Challenge 2 : Surviving the zombie apocalypse

Luther Ollier, Claribel Manzano

A zombie epidemic started in Rize, a Turkish town, on August 18th, 2019. The world goes through its worst crisis ever. Zombies have no mercy. The entire population in Rize, 94,800 inhabitants, were transformed into zombies. They are hungry for brains and are willing to do anything to get them. The epidemic will inevitably spread throughout Europe and we would like to make some predictions about the progress of this epidemic to monitor the best countermeasures.

We will work with a graph of Europe that will represent the zombies ability to spread through the continent according to the human density and the ground level.

Synopsis :

- The challenge
- The Data
- Epidemic model
- Stop the epidemic : with soldiers
- Stop the epidemic : with nuclear bombs

The challenge :

A zombie pandemic has started, the world will no longer be the same as before, people die and transform quickly, a single bite is deadly..

However, we are not afraid, let's counter attack and save what can be save !

We are currently in a bunker located in Brest. Our first main concern will then be whether or not this pandemic will reach us and, if so, when it will happen. To find out this, we will be studying the behavior of zombies : how they move from one place to another, how many humans they kill every day, how many zombies are killed by humans, how their population increases and finally, we will take into account that they die 15 days after starting to live.

After two months, if this situation continues, we will apply different strategies to help the European Union to locate troops of soldiers in different places of Europe and thus, try to save as many humans as possible.

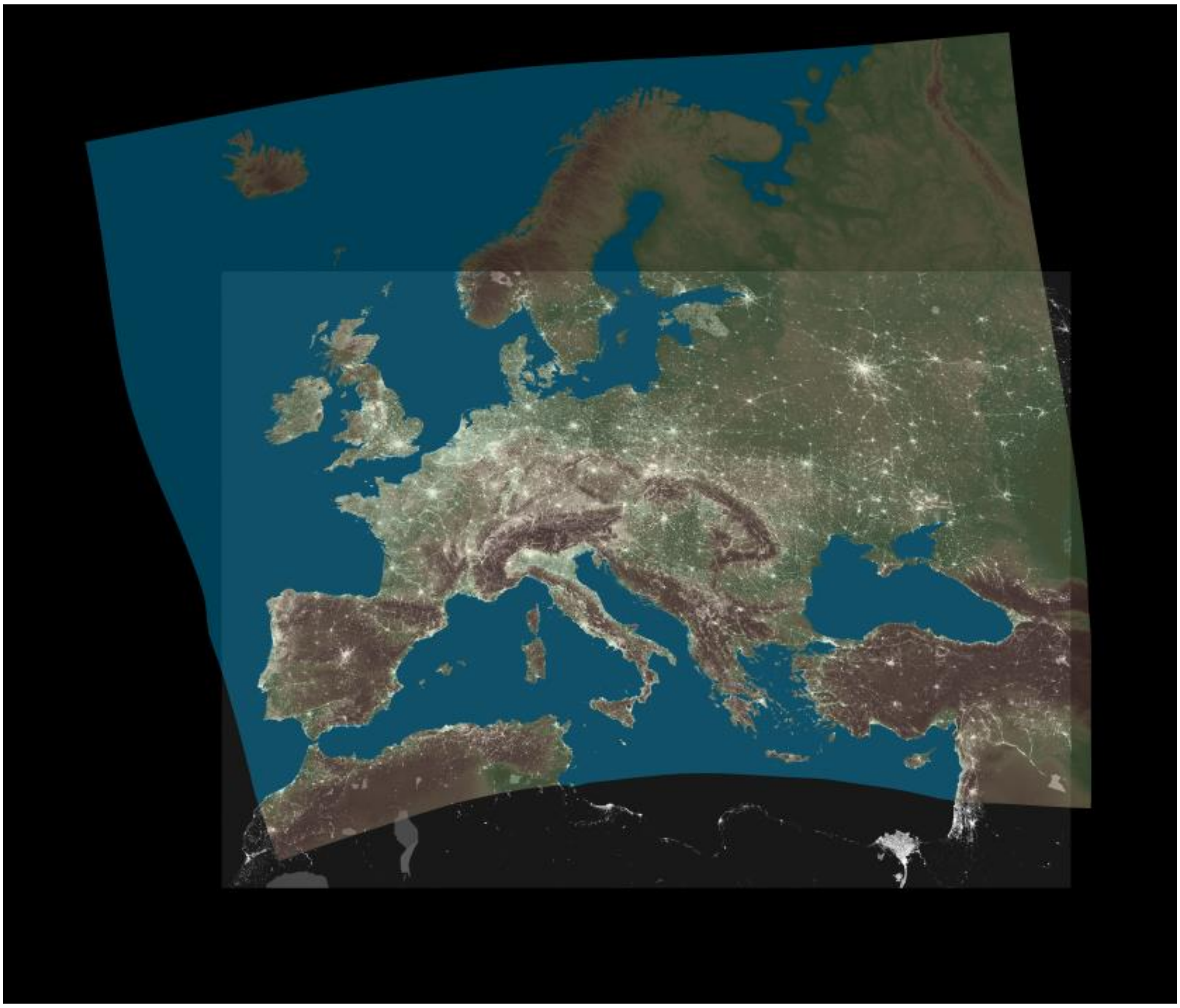
Lastly, if the pandemic does not stop, after four months there will be no other choice but to drop nuclear bombs in strategic locations and our job will be to find out exactly where and how many to drop to stop the pandemic, killing as many zombies as possible and as few humans as possible. And all this work will always make us wonder: Will Brest be preserved?

