EDISON Data Science Framework (EDSF): Facilitating Data Science Curricula Development and organisational capacity building

Working with EDSF: Curricula Analysis

EDISON
building the data science profession

Yuri Demchenko, EDISON Project
University of Amsterdam

HPCS2019

16 July 2019, Dublin

EDISON – **E**ducation for **D**ata **I**ntensive **S**cience to **O**pen **N**ew science frontiers

H2020 INFRASUPP Grant 675419 (2015-2017)

Workshop Agenda

Part 1 13:30 – 15:00

1. Introduction and bootstrapping (10 min)
   - EDISON Data Science Framework and EDISON Community Initiative
   - Workshop goals and attendees interests
2. Background information and how the EDSF has been done (20 min)
   - EU and International studies on data related competences and skills
   - EU standards, projects, initiatives, associations
3. EDISON Data Science Framework (EDSF) in details and walk through customised curriculum design (60 min)
   - EDSF components: Competence Framework (CF-DS), Body of Knowledge (DS-BoK), Model Curriculum (MC-DS), Data Science Professional Profiles (DSPP)
   - Walk through customised curriculum design: from target professional groups or competences to curriculum suggestions

Part 2 15:00-15:30

1. Data Scientist and Data Steward vacancy and CV analysis and matching from indeed.com – Interactive session and practice (in groups)
   - Analysis of Data Science and Data Steward CVs and Job vacancies
   - Matching exercise
2. Assessment and discussion on Data Science soft skills: workplace and 21st Century
3. Closing: comments, questions

EDSF for Data Science programmes Assessment

- Ideal/Balanced Data Science programme need to include all KAG's
  - DSDA – approx. 40-50%
  - DSENG – approx. 30-40%
  - DSDM – approx. 30-20%
  - DSRMPM – 10-20%
  - Domain Knowledge is acquired as an initial profession education or as practical work
    - May be included in the curriculum as project or staging

