Part 1: Veris Data

Thomas Reid Zuk

March 12, 2016

Github Code for this part: https://github.com/Freakazoidile/SRT_Assignment_Graphs/tree/master/Part1

The textbook code examines VCDB data about computer security related incidents from 2013. This data is now nearly 3 years old and the current VCDB contains incidents from as recent as roughly 30 days. "We have more contributors and more incidents than ever before."[1] Between 2013 and 2016 the number of contributors to VCDB has increased greatly which helps to increase the knowledge surrounding security incidents and react accordingly.

This report will examine both the old and current data to see how security incidents have changed and explore some of the following questions.

How have threat actors changed? Are incidents from outside sources, inside sources or other?

With internet being an ever evolving entity and with people around the world becoming more and more connected everyday surely where from attacks come from will change. Are the number of external 'bad guys' increasing, are companies becoming more vulnerable from the inside and how has the ability to remain stealthy/unknown in terms of the actor changed?

How have the type of attacks changed?

What type of attacks are being carried out, what is on the rise, and what is becoming less popular. As companies learn to deal with computer security it is going to force attackers to change how they operate. Viruses warning chain e-mails still make their rounds around the internet from time to time but in the past they were much more popular and in the news. Social media and data dumps seem to be a large target so will the data gathered reflect the medias reporting on the issues?

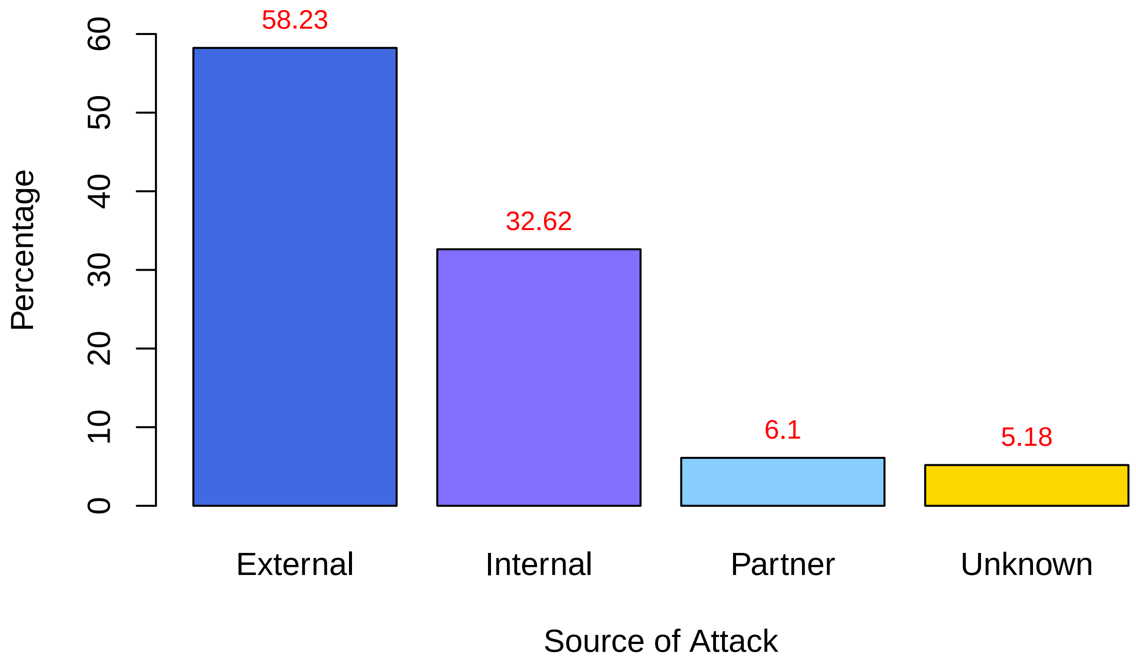
How have the targets of the attacks changed?

As people, companies, and even things get connected the targets will likely change. Attackers will likely look for the easiest, unsuspecting and uninformed to prey on. So how has this changed, are companies servers becoming a target, or is it every day users checking their e-mail and downloading files.

