

Parliament Data Analysis

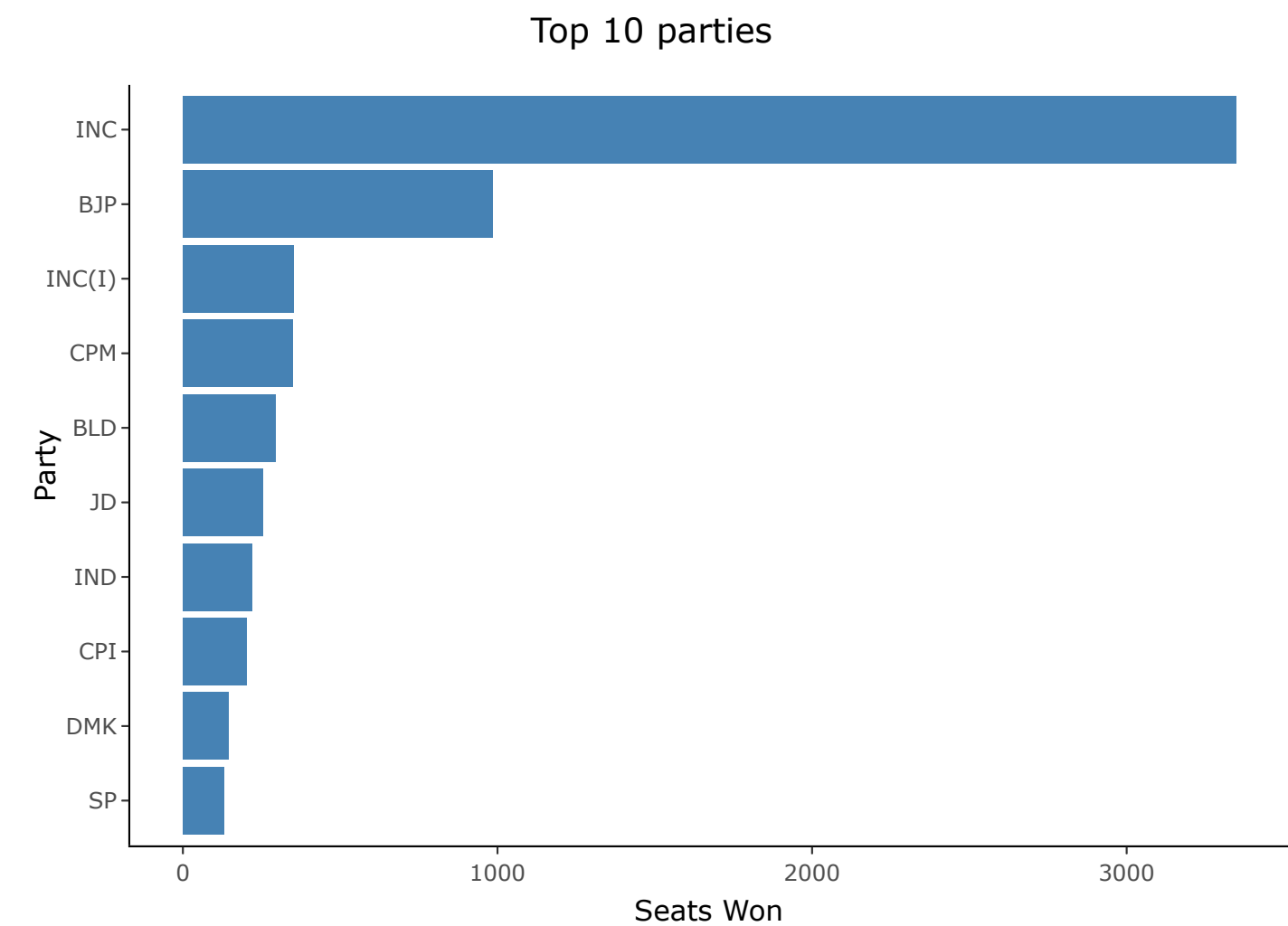
Ashok Kumar
7/22/2018

Top 10 Parties across the years

```
summary = parliamentdataset %>% group_by(PARTY) %>% filter(Position == 1) %>% summarise(SeatsWon = n()) %>% select
(PARTY, SeatsWon) %>% arrange_(-desc(SeatsWon))
top10 = head(summary, 10)

plot = ggplot(top10, aes(x= reorder(PARTY, SeatsWon), y= SeatsWon)) + geom_bar(stat = "identity", fill = "steelbl
ue") + coord_flip() + labs(y = "Seats Won", x = "Party", title = "Top 10 parties") + theme_classic()

ggplotly(plot)
```



