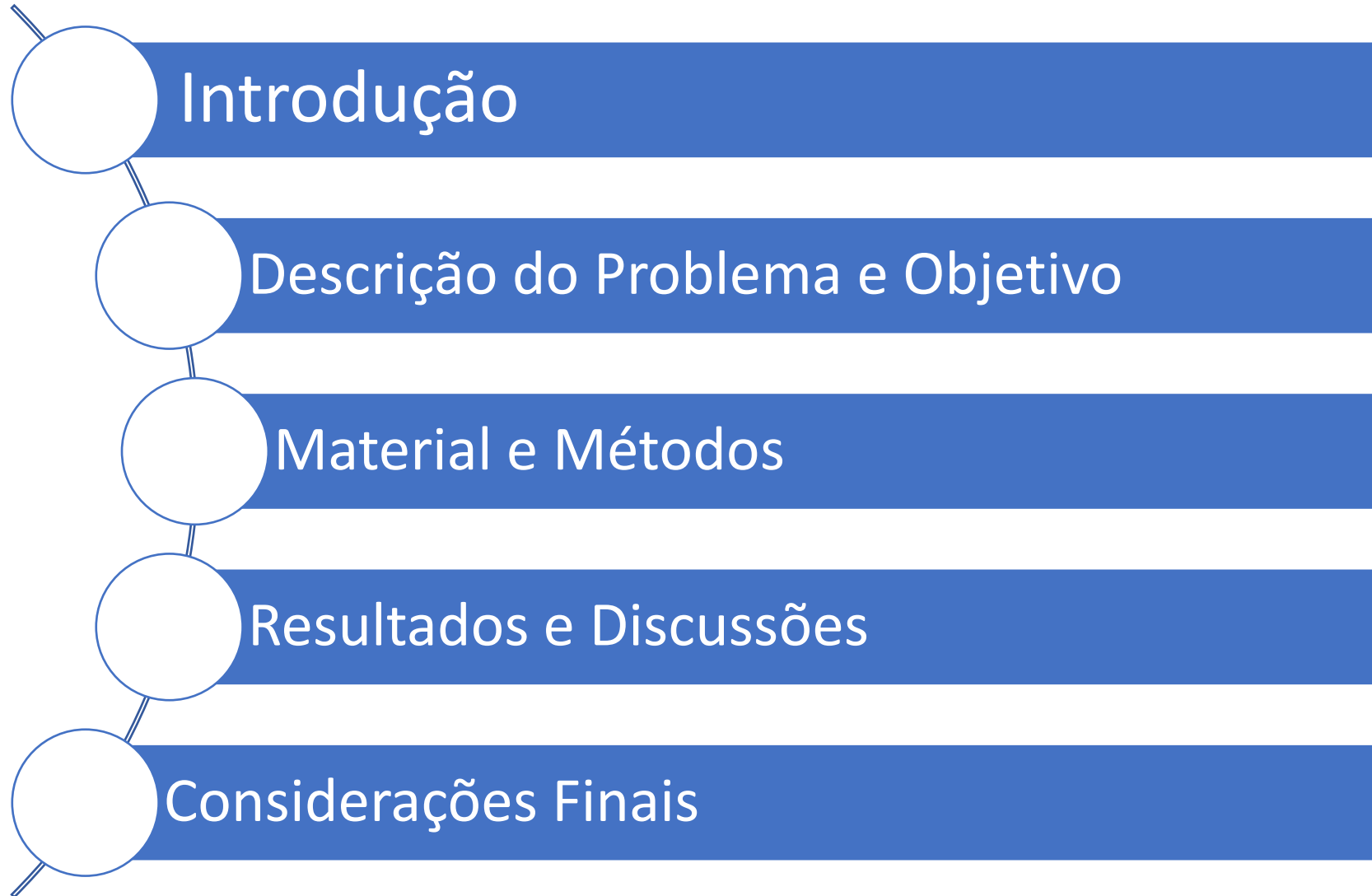
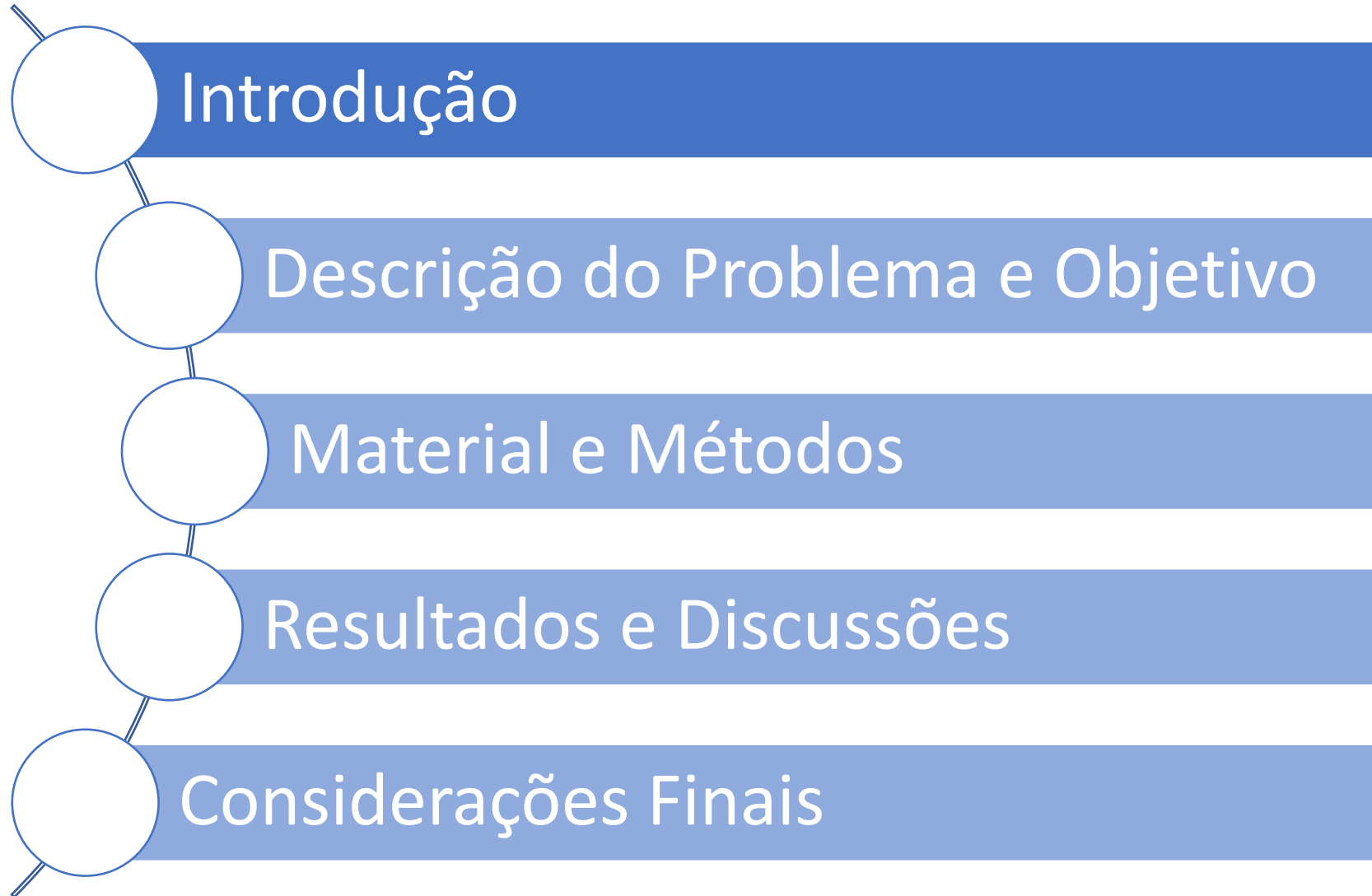


