Pull fundraising data

Pull the data from 2018

The data are available here: https://www.dailykos.com/stories/2018/7/17/1781034/-Daily-Kos-Elections-2Q-2018-House-fundr

Load libraries

```r
library(googlesheets)
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```r
library(devtools)
library(stringr)
library(purrr)
library(tidyr)
library(readr)
```

Pull data

```r
fundraise_2018_url = gs_url("https://docs.google.com/spreadsheets/d/1zp3vv0pHUVeA_GJ9F9tpsa01dwu1h7HJiU8
```

```
## Sheet-identifying info appears to be a browser URL.
## googlesheets will attempt to extract sheet key from the URL.

## Putative key: 1zp3vv0pHUVeA_GJ9F9tpsa01dwu1h7HJiU809uZ683E

## Worksheets feed constructed with public visibility
```

```r
fundraise_2018 = gs_read(fundraise_2018_url,col_names=TRUE)
```

```
## Accessing worksheet titled 'List'.

## Parsed with column specification:
## cols(
##   CD = col_character(),
##   Name = col_character(),
##   Party = col_character(),
##   `2Q
## Raised` = col_character(),
##   `Self
## Fund` = col_character(),
##   `Self
```

```
## Loan` = col_character(),
##   Spent = col_character(),
##   `Raised
## CTD` = col_character(),
##   CoH = col_character(),
##   `Self Fund
## CTD` = col_character(),
##   Transfer = col_character(),
##   `Candidate ID` = col_character(),
##   `Committee ID` = col_character()
## )

## Warning in rbind(names(probs), probs_f): number of columns of result is not
## a multiple of vector length (arg 2)

## Warning: 1 parsing failure.
## row # A tibble: 1 x 5 col     row col   expected   actual     file         expected   <int> <chr> <cl
```

```
glimpse(fundraise_2018)
```

```
## Observations: 424
## Variables: 13
## $ CD              <chr> "Raised\"", "AK-AL", "AK-AL", "AL-02", "AL-02...
## $ Name            <chr> "Self\nFund", "Alyse Galvin", "Don Young", "M...
## $ Party           <chr> "Self\nLoan", "(I/D)", "(R-inc)", "(R-inc)", ...
## $ `2Q\nRaised`    <chr> "Spent", "$242", "$177", "$1,010", "$62", "$7...
## $ `Self\nFund`    <chr> "Raised\nCTD", "$4", "$0", "$0", "$0", "$0", ...
## $ `Self\nLoan`    <chr> "CoH", "$0", "$0", "$0", "$0", "$0", "$0", "$...
## $ Spent           <chr> "Self Fund\nCTD", "$179", "$170", "$1,100", "...
## $ `Raised\nCTD`   <chr> "Transfer", "$517", "$768", "$2,101", "$97", ...
## $ CoH             <chr> "Candidate ID", "$267", "$470", "$731", "$163...
## $ `Self Fund\nCTD` <chr> "Committee ID", "$14", "$0", "$0", "$0", "$0"...
## $ Transfer        <chr> NA, "$0", "$0", "$31", "$0", "$0", "$0", "$0"...
## $ `Candidate ID`  <chr> NA, "H8AK00140", "H6AK00045", "H0AL02087", "H...
## $ `Committee ID`  <chr> NA, "C00665711", "C00012229", "C00462143", "C...
```

Data cleaning

Change names

```
colnames(fundraise_2018) = c("cd","name",   "party", "q2_raised",
                             "self_fund", "self_loan", "spent",
                             "raised_ctd","coh","self_fund_ctd","transfer",
                             "candidate_id","committee_id")
fundraise_2018 = fundraise_2018[-1,]
```

Make dollar values numeric:

```
dol_numeric = function(x){
  return(as.numeric(str_replace_all(x,"[,\\$]","")))
}
fundraise_2018 = fundraise_2018 %>%
  mutate(q2_raised  = dol_numeric(q2_raised)) %>%
  mutate(self_fund  = dol_numeric(self_fund)) %>%
  mutate(self_loan  = dol_numeric(self_loan)) %>%
  mutate(spent  = dol_numeric(spent)) %>%
```

```
  mutate(raised_ctd  = dol_numeric(raised_ctd)) %>%
  mutate(coh  = dol_numeric(coh)) %>%
  mutate(self_fund_ctd  = dol_numeric(self_fund_ctd)) %>%
  mutate(transfer  = dol_numeric(transfer))
```

Warning in dol_numeric(q2_raised): NAs introduced by coercion

Make candidate information a little easier

```
fundraise_2018 = fundraise_2018 %>%
  mutate(incumbant = str_detect(fundraise_2018$party,"inc")) %>%
  mutate(party_ind = str_detect(party,"I")) %>%
  mutate(party_dem = str_detect(party,"D")) %>%
  mutate(party_rep = str_detect(party,"R"))
```

Finally separate state/district

```
fundraise_2018 = fundraise_2018 %>%
  separate(cd,into=c("state","district"),sep="-")
```

Write data

```
write_csv(fundraise_2018,"fundraise_2018_dailykos.csv")
```

Session Info

Here is some information about the document so we can debug.

```
Sys.time()
```

[1] "2018-08-22 17:02:15 UTC"

```
session_info()
```

Warning in system("timedatectl", intern = TRUE): running command
'timedatectl' had status 1

Session info --

```
##  setting  value
##  version  R version 3.5.0 (2018-04-23)
##  system   x86_64, linux-gnu
##  ui       X11
##  language (EN)
##  collate  C.UTF-8
##  tz       Etc/UTC
##  date     2018-08-22
```

Packages --

```
##  package    * version date       source
##  assertthat   0.2.0   2017-04-11 RSPM (R 3.5.0)
##  backports    1.1.2   2017-12-13 RSPM (R 3.5.0)
##  base       * 3.5.0   2018-06-11 local
##  bindr        0.1.1   2018-03-13 RSPM (R 3.5.0)
##  bindrcpp   * 0.2.2   2018-03-29 RSPM (R 3.5.0)
```

```
##   cellranger      1.1.0   2016-07-27 RSPM (R 3.5.0)
##   cli             1.0.0   2017-11-05 RSPM (R 3.5.0)
##   compiler        3.5.0   2018-06-11 local
##   crayon          1.3.4   2017-09-16 RSPM (R 3.5.0)
##   curl            3.2     2018-03-28 RSPM (R 3.5.0)
##   datasets      * 3.5.0   2018-06-11 local
##   devtools      * 1.13.6  2018-06-27 RSPM (R 3.5.0)
##   digest          0.6.15  2018-01-28 RSPM (R 3.5.0)
##   dplyr         * 0.7.6   2018-06-29 RSPM (R 3.5.0)
##   evaluate        0.11    2018-07-17 RSPM (R 3.5.0)
##   fansi           0.3.0   2018-08-13 RSPM (R 3.5.0)
##   glue            1.3.0   2018-07-17 RSPM (R 3.5.0)
##   googlesheets  * 0.3.0   2018-06-29 RSPM (R 3.5.0)
##   graphics      * 3.5.0   2018-06-11 local
##   grDevices     * 3.5.0   2018-06-11 local
##   hms             0.4.2   2018-03-10 RSPM (R 3.5.0)
##   htmltools       0.3.6   2017-04-28 RSPM (R 3.5.0)
##   httr            1.3.1   2017-08-20 RSPM (R 3.5.0)
##   knitr           1.20    2018-02-20 RSPM (R 3.5.0)
##   magrittr        1.5     2014-11-22 RSPM (R 3.5.0)
##   memoise         1.1.0   2017-04-21 RSPM (R 3.5.0)
##   methods       * 3.5.0   2018-06-11 local
##   pillar          1.3.0   2018-07-14 RSPM (R 3.5.0)
##   pkgconfig       2.0.2   2018-08-16 RSPM (R 3.5.0)
##   purrr         * 0.2.5   2018-05-29 RSPM (R 3.5.0)
##   R6              2.2.2   2017-06-17 RSPM (R 3.5.0)
##   Rcpp            0.12.18 2018-07-23 RSPM (R 3.5.0)
##   readr         * 1.1.1   2017-05-16 RSPM (R 3.5.0)
##   rlang           0.2.2   2018-08-16 RSPM (R 3.5.0)
##   rmarkdown       1.10    2018-06-11 RSPM (R 3.5.0)
##   rprojroot       1.3-2   2018-01-03 RSPM (R 3.5.0)
##   stats         * 3.5.0   2018-06-11 local
##   stringi         1.2.4   2018-07-20 RSPM (R 3.5.0)
##   stringr       * 1.3.1   2018-05-10 RSPM (R 3.5.0)
##   tibble          1.4.2   2018-01-22 RSPM (R 3.5.0)
##   tidyr         * 0.8.1   2018-05-18 RSPM (R 3.5.0)
##   tidyselect      0.2.4   2018-02-26 RSPM (R 3.5.0)
##   tools           3.5.0   2018-06-11 local
##   utf8            1.1.4   2018-05-24 RSPM (R 3.5.0)
##   utils         * 3.5.0   2018-06-11 local
##   withr           2.1.2   2018-03-15 RSPM (R 3.5.0)
##   xml2            1.2.0   2018-01-24 RSPM (R 3.5.0)
##   yaml            2.2.0   2018-07-25 RSPM (R 3.5.0)
```