Lost Mail Appendix

JL, ZB, OX, WM, and ED

4/13/17

Data Import + Cleaning

```
datainfo <- read.table("http://pmatheson.people.amherst.edu/c62.surinfo.dat")
datasurvey <- read.table("http://pmatheson.people.amherst.edu/c62.survey.dat")
colnames(datasurvey) <- c("ID", "Q1", "Q2", "Q3a", "Q3b", "Q3c", "Q4") #change column names
colnames(datainfo) <- c("i", "Group", "ID", "NumberMag", "State")
# merge datasets by ID number
alldata <- merge(x = datasurvey, y = datainfo, by = "ID", all.x = TRUE)
# omit NA's
alldata <- na.omit(alldata)
```

Dataset #1: Only considering those who responded to Question 3a

```
# total magazines in dataset
totalmag <- as.numeric(sum(alldata$NumberMag))

# filter for Q3a responses only
minidata <- filter(alldata, Q3a != ".")
percentdel <- as.vector(minidata$Q3a)
percentdel <- as.numeric(percentdel)/100
numbermag <- as.vector(minidata$NumberMag)
numbermag <- as.numeric(numbermag)
groups <- as.vector(minidata$Group)

numberdel <- percentdel * numbermag
totaldel <- sum(numberdel)
percentdelivered <- totaldel/totalmag
# % delivered correctly = 5.41319
#percentdelivered*100
```

Finding percent of mail delivered for each group for Dataset #1

```
bygroups <- as.data.frame(cbind(groups,percentdel, numbermag))
bygroupsdel <- function(){
  for (i in 1:23){
    if (bygroups$groups[i] == 1){ #group 1
      group1 <- 0
      group1temp <- bygroups$percentdel[i] * bygroups$numbermag[i]
      group1 <- group1 + group1temp
    }
    if (bygroups$groups[i] == 2){ #group 2
      group2 <- 0
      group2temp <- bygroups$percentdel[i] * bygroups$numbermag[i]
      group2 <- group2 + group2temp
    }
    #no group 3 because there were no "a" responses to question 3 for group 3
    if (bygroups$groups[i] == 4){ #group 4
      group4 <- 0
      group4temp <- bygroups$percentdel[i] * bygroups$numbermag[i]
```

```
      group4 <- group4 + group4temp
    }
    if (bygroups$groups[i] == 5){ #group 5
      group5 <- 0
      group5temp <- bygroups$percentdel[i] * bygroups$numbermag[i]
      group5 <- group5 + group5temp
    }
  }
  all <- c(group1, group2, 0, group4, group5)
  return (all)
}
totalminimag <- sum(bygroups$numbermag)
PercentDelivered <- bygroupsdel()/totalminimag
Group <- c(1:5)
groupsdel <- as.data.frame(cbind(Group, PercentDelivered *100))
colnames(groupsdel) <- c("Group", "PercentDelivered")
#xtable(groupsdel,caption="Worst-Case Scenario")
```

Dataset #2: Considering those who answered "No" to Question 1 as fully delivering all mail correctly plus those who answered Question 3a

```
q1no <- filter(alldata, Q1 == 0)
q1no$Q3a[q1no$Q3a == "."] = 100
biggerdata <- rbind(q1no, minidata)
percentdel <- as.vector(biggerdata$Q3a)
percentdel <- as.numeric(percentdel)/100
numbermag <- as.vector(biggerdata$NumberMag)
numbermag <- as.numeric(numbermag)
new <- percentdel * numbermag
new <- sum(new)
groups <- as.vector(biggerdata$Group)
# % delivered correctly = 97.9448
#new/sum(biggerdata$NumberMag) * 100
```

```
bygroups2 <- as.data.frame(cbind(groups,percentdel, numbermag))
new <- bygroups2$percentdel * bygroups2$numbermag
bygroups2 <- cbind(bygroups2, new)
#breaking up % delivered correctly by group
group1 <- filter(bygroups2, groups == 1)
group2 <- filter(bygroups2, groups == 2)
group3 <- filter(bygroups2, groups == 3)
group4 <- filter(bygroups2, groups == 4)
group5 <- filter(bygroups2, groups == 5)
group1tot <- sum(group1$new)/sum(group1$numbermag)
group2tot <- sum(group2$new)/sum(group2$numbermag)
group3tot <- sum(group3$new)/sum(group3$numbermag)
group4tot <- sum(group4$new)/sum(group4$numbermag)
group5tot <- sum(group5$new)/sum(group5$numbermag)
all <- c(group1tot, group2tot, group3tot, group4tot, group5tot)
Group <- c(1:5)
groupsdel2 <- as.data.frame(cbind(Group/100, all))
colnames(groupsdel2) <- c("Group", "PercentDelivered")
groupsdel2<-groupsdel2*100
#xtable(groupsdel2,caption="Best-Case Scenario")
```