GSA Competition

Ryan Tuggle and Jessica Bonnie

April 13, 2016

This R Markdown document shows an initial exploration of GSA contracting competition data. See end for summary section containing notes and questions for next iteration.

```
#load libraries silently.

require(dplyr)
require(ggplot2)
require(pander)
require(knitr)


opts_chunk$set(tidy=TRUE)
```

Kristin's Data

Currently this section is merely trying to reproduce Ryan's steps from his data set, which are included in the next section.

```
# Load data into single frame from multiple CSV files in directory.


# Directory of files from Kristin, filenames contain numerical date
# information that we will want to capture.
datadir <- "./data/FPDS_20160413"

reading_helper <- function(infile, datadir, ...) {
    ## Function to read in csv and strip date information from file name
    temp <- read.csv(paste(datadir, infile, sep = "/"), stringsAsFactors = F,
        ...)
    temp$period <- gsub("FPDS", "", infile) %>% gsub("\\.csv", "", .)
    return(temp)
}

# if (exists(fpds)){rm(fpds)}


files <- list.files(datadir, pattern = "FPDS[0-9_]*.csv")
fpds <- do.call(rbind, lapply(files, function(x) reading_helper(x, datadir)))

# Here are the column headers in the data.frame
names(fpds)
```

```
##  [1] "Contracting.Agency.ID"
##  [2] "Contracting.Agency.Name"
##  [3] "Contracting.Department.ID"
```

```
##  [4] "Contracting.Department.Name"
##  [5] "Contracting.Office.ID"
##  [6] "Contracting.Office.Name"
##  [7] "Contracting.Office.Region"
##  [8] "PIID"
##  [9] "Modification.Number"
## [10] "Referenced..IDV.PIID"
## [11] "Referenced.IDV.Mod.Number"
## [12] "Date.Signed"
## [13] "Effective.Date"
## [14] "Est..Ultimate.Completion.Date"
## [15] "Completion.Date"
## [16] "Funding.Agency.ID"
## [17] "Funding.Agency.Name"
## [18] "Funding.Office.ID"
## [19] "Funding.Office.Name"
## [20] "Type.of.Contract"
## [21] "NAICS.Code"
## [22] "NAICS.Description"
## [23] "Product.or.Service.Code"
## [24] "Product.or.Service.Description"
## [25] "Place.of.Performance.Zip.Code"
## [26] "Type.of.Set.Aside"
## [27] "Award.or.IDV.Type"
## [28] "Last.Modified.Date"
## [29] "DUNS.Number"
## [30] "Global.DUNS.Number"
## [31] "Congressional.District...Contractor"
## [32] "Contractor.Name"
## [33] "Vendor.Name"
## [34] "Base.and.Exercised.Options.Value"
## [35] "Action.Obligation"
## [36] "Base.and.All.Options.Value"
## [37] "period"
```

This data has already been filtered to have only GSA as the Contracting Department.

```
## Adapt numeric columns to be numbers without all the extra garbage
## characters

fpds$Action.Obligation <- gsub("[^0-9.]", "", fpds$Action.Obligation) %>% as.numeric()


## Select Desired Variables

fpds <- fpds %>% select(Referenced..IDV.PIID, PIID, Referenced.IDV.Mod.Number,
    Modification.Number, Date.Signed, Action.Obligation, Contracting.Agency.ID,
    Contracting.Office.ID, Contracting.Office.Name, Award.or.IDV.Type, Contracting.Agency.Name,
    Funding.Office.Name, Funding.Agency.Name, Place.of.Performance.Zip.Code,
    Type.of.Contract, Global.DUNS.Number, Type.of.Set.Aside, period)


## Create any new variables we might need
fpds <- fpds %>% mutate(cat = paste(Referenced..IDV.PIID, PIID, sep = "-"))
```

```
# The following variables removed from selection list due to not being
# present in the current pull: Extent.Competed, Number.of.Offers.Received,
# Other.Than.Full.and.Open.Competition, Actions, Fiscal.Year,
# Number.of.Records
```

Take a look at how the extent competed variable is aggregated across mods. One goal is to understand how to sum up the competition variables such as the Number of Offers Received.

```
# We do not currently have the columns necessary for this: fpds.test1 <-
# fpds %>% filter(Number.of.Records > 1 | Actions > 1)


# group variables by PIIDs
fpds.test2 <- fpds %>% group_by(cat) %>% summarize(cnt = n_distinct(cat)) %>%
    arrange(desc(cnt)) %>% filter(cnt > 1)

# This is empty just like in Ryan's set
head(fpds.test2)
```

```
## Source: local data frame [0 x 2]
##
## Variables not shown: cat (chr), cnt (int)
```

```
# does any PIID have a mod in this set?
fpds.test3 <- fpds %>% filter(Modification.Number != 0) %>% select(-Date.Signed,
    -Contracting.Office.ID, -Contracting.Agency.ID, -Funding.Agency.Name, -Funding.Office.Name,
    -Contracting.Agency.Name, -Global.DUNS.Number, -Place.of.Performance.Zip.Code,
    -Type.of.Set.Aside)
pander(tbl_df(sample_n(fpds.test3, 20)), split.cells = 15)
```

