



**EDUCACIÓN**  
SECRETARÍA DE EDUCACIÓN PÚBLICA



TECNOLÓGICO  
NACIONAL DE MÉXICO®



**Tecnológico Nacional de México  
Instituto Tecnológico de Tijuana**

**Subdirección Académica  
Departamento de Sistemas y Computación**

**Semestre:**

Febrero – Junio 2021

**Carrera:**

Ingeniería en Tecnologías de la Información y Comunicaciones

**Materia y serie:**

Minería de datos

BDD-1703TI9A

**Unidad a evaluar:** Unidad II

**Nombre de la Tarea:**

Origen y gramática de las gráficas

**Nombre del Alumno:**

Sifuentes Martinez Manuel Javier 17212934

**Nombre del docente:**

José Christian Romero Hernández

## Origin and grammar of the graphs

### What is the origin of the graphs?

It is not easy to pinpoint the exact moment when statistical charts were born. The visual representation of information was used since very ancient times, for example in the form of geographical or celestial maps. But it could be said that it was not until the 18th century that statistical graphics as such emerged. Scientific, economic, geographical, social, demographic studies, etc., were increasingly important, and the different scientific disciplines were progressing more and more. Measurement instruments and techniques were becoming more precise and better, and ultimately, as a result of experiments and observations, a set of data was obtained, which was recorded and analyzed. Graphing those data was a natural step.

For a brief description, the evolution and development of statistical graphics can be divided into four major stages:

#### **First stage: 1600 - 1699**

In this first stage, graphic representations of the results of observations and scientific experiments arise. Noteworthy is the 1644 graph by Michael van Langren, a Dutch astronomer who worked at the court of Philip IV, which collected the different estimates that had been made of the distance between Toledo and Rome (12 in total).

In 1669 Christiaan Huygens made a graph showing the remaining years of the Second Stage: 1700 - 1799 given as a function of age.

#### **Second stage: 1700 - 1799**

The graphics that already existed (mainly maps) continue to develop, and new forms of representation appear. Joseph Priestley, Johann Heinrich Lambert and especially William Playfair are the most notable figures of this period.

- Joseph Priestley (1733 - 1804): He was one of the first to create graphs of historical events using the time line. His intention was to offer a more analytical and comprehensive vision of History than what had been done to date. Later his charts inspired Playfair to create the bar chart.
- Johann Heinrich Lambert (1728 - 1777): Introduced the idea of interpolation from empirical data. He used various types of graphs to represent the periodic variations in air and ground temperatures.
- William Playfair (1759 - 1823): He was the inventor of the bar and pie chart, and also a pioneer in using the line graph to represent time series of data.

### **Third stage: 1800 - 1899**

The innovations that appeared in the previous century are taken advantage of. Among others, the graphics of Charles Minard, Florence Nightingale and Francis Amasa Walker stand out. In addition, the genetic inheritance studies of Francis Galton and the astronomy studies of John FW Herschel facilitated the creation of the scatter plot.

- Charles Joseph Minard (1781-1880): In the field of graphics, he stands out for the use of cartograms, or maps on which statistical data are represented.
- Florence Nightingale (1820 - 1910): She is the inventor of the polar area graph, which she used to represent mortality data for British Army soldiers in the Crimean War.
- Francis Amasa Walker (1840 - 1897): He published, in 1874, the work The Statistical Atlas of the United States, based on the data of the 1870 census. This atlas was a compendium of graphs and tables that offered geological, economic and social information .
- Francis Galton (1822 - 1911): Thanks to his studies and innovative data analysis techniques and graphical representations included, and to the astronomy works of JohnF. W. Herschel devised and extended the use of the scatter graph, used to study the relationships that may exist between two variables referring to the same set of individuals.

### **Fourth stage: 1900 - 2010**

During the first half of the 20th century, there were not too significant advances in the sense of inventing new forms of graphic representation, although its use began to become popular in all areas (scientific, economic, political, social, etc.).

The second half, especially after 1975, was undoubtedly marked by the advancement of computing. The development of statistical programs, new data manipulation and visualization models, the reinvention of graphical techniques to represent discrete and categorical data, etc. they fostered an explosive growth in the field of visual representation. Highlights the role of John Wilder Tukey, American statistician creator of the box and whisker diagram and Joseph Juran, Romanian engineer expert in quality management, creator of the Pareto diagram.

- Joseph Juran (1904 - 2008): Devised the Pareto chart.
- John W. Tukey (1915 - 2000): Creation of the box-and-whisker plot, which first appeared in his 1977 book Exploratory data analysis.

## What is the grammar of graphs?

An analogy (not to be interpreted literally) was established between the grammatical construction and the structure of a graph. In this way, equivalences such as those shown in table 1 were reached.

Analogy between elements of graphic constructions and linguistic concepts.

<b>Graphic</b>	<b>text Text</b>
Graphic	Sentence
Infographic (group of graphs)	Paragraph
Magnitude represented	Subject of the sentence
Graph with a single data series	Simple
Graph with a data series	Compound
Graph related to others in an infographic	Subordinate

From the analysis of many graphs, They drew some conclusions to recommend some tips with which to build a graph that would not give rise to confusion and would be faithful to its original message. They are the ones that we can see in the following table.

<b>Spelling</b>	The subject must be defined There must be agreement between the type of graph and the type of data They must include explicit, complete and well-referenced axes
<b>Syntax</b>	The magnitude represented must be the same for all the data Axes and scales must keep a certain proportion Adequacy information density (with respect to space or ink) A minimal composition should be incorporated and visual effects avoided
<b>Semantics</b>	With the same data there are different representations depending on the message Demagogy can be done with impeccable graphics Small formal errors help the <demagogic> effects The choice of the magnitude to present must be in accordance with the message
<b>Literary styles</b>	And in all of the above there is much room for personal taste

## References

Silo (2017) History of Statistical Graphics. April 28, 2021. Website: <https://silo.tips/download/indice-1-26-continua>

Barbeira, S (sf) The graphs also have a subject and a predicate. April 28, 2021. Website: <https://www.sergiobarbeira.com/gramatica-graficas/>