GIE MODULE 3 REPORT: RADIOLOGICAL EVENTS

Andreu Schoenenberger López

June 6, 2017

Contents

Introduction

1 Goal of the project

It's well known that use of civil nuclear energy in the last 30 years has been dramatically increasing since it's a very effective way to obtain large amounts of energy, besides its military use. Of course there are drawbacks like nuclear waste but, most important, dangerous radiation emitions when accidents or events occur, leading to a very unclear future for those that absorved significant amounts of radiation during those evenets. Furthermore, soil and wild life is also severly affected when accidents happen.

With that in mind, we will extract information about the Database of Radiological Incidents and Related Events collected from 1896 to 2013 approx. (last major event being Fukushima incident) by Wm. Robert Johnston and stored in the following url:
http://www.johnstonsarchive.net/nuclear/radevents/

Wm. Robert Johnston is a research physicist in the field of space physics: the study of the space environment, encompassing realms from the ionosphere to the magnetosphere to interplanetary space. His current concentration is in the study of the Earth's radiation belts. We holds a B.A. in astronomy from University of Texas (Austin), an M.S. in physics also from University of Texas (El Paso) and a Ph.D. in physics from University of Texas (Dallas). More information about his publications in: http://www.johnstonsarchive.net/about.html

The goal of the project is then:

- To see the overall (or per country) evolution of deaths due to radiological events (direct deaths in the incident, not posterior deaths) per year.
- To see the overall (or per country) evolution of leaked radiation that can lead to posterior deaths or health issues related to exposure of radiation per year.
- Also, less obvious, if number of incidents worldwide follow a tendency of decrease or increase per year. Knowing that there is more use of nuclear energy but also more safety protocols and knowledge of consequences of nuclear incidents.

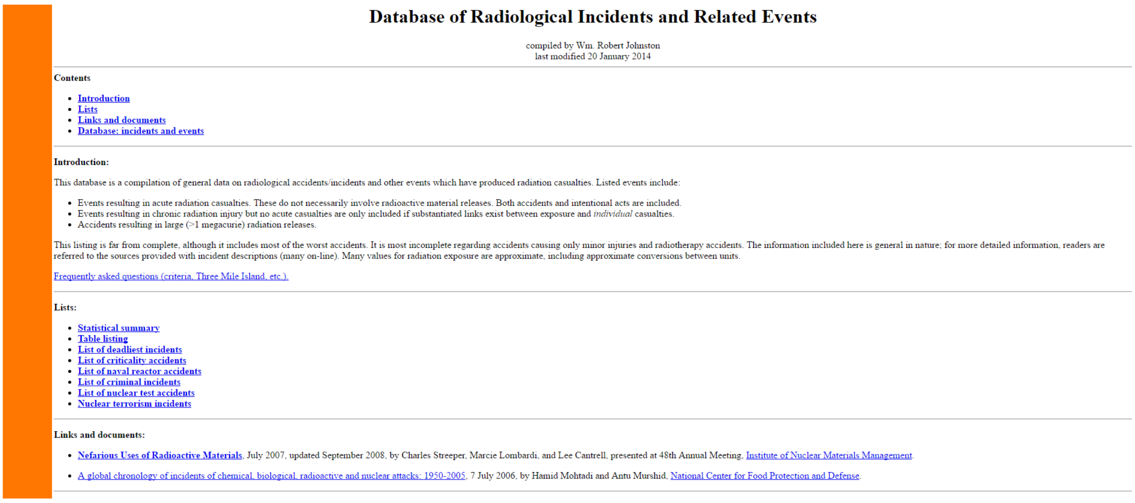
To achieve this goal we will use the list of deadliest incidents, list of criticality accidents, list of naval reactor accidents, list of criminal incidents and list of nuclear test accidents. All lists are stored in the previous url, as well as the table listing that describe the codes used to classify incidents in the previous tables.

2 Web Technology and Data Source

2.1 Web Technology

In figure 1 we can see the main web page of Wm. Robert Johnston regarding the Database of Radiological Incidents and Related Events. At first sight we can't see what technology the web is based on (html, xml, ...). But we have severals ways to know that:

- Since we are interested in list of deadliest incidents, list of criticality accidents, list of naval reactor accidents, list of criminal incidents and list of nuclear test accidents, we can enter those links and see if the web has an ".html" format at the end. For example, accessing the list of deadliest incidents link from the main page, we can see the following url at our browser: http://www.johnstonsarchive.net/nuclear/radevents/radevents1.html. Meaning that the site is based on HTML technology.
- Another easiest way of knowing the technology is simply inspecting the site with our internet browser (Chrome in this case) and going to the top. As we can see in figure 2 the top line is HTML, indicating that the site is written in that programming language.