Table 1: Table continues below

Referenced..IDV.PIID	PIID	Referenced.IDV.Mod.Number
	GS23F0168L	
GS07P11HHD0020	GSP0713UF5028	0
	GS02F0195X	
GS07F9462G	GSP0812JA0040	14
	GS07F9428S	
	GS21F0085U	
	GS35F0347R	
GS07P99HHD0080	GSP0713UT5079	0
	GS35F0678P	
	GS02Q14DCR0008	
	GS07F0377X	
	GS00F0012S	
	GS35F0235Y	
	GS21F0117W	
	GS35F0638J	
GS07P11HHD0004	GSP0714UC5013	0
	GS02F0105X	
GS09P14KSD0022	GSP0915WB7007	0

Referenced..IDV.PIID	PIID	Referenced.IDV.Mod.Number
GS03P11CDD0036	GSP0312DX5069	0
	GS04P11BVC0027	0

Table 2: Table continues below

Modification.Number	Action.Obligation	Contracting.Office.Name
2	0	GSA/FAS MANAGEMENT SERVICES CENTER
PS03	0	GSA/PBS OK/NM SECTION
PO0036DELPROD	0	GSA/FAS NORTHEAST&CARIBBEAN SUPPLY AND ACQUISITION CENTER (2QSB)
PS02	0	GSA/PBS/DENVER FEDERAL CENTER SERVICE DIVISION
PO0022ADDPROD	0	GSA/FAS GREATER SOUTHWEST ACQUISITION CTR (7FCO)
PS0027EPACPL	0	GSA/FAS TOOLS ACQUISITION DIVISION I
PO0005EXERCISEOPTION	0	GSA/FAS CENTER FOR IT SCHEDULE PROGRAM
PS02	0	GSA/PBS ARK/LA/TX SECTION
PO0303DELPROD	0	GSA/FAS CENTER FOR IT SCHEDULE PROGRAM
PO0007PRCREDINDUST	0	GSA/FAS NORTHEAST&CARIBBEAN SUPPLY AND ACQUISITION CENTER (2QSB)
PA0027AUTHNEG	0	GSA/FAS GREATER SOUTHWEST ACQUISITION CTR (7FCO)
CMA473REVTERMCOND	0	GSA/FAS MANAGEMENT SERVICES CENTER
PS0010ADDSINS	0	GSA/FAS CENTER FOR IT SCHEDULE PROGRAM
PA0039MANUFDEALER	0	GSA/FAS TOOLS ACQUISITION DIVISION I
PO0155DELPROD	0	GSA/FAS CENTER FOR IT SCHEDULE PROGRAM
PS03	0	GSA BORDER STATION SUPPORT GSA-WIDE
PA0004TERMSINS	0	GSA/FAS NORTHEAST&CARIBBEAN SUPPLY AND ACQUISITION CENTER (2QSB)
PA03	0	SAN FRANCISCO SERVICE CENTER, CENTRAL CA FIELD OFFICE

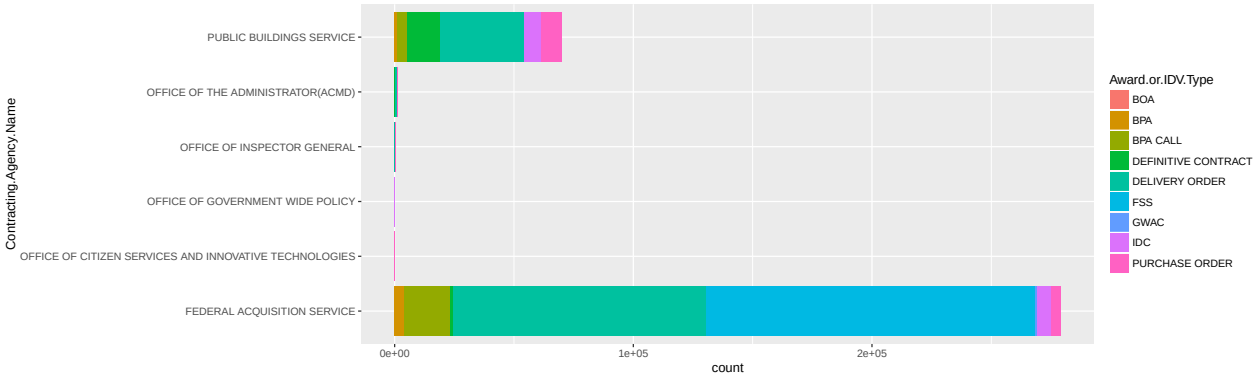
Modification.Number	Action.Obligation	Contracting.Office.Name
PA02	0	GSA/PBS/R03 REGIONALCONTRACTS SUPPORT SERVICES SECTION
PS67	11117	ACQUISITION DIV/PROGRAM SUPPORT&CAPITAL PROJECTS