**MBA
USP
ESALQ**

**Desafios na identificação
automática de grupos econômicos
para aprimorar a gestão de riscos e
controle de crédito.**

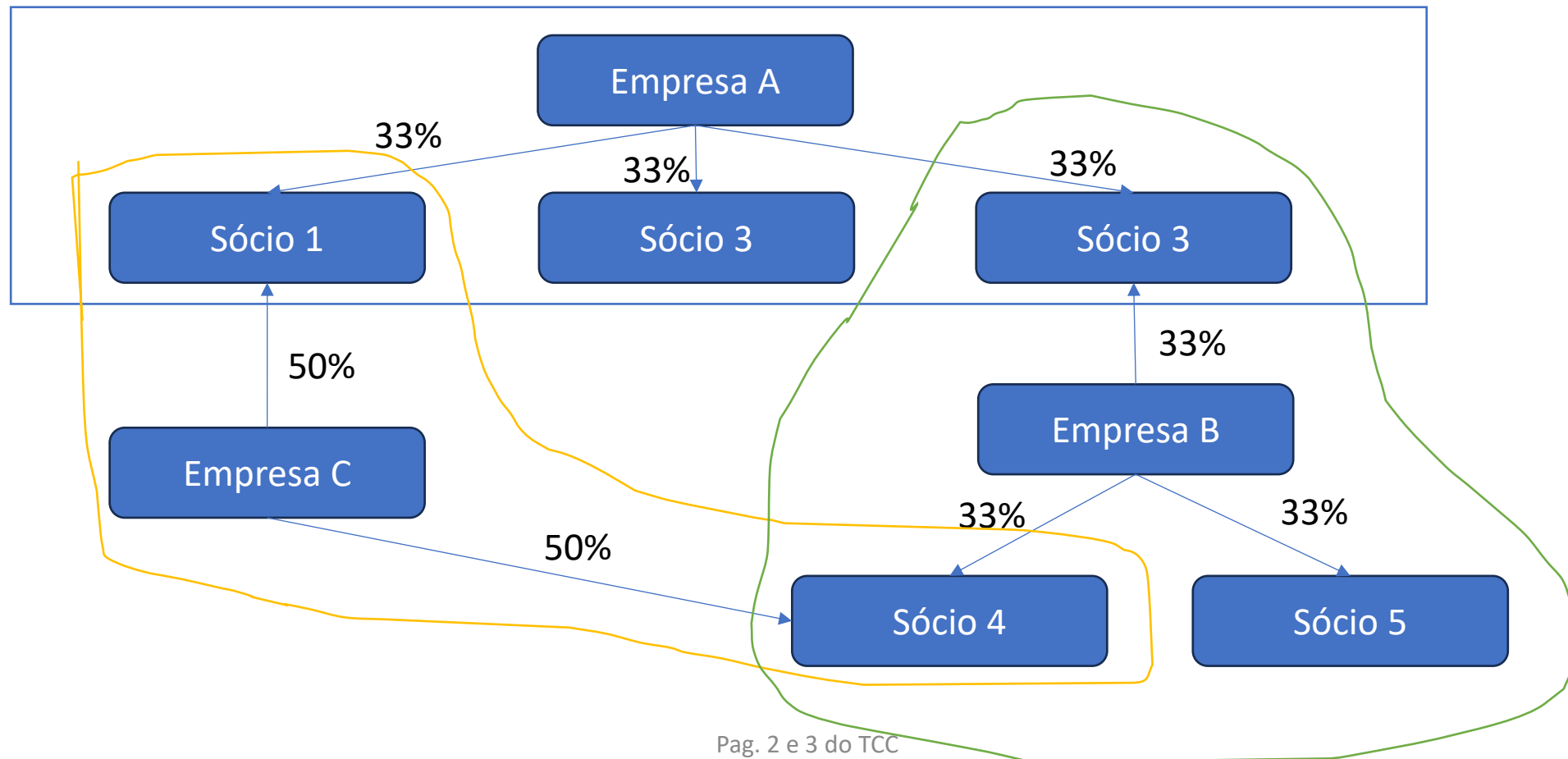
Aluno: Gustavo de Paula Aguiar Araújo
Orientador: Prof. Diego de Oliveria Cunha

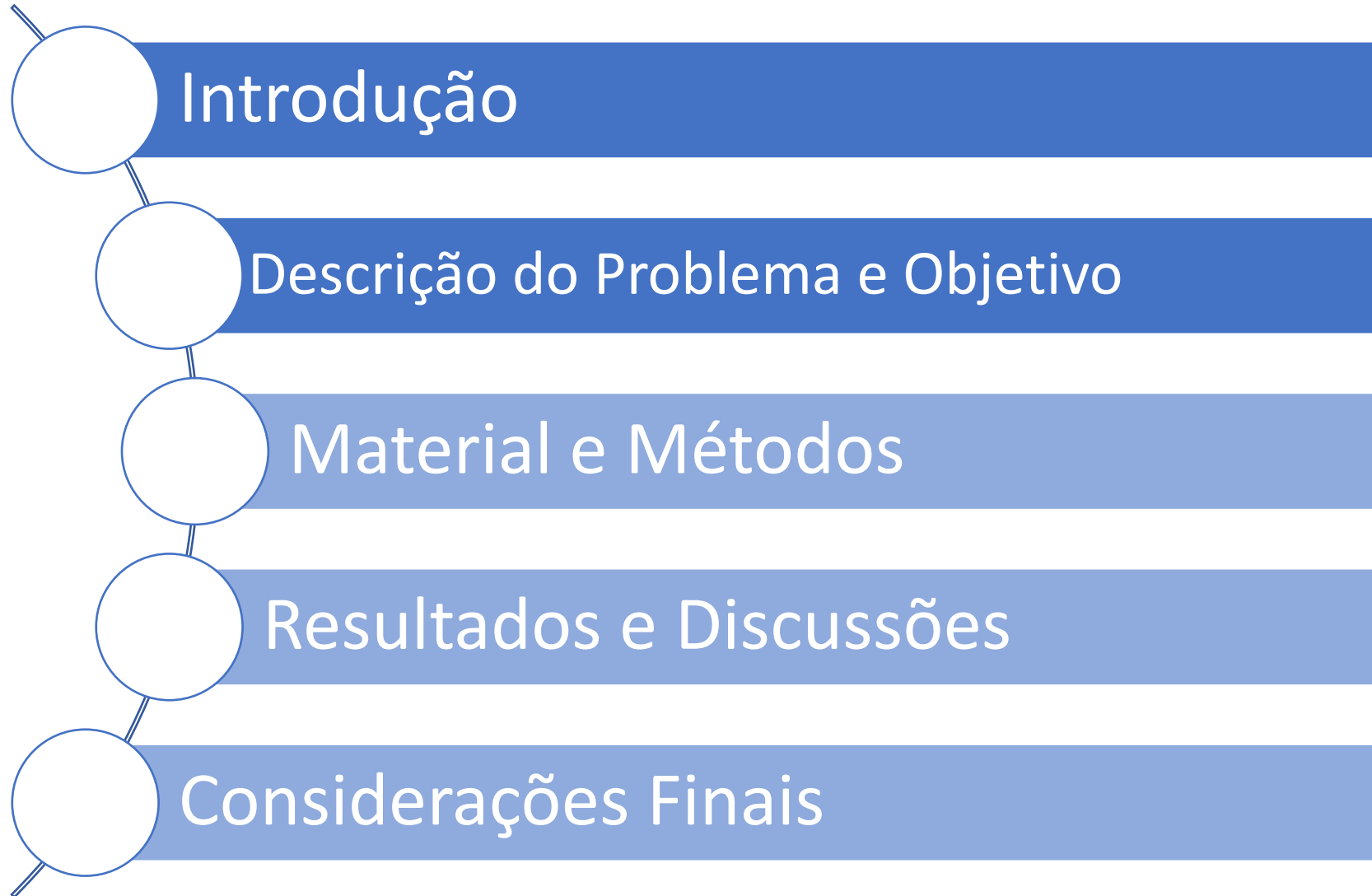




- O trabalho demonstra que existem **desafios as Instituições Financeiras [IFs]** para controle em sua principal atividade – Concessão de crédito;
 - **Crises financeiras e instabilidade econômica;**
 - **A gestão de risco e controle na concessão de crédito;**
 - **Controle de inadimplência;**
 - **Quem é o meu cliente?**
 - **Controle regulatório – Bacen;**
 - 2017 o Banco Central Brasileiro [**BACEN**] propõe a **regulamentação do controle de contrapartes** para fins de gerenciamento de risco.
- Com o advento de **novas tecnologias e modelos estatísticos** possibilitou-se melhorar os controles bancários e consequentemente **conhecer melhor o cliente;**
- É nesse contexto que o trabalho introduz o conceito de **Grupos Econômicos;**

