

MATES ED2MIT

Education and Training for Data Driven Maritime Industry

Introduction to Data Management and Governance

Introduction: MATES Project and course overview

Yuri Demchenko MATES Project
University of Amsterdam

**Maritime Alliance for fostering the
European Blue economy through a
Marine Technology Skilling Strategy**



Co-funded by the
Erasmus+ Programme
of the European Union

The European Commission support for the production of this publication does not constitute an endorsement of the contents which reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.



3 Days/Sessions course – Logistics

- 3 Sessions of 3 hours, including 2 breaks 10 min
- Time 15:00-18:00, online
- Training course information page
 - To be created at Github
- Zoom class
 - <https://uva-live.zoom.us/j/88358719282?pwd=SkFOdi9qUm9WOG93OFhpd1pvY2Y1UT09>
- Course materials uploaded to GoogleDrive folder
 - <https://drive.google.com/drive/folders/1Hjevyvnm8we2lgtieLMtU3WtD2KvPaMH>
- Lectures will be recorded and uploaded after lecture overnight, and updated after processing - approx. 2-3 days
- Discussion and Q&A via Zoom chat





Daily Topics

- **Day 1**
 - **Course introduction: MATES project, Industry 4.0, Industry 5.0 and digitalisation in maritime industry**
 - **Industrial Data Spaces: Definitions, Use cases**
 - **European policy on Data Governance**
 - **Data Protection, GDPR**
- **Day 2**
 - Data Management and Governance, Best Practices
 - DAMA Body of Knowledge (DAMA-BOK) and DAMA Data Architecture
 - Data Governance and Data Management
 - Data Management Maturity frameworks
- **Day 3**
 - RDM in Europe: Initiatives and projects
 - Open Science, Open Data, Open Access, European Open Data Pilot
 - FAIR Data Principles
 - Responsible Data Use (citation, copyright, data restrictions)
 - Data Management Plan
 - Data Stewardship, Competences and demand



Expected Outcome

- End of module review and feedback
- Certificate of attendance upon completed assignment and feedback
- All videos will be available overnight and after the course
- All materials will be available at the github workshop page and linked from the MATES project website
- Your contribution to **Blue Survey** by Evelyn Parades, Uni Gent, available in 11 languages
 - <http://thebluesurvey.eu/>
 - To investigate the link between maritime professionals and the ocean, with a focus on sustainability



Outline Day 1, Part 1

- Industry 4.0 and Maritime Industry Digitalisation
- Demand for Digital and Data competences and Skills
 - Competences and skills for data driven organisations
 - Data literacy and Transversal skills
- MATES Project contribution
 - ED2MIT – Training on Digital and Data Skills and Literacy
- Discussion



This work is licensed under the Creative Commons Attribution 4.0 International License.
To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

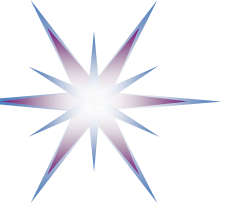


Co-funded by the
Erasmus+ Programme
of the European Union

The European Commission support for the production of this publication does not constitute an endorsement of the contents which reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.



- Graduated and PhD from National Technical University of Ukraine “Kiev Polytechnic Institute”
 - University of Amsterdam – since 2003
- Research areas
 - Big Data Infrastructure and Data Science platforms
 - Cloud architecture framework, DevOps and cloud automation platform
 - Cloud security and compliance
- Teaching courses (on campus and online)
 - **Big Data Infrastructure and Technologies for Data Analytics (BDIT4DA)**
 - Cloud Computing Engineering (CCENG), Security Engineering (SECENG))
 - DevOps and Cloud based Software Development (DevOps)
 - **Web technologies and Databases (WebDB)**
 - **Data Science and Analytics Foundations (DSAF), Professional Issues in Data Science**
- Recent projects
 - **EDISON: Building the Data Science Profession for Europe**
 - **MATES: Digitalisation of the European Blue Economy**
 - **FAIRsFAIR/EOSC: FAIR Data Management and Data Stewardship**
 - GEANT4 Research: Cloud aware networking infrastructure provisioning on-demand
 - SLICES-DS: Research Infrastructure for ICT



