Data Challenge 1

Team 2

February 21, 2017

Question 1.

1.1. Civilian Deaths on Perfect Lethality Incidents

As it stands, we were unable to replicate the numbers that appear on the PPD document. We tried different assumptions trying to get as close as possible; these are their descriptions and the results of adopting them.

```
(d1 <- sum((d[d$total.people.dead > 0 & d$total.people.wounded == 0,
              ])$civilian.dead) / sum(d$civilian.dead))
```

```
## [1] 0.393617
```

The first set of assumptions is a literal interpretation of the conclusions of the director of PPD. Given the data, this is the result of civilian deaths in Perfect Lethality confrontations as a proportion of total civilian deaths. With a 39.4% it is far from the 86.1% described in the document.

```
(d2 <- sum((d[d$civilian.dead > 0 & d$civilian.wounded == 0, ])$civilian.dead)
 / sum(d$civilian.dead))
```

```
## [1] 0.6755319
```

```
(d3 <- sum((d[d$civilian.dead > 0 & d$civilian.wounded == 0 |
                d$organized.crime.dead > 0 &
                d$organized.crime.wounded == 0, ])$civilian.dead,
           (d[d$civilian.dead > 0 & d$civilian.wounded == 0 |
                d$organized.crime.dead > 0 &
                d$organized.crime.wounded == 0, ])$organized.crime.dead)

 / sum(d$civilian.dead, d$organized.crime.dead))
```

```
## [1] 0.7823254
```

If we assume that the ratio is referring to cases of perfect civilian fatality over total civilian deaths, the ratio come to a 67.6%. If we assume that the the index if for cases of perfect fatality of civilians and organized crime, and that organized crime members are civilians the ratio comes to 78.2%. Still not quite the number offered by PPD.

```
federal <- d[d$federal.police.dead > 0|
               d$federal.police.wounded >0|
                 d$afi.dead >0 |
                 d$afi.wounded >0 |
               d$ministerial.police.dead > 0 |
                 d$ministerial.police.wounded > 0, ]


D <- federal[federal$civilian.dead >0 &
                d$federal.police.wounded >0|
                d$afi.dead >0 |
                d$afi.wounded >0 |
              d$ministerial.police.dead > 0 |
                d$ministerial.police.wounded > 0, ]

D <- federal[federal$civilian.dead >0 &
                federal$civilian.wounded >= 0 , ]

N <- federal[federal$civilian.dead >0 &
                federal$civilian.wounded == 0, ]
```

```
sum(N$civilian.dead)/sum(D$civilian.dead)
```

```
## [1] 0.7111111
```

This is assuming that, as the document says, the percentage of civilian deaths is solely on confrontations with Federal Armed Forces. At the same time, it assumes that these Federal Armed Forces exclude Army and Navy bodies and include only federal police forces. The third assumption here is that we can distinguish cases where these forces were exclusively confronted if a representative of the forces was either willed or wounded in the confrontation. This last assumption is weak, for there is no reason to consider that just because a representative was hurt, this is an indicator that the body it represents was exclusively engaged.

In this case the ratio refers to cases of total civilian fatality over non-total fatality cases in the aforementioned subset. A 71.1% is still not close to the 86.1% we were trying to replicate.

```
federal <- d[d$federal.police.dead > 0|
               d$federal.police.wounded >0|
                 d$afi.dead >0 |
                 d$afi.wounded >0 |
               d$ministerial.police.dead > 0 |
                 d$ministerial.police.wounded > 0, ]


D <- federal[federal$civilian.dead >0 &
                 federal$civilian.wounded >= 0 &
                federal$organized.crime.dead > 0 &
                federal$organized.crime.wounded >= 0, ]

N <- federal[federal$civilian.dead >0 &
                 federal$civilian.wounded == 0 &
                federal$organized.crime.dead > 0 &
                federal$organized.crime.wounded == 0, ]


sum(N$civilian.dead)/sum(D$civilian.dead)
```

```
## [1] 0.8888889
```

Finally, this is the same subset as before, but assuming organized crime members are considered civilian population as they are not part of any Law Enforcement or Armed Forces body. They are legally civilians. 88.1% of civilians so defined were killed in total civilian fatality confrontations; this is above the number offered by PPD.