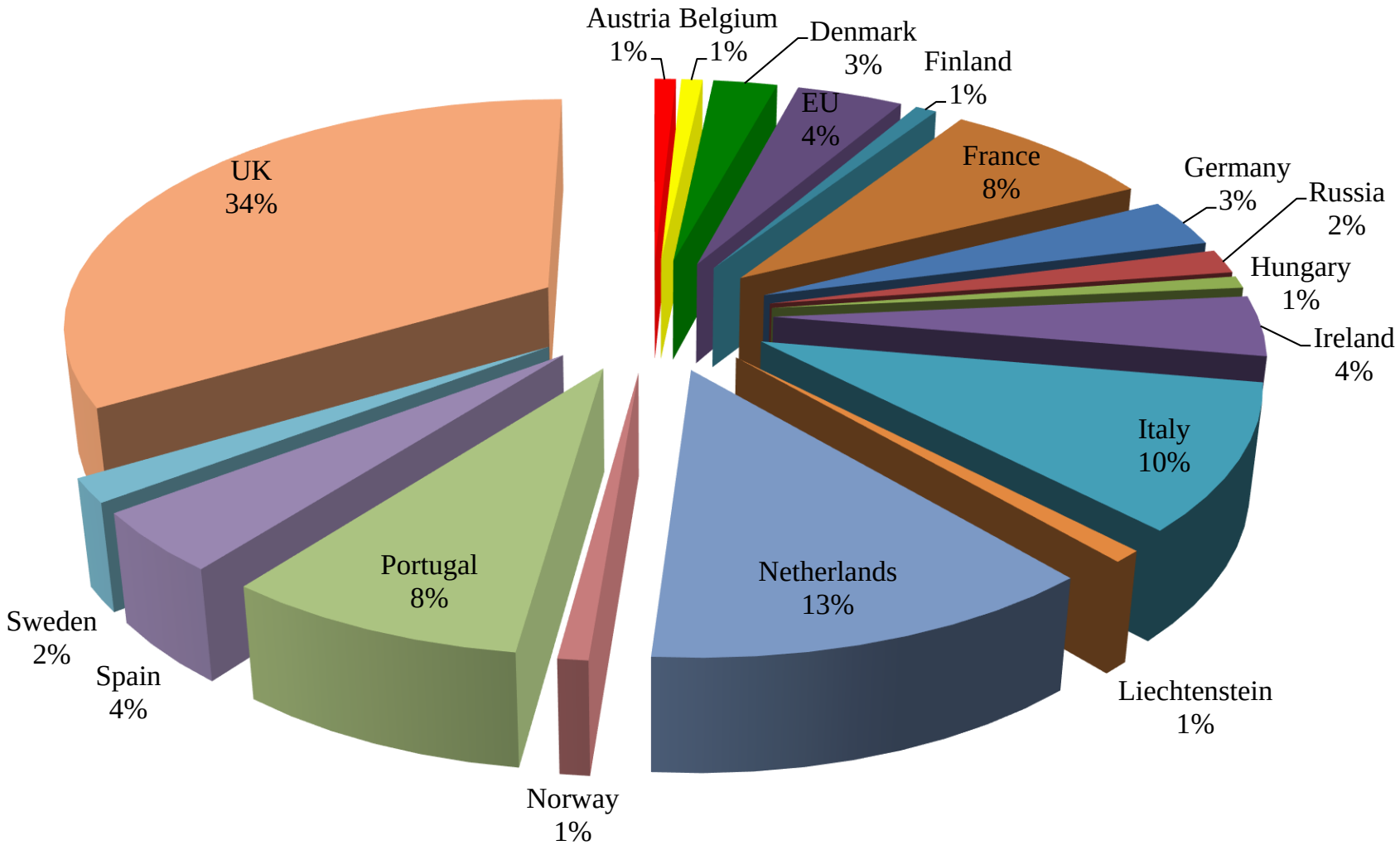
Analysis Data Science programmes 2016

- More than 200 programmes reviewed
- Inventory is available at http://edison-project.eu/university-programs-list
- Content of the programmes analysed on included courses by 4 KAGs

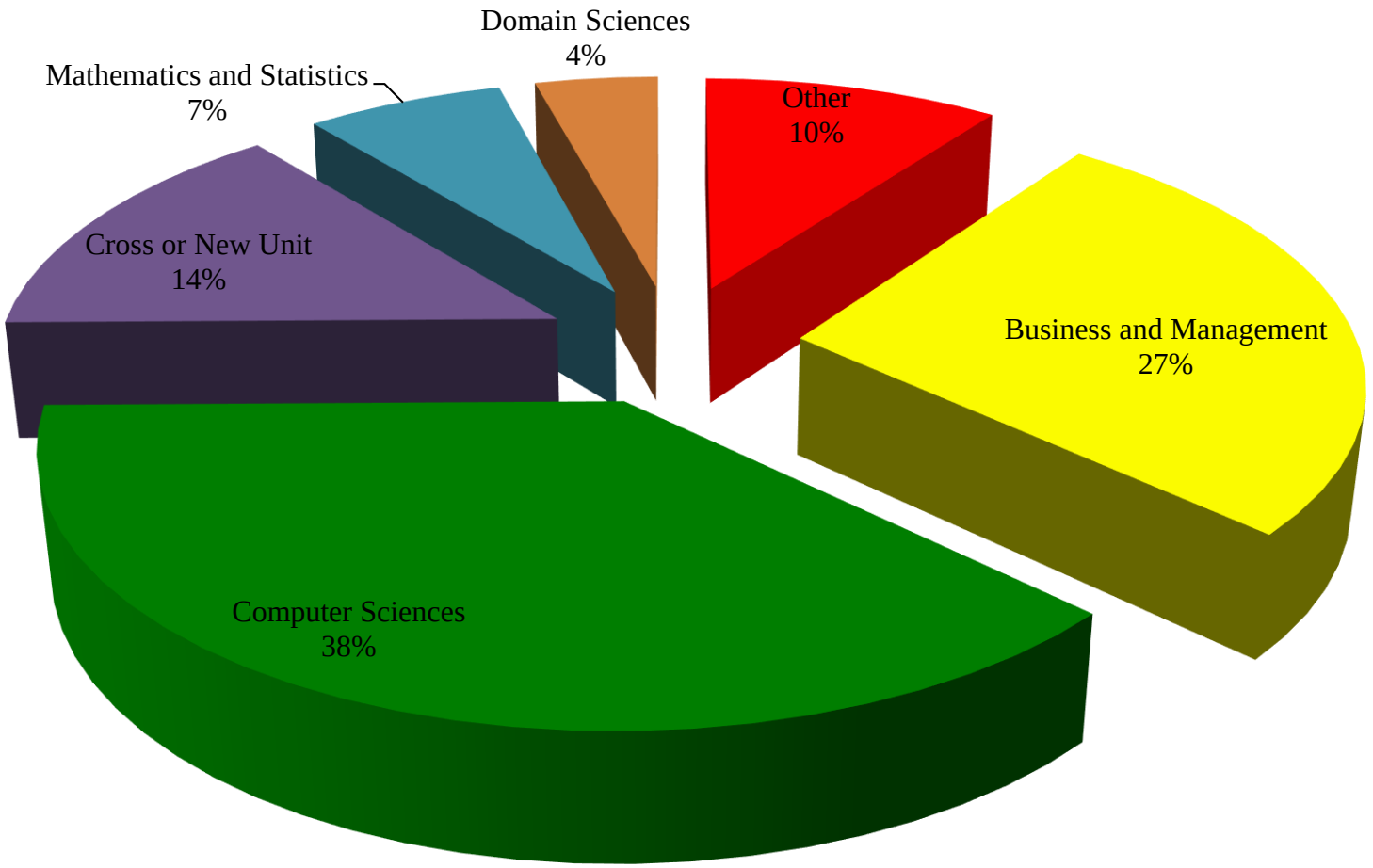
Approach to analysis

- We should expect equal coverage of each competence group
- We established that difference between most and least covered group should not exceed 20% for program to be considered balanced
- Between 20% and 30% we classified program as having a small imbalance
- Over 30% we classified program as imbalanced
  - 30% usually means one of the competence groups is almost not covered at all
- Results are presented as a 2 digit percentage for convenience
  - Quantitative differences of just a few pp should not be over-interpreted
  - Focus is on qualitative differences

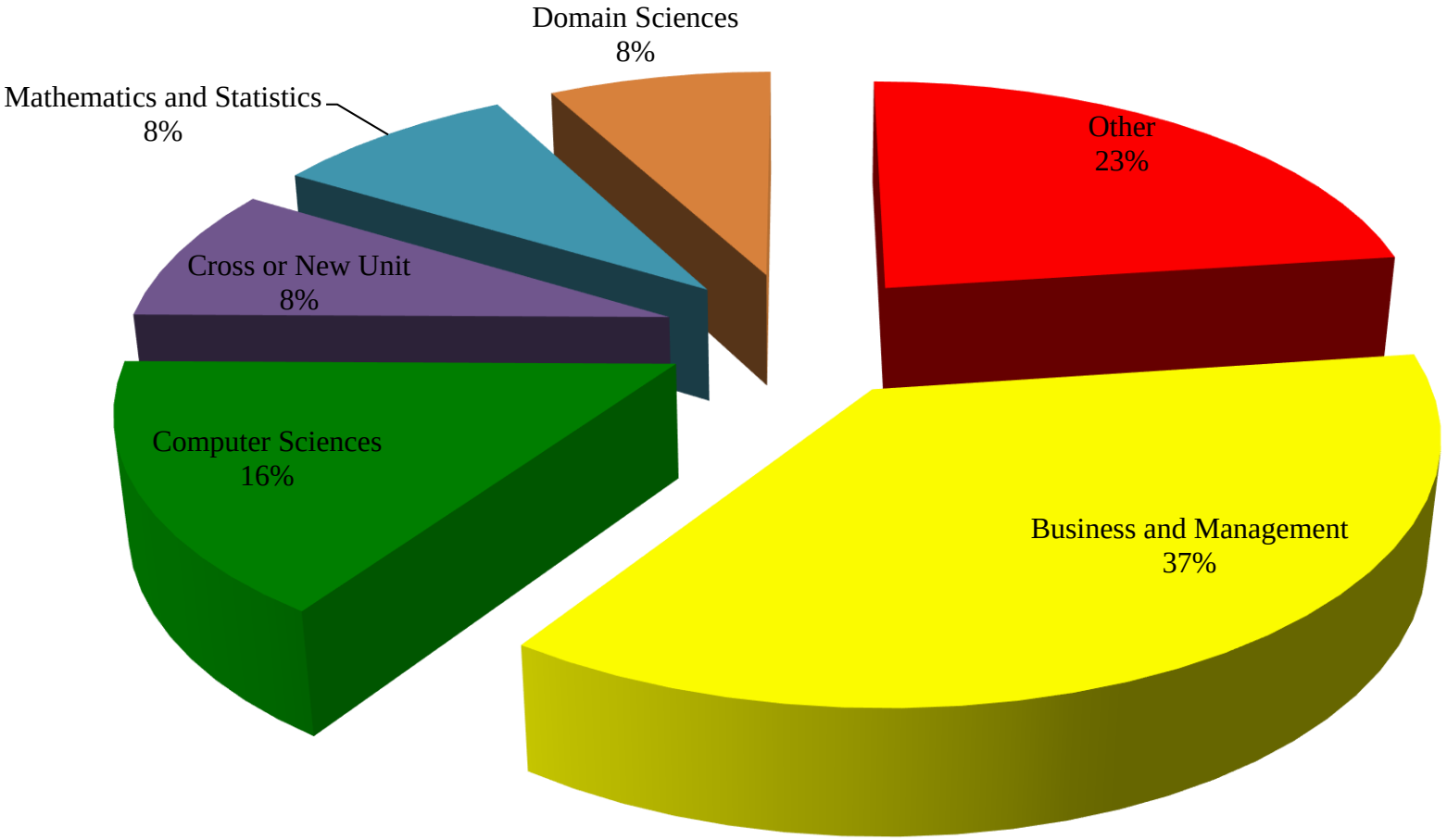
Origin of European Programs 2016



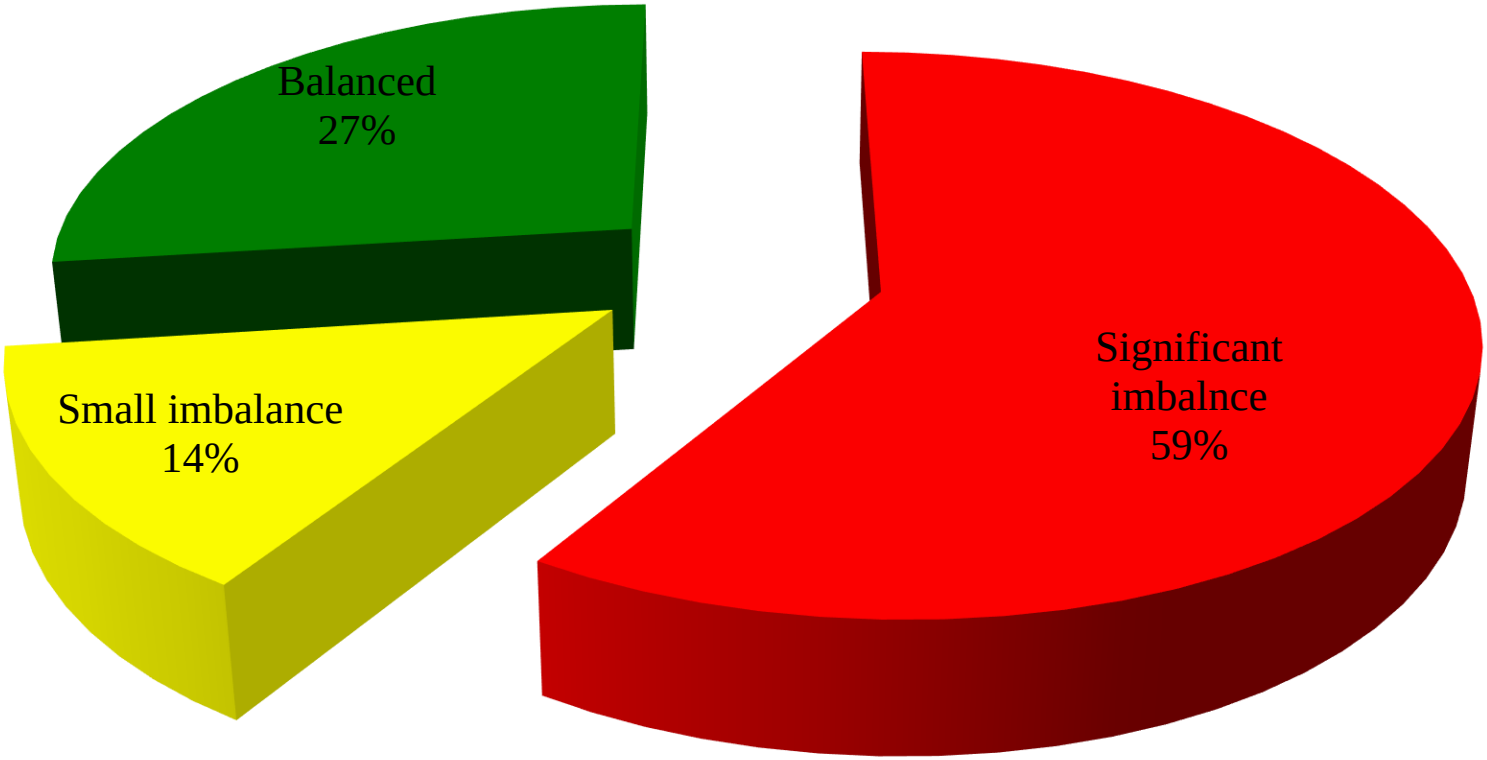
Source of European Programs



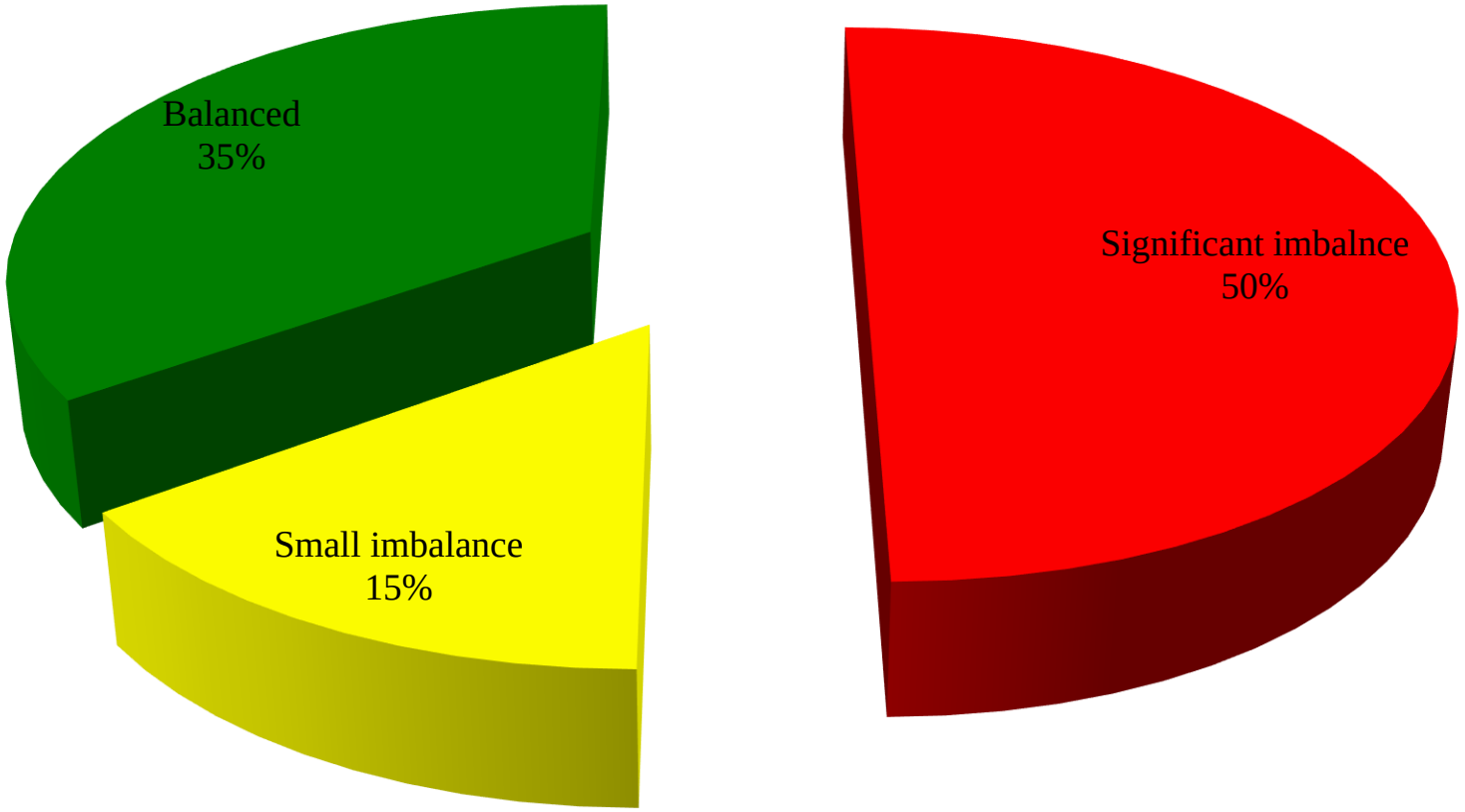
Source of Non-European Programs



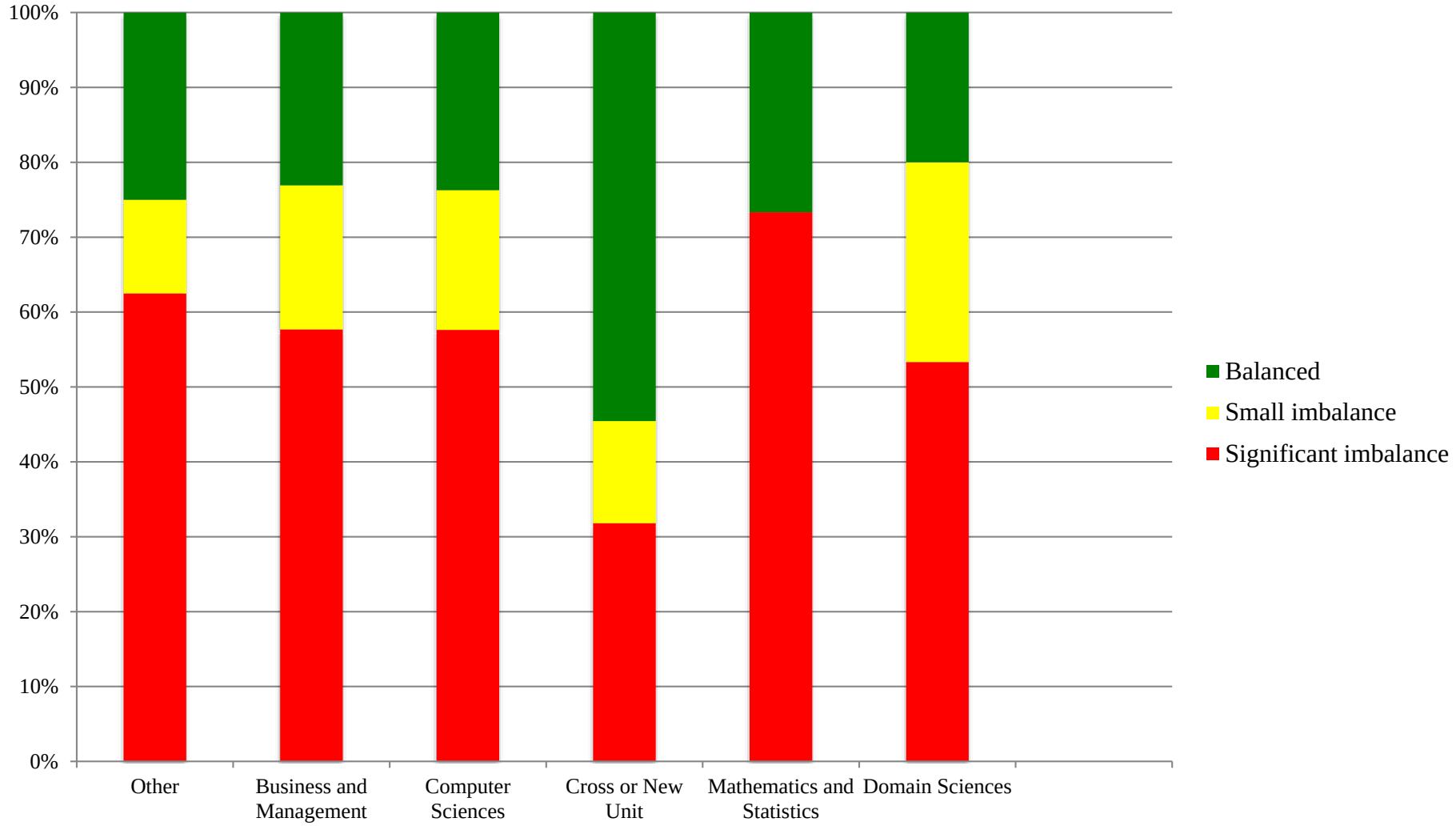
Balance of European Programs



Balance of Non-European Programs



Balance of Programs w.r.t. Source Department



Qualitative analysis of coverage of Data Science Competences

- Only 8% of European programs have learning outcomes formally defined
- Outside Europe it is 16%, mostly due to US influence.
