

A first view on a data driven approach to tackle frauds: Italy as a case study

Maria Felice Arezzo

mariafelice.arezzo@uniroma1.it

DIPARTIMENTO DI METODI E MODELLI
PER L'ECONOMIA IL TERRITORIO E LA FINANZA
MEMOTEF

CENTRO DI RICERCA
IMPRESAPIENS



SAPIENZA
UNIVERSITÀ DI ROMA

Athens, 5 July 2024

1st Study visit of the project “101101784 — 2022-IT-FRED2” Fraud Repression through Education2

Project summary

The project aims to improve the cooperation between the anti-irregularities/anti-fraud units and the academic environment, with reference in particular to the Italian AFCOS with the Sapienza University of Rome, on the subject of European Funds including recent procedural developments (penal and administrative), by strengthening the exchange of information, experiences and good practices.

The project foresees the following objectives:

- 1) An **experiential co-lab-learning** (WP2) itinerary **participated by academic and practitioners** for **developing innovative cost-effective approaches** that will fit out participants with a greater know-how effectiveness **for determining fraud risks** and comparing control strategies in the sector of public budget support.
- 2) A mixed European academic-practitioner group, which will form a specific **task force**, to produce a **pilot study useful to create the map of concepts for detecting and predicting behavior and risk profiles** that characterize the deviated qualitative and financial use of European funds and produce. (WP3)
- 3) An **Anti-Fraud Observatory** (WP4) with a European transnational dimension perspective, which will **give sustainability to the new network** through relations aimed at **improving the awareness of fraud and other illegal activities** by enhancing transnational and multi-disciplinary cooperation.
- 4) Promotional and **dissemination** material planning and realization for supporting the previously mentioned goals.

The map of concepts to «detect and predict behavior and risk profiles»

- The reasoning is grounded on the fact that firm which put in place frauds have specific characteristics. Some of these can be caught from the balance sheet and financial statement;
- A key point is to understand these characteristics and express them into measurable indicators;
- Statistical and machine learning algorithms help in comparing a sample of firms and detect the anomalies
- Other statistical methods help in refining the search

The map of concepts to «detect and predict behavior and risk profiles»

