fec-r-breakout-analysis.R

julian

Tue Jun 9 11:18:55 2015

```
#!/usr/bin/env Rscript

# FEC data analysis:
# 1. time series plotting and rough regression fits of data created using Python
# 2. Breakout Detection applications of cycle (spikes in cash flow)


# import files ------------------------------------------------------------

rm(list = ls())
options(digits = 10)

file.dir <- '~/GitHub/election-simulations/basic-exploration-python'

# all FEC data

fec.data.net.contributions <-
  read.csv(file.path(file.dir, 'fec-data-net-contributions.csv'))

fec.data.census.regions.net.contributions <-
  read.csv(file.path(file.dir, 'fec-data-census-regions-net-contributions.csv'))

fec.data.economic.regions.net.contributions <-
  read.csv(file.path(file.dir, 'fec-data-economic-regions-net-contributions.csv'))

# donations
donation.data <-
  read.csv(file.path(file.dir,'donation-data-candidate-idx.csv'))

donation.data.census.regions <-
  read.csv(file.path(file.dir, 'donation-data-census-regions-candidate-idx.csv'))

donation.data.economic.regions <-
  read.csv(file.path(file.dir, 'donation-data-economic-regions-candidate-idx.csv'))

# expenditures
expenditure.data <-
  read.csv(file.path(file.dir, 'expenditure-data-candidate-idx.csv'))

expenditure.data.census.regions <-
  read.csv(file.path(file.dir, 'expenditure-data-census-regions-candidate-idx.csv'))

expenditure.data.economic.regions <-
  read.csv(file.path(file.dir, 'expenditure-data-economic-regions-candidate-idx.csv'))

# remove X column
fec.data.net.contributions$X <- NULL
```

```
fec.data.census.regions.net.contributions$X <- NULL
fec.data.economic.regions.net.contributions$X <- NULL

donation.data$X <- NULL
donation.data.census.regions$X <- NULL
donation.data.economic.regions$X <- NULL

expenditure.data$X <- NULL
expenditure.data.census.regions$X <- NULL
expenditure.data.economic.regions$X <- NULL

# rename columns to something friendlier

fec.data.net.contributions <-
  dplyr::rename(fec.data.net.contributions, Candidate = cand_nm,
                Date = contb_receipt_dt_format, Dollars = contb_receipt_amt)

fec.data.census.regions.net.contributions <-
  dplyr::rename(fec.data.census.regions.net.contributions, Candidate = cand_nm,
                Date = contb_receipt_dt_format, Dollars = contb_receipt_amt,
                Region = census_region)

fec.data.economic.regions.net.contributions <-
  dplyr::rename(fec.data.economic.regions.net.contributions, Candidate = cand_nm,
                Date = contb_receipt_dt_format, Dollars = contb_receipt_amt,
                Region = economic_region)

donation.data <-
  dplyr::rename(donation.data, Candidate = cand_nm,
                Date = contb_receipt_dt_format, Dollars = contb_receipt_amt)

donation.data.census.regions <-
  dplyr::rename(donation.data.census.regions, Candidate = cand_nm,
                Date = contb_receipt_dt_format, Dollars = contb_receipt_amt,
                Region = census_region)

donation.data.economic.regions <-
  dplyr::rename(donation.data.economic.regions, Candidate = cand_nm,
                Date = contb_receipt_dt_format, Dollars = contb_receipt_amt,
                Region = economic_region)


expenditure.data <-
  dplyr::rename(expenditure.data, Candidate = cand_nm,
                Date = contb_receipt_dt_format, Dollars = contb_receipt_amt)

expenditure.data.census.regions <-
  dplyr::rename(expenditure.data.census.regions, Candidate = cand_nm,
                Date = contb_receipt_dt_format, Dollars = contb_receipt_amt,
                Region = census_region)

expenditure.data.economic.regions <-
  dplyr::rename(expenditure.data.economic.regions, Candidate = cand_nm,
```

```
              Date = contb_receipt_dt_format, Dollars = contb_receipt_amt,
              Region = economic_region)


# converting Date colums to actual dates


fec.data.net.contributions$Date <- as.Date(fec.data.net.contributions$Date)

fec.data.census.regions.net.contributions$Date <-
  as.Date(fec.data.census.regions.net.contributions$Date)

fec.data.economic.regions.net.contributions$Date <-
  as.Date(fec.data.economic.regions.net.contributions$Date)

donation.data$Date <- as.Date(donation.data$Date)

donation.data.census.regions$Date <- as.Date(donation.data.census.regions$Date)

donation.data.economic.regions$Date <-
  as.Date(donation.data.economic.regions$Date)

expenditure.data$Date <- as.Date(expenditure.data$Date)

expenditure.data.census.regions$Date <-
  as.Date(expenditure.data.census.regions$Date)

expenditure.data.economic.regions$Date <-
  as.Date(expenditure.data.economic.regions$Date)


# absolute value of expenditures
expenditure.data$absDollars <- abs(expenditure.data$Dollars)

expenditure.data.census.regions$absDollars <-
  abs(expenditure.data.census.regions$Dollars)

expenditure.data.economic.regions$absDollars <-
  abs(expenditure.data.economic.regions$Dollars)


# color palettes for plotting ---------------------------------------------

cbPalette <-
  c("#999999", "#E69F00", "#56B4E9",
    "#009E73", "#F0E442", "#0072B2", "#D55E00", "#CC79A7")

library(RColorBrewer)

qualEightPalette <- brewer.pal(8, 'Accent')

# time series plotting ----------------------------------------------------
```

```
library(ggplot2)
library(scales)
library(mgcv)
```