- Afinal, o que é grupos econômicos?
 - *Conjunto de empresas que, estão interligadas por relações contratuais, cuja propriedade pertence a indivíduos ou instituições, que exercem o controle efetivo sobre essas empresas (BACEN, 2017, Art. 22) ;*





DESCRIÇÃO DO PROBLEMA

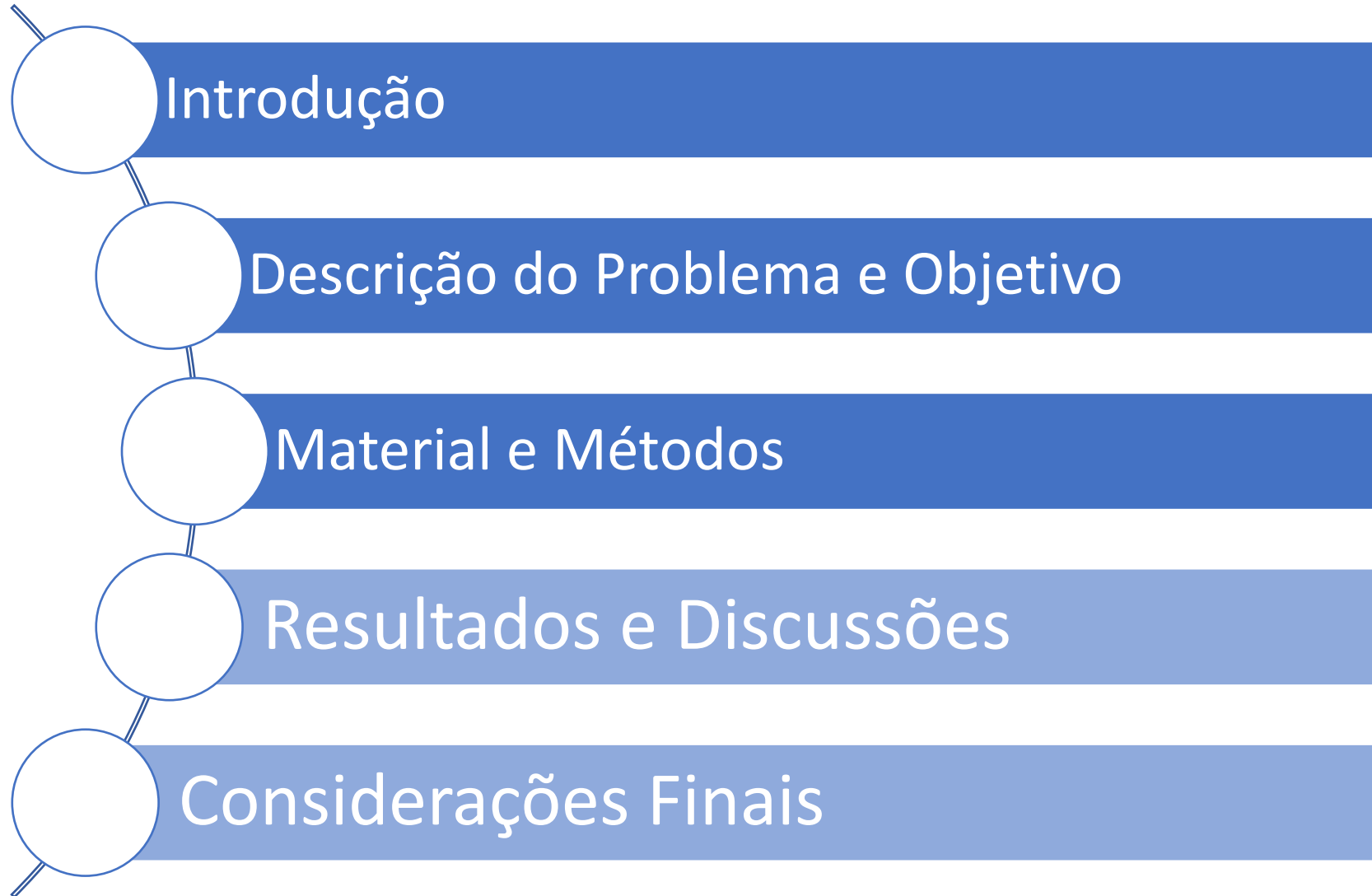
- A **norma** do BACEN e a **exigência** do mercado imprimem a necessidade de **maior controle de crédito** nas IFs;
- Análise tradicional de avaliação de crédito **sem conhecer o cliente**, pode gerar muita inconsistências e **falta de precisão**;
- Necessidade de **maior segurança e confiabilidade nas avaliações** para minimizar a exposição a riscos financeiros, **exige métodos e ferramentas mais automatizadas**;
- Quem é o meu cliente?
 - **Dificuldades** na identificação de **vínculos entre empresas e seus sócios**.

- **Objetivo do Estudo**

- Demonstrar que formar proativamente e automaticamente GEs pode possibilitar a expansão do escopo da análise de risco de crédito, incluindo vínculos indiretos entre sócios e empresas com interdependências financeiras.

- **Relevância do algoritmo proposto**

- Melhoria na capacidade das IFs em antecipar riscos e tomar decisões mais fundamentadas na concessão e gestão de crédito.
- Acelerar e automatizar a análise de risco nas IFs.
- Pode minimizar riscos de inadimplência e perdas financeiras.
- Maior abrangência ao identificar interconexões de empresas/sócios, consequentemente imprime uma alocação de crédito de forma mais estratégica e segura.



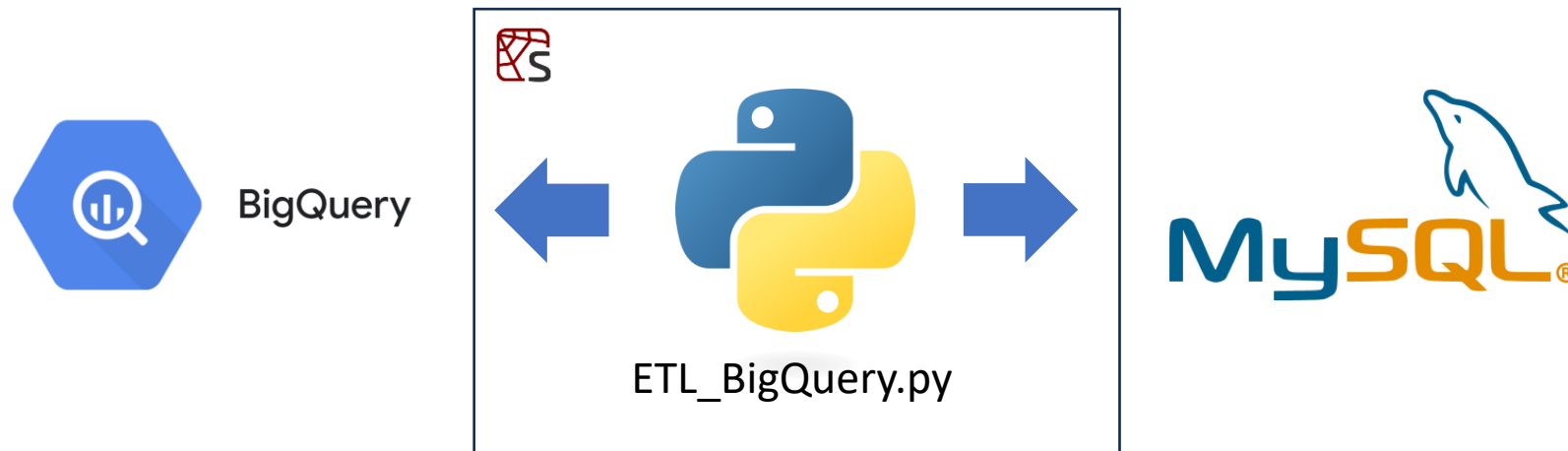
- **Metologia:**
 - **Tratamento dos dados de empresas públicas:**
 - **Data Wrangling – tratamento dos dados de entrada;**
 - **Criptografia – mascaramento dos dados;**
 - **Utilização e adaptação da teoria de grafos para identificação de grupos econômicos;**
 - **Identificação de padrões entre grafos não direcionados e grupos econômicos;**
 - **Criação de algoritmo para controle e identificação massiva de possíveis grupos econômicos;**
 - **Aplicação do algoritmo e validação da saída do modelo;**

