

Exploration2

Rohan Fray

```
library(plyr)
```

```
nyopth <- read.csv("~/cunyMsda2015/nyopth.csv")  
nyopth.92014 <- nyopth[nyopth$hcpcs_code==92014,]
```

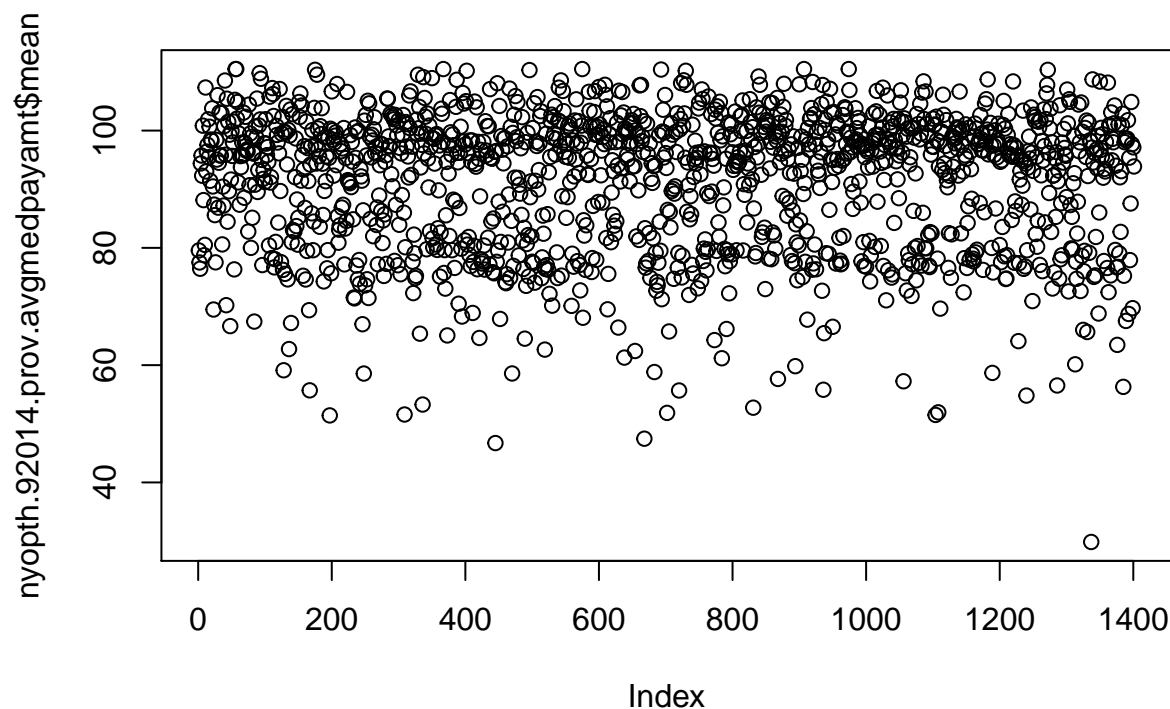
The thinking behind the fraud detection is to find some features about the providers on which we can grade them on. First we will look at the hcpcs code 92014 which stands for “Eye and medical examination for diagnosis and treatment, established patient, 1 or more visits”

```
length(unique(nyopth.92014$npi))
```

```
## [1] 1401
```

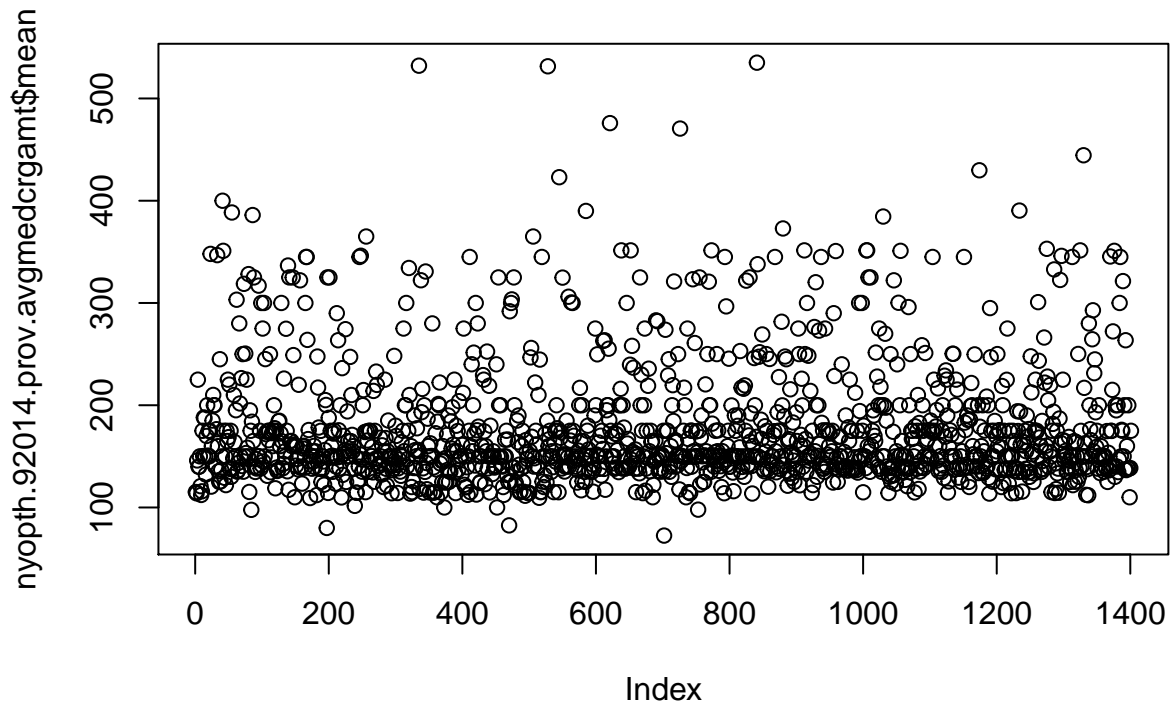
There are 1401 unique providers.

```
nyopth.92014.prov.avgmedpayamt <- ddply(nyopth.92014,~npi,summarise,mean=mean(average_medicare_payment_amt))  
plot(nyopth.92014.prov.avgmedpayamt$mean)
```