Database of Radiological Incidents and Related Events

compiled by Wm. Robert Johnston
last modified 20 January 2014

Contents

- **Introduction**
- **Lists**
- **Links and documents**
- **Database: incidents and events**

Introduction:

This database is a compilation of general data on radiological accidents/incidents and other events which have produced radiation casualties. Listed events include:

- Events resulting in acute radiation casualties. These do not necessarily involve radioactive material releases. Both accidents and intentional acts are included.
- Events resulting in chronic radiation injury but no acute casualties are only included if substantiated links exist between exposure and *individual* casualties.
- Accidents resulting in large (>1 megacurie) radiation releases.

This listing is far from complete, although it includes most of the worst accidents. It is most incomplete regarding accidents causing only minor injuries and radiotherapy accidents. The information included here is general in nature; for more detailed information, readers are referred to the sources provided with incident descriptions (many on-line). Many values for radiation exposure are approximate, including approximate conversions between units.

Frequently asked questions (criteria, Three Mile Island, etc.).

Lists:

- **Statistical summary**
- **Table listing**
- **List of deadliest incidents**
- **List of criticality accidents**
- **List of naval reactor accidents**
- **List of criminal incidents**
- **List of nuclear test accidents**
- **Nuclear terrorism incidents**

Links and documents:

- **Nefarious Uses of Radioactive Materials**, July 2007, updated September 2008, by Charles Streeper, Marcie Lombardi, and Lee Cantrell, presented at 48th Annual Meeting, Institute of Nuclear Materials Management.
- A global chronology of incidents of chemical, biological, radioactive and nuclear attacks: 1950-2005, 7 July 2006, by Hamid Mohtadi and Antu Murshid, National Center for Food Protection and Defense.

Figure 1: Main Web Page

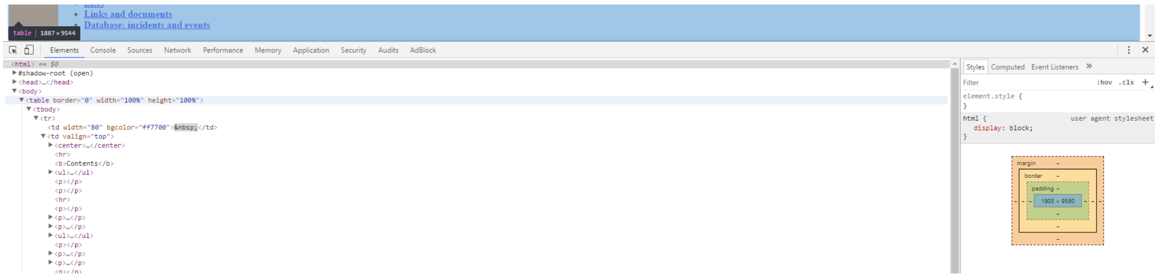
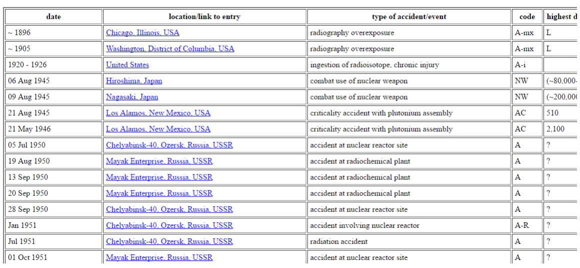


Figure 2: Code inspection of the main page

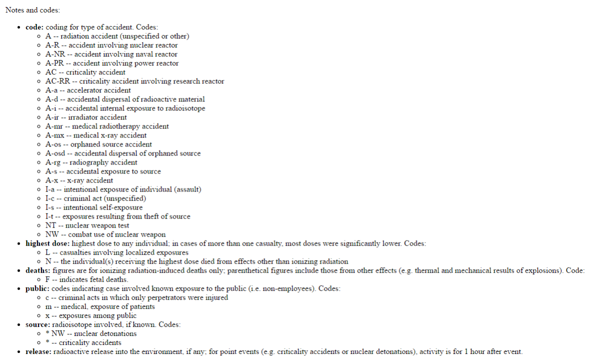
2.2 Data Sources

In this study we will have two types of data sources. The main one is HTML tables stored in the web corresponding to: list of deadliest incidents, list of criticality accidents, list of naval reactor accidents, list of criminal incidents and list of nuclear test accidents. They contain different variables like location, number of deaths, injured, etc. (see figure 3 for sample). The other type of data that we have to extract is the codes that are assigned to most of the events, those codes are usefull to obtain a description of what happened (in general terms) in each incident (see figure 4 for a quick view). Codes are just plain HTML listing that will need a different approach than the tables listed before.