Data Science, graph theory and social network studies

The Data :

First, we will make some predictions on the advancement of this epidemic, trying to know if it will reach Brest or not, and when it will occur.

To begin with our investigation, we have two maps of Europe and part of Asia.
One of them represents the human population, each pixel on this map takes a value from 0 to 255 according to the grayscale. The whitest pixel, value 255, corresponds to a population of 3,000 inhabitants, and the blackest pixel corresponds to 0 inhabitants.
Our second map is the relief and shows us, by colors, the different heights of the terrain with respect to the sea, the highest point will be Mont Blanc and the lowest, any light blue point corresponding to the sea.
All these values will be later normalized in order to get 500 millions inhabitants in Europe and the elevation values will be smooth in order to have a continuous ground.
We needed also to overlay the two map (we used photoshop to obtain the result below):



Due to the poor quality of the data (internet has been shut down since the beginning of the epidemic so we do with what we have) we have been forced to make some choices (smoothing the height of the cells, the way degree is compute, the number of humans) in the preprocess of the data in order to get the more trustful results (see the notebook).

The next step after overlaying our maps and create cells of 20x20 pixels that will be the nodes in our graph. These cells will be connected to each other by edges only if the elevation angle between both cells is less than 10, otherwise, we can say that the zombies will not be able to cross from one cell to the other due to this height difference and we will not connect these cells. To be more accurate, 20x20 pixels represents 400 km^2 , so to compute the slope between two cells we take the mean height of both cells and consider a squared triangle of side 400 km^2 and the difference between the two height of the two cells.

A node will have eight linked nodes maximum : the eight neighboring cells. It will also have three attributes: the number of zombies, the number of humans and the height of the cell. It will allow us to simulate the propagation of the outbreak. The two first attribute of the graph will be function of time and we will simulate the spreading until convergence, as humans don’t reproduce, we can assume that the outbreak will inevitably converge.

Of course, as zombies can’t go through sea, we deleted from our graph the smallest non-connected part as the islands (Great Britain for example).

This first part will allow us to get a good understanding of the outbreak spreading and to build a model helping us to test the efficiency of countermeasures.

Epidemic model :

Propagation of Zombies

To determine if the pandemic will reach Brest, we are going to investigate the movement of zombies in Europe. We will use our map divided into cells. Each of them will have three attributes, the human population (function of time), the zombie population (function of time) and the height of the area with respect to sea level. Zombies can move each day to a neighboring cell, so if we consider a specific cell, it will have a maximum of eight others around it that a zombie can reach. All this behavior will give us a different graph every day, where the population of humans and zombies will change.