```
## Loading required package: nlme
## This is mgcv 1.8-6. For overview type 'help("mgcv-package")'.
```

```
source(file.path(file.dir, 'plot-facet-parameters.R'))


# important dates ------------------------------------------------------


first.debate <- '2012-10-03'
vp.debate <- '2012-10-11'
second.debate <- '2012-10-16'
third.debate <- '2012-10-22'
republican.convention.start <- '2012-08-27'
republican.convention.end <- '2012-08-30'
democratic.convention.start <- '2012-09-03'
democratic.convention.end <- '2012-09-06'



# net cash flow plots --------------------------------------------------


# plots with smoothers overlaid
# plot sequence:
# 1. net flow, donations, expenditures
# 2. broken down by convention season, then debates
# 3. entire series and then facetting by region

cash.flow.overall.plot <-
  ggplot(fec.data.net.contributions, aes(x = Date, y = Dollars)) +
  geom_line(size = 1, aes(colour = Candidate, linetype = Candidate),
            linesize = 2) +
  theme_bw() +
  labs(title = 'Net cash flow: July 31st to December 21st 2012',
       x = 'Date', y= ' Dollars') +
  theme(axis.text.x = element_text(angle = 30,hjust=1)) +
  scale_x_date(limits = as.Date(c('2012-07-31', '2012-12-31')),
               breaks = seq(as.Date('2012-07-31'),as.Date('2012-12-31'),
                            by ='2 weeks'),
               labels = date_format('%Y %b')) +
  large.bold.bottom.legend.facet +
  scale_colour_manual(values = c('#1f78b4', '#33a02c')) +
  scale_linetype_manual(values = c('solid', 'dashed'))

ggsave('overall-cash-flow.png', cash.flow.overall.plot,
       width = 9.2, height = 9.2, units = 'in', dpi = 600)

cash.flow.census.region.plot <-
```

```
  ggplot(fec.data.census.regions.net.contributions,
         aes(x = Date, y = Dollars)) +
  geom_line(size = 1, aes(colour = Candidate, linetype = Candidate),
            linesize = 2) +
  theme_bw() +
  labs(title = 'Net cash flow: July 31st to December 21st 2012',
       x = 'Date', y= ' Dollars') +
  theme(axis.text.x = element_text(angle = 30,hjust=1)) +
  scale_x_date(limits = as.Date(c('2012-07-31', '2012-12-31')),
               breaks = seq(as.Date('2012-07-31'),as.Date('2012-12-31'),
                            by ='2 weeks'),
               labels = date_format('%Y %b')) +
  large.bold.bottom.legend.facet +
  scale_colour_manual(values = c('#1f78b4', '#33a02c')) +
  scale_linetype_manual(values = c('solid', 'dashed')) +
  facet_grid(Region ~ .)

ggsave('cash-flow-census-region.png', cash.flow.census.region.plot,
       width = 9.2, height = 9.2, units = 'in', dpi = 600)

cash.flow.economic.region.plot <-
  ggplot(fec.data.economic.regions.net.contributions,
         aes(x = Date, y = Dollars)) +
  geom_line(size = 1, aes(colour = Candidate, linetype = Candidate),
            linesize = 2) +
  theme_bw() +
  labs(title = 'net cash flow: July 31st to December 21st 2012',
       x = 'Date', y = ' Dollars') +
  theme(axis.text.x = element_text(angle = 30,hjust=1)) +
  scale_x_date(limits = as.Date(c('2012-07-31', '2012-12-31')),
               breaks = seq(as.Date('2012-07-31'),as.Date('2012-12-31'),
                            by ='2 weeks'),
               labels = date_format('%Y %b')) +
  large.bold.bottom.legend.facet +
  scale_colour_manual(values = c('#1f78b4', '#33a02c')) +
  scale_linetype_manual(values = c('solid', 'dashed')) +
  facet_grid(Region ~ .)

ggsave('cash-flow-economic-region.png', cash.flow.economic.region.plot,
       width = 9.2, height = 9.2, units = 'in', dpi = 600)

cash.flow.convention.plot <-
  ggplot(fec.data.net.contributions, aes(x = Date, y = Dollars)) +
  geom_line(size = 1, aes(colour = Candidate, linetype = Candidate),
            linesize = 2) +
  theme_bw() +
  labs(title = 'Net cash flow: party conventions',
       x = 'Date', y= ' Dollars') +
  theme(axis.text.x = element_text(angle = 30,hjust=1)) +
  scale_x_date(limits = as.Date(c('2012-08-15', '2012-09-15')),
               breaks = seq(as.Date('2012-08-15'),as.Date('2012-09-15'),
                            by ='2 weeks'),
               labels = date_format('%Y %b')) +
```

```
  large.bold.bottom.legend.facet +
  scale_colour_manual(values = c('#1f78b4', '#33a02c')) +
  scale_linetype_manual(values = c('solid', 'dashed')) +
  coord_cartesian(ylim = c(0, 3e7)) +
  annotate('text', x = as.Date(republican.convention.end), y = 2.2e07,
           label = 'Republican convention', angle = 30, vjust = 1, hjust = 0.1) +
  annotate('text', x = as.Date(democratic.convention.end), y = 1.2e07,
           label = 'Democratic convention', angle = 30, vjust = 1, hjust = 0.1)

ggsave('cash-flow-party-convention.png', cash.flow.convention.plot,
       width = 9.2, height = 9.2, units = 'in', dpi = 600)

cash.flow.convention.census.plot <-
  ggplot(fec.data.census.regions.net.contributions, aes(x = Date, y = Dollars)) +
  geom_line(size = 1, aes(colour = Candidate, linetype = Candidate),
            linesize = 2) +
  theme_bw() +
  labs(title = 'Net cash flow:: party conventions',
       x = 'Date', y= ' Dollars') +
  theme(axis.text.x = element_text(angle = 30,hjust=1)) +
  scale_x_date(limits = as.Date(c('2012-08-15', '2012-09-15')),
               breaks = seq(as.Date('2012-08-15'),as.Date('2012-09-15'),
                            by ='2 weeks'),
               labels = date_format('%Y %b')) +
  large.bold.bottom.legend.facet +
  scale_colour_manual(values = c('#1f78b4', '#33a02c')) +
  scale_linetype_manual(values = c('solid', 'dashed')) +
  coord_cartesian(ylim = c(0, 3e7)) +
  annotate('text', x = as.Date(republican.convention.end), y = 1e07,
           label = 'RC', angle = 30, vjust = 1, hjust = 0.1) +
  annotate('text', x = as.Date(democratic.convention.end), y = 1e07,
           label = 'DC', angle = 30, vjust = 1, hjust = 0.1) +
  facet_grid(Region ~ .)