date	location/link to entry	type of accident/event	code	highest d
~ 1896	Chicago, Illinois, USA	radiography overexposure	A-mx	L
~ 1905	Washington, District of Columbia, USA	radiography overexposure	A-mx	L
1920 - 1926	United States	ingestion of radioisotope, chronic injury	A-i	
06 Aug 1945	Hiroshima, Japan	combat use of nuclear weapon	NW	(~80,000-
09 Aug 1945	Nagasaki, Japan	combat use of nuclear weapon	NW	(~200,00
21 Aug 1945	Los Alamos, New Mexico, USA	criticality accident with plutonium assembly	AC	510
21 May 1946	Los Alamos, New Mexico, USA	criticality accident with plutonium assembly	AC	2,100
05 Jul 1950	Chelyabinsk-40, Ozersk, Russia, USSR	accident at nuclear reactor site	A	?
19 Aug 1950	Mayak Enterprise, Russia, USSR	accident at radiochemical plant	A	?
13 Sep 1950	Mayak Enterprise, Russia, USSR	accident at radiochemical plant	A	?
20 Sep 1950	Mayak Enterprise, Russia, USSR	accident at radiochemical plant	A	?
28 Sep 1950	Chelyabinsk-40, Ozersk, Russia, USSR	accident at nuclear reactor site	A	?
Jan 1951	Chelyabinsk-40, Ozersk, Russia, USSR	accident involving nuclear reactor	A-R	?
Jul 1951	Chelyabinsk-40, Ozersk, Russia, USSR	radiation accident	A	?
01 Oct 1951	Mayak Enterprise, Russia, USSR	accident at nuclear reactor site	A	?

Figure 3: Sample of tables for the Radiological Incidents



Notes and codes:

- **code:** coding for type of accident. Codes:
  - A -- radiation accident (unspecified or other)
  - A-R -- accident involving nuclear reactor
  - A-NR -- accident involving naval reactor
  - A-PR -- accident involving power reactor
  - AC -- criticality accident
  - AC-RR -- criticality accident involving research reactor
  - A-a -- accelerator accident
  - A-d -- accidental dispersal of radioactive material
  - A-i -- accidental internal exposure to radioisotope
  - A-ir -- irradiator accident
  - A-mr -- medical radiotherapy accident
  - A-mx -- medical x-ray accident
  - A-os -- orphaned source accident
  - A-osd -- accidental dispersal of orphaned source
  - A-rg -- radiography accident
  - A-s -- accidental exposure to source
  - A-x -- x-ray accident
  - I-a -- intentional exposure of individual (assault)
  - I-c -- criminal act (unspecified)
  - I-s -- intentional self-exposure
  - I-t -- exposures resulting from theft of source
  - NT -- nuclear weapon test
  - NW -- combat use of nuclear weapon
- **highest dose:** highest dose to any individual; in cases of more than one casualty, most doses were significantly lower. Codes:
  - L -- casualties involving localized exposures
  - N -- the individual(s) receiving the highest dose died from effects other than ionizing radiation
- **deaths:** figures are for ionizing radiation-induced deaths only; parenthetical figures include those from other effects (e.g. thermal and mechanical results of explosions). Code:
  - F -- indicates fetal deaths.
- **public:** codes indicating case involved known exposure to the public (i.e. non-employees). Codes:
  - c -- criminal acts in which only perpetrators were injured
  - m -- medical, exposure of patients
  - x -- exposures among public
- **source:** radioisotope involved, if known. Codes:
  - * NW -- nuclear detonations
  - * -- criticality accidents
- **release:** radioactive release into the environment, if any; for point events (e.g. criticality accidents or nuclear detonations), activity is for 1 hour after event.