- **Material: Metadados**

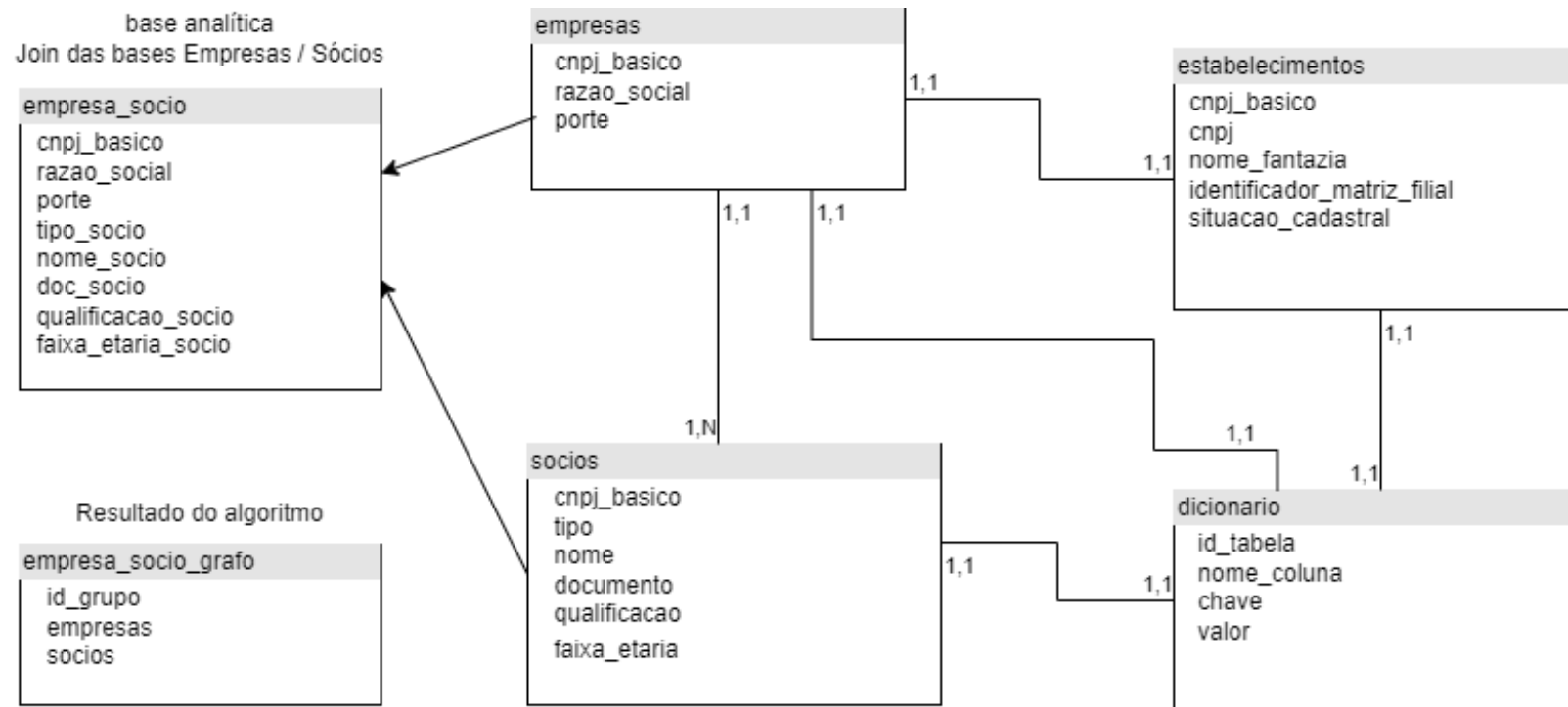
- Dados públicos de participação societária de empresas disponibilizados pelo Ministério da Economia - RFB;
- Base com aproximadamente 55MM de empresas;
 - Desafio de processamento visto o tamanho da base;
- Consumidos através do site (<https://basedosdados.org/>) base dos dados através da ferramenta BigQuery do Google;

- **Material: Data Wrangling- tratamento de dados**

- Ingestão dos dados através do BigQuery com persistência em base local em MySQL;
- Criptografia dos dados de empresas;



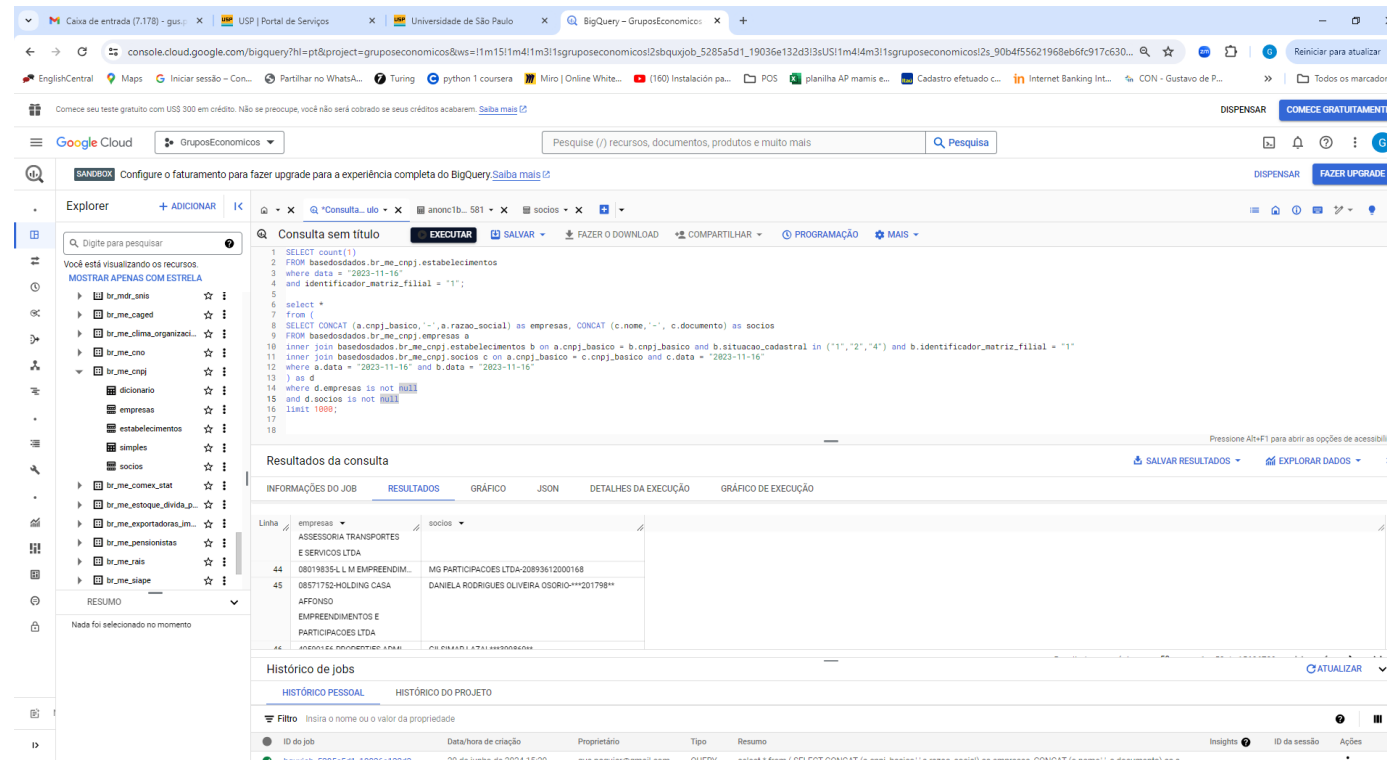
- **Material: Data Wrangling- Modelagem de dados**
 - Modelagem local dos dados para possibilitar a criação do algoritmo