Digital and Data Training: General goals

- Meet maritime experts and community and obtain feedback how to deliver digital and data technologies to wide community of maritime experts and community – to achieve digital and data literacy
 - Goal: Obtain feedback from professional maritime community via interactive teaching
 - Potentially find cooperators to transfer knowledge: train the trainers
- Provide basic information on Big Data and Data Analytics technology and tools
- Introduce into Data Management and Governance practices
- Build trust in data-driven and digitalized operations by providing assurance of data quality, algorithms, sensors, systems and cyber security
- Facilitate digital readiness and transformation, understand benefits
- **MATES ED2MIT Course on Data Governance and Data Management**
 - This is the second in series of ED2MIT training webinars
- Provide general information on Data Governance and Management in the context of the European Data Governance policy
- Introduce Data Management best practice and Data Management Plan template



Industry 4.0 and demand for new skills

WORLD
ECONOMIC

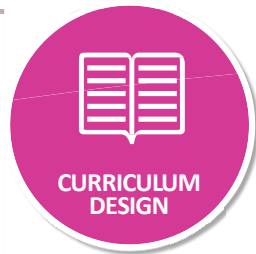


The Fourth Industrial Revolution, which includes developments in previously disjointed fields such as

- artificial intelligence and machine-learning, robotics, nanotechnology, 3-D printing, and genetics and biotechnology,
- will cause widespread disruption not only to business models but also to labour markets over the next five years,
- with enormous change predicted in the skill sets needed to thrive in the new landscape.
- Ref report *The Future of Jobs*, published by the World Economic Forum.