Figure 4: Codes for the Radiological Incidents

Web Scraping

3 Downloading data

Downloading and reading the data from the Web into R is fairly simple in this case. We are using mainly packages *RCurl*[1] and *XML*[2]. The first one just to download the URL using function *getURL* and the second one is used in a more fancy way like for instance parsing HTML code with function *htmlParse* or to read HTML tables into data frames using funcion *readHTMLTable*.
An example of this can be found below where we are reading the list of deadliest incidents, first downloading the URL, then obtaining all the tables in that URL with function *readHTMLTable* and extracting the table that we want into the object ”dead”:

```
> html <- getURL("http://www.johnstonsarchive.net/nuclear/radevents/radevents1.html")
> tables <- readHTMLTable(html, stringsAsFactors = FALSE)
> dead <- tables[[1]]
> dead <- arrange_table(dead)
```

As a result, we obtain the following data frame:

date	location	event	code	deaths	injuries	dose
06 Aug 1945	Hiroshima, Japan	combat use of nuclear weapon	NW	45,000 (130,000)	60,000? (86,000)	(˜80,000–N)
09 Aug 1945	Nagasaki, Japan	combat use of nuclear weapon	NW	20,000 (65,000)	50,000? (75,000)	(˜200,000–N)
04 Jul 1961	K-19 submarine, North Atlantic	reactor accident	A-NR	8	31	6,000
21 Mar 1962 - Aug 1962	Mexico City, Mexico	lost radiography source	A-os	4	1	5,200
24 May 1968	K-27 submarine, Barents Sea	naval reactor accident	A-NR	9	83	?
1974 - 1976	Columbus, Ohio, USA	radiotherapy accident	A-mr	10	88	L
1980	Houston, Texas, USA	radiotherapy accident	A-mr	7	?	L?
05 Oct 1982	Baku, Azerbaidjan, USSR	lost source	A-os	5	13	?
19 Mar 1984	Casablanca, Morocco	lost radiography source	A-os	8	3	?
10 Aug 1985	K-431 submarine, Chazhma Bay, Vladivostok, Russia, USSR	reactor accident during refueling	A-NR	0 (10)	49	220
26 Apr 1986 - 06 May 1986	Chernobyl, Ukraine, USSR	steam explosion and fire in graphite-moderated power reactor	A-PR	28 (31)	238+	1,600
12 Sep 1987 - 29 Sep 1987	Goiania, Goias, Brazil	accidental dispersal of lost radiography source	A-osd	5	20	700
10 Dec 1990 - 20 Dec 1990	Zarragosa, Spain	radiotherapy accident	A-mr	18	9	L
22 Aug 1996 - 27 Sep 1996	San Jose, Costa Rica	radiotherapy accident	A-mr	7	81	L
Aug 2000 - 24 Mar 2001	Panama City, Panama	radiotherapy accident	A-mr	17	11	L

Table 1: Table from list of deadliest incidents

For the rest of lists of interest, the process is more or less the same, with little variation between them. At the end, we will have a common table containing all the previous tables with the most relevant variables: date of the event, location, brief description, code of the event, deaths, injuries and radioactive dose emitted.

4 Scraping data

In this section we will make use of Regular Expressions in order to clean the data coming from the HTML code. The use of Regular Expressions in this study is used basically in three different ways:

[1]https://cran.r-project.org/web/packages/RCurl/index.html
[2]https://cran.r-project.org/web/packages/XML/index.html

1. To obtain a table for the Codes of Radiological Incidents. This table is constructed from plain HTML text (list in the web, not an HTML table). The process consists of obtaining the HTML text, then extracting the list, drop parts of the list that are not of interest and then clean the remaining data with the use of functions like *grep*, *strsplit*, *sub*, *gsub*, etc.

2. The second way is to extract the data columns that are of interest in the study when reading the HTML tables. This is done with an own-implemented function called "arrange_table", this function recieves a read table and outputs an arranged table that we can use. It's a form of standarizing the extracted data:

```
> arrange_table <- function(one_table){
+   columns.use <- c("date", "location", "accident", "event", "code", "deaths",
+                    "injuries", "dose")
+   columns.use.pat <- paste0(columns.use, collapse = "|")
+   columns.select <- c("date", "location", "event", "code", "deaths",
+                       "injuries", "dose")
+   one_table <- one_table[, grep(columns.use.pat, colnames(one_table))]
+   colnames(one_table)[grep("location|event|accident|dose", colnames(one_table))] <-
+     c("location", "event", "dose")
+   one_table <- one_table[, columns.select]
+   return(one_table)
+ }
>
```

3. Once we have all the data standarized, we still need to clean the values inside the data frame (as seen in figure 3), where numbers of deaths or injured have different characters like "(" or "?". This need cleaning in order to obtain a single number that can be used by R. This is done after binding all tables into one large table (thanks to standarized form of table made before), then we clean the data using Regular Expressions basically to obtain the digits date, injuries and deaths.