Step 1

- Key indicators of frauds (KIF)
- Validation of KIF

Step 2

- Data gathering

Step 3

- Unsupervised anomaly detection algorithm
- Benford law

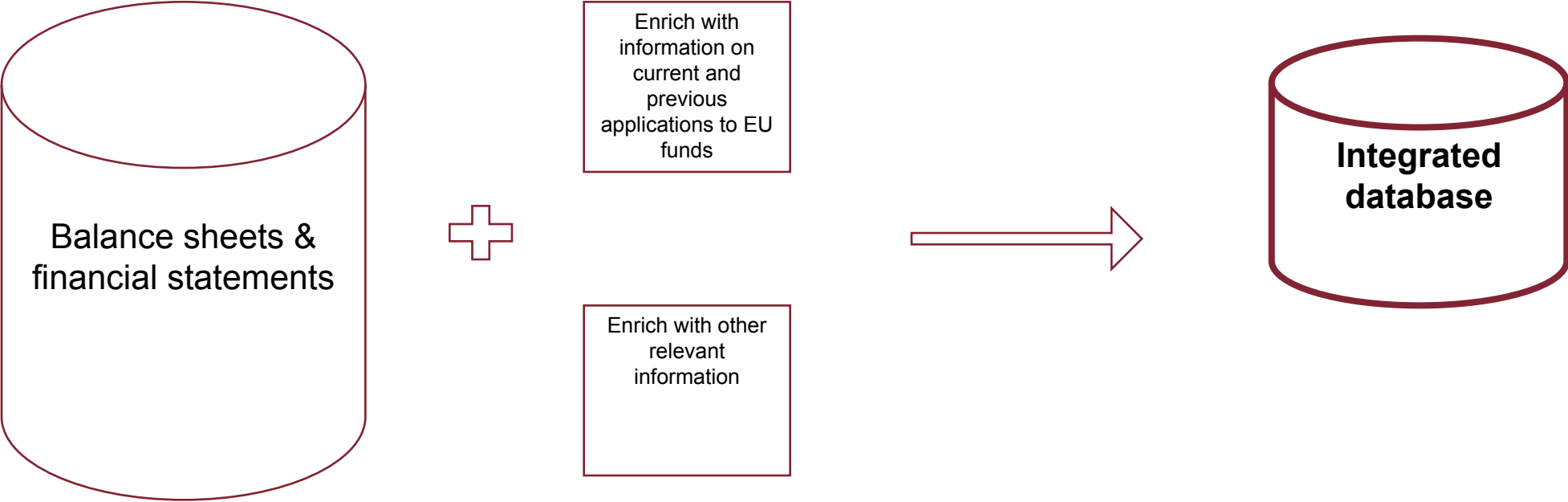
Step 4

- Validation: verify if the anomalous firms are fraudulent

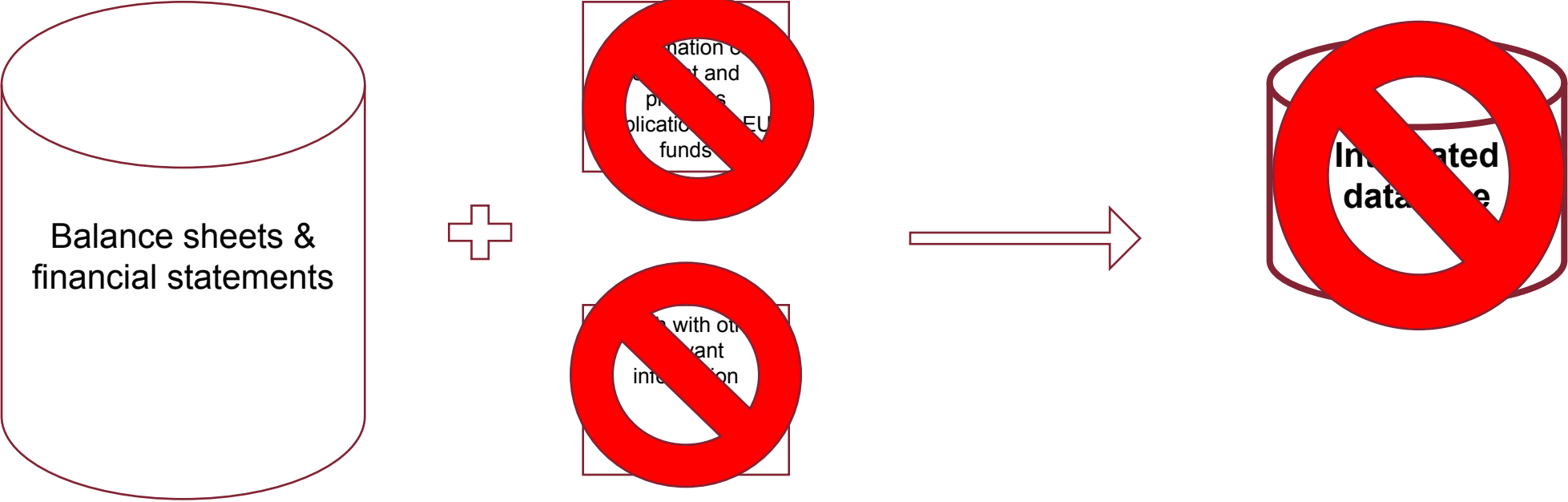
Step1: Key indicators of risk of fraud (KIF)

- The work group of Sapienza defined a set of indicators that can be built from firms' balance sheet and financial statement. In so doing we used all the information and knowledge acquired during the FRED2 workshops
- We are acquiring the opinions of a sample of accountants on such indicators to gather feedback and suggestions.
- A focus group (mixed accountants and professionals of the field) could help improving the selection of KIF gathered from firms data

