Analysis - Gender Equity

Chelsea Ursaner

February 2, 2017

Background: EXECUTIVE DIRECTIVE ON GENDER EQUITY IN CITY OPERATIONS

Objective: Analyze 2/2 employee database extract to identify areas of gender bias in City gov hiring based on Executive Directive areas of focus.

Read data

```
setwd("~/R_Datasets/DataLA")
data <- read.csv("employee_demographics.csv")
```

Data Prepping

```
#See a sample of the data, column formats, and overall size of the dataset
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
```

```
glimpse(data)
```

```
## Observations: 47,557
## Variables: 23
## $ DEPT.No            (int) 40, 94, 38, 88, 70, 70, 38, 38, 70, 88, 8, ...
## $ DEPT_DESC          (fctr) General Services, Transportation, Fire, Re...
## $ PAYROLL_DEPT.NO     (int) 3371, 8501, 3001, 7903, 4301, 4301, 3001, 3...
## $ PAYROLL_DEPT_DESC  (fctr) GSD STANDARDS, TRANSPORTATION, FIRE - FIRE...
## $ DIVISION_NBR       (int) 160, 613, 3, 331, 411, 180, 41, 311, 180, 1...
## $ DIVISION_NAME      (fctr) , , , , , , , , , , , , , , , , , , , , , , ...
## $ RECORD_ID          (dbl) 237161, 287426, 383335, 424248, 591587, 599...
## $ JOB_CLASS          (int) 7967, 3430, 2112, 2498, 2214, 2214, 2112, 1...
## $ PAY_GRADE          (fctr) 1, 1, 2, 0, 3, 2, 3, 2, 2, A, 0, 3, 0, A, ...
## $ JOB_CLASS_TITLE    (fctr) MATL TST ENGRG ASSC I, TRAF MARK/SIGN SUPT...
## $ LEGACY_JOB_STATUS   (int) 11, 11, 11, 17, 11, 11, 11, 11, 11, 7, 11, ...
## $ LEGACY_DESCRIPTION (fctr) Full-Time Regular, Full-Time Regular, Full...
## $ GENDER             (fctr) Male, Male, Male, Female, Male, Male, Male...
## $ ETHNICITY          (fctr) Filipino, Black, Caucasian, Black, Caucasi...
## $ ORIGINAL_HIRE_DATE (fctr) 11/29/99, 2/3/85, 6/27/16, 9/25/91, 1/2/90...
## $ EMPLOYMENT_DATE    (fctr) 11/29/99, 2/3/85, 6/27/16, 9/25/91, 1/2/90...
## $ MOU                (int) 8, 12, 23, 7, 24, 24, 23, 1, 20, 7, 5, 24, ...
## $ MOU_DESCRIPTION    (fctr) PROFES.ENGR. & SCIENTIF, SUPV BLUE COLLAR,...
## $ EEO                (fctr) B, G, D, B, D, D, D, B, F, B, G, D, G, H, ...
## $ EEO_DESCRIPTION    (fctr) Professionals, Skilled Craft, Protective S...
## $ HOURLY_BASE_RATE   (fctr) , , ,      $16.50, , , , , ,      $17.25, ...
```

```
## $ BIWEEKLY_BASE_RATE (fctr)    $3,328.80,   $3,083.20,   $2,540.80, , ...
## $ RATE_TYPE          (fctr) Biweekly, Biweekly, Biweekly, Hourly, Biwe...
#perform transformations to improve dataset
#filter to only full time employees
data <- filter(data, LEGACY_DESCRIPTION == "Full-Time Regular")
#remove the dollar sign from hourly base rate so R reads it as numeric
library(stringr)
data$HOURLY_BASE_RATE <- as.numeric(str_replace_all(data$HOURLY_BASE_RATE, "\\$", ""))
library(lubridate) #fix date formats as needed
```

Warning: package 'lubridate' was built under R version 3.2.3

```
data$ORIGINAL_HIRE_DATE <- mdy(data$ORIGINAL_HIRE_DATE)
data$EMPLOYMENT_DATE <- mdy(data$EMPLOYMENT_DATE)
```

Enrichment Create new groupings/variables in order to understand female representation in different jobs

```
#Aggregating to look at female representation by job (3 options)
#get counts of females and males - looking at just representation
data$Female <- ifelse(data$GENDER == "Female",1,0)
data$Male <- ifelse(data$GENDER == "Male",1,0)
```