Firstly, if there are zombies on a cell, the eight neighboring cells will receive a certain contribution of zombies from the central cell, depending on different factors such as population density of humans, number of zombies, the slope between both cells, etc.

When there are no humans around a cell, zombies just stay on it. When there are humans in its neighboring cells, all the zombies go there even if there are already humans in this cell because then others zombies from the neighbors will come to them.

At the end of this displacement, the population of zombies in a cell will be the sum of all the contributions of its eight neighbors, taking into account that this amount will depend only on the number of zombies that were in them the previous day, not the actual one.

Secondly, once the zombies have already moved, they will now start attacking humans. Each one who has entered into the cell will kill ten humans who will eventually convert themself and join the total population of zombies. We will take in account the scale of the population to do this correctly.

Finally, once the zombies have attacked and killed, it will be the remaining humans' turn to destroy each one ten zombies, leading to a reduction in their population.

There are more information that we will have to take into account too, for example, that zombies have a life of 15 days and, after that time, they will self-destruct. And to do this we will eliminate the same quantity of zombies that were in the cell 15 days before.

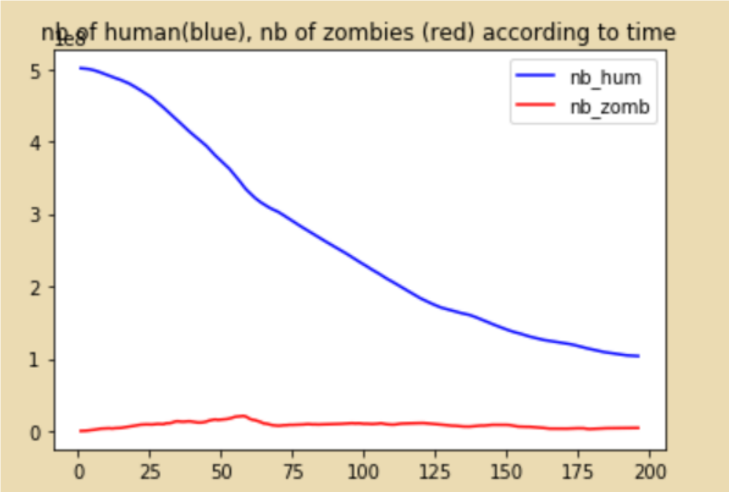
We also know that zombies cannot go from one cell to another if the elevation angle between them is equal to or greater than ten degrees, so we will not make them to be neighbors, they will not be connected, the movement will not happen.

Let's run the first simulation of the potential end of human kind !

First simulation

After running a first simulation, let's observe the havoc wreaked by zombies in Europe :

- 73 millions of survivors (83 % of Europe population has been exterminate)
- All people in Brest are dead at day 198 (too bad for us)
- The epidemic converges after 263 days



We can see that the number of zombies stays almost the same all the long (as they die after 15 days this is pretty logical) and that the number of humans steadily decreases. The survivors are in areas that can't be reached by zombies because of the elevation.

What we can deduce from this simulation is that the zombies progress as if the Europe population was equally distributed. This may come from the fact that the scale of the map giving the Europe inhabitants density has not enough values to really state the difference between cities and country.

Stop the epidemic : with soldiers

Two months after the start of the epidemic, the European Union will be able to send military troops. We work on finding twenty cells that will be secured by them, considering that they have to allow the protection of the maximum amount of persons. Troops will be transferred instantly to those cells, where all zombies will be destroyed and no zombie can enter anymore. Once we have selected the cells, we aim to analyze the impact of our choice to the spreading of the epidemics and if Brest will be preserved. To this end, we will use the model built in the first part. It will teach us the state of the outbreak two months after the start of the epidemic and we will use it to find the best twenty cells to “clean” in order to save the most humans we can.