1.2 Total Lethality

```
sum(d$total.people.dead) / sum(d$total.people.wounded)
```

```
## [1] 1.438211
```

This is the index assuming the overall index for Mexico refers to the literal ratio of total deaths per people wounded in a confrontation. 1.4 is not the 2.6 of the document.

```
sum(d$civilian.dead, d$organized.crime.dead) /
  sum(d$civilian.wounded, d$organized.crime.wounded)
```

```
## [1] 2.309972
```

If we were to assume that the index refers only to non Law Enforcement or Armed Forces victims (because

the context is lethality of these bodes), and thus includes organized crime and civilian people, the index comes close to the one trying to be replicated with a 2.3.

1.3 Lethality Index by Law Enforcement Body

```
##Federal Police
sum(d$federal.police.dead) /sum(d$federal.police.wounded)
```

```
## [1] 0.3033708
```

```
## Navy
sum(d$navy.dead) /sum(d$navy.wounded)
```

```
## [1] 0.3064516
```

```
## Army
sum(d$military.dead) /sum(d$military.wounded)
```

```
## [1] 0.1784038
```

This is assuming the index refers to fatalities of the specific bodies mentioned and not of the number of killed and wounded by the bodies. Because the variables refer to deaths and not kills, there is no direct way to measure with these data what the PPD characterizes as an index per Law Enforcement body. The indexes reflex that this assumption is not what is behind the numbers in the document.

```
##Federal Police
federal_police <- d[d$federal.police.dead > 0|
              d$federal.police.wounded > 0, ]

sum(federal_police$total.people.dead)/sum(federal_police$total.people.wounded)
```

```
## [1] 0.8563218
```

```
##Navy
navy <- d[d$navy.dead > 0|
              d$navy.wounded > 0, ]

sum(navy$total.people.dead)/sum(navy$total.people.wounded)
```

```
## [1] 1.348315
```

```
##Army
army <- d[d$military.dead > 0|
              d$military.wounded > 0, ]

sum(army$total.people.dead)/sum(army$total.people.wounded)
```

