Show Me the Money

Using tidyverse tools to analyze fundraising data

The dataset of my dreams

DonorsChoose.org

Data   Kernels (149)   Discussion (39)   Activity   Metadata

Data Sources

⊞ Donations.csv	4.69m x 7	
⊞ Donors.csv	2.12m x 5	
⊞ Projects.csv		
⊞ Resources.csv	7.21m x 5	
⊞ Schools.csv	73.0k x 9	
⊞ Teachers.csv	403k x 3	

About this file

For every project in the Projects.csv dataset, there are one or more donations. This dataset contains each donation from a citizen donor and is joined with the dataset above using the "Project ID" column.testt test1

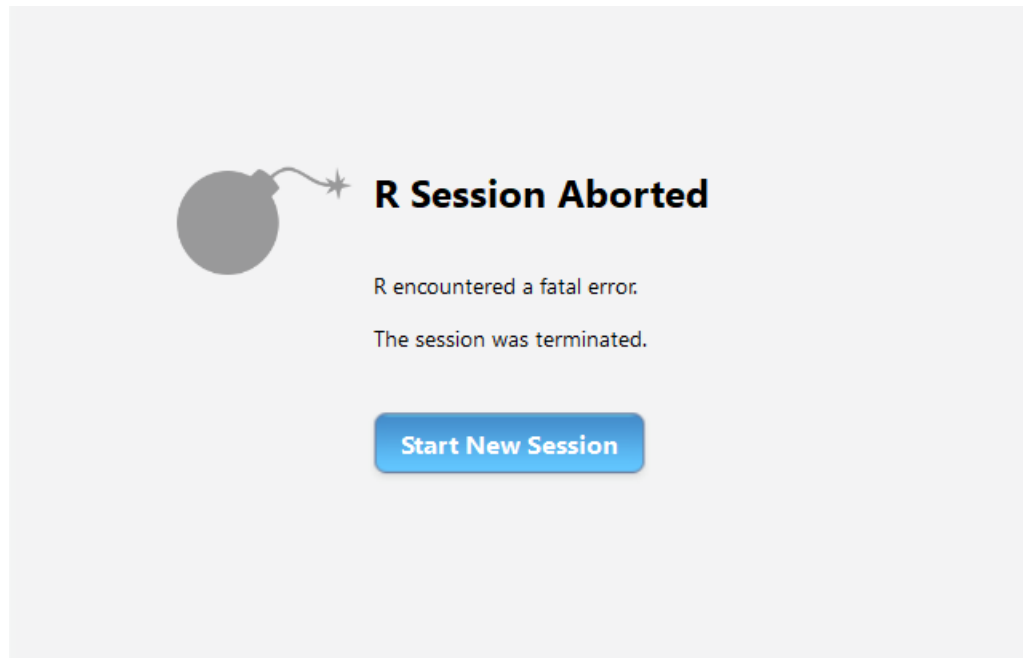# Phase 1: This is Fun!

Phase 1: This is Fun!

Phase 1: This is Fun!

Phase 2: Getting to the Money

```r
donations_plus <- donations %>%
  left_join(projects, by = "project_id") %>%
  left_join(donors, by = "donor_id") %>%
  left_join(schools, by = "school_id") %>%
  left_join(teachers, by = "teacher_id")
```

Phase 2: Getting to the Money

```r
donations_plus <- donations %>%
  left_join(projects, by = "project_id") %>%
  left_join(donors, by = "donor_id") %>%
  left_join(schools, by = "school_id") %>%
  left_join(teachers, by = "teacher_id")

revenue_per_year_plot <- donations_plus %>%
  filter(year(donation_received_date) > 2012 & year(donation_received_date) < 2018) %>%
  group_by(year = year(donation_received_date), retention_status) %>%
  summarize(donation_millions = sum(donation_amount)/1000000) %>%
  ggplot(aes(x = year, y = donation_millions, fill = "pink")) +
  geom_col() +
  labs(y = "Total donation value (in millions)") +
  theme(legend.position = "none",
        axis.title.x = element_blank()) +
  scale_y_continuous(labels = dollar_format())
```

