

GRAPH AND SOCIAL NETWORK VISUALIZATION

STEFANO FOSSEN

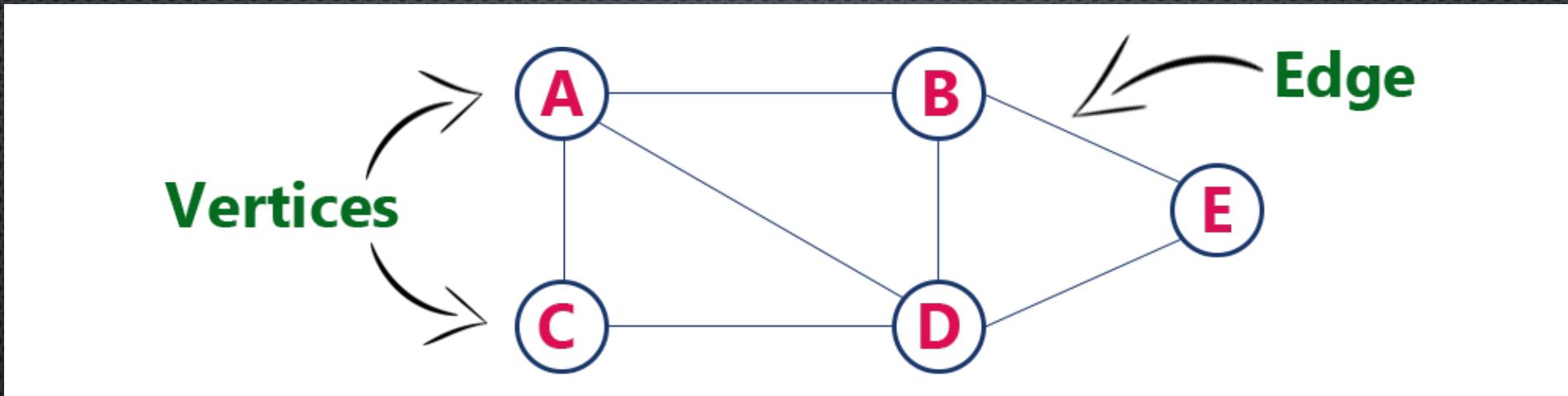
DIPARTIMENTO DI INFORMATICA
UNIVERSITÀ DEGLI STUDI DI MILANO

INDICE:

- GRAPH
- GRAPH THEORY
- GRAPH VISUALIZATION

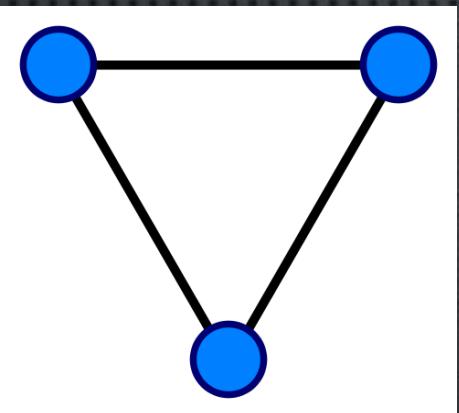
GRAPH

WHAT IS A GRAPH?

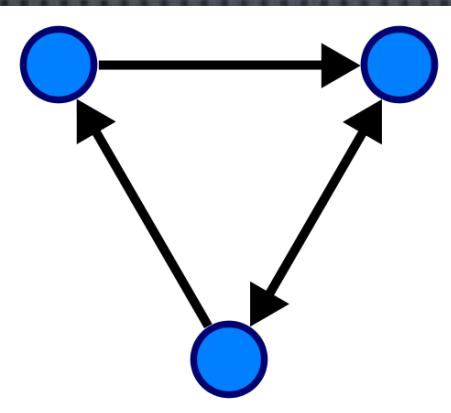


- A GRAPH, ALSO CALLED NETWORK, IS A COLLECTION OF VERTICES, OR NODES/POINTS, AND EDGES, OR LINKS/LINE
- EACH NODE REPRESENTS A SINGLE DATA POINT AND EACH EDGE REPRESENTS HOW TWO NODES ARE CONNECTED.

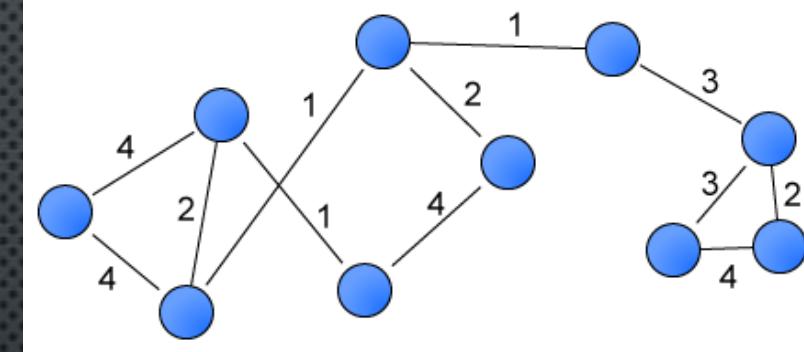
TYPES OF GRAPH



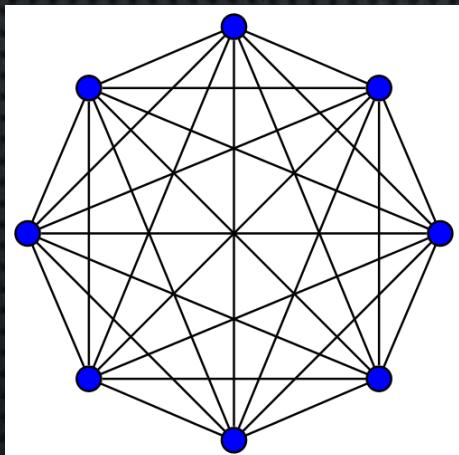
UNDIRECTED GRAPH



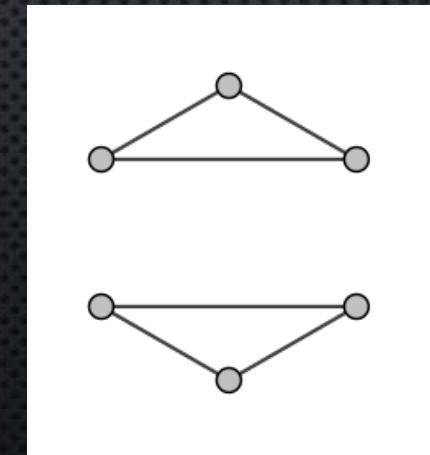
DIRECTED GRAPH



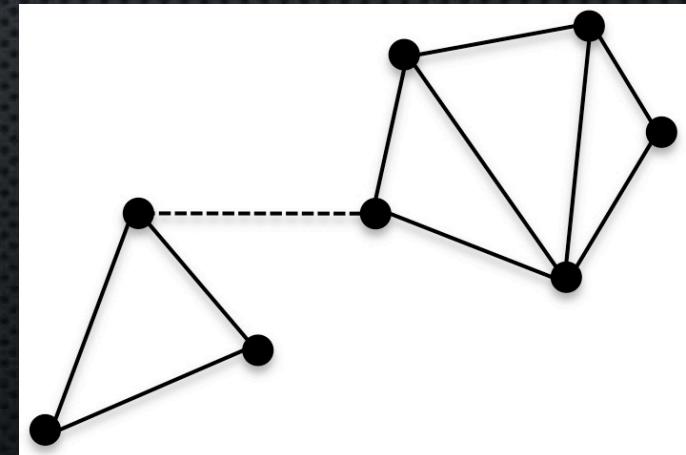
WEIGHTED GRAPH



COMPLETE GRAPH

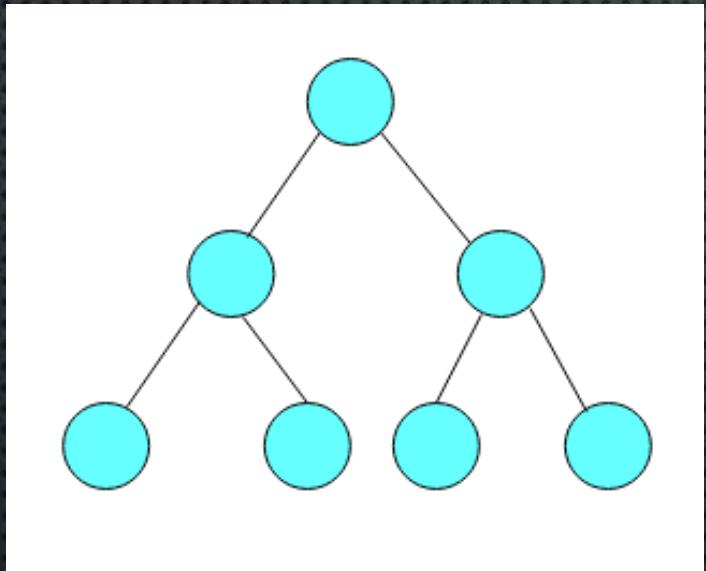


REGULAR GRAPH

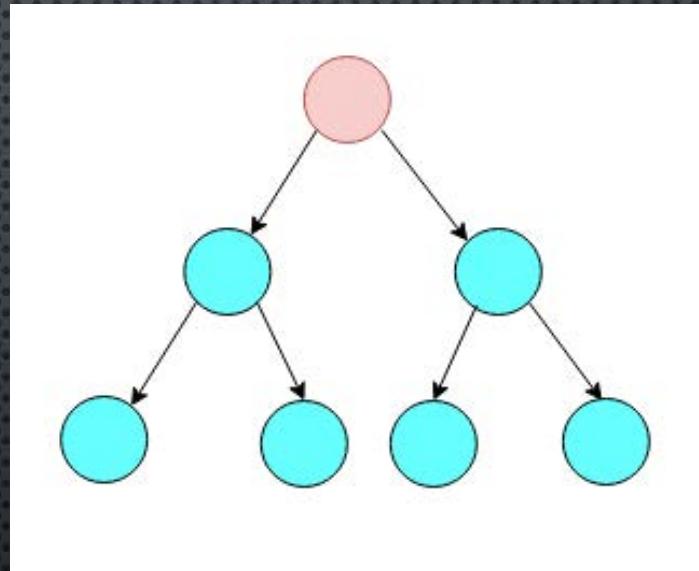


CONNECTED GRAPH

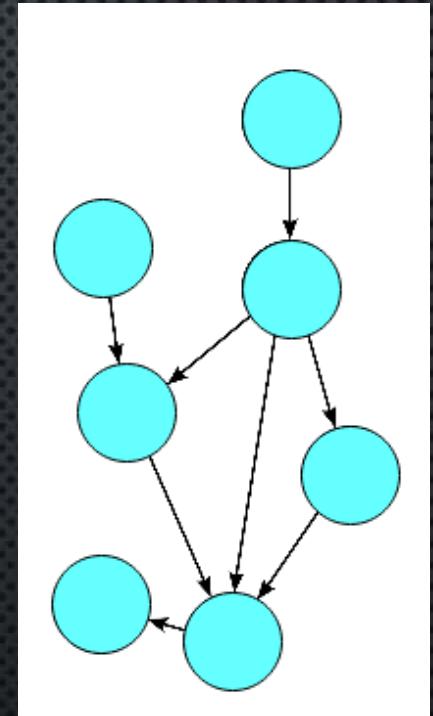
SPECIAL GRAPH



TREE



ROOTED TREE



DAG
(DIRECTED
ACYCLIC GRAPH)

WHY VISUALIZE A GRAPH AT ALL?

1. TO FIND WHAT TO LOOK FOR:

A GOOD VISUALIZATION CAN SHOW IF THERE ARE SOME CLUSTERS OR BRIDGES IN A GRAPH.

2. TO IMPRESS THE PUBLIC:

THAT DATA VISUALIZATIONS ARE USED FOR PRESENTATION.

3. TO GET FEATURES:

A GRAPH THAT REPRESENTED AS AN ADJACENCY MATRIX IS DATA IN HIGH DIMENSIONAL SPACE. WHEN DRAWING IT WE GET TWO COORDINATES FOR EACH VERTEX. THESE COORDINATES CAN ALSO BE USED AS FEATURES.

WHEN CREATING A GRAPH MODEL IS A BAD IDEA

- SOME DATASETS DON'T FEATURE CONNECTIONS, SO THERE ISN'T MUCH VALUE IN APPLYING THE NODE-LINK STRUCTURE TO THEM.
- FOR EXAMPLE IF A TABLE IS RESTRICTED TO UNIQUE DATA ONLY, GRAPH MODELING ISN'T EFFECTIVE.

GRAPH THEORY

WHAT IS GRAPH THEORY?

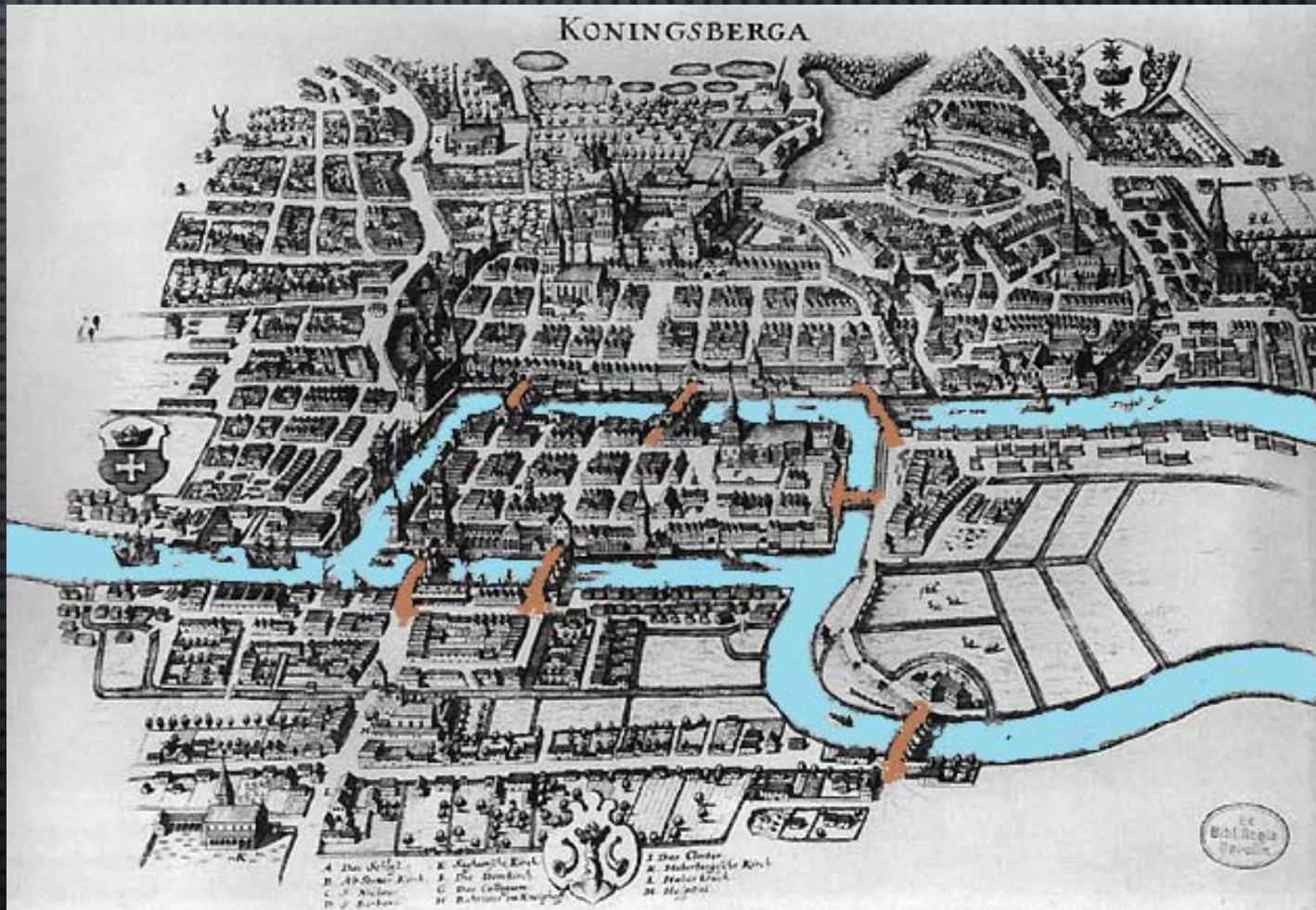
GRAPH THEORY IS THE BRANCH OF
MATHEMATICS AND COMPUTER SCIENCE
INVOLVED WITH THE STUDY OF SETS OF OBJECTS
LINKED BY RELATIONSHIPS, THE GRAPH.

WHERE GRAPH THEORY IS APPLIED

- COMPUTER SCIENCE:
GRAPHS ARE USED TO REPRESENT NETWORKS OF COMMUNICATION OR DATA ORGANIZATION.
- LINGUISTICS:
SYNTAX AND COMPOSITIONAL SEMANTICS FOLLOW TREE-BASED STRUCTURES MODELED IN A HIERARCHICAL GRAPH.
- PHYSICS AND CHEMISTRY:
GRAPH THEORY IS USED TO STUDY MOLECULES IN CHEMISTRY AND PHYSICS.

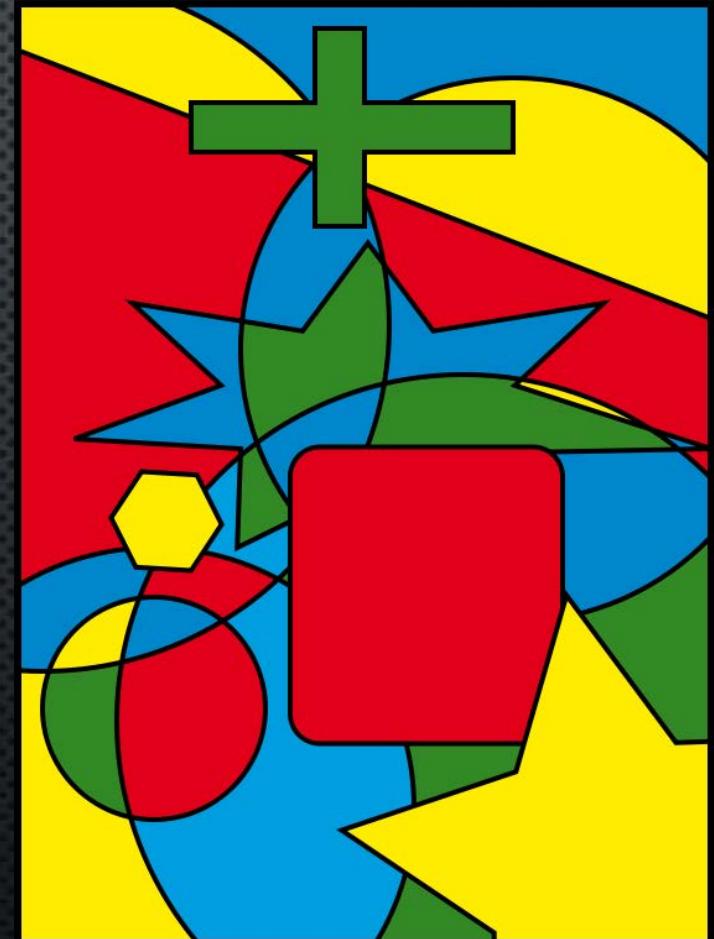
THE SEVEN BRIDGES OF KÖNIGSBERG

- THE PAPER WRITTEN BY LEONHARD EULER ON THE SEVEN BRIDGES OF KÖNIGSBERG AND PUBLISHED IN 1736 IS REGARDED AS THE FIRST PAPER IN THE HISTORY OF GRAPH THEORY.
- THE PROBLEM WAS TO DEVISE A WALK THROUGH THE CITY THAT WOULD CROSS EACH OF THOSE BRIDGES ONCE AND ONLY ONCE.
- EULER PROVED THAT THE PROBLEM HAS NO SOLUTION.



FOUR COLOR THEOREM

- ONE OF THE MOST FAMOUS AND STIMULATING PROBLEMS IN GRAPH THEORY IS «THE FOUR COLOR PROBLEM»
- IN MATHEMATICS, THE FOUR COLOR THEOREM STATES THAT, GIVEN ANY SEPARATION OF A PLANE INTO CONTIGUOUS REGIONS, PRODUCING A FIGURE CALLED A MAP, NO MORE THAN FOUR COLORS ARE REQUIRED TO COLOR THE REGIONS OF THE MAP SO THAT NO TWO ADJACENT REGIONS HAVE THE SAME COLOR.



GRAPH VISUALIZATION

WHAT IS GRAPH VISUALIZATION?

- GRAPH VISUALIZATION, ALSO CALLED ‘NETWORK VISUALIZATION’ OR ‘LINK ANALYSIS’, IS AN AREA OF MATHEMATICS AND COMPUTER SCIENCE, AT THE INTERSECTION OF GEOMETRIC GRAPH THEORY AND INFORMATION VISUALIZATION.
- IT IS THE PROCESS OF VISUALLY PRESENTING NETWORKS OF CONNECTED ENTITIES USING A NODE-LINK MODEL.
- ANY TYPE OF DATA ASSET THAT CONTAINS INFORMATION ABOUT CONNECTIONS CAN BE MODELED AND VISUALIZED AS A GRAPH, EVEN DATA INITIALLY STORED IN A TABULAR WAY.

WHY GRAPH VISUALIZATION?

- GRAPH VISUALIZATION IS OFTEN OVERLOOKED OR VIEWED AS A NON-ESSENTIAL, BUT IT SHOULD BE A MUST-HAVE AND NOT A NICE-TO-HAVE. IT'S THE MOST EFFECTIVE WAY TO UNCOVER, EXPLORE, UNDERSTAND AND COMMUNICATE INSIGHT FROM COMPLEX DATA.
- GRAPH VISUALIZATION IS SCALABLE. REPRESENTING CONNECTED DATA IN TABLES IS NOT INTUITIVE AND OFTEN HIDES THE CONNECTIONS IN WHICH LIES THE VALUE. YOU CAN STILL FIND AND EXPLORE THE CONNECTIONS THAT EXIST EVEN IN YOUR LARGEST DATASETS. THIS WOULD BE A CHALLENGE IF YOUR DATA WAS CONFINED TO TABLES OF ROWS AND COLUMNS.
- WHEN YOU'RE WORKING WITH LARGE DATASETS, FILTERING OUT THE NOISE LETS YOU FOCUS ON THE NODES AND LINKS OF MOST INTEREST.

THE BENEFITS OF GRAPH VISUALIZATION

THERE ARE A NUMBER OF REASONS WHY GRAPH VISUALIZATION IS USEFUL:

1. OUR BRAINS STRUGGLE WITH PROCESSING DATA, BUT ARE GREAT AT SPOTTING PATTERNS IF THE DATA IS PRESENTED IN A TANGIBLE FORMAT. YOU WILL SPEND LESS TIME ASSIMILATING INFORMATION BECAUSE THE HUMAN BRAIN PROCESSES VISUAL INFORMATION MUCH FASTER THAN WRITTEN ONE. VISUALLY DISPLAYING DATA ENSURES A FASTER COMPREHENSION WHICH REDUCES THE TIME TO ACTION.
2. YOU HAVE A HIGHER CHANCE TO DISCOVER INSIGHTS BY INTERACTING WITH DATA. GRAPH VISUALIZATION TOOLS OFFER THE POSSIBILITY TO MANIPULATE THE DATA. IT ENCOURAGES DATA APPROPRIATION, ITS QUESTIONING AND IN THE END INCREASES THE POSSIBILITY TO DISCOVER ACTIONABLE INSIGHTS.