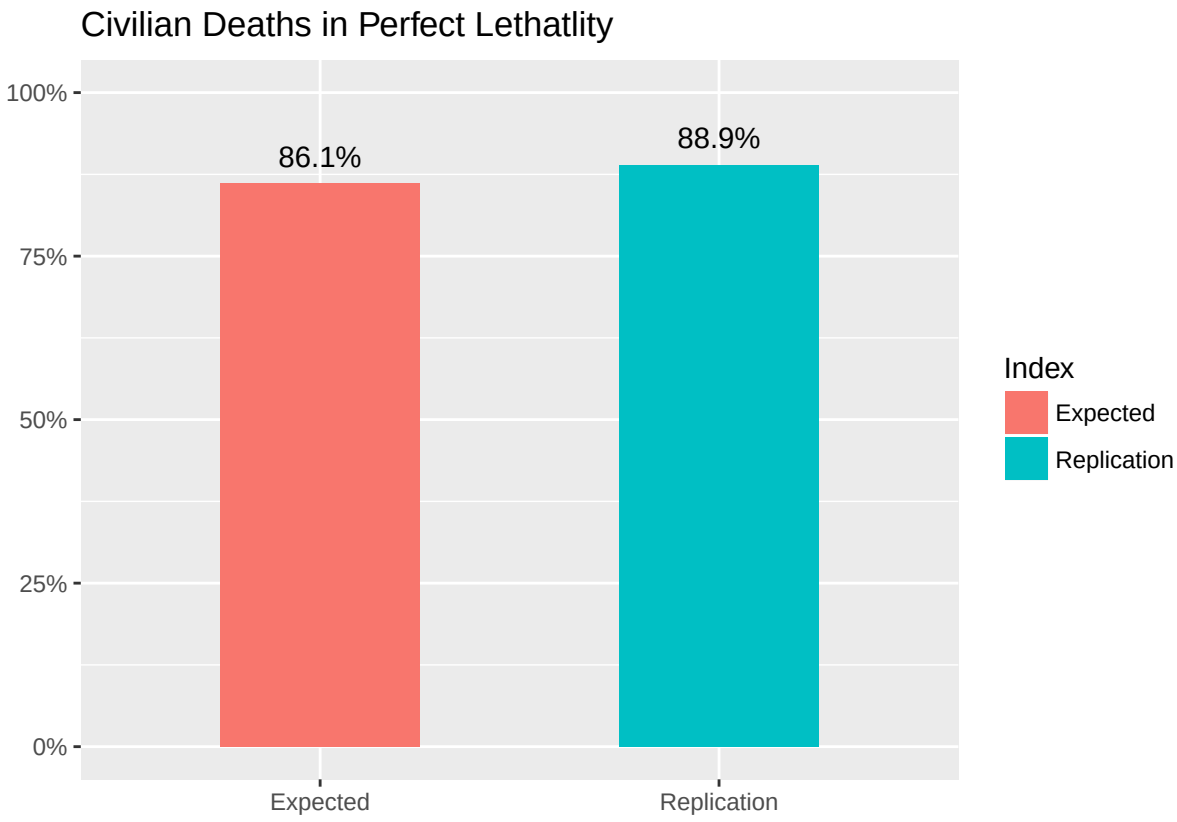
```
## [1] 0.9666255
```

If we assume once more that death or wound o a particular body is an indicator of their sole action, these are the indexes of lethality for total people wounded and killed in each subset of particular Law Enforcement body incidence. It is still very far from the numbers offered by PPD.