ggsave('cash-flow-party-convention-census.png', cash.flow.convention.census.plot,
       width = 9.2, height = 9.2, units = 'in', dpi = 600)

cash.flow.convention.economic.plot <-
  ggplot(fec.data.economic.regions.net.contributions,
         aes(x = Date, y = Dollars)) +
  geom_line(size = 1, aes(colour = Candidate, linetype = Candidate),
            linesize = 2) +
  theme_bw() +
  labs(title = 'Net cash flow: party conventions',
       x = 'Date', y= ' Dollars') +
  theme(axis.text.x = element_text(angle = 30,hjust=1)) +
  scale_x_date(limits = as.Date(c('2012-08-15', '2012-09-15')),
               breaks = seq(as.Date('2012-08-15'),as.Date('2012-09-15'),
                            by ='2 weeks'),
               labels = date_format('%Y %b')) +
  large.bold.bottom.legend.facet +
  scale_colour_manual(values = c('#1f78b4', '#33a02c')) +
  scale_linetype_manual(values = c('solid', 'dashed')) +
```

```
  coord_cartesian(ylim = c(0, 3e7)) +
  annotate('text', x = as.Date(republican.convention.end), y = 2.2e07,
           label = 'RC', angle = 30, vjust = 1, hjust = 0.1) +
  annotate('text', x = as.Date(democratic.convention.end), y = 1.2e07,
           label = 'DC', angle = 30, vjust = 1, hjust = 0.1) +
  facet_grid(Region ~ .)

ggsave('cash-flow-party-convention-economic.png',
       cash.flow.convention.economic.plot,
       width = 9.2, height = 9.2, units = 'in', dpi = 600)

cash.flow.debate.plot <-
  ggplot(fec.data.net.contributions, aes(x = Date, y = Dollars)) +
  geom_line(size = 1, aes(colour = Candidate, linetype = Candidate),
            linesize = 2) +
  theme_bw() +
  labs(title = 'Net cash flow: debate season',
       x = 'Date', y= ' Dollars') +
  theme(axis.text.x = element_text(angle = 30,hjust=1)) +
  scale_x_date(limits = as.Date(c('2012-10-01', '2012-10-31')),
               breaks = seq(as.Date('2012-10-01'),as.Date('2012-10-31'),
                            by ='5 days'),
               labels = date_format('%Y %b %d')) +
  large.bold.bottom.legend.facet +
  scale_colour_manual(values = c('#1f78b4', '#33a02c')) +
  scale_linetype_manual(values = c('solid', 'dashed')) +
  coord_cartesian(ylim = c(0, 3e7)) +
  annotate('text', x = as.Date(first.debate), y = 1e07,
           label = 'First debate', angle = 30, vjust = 1, hjust = 0.1) +
  annotate('text', x = as.Date(vp.debate), y = 1.1e07,
           label = 'VP debate', angle = 30, vjust = 1, hjust = 0.1) +
  annotate('text', x = as.Date(second.debate), y = 1.8e07,
           label = 'Second debate', angle = 30, vjust = 1, hjust = 0.1) +
  annotate('text', x = as.Date(third.debate), y = 0.9e07,
           label = 'Third debate', angle = 30, vjust = 1, hjust = 0.1)

ggsave('cash-flow-debates.png', cash.flow.debate.plot,
       width = 9.2, height = 9.2, units = 'in', dpi = 600)

cash.flow.debate.census.plot <-
  ggplot(fec.data.census.regions.net.contributions,
         aes(x = Date, y = Dollars)) +
  geom_line(size = 1, aes(colour = Candidate, linetype = Candidate),
            linesize = 2) +
  theme_bw() +
  labs(title = 'Net cash flow: debate season',
       x = 'Date', y= ' Dollars') +
  theme(axis.text.x = element_text(angle = 30,hjust=1)) +
  scale_x_date(limits = as.Date(c('2012-10-01', '2012-10-31')),
               breaks = seq(as.Date('2012-10-01'),as.Date('2012-10-31'),
                            by ='5 days'),
               labels = date_format('%Y %b %d')) +
  large.bold.bottom.legend.facet +
```

```
  scale_colour_manual(values = c('#1f78b4', '#33a02c')) +
  scale_linetype_manual(values = c('solid', 'dashed')) +
  coord_cartesian(ylim = c(0, 3e7)) +
  annotate('text', x = as.Date(first.debate), y = 1e07,
           label = 'First debate', angle = 30, vjust = 1, hjust = 0.1) +
  annotate('text', x = as.Date(vp.debate), y = 1.1e07,
           label = 'VP debate', angle = 30, vjust = 1, hjust = 0.1) +
  annotate('text', x = as.Date(second.debate), y = 1.8e07,
           label = 'Second debate', angle = 30, vjust = 1, hjust = 0.1) +
  annotate('text', x = as.Date(third.debate), y = 0.9e07,
           label = 'Third debate', angle = 30, vjust = 1, hjust = 0.1) +
  facet_grid(Region ~ .)

ggsave('cash-flow-debates-census.png', cash.flow.debate.census.plot,
       width = 9.2, height = 9.2, units = 'in', dpi = 600)

cash.flow.debate.economic.plot <-
  ggplot(fec.data.economic.regions.net.contributions,
         aes(x = Date, y = Dollars)) +
  geom_line(size = 1, aes(colour = Candidate, linetype = Candidate),
            linesize = 2) +
  theme_bw() +
  labs(title = 'Net cash flow: debate season',
       x = 'Date', y= ' Dollars') +
  theme(axis.text.x = element_text(angle = 30,hjust=1)) +
  scale_x_date(limits = as.Date(c('2012-10-01', '2012-10-31')),
               breaks = seq(as.Date('2012-10-01'),as.Date('2012-10-31'),
                            by ='5 days'),
               labels = date_format('%Y %b %d')) +
  large.bold.bottom.legend.facet +
  scale_colour_manual(values = c('#1f78b4', '#33a02c')) +
  scale_linetype_manual(values = c('solid', 'dashed')) +
  coord_cartesian(ylim = c(0, 3e7)) +
  annotate('text', x = as.Date(first.debate), y = 1e07,
           label = 'First debate', angle = 30, vjust = 1, hjust = 0.1) +
  annotate('text', x = as.Date(vp.debate), y = 1.1e07,
           label = 'VP debate', angle = 30, vjust = 1, hjust = 0.1) +
  annotate('text', x = as.Date(second.debate), y = 1.8e07,
           label = 'Second debate', angle = 30, vjust = 1, hjust = 0.1) +
  annotate('text', x = as.Date(third.debate), y = 0.9e07,
           label = 'Third debate', angle = 30, vjust = 1, hjust = 0.1) +
  facet_grid(Region ~ .)