Exploration and Visualization How do the findings change depending on how I define a job?

```
#Option 1 - Job and department
jobs1 <- data %>% group_by(PAYROLL_DEPT_DESC, JOB_CLASS_TITLE) %>%
  summarize(CountF = sum(Female),
            CountM = sum(Male),
            Total = (sum(Female) + sum(Male)),
            PercentFemale = CountF/Total)
library(ggplot2)
```

Warning: package 'ggplot2' was built under R version 3.2.3

```
sum(data$Female)/nrow(data)
```

[1] 0.2840202

```
p1 <- qplot(jobs1$PercentFemale,
            main = "Level of detail: Job and department. I.e. Mgmt assistant in LAWA",
            ylab = "Number of Records",
            xlab = "Percent Female")


#Option 2 - Job and department, filtering to jobs that have more than 8 people
jobs2 <- filter(jobs1, Total > 5)
p2 <- qplot(jobs2$PercentFemale,
            main = "Same as #1 but filtered to jobs that have >8 people",
            ylab = "Number of Records",
            xlab = "Percent Female")

#Option 3 - just job class title and filtering to > 10 people
jobs3 <- data %>% group_by(JOB_CLASS_TITLE) %>%
  summarize(CountF = sum(Female),
            CountM = sum(Male),
            Total = (sum(Female) + sum(Male)),
            PercentFemale = CountF/Total) %>%
```

```
   filter(Total > 10)

p3 <- qplot(jobs3$PercentFemale,
        main = "LOD Job Class (i.e. Mgmt Assistant) filtered to classes with >10 people",
        ylab = "Number of Records",
        xlab = "Percent Female")

library(gridExtra)
grid.arrange(p1, p2, p3)
```
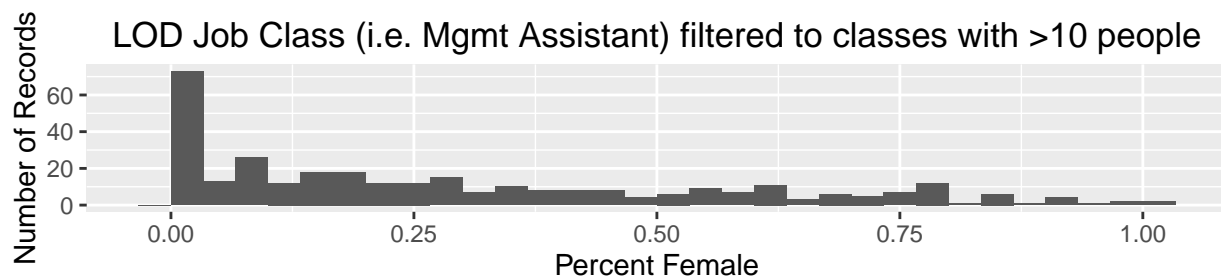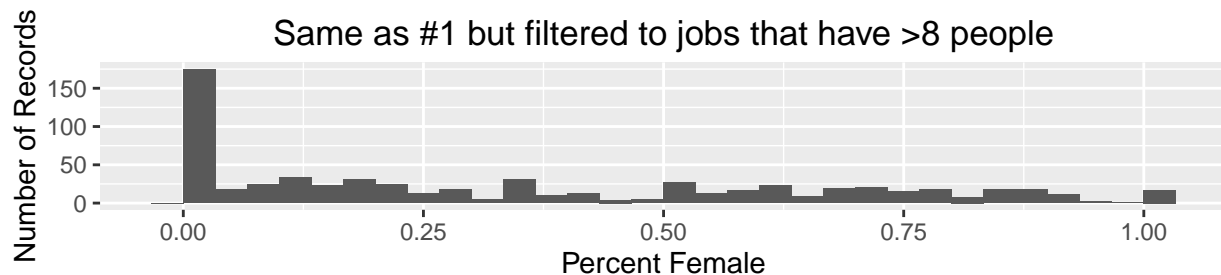
```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```
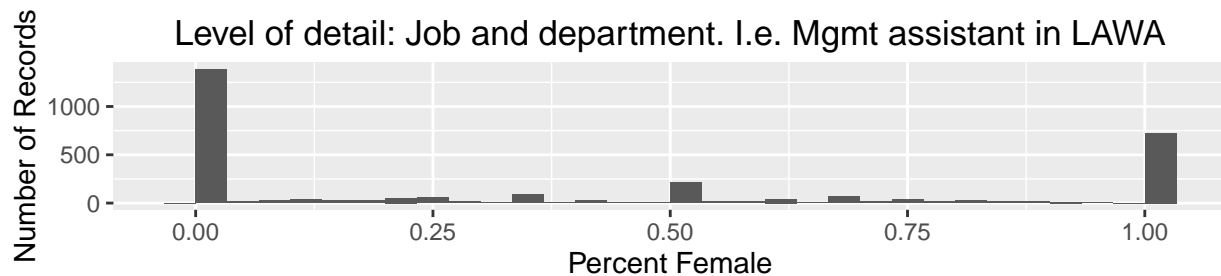

Results/Next Steps Creating two lists - one of jobs that are heavily male-dominated and another of jobs that are heavily female-dominated. From here, I'll work with the

