Dataset consists of complaints files by consumers against financial institutions between years 2011-2016.

Initial exploratory analysis reveals:

1. In a span of 5 years complaint volume has increased more than 65 times implying that this is a significant and impactful tool in our hand
2. Total complaints submitted to government = 550941 which were then directed back to company for their response. Among them 20% of company responses were further disputed (FP) by consumers. For initial analysis I have selected this subset of the data to give me a sample of completely dissatisfied customers.
3. The data lists total of 3591 companies of which I select 30 with > 500 complaints (FP) enlisted under their names.
4. The most commonly seen products about which consumers expressed grievance are 5 (from 11 listed products) that share 90% of the entries.
5. Identified the top 5 issues that customers had with every financial product
6. The trust metric will be designed on the basis of these 5 producs : Mortgage, Debt Collection, Credit Reporting, Credit card and bank account services. All the loan categories are grouped together under one common head.

CODE :

ccom<-read.csv(file="Consumer\_Complaints.csv")

ccom<-ccom %>% mutate(year=year(mdy(Date.received)))

ccom %>% group\_by(year) %>% summarise(ncomplain=n())

ccom %>% filter(Consumer.disputed.=="Yes")%>% nrow()

s1<-ccom %>% group\_by(Company) %>% filter(Consumer.disputed.=="Yes")%>%summarise(ncount=n()) %>% arrange(desc(ncount))

s2<-s1%>% filter(ncount>500)

ccom %>% filter(Consumer.disputed.=="Yes") %>% group\_by(Product) %>% tally(sort=TRUE) %>% mutate (frac\_percent = 100.\*n/sum(n))

s3<- ccom %>% group\_by(Product,Issue) %>% summarise(ncases=n()) %>% top\_n(5,ncases) %>% ungroup() %>% mutate(nfrac=100\*ncases/sum(ncases))%>% arrange(Product)

plot#1

c<-ccom%>%filter(year!=2016) %>% filter(Consumer.disputed.=="Yes") %>% group\_by(year,Product) %>%summarise(count=n())

data<-tbl\_df(c)

ggplot(data) + aes(x=year,y=count,color=Product) + geom\_point(size=4,shape=17) + geom\_line(size=1) +xlab('Year')+ylab('Number of Complaints')+theme\_bw()

plot#2

Mortgage <-ccom %>% group\_by(Company) %>% filter(Product=="Mortgage") %>% filter(Consumer.disputed.=="Yes") %>% summarise(Mortgage=n())

CreditRep <-ccom %>% group\_by(Company) %>% filter(Product=="Credit reporting") %>% filter(Consumer.disputed.=="Yes") %>% summarise(CreditRep=n())

Debt <-ccom %>% group\_by(Company) %>% filter(Product=="Debt collection") %>% filter(Consumer.disputed.=="Yes") %>% summarise(Debt=n())

BankAccnt <-ccom %>% group\_by(Company) %>% filter(Product=="Bank account or service") %>% filter(Consumer.disputed.=="Yes") %>% summarise(BankAccnt=n())

CreditCard <-ccom %>% group\_by(Company) %>% filter(Product=="Credit card") %>% filter(Consumer.disputed.=="Yes") %>% summarise(CreditCard=n())

lcomp<-s1[1:5,1]

bplot<-inner\_join(lcomp,Mortgage)

> bplot<-inner\_join(bplot,CreditRep)

> bplot<-inner\_join(bplot,Debt)

> bplot<-inner\_join(bplot,BankAccnt)

> bplot<-inner\_join(bplot,CreditCard)

bplot<-bplot[,-1]

row.names(bplot)<-c("BankOAmerica", "Wells\_Fargo", "JPMorgan", "Equifax", "Citibank")

dt<-as.table(as.matrix(bplot))

balloonplot(t(dt),main="Product Complaint Distribution",show.margins = FALSE,label=FALSE,xlab="",ylab="",cum.margins=FALSE,rowmar=2.5,colmar=.75,rowsrt=45,dotcolor='red',text.size=.75)