Apêndice A – Imagem 1. Modelo de dados proposto - Data Wrangling

MATERIAL E MÉTODOS

- **Material: Data Wrangling- Google BigQuery**
 - *O BigQuery é um serviço de armazenamento de dados do Google Cloud que permite realizar análises em grande escala e com baixo custo. Ele é um data warehouse corporativo totalmente gerenciado e sem servidor*



Apêndice A – Imagem 2 – Data Wrangling – Plataforma Google Bigquery

- Material: Data Wrangling - Desafios

- Desafios de hardware;

- **Utilização de Chunks** = divisões de dados em partes menores e mais gerenciáveis

```
def insert_data_in_chunks(engine, dataframe, table_name, chunk_size=1000):  
    #captura horario do comeco do processo  
    ini = datetime.now()  
  
    print(f'insert_data_in_chunks: Inicio do processo de inclusao {table_name}: ',ini.strftime('%H:%M:%S'))  
    with engine.connect() as connection:  
        for start in range(0, len(dataframe), chunk_size):  
            end = start + chunk_size  
  
            chunk = dataframe.iloc[start:end] #recupera 1000 linhas  
  
            chunk.to_sql(name=table_name, con=connection, if_exists='append', index=False)  
            #print(f"Inserção via chunk {start} to {end}")  
  
    fim = datetime.now()  
    dif = (fim - ini)  
    print(f'insert_data_in_chunks: Fim do processo de inclusao {table_name}: ',dif)
```

Possibilitou gerenciar a base de 55 MM de empresas e persistir os dados no MySQL localmente.

Fonte: Resultados originais da pesquisa

- **Material: Data Wrangling - Criptografia**
 - Para controle dos dados foi incluído também um algoritmo de criptografia para dispor as informações públicas das empresas, sem necessariamente identifica-las.

| | empresas | socios |
|---|--|------------------------|
| ▶ | 08169126-FR COMERCIO DE CALCADOS E ACESSORI... | ACMDO*****_***364671** |
| | 08169126-FR COMERCIO DE CALCADOS E ACESSORI... | AJM*****_***599401** |
| | 24705950-RESTAURANTE CANTINHO DO SABOR LTDA | AJM*****_***599401** |
| | 24705950-RESTAURANTE CANTINHO DO SABOR LTDA | DM*****_***643061** |

Figura - Chaves únicas de referência entre empresas e sócios

Fonte: Resultados originais da pesquisa

- **Material: Grafos**

- Semelhança de da teoria dos grafos e as ligações entre empresas e sócios coligados