Comparing Threat Actors Old versus New

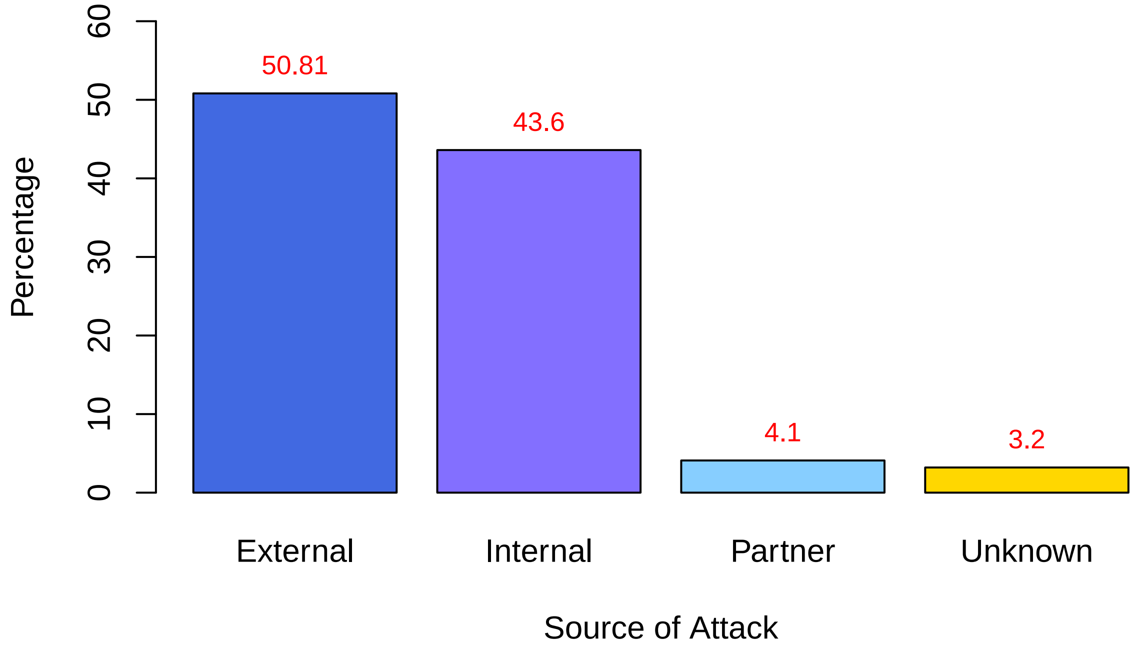
Lets examing where attacks come from in the old data

```r
oldactors <- getenum(oldvcdb, "actor", add.n=TRUE, add.freq=TRUE)

oldActorsPlot <- barplot(
                  oldactors$freq*100,
                  names = oldactors$enum,
                  col=c("royalblue", "slateblue1", "skyblue1","gold"),
                  xlab="Source of Attack",
                  ylab="Percentage",
                  ylim=range(0:65)
                )

text(x = oldActorsPlot, y = oldactors$freq*100, label = round(oldactors$freq*100,2), pos = 3, cex = 0.8
```