Somewhat unsurprisingly, there don't seem to be any gross outliers when it comes to average medicare payment amount

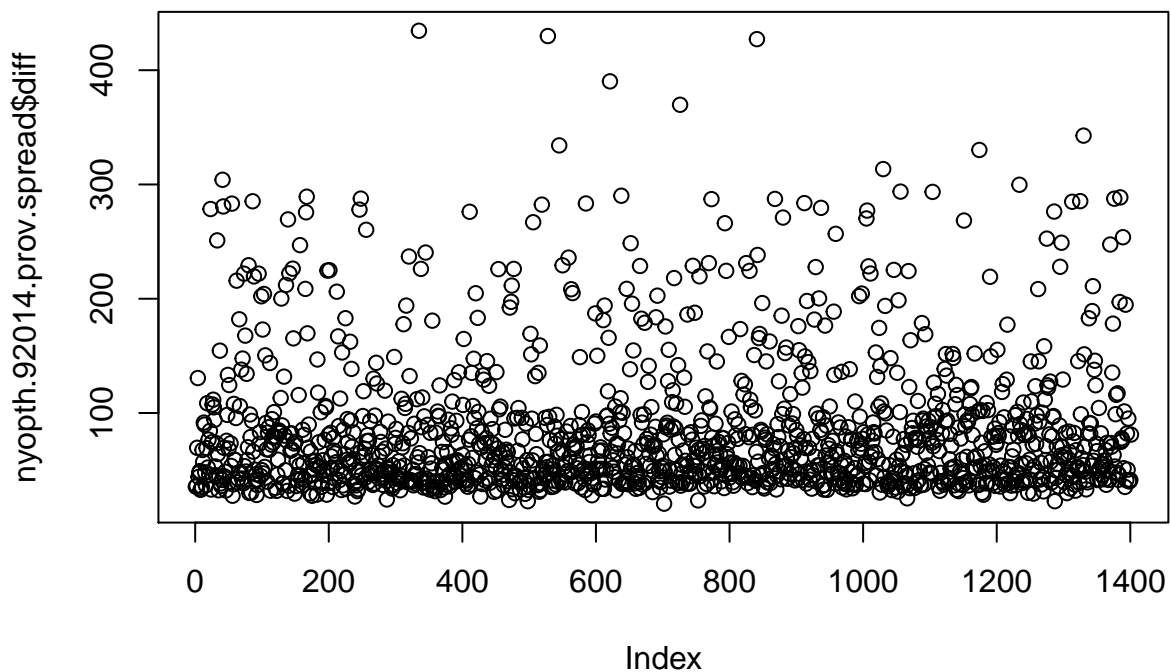
```
nyopth.92014.prov.avgmedcrgamt <- ddply(nyopth.92014,~npi,summarise,mean=mean(average_submitted_chrg_amt))  
plot(nyopth.92014.prov.avgmedcrgamt$mean)
```



Also not surprising, The amount that the providers charge (that is the amount they want for the procedure) has outliers

I believe one of our features should be the spread between the average medicare payment and the average submitted charged amounts

```
nyoph.92014.prov.spread <- merge(nyoph.92014.prov.avgmedpayamt,nyoph.92014.prov.avgmedcrgamt,by.x =
nyoph.92014.prov.spread$diff = nyoph.92014.prov.spread$mean.y - nyoph.92014.prov.spread$mean.x
plot(nyoph.92014.prov.spread$diff)
```



There are a good many outliers here.

I will focus on one provider for a moment who has a large difference

```
prov <- nyopth.92014[nyopth.92014$npi==1235114026,]  
prov
```

```
##          X          npi nppes_provider_last_org_name  
## 4766 4766 1235114026          ABRAMSON  
##      nppes_provider_first_name nppes_provider_mi nppes_credentials  
## 4766          DAVID H          H          MD  
##      nppes_provider_gender nppes_entity_code nppes_provider_street1  
## 4766          M          I          1275 YORK AVE  
##      nppes_provider_street2 nppes_provider_city nppes_provider_zip  
## 4766          NEW YORK          100216007  
##      nppes_provider_state nppes_provider_country provider_type  
## 4766          NY          US Ophthalmology  
##      medicare_participation_indicator place_of_service hcpcs_code  
## 4766          Y          0          92014  
##                                          hcpcs_description  
## 4766 Eye and medical examination for diagnosis and treatment, established patient, 1 or more visits  
##      hcpcs_drug_indicator line_srvc_cnt bene_unique_cnt bene_day_srvc_cnt  
## 4766          N          273          205          273  
##      average_medicare_allowed_amount stdev_medicare_allowed_amount  
## 4766          134.76          0  
##      average_submitted_chrg_amt stdev_submitted_chrg_amt  
## 4766          532.07          11.7445  
##      average_medicare_payment_amt stdev_medicare_payment_amt  
## 4766          97.5562          29.91177
```

According to <http://www.city-data.com/zip/10021.html> This area has a very high cost of living and AGI, so it is not inconceivable that the provider would submit a higher charge. Thus we should factor in COL or AGI when look at the spread.