- All programs should have formal learning outcomes
- We also evaluated the quality of learning outcomes w.r.t. Blooms taxonomy
  - very few programs explicitly distribute learning outcomes across various learning levels
  - usually, learning outcomes seemed very generic and offer little useful information

Practice: EDISON Champions Universities

- University of Amsterdam (UvA)
- Univ of Stavanger (UiS)
- Goethe University Frankfurt (GOE)
- University of Bedfortshire (BEDS)
- Lodz University of Technology (TUL)
- University of Perugia (UoP)
- Lucerne School of Information Technology (HSLU)

Links to EDISON Resources

- EDISON Data Science Framework Release 3 (EDSF)
  https://github.com/EDISONcommunity/EDSF

Component EDSF documents

CF-DS – Data Science Competence Framework
https://github.com/EDISONcommunity/EDSF/blob/master/EDISON_CF-DS-release3-v09.pdf

DS-BoK – Data Science Body of Knowledge
https://github.com/EDISONcommunity/EDSF/blob/master/EDISON_DS-BoK-release3-v04.pdf

MC-DS – Data Science Model Curriculum
https://github.com/EDISONcommunity/EDSF/blob/master/EDISON_MC-DS-release3-v04.pdf

DSPP – Data Science Professional profiles
https://github.com/EDISONcommunity/EDSF/blob/master/EDISON_DSPP-release3-v05.pdf

EDISON Initiative Online Presence

- EDSF github project - https://github.com/EDISONcommunity/EDSF
  - Component documents CF-DS, DS-BoK, MC-DS, DSPP

- EDISON Community work area and discussions - https://github.com/EDISONcommunity/EDSF/wiki/EDSFhome
- Mailing list - edison-net@list.uva.nl

- EDISON project website - old domain ***edison-project.eu*** expired: Legacy information to be moved to http://edison-project.net/