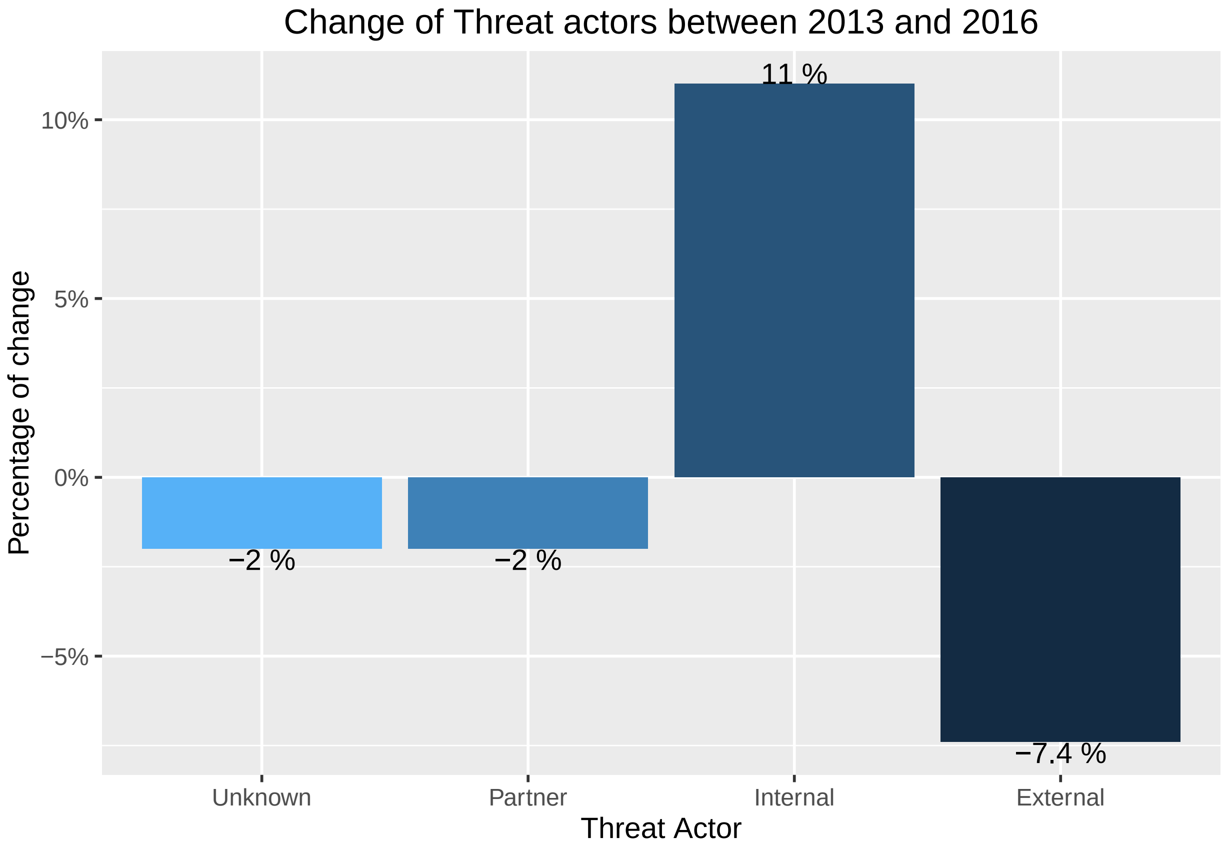
And now the new/current data

```
newactors <- getenum(newvcdb, "actor", add.n=TRUE, add.freq=TRUE)

newActorsPlot <- barplot(
                  newactors$freq*100,
                  names = newactors$enum,
                  col=c("royalblue", "slateblue1", "skyblue1","gold"),
                  xlab="Source of Attack",
                  ylab="Percentage",
                  ylim=range(0:65)
                )

text(x = newActorsPlot, y = newactors$freq*100, label = round(newactors$freq*100,2), pos = 3, cex = 0.8
```



Now that we have looked at the sources of the attack, lets see how they changed between the 2 data sets.

```
enums <- c("Unknown", "Partner", "Internal", "External")
values <- round((newactors$freq - oldactors$freq),3)

dtf <- data.frame(x = c(enums), y = c(values))

ggplot(dtf, aes(x, y)) +
  geom_bar(stat = "identity", aes(fill = c(1,2,3,4))) +
  geom_text(aes(label = paste(y * 100, "%"),
                vjust = ifelse(y >= 0, 0, 1))
            ) +
  scale_y_continuous("Percentage of change", labels = percent_format()) +
  scale_x_discrete("Threat Actor", labels=c("Unknown", "Partner", "Internal", "External")) +
  theme(legend.position="none") +
  ggtitle("Change of Threat actors between 2013 and 2016")
```

