

Basics of Data Science

Assignment

Rommel Labastida (Cohort 1)



Why are so many people dying of COVID-19 in Mexico?

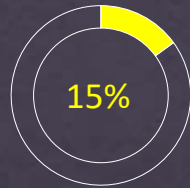
Source: [Forbes](#)

Summary

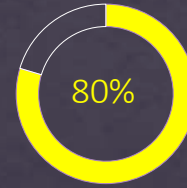
65,241

Deaths

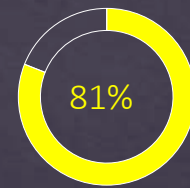
Private hospitals deaths



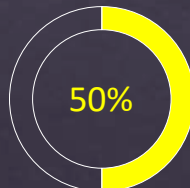
Died w/o access to ventilators



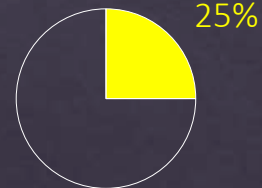
35-55yrs – Died without connection to ventilators



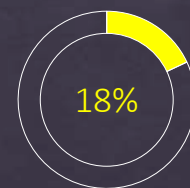
Public hospitals deaths



Died who are obese



Healthy 35-55yrs – connected to ventilators



Why Data Science is best suited for this article



Data Science can help answer questions like the following:

- Are there really more people dying of the virus?
- How are the public hospitals doing compared to private hospitals?
- Does the virus increase the probability of dying of an obese or diabetic person?
- Can some deaths be prevented?
- How is Mexico's managing the fight against the virus compared to the world?

A little bit of Data Science

606,036

Infection

65,241

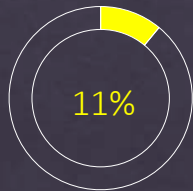
Deaths

Tests 3 per 100,000
people



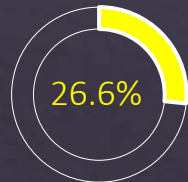
Comparison: US tests 150 per 100,000

Mortality rate

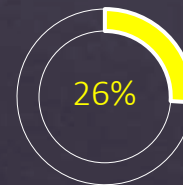


World: 3.3%

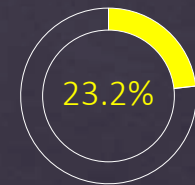
Survived with
ventilators



Survival of obese person

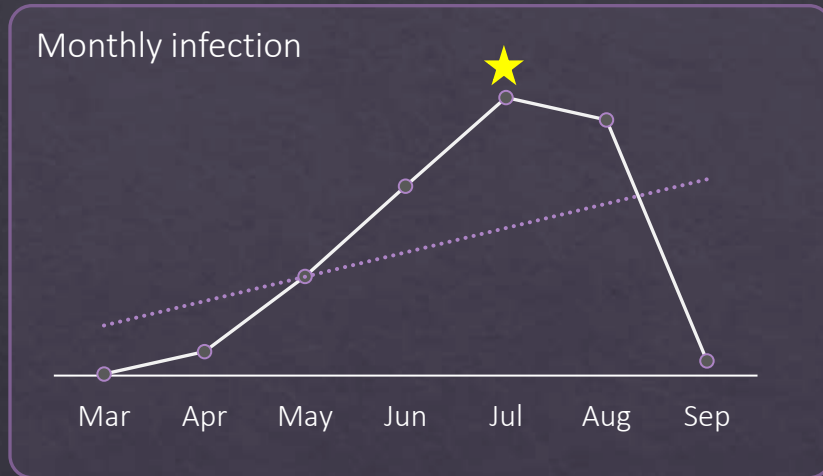


Survival of a diabetic

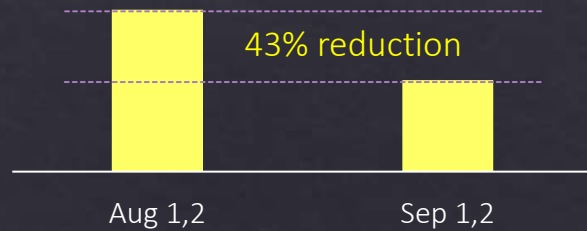


So, are there really more people
dying in Mexico of Covid-19?