Phase 2: Getting to the Money

```r
donations_plus <- donations %>%
  left_join(projects, by = "project_id") %>%
  left_join(donors, by = "donor_id") %>%
  left_join(schools, by = "school_id") %>%
  left_join(teachers, by = "teacher_id")


revenue_per_year_plot <- donations_plus %>%
  filter(year(donation_received_date) > 2012 & year(donation_received_date) < 2018) %>%
  group_by(year = year(donation_received_date), retention_status) %>%
  summarize(donation_millions = sum(donation_amount)/1000000) %>%
  ggplot(aes(x = year, y = donation_millions, fill = "pink")) +
  geom_col() +
  labs(y = "Total donation value (in millions)") +
  theme(legend.position = "none",
        axis.title.x = element_blank()) +
  scale_y_continuous(labels = dollar_format())


plot_grid(perc_growth_per_year_plot, revenue_per_year_plot,
          align = "v", nrow = 2, rel_heights = c(1/4, 3/4))
```

Phase 2: Getting to the Money
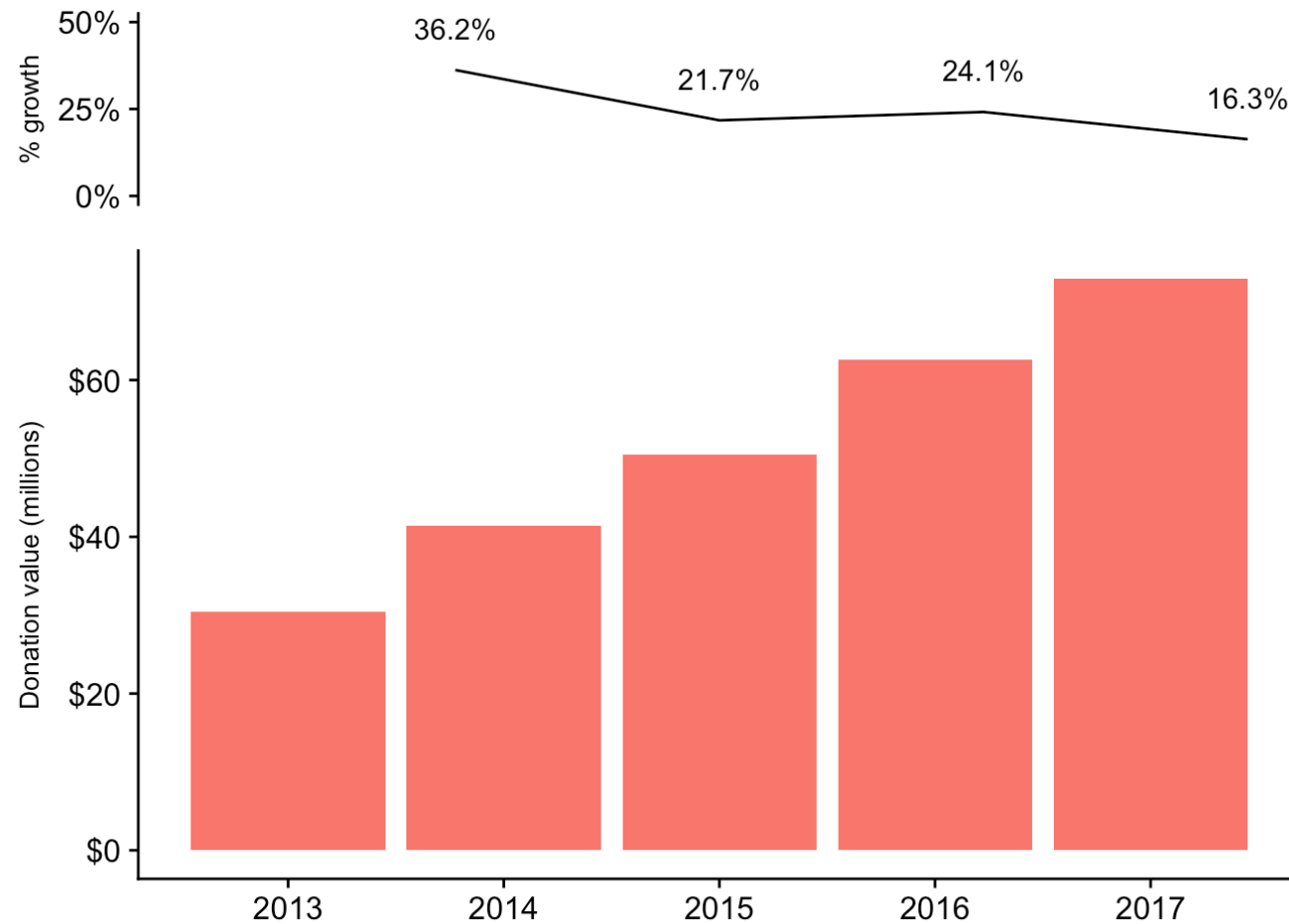
Are donors coming back?