Step 2: data gathering



Step 2: data gathering so far



Step 3: anomaly detection algorithms

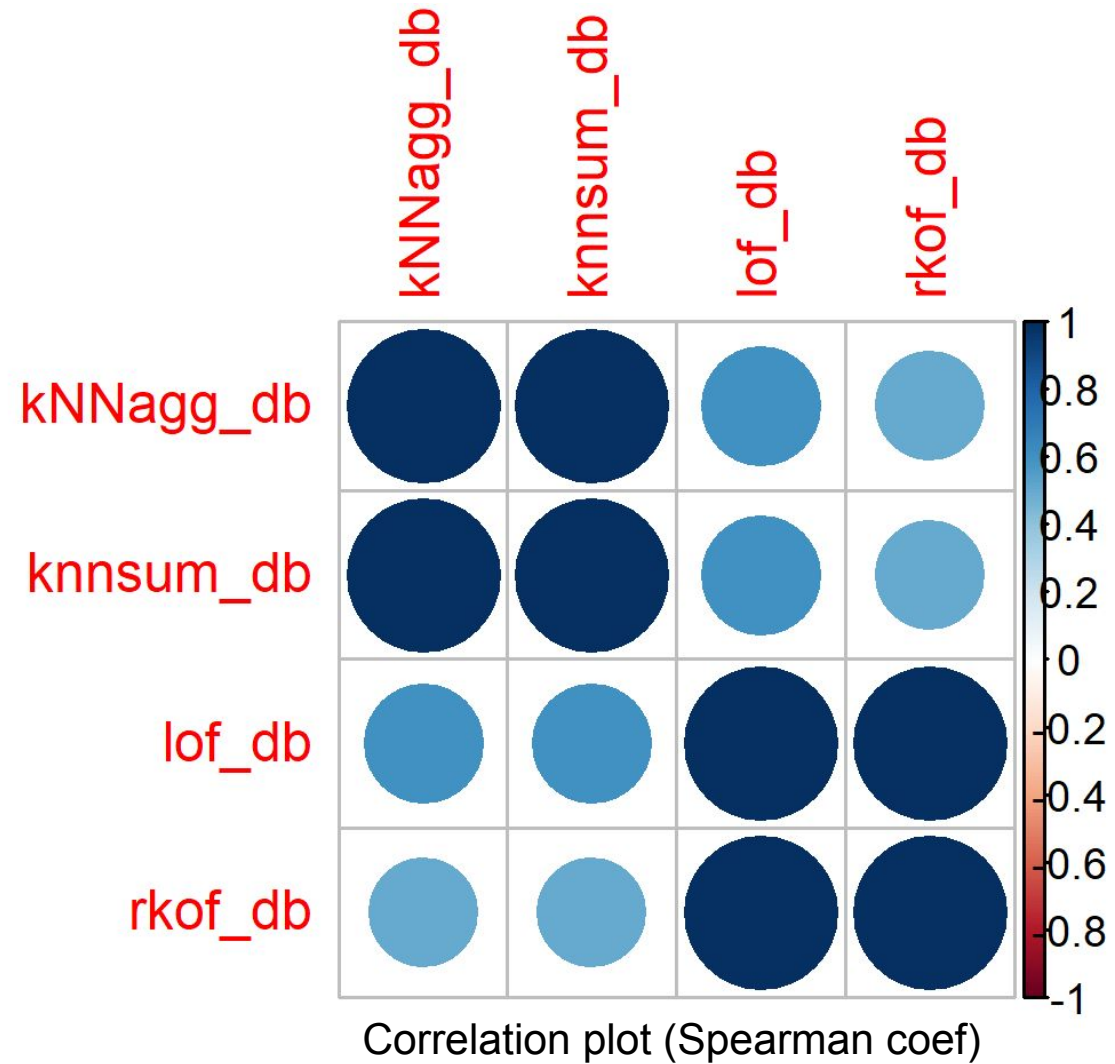
- The key idea is that firms that do frauds tend to be different from the others and the indicators of step 1 can be helpful in detecting the differences.
- A complete overview of the statistical techniques to capture anomalies was offered in the webinar held on 27 June 2024

Case study: finding anomalies in building and construction companies

- We constructed the indicators of fraud risk for a sample of Italian firms
- We ran 12 of the main algorithms of unsupervised anomaly detection on such data. We analyzed the results and decided to base the decision on 4 algorithms.
- We obtained a ranking of firms based on the anomaly score (the higher the score the higher the degree of anomaly of the firm)

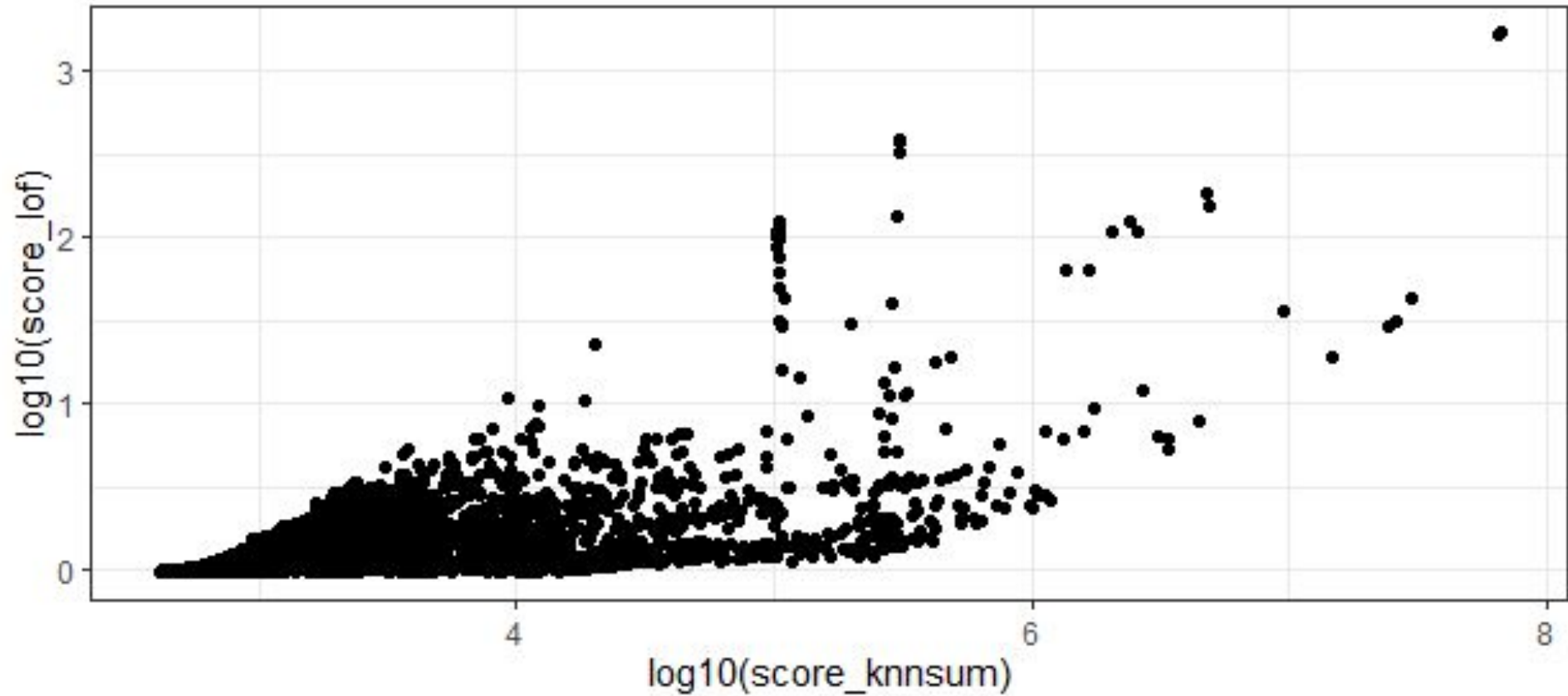