   Obtaining all data, removing events with missing date:

```
> ## ALL IN ONE TABLE
> all.data <- rbind(criminal, critical, dead, naval,tests)
> all.data <- all.data[!(all.data$date == ""),]
```

   Obtaining the date, first 4 digits found in all string:

```
> # Last 4 digits in date
> years.in.order <- sort(unique(as.numeric(sub('.*(\\d{4}).*', '\\1',
+                                              all.data$date))))
```

```
> all.data$date <- factor(sub('.*(\\d{4}).*', '\\1',
+                              all.data$date))
```

Obtaining the number of deaths, first change comma with nothing, then remove all after "(" if found, all other characters like "?" that has no numbers are coerced to NA:

```
> # First digits in deaths
> all.data$deaths2 <- gsub(",", "", all.data$deaths)
> all.data$deaths2 <- as.numeric(gsub(" \\(.*", "",
+                                      all.data$deaths2))
```

Similar as deaths, but a little bit more complex for number of injured:

```
> # First digits in injuries
> all.data$injuries2 <- gsub(",", "", all.data$injuries)
> all.data$injuries2 <- as.numeric(sub("([0-9]*).*$", "\\1",
+                                       gsub(" \\(.*", "",
+                                            all.data$injuries2)))
```

Results

5 Results of Web Scraping

In this case, results from Web Scraping are 2 main data frames, as commented before.

Data from Codes (sample of actual table)

code	description
A	radiation accident
A-R	accident involving nuclear reactor
A-NR	accident involving naval reactor
A-PR	accident involving power reactor
AC	criticality accident
AC-RR	criticality accident involving research reactor

Table 2: Table of Codes for Radiological Incidents

Data from Radiological Incidents (sample of actual data)

date	location	event	code	deaths2	injuries2
1960	Moscow, Russia, USSR	intentional overexposure	I-s	1.00	0.00
1961	SL-1 reactor, Idaho RTA, Idaho, USA	criticality excursion with uranium research reactor	AC-RR	3.00	0.00
1968	Pennsylvania, USA	attempt to self-induce abortion using x-ray machine	I-s	0.00	1.00
1972	Harris county, Texas, USA	intentional exposure to individual	I-a	0.00	1.00
1972	Primorsky region, Russia, USSR	criminal act using radioactive source	I-c	0.00	1.00
1972	Bulgaria	self-inflicted radiation exposure	I-s	1.00	0.00

Table 3: Table of all data

6 Results of Analysis

For the analysis we focused mainly on number of events per year, number of deaths and number of injured.