$$\text{Retention for this year} = \frac{\text{Value of this year's donations from donors who also gave last year}}{\text{Total value of last year's donations}}$$
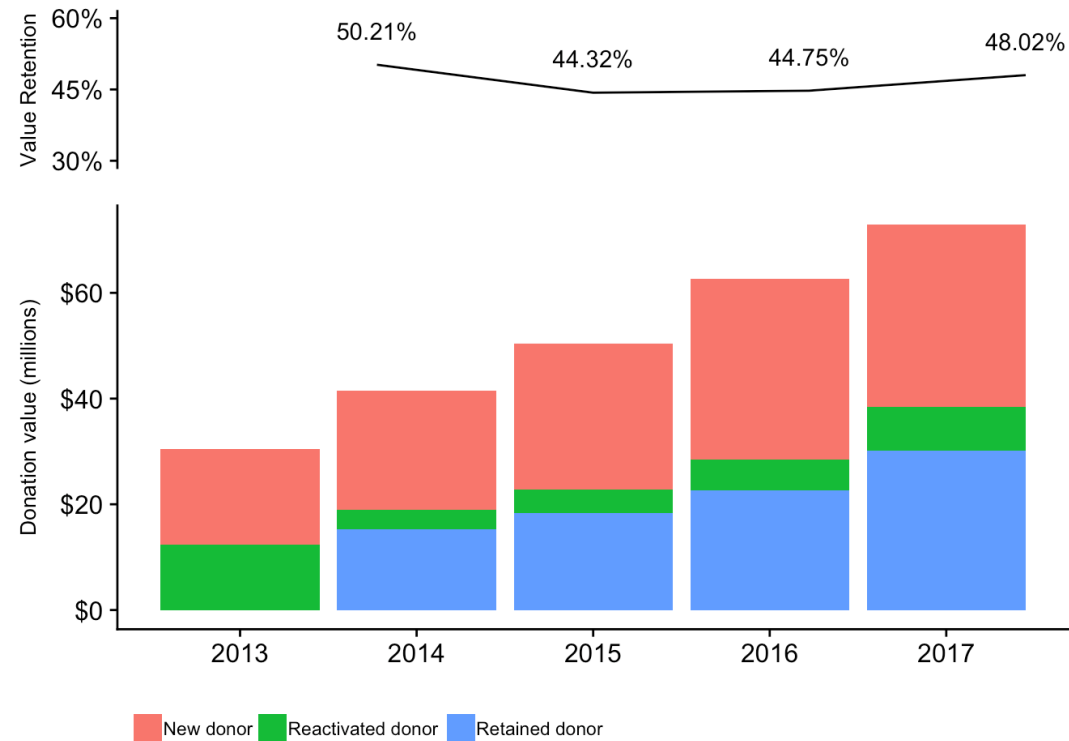
Are donors coming back?

```r
donations_plus <- donations_plus %>%
  mutate(donor_gave_last_year = case_when(
    year(donation_received_date) == 2013 ~ as.character(donor_id) %in% pull(donors_2012),
    year(donation_received_date) == 2014 ~ as.character(donor_id) %in% pull(donors_2013),
    year(donation_received_date) == 2015 ~ as.character(donor_id) %in% pull(donors_2014),
    year(donation_received_date) == 2016 ~ as.character(donor_id) %in% pull(donors_2015),
    year(donation_received_date) == 2017 ~ as.character(donor_id) %in% pull(donors_2016),
    year(donation_received_date) == 2018 ~ as.character(donor_id) %in% pull(donors_2017)))


donations_plus <- donations_plus %>%
  mutate(retention_status = case_when(
    donor_gave_last_year ~ "Retained donor",
    !donor_gave_last_year & donor_gave_ever_before ~ "Reactivated donor",
    !donor_gave_ever_before ~ "New donor"
  ))
```