Award.or.IDV.Type	Type.of.Contract	period	cat
FSS	FIRM FIXED PRICE	052014_062014	-GS23F0168L
DELIVERY ORDER	FIRM FIXED PRICE	102013_122013	GS07P11HHD0020-GSP0713UF5028
FSS	FIXED PRICE WITH ECONOMIC PRICE ADJUSTMENT	052015_062015	-GS02F0195X
DELIVERY ORDER	FIRM FIXED PRICE	102013_122013	GS07F9462G-GSP0812JA0040
FSS	FIXED PRICE WITH ECONOMIC PRICE ADJUSTMENT	102013_122013	-GS07F9428S
FSS	FIXED PRICE WITH ECONOMIC PRICE ADJUSTMENT	032014	-GS21F0085U
FSS	FIRM FIXED PRICE	012015_022015	-GS35F0347R
DELIVERY ORDER	FIRM FIXED PRICE	102013_122013	GS07P99HHD0080-GSP0713UT5079
FSS	FIRM FIXED PRICE	072015_082015	-GS35F0678P
IDC	FIXED PRICE WITH ECONOMIC PRICE ADJUSTMENT	052015_062015	-GS02Q14DCR0008
FSS	FIXED PRICE WITH ECONOMIC PRICE ADJUSTMENT	072015_082015	-GS07F0377X
FSS	FIRM FIXED PRICE	102015_112015	-GS00F0012S
FSS	FIRM FIXED PRICE	052015_062015	-GS35F0235Y
FSS	FIXED PRICE WITH ECONOMIC PRICE ADJUSTMENT	092014	-GS21F0117W
FSS	FIRM FIXED PRICE	102014_122014	-GS35F0638J
DELIVERY ORDER	FIRM FIXED PRICE	072015_082015	GS07P11HHD0004-GSP0714UC5013
FSS	FIRM FIXED PRICE	122015	-GS02F0105X
DELIVERY ORDER	FIRM FIXED PRICE	102015_112015	GS09P14KSD0022-GSP0915WB7007
DELIVERY ORDER	FIRM FIXED PRICE	012014_022014	GS03P11CDD0036-GSP0312DX5069
DEFINITIVE CONTRACT	FIRM FIXED PRICE	012015_022015	-GS04P11BVC0027

Next let's look at a few dimensions of the dataset.

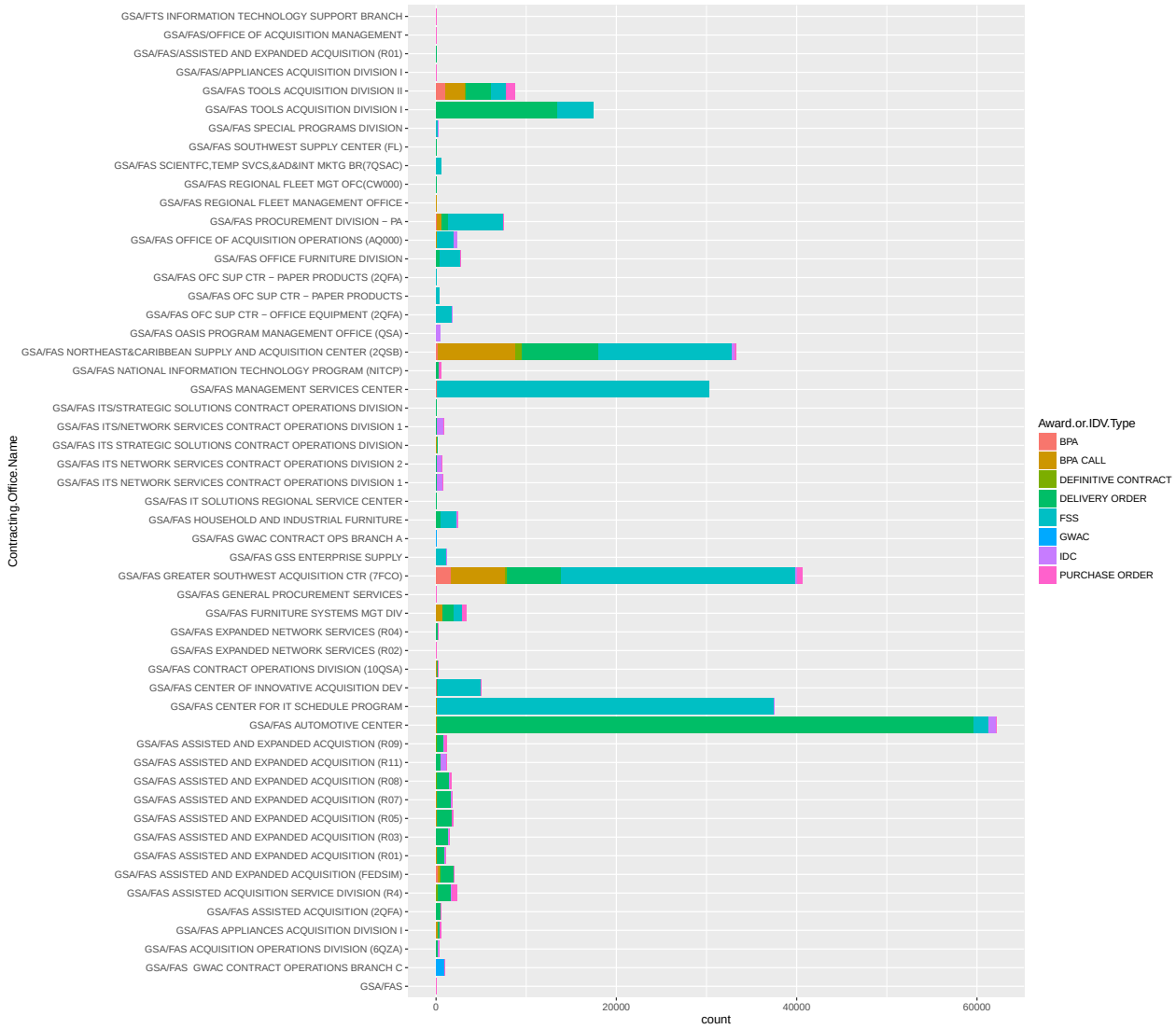
Start with the count of records by contracting Agency

```
# contracting agency

ggplot(fpds, aes(Contracting.Agency.Name)) + geom_bar(aes(fill = Award.or.IDV.Type)) +
    coord_flip()
```