Top 10 Parties and seats won - Stack bar

Note - Combined INC(I) and INC for the year 1980

```
summaryByYear = parliamentdataset %>% group_by(PARTY, YEAR) %>% filter(Position == 1) %>% summarise(seats_won = n(
)) %>% arrange(YEAR, -seats_won) %>% group_by(PARTY, YEAR)

filtered = summaryByYear %>% select(PARTY, YEAR, seats_won) %>% group_by(YEAR) %>% top_n(10)

## Selecting by seats_won

congress = filtered %>% filter(PARTY == "INC" | PARTY == "INC(I)")
congress$PARTY = "INC"

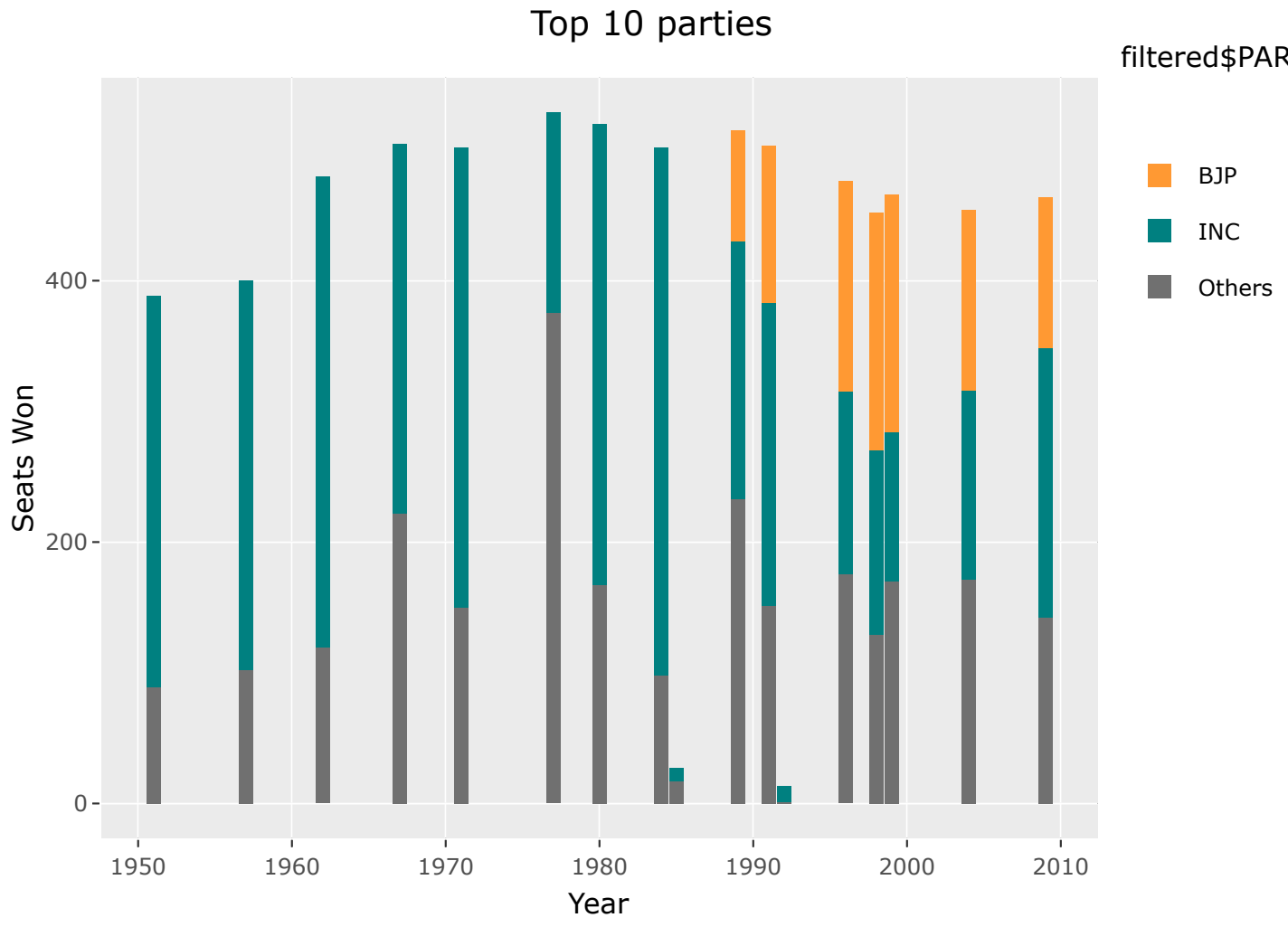
bjp = filtered %>% filter(PARTY == "BJP")