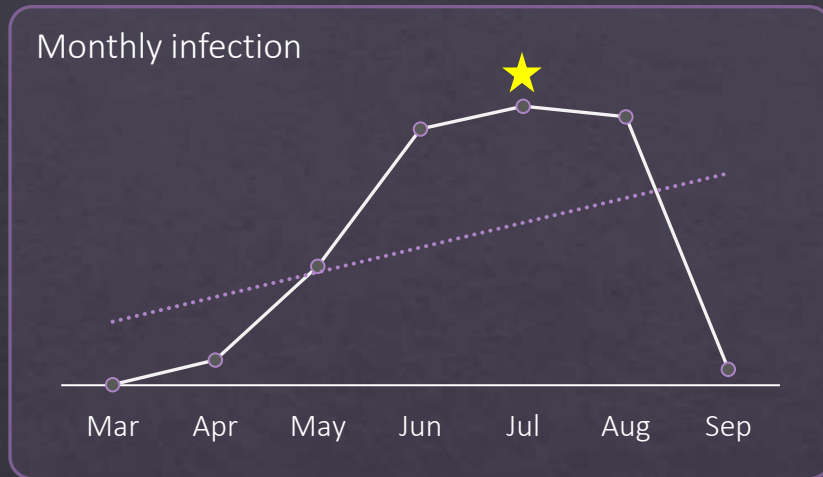
Are there really more people dying in Mexico? (Infection rate)



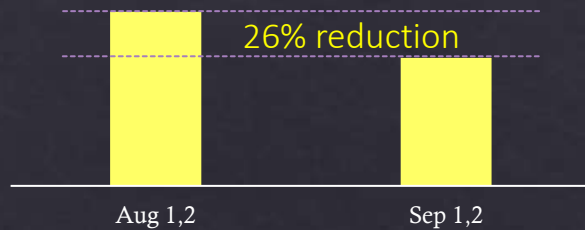
- Trend slowly decreasing
- ★ ➤ The infection has peaked in July
- 18,014 infection in Aug 1, 2
- 10,195 infection in Sep 1, 2



Are there really more people dying in Mexico? (Death rate)



- Trend started to decline
- ★ ➤ Number of death has peaked in July
- 1,472 deaths in Aug 1, 2
- 1,083 deaths in Sep 1, 2



Conclusion



So, did Data Science helped answers these questions?

- Are there really more people still dying of the virus?
- How is Mexico's managing the fight against the virus compared to the world?
- Does the virus increase the probability of dying of an obese or diabetic person?
- How are the public hospitals doing compared to private hospitals?
- Can some deaths be prevented?

Data Science aspects



Predictive Analytics – forecasting on what situation an infected person may die

Descriptive Analytics – describing the data by the numbers

Prescriptive Analytics – data shows what the Mexico government can do to reduce at least the deaths due to non-access to ventilators

Correlation – dying without having connected to the life-saving ventilators and dying without even getting admitted to hospital

Insight – being obese or diabetic does not mean death sentence when infected. The biggest factor of survival of both healthy and obese/diabetic patient is access to ventilators

Noise data – skipping all the political issue where beds may be available or deflecting the blame to chronic diseases

Why data science is not suited for this article



- No reason why data science can't be used in this situation
- This is a typical example of data science application
- Data science is almost always best used in this kind of information

Pitfalls



- Inaccurate data due to complexity of data collection
- Other data not related to Covid-19 that can affect the number of deaths are ignored
i.e. deaths due to chronic diseases in the previous year
- Data is influenced by noise like the government propaganda

Success factors



- Noise data is ignored
- Error variance is recognized and accepted

Q & A