ggsave('cash-flow-debates-economic.png',
       cash.flow.debate.economic.plot,
       width = 9.2, height = 9.2, units = 'in', dpi = 600)

# donation plots ------------------------------------------------------------

donation.overall.plot <-
  ggplot(donation.data, aes(x = Date, y = Dollars)) +
  geom_line(size = 1, aes(colour = Candidate, linetype = Candidate),
            linesize = 2) +
```

```
  theme_bw() +
  labs(title = 'Donation amounts: July 31st to December 21st 2012',
       x = 'Date', y= ' Dollars') +
  theme(axis.text.x = element_text(angle = 30,hjust=1)) +
  scale_x_date(limits = as.Date(c('2012-07-31', '2012-12-31')),
               breaks = seq(as.Date('2012-07-31'),as.Date('2012-12-31'),
                            by ='2 weeks'),
               labels = date_format('%Y %b')) +
  large.bold.bottom.legend.facet +
  scale_colour_manual(values = c('#1f78b4', '#33a02c')) +
  scale_linetype_manual(values = c('solid', 'dashed'))

ggsave('donations-overall-annotated.png', donation.overall.plot,
       width = 9.2, height = 9.2, units = 'in', dpi = 600)

donation.census.region.plot <-
  ggplot(donation.data.census.regions, aes(x = Date, y = Dollars)) +
  geom_line(size = 1, aes(colour = Candidate, linetype = Candidate),
            linesize = 2) +
  theme_bw() +
  labs(title = 'Donation amounts: July 31st to December 21st 2012',
       x = 'Date', y= ' Dollars') +
  theme(axis.text.x = element_text(angle = 30,hjust=1)) +
  scale_x_date(limits = as.Date(c('2012-07-31', '2012-12-31')),
               breaks = seq(as.Date('2012-07-31'),as.Date('2012-12-31'),
                            by ='2 weeks'),
               labels = date_format('%Y %b')) +
  large.bold.bottom.legend.facet +
  scale_colour_manual(values = c('#1f78b4', '#33a02c')) +
  scale_linetype_manual(values = c('solid', 'dashed')) +
  facet_grid(Region ~ .)

ggsave('donations-overall-census-annotated.png',
       donation.census.region.plot,
       width = 9.2, height = 9.2, units = 'in', dpi = 600)

donation.economic.region.plot <-
  ggplot(donation.data.economic.regions, aes(x = Date, y = Dollars)) +
  geom_line(size = 1, aes(colour = Candidate, linetype = Candidate),
            linesize = 2) +
  theme_bw() +
  labs(title = 'Donation amounts: July 31st to December 21st 2012',
       x = 'Date', y = ' Dollars') +
  theme(axis.text.x = element_text(angle = 30,hjust=1)) +
  scale_x_date(limits = as.Date(c('2012-07-31', '2012-12-31')),
               breaks = seq(as.Date('2012-07-31'),as.Date('2012-12-31'),
                            by ='2 weeks'),
               labels = date_format('%Y %b')) +
  large.bold.bottom.legend.facet +
  scale_colour_manual(values = c('#1f78b4', '#33a02c')) +
  scale_linetype_manual(values = c('solid', 'dashed')) +
  facet_grid(Region ~ .)
```

```
ggsave('donations-overall-economic-annotated.png',
       donation.economic.region.plot,
       width = 9.2, height = 9.2, units = 'in', dpi = 600)

donation.convention.plot <-
  ggplot(donation.data, aes(x = Date, y = Dollars)) +
  geom_line(size = 1, aes(colour = Candidate, linetype = Candidate),
            linesize = 2) +
  theme_bw() +
  labs(title = 'Donation amounts: party conventions',
       x = 'Date', y= ' Dollars') +
  theme(axis.text.x = element_text(angle = 30,hjust=1)) +
  scale_x_date(limits = as.Date(c('2012-08-15', '2012-09-15')),
               breaks = seq(as.Date('2012-08-15'),as.Date('2012-09-15'),
                            by ='2 weeks'),
               labels = date_format('%Y %b')) +
  large.bold.bottom.legend.facet +
  scale_colour_manual(values = c('#1f78b4', '#33a02c')) +
  scale_linetype_manual(values = c('solid', 'dashed')) +
  coord_cartesian(ylim = c(0, 3e7)) +
  annotate('text', x = as.Date(republican.convention.end), y = 2.2e07,
           label = 'Republican convention', angle = 30, vjust = 1, hjust = 0.1) +
  annotate('text', x = as.Date(democratic.convention.end), y = 1.2e07,
           label = 'Democratic convention', angle = 30, vjust = 1, hjust = 0.1)

ggsave('donations-convention-annotated.png',
       donation.convention.plot,
       width = 9.2, height = 9.2, units = 'in', dpi = 600)

donation.convention.census.plot <-
  ggplot(donation.data.census.regions, aes(x = Date, y = Dollars)) +
  geom_line(size = 1, aes(colour = Candidate, linetype = Candidate),
            linesize = 2) +
  theme_bw() +
  labs(title = 'Donation amounts: party conventions',
       x = 'Date', y= ' Dollars') +
  theme(axis.text.x = element_text(angle = 30,hjust=1)) +
  scale_x_date(limits = as.Date(c('2012-08-15', '2012-09-15')),
               breaks = seq(as.Date('2012-08-15'),as.Date('2012-09-15'),
                            by ='2 weeks'),
               labels = date_format('%Y %b')) +
  large.bold.bottom.legend.facet +
  scale_colour_manual(values = c('#1f78b4', '#33a02c')) +
  scale_linetype_manual(values = c('solid', 'dashed')) +
  coord_cartesian(ylim = c(0, 3e7)) +
  annotate('text', x = as.Date(republican.convention.end), y = 1e07,
           label = 'RC', angle = 30, vjust = 1, hjust = 0.1) +
  annotate('text', x = as.Date(democratic.convention.end), y = 1e07,
           label = 'DC', angle = 30, vjust = 1, hjust = 0.1) +
  facet_grid(Region ~ .)

ggsave('donations-convention-census-annotated.png',
       donation.convention.census.plot,
```

```
       width = 9.2, height = 9.2, units = 'in', dpi = 600)

donation.convention.economic.plot <-
  ggplot(donation.data.economic.regions, aes(x = Date, y = Dollars)) +
  geom_line(size = 1, aes(colour = Candidate, linetype = Candidate),
            linesize = 2) +
  theme_bw() +
  labs(title = 'Donation amounts: party conventions',
       x = 'Date', y= ' Dollars') +
  theme(axis.text.x = element_text(angle = 30,hjust=1)) +
  scale_x_date(limits = as.Date(c('2012-08-15', '2012-09-15')),
               breaks = seq(as.Date('2012-08-15'),as.Date('2012-09-15'),
                            by ='2 weeks'),
               labels = date_format('%Y %b')) +
  large.bold.bottom.legend.facet +
  scale_colour_manual(values = c('#1f78b4', '#33a02c')) +
  scale_linetype_manual(values = c('solid', 'dashed')) +
  coord_cartesian(ylim = c(0, 3e7)) +
  annotate('text', x = as.Date(republican.convention.end), y = 2.2e07,
           label = 'RC', angle = 30, vjust = 1, hjust = 0.1) +
  annotate('text', x = as.Date(democratic.convention.end), y = 1.2e07,
           label = 'DC', angle = 30, vjust = 1, hjust = 0.1) +
  facet_grid(Region ~ .)

ggsave('donations-convention-economic-annotated.png',
       donation.convention.economic.plot,
       width = 9.2, height = 9.2, units = 'in', dpi = 600)

donation.debate.plot <-
  ggplot(donation.data, aes(x = Date, y = Dollars)) +
  geom_line(size = 1, aes(colour = Candidate, linetype = Candidate),
            linesize = 2) +
  theme_bw() +
  labs(title = 'Donation amounts: debate season',
       x = 'Date', y= ' Dollars') +
  theme(axis.text.x = element_text(angle = 30,hjust=1)) +
  scale_x_date(limits = as.Date(c('2012-10-01', '2012-10-31')),
               breaks = seq(as.Date('2012-10-01'),as.Date('2012-10-31'),
                            by ='5 days'),
               labels = date_format('%Y %b %d')) +
  large.bold.bottom.legend.facet +
  scale_colour_manual(values = c('#1f78b4', '#33a02c')) +
  scale_linetype_manual(values = c('solid', 'dashed')) +
  coord_cartesian(ylim = c(0, 3e7)) +
  annotate('text', x = as.Date(first.debate), y = 1e07,
           label = 'First debate', angle = 30, vjust = 1, hjust = 0.1) +
  annotate('text', x = as.Date(vp.debate), y = 1.1e07,
           label = 'VP debate', angle = 30, vjust = 1, hjust = 0.1) +
  annotate('text', x = as.Date(second.debate), y = 1.8e07,
           label = 'Second debate', angle = 30, vjust = 1, hjust = 0.1) +
  annotate('text', x = as.Date(third.debate), y = 0.9e07,
           label = 'Third debate', angle = 30, vjust = 1, hjust = 0.1)
```

```
ggsave('donations-debate-annotated.png',
       donation.debate.plot,
       width = 9.2, height = 9.2, units = 'in', dpi = 600)

donation.debate.census.plot <-
  ggplot(donation.data.census.regions, aes(x = Date, y = Dollars)) +
  geom_line(size = 1, aes(colour = Candidate, linetype = Candidate),
            linesize = 2) +
  theme_bw() +
  labs(title = 'Donation amounts: debate season',
       x = 'Date', y= ' Dollars') +
  theme(axis.text.x = element_text(angle = 30,hjust=1)) +
  scale_x_date(limits = as.Date(c('2012-10-01', '2012-10-31')),
               breaks = seq(as.Date('2012-10-01'),as.Date('2012-10-31'),
                            by ='5 days'),
               labels = date_format('%Y %b %d')) +
  large.bold.bottom.legend.facet +
  scale_colour_manual(values = c('#1f78b4', '#33a02c')) +
  scale_linetype_manual(values = c('solid', 'dashed')) +
  coord_cartesian(ylim = c(0, 3e7)) +
  annotate('text', x = as.Date(first.debate), y = 1e07,
           label = 'First debate', angle = 30, vjust = 1, hjust = 0.1) +
  annotate('text', x = as.Date(vp.debate), y = 1.1e07,
           label = 'VP debate', angle = 30, vjust = 1, hjust = 0.1) +
  annotate('text', x = as.Date(second.debate), y = 1.8e07,
           label = 'Second debate', angle = 30, vjust = 1, hjust = 0.1) +
  annotate('text', x = as.Date(third.debate), y = 0.9e07,
           label = 'Third debate', angle = 30, vjust = 1, hjust = 0.1) +
  facet_grid(Region ~ .)

ggsave('donations-debate-census-annotated.png',
       donation.debate.census.plot,
       width = 9.2, height = 9.2, units = 'in', dpi = 600)

donation.debate.economic.plot <-
  ggplot(donation.data.census.regions, aes(x = Date, y = Dollars)) +
  geom_line(size = 1, aes(colour = Candidate, linetype = Candidate),
            linesize = 2) +
  theme_bw() +
  labs(title = 'Donation amounts: debate season',
       x = 'Date', y= ' Dollars') +
  theme(axis.text.x = element_text(angle = 30,hjust=1)) +
  scale_x_date(limits = as.Date(c('2012-10-01', '2012-10-31')),
               breaks = seq(as.Date('2012-10-01'),as.Date('2012-10-31'),
                            by ='5 days'),
               labels = date_format('%Y %b %d')) +
  large.bold.bottom.legend.facet +
  scale_colour_manual(values = c('#1f78b4', '#33a02c')) +
  scale_linetype_manual(values = c('solid', 'dashed')) +
  coord_cartesian(ylim = c(0, 3e7)) +
  annotate('text', x = as.Date(first.debate), y = 1e07,
           label = 'First debate', angle = 30, vjust = 1, hjust = 0.1) +
  annotate('text', x = as.Date(vp.debate), y = 1.1e07,
```

```
          label = 'VP debate', angle = 30, vjust = 1, hjust = 0.1) +
  annotate('text', x = as.Date(second.debate), y = 1.8e07,
           label = 'Second debate', angle = 30, vjust = 1, hjust = 0.1) +
  annotate('text', x = as.Date(third.debate), y = 0.9e07,
           label = 'Third debate', angle = 30, vjust = 1, hjust = 0.1) +
  facet_grid(Region ~ .)

ggsave('donations-debate-economic-annotated.png',
       donation.debate.economic.plot,
       width = 9.2, height = 9.2, units = 'in', dpi = 600)


# expenditure plots ----------------------------------------------------

expenditure.overall.plot <-
  ggplot(expenditure.data, aes(x = Date, y = absDollars)) +
  geom_line(size = 1, aes(colour = Candidate, linetype = Candidate),
            linesize = 2) +
  theme_bw() +
  labs(title = 'Expenditure amounts: July 31st to December 21st 2012',
       x = 'Date', y= ' Dollars') +
  theme(axis.text.x = element_text(angle = 30,hjust=1)) +
  scale_x_date(limits = as.Date(c('2012-07-31', '2012-12-31')),
               breaks = seq(as.Date('2012-07-31'),as.Date('2012-12-31'),
                            by ='2 weeks'),
               labels = date_format('%Y %b')) +
  large.bold.bottom.legend.facet +
  scale_colour_manual(values = c('#1f78b4', '#33a02c')) +
  scale_linetype_manual(values = c('solid', 'dashed'))

ggsave('expenditures-overall-annotated.png',
       expenditure.overall.plot,
       width = 9.2, height = 9.2, units = 'in', dpi = 600)

expenditure.census.region.plot <-
  ggplot(expenditure.data.census.regions, aes(x = Date, y = absDollars)) +
  geom_line(size = 1, aes(colour = Candidate, linetype = Candidate),
            linesize = 2) +
  theme_bw() +
  labs(title = 'Expenditure amounts: July 31st to December 21st 2012',
       x = 'Date', y= ' Dollars') +
  theme(axis.text.x = element_text(angle = 30,hjust=1)) +
  scale_x_date(limits = as.Date(c('2012-07-31', '2012-12-31')),
               breaks = seq(as.Date('2012-07-31'),as.Date('2012-12-31'),
                            by ='2 weeks'),
               labels = date_format('%Y %b')) +
  large.bold.bottom.legend.facet +
  scale_colour_manual(values = c('#1f78b4', '#33a02c')) +
  scale_linetype_manual(values = c('solid', 'dashed')) +
  facet_grid(Region ~ .)

ggsave('expenditures-census-annotated.png',
       expenditure.census.region.plot,
```

```
       width = 9.2, height = 9.2, units = 'in', dpi = 600)

expenditure.economic.region.plot <-
  ggplot(expenditure.data.economic.regions, aes(x = Date, y = absDollars)) +
  geom_line(size = 1, aes(colour = Candidate, linetype = Candidate),
            linesize = 2) +
  theme_bw() +
  labs(title = 'Expenditure amounts: July 31st to December 21st 2012',
       x = 'Date', y = ' Dollars') +
  theme(axis.text.x = element_text(angle = 30,hjust=1)) +
  scale_x_date(limits = as.Date(c('2012-07-31', '2012-12-31')),
               breaks = seq(as.Date('2012-07-31'),as.Date('2012-12-31'),
                            by ='2 weeks'),
               labels = date_format('%Y %b')) +
  large.bold.bottom.legend.facet +
  scale_colour_manual(values = c('#1f78b4', '#33a02c')) +
  scale_linetype_manual(values = c('solid', 'dashed')) +
  facet_grid(Region ~ .)

ggsave('expenditures-economic-annotated.png',
       expenditure.economic.region.plot,
       width = 9.2, height = 9.2, units = 'in', dpi = 600)

expenditure.stretch.plot <-
  ggplot(expenditure.data, aes(x = Date, y = absDollars)) +
  geom_line(size = 1, aes(colour = Candidate, linetype = Candidate),
            linesize = 2) +
  theme_bw() +
  labs(title = 'Expenditure amounts: Sept to Nov',
       x = 'Date', y= ' Dollars') +
  theme(axis.text.x = element_text(angle = 30,hjust=1)) +
  scale_x_date(limits = as.Date(c('2012-09-01', '2012-11-15')),
               breaks = seq(as.Date('2012-09-01'),as.Date('2012-11-15'),
                            by ='2 weeks'),
               labels = date_format('%Y %b')) +
  large.bold.bottom.legend.facet +
  scale_colour_manual(values = c('#1f78b4', '#33a02c')) +
  scale_linetype_manual(values = c('solid', 'dashed')) +
  coord_cartesian(ylim = c(0, 5e5)) +
  annotate('text', x = as.Date('2012-11-06'), y = 4.5e5,
           label = 'Election', vjust = 1, hjust = 0.1)

ggsave('expenditures-stretch-run-annotated.png',
       expenditure.stretch.plot,
       width = 9.2, height = 9.2, units = 'in', dpi = 600)

expenditure.stretch.census.plot <-
  ggplot(expenditure.data.census.regions, aes(x = Date, y = absDollars)) +
  geom_line(size = 1, aes(colour = Candidate, linetype = Candidate),
            linesize = 2) +
  theme_bw() +
  labs(title = 'Expenditure amounts: Sept to Nov',
       x = 'Date', y= ' Dollars') +
```

```
  theme(axis.text.x = element_text(angle = 30,hjust=1)) +
  scale_x_date(limits = as.Date(c('2012-09-01', '2012-11-15')),
               breaks = seq(as.Date('2012-09-01'),as.Date('2012-11-15'),
                            by ='2 weeks'),
               labels = date_format('%Y %b')) +
  large.bold.bottom.legend.facet +
  scale_colour_manual(values = c('#1f78b4', '#33a02c')) +
  scale_linetype_manual(values = c('solid', 'dashed')) +
  coord_cartesian(ylim = c(0, 2e5)) +
  annotate('text', x = as.Date('2012-11-06'), y = 1.2e5,
           label = 'Election', vjust = 1, hjust = 0.1) +
  facet_grid(Region ~ .)

ggsave('expenditures-stretch-run-census-annotated.png',
       expenditure.stretch.census.plot,
       width = 9.2, height = 9.2, units = 'in', dpi = 600)


expenditure.stretch.economic.plot <-
  ggplot(expenditure.data.economic.regions, aes(x = Date, y = absDollars)) +
  geom_line(size = 1, aes(colour = Candidate, linetype = Candidate),
            linesize = 2) +
  theme_bw() +
  labs(title = 'Expenditure amounts: Sept to Nov',
       x = 'Date', y= ' Dollars') +
  theme(axis.text.x = element_text(angle = 30,hjust=1)) +
  scale_x_date(limits = as.Date(c('2012-09-01', '2012-11-15')),
               breaks = seq(as.Date('2012-09-01'),as.Date('2012-11-15'),
                            by ='2 weeks'),
               labels = date_format('%Y %b')) +
  large.bold.bottom.legend.facet +
  scale_colour_manual(values = c('#1f78b4', '#33a02c')) +
  scale_linetype_manual(values = c('solid', 'dashed')) +
  coord_cartesian(ylim = c(0, 2e5)) +
  annotate('text', x = as.Date('2012-11-06'), y = 1.2e5,
           label = 'Election', vjust = 1, hjust = 0.1) +
  facet_grid(Region ~ .)

ggsave('expenditures-stretch-run-economic-annotated.png',
       expenditure.stretch.economic.plot,
       width = 9.2, height = 9.2, units = 'in', dpi = 600)

expenditure.debate.plot <-
  ggplot(expenditure.data, aes(x = Date, y = absDollars)) +
  geom_line(size = 1, aes(colour = Candidate, linetype = Candidate),
            linesize = 2) +
  theme_bw() +
  labs(title = 'Expenditure amounts: debate season',
       x = 'Date', y= ' Dollars') +
  theme(axis.text.x = element_text(angle = 30,hjust=1)) +
  scale_x_date(limits = as.Date(c('2012-10-01', '2012-10-31')),
               breaks = seq(as.Date('2012-10-01'),as.Date('2012-10-31'),
                            by ='5 days'),
```

```
              labels = date_format('%Y %b %d')) +
  large.bold.bottom.legend.facet +
  scale_colour_manual(values = c('#1f78b4', '#33a02c')) +
  scale_linetype_manual(values = c('solid', 'dashed')) +
  coord_cartesian(ylim = c(0, 4e5)) +
  annotate('text', x = as.Date(first.debate), y = 1e5,
           label = '1st', angle = 30, vjust = 1, hjust = 0.1) +
  annotate('text', x = as.Date(vp.debate), y = 0.7e5,
           label = 'VP', angle = 30, vjust = 1, hjust = 0.1) +
  annotate('text', x = as.Date(second.debate), y = 1e5,
           label = '2nd', angle = 30, vjust = 1, hjust = 0.1) +
  annotate('text', x = as.Date(third.debate), y = 2.5e5,
           label = '3rd', angle = 30, vjust = 1, hjust = 0.1)

ggsave('expenditures-debate-annotated.png',
       expenditure.debate.plot,
       width = 9.2, height = 9.2, units = 'in', dpi = 600)

expenditure.debate.census.plot <-
  ggplot(expenditure.data.census.regions, aes(x = Date, y = absDollars)) +
  geom_line(size = 1, aes(colour = Candidate, linetype = Candidate),
            linesize = 2) +
  theme_bw() +
  labs(title = 'Expenditure amounts: debate season',
       x = 'Date', y= ' Dollars') +
  theme(axis.text.x = element_text(angle = 30,hjust=1)) +
  scale_x_date(limits = as.Date(c('2012-10-01', '2012-10-31')),
               breaks = seq(as.Date('2012-10-01'),as.Date('2012-10-31'),
                            by ='5 days'),
               labels = date_format('%Y %b %d')) +
  large.bold.bottom.legend.facet +
  scale_colour_manual(values = c('#1f78b4', '#33a02c')) +
  scale_linetype_manual(values = c('solid', 'dashed')) +
  coord_cartesian(ylim = c(0, 2e5)) +
  annotate('text', x = as.Date(first.debate), y = 1e5,
           label = '1st', angle = 30, vjust = 1, hjust = 0.1) +
  annotate('text', x = as.Date(vp.debate), y = 0.7e5,
           label = 'VP', angle = 30, vjust = 1, hjust = 0.1) +
  annotate('text', x = as.Date(second.debate), y = 1e5,
           label = '2nd', angle = 30, vjust = 1, hjust = 0.1) +
  annotate('text', x = as.Date(third.debate), y = 1e5,
           label = '3rd', angle = 30, vjust = 1, hjust = 0.1) +
  facet_grid(Region ~ .)

ggsave('expenditures-debate-census-annotated.png',
       expenditure.debate.census.plot,
       width = 9.2, height = 9.2, units = 'in', dpi = 600)

expenditure.debate.economic.plot <-
  ggplot(expenditure.data.census.regions, aes(x = Date, y = absDollars)) +
  geom_line(size = 1, aes(colour = Candidate, linetype = Candidate),
            linesize = 2) +
  theme_bw() +
```

```
  labs(title = 'Expenditure amounts: debate season',
       x = 'Date', y= ' Dollars') +
  theme(axis.text.x = element_text(angle = 30,hjust=1)) +
  scale_x_date(limits = as.Date(c('2012-10-01', '2012-10-31')),
               breaks = seq(as.Date('2012-10-01'),as.Date('2012-10-31'),
                            by ='5 days'),
               labels = date_format('%Y %b %d')) +
  large.bold.bottom.legend.facet +
  scale_colour_manual(values = c('#1f78b4', '#33a02c')) +
  scale_linetype_manual(values = c('solid', 'dashed')) +
  coord_cartesian(ylim = c(0, 2e5)) +
  annotate('text', x = as.Date(first.debate), y = 1e5,
           label = '1st', angle = 30, vjust = 1, hjust = 0.1) +
  annotate('text', x = as.Date(vp.debate), y = 0.7e5,
           label = 'VP', angle = 30, vjust = 1, hjust = 0.1) +
  annotate('text', x = as.Date(second.debate), y = 1e5,
           label = '2nd', angle = 30, vjust = 1, hjust = 0.1) +
  annotate('text', x = as.Date(third.debate), y = 1e5,
           label = '3rd', angle = 30, vjust = 1, hjust = 0.1) +
  facet_grid(Region ~ .)

ggsave('expenditures-debate-economic-annotated.png',
       expenditure.debate.economic.plot,
       width = 9.2, height = 9.2, units = 'in', dpi = 600)


# breakout detection data setup ------------------------------------------


library(BreakoutDetection)
library(lubridate)


# net contributions by candidate
obama.net.contributions  <-
  dplyr::filter(fec.data.net.contributions, Candidate == 'Obama, Barack')

obama.net.contributions <-
  dplyr::rename(obama.net.contributions, timestamp = Date, count = Dollars)

obama.net.contributions$timestamp <- ymd(obama.net.contributions$timestamp)

obama.net.contributions <- droplevels(obama.net.contributions)

romney.net.contributions  <-
  dplyr::filter(fec.data.net.contributions, Candidate == 'Romney, Mitt')

romney.net.contributions <-
  dplyr::rename(romney.net.contributions, timestamp = Date, count = Dollars)

romney.net.contributions$timestamp <- ymd(romney.net.contributions$timestamp)

romney.net.contributions <- droplevels(romney.net.contributions)
```

```
# donations by candidate

obama.donations <-
  dplyr::filter(donation.data, Candidate == 'Obama, Barack')

obama.donations <-
  dplyr::rename(obama.donations, timestamp = Date, count = Dollars)

obama.donations$timestamp <- ymd(obama.donations$timestamp)

obama.donations <- droplevels(obama.donations)


romney.donations <-
  dplyr::filter(donation.data, Candidate == 'Romney, Mitt')

romney.donations <-
  dplyr::rename(romney.donations, timestamp = Date, count = Dollars)

romney.donations$timestamp <- ymd(romney.donations$timestamp)

romney.donations <- droplevels(romney.donations)

# expenditures by candidate
obama.expenditures <-
  dplyr::filter(expenditure.data, Candidate == 'Obama, Barack')

obama.expenditures <-
  dplyr::rename(obama.expenditures, timestamp = Date, count = absDollars)

obama.expenditures$timestamp <- ymd(obama.expenditures$timestamp)

obama.expenditures <- droplevels(obama.expenditures)


romney.expenditures <-
  dplyr::filter(expenditure.data, Candidate == 'Romney, Mitt')

romney.expenditures <-
  dplyr::rename(romney.expenditures, timestamp = Date, count = absDollars)

romney.expenditures$timestamp <- ymd(romney.expenditures$timestamp)

romney.expenditures <- droplevels(romney.expenditures)

# breakout detection calculations ----------------------------------------

# net data
obama.net.breakout.week <-
  breakout(obama.net.contributions, min.size = 7, method = 'multi',
           beta = 1e-3, degree = 1, plot = T,
           xlab = 'Date', ylab = 'Dollars',
```

```
          title = 'Obama campaign net expense breakouts - 7 day period')

obama.net.breakout.2day <-
  breakout(obama.net.contributions, min.size = 2, method = 'multi',
           beta = 1e-3, degree = 1, plot = T,
           xlab = 'Date', ylab = 'Dollars',
           title = 'Obama campaign net expense breakouts - 2 day period')

romney.net.breakout.week <-
  breakout(romney.net.contributions, min.size = 7, method = 'multi',
           beta = 1e-3, degree = 1, plot = T,
           xlab = 'Date', ylab = 'Dollars',
           title = 'Romney campaign net expense breakouts - 7 day period')

romney.net.breakout.2day <-
  breakout(romney.net.contributions, min.size = 2, method = 'multi',
           beta = 1e-3, degree = 1, plot = T,
           xlab = 'Date', ylab = 'Dollars',
           title = 'Romney campaign net expense breakouts - 2 day period')

# donations
obama.donation.breakout.week <-
  breakout(obama.donations, min.size = 7, method = 'multi',
           beta = 1e-3, degree = 1, plot = T,
           xlab = 'Date', ylab = 'Dollars',
           title = 'Obama campaign donation breakouts - 7 day period')

obama.donation.breakout.2day <-
  breakout(obama.donations, min.size = 2, method = 'multi',
           beta = 1e-3, degree = 1, plot = T,
           xlab = 'Date', ylab = 'Dollars',
           title = 'Obama campaign donation breakouts - 2 day period')

romney.donation.breakout.week <-
  breakout(romney.donations, min.size = 7, method = 'multi',
           beta = 1e-3, degree = 1, plot = T,
           xlab = 'Date', ylab = 'Dollars',
           title = 'Romney campaign donation breakouts - 7 day period')

romney.donation.breakout.2day <-
  breakout(romney.donations, min.size = 2, method = 'multi',
           beta = 1e-3, degree = 1, plot = T,
           xlab = 'Date', ylab = 'Dollars',
           title = 'Romney campaign donation breakouts - 2 day period')

# expenditures

obama.expenditure.breakout.week <-
  breakout(obama.expenditures, min.size = 7, method = 'multi',
           beta = 1e-3, degree = 1, plot = T,
           xlab = 'Date', ylab = 'Dollars',
           title = 'Obama campaign expenditure breakouts - 7 day period')
```

```
obama.expenditure.breakout.2day <-
  breakout(obama.expenditures, min.size = 2, method = 'multi',
           beta = 1e-3, degree = 1, plot = T,
           xlab = 'Date', ylab = 'Dollars',
           title = 'Obama campaign expenditure breakouts - 2 day period')

romney.expenditure.breakout.week <-
  breakout(romney.expenditures, min.size = 7, method = 'multi',
           beta = 1e-3, degree = 1, plot = T,
           xlab = 'Date', ylab = 'Dollars',
           title = 'Romney campaign expenditure breakouts - 7 day period')

romney.expenditure.breakout.2day <-
  breakout(romney.expenditures, min.size = 2, method = 'multi',
           beta = 1e-3, degree = 1, plot = T,
           xlab = 'Date', ylab = 'Dollars',
           title = 'Romney campaign expenditure breakouts - 2 day period')


# save workspace ------------------------------------------------------------

save.image('fec-breakout-detection-analysis.rda')
```