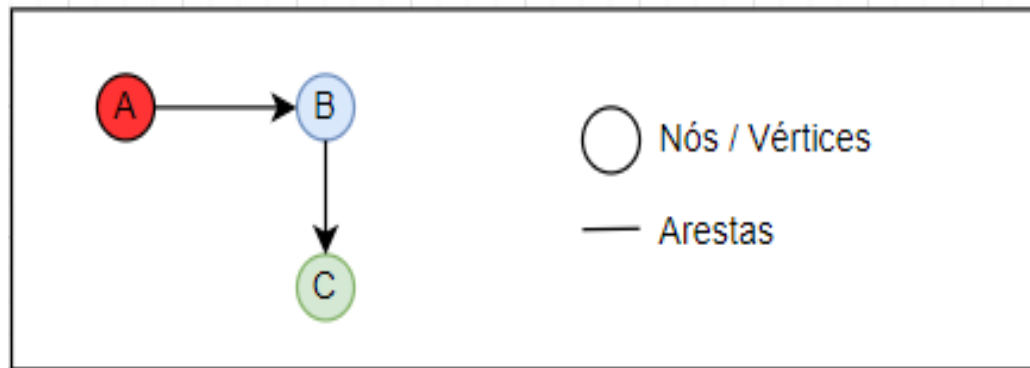


Figura 2. Grafo direcionado

Fonte: Resultados originais da pesquisa

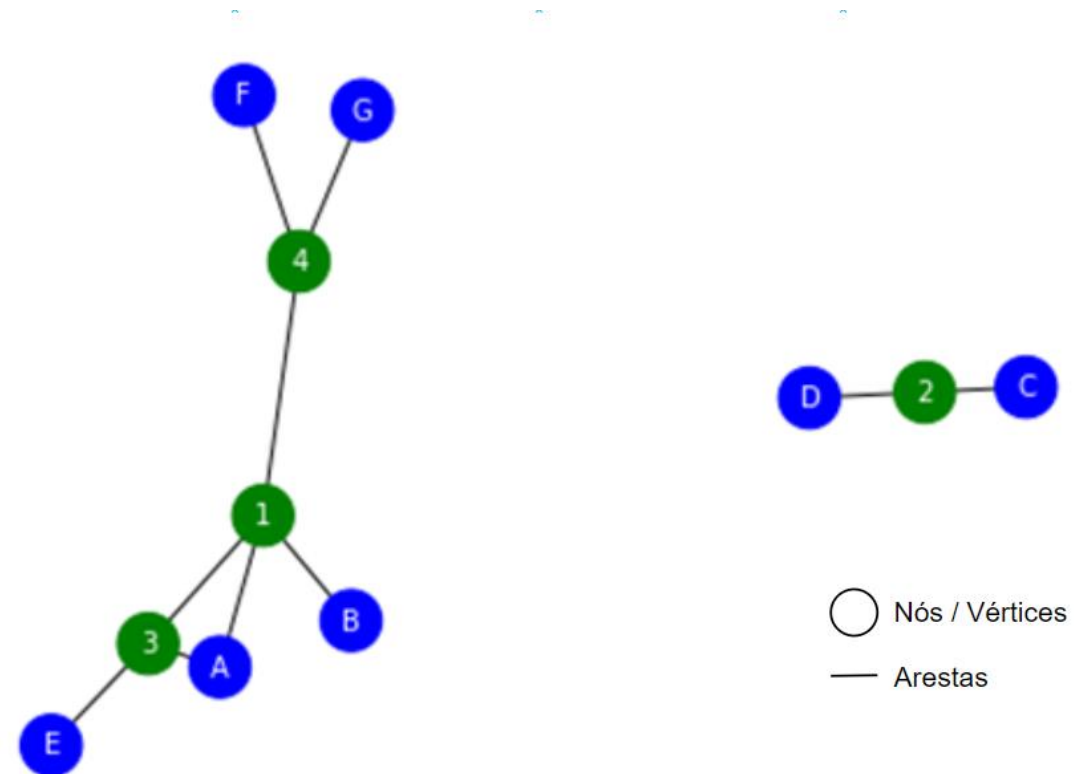
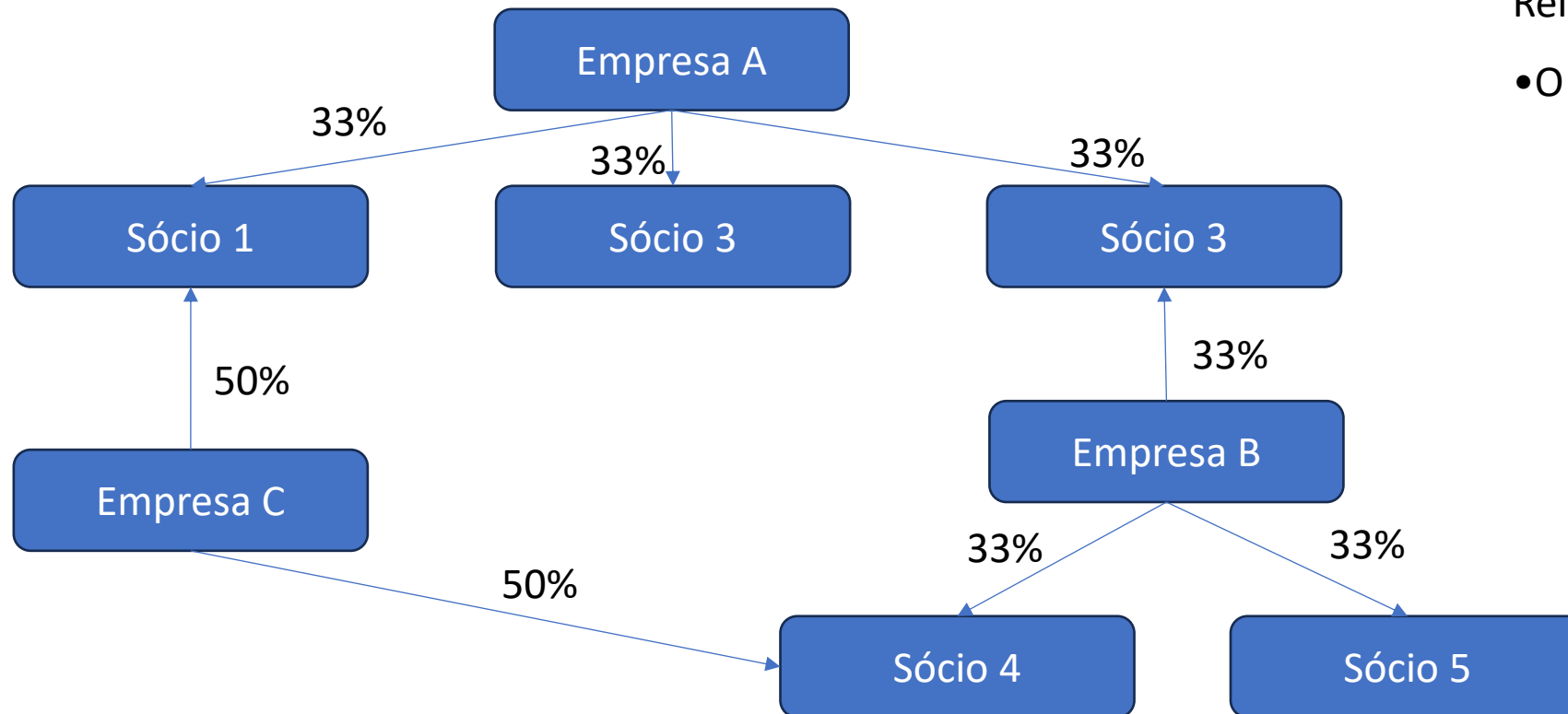


Figura 3. Estrutura de Grafo não direcionado

Fonte: Resultados originais da pesquisa

- **Método: Grafos**

- Percebe-se similaridade em grafos o processo de identificação e formação de Grupos Econômicos;



Relembrando

- O que é grupos econômicos?
 - *Conjunto de empresas que, estão interligadas por relações contratuais, cuja propriedade pertence a indivíduos ou instituições, que exercem o controle efetivo sobre essas empresas;*

- **Método: Grafos e Grupos Econômicos**

- Identificam-se como empresas as os nós vinculados as empresas por arestas;
- Desta forma é possível identificar de forma automática vínculos entre as empresas e sócios coligados;

```
def criar_grafo(socios, empresas, relacoes, chunk_size=1000):  
    """  
    Cria um grafo que representa as relações entre sócios e empresas.  
    :param socios: Lista de sócios.  
    :param empresas: Lista de empresas.  
    :param relacoes: Lista de tuplas representando relações (socio, empresa).  
    :return: Um grafo NetworkX.  
    """  
    G = nx.Graph()  
  
    """  
    #cria os nós do tipo sócios  
    for index, row in socios.iterrows():  
        G.add_node(row['socios'], tipo='socio')  
  
    #cria os nós do tipo empresas  
    for index, row in empresas.iterrows():  
        G.add_node(row['empresas'], tipo='empresa')  
    """  
    # Criar threads para adicionar nós de sócios e empresas  
    with ThreadPoolExecutor(max_workers=2) as executor:  
        #.result é utilizado para esperar a finalização da thread  
        executor.submit(adicionar_nos_socios, G, socios).result()  
        executor.submit(adicionar_nos_empresas, G, empresas).result()  
  
    #vincula os nós e cria as arestas do grafo  
    for row in relacoes.itertuples(index=False):  
        G.add_edge(row.socios, row.empresas)  
  
    return G
```