others = filtered %>% filter(PARTY != "INC", PARTY != "INC(I)", PARTY != "BJP") %>% group_by(YEAR) %>% summarise
(seats_won = sum(seats_won))
others$PARTY <- "Others"

filtered <- rbind.data.frame(congress, bjp, others)
filtered = filtered %>% group_by(YEAR)

plot1 = ggplot(filtered, aes(x= filtered$YEAR, y= filtered$seats_won, fill = filtered$PARTY)) + geom_bar(stat =
"identity", position = "stack") + labs(y = "Seats Won", x = "Year", title = "Top 10 parties") + scale_fill_manual
(values=c("#FF9933", "#008080", "#707070"))

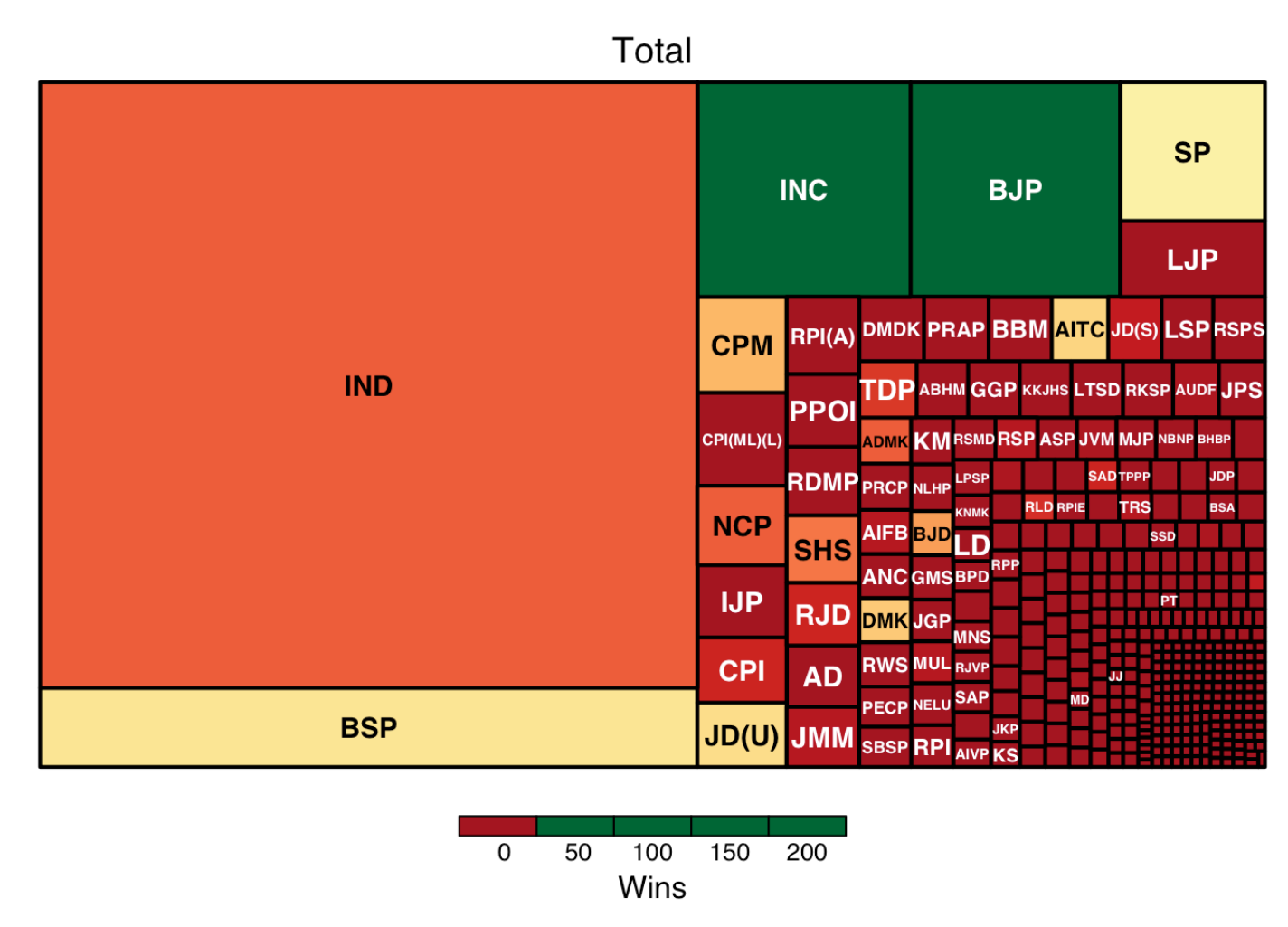
ggplotly(plot1)
```



