# Introduction

Welcome to the Belgian Business Intelligence Hackathon!

We're excited to present a unique opportunity for student data scientists to dive into the world of business analytics and risk assessment. Over the next two days, you'll have the chance to work with an exclusive dataset containing years of comprehensive financial and demographic data for Belgian companies.

# Challenges

**Develop an innovative solution for one of the following:**

1. Comprehensive Company Health Score
2. Company Clustering and Lookalike Analysis
3. Financial Anomaly Detection and Risk Alert System

### 1. Comprehensive Company Health Score

Challenge: Develop a holistic scoring model that assesses a company's overall health and risk profile.

Details:

- Create a composite score (e.g., 0-100) that reflects a company's financial stability, growth potential, and risk factors.
- Incorporate various financial ratios, growth trends, and industry benchmarks into the scoring algorithm.
- Consider both company-specific factors and industry health indicators in the model.
- Design the score to be easily interpretable, possibly with sub-scores for different aspects (e.g., liquidity, profitability, growth).

Value: This score could serve as a quick reference point for users of your business information application, allowing them to gauge a company's overall standing at a glance. It could be particularly useful for initial screening in investment decisions, credit assessments, or partnership evaluations.

### 2. Company Clustering and Lookalike Analysis

Challenge: Develop a clustering model to group similar companies and create a "lookalike" identification system.

Details:

- Use machine learning techniques (e.g., K-means, hierarchical clustering, or better ^^) to group companies based on financial attributes, growth patterns, and industry factors.
- Identify key characteristics of each cluster (e.g., high-growth tech startups, stable manufacturing firms).
- Create a "lookalike" algorithm that can find similar companies given a reference company.
- Incorporate a time dimension to track how companies move between clusters over the years.

Value: This feature could help users identify peer groups for benchmarking, find potential competitors or acquisition targets, and understand common trajectories for companies in different sectors. The lookalike functionality could be particularly useful for identifying potential business partners or investment opportunities similar to known successful companies.

3. **Financial Anomaly Detection and Risk Alert System**

Challenge: Build an advanced anomaly detection system to identify unusual financial patterns and potential risks.

Details:

- Develop algorithms to detect anomalies in financial ratios, growth rates, and other key metrics, considering both historical data and industry norms.
- Create a classification system for different types of anomalies (e.g., potential fraud, rapid growth, financial distress).
- Implement a time-series analysis to identify sudden changes or developing trends that could indicate increasing risk.
- Design an alert system that flags companies with significant anomalies and provides context for the detected issues.

Value: This system could serve as an early warning mechanism for various stakeholders. It could help investors identify potential fraud or misreporting, alert creditors to increasing default risks, and help businesses monitor their own financial health in comparison to industry norms. The system could also be used to identify companies with unusually positive performance, potentially highlighting investment opportunities.

# Dataset

To support your work, we're providing access to a rich dataset encompassing all registered Belgian companies over the past decade. This dataset includes:

Demographic Information:

- VAT number (unique identifier)
- Zip code
- NACE code (industry classification)
- Company type
- Start date
- End date (if applicable)
- Current status
- Legal situation

Financial Data (annual):

- Number of employees
- Turnover (when available)
- Current assets
- Gross operating margin
- Tangible fixed assets
- Equity
- Profit/Loss
- Current ratio
- Net cash
- Self-financing degree
- Return on equity
- Added value
- …

This dataset offers a comprehensive view of the Belgian business landscape, allowing you to analyse trends, identify patterns, and develop sophisticated models to address the challenges at hand.

# Goal

Your task is to leverage this data to create a solution that not only tackles the chosen challenge but also demonstrates potential for integration into a cutting-edge business information application. We're looking for innovative approaches that provide actionable insights and add significant value for businesses, investors, and financial analysts.

Remember, the key to success lies not just in the technical implementation, but also in how your solution can be applied to real-world business scenarios. Good luck, and we look forward to seeing your creative and impactful solutions!