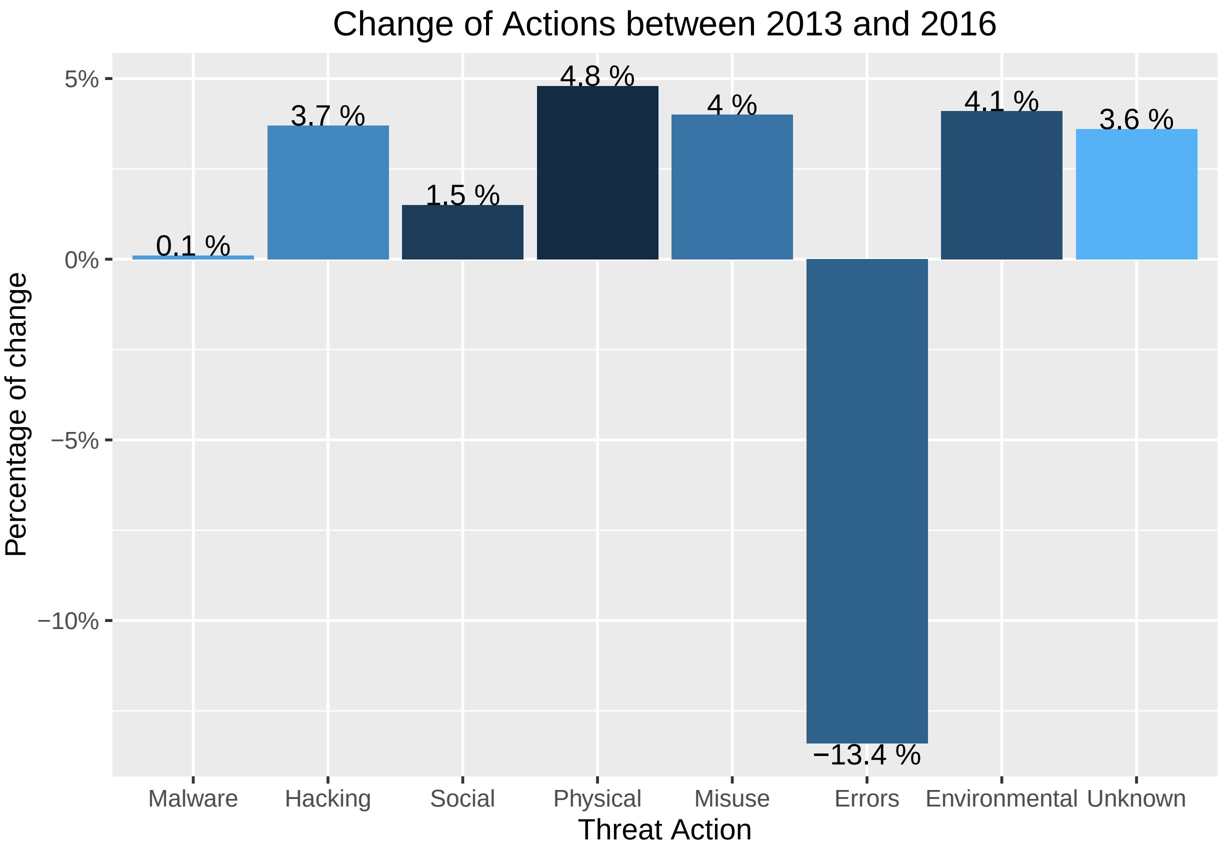


After examining the data provided we can see that attacks are interestingly coming more from Internal Sources. With security breachs from personal information data dumps, to credit card numbers being stolen in the news I was under the impression internal security, auditing and monitoring would be sufficient. Atleast there is some comfort in realizing that external attacks have declined, maybe because companies have paid attention and reacted accordingly. Now it is the people on the inside that the ability to access internal systems and security policies to find vulnerabilties how are in the spotlight.

```
newactions <- getenum(newvcdb, "action", add.n=TRUE, add.freq=TRUE)

oldactions <- getenum(oldvcdb, "action", add.n=TRUE, add.freq=TRUE)


enums <- c("Malware", "Hacking", "Social", "Physical", "Misuse", "Errors", "Environmental", "Unknown")
values <- round((newactions$freq - oldactions$freq),3)

dtf <- data.frame(x = c(enums), y = c(values))
```