Year 2009 - Seats contested and won - Tree Map

```
library(treemap)
year2009 = parliamentdataset %>% filter(YEAR == 2009) %>% group_by(PARTY) %>% summarise(Total = n(), Wins = sum(Po
sition == 1))

treemap(year2009,
  index=c('PARTY'),
  vSize=c('Total'),
  vColor=c('Wins'),
  types="value",
  palette = 'RdYlGn',
  mapping = c(0, 25, 50))
```



Top 5 Parties of all time and their 2004, 2009 - Win Percentage

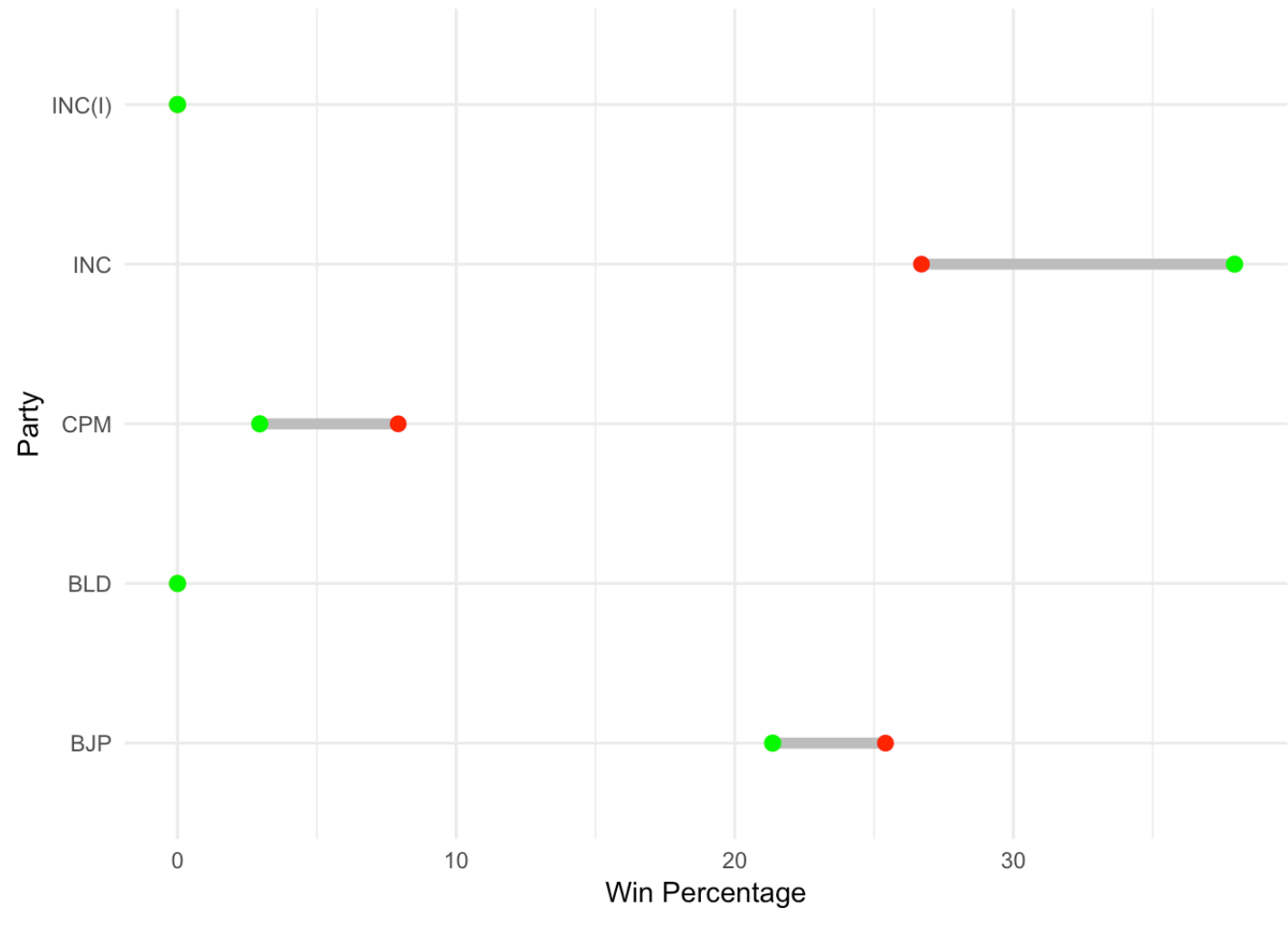
Note - Took All time top 5 parties and created the chart with 2004 and 2009 data. INC(I) and BLD are one of the all time top 5 parties, but they didn't contested in 2004 and 2009 elections

```
library(ggalt)

total_seats_2009 = parliamentdataset %>% filter(Position == 1, YEAR == 2009) %>% summarise(rows = n())
total_seats_2004 = parliamentdataset %>% filter(Position == 1, YEAR == 2004) %>% summarise(rows = n())

dumbbell = parliamentdataset %>% filter(Position == 1) %>% group_by(PARTY) %>% summarise(SeatsWon = n(), Year2009
= round((sum(YEAR == 2009)*100/total_seats_2009$rows), digits = 2), Year2004 = round((sum(YEAR == 2004)*100/total