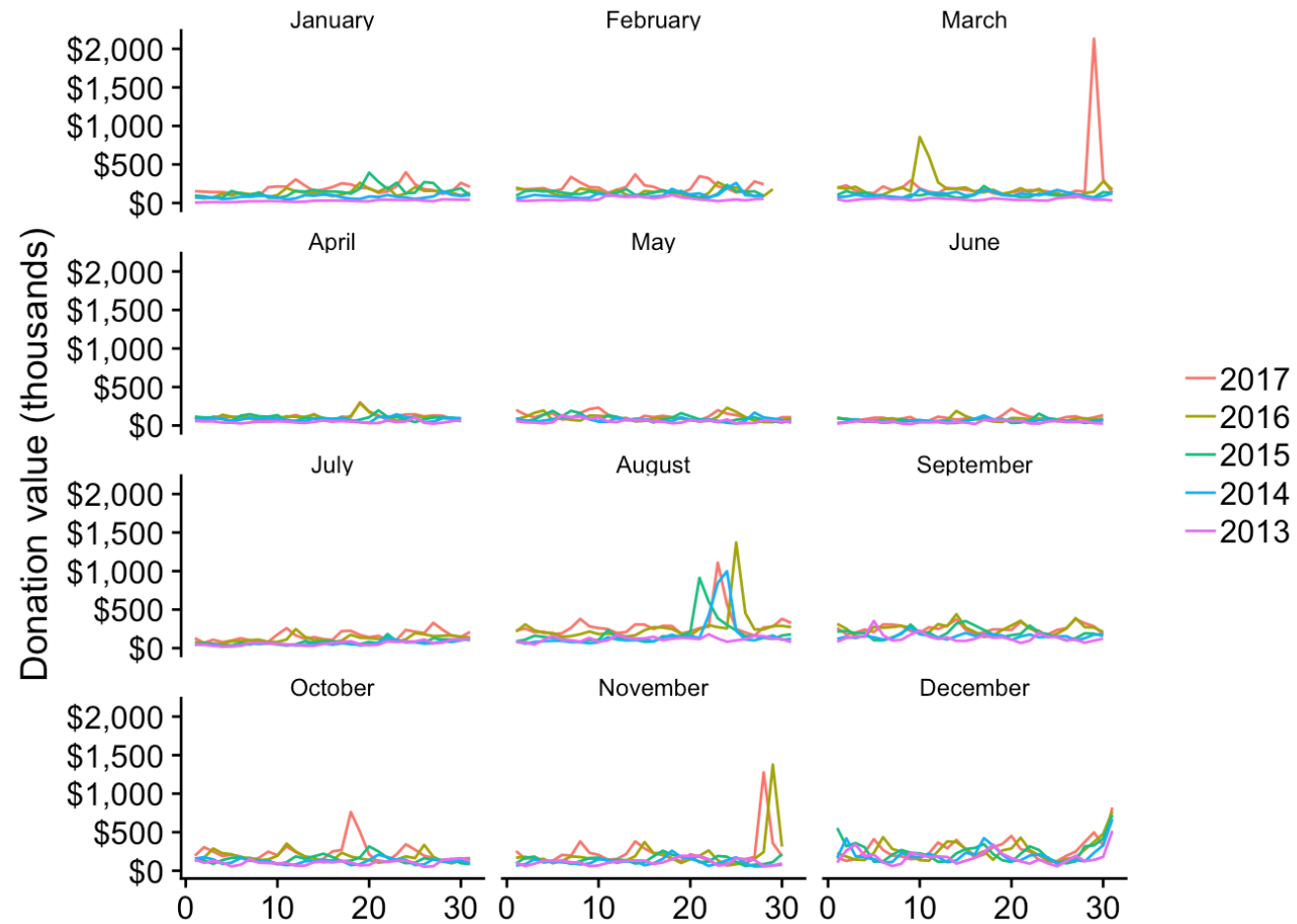# Are donors coming back?

