congressNumber, count(*) as fail
           from senateRollCalls
           where  CAST(yeas as INTEGER) <= 5
           group by congressNumber) as f
          on p.congressNumber == f.congressNumber"
unanimous <- queryDB(query, "data.sqlite")
unanimous$congressNumber <- factor(unanimous$congressNumber)
unanimous <- melt(unanimous[,c('congressNumber','pass','fail')],id.vars = 1)
ggplot(unanimous,aes(x = congressNumber,y = value)) +
  geom_bar(stat='identity',aes(fill = variable),position = "dodge")
```



```
stateMap = map_data("state")
stateVote = queryDB(sprintf("SELECT state as stateAbrev, avg(votes_with_party_pct) as withParty FROM members GROU
stateVote$state = apply(stateVote, 1, FUN=function(x){stateAbrevToFull(x["stateAbrev"])})
ggplot()+geom_map(data=stateMap, map=stateMap, aes(x=long, y=lat, map_id=region), fill="#ffffff", color="grey10")
```