OCCUPATIONAL
PROFILES



CURRICULUM
DESIGN



EDUCATION
& TRAINING



KNOWLEDGE
TRANSFER



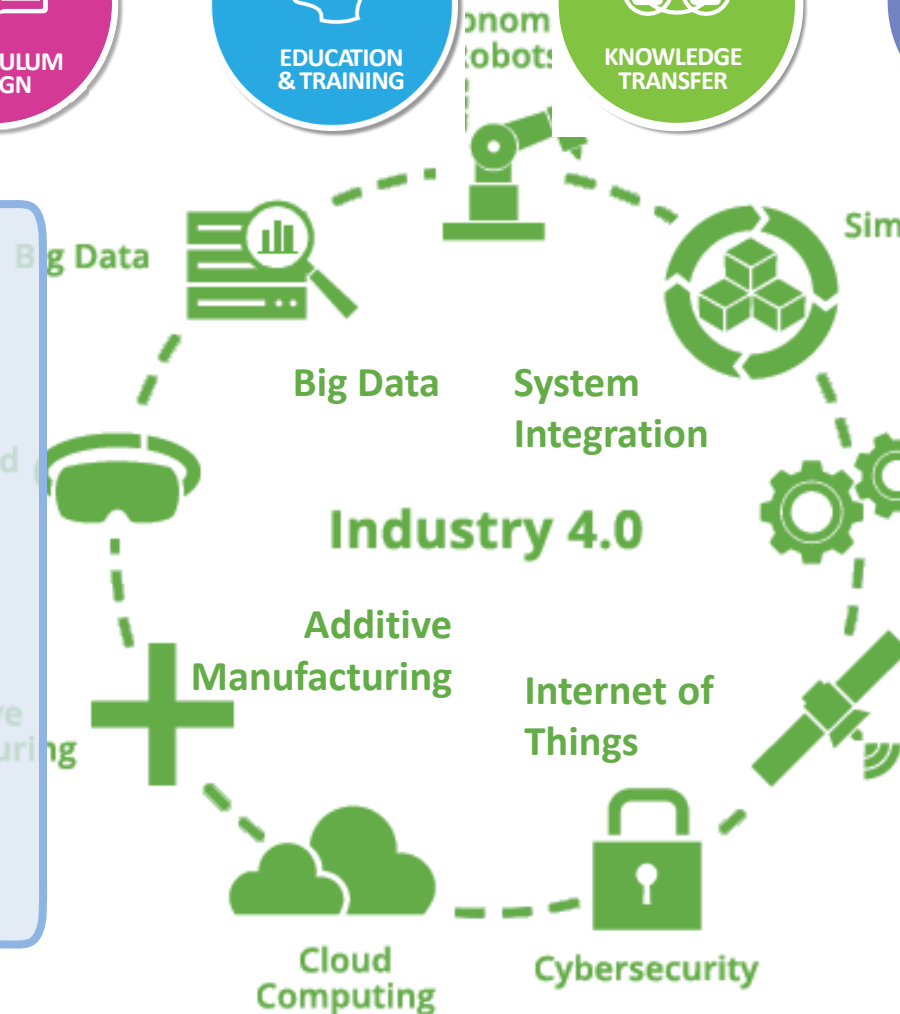
LONG-TERM
SKILLS STRATEGY



OCEAN
LITERACY

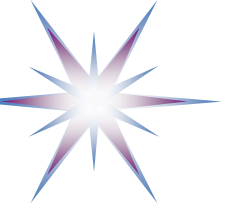
Digital Transformation

- Digitalisation and IoT
- Intelligent Information
- Data Management
- Digital Assets Manage
- Data Driven Optimisation
- Agile Continuous Improvement
- Customer Experience
- People and skills



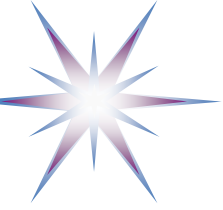
Digital Competences/Skills

- Automation, robotics, electrical vehicles
- Information and data literacy
- Communication and collaboration
- Digital content creation, safety
- Problem solving and critical thinking



Industry 5.0 – Human centric Industry

- The term Industry 5.0 refers to people working alongside robots and smart machines.
 - Humans and robots co-working
 - It's about robots helping humans work better and faster by leveraging advanced technologies like the Internet of Things (IoT) and big data.
 - It adds a personal human touch to the Industry 4.0 pillars of automation and efficiency.
- #1 Industry 5.0 is aimed at supporting – not superseding – humans.
- #2 Industry 5.0 is about finding the optimal balance of efficiency and productivity.
- #3 The progress of Industry 5.0 is unavoidable.



Industry 5.0: Towards more sustainable, resilient and human-centric industry

- Industry 5.0: Towards more sustainable, resilient and human-centric industry
- https://ec.europa.eu/info/news/industry-50-towards-more-sustainable-resilient-and-human-centric-industry-2021-jan-07_en
- Elements pertinent to Industry 5.0 are already part of major Commission policy initiatives
 - adopting a human-centric approach for digital technologies including artificial intelligence ([AI White Paper](#))
 - **up-skilling and re-skilling European workers, particularly digital skills** ([Skills Agenda](#) and [Digital Education Action plan](#))
 - modern, resource-efficient and sustainable industries and transition to a circular economy ([Green Deal](#))
 - a globally competitive and world-leading industry, speeding up investment in research and innovation ([Industrial Strategy](#))



Industry 5.0 website

- Industry 5.0 website https://ec.europa.eu/info/research-and-innovation/research-area/industrial-research-and-innovation/industry-50_en
- Industry 5.0 is characterised by going beyond producing goods and services for profit. It shifts the focus from the shareholder value to stakeholder value and reinforces the role and the contribution of industry to society.
- **Industry 5.0** – Report January 2021
 - <https://op.europa.eu/en/publication-detail/-/publication/468a892a-5097-11eb-b59f-01aa75ed71a1/>
 - Towards a sustainable, human-centric and resilient European industry
- **Enabling Technologies for Industry 5.0** - Sept 2020
 - https://ec.europa.eu/info/publications/enabling-technologies-industry-50_en
 - Results of a workshop with Europe's technology leaders





Stages of digitalization in maritime transport

The effects of digitalization on maritime transport can be divided into the following three stages:

1. Optimization – maximizing efficiency and reliability in existing processes to reduce the costs of trading.
2. Extension – moving beyond efficiency to produce opportunities for new services and businesses.
3. Transformation – reinventing logistics, trade and business models, based on data-driven revenue streams and shifts in trade flows.

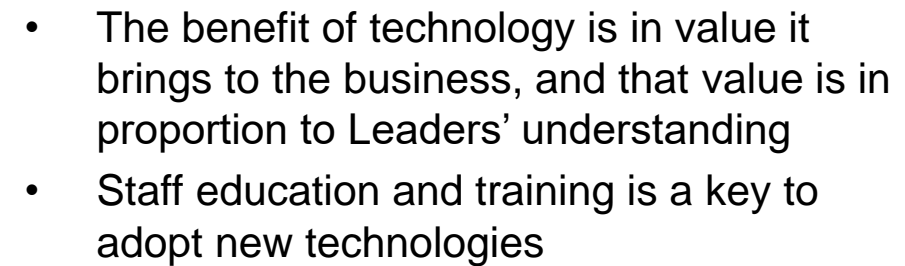


The infographic consists of two rows. The top row features a circular icon with a cloud and a smartphone, followed by a text box stating 'Technical and nautical operation improvements using digitalisation (data driven technology for safety, security and environmental concerns)', a circle with '25%', and a blue horizontal bar. The bottom row features a circular icon of a person in a uniform, followed by a text box stating 'Crew and shoreside staff training', a circle with '19%', and a blue horizontal bar.

Area of Improvement	Percentage
Technical and nautical operation improvements using digitalisation (data driven technology for safety, security and environmental concerns)	25%
Crew and shoreside staff training	19%

The infographic displays the top 10 digital technologies in maritime transport, ranked by their usage percentage. Each technology is represented by an icon, a percentage value, and a corresponding blue bar.

Technology	Percentage
Fleet management data analysis	20%
Technical operations	18%
Nautical operations (for example route optimisation)	15%
Cyber security	14%
Predictive maintenance	11%
Artificial intelligence	7%
Electronic log book	6%
Digital twins	5%
Blockchain	3%
Other	1%



MA



Opportunities with Digitalisation

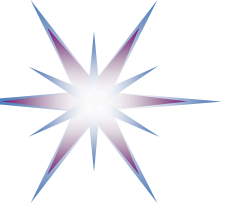
- Cross-sector sharing
- Catch up with the IoT data tsunami
 - Sensors network
 - Robotics
 - Vehicles connectivity
- Starting data pipeline from IoT and edge computing
 - Early problems identification



Data driven culture in enterprise

Building a Data-Driven Culture in Enterprise: Introducing the 4 Pillars

- **Data maturity.** Solid data maturity is foundational to a data culture.
 - Organization's data maturity manifests itself in every individual in your organization having an easy and appropriate level of access to the clean and accurate data they need.
 - Importance of a well-defined CDO role and other related roles
- **Data-driven leadership.** Leaders define the culture of their organization.
 - A data-driven leader supports a culture of data by demonstrating data-driven decision making and involve the team members
 - A data-driven leader sees data as a strategic asset and makes "think and act data" a key strategic priority.
- **Data literacy.** Individual decision makers must be data literate to leverage their data
 - The CDO office needs to invest in enterprise wide data literacy, where every role is upgraded with the right level of data science skills.
- **Data-driven decision-making processes.** Establish a structured process of forward-looking decision making and backward-looking reviews of decisions.
 - Build experience of aligning data analytics, insight and data-driven decision-making processes.



Five Characteristics of a Data-Driven Company

#1. Creative executives who run their businesses with passion and curiosity

- Being data-driven requires a bit of a researcher's mindset -- the curiosity to dig into the data and glean insights from it that can be of use for the business.

#2. Data democratization

- Data-driven organizations emphasize the importance of broad data access for all employees.

#3. Data literacy

- An organization's ability to succeed in the digital era is heavily dependent on its employees' data literacy: the ability to read, work, analyze, and argue with data.
- Example of how to respond to a data literacy problem is Data University at Airbnb. Airbnb could not have a data scientist in every room to inform every decision with data.

#4. Automation of data management workloads

- A core criterion for a data-driven organization is how much data analytics tools are automated and provide information in a form that can be easy for decision making

#5. A companywide, data-driven culture

- Becoming data-driven involves more than technology and tools. It also requires a shift in the enterprise's mindset and culture.
- IBM example: Require all analytics blending with open data (social media, weather, climate etc.)



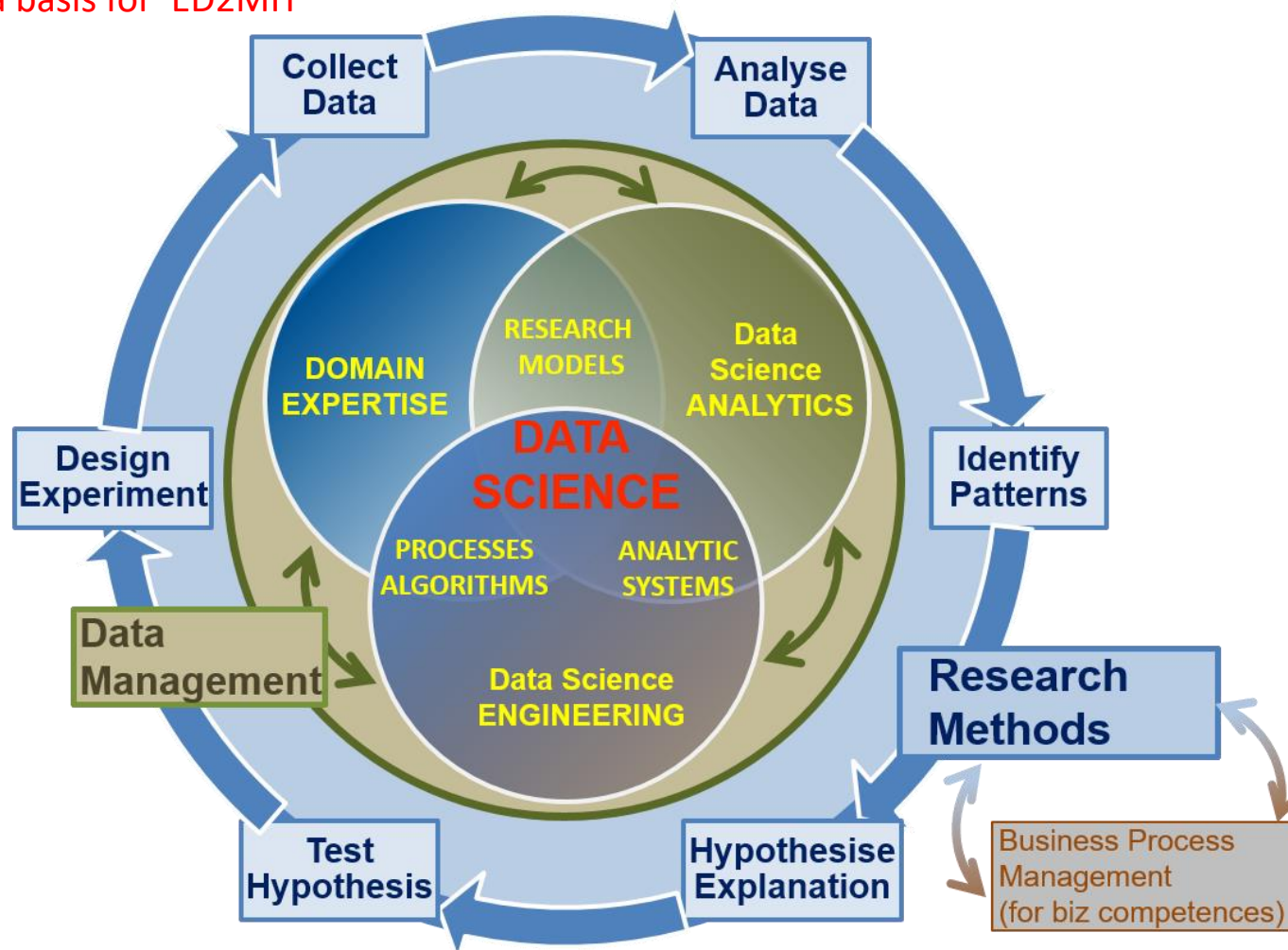
MATES Project contribution

- Study and report on demanded skills in maritime industry
- Pilot experiences and Roadmap
 - Pilot Experience ED2MIT - Education and Training for Data Driven Maritime Industry
- Leverage EDISON Data Science Framework (EDSF) that created a basis for Big Data, Data Science and digital and data skills education and training
 - EDSF provides a basis for ED2MIT training program
- Leverage European Digital Competence framework DigCom2.1
- Provide training on digital and data technologies
 - Train trainers, create pool of reference training materials



EDISON Project (2015-2017): Data Science Competence Groups - Research

EDISON Data Science Framework (EDSF) provides a basis for ED2MIT



Data Science Competences include 5 groups

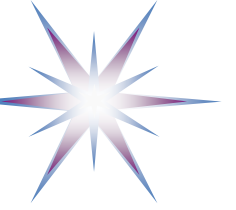
- Data Science Analytics
- Data Science Engineering
- Domain Knowledge and Expertise
- Data Management
- Research Methods and Project Management
 - Business Process Management (biz)

Scientific Methods

- Design Experiment
- Collect Data
- Analyse Data
- Identify Patterns
- Hypothesis Explanation
- Test Hypothesis

Business Operations

- Operations Strategy
- Plan
- Design & Deploy
- Monitor & Control
- Improve & Re-design



Discussion

- Digitalisation aspects in your organisations
 - Go to **www.menti.com** and use the provided code
- Experience and aspects of digitalisation in your organosations



Contribute to Blue Survey

- **Blue Survey** by Evelyn Parades, Uni Gent, available in 11 languages - <http://thebluesurvey.eu/>
 - To investigate the link between maritime professionals and the ocean, with a focus on sustainability



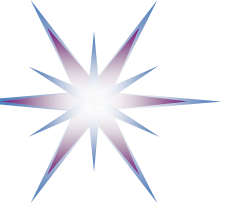
Acknowledgement

- This work is supported by the ERASMUS+ MATES project
- The work is committed to the Open Source under Creative Commons 4.0 CC BY License



This work is licensed under the Creative Commons Attribution 4.0 International License.
To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>





Additional materials

- ED2MIT Training on digital and data skills



ED2MIT: Digital and Data Competence Groups

A. Data – B. Cloud – C. Digital Content – D. Data Science & Analytics

A. Data related competences and technologies

B. Cloud services and cloud economics

C. Digital content creation, access and management

D. Data Science and Big Data Analytics



A. Data related competences and technologies

A. Data – B. Cloud – C. Digital Content – D. Data Science & Analytics

A. Data related competences and technologies

- A.1. Big Data definition and technologies: 6V of Big Data and challenges for companies. Big Data examples from research and industry
- A.2. Data collection, access and sharing
- A.3. Data formats, data models, metadata
- A.4. Data Storage and databases, SQL scripting and simple commands
- A.5. Data inspection, Data protection, data backup and archiving
- A.6. Cloud based services and tools for data storage, sharing and management
- A.7. Open Data repositories, test datasets, developer communities
- A.8. Organisational and private Data Management, FAIR Data Principles, organisational roles, Data Stewards



B. Cloud services and cloud economics

A. Data – **B. Cloud** – C. Digital Content – D. Data Science & Analytics

B. Cloud services and cloud economics

- B.1. Cloud service models: IaaS, PaaS, SaaS, Apps. Use examples and Cloud Service Providers. Cost model of cloud services.
- B.2. Company IT infrastructure migrating to cloud: benefits and problems
- B.3. Cloud and Big Data, Cloud based Big Data platform and services
- B.4. Data storing, backing up, sharing and processing in clouds (for organisational and private data)
- B.5. Practical exercises with cloud services: Cloud management console and cloud services deployment and access.



C. Digital content creation, access and management

A. Data – B. Cloud – **C. Digital Content** – D. Data Science & Analytics

C. Digital content creation, access and management

- To be acquired as self-study or expected to be known

C.1. Internet and World Wide Web

C.2. HTML, CSS, JavaScript for active pages

C.3. UX design and web portal services

C.4. Web applications security

C.5. PHP and interactive web sites (advanced)

To be provided as self-study materials



D. Data Science and Big Data Analytics

A. Data – B. Cloud – C. Digital Content – D. Data Science & Analytics

D. Data Science and Big Data Analytics

This course is provided as a general overview of the listed below topics. More in-depth training and learning will require more time commitment and pre-requisite knowledge.

D.1. Statistical methods and Probability theory

D.2. Data description and Statistical Data Analysis

D.3. Data preparation: data loading, data cleaning, data pre-processing, parsing, transforming, merging, and storing data

D.4. Qualitative and Quantitative data analysis

D.5. Classification: methods and algorithms

D.6. Cluster analysis basics and algorithms

D.7. Performance of data analytics algorithms and tools

D.8. Visualizations of data analysis and dashboards

D.9. Organizing data analytics process following CRISP-DM and Data Science Process