Let’s take a look at the change in the type of attacks.

```
ggplot(dtf, aes(x, y)) +
  geom_bar(stat = "identity", aes(fill = c(1,2,3,4,5,6,7,8))) +
  geom_text(aes(label = paste(y * 100, "%"),
              vjust = ifelse(y >= 0, 0, 1))
  ) +
  scale_y_continuous("Percentage of change", labels = percent_format()) +
  scale_x_discrete("Threat Action", labels=c("Malware", "Hacking", "Social", "Physical", "Misuse", "Err
  theme(legend.position="none") +
  ggtitle("Change of Actions between 2013 and 2016")
```

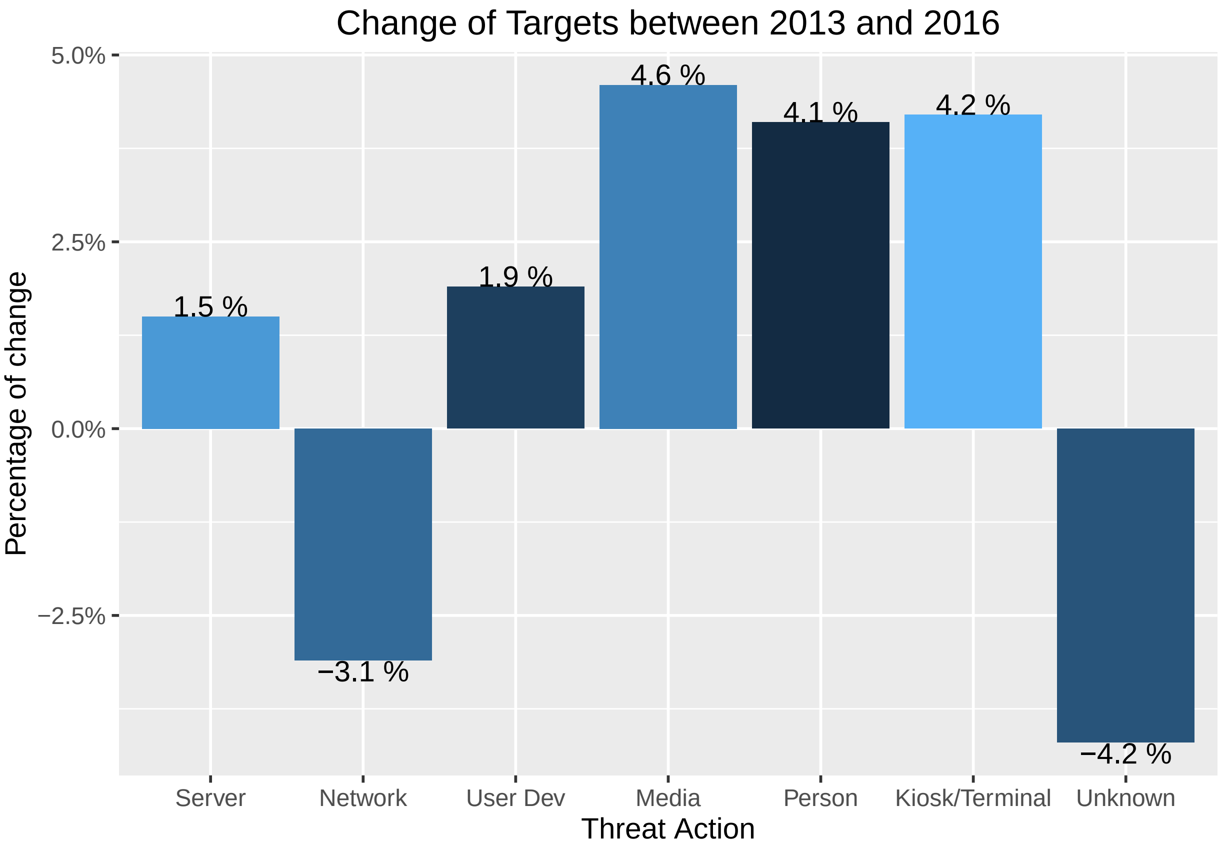


How the attacks are being performed is reassuring, with Errors decreasing by 13%, the most change in any area of security incidents examined. Knowing that software and applications have less bugs means companies are putting more effort into testing, listening to users feedback and in general have better software development. Enviornmental and Physical is interesting, if these factors keep increasing in years to come it would suggest both good and bad actions of a company. On one hand maybe companies are becoming so secure to online threats that physical is the only way to exploit, but on the other hand maybe companies have neglected physical related concerns in an effort to improve digital. Either way it will be interesting to see how these statistics change in the years to come.

Types of Targets

```
newasset <- getenum(newvcdb, "asset.assets", add.n=TRUE, add.freq=TRUE)

oldasset <- getenum(oldvcdb, "asset.assets", add.n=TRUE, add.freq=TRUE)



enums <- c("Server", "Network", "User Dev", "Media", "Person", "Kiosk/Terminal", "Unknown")
```

```r
values <- round((newasset$freq - oldasset$freq),3)

dtf <- data.frame(x = c(enums), y = c(values))

ggplot(dtf, aes(x, y)) +
  geom_bar(stat = "identity", aes(fill = c(1,2,3,4,5,6,7))) +
  geom_text(aes(label = paste(y * 100, "%"),
              vjust = ifelse(y >= 0, 0, 1))
  ) +
  scale_y_continuous("Percentage of change", labels = percent_format()) +
  scale_x_discrete("Threat Action", labels=c("Server", "Network", "User Dev", "Media", "Person", "Kiosk/
  theme(legend.position="none") +
  ggtitle("Change of Targets between 2013 and 2016")
```