_seats_2004$rows), digits = 2)) %>% arrange(-SeatsWon) %>% head(5)

ggplot(dumbbell, aes(x = Year2004, xend = Year2009, y = PARTY, group = PARTY)) + geom_dumbbell(color = "gray", si
ze = 2, colour_x = "red", colour_xend = "green") + labs(y = "Party", x = "Win Percentage") + theme_minimal()
```



Top 5 Parties of 2004 and 2009 - Win Percentage

Note - Took top 5 parties for the years 2004 and 2009 alone and created the chart

```
total_seats_2009 = parliamentdataset %>% filter(Position == 1, YEAR == 2009) %>% summarise(rows = n())
total_seats_2004 = parliamentdataset %>% filter(Position == 1, YEAR == 2004) %>% summarise(rows = n())

dumbbell = parliamentdataset %>% filter(Position == 1) %>% group_by(PARTY) %>% summarise(Year2009 = round((sum(YEA
R == 2009)*100/total_seats_2009$rows), digits = 2), Year2004 = round((sum(YEAR == 2004)*100/total_seats_2004$rows
), digits = 2)) %>% arrange(-Year2009, -Year2004) %>% head(5)

ggplot(dumbbell, aes(x = Year2004, xend = Year2009, y = PARTY, group = PARTY)) + geom_dumbbell(color = "gray", si
ze = 2, colour_x = "red", colour_xend = "green") + labs(y = "Party", x = "Win Percentage") + theme_minimal()
```