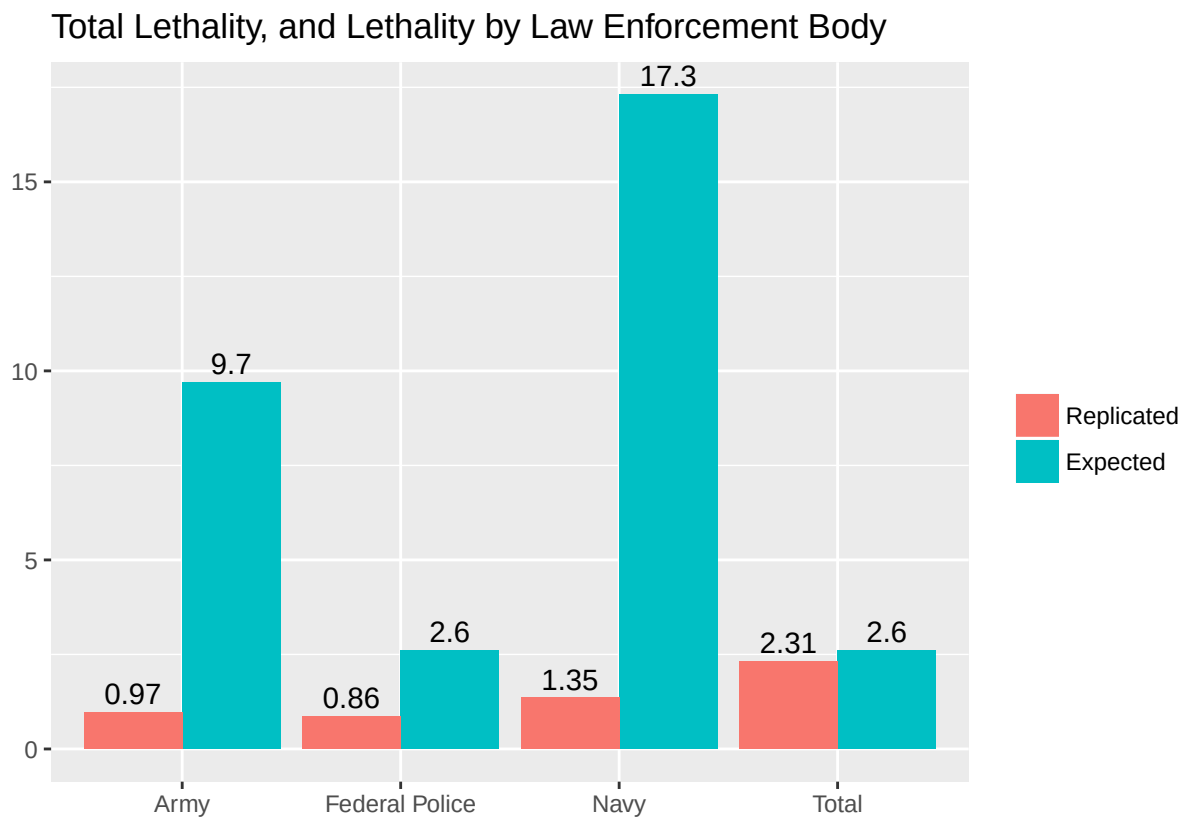
Plots

We have decided to plot our closest replications of the original values as shown in the reference document, compared to the expected value to be replicated.

Civilian Deaths in Perfect Lethatlity



Total Lethality, and Lethality by Law Enforcement Body



Question 2

Is this the right metric to look at? Why or why not?

We believe that the metric is not the right one. Though it is important to measure the lethality of encounters and particular bodies (specially if we are to believe that violence brings forth violence, and death motivates vindictive behavior), a more appropriate measure would put the death and wounded in relation to arrests made. Even if the conflict is at the point of 'by any means necessary', it can be argued that extermination is probably not the best way towards a permanent resolution. Considering this, the effectiveness of confrontations should account for how many arrests are made and what the ratio is between this and use of force (deaths, lesions). Arrests can lead to due process and new information, whereas violence can lead to an obfuscation process. Perhaps controlling behavior like this is better.

What is the "lethality index" showing explicitly?

In the document, Lethality seems to literally refer to the amount of deaths per people wounded in a given confrontation; in particular, the amount of combatants killed over combatants murdered. It is clear from our attempts to replicate the index that this is not exactly how it is measured. Given our best attempts, it seems that the index measures civilian deaths and wounded (with organized crime members counted as civilians), and/or encounters where representatives of particular Law Enforcement bodies were killed or wounded. A third option is that it measures a ratio of wounded Law Enforcement vs death civilians, but this is not clear either.

What is it not showing?

The variables provided fail to show the causation of the death, in which some civilian deaths may have been caused by armed forces, civilians and/or criminals. At the same time, the data is not showing how many total people participated in the confrontation, just the dead and wounded.

What is the definition assuming?

The definition, in general, assumes that the combatants are killed exclusively by Armed Forces, obviating the possibility of inter-gang killings and gang-member-civilian killings. Further, because there are no variables that indicate clearly who was involved in the confrontation and how many people were involved, it appears to assume that Armed Forces representative deaths or wounds are indicators of their sole involvement in the confrontation; hence the lethality indexes for each body being calculated.

With the same available data, can you think of an alternative way to capture the same construct? Is it "better"?

There are three things we can calculate. One is the number of dead and wounded civilians and criminals comparing with dead and wounded government officers to assess the efficiency of the armed forces. The second one is calculating the ratio of death of both civilians and organized crime members vs deaths of these two groups to generate a better Lethality Index that reflexes the bluntness of action of Armed Forced in Mexico. Finally, a better index for purposes of understanding actual effectiveness would be to compare the number of civilians and criminals deaths and wounded to detentions made, for the reasons outlined above. A way to control for if blunt force was necessary, is to include in this index the number of death and wounded Armed Forces; if there is none, perhaps the magnitude of the force was not necessary.

What additional information would you need to better understand the data? What additional information could help you better capture the construct behind the "lethality index"

An additional variable that examines the causation of death can provide a more accurate understanding of the death variables. Also, a variable that clearly states how many participants of each body (from civilian to Navy) were involved in each conflict would clear many of our doubts. At the same time, understanding if the confrontation was as an immediate response to a situation on the street or if it was the product of a planned raid could also help better understand how the Armed Forces are using their strength. Finally, a variable that specifies if the detentions led to a conviction could help better understand, together with the other variables, the overall necessity and effectiveness of confrontations.