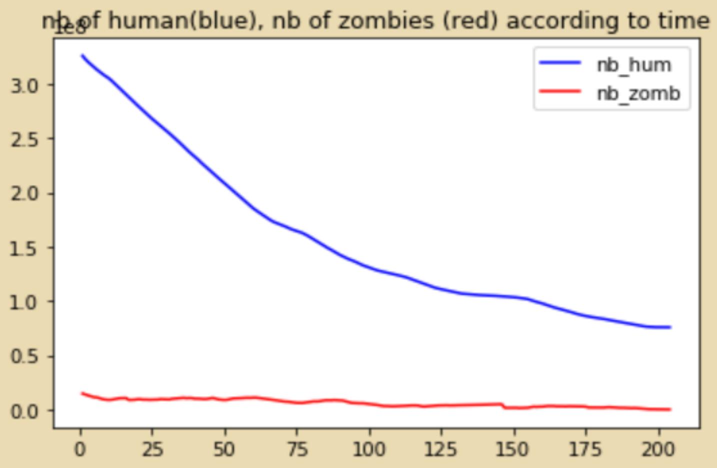
First, let’s define a metric : We need to save the more humans we can, so let’s try to minimize the following ratio :

$$\frac{\text{Number of humans death between 60 days and the end of the simulation with soldiers}}{\text{Number of humans death between 60 days and the end of the simulation without soldiers}}$$

After trying a few solutions as securing the 20 cells where there is the most zombies, computing the approximate betweenness of cells and then choose the 20 with both the highest betweenness and the highest number of zombies etc.. we had to give up to save humanity with soldiers.

Indeed, we didn’t even succeed to slow the outbreak convergence. Below a result of an attempt where we changed the simulation rules in order to keep zombies attracted by cells with soldiers. You can see that only 2 millions more people are saved and that zombies reach Brest on day 198 (138+60).

```
number of humans at Brest on day 138 : 9487.0
number of humans at Brest on day 139 : 0
number of human still alive : 75754876.0
number of human killed : 429381347.0
ratio from the first simulation : 0.9899470738315811
```



We can explain theses poor results by the fact that the outbreak is already too spread. Indeed there is 116 cells with zombies on day 60, even if we destroy 20 of them, it will still remains enough zombies to “eat” the remaining humans.

Let's move on the last solution to save humanity: Nuclear bomb !!

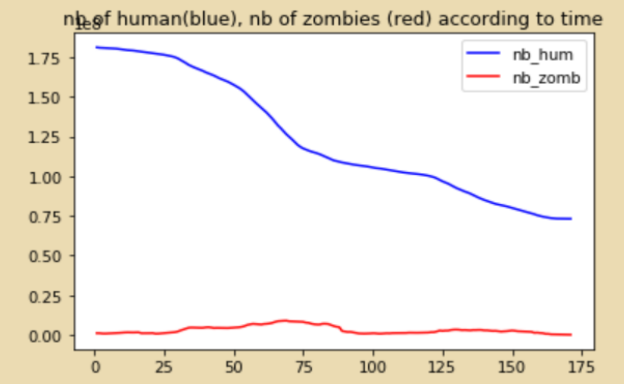
Stop the epidemic : with nuclear bombs

Four months after the start of the epidemics, the European Union takes the decision of using nuclear weapons to save the rest of humanity. Nuclear bombs not only destroy entire cells, meaning all life forms including all zombies, but also, after a cell has been bombed, it cannot be traversed by zombies anymore. We will work to determine how many nuclear bombs will be necessary to drop, considering that there are an unlimited quantity. Furthermore, it will be also necessary to know in which cells they should be dropped to obtain the best possible ratio of (humans killed by bombs) / (humans saved by total destruction of zombies).

What remains of the solution with the soldiers is that if there still zombies on a cell at the end of our nuclear attack, they will spread through Europe whatever cells we focused. After 120 days there is only 41 cells occupied by zombies and among theses cells 0 humans (because they all have been eaten).