6.1 Analysis on number of events

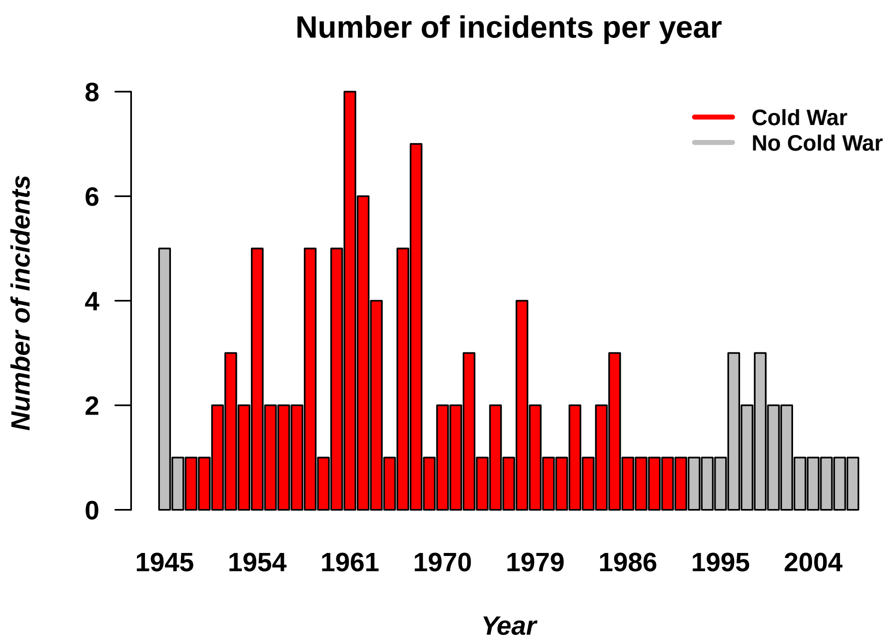


Figure 5: Number of incidents per year

In figure 5 we can see a peak in year 1945, due to nuclear attacks to Hiroshima and Nagasaki. Then, from 1947 to 1965 approx. there is a clear and steep tendency due to the cold war between US and USSR. Due to this ”Nuclear Revolution” in both countries and the lack of knowledge of nuclear energy most probably led to an increase of nuclear incidents. From 1965 to nowadays, the number of incidents is decreasing, leading to a more safe use of nuclear energy.

6.2 Analysis on number of deaths and injured

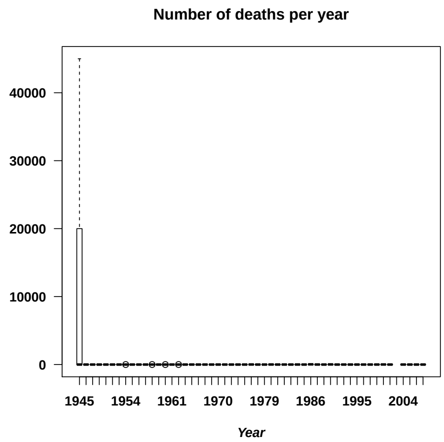


Figure 6: Number of deaths per year

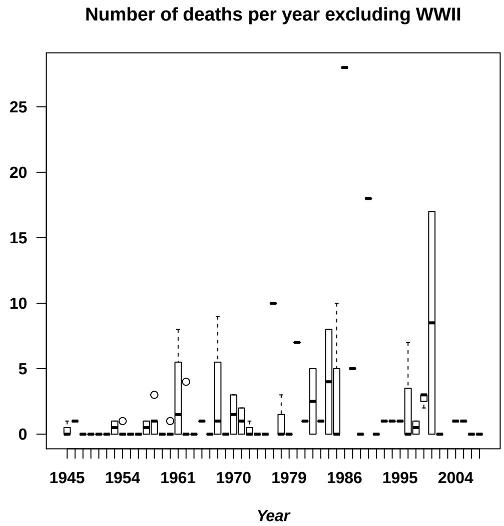


Figure 7: Number of deaths per year without WWII

In figure 6 we can actually see the magnitudes of Hiroshima and Nagasaki nuclear bombs in 1945 compared to the number of deaths in all the history of nuclear energy usage. So in figure 7 we excluded the Japan bombings in WWII and we can see a clear peak in 1986 corresponding to Chernobyl incident. Another large box in 2001 that was due to two incidents in that year, one in Panama City (radiotherapy accident, code A-mr) and another one in Russia (theft of nuclear source). Second largest peak corresponds to an incident in Spain in 1990:

date	location	event	code	deaths	injuries	dose	deaths2	injuries2
1990	Zarragosa, Spain	radiotherapy accident	A-mr	18	9	L	18.00	9.00

Table 4: Incident in 1990 (Spain)

And if we take a look at code A-mr:

code	description
A-mr	medical radiotherapy accident

Table 5: Code for incident in 1990 (Spain)

Meaning that something went wrong using medical nuclear energy, leaving 18 deaths and 9 injuries.
We can also take a look at the aggregation of data by year, summing then the number of deaths and injured:

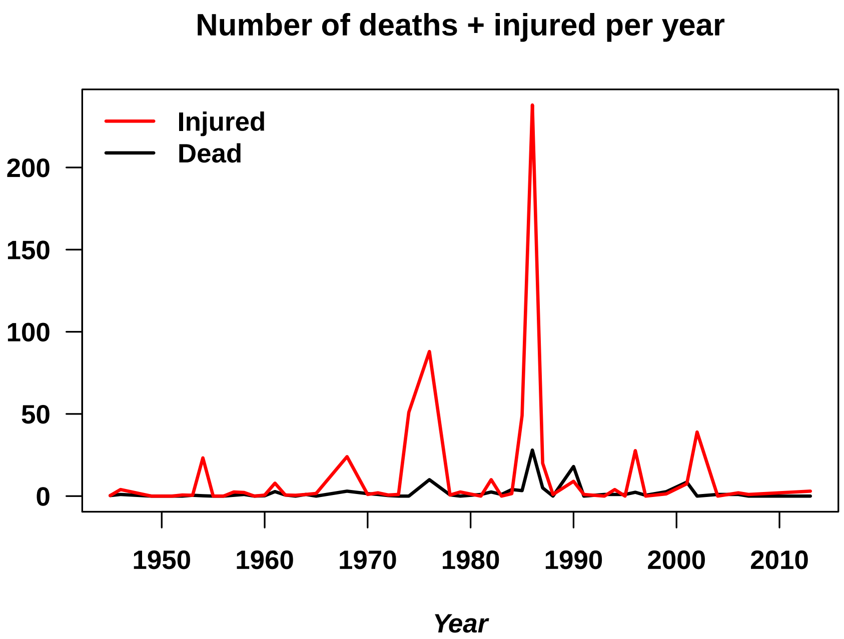


Figure 8: Number of death and injured per year, excluding WWII

The largest peak corresponding to Chernobyl (1986) and the second one is from a radiotherapy accident in Ohio, USA at 1976.

6.3 Analysis on radiology emitted

Another interesting analysis would be to also into nuclear emissions across all years, since radioactivity is known to affect biological tissue in the long term if the exposure is not so intense. That would need the use of regular expressions to clean the data, find the location of peaks of emission (now the location of the incident is important) and cross-reference with future mortality rates in that location.

7 Discussion

The limitations found in this study were mainly regarding the read of data, since we are reading several URLs and we had to put them inside the code. A possible solution would be to read those hyperlinks directly from the main page and "browse" the page from the code, accessing then the links of interest. Other limitations were found when trying to organize the data since different tables had different order of columns, number of columns, etc. So that's why the "standarize table" function was made, although some information was lost in the process that we actually can use (columns of certain tables were dropped because they were not shared across all tables). Another thing that could be done in the future is also look and clean data from radiation emitted, since it's time consuming and a little bit more complex (in terms of regular expressions) than number of deaths and injured.

