<https://github.com/GToritto/Homework12>

Big Data is extremely large and complicated datasets that cannot be easily managed or processed using regular data processing methods and tools. It is characterized by its volume, velocity, and variety, requiring advanced tools and techniques for processing and analysis. Big Data is very important because it allows for many things like data based decision making, innovation, custom experiences, and healthcare advancements. The three main types of Big Data are structured, unstructured, and semi-structured data. Structured data is well-organized and fits into traditional databases, while unstructured data is diverse and lacks organization. Semi-structured data lies between these two types. The three V’s are Volume, Velocity, and Variety. Volume involves: the processing and analysis of massive volumes of data that have higher capabilities of traditional databases and storage systems. These datasets go from terabytes to petabytes, exabytes and higher. Velocity means Big Data is generated at high velocity and requires real time or near real time processing. Data is continuously flowing in from different places like social media updates, website clicks, sensor readings, etc. The ability to process and analyze data quickly is important to extract meaningful insights and take smart actions. The third V is Variety. This involves the diverse data types and formats like structured, semi structured, and unstructured data. It contains numbers, text, images, videos, audio files, social media posts, emails, etc. Viewing and extracting value from these different types of data requires very highly advanced techniques and tools. Understanding Big Data has big potential for learning insights and having growth in different domains.

<https://github.com/GToritto/Homework12>