Case study: preliminary results

The algorithms give consistent responses



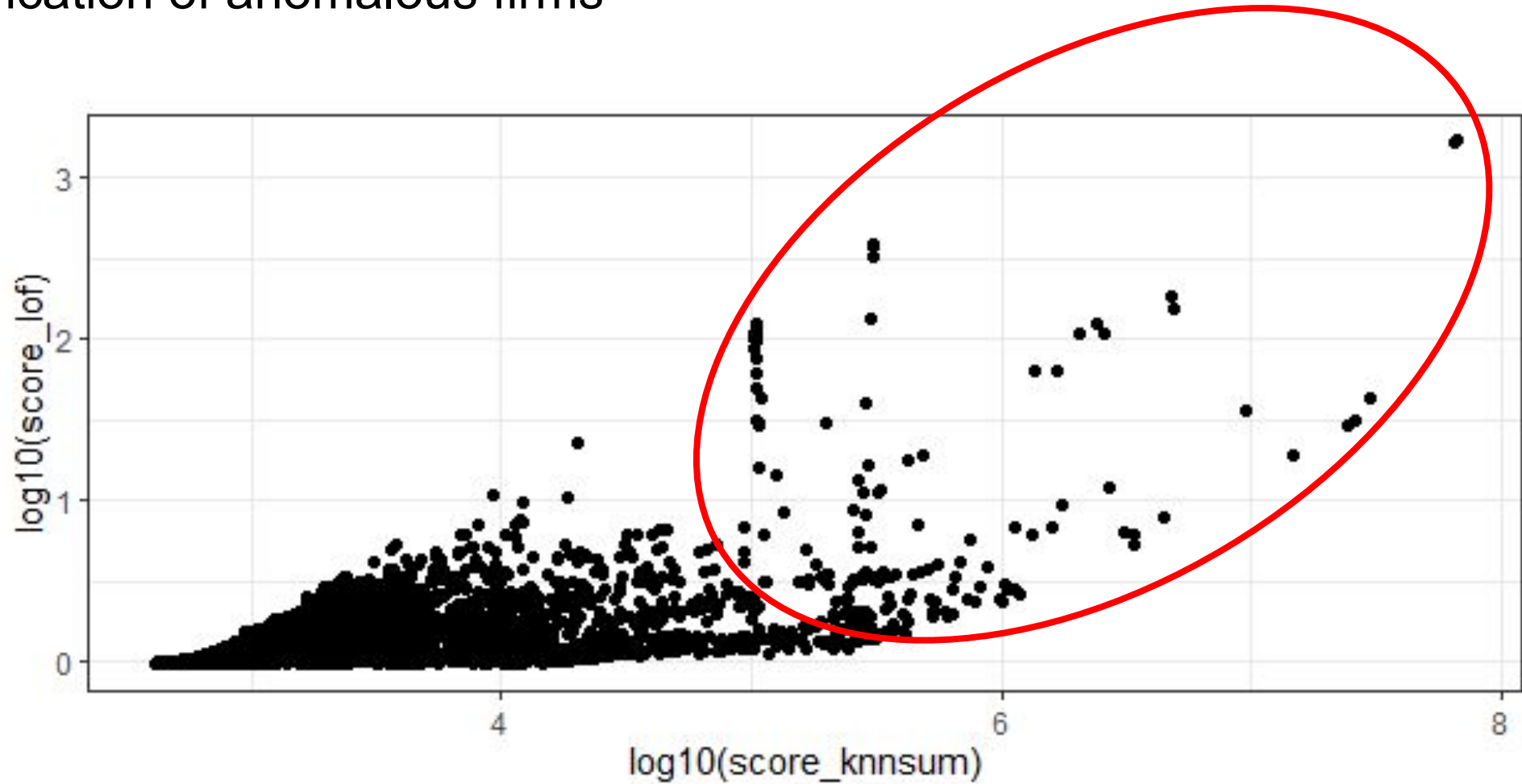
Case study: preliminary results

The identification of anomalous firms



Case study: preliminary results

The identification of anomalous firms



Step 4: validation

- Are the anomalous companies also fraudulent? Not necessarily!
- The only way to validate the procedure is to run it on a sample of firms for which is known which ones are fraudulent and which ones are not

Thank you