A R Code used

```
> ##### PACKAGES
> if (!require(rvest)) install.packages("rvest", dep=TRUE); require(rvest)
> if (!require(stringr)) install.packages("stringr", dep=TRUE); require(stringr)
> if (!require(XML)) install.packages("XML", dep=TRUE); require(XML)
> if (!require(RCurl)) install.packages("maps", dep=TRUE); require(RCurl)
> ##### FUNCTIONS
> arrange_table <- function(one_table){
+   columns.use <- c("date", "location", "accident", "event", "code", "deaths",
+                    "injuries", "dose")
+   columns.use.pat <- paste0(columns.use, collapse = "|")
+   columns.select <- c("date", "location", "event", "code", "deaths",
+                       "injuries", "dose")
+   one_table <- one_table[, grep(columns.use.pat, colnames(one_table))]
+   colnames(one_table)[grep("location|event|accident|dose", colnames(one_table))] <-
+     c("location", "event", "dose")
+   one_table <- one_table[, columns.select]
+   return(one_table)
+ }
> odd <- function(x) x%%2 != 0
> even <- function(x) x%%2 == 0
> ########### CODING
> html <- getURL("http://www.johnstonsarchive.net/nuclear/radevents/radaccidents.html")
> doc = htmlParse(html, asText=TRUE)
> plain.text <- xpathSApply(doc, "//li", xmlValue)
> table.codes <- plain.text[1:(min(grep("highest", plain.text)) - 1)][-1]
> table.codes <- unlist(strsplit(table.codes, "\r\n"))
> table.codes <- table.codes[grep("--", table.codes)]
> table.codes <- unlist(strsplit(table.codes, "--"))
> codes <- gsub(" ", "", table.codes[odd(1:length(table.codes))])
> descr <- sub("^ ", "", table.codes[even(1:length(table.codes))])
> descr <- gsub(" \\(.*", "", descr)
> table.codes <- data.frame(code = codes, description = descr)
> ########### List of deadliest incidents
> html <- getURL("http://www.johnstonsarchive.net/nuclear/radevents/radevents1.html")
> tables <- readHTMLTable(html, stringsAsFactors = FALSE)
```

```
> dead <- tables[[1]]
> dead <- arrange_table(dead)
> ########### List of criticality accidents
> html <- getURL("http://www.johnstonsarchive.net/nuclear/radcrit.html")
> tables <- readHTMLTable(html, stringsAsFactors = FALSE)
> critical <- tables[[1]]
> critical$code <- "AC"
> critical <- arrange_table(critical)
> ########### List of naval reactor accidents
> html <- getURL("http://www.johnstonsarchive.net/nuclear/radevents/radevents3.html")
> tables <- readHTMLTable(html, stringsAsFactors = FALSE)
> naval <- tables[[1]]
> naval <- arrange_table(naval)
> ########### List of criminal incidents
> html <- getURL("http://www.johnstonsarchive.net/nuclear/radevents/radevents2.html")
> tables <- readHTMLTable(html, stringsAsFactors = FALSE)
> criminal <- tables[[1]]
> criminal <- arrange_table(criminal)
> ########### List of nuclear test accidents
> html <- getURL("http://www.johnstonsarchive.net/nuclear/radevents/radevents4.html")
> tables <- readHTMLTable(html, stringsAsFactors = FALSE)
> tests <- tables[[1]]
> tests <- arrange_table(tests)
> ## ALL IN ONE TABLE
> all.data <- rbind(criminal, critical, dead, naval,tests)
> all.data <- all.data[!(all.data$date == ""),]
> # Last 4 digits in date
> years.in.order <- sort(unique(as.numeric(sub('.*(\\d{4}).*', '\\1', all.data$date))))
> all.data$date <- factor(sub('.*(\\d{4}).*', '\\1', all.data$date))
> par(font = 2, font.lab = 4, font.axis = 2, las = 1)
> cols <- c("grey", "red")[(names(table(all.data$date)) >= 1947 &
+                              names(table(all.data$date)) <= 1991) + 1]
> barplot(table(all.data$date), xlab = "Year",
+         ylab = "Number of incidents",
+         main = "Number of incidents per year",
+         col = cols)
> legend("topright", c("Cold War", "No Cold War"),
+        col = c("red", "grey"), lty = 1, lwd = 3, cex = 0.8,
+        bty = "n")
> # First digits in deaths
```

```r
> all.data$deaths2 <- gsub(",", "", all.data$deaths)
> all.data$deaths2 <- as.numeric(gsub(" \\(.*", "", all.data$deaths2))
> par(font = 2, font.lab = 4, font.axis = 2, las = 1)
> plot(all.data$date, all.data$deaths2, main = "Number of deaths per year",
+      xlab = "Year")
> with(all.data[all.data$deaths2 < 10000,],
+      plot(date, deaths2,
+           type = "p", xlab = "Year",
+           main = "Number of deaths per year excluding WWII"))
> agg.by.year.d <- aggregate(all.data$deaths2[all.data$deaths2 < 10000],
+                            by = list(all.data$date[all.data$deaths2 < 10000]),
+                            FUN = mean, na.rm = T)
> agg.by.year.d$Group.1 <- as.character(agg.by.year.d$Group.1)
> colnames(agg.by.year.d) <- c("date", "deaths")
> plot(agg.by.year.d$date, agg.by.year.d$deaths, type = "l")
> # First digits in injuries
> all.data$injuries2 <- gsub(",", "", all.data$injuries)
> all.data$injuries2 <- as.numeric(sub("([0-9]*).*$", "\\1",
+                                      gsub(" \\(.*", "", all.data$injuries2)))
> plot(all.data$date, all.data$injuries2)
> with(all.data[all.data$injuries2 < 10000,], plot(date, injuries2, type = "p"))
> agg.by.year.i <- aggregate(all.data$injuries2[all.data$injuries2 < 10000],
+                            by = list(all.data$date[all.data$injuries2 < 10000]),
+                            FUN = mean, na.rm = T)
> agg.by.year.i$Group.1 <- as.character(agg.by.year.i$Group.1)
> colnames(agg.by.year.i) <- c("date", "injuries")
> plot(agg.by.year.i$date, agg.by.year.i$injuries, type = "l")
> # Deaths + Injuries
> agg.by.year <- merge(agg.by.year.d, agg.by.year.i, by = "date")
> plot(agg.by.year$date, agg.by.year$deaths, type = "l", col = "black",
+      ylim = c(0, max(agg.by.year$injuries, na.rm = T)), lwd = 2,
+      main = "Number of deaths + injured per year",
+      xlab = "Year", ylab = "")
> lines(agg.by.year$date, agg.by.year$injuries, type = "l", col = "red", lwd = 2)
```