

**Iaac**

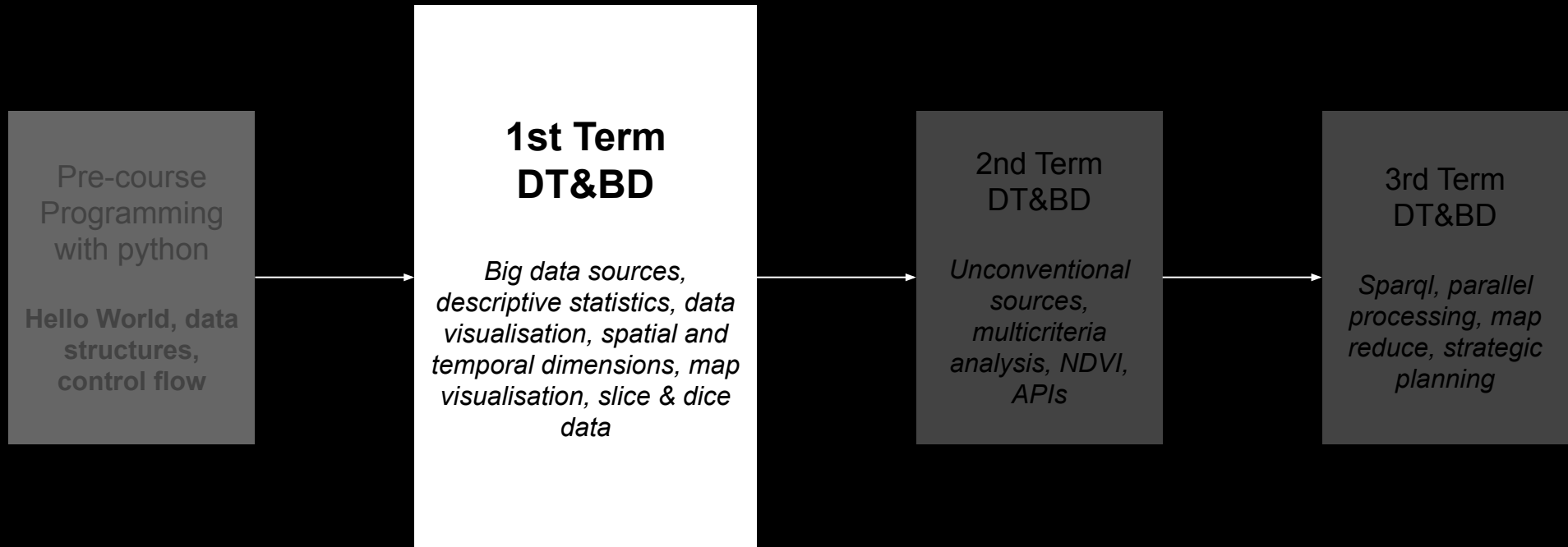
Institute for  
advanced  
architecture  
of Catalonia

BARCELONA

**MASTER IN CITY & TECHNOLOGY  
DIGITAL TOOLS AND BIG DATA  
2020/2021**

**FACULTY** DIEGO PAJARITO

# From data to Big data



Master in City and Technology

# Data / Datum

Facts and statistics collected together for reference or analysis.

**1.1 The quantities, characters, or symbols on which operations are performed by a computer, which may be stored and transmitted in the form of electrical signals and recorded on magnetic, optical, or mechanical recording media.**

**1.2 Philosophy Things known or assumed as facts, making the basis of reasoning or calculation.**

\* Geodetic datum

Definition from: [https://en.wikipedia.org/wiki/Geodetic\\_datum](https://en.wikipedia.org/wiki/Geodetic_datum)

# A definition made out of V's

Volume, Variety, Velocity and Value (+)

Paper published by ACM at: [The 10 Vs, Issues and Challenges of Big Data](#)

01	<b>Volume</b>	<ul style="list-style-type: none"> <li>• Data scale</li> </ul>
02	<b>Value</b>	<ul style="list-style-type: none"> <li>• Data usefulness in decision making</li> </ul>
03	<b>Velocity</b>	<ul style="list-style-type: none"> <li>• Data Processing: Batch and Stream</li> </ul>
04	<b>Veracity</b>	<ul style="list-style-type: none"> <li>• Data quality and accuracy</li> </ul>
05	<b>Viscosity</b>	<ul style="list-style-type: none"> <li>• Data complexity</li> </ul>
06	<b>Variability</b>	<ul style="list-style-type: none"> <li>• Data flow inconsistency</li> </ul>
07	<b>Volatility</b>	<ul style="list-style-type: none"> <li>• Data durability</li> </ul>
08	<b>Viability</b>	<ul style="list-style-type: none"> <li>• Data activeness</li> </ul>
09	<b>Validity</b>	<ul style="list-style-type: none"> <li>• Data properly understandable</li> </ul>
10	<b>Variety</b>	<ul style="list-style-type: none"> <li>• Data heterogeneity: Structured, Semi-structured, Un-structured</li> </ul>

## Some Facts

**3.8 billion connected**  
**7.6 billion population**

**8.06 billion connected**  
**13.4 billion devices**

¿2021?

Paper published by ACM at: [The 10 Vs, Issues and Challenges of Big Data](#)

**To provide an experience handling common tasks of big data, data science or data analytics.**

**The course provides a practical perspective of the main activities developed for urban analytics. From data collection, ingestion, analysis and visualization, the students will experience the workflow while getting their hands on extracting information from massive datasets.**



**Getting familiar with data repositories for Big Data**

**Spatial and temporal dimensions**

**Multilayer data management and visualisation**

**Describing our data using statistics**

**Data visualisation and statistical plots**

**Studio integration, GIS and pandas**



**Source Code**

**Examples of the tasks performed during the course**



**Poster / infographic**

**Windows/Linux/Mac Computer (Dual-Core + processor, RAM 8Gb)**

**Python 3.7.x or later (Get the installer at <https://www.python.org/downloads/>)**

**Anaconda (Get the installer at <https://www.anaconda.com/distribution/>)**

**Pycharm Community (Get the installers at <https://www.jetbrains.com/pycharm/download>)**

**QGIS Desktop (Get the installers at <https://qgis.org/en/site/forusers/download.html>)**

**Google Earth Pro (<https://www.google.com/earth/versions/#earth-pro>)**

**pandas <https://pandas.pydata.org/>**

**geopandas <http://geopandas.org/>**

**matplotlib <https://matplotlib.org/>**

**seaborn <https://seaborn.pydata.org/>**

**Missingno <https://github.com/ResidentMario/missingno>**

**Sublime text editor (get the installer at <https://www.sublimetext.com/3>)**

**Iaac**

Institute for  
advanced  
architecture  
of Catalonia

BARCELONA

**MASTER IN CITY & TECHNOLOGY  
DIGITAL TOOLS AND BIG DATA  
2020/2021**

**FACULTY** DIEGO PAJARITO