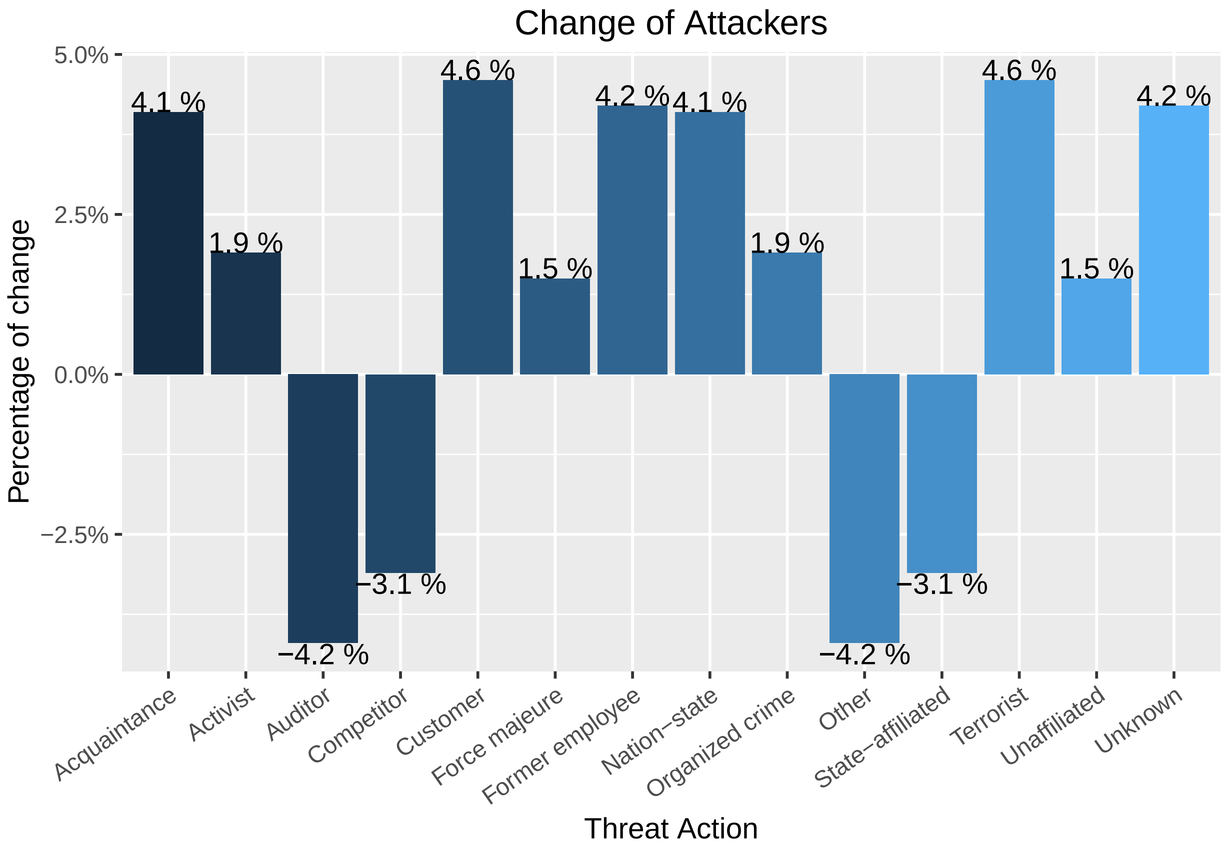


The choice of targets is about as expected, given the various credit card attacks at large retails Person and Kiosk incidents increasing reflects the news. Media and Person increasing also relate to what has been in the news, with all the social media and information put online it provides a very large and sometimes easy target. In the past few years there was the iCould breach which say many clebrities personal information that was backed up on Apples iCloud being released online.

The type of attackers

```r
newCause <- getenum(newvcdb, "actor.external.variety", add.n=TRUE, add.freq=TRUE)
```

```
oldCause<- getenum(oldvcdb, "actor.external.variety", add.n=TRUE, add.freq=TRUE)


enums <- c("Acquaintance", "Activist", "Auditor", "Competitor", "Customer", "Force majeure", "Former emp
values <- round((newasset$freq - oldasset$freq),3)

dtf <- data.frame(x = c(enums), y = c(values))

ggplot(dtf, aes(x, y)) +
  geom_bar(stat = "identity", aes(fill = c(1,2,3,4,5,6,7,8,9,10,11,12,13,14))) +
  geom_text(aes(label = paste(y * 100, "%"),
                vjust = ifelse(y >= 0, 0, 1))
  ) +
  scale_y_continuous("Percentage of change", labels = percent_format()) +
  scale_x_discrete("Threat Action", labels=c("Acquaintance", "Activist", "Auditor", "Competitor", "Custo
  theme(legend.position="none", axis.text.x = element_text(angle = 35, hjust = 1)) +
  ggtitle("Change of Attackers")
```



The type of attackers is also interesting to look at and provides some informative results. Customer, Nation-state (country), and Terrorist all rose just over 4% and some what align with what people might expect given the past topics in computer security. A break out in a cyber war, or terrorists attacking infastructure have always been theorized but many experts and presented in the media both in movies and the news, so seeing Terrorist and Nation-state rise may suggest the thoughts have some credit.

Not much surprise in the decreases of who the attackers were. Auditors, competitors and government related decreasing seems to align well with a stronger computer security mindset.

With more data than ever being gathered, and better, more efficient ways to analyze the data watching how computer and technology security relate incidents change is going to be intruging. It will also provide crucial data on how security increase or decreases, and whether professionals in the technology world predicted the future of security accurately or not.

References

[1] Verizon Data Breach Investigations Report 2015 http://www.verizonenterprise.com/DBIR/

[2] Course text book: Jacobs, Jay. Data-Driven Security Analysis, Visualization and Dashboards.

http://stackoverflow.com/questions/11938293/how-to-label-a-barplot-bar-with-positive-and-negative-bars-with-ggplot2

https://github.com/vz-risk/VCDB - Veris Community Data Base