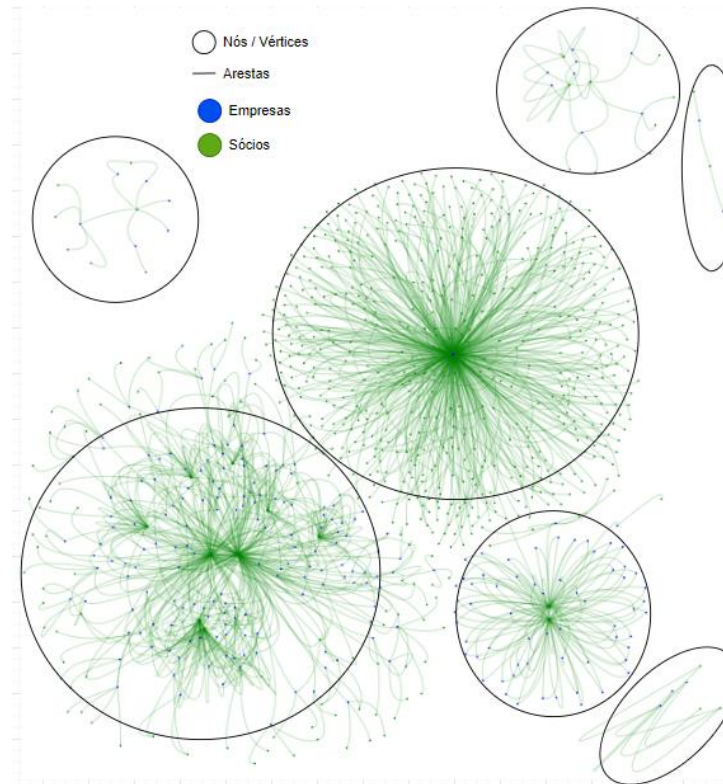


Figura 4. Grafo de referência

Fonte: Resultados originais da pesquisa

- Método: Grafos e Grupos Econômicos

Sócio AJM possui vínculo societário com a FR Comercio e o Restaurante Cantinho

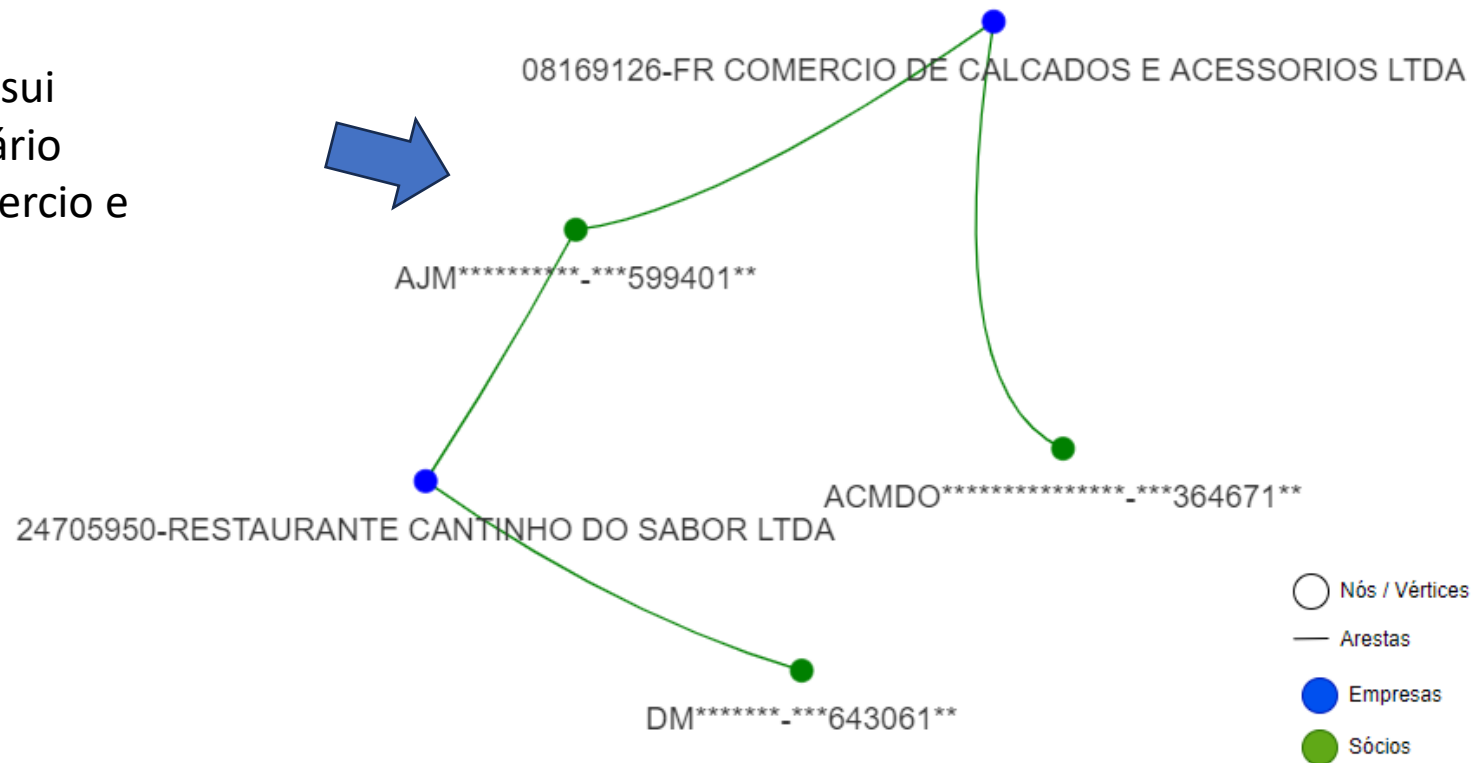
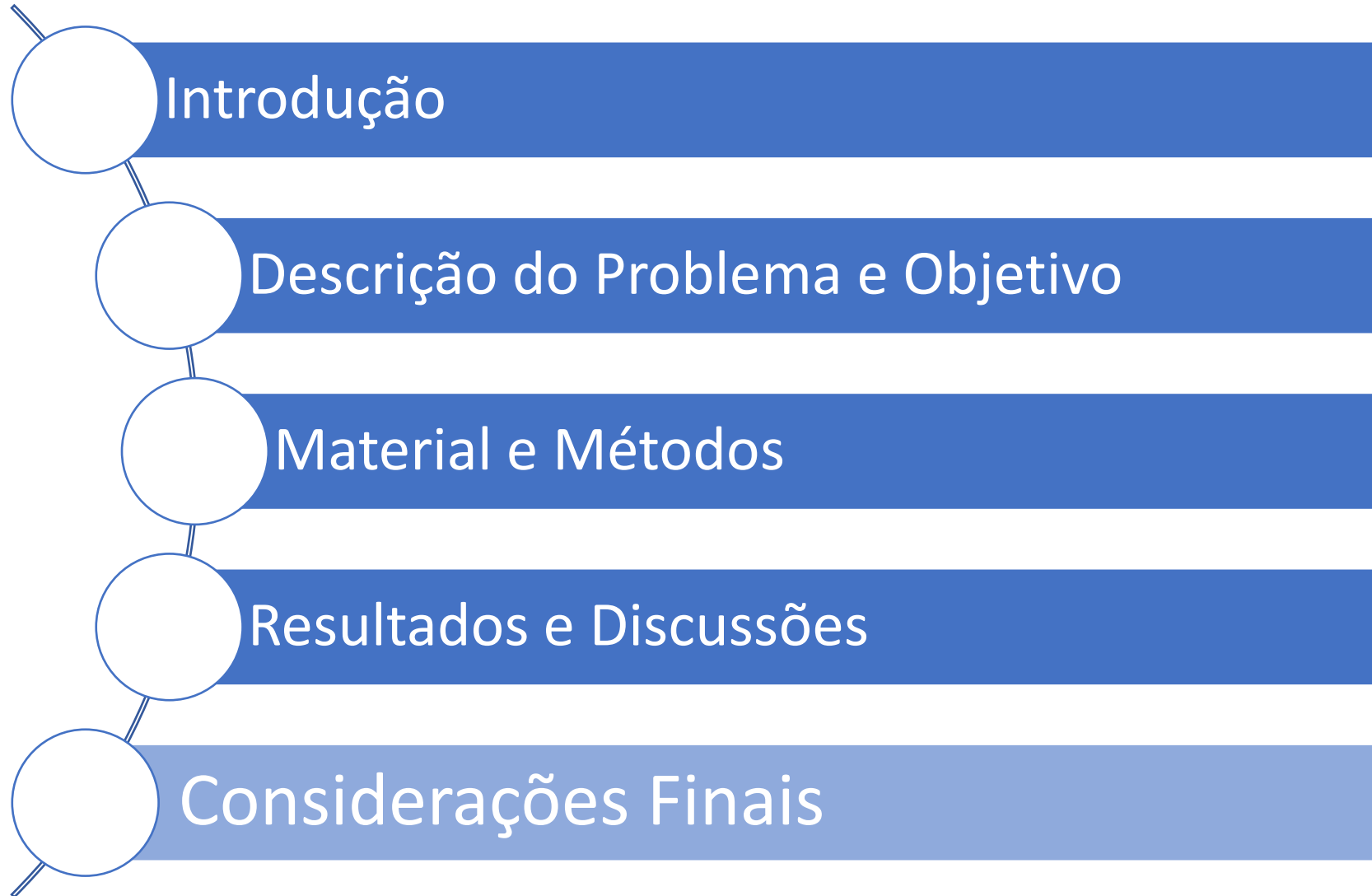


Figura 5. Grupo Econômico – Vínculo Societário

Fonte: Resultados originais da pesquisa



• Resultados

- O algoritmo conseguiu identificar automaticamente possíveis estruturas de GEs, oferecendo uma visão mais clara dos vínculos entre empresas e sócios.
- Demonstrou de forma gráfica os vínculos entre as empresas e seus sócios;
- Conseguiu identificar todas as relações entre empresas e sócios da base de forma automática e em menos de 30 min em hardware simples. Demonstrando que é possível escalar a solução;