```
#Using option 3, get subsets of gender biased jobs
male_bias <- jobs3 %>% filter(PercentFemale < .1)
female_bias <- jobs3 %>% filter(PercentFemale > .9)

#Display as tables
library(knitr)
```

```
## Warning: package 'knitr' was built under R version 3.2.5
```

```
kable(female_bias, digits = 2, caption = 'Female-Dominated Jobs (where women represent > 90%)')
```

Table 1: Female-Dominated Jobs (where women represent > 90%)

JOB_CLASS_TITLE	CountF	CountM	Total	PercentFemale
CH CLERK	11	0	11	1.00
EXEC ADMIN ASST II	57	1	58	0.98
EXEC ADMIN ASST III	31	0	31	1.00
PAYROLL SUPERVISOR I	19	2	21	0.90
PERS RECORDS SUPV	21	1	22	0.95
SECRETARY	128	3	131	0.98
SR ADMIN ANALYST I	10	1	11	0.91
WORKERS COMP CLAIMS AST	13	1	14	0.93

```
kable(male_bias, digits = 2, caption = 'Male-Dominated Jobs (where men represent > 90%)')
```

Table 2: Male-Dominated Jobs (where men represent > 90%)

JOB_CLASS_TITLE	CountF	CountM	Total	PercentFemale
AIR COND MECHANIC	0	46	46	0.00
AIRPORT SAFETY OFFICER II	0	37	37	0.00
AIRPORTS MTCE SUPVR III	1	17	18	0.06
APPARATUS OPERATOR	2	136	138	0.01
ARCHITECT	1	11	12	0.08
ARPT SUPT OF OPER I	3	30	33	0.09
ARPT SUPT OF OPER II	3	38	41	0.07
ASST SIGNL SYS ELECTRCN	0	12	12	0.00
ASST ST LTG ELECTRCN	0	22	22	0.00
AUTO BODY BLDR/REPAIRER	0	22	22	0.00
AUTOMOTIVE SUPERVISOR	0	21	21	0.00
BUILD INSPECTOR	1	44	45	0.02
BUILD MECH INSPECTOR	0	144	144	0.00
BUILD OPERATING ENGR	2	34	36	0.06
BUILDING REPAIRER I	1	13	14	0.07
CARPENTER	0	70	70	0.00
CEMENT FINISHER	0	75	75	0.00
CH OF OPERATIONS I	1	11	12	0.08
CHEMIST I	1	10	11	0.09
COMMUN ELECTRICIAN	3	99	102	0.03
CONSTR INSPECTOR	6	153	159	0.04
DELIVERY DRIVER II	0	12	12	0.00
ELECTRCL CRAFT HELPER	1	60	61	0.02
ELECTRCL ENGRG ASSC I	1	12	13	0.08
ELECTRCL INSPECTOR	0	15	15	0.00
ELECTRICIAN	0	76	76	0.00
ELEVATOR MECHANIC	0	22	22	0.00
ELEVATOR MECHANIC HLPR	0	15	15	0.00
ENGINEER OF FIRE DEPT	6	398	404	0.01
EQUIP REPAIR SUPVR	0	12	12	0.00
EQUIPMNT MECHANIC	0	297	297	0.00
EQUIPMNT OPERATOR	1	144	145	0.01
FIELD ENGINEER AIDE	2	38	40	0.05
FIRE ASSISTANT CHIEF	1	16	17	0.06
FIRE BATTALION CHIEF	6	59	65	0.09

JOB_CLASS_TITLE	CountF	CountM	Total	PercentFemale
FIRE CAPTAIN I	10	339	349	0.03
FIRE CAPTAIN II	4	156	160	0.02
FIRE INSPECTOR II	3	28	31	0.10
FIREBOAT MATE	0	12	12	0.00
FIREFIGHTER II	14	307	321	0.04
FIREFIGHTER III	32	1495	1527	0.02
GARAGE ATTENDANT	0	59	59	0.00
HEAVY DUTY EQUIP MECH	0	112	112	0.00