3. YOU CAN ACHIEVE A BETTER UNDERSTANDING OF A PROBLEM BY VISUALIZING PATTERNS AND CONTEXT. GRAPH VISUALIZATION TOOLS ARE PERFECT TO VISUALIZE RELATIONSHIPS BUT ALSO TO UNDERSTAND THE CONTEXT OF THE DATA. YOU GET A COMPLETE OVERVIEW OF HOW EVERYTHING IS CONNECTED WHICH ALLOWS TO IDENTIFY TRENDS AND CORRELATIONS IN YOUR DATA.
4. IT'S AN EFFECTIVE FORM OF COMMUNICATION. VISUAL REPRESENTATIONS OFFER A MORE INTUITIVE WAY TO UNDERSTAND THE DATA AND ARE AN IMPACTFUL MEDIUM TO SHARE YOUR FINDINGS WITH DECISION-MAKERS.
5. EVERYBODY CAN WORK WITH GRAPH VISUALIZATION, NOT ONLY TECHNICAL USERS. MORE USERS CAN ACCESS THE INSIGHTS SINCE SPECIFIC PROGRAMMING SKILLS ARE NOT REQUIRED TO INTERACT WITH GRAPH VISUALIZATIONS. THIS INCREASES THE VALUE CREATION POTENTIAL.

HOW GRAPH VISUALIZATION IS BEING USED

- CYBERSECURITY
ORGANIZATIONS NEED TO PROTECT THEMSELVES FROM VULNERABILITIES LIKE ZERO-DAY VULNERABILITIES AND DDoS OR PHISHING ATTACKS. THEY COLLECT DATA FROM SERVERS, ROUTERS OR APPLICATION LOGS AND NETWORK STATUS IN ORDER TO DETECT SUSPICIOUS ACTIVITY. GRAPH VISUALIZATION IS A GREAT TOOL TO DIGEST THIS DATA AND DETECT SUSPICIOUS PATTERNS IN A GLIMPSE.
- INTELLIGENCE
ALMOST EVERY GOVERNMENT HAS ITS INTELLIGENCE AGENCY. TO SUPPORT LAW ENFORCEMENT, NATIONAL SECURITY OR MILITARY OBJECTIVES, THESE ORGANIZATIONS COLLECT AND ANALYZE DATA FROM VARIOUS SOURCES. VISUALIZING CONNECTIONS BETWEEN PEOPLE, EMAILS, TRANSACTIONS OR PHONE RECORDS IS A KEY TO EASE SUCH INVESTIGATIONS.
- LIFE SCIENCE
PROTEIN INTERACTIONS, DRUG COMPOSITIONS, DISEASE NETWORKS: FOR LIFE SCIENCE DATA ANALYSIS ALMOST EVERYTHING IS ABOUT CONNECTIONS AND DEPENDENCIES. GRAPH VISUALIZATION MAKES LARGE AMOUNTS OF DATA MORE ACCESSIBLE AND EASIER TO READ.

REFERENCES:

- [HTTPS://TOWARDSDATASCIENCE.COM](https://towardsdatascience.com)
- [HTTPS://LINKURIO.US/BLOG/](https://linkurio.us/blog/)
- [HTTPS://CAMBRIDGE-INTELLIGENCE.COM/BLOG/](https://cambridge-intelligence.com/blog/)
- [HTTPS://WIKIPEDIA.COM](https://wikipedia.com)