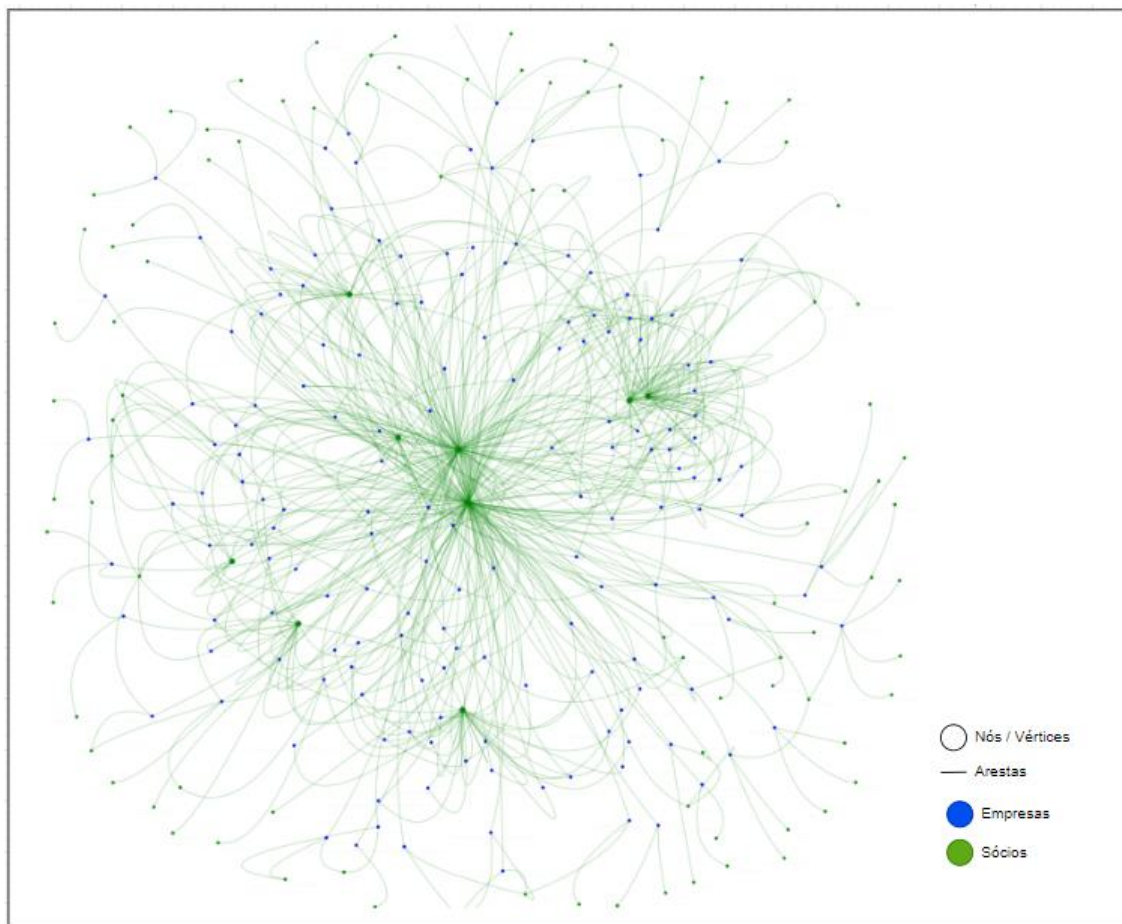


Figura 6. Grupo Econômico – Vínculo Societário com grande ramificação

Fonte: Resultados originais da pesquisa

- Benefícios:
 - Percebe-se que há correlações entre sócios e empresas coligadas;
 - Esse vínculo entre as empresas, demonstra que a análise isoladamente das empresas e sócios pode ser insuficiente ao avaliar o risco de crédito envolvido
 - Possibilita acelerar o processo de análise e redução do risco de exposição em carteiras de crédito;

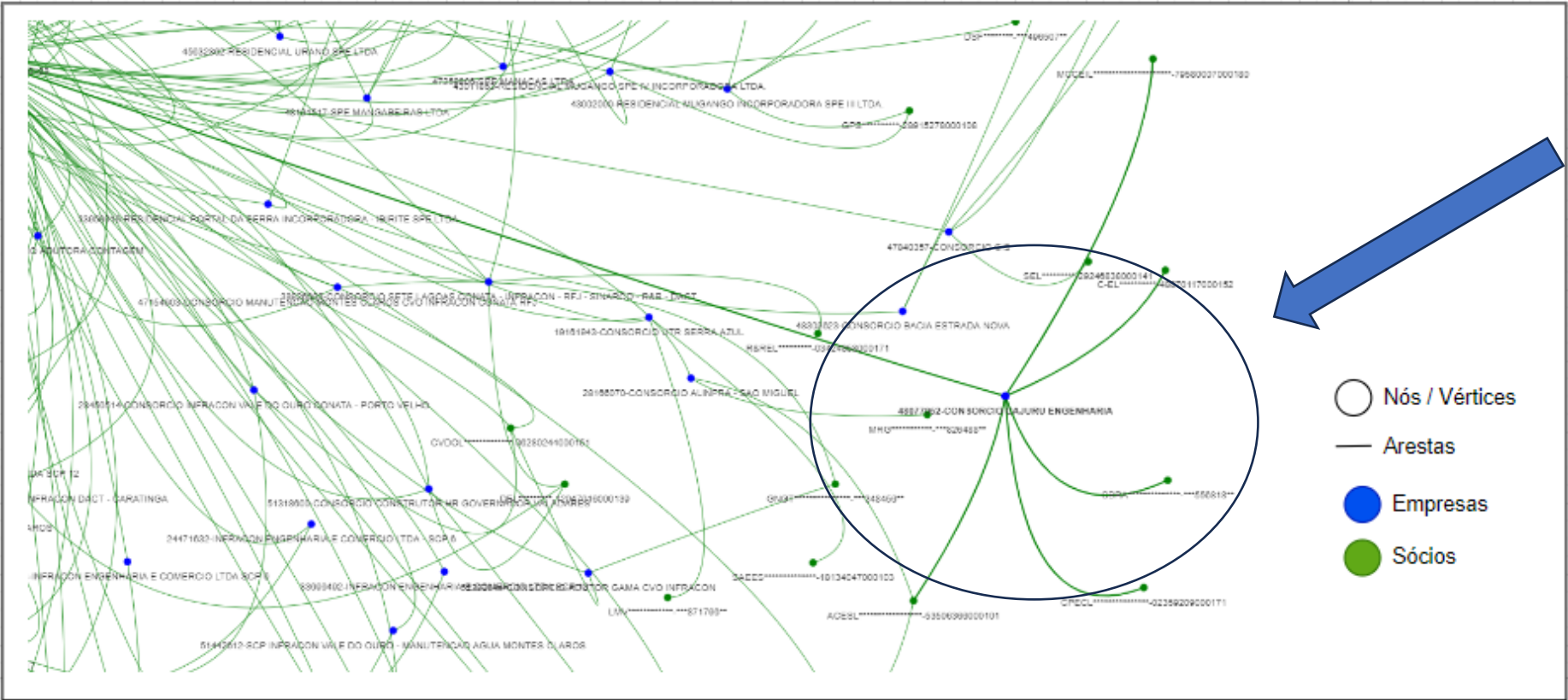
- **Discussões:**

- A identificação de relações societárias permite conhecer melhor os clientes envolvidos na concessão de crédito;

- **Limitações:**

- A consideração exclusiva dos vínculos entre empresas e sócios pode gerar grupos complexos, com muitos vínculos.
 - Distâncias grandes entre os nós podem obscurecer a visão do risco agregado, reduzindo a relevância de certos vínculos para a avaliação do risco de crédito.

RESULTADOS E DISCUSSÕES



Positivos:

