Predictive Data Analysis

Lecturer: Marina Iantorno

E-mail: miantorno@cct.ie

cct | College Dublin

Computing • IT • Business


cct
College Dublin
Computing • IT • Business
WELCOME
TO THE
COURSE
SAMSUNG

What is Statistics?

Statistics is a form of mathematical analysis that uses quantified models, representations and synopses for a given set of experimental data or real-life studies.

Statistics include numerical facts and figures.

For example: Compared to women, men are three times more likely to die by suicide.

Statistics is NOT only numbers and figures. The numbers may be right, but the interpretation may be wrong.

For example: The more churches in the city, the more crime there is. Thus, churches lead to crime.

Statistics refer to a range of techniques and procedures for collecting, analysing, interpreting, displaying and making decision based on data.

What is Statistics?

We can find two different kind of events:

- **Deterministic event:** This is an experiment that could tell us precisely what will happened under certain conditions. For instance, if we put in a jar with one litre of water in it on the fire at 100 degrees Celsius, we know that it will boil after 3 minutes, and if we repeat this experiment many times, we will always get the same result.
- **Probabilistic event:** This is an experiment that depend on chance, and therefore we could get always different results even though we repeat the same conditions. For instance, we can roll a dice several times and although we shake the glass three times and roll it on the same surface, we can get always different results.

What is Statistics?

We talk about probabilities every day and maybe we do not realize. We are talking about probabilities when:

- We check the weather on our phone and it says that today there is a 70% chances of rain/have a sunny day.
- We mention that 9% of the global population live in poverty.
- We play the lottery.
- Choose a card from a deck.
- Discuss about possible results in sports.
- When we discuss about the prices.

What is Data?

When we talk about data, we are talking about units of information. A sole value by itself does not really explain anything, but when we observe that value in a context, it could mean different things. In this sense, it is important to know that there are two different kind of data:

- Numerical data.
- Categorical data.

In this course we will focus on numerical data, because this is the one that allow us making predictions.

Exploring Predictive Analytics

Imagine a world in which you can maximise the potential of every moment of your life. This is what organisations try to find. Companies spend million or euros yearly to manage the risks, optimise their operations and maximise their profits. This is done through predictions, and here is where Predictive Analytics takes place.

Nowadays, data is the most valuable asset and that is why the richest companies in the world are those that hold the maximum number of data, for example Facebook and Google.

Exploring Predictive Analytics

Companies collect, examine and process data to search ways to better understand their customers, outperform their competitors and get a better position in the market.

Governments use data to understand the society and thus develop politic campaigns, economic plans, budgeting and investing.

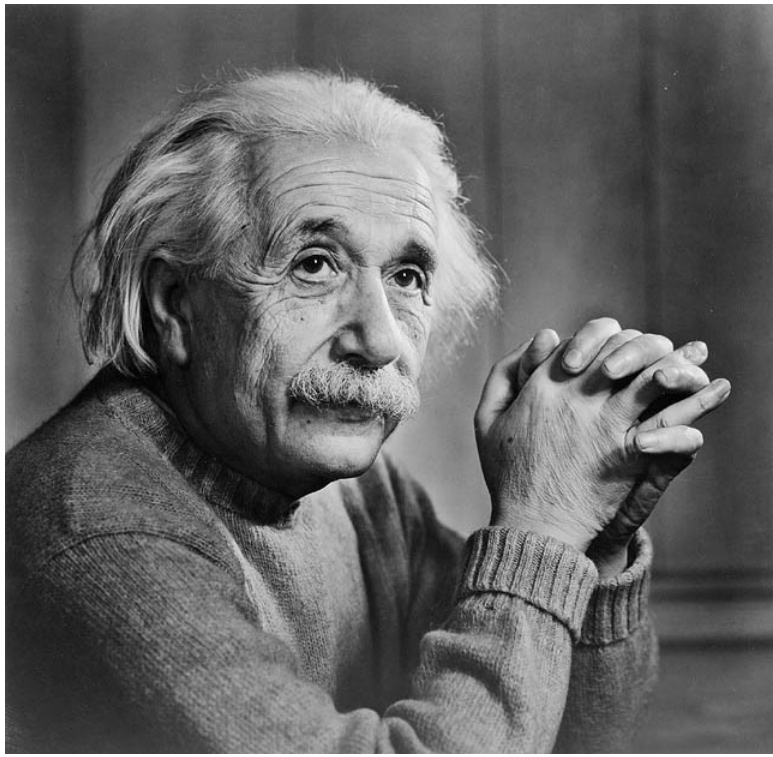
Research uses data to prove some studies and analysis, and this is useful for all possible fields.

Exploring Predictive Analytics

We collect data from past occurrences, look at the history and build models –mathematical representations- and ultimately make predictions based on that.

Albert Einstein said: “Know where to find information and how to use it; that is the secret of success”.

Predictive Analytics will help us to succeed. It will lead us to see what is invisible or not so clear to others, and therefore we will understand the patterns of our data.



What to expect from this course?

After successful completion of this course you will be able to:

- Understand different models to analyse data.
- Make predictions in different fields.
- Use concepts of Statistics, Machine Learning and Data Science to understand the data and transform it into information.
- Work with large datasets.
- Use Software like Python.

Predictive Data Analytics