The first solution we tried is to hit all the cell with zombies and that have 1000 humans or more in the neighboring cells, see results below (the ratio value is 1, the approximation of the computation added some undesired significant numbers).

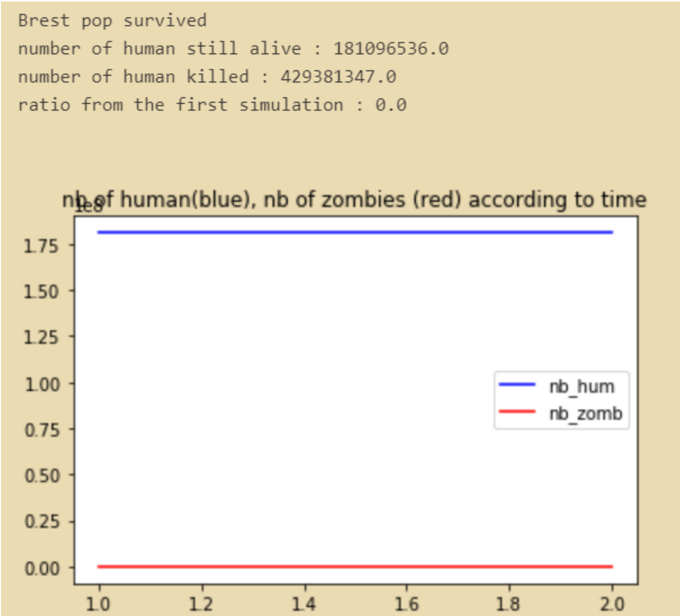
```
number of humans at Brest on day 105 : 9487.0
number of humans at Brest on day 106 : 0
number of human still alive : 73132295.0
number of human killed : 429381347.0
ratio from the first simulation : 1.0007828714156224
```



Once again, the number of human killed remains very high and we don’t manage to save lots of life. However this time we succeeded in slowing the outbreak : 290 days to converge and Brest is reached on day 225 (120 + 105).

The problem is that even if we forget that the Europe inhabitants density is not very accurate, the continent is very smooth. We mean by that that it always exist several ways for zombies to reach a point (even Brest) and trying to block theses ways is the same as killing all the zombies.

It gives us no other choice but to hit all cells with zombies.



The solution is extreme but worked. The outbreak has been eradicated and we saved 110 millions of Europeans. The only good news is that we needed only 41 bombs to achieve this and we can reduce this number to 21 if we use the soldiers of the first solution.

Conclusion

As Europe is attacked by the worst outbreak in its history, we tried to modelized the propagation of the epidemic and to explore solutions in order to protect life. Several challenge had been faced.

First the complexity of Data preprocessing. In real life, we don't always have a very clean dataset or a graph prebuild. We needed to take decisions on the data design in order to build the model. We aimed to get the more accurate data but as we had to balance with the computation time, we had to make several simplifications. In the end, it's probably the fact that we couldn't achieve to make a good representation of Europe inhabitants density and Europe elevation that made all the solutions we tried so bad.

Second the complexity of the model. Due to time computation we had to make simplifications (for instance in our model the way we implement the fact that zombies die after 15 days is not very efficient).

However, to conclude, may be the model we choose is a good approximation and it is impossible to save most humans from a zombie epidemic. We have to keep in mind this possibility. Let's follow the advancement of the epidemic in real life and if our predictions are correct after a few days, we can deduce that the only way of stopping the zombies is to destroy all of them. To this end, 21 nuclear bombs are requires and 20 military interventions groups as we showed in this report.

Useful links:

[CRISP-DM method] : https://medium.com/@christophberns/using-crisp-dm-to-predict-car-prices-f15eb5b14025

Centrality measure :

- Betweenness : https://en.wikipedia.org/wiki/Betweenness_centrality

Centrality measure on big graph :

http://matteo.rionda.to/centrtutorial/BonchiDeFrancisciMoralesRiondato-CentralityBigGraphsTutorial-Slides.pdf