HEAVY DUTY TRUCK OPER	4	94	98	0.04
HELICOPTER MECH	1	25	26	0.04
HOUSING INSPECTOR	6	82	88	0.07
INSTRUMENT MECH	0	34	34	0.00
IRRIGATION SPECIALIST	1	35	36	0.03
LAND SURVEYING ASST	0	24	24	0.00
LIGHT EQUIP OPERATOR	0	34	34	0.00
MAINT & CONSTR HELPER	1	68	69	0.01
MECH HELPER	0	38	38	0.00
MECH REPAIRER	0	25	25	0.00
MOTOR SWEEPER OPERATOR	8	93	101	0.08
MUNICIPAL POLICE OFFICER III	2	21	23	0.09
PAINTER	0	49	49	0.00
PARKING MTR TECHNICIAN	0	20	20	0.00
PIPEFITTER	0	12	12	0.00
PLUMBER	1	76	77	0.01
PLUMBER SUPERVISOR	1	15	16	0.06
PLUMBING INSPECTOR	0	15	15	0.00
POL SURVLLNCE SPEC I	0	13	13	0.00
PORT ELECTRICAL MECHANIC	0	21	21	0.00
PORT POLICE SERGEANT	0	18	18	0.00
PR CONSTR INSPECTOR	1	18	19	0.05
PR INSPECTOR	1	23	24	0.04
REF COLL TRUCK OPER II	14	628	642	0.02
ROOFER	0	25	25	0.00
SAFETY ENGR ELEVATORS	0	14	14	0.00
SIGNAL SYSTEM ELECTRCN	1	78	79	0.01
SR BUILD INSPECTOR	1	44	45	0.02
SR BUILD MECH INSPECTR	0	41	41	0.00
SR CARPENTER	0	12	12	0.00
SR COMMUN ELECTRICIAN	1	12	13	0.08
SR COMMUN ENGINEER	1	11	12	0.08
SR CONSTR INSPECTOR	2	63	65	0.03
SR CUSTODIAN II	1	13	14	0.07
SR ELECTRCL INSPECTOR	2	25	27	0.07
SR EQUIPMENT MECHANIC	0	28	28	0.00
SR HOUSING INSPECTOR	3	29	32	0.09
SR PLUMBING INSPECTOR	1	10	11	0.09
SR STRUCTURAL ENGINEER	0	11	11	0.00
ST LTG ELECTRCN	0	30	30	0.00
ST LTG ENGRG ASSC I	1	13	14	0.07
ST SVCS SUPT I	1	10	11	0.09
ST SVCS SUPVR I	1	53	54	0.02
ST SVCS SUPVR II	0	24	24	0.00

JOB_CLASS_TITLE	CountF	CountM	Total	PercentFemale
ST SVCS WORKER I	4	60	64	0.06
ST SVCS WORKER II	2	19	21	0.10
STOREKEEPER II	8	75	83	0.10
SURVEY PARTY CHIEF I	0	17	17	0.00
SURVEY PARTY CHIEF II	0	14	14	0.00
TRAF PAINT SIGN POST I	0	23	23	0.00
TRAF PAINT SIGN POST II	0	30	30	0.00
TRAF PNT SIGN POST III	0	12	12	0.00
TRANSP ENGINEER	3	34	37	0.08
TREE SURGEON	1	54	55	0.02
TREE SURGEON ASST	1	10	11	0.09
TREE SURGEON SUPVSR I	0	28	28	0.00
TRUCK OPERATOR	1	65	66	0.02
W/WTR COLL SUPERVISOR	0	15	15	0.00
W/WTR COLL WORKER I	0	43	43	0.00
W/WTR COLL WORKER II	2	136	138	0.01
W/WTR TRMT ELEC I	0	25	25	0.00
W/WTR TRMT MECH	0	52	52	0.00
W/WTR TRMT OPER I	1	63	64	0.02
W/WTR TRMT OPER II	2	23	25	0.08
W/WTR TRMT OPER III	1	29	30	0.03
WAREHOUSE & T/R WKR I	2	21	23	0.09
WAREHOUSE & T/R WKR II	0	28	28	0.00
WELDER	0	38	38	0.00
WINDOW CLEANER - AIRPORT	0	13	13	0.00

```r
#as new files
write.csv(male_bias, "male_bias.csv", row.names = F)
write.csv(female_bias, "female_bias.csv", row.names = F)
```