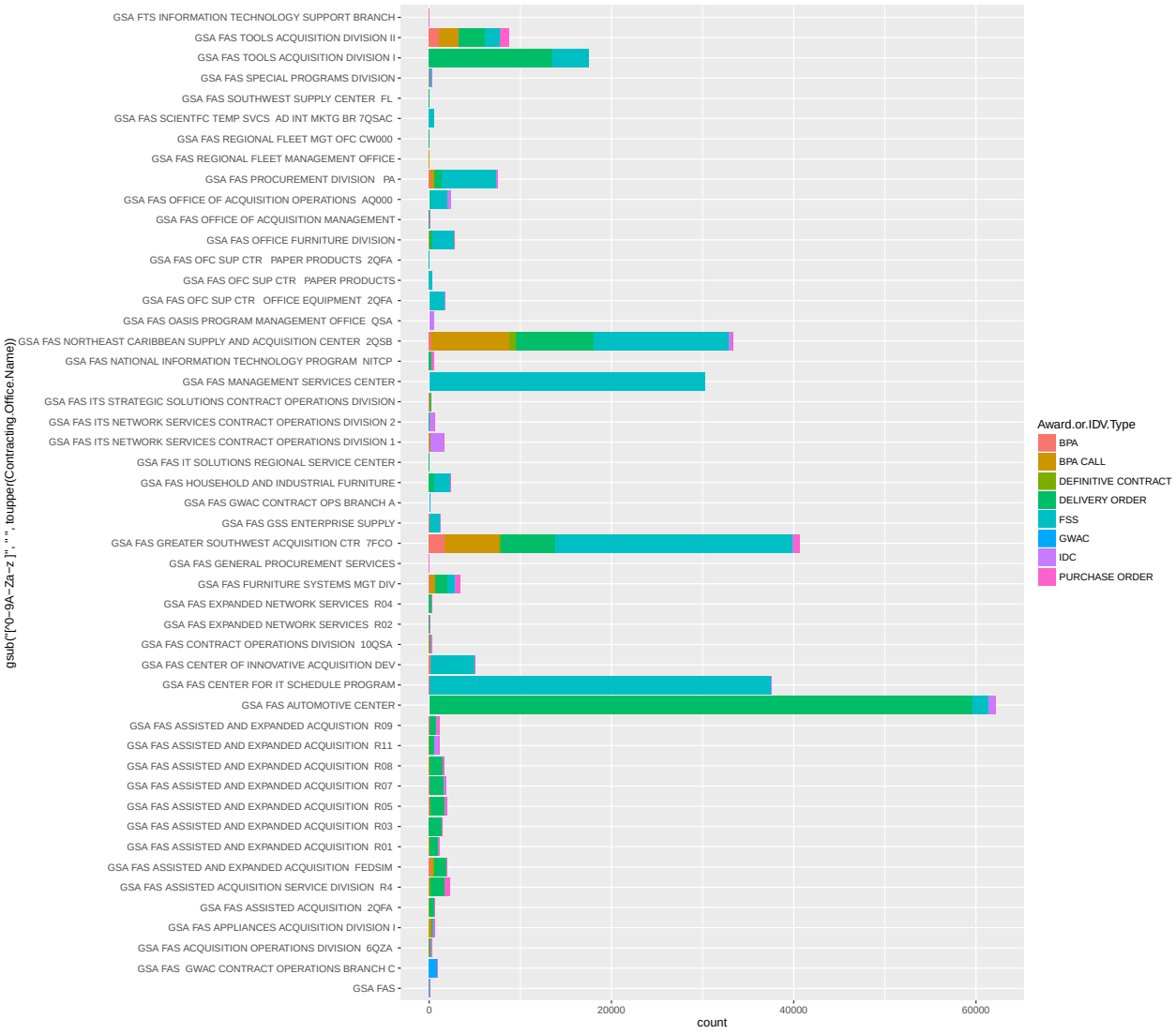


Drill down to the offices within the Federal Acquisition Service.

```
# contracting agency

ggplot(filter(fpds, Contracting.Agency.Name == "FEDERAL ACQUISITION SERVICE"),
    aes(Contracting.Office.Name)) + geom_bar(aes(fill = Award.or.IDV.Type)) +
    coord_flip()
```




```
# Let's try again, but this time strip the non-alphanumeric characters so
# that the contracting office names group even when one of them has a '/' in
# it.
ggplot(filter(fpds, Contracting.Agency.Name == "FEDERAL ACQUISITION SERVICE"),
    aes(gsub("[^0-9A-Za-z ]", " ", toupper(Contracting.Office.Name)))) + geom_bar(aes(fill = Award.or.I
    coord_flip()
```



Summary

NOTES:

1. Need **Fiscal Year** Column to be be pulled... can be reconstructed from filenames, but that is problematic if more than one month is being included.
2. If there is a **Month** Column, please also include that. Failing that, let's pull each month individually.
3. We may need these fields to be included in the next pull?

- `Extent.Competed`
- `Number.of.Offers.Received`
- `Other.Than.Full.and.Open.Competition`
- `Actions`
- `Number.of.Records`

Questions:

1. What belongs in the `Referenced.IDV.Mod.Number` field? Because there are alphanumeric strings and also two digit numbers.
2. Can we determine anything from the lengths of the strings in the `PIID` column? We have strings of the following lengths 7, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 33, 36, 37, 38.
3. What belongs in the `Modification.Number` field? We have a combination of numeric values (e.g. 0, 1, 10, 101, 1061, 11) and alphanumeric strings (e.g. S185, S184, S183, S182, S181, S180). Some of them appear to have words embedded inside of them (e.g. PS0025TEMPEXTMAS, PA0005REASSIGNMENT, PO0004TEMPEXTMAS, PA0034REASSIGNMENT, PS0043TEMPEXTMAS, PA0008REASSIGNMENT)

Ryan's Data Extraction

*This is largely unchanged from Ryan's approach. I adapted the `Obligation.Amount` field to be properly numeric and changed some of the aesthetics.

Take a look at how the extent competed variable is aggregated across mods. One goal is to understand how to sum up the competition variables such as the Number of Offers Received.

```
df <- read.csv("./data/GSA_Competition_Q1FY14.csv", stringsAsFactors = F)
names(df)
```

```
##  [1] "Modification.Number"
##  [2] "PIID"
##  [3] "PIID.Agency.ID"
##  [4] "Referenced..IDV.PIID"
##  [5] "Referenced.IDV.Agency.ID"
##  [6] "Referenced.IDV.Mod.Number"
##  [7] "Solicitation.ID"
##  [8] "Transaction.Number"
##  [9] "Date.Signed"
## [10] "Effective.Date"
## [11] "Est..Ultimate.Completion.Date"
## [12] "Fiscal.Year"
## [13] "Funding.Agency.ID"
## [14] "Funding.Agency.Name"
## [15] "Funding.Department.ID"
## [16] "Funding.Department.Name"
## [17] "Funding.Office.ID"
## [18] "Funding.Office.Name"
## [19] "Congressional.District.Place.of..Performance"
## [20] "Place.of.Performance.Zip.Code"
## [21] "Principal.Place.of.Performance.City.Name"
## [22] "Principal.Place.of.Performance.Country.Code"
## [23] "Principal.Place.of.Performance.Country.Name"
## [24] "Principal.Place.of.Performance.County.Code"
## [25] "Principal.Place.of.Performance.County.Name"
## [26] "Principal.Place.of.Performance.Location.Code"
## [27] "Principal.Place.of.Performance.State.Code"
## [28] "Country.of.Product.or.Service.Origin.Code"
## [29] "Alternative.Advertising"
## [30] "Commercial.Item.Acquisition.Procedures.Code"
```

```
## [31] "Commercial.Item.Acquisition.Procedures.Description"
## [32] "Commercial.Item.Test.Program"
## [33] "Commercial.Item.Test.Program.Description"
## [34] "Evaluated.Preference"
## [35] "Extent.Competed"
## [36] "Fair.Opportunity.Limited.Sources"
## [37] "FedBiz.Opps.Code"
## [38] "FedBiz.Opps.Description"
## [39] "Local.Area.Set.Aside.Code"
## [40] "Local.Area.Set.Aside.Description"
## [41] "Number.of.Offers.Received"
## [42] "Other.Than.Full.and.Open.Competition"
## [43] "Pre.Award.FBO.Synopsis"
## [44] "Price.Evaluation.Percent.Difference"
## [45] "SBA.or.OFPP.Synopsis.Waiver.Pilot"
## [46] "SBIR.STTR"
## [47] "Small.Business.Competitiveness..Demonstration.Program"
## [48] "Solicitation.Procedures"
## [49] "Type.of.Set.Aside"
## [50] "Award.or.IDV.Type"
## [51] "Reason.For.Modification.Code"
## [52] "Reason.For.Modification.Description"
## [53] "Congressional.District...Contractor"
## [54] "Contractor.Name"
## [55] "Doing.Business.As.Name"
## [56] "DUNS.Number"
## [57] "Global.DUNS.Number"
## [58] "Global.Vendor.Name"
## [59] "Vendor.City"
## [60] "Vendor.Country"
## [61] "Vendor.Name"
## [62] "SAM.Exception"
## [63] "Street"
## [64] "Zip"
## [65] "Contracting.Agency.ID"
## [66] "Contracting.Agency.Name"
## [67] "Contracting.Department.ID"
## [68] "Contracting.Department.Name"
## [69] "Contracting.Office.ID"
## [70] "Contracting.Office.Name"
## [71] "Contracting.Office.Region"
## [72] "Base.and.Exercised.Options.Value"
## [73] "Action.Obligation"
## [74] "Base.and.All.Options.Value"
## [75] "Actions"
## [76] "Number.of.Records"
```

```
# select variables for exploration
dfs <- select(df, Referenced..IDV.PIID, PIID, Referenced.IDV.Mod.Number, Modification.Number,
    Date.Signed, Extent.Competed, Number.of.Offers.Received, Other.Than.Full.and.Open.Competition,
    Action.Obligation, Actions, Number.of.Records, Contracting.Office.Name,
    Award.or.IDV.Type, Contracting.Agency.Name)

dfs$Action.Obligation <- gsub("[^0-9.]", "", dfs$Action.Obligation) %>% as.numeric()
```

```
dfs <- mutate(dfs, cat = paste(Referenced..IDV.PIID, PIID, sep = "-"))

test1 <- filter(dfs, Number.of.Records > 1 | Actions > 1)


# group variables by PIIDs
dfg <- group_by(dfs, cat)

test2 <- dfg %>% summarize(cnt = n_distinct(cat)) %>% arrange(desc(cnt)) %>%
    filter(cnt > 1)

head(test2)
```

```
## Source: local data frame [0 x 2]
##
## Variables not shown: cat (chr), cnt (int)
```

```
# does any PIID have a mod in this set?
test3 <- filter(dfs, Modification.Number != 0)
pander(tbl_df(sample_n(test3, 20)), split.cells = 15)
```

Table 4: Table continues below

Referenced..IDV.PIID	PIID	Referenced.IDV.Mod.Number
	GS10F0315P	
	GS00F0011N	
	GS02F0009S	
	GS03P10QKC0069	0
	GS00I10AAC0042058	0
SP060011D4005	GSP1013XJ0009	0
	GS02F0233X	
	GS14F0043M	
GS07P10HHD0011	GSP0713HH5042	PS01
GS00Q12NSD0020	GST0513BM0114	0
	GS35F0504W	
	GS35F0251V	
	GS35F0673J	
	GS35F140AA	
	GS10F078AA	
	GS07F9160D	
	GS07F5514P	
	GS10F0252X	
	GS35F0265X	
	GS02F180AA	

Table 5: Table continues below

Modification.Number	Date.Signed	Extent.Competed
PA0010REASSIGNMENT	12/21/13	FULL AND OPEN COMPETITION

Modification.Number	Date.Signed	Extent.Competed
PO0017CANCELCONT	10/3/13	FULL AND OPEN COMPETITION
PO0067DELPROD	12/12/13	FULL AND OPEN COMPETITION
PO03	12/11/13	COMPETED UNDER SAP
1	10/31/13	NOT COMPETED UNDER SAP
PC08	10/28/13	FULL AND OPEN COMPETITION
PA0003REASSIGNMENT	11/27/13	FULL AND OPEN COMPETITION
PO0110DELPROD	10/31/13	FULL AND OPEN COMPETITION
PS01	11/12/13	FULL AND OPEN COMPETITION
AA002	11/14/13	FULL AND OPEN COMPETITION
PO0017ADDPROD	10/9/13	FULL AND OPEN COMPETITION
PO0038ADDPROD	11/8/13	FULL AND OPEN COMPETITION
PA0022REASSIGNMENT	12/19/13	FULL AND OPEN COMPETITION
PA0003REASSIGNMENT	11/26/13	FULL AND OPEN COMPETITION
PA0003REASSIGNMENT	10/22/13	FULL AND OPEN COMPETITION
PO0093ADDPROD	10/24/13	FULL AND OPEN COMPETITION
PS0047DELPROD	12/20/13	FULL AND OPEN COMPETITION
PA0003REASSIGNMENT	11/20/13	FULL AND OPEN COMPETITION
PS1007ADDPROD	11/11/13	FULL AND OPEN COMPETITION
PO0003AUTHNEG	10/10/13	FULL AND OPEN COMPETITION

Table 6: Table continues below

Number.of.Offers.Received	Other.Than.Full.and.Open.Competition
0	
51	
0	
3	
1	AUTHORIZED BY STATUTE
13	
999	
1	
13	

Number.of.Offers.Received	Other.Than.Full.and.Open.Competition
6	
NA	
0	
999	
999	
999	
2	
0	
999	
999	
999	

Table 7: Table continues below

Action.Obligation	Actions	Number.of.Records	Contracting.Office.Name
0.00	1	1	GSA/FAS MANAGEMENT SERVICES CENTER
0.00	1	1	GSA/FAS MANAGEMENT SERVICES CENTER
0.00	1	1	GSA/FAS NORTHEAST&CARIBBEAN SUPPLY AND ACQUISITION CENTER (2QSB)
1260.00	1	1	GSA/PBS/R03 PITTSBURG FO
0.00	1	1	THE INTERNAL ACQUISITION DIVISION (IAD)
2160.58	1	1	GSA/PBS
0.00	1	1	GSA/FAS NORTHEAST&CARIBBEAN SUPPLY AND ACQUISITION CENTER (2QSB)
0.00	1	1	GSA/FAS OFC SUP CTR - OFFICE EQUIPMENT (2QFA)
0.00	1	1	SPECIAL PROGRAMS DIVISION - LPOE
0.00	1	1	GSA/FAS ASSISTED AND EXPANDED ACQUISITION (R05)
0.00	1	1	GSA/FAS CENTER FOR IT SCHEDULE PROGRAM
0.00	1	1	GSA/FAS CENTER FOR IT SCHEDULE PROGRAM
0.00	1	1	GSA/FAS CENTER FOR IT SCHEDULE PROGRAM
0.00	1	1	GSA/FAS CENTER FOR IT SCHEDULE PROGRAM
0.00	1	1	GSA/FAS MANAGEMENT SERVICES CENTER

Action.Obligation	Actions	Number.of.Records	Contracting.Office.Name
0.00	1	1	GSA/FAS GREATER SOUTHWEST ACQUISITION CTR (7FCO)
0.00	1	1	GSA/FAS GREATER SOUTHWEST ACQUISITION CTR (7FCO)
0.00	1	1	GSA/FAS MANAGEMENT SERVICES CENTER
0.00	1	1	GSA/FAS CENTER FOR IT SCHEDULE PROGRAM
0.00	1	1	GSA/FAS NORTHEAST&CARIBBEAN SUPPLY AND ACQUISITION CENTER (2QSB)

Award.or.IDV.Type	Contracting.Agency.Name	cat
FSS	FEDERAL ACQUISITION SERVICE	-GS10F0315P
FSS	FEDERAL ACQUISITION SERVICE	-GS00F0011N
FSS	FEDERAL ACQUISITION SERVICE	-GS02F0009S
DEFINITIVE CONTRACT	PUBLIC BUILDINGS SERVICE	-GS03P10QKC0069
PURCHASE ORDER	OFFICE OF THE ADMINISTRATOR(ACMD)	-GS00I10AAC0042058
DELIVERY ORDER	PUBLIC BUILDINGS SERVICE	SP060011D4005-GSP1013XJ0009
FSS	FEDERAL ACQUISITION SERVICE	-GS02F0233X
FSS	FEDERAL ACQUISITION SERVICE	-GS14F0043M
DELIVERY ORDER	PUBLIC BUILDINGS SERVICE	GS07P10HHD0011-GSP0713HH5042
DELIVERY ORDER	FEDERAL ACQUISITION SERVICE	GS00Q12NSD0020-GST0513BM0114
FSS	FEDERAL ACQUISITION SERVICE	-GS35F0504W
FSS	FEDERAL ACQUISITION SERVICE	-GS35F0251V
FSS	FEDERAL ACQUISITION SERVICE	-GS35F0673J
FSS	FEDERAL ACQUISITION SERVICE	-GS35F140AA
FSS	FEDERAL ACQUISITION SERVICE	-GS10F078AA
FSS	FEDERAL ACQUISITION SERVICE	-GS07F9160D
FSS	FEDERAL ACQUISITION SERVICE	-GS07F5514P
FSS	FEDERAL ACQUISITION SERVICE	-GS10F0252X

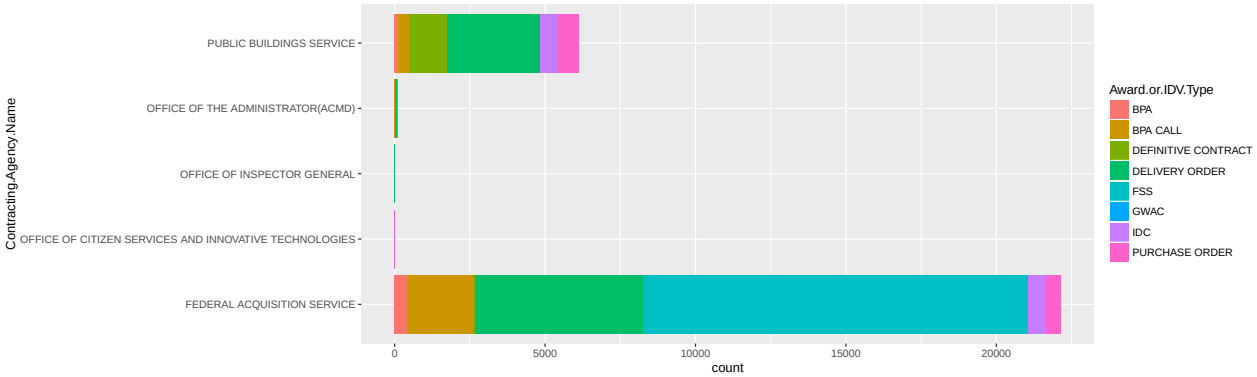
Award.or.IDV.Type	Contracting.Agency.Name	cat
FSS	FEDERAL ACQUISITION SERVICE	-GS35F0265X
FSS	FEDERAL ACQUISITION SERVICE	-GS02F180AA

These initial tests of aggregation show that all PIID-pair (award + reference IDV) entries are unique, i.e. none have more than one entry in the dataset. This is surprising because the same award tends to have multiple modifications. Assume that each entry is unique because of the time period. Leaves unresolved the question of whether to sum the number of competitors PIIDs or take the max, will default to the latter.

Next let's look at a few dimensions of the dataset.

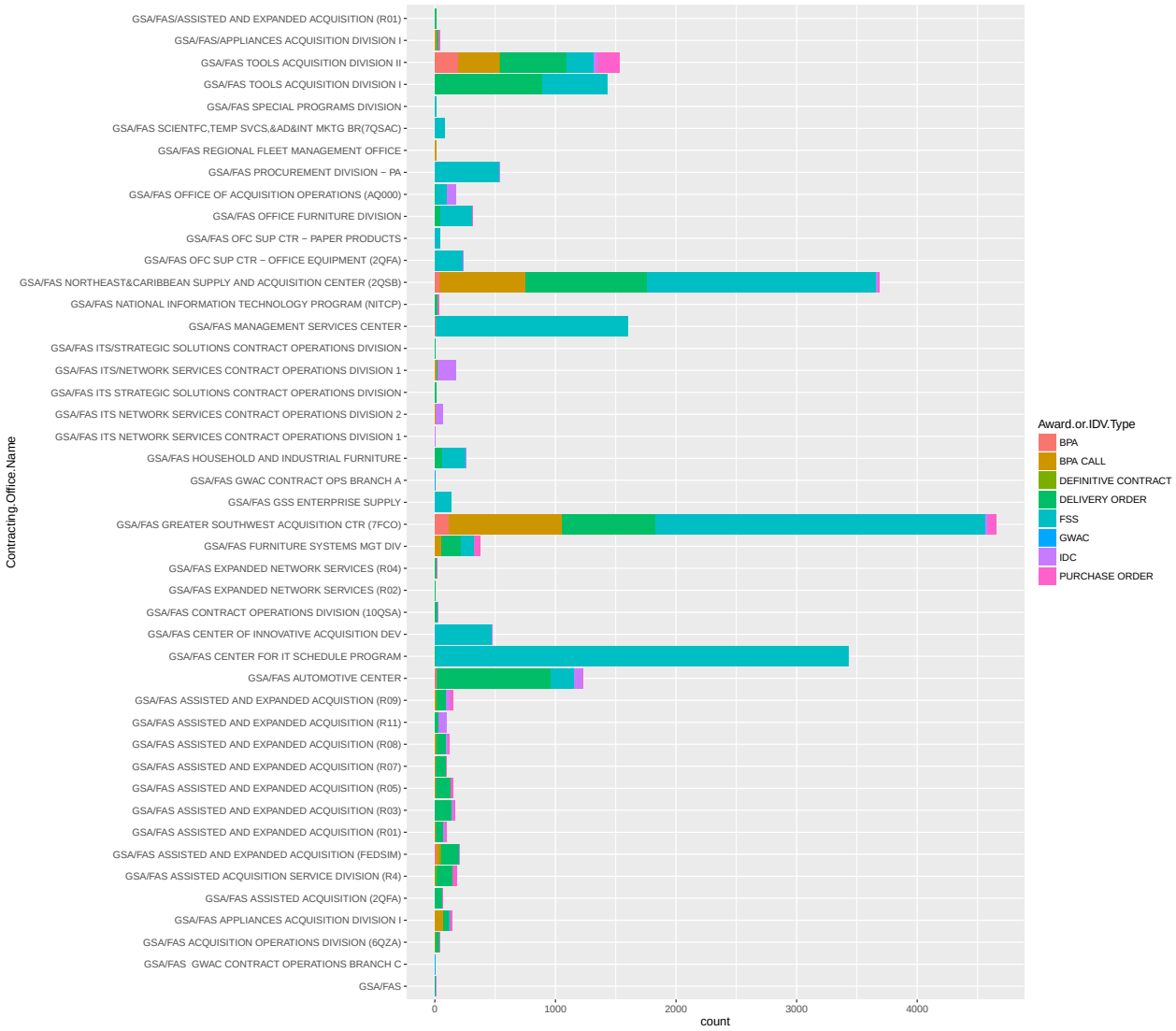
Start with the count of records by contracting Agency

```
# contracting agency

ggplot(dfs, aes(Contracting.Agency.Name)) + geom_bar(aes(fill = Award.or.IDV.Type)) +
    coord_flip()
```



Drill down to the offices within the Federal Acquisition Service.

```
# contracting agency
ggplot(filter(dfs, Contracting.Agency.Name == "FEDERAL ACQUISITION SERVICE"),
    aes(Contracting.Office.Name)) + geom_bar(aes(fill = Award.or.IDV.Type)) +
    coord_flip()
```



Looks like Federal Supply Schedules (FSS) make up the majority of the contract actions. Believe these are a vehicle and so too BPA and IDC, probably want to add category for Award or vehicle type. First, let's get the overall percentages by type.

```
pct <- dfs %>% group_by(Award.or.IDV.Type) %>% summarise(cnt = n(), value = sum(Action.Obligation)) %>%
    mutate(propCnt = cnt/sum(cnt), propSum = value/sum(value))

pander(tbl_df(pct), split.cells = 10)
```

Award.or.IDV.Type	cnt	value	propCnt	propSum
BPA	518	335673.6	0.0182529335	0.0002612679
BPA CALL	2618	98202013.4	0.0922513126	0.0764344819
DEFINITIVE CONTRACT	1302	191924816.0	0.0458789950	0.1493826183
DELIVERY ORDER	8742	858717504.5	0.3080446809	0.6683735424
FSS	12781	0.0	0.4503682300	0.0000000000
GWAC	4	0.0	0.0001409493	0.0000000000
IDC	1192	73195010.4	0.0420028895	0.0569705498

Award.or.IDV.Type	cnt	value	propCnt	propSum
PURCHASE ORDER	1222	62411781.6	0.0430600092	0.0485775396

Seventy five percent of the records are FSS (45%) or Delivery Order (30%). Seventy percent of the funding comes from Delivery Order (50%) and BPA Calls (19%).

While we're looking at percentages let's take a similar look at the break down of competition types.

```
pct <- dfs %>% group_by(Extent.Competed) %>% summarise(cnt = n(), value = sum(Action.Obligation)) %>%
    mutate(propCnt = cnt/sum(cnt), propSum = value/sum(value))

pander(tbl_df(pct), hyphenate = T, split.cells = 15)
```

Extent.Competed	cnt	value	propCnt	propSum
	279	4945940.5	0.0098312132	0.0038496196
COMPETED UNDER SAP	1408	16556507.9	0.0496141513	0.0128865800
COMPETITIVE DELIVERY ORDER	107	39616089.2	0.0037703936	0.0308347573
FOLLOW ON TO COMPETED ACTION	8	771323.5	0.0002818986	0.0006003514
FULL AND OPEN COMPETITION	21509	865256038.7	0.7579195884	0.6734627403
FULL AND OPEN COMPETITION AFTER EXCLUSION OF SOURCES	2162	173218617.7	0.0761830931	0.1348228498
NON-COMPETITIVE DELIVERY ORDER	7	3620636.7	0.0002466613	0.0028180837
NOT AVAILABLE FOR COMPETITION	2010	132844255.9	0.0708270200	0.1033978991
NOT COMPETED	256	27726767.7	0.0090207548	0.0215808317
NOT COMPETED UNDER SAP	633	20230621.8	0.0223052257	0.0157462871

Seventy five percent of the records are listed as completely full and open. Only sixty six percent when weighted by dollars.

Now we should start to explore patterns in the number of offers...