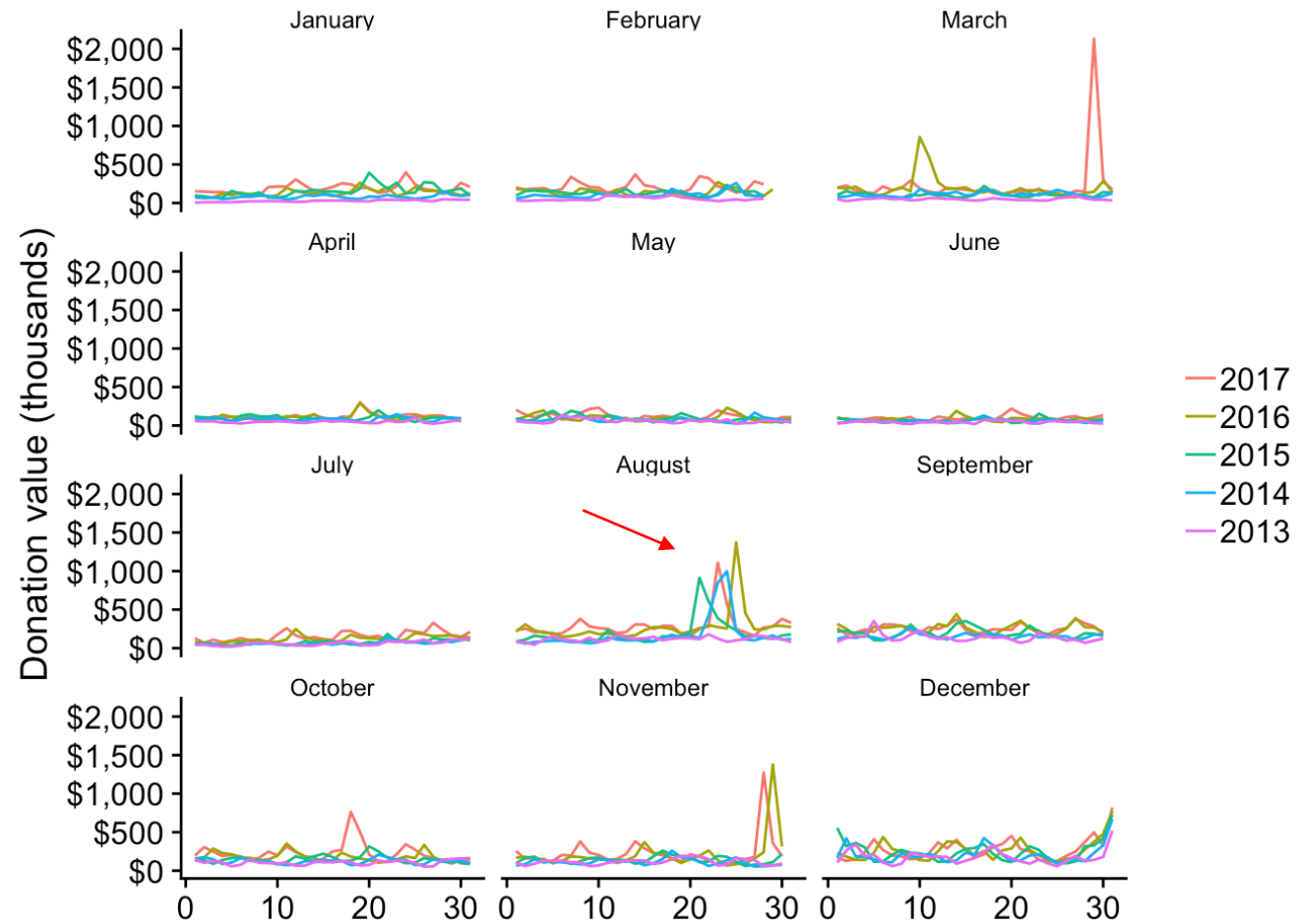Value Retention
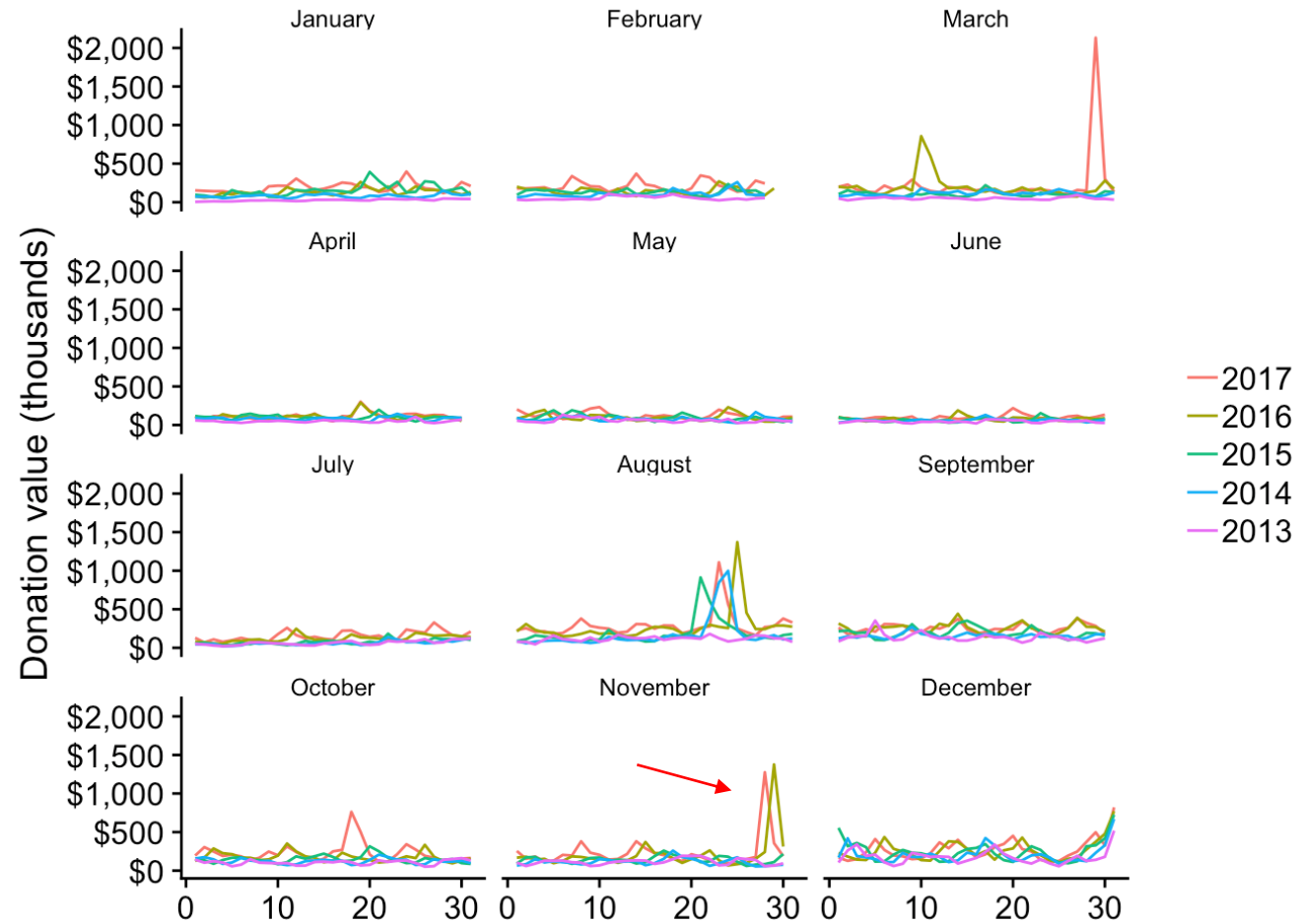
Are donors coming back?

Value Retention
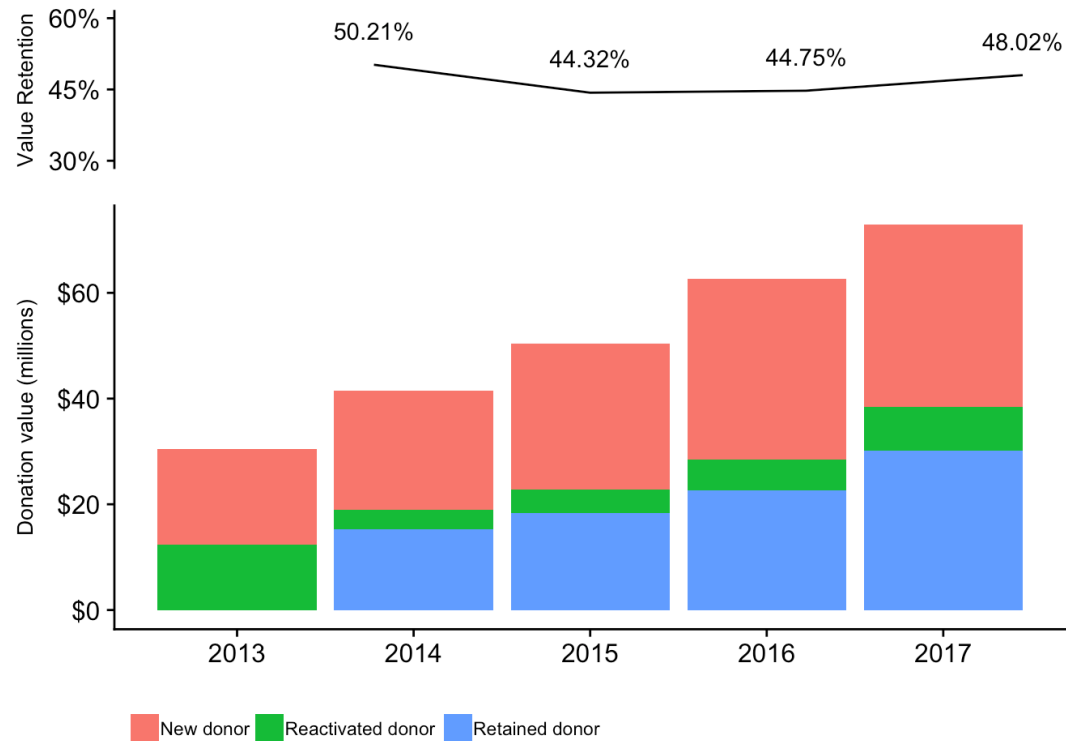
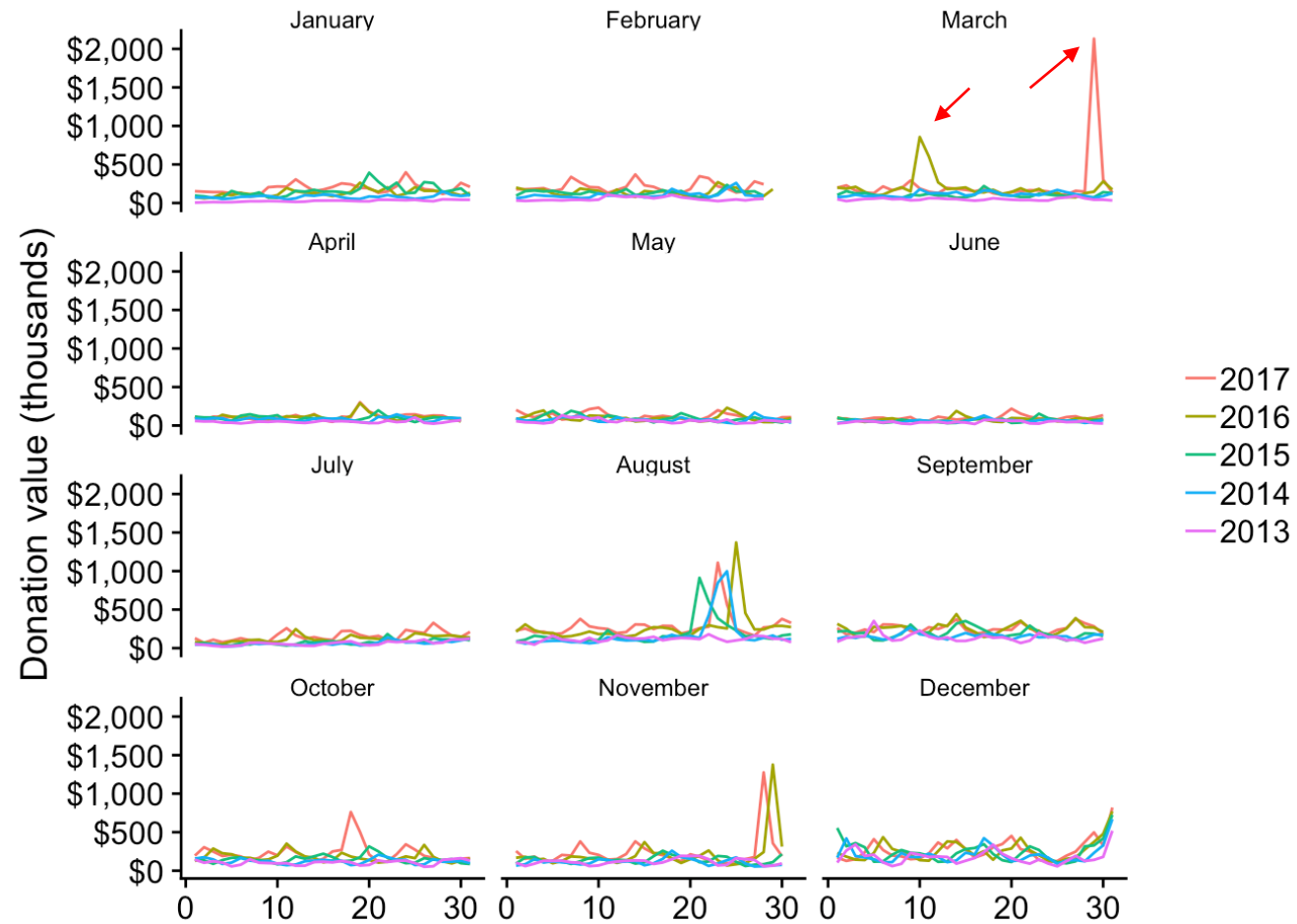Donor Retention
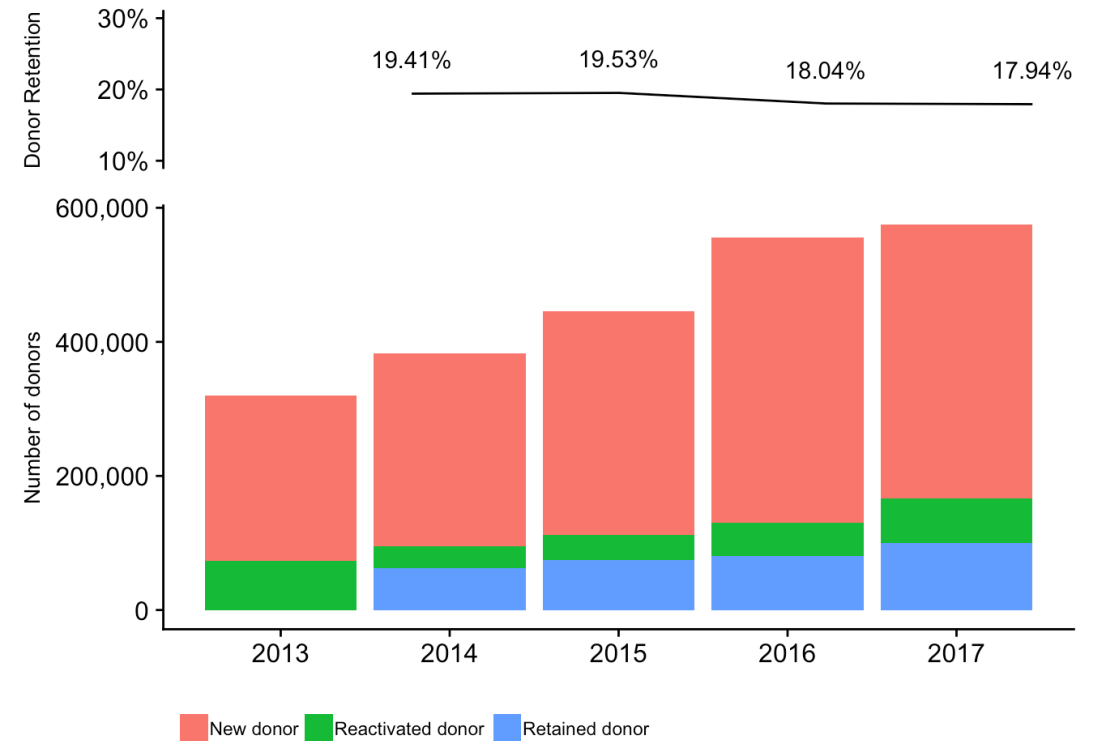
Well–timed campaigns

Well-timed campaigns

Well–timed campaigns

Well-timed campaigns

Phase 3: Bringing it all together

Phase 3: Bringing it all together

Alejandra Gerosa

About me

December 19, 2018   #Fundraising Analytics | #R | #retention
Fundraising Analytics on DonorsChoose.org's Data

This project analyses data released by DonorsChoose.org for their mid-2018 Kaggle competition. The dataset consists of 6 tables containing information on more than 4 million donations, and was analyzed using R. In this writeup, I set out to summarize a few key insights regarding their fundrasing performance between 2013 and 2017. For a more comprehensive approach, the full code of my ongoing exploration of the dataset can be found here.

Key Insights

- Donations are growing by an average of 24.6% every year, and growth has been slowing down.
- Most donors have given only once, showing uncaptured opportunities for long term value.
- Back-to-school and end-of-year donations mark the most important yearly-consistent giving times, with Giving Tuesday and March campaigns gaining ground.

Bottom line: Over the covered period DonorsChoose.org has shown exceptional success in growth and new donor acquisition through well-performing and well-timed campaigns. They have not yet been able to similarly outperform industry averages when it comes to retention.

Thank you! + Shameless plug for my just-for-fun analysis of your Kaggle data ⟩ Inbox
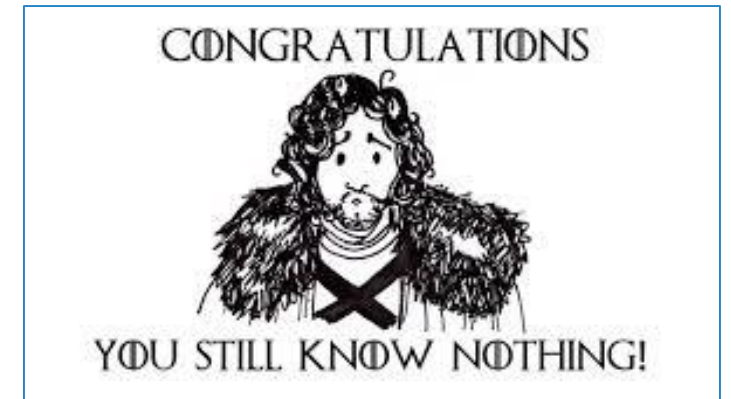
me   May 16
Hi Stephen, Thanks for sharing your experie...

S   Stephen ▇▇▇ May 28
to me ⌄

Hi Alejandra! Thanks for reaching

Phase 3: Bringing it all together

Thank you!!

Alejandra Gerosa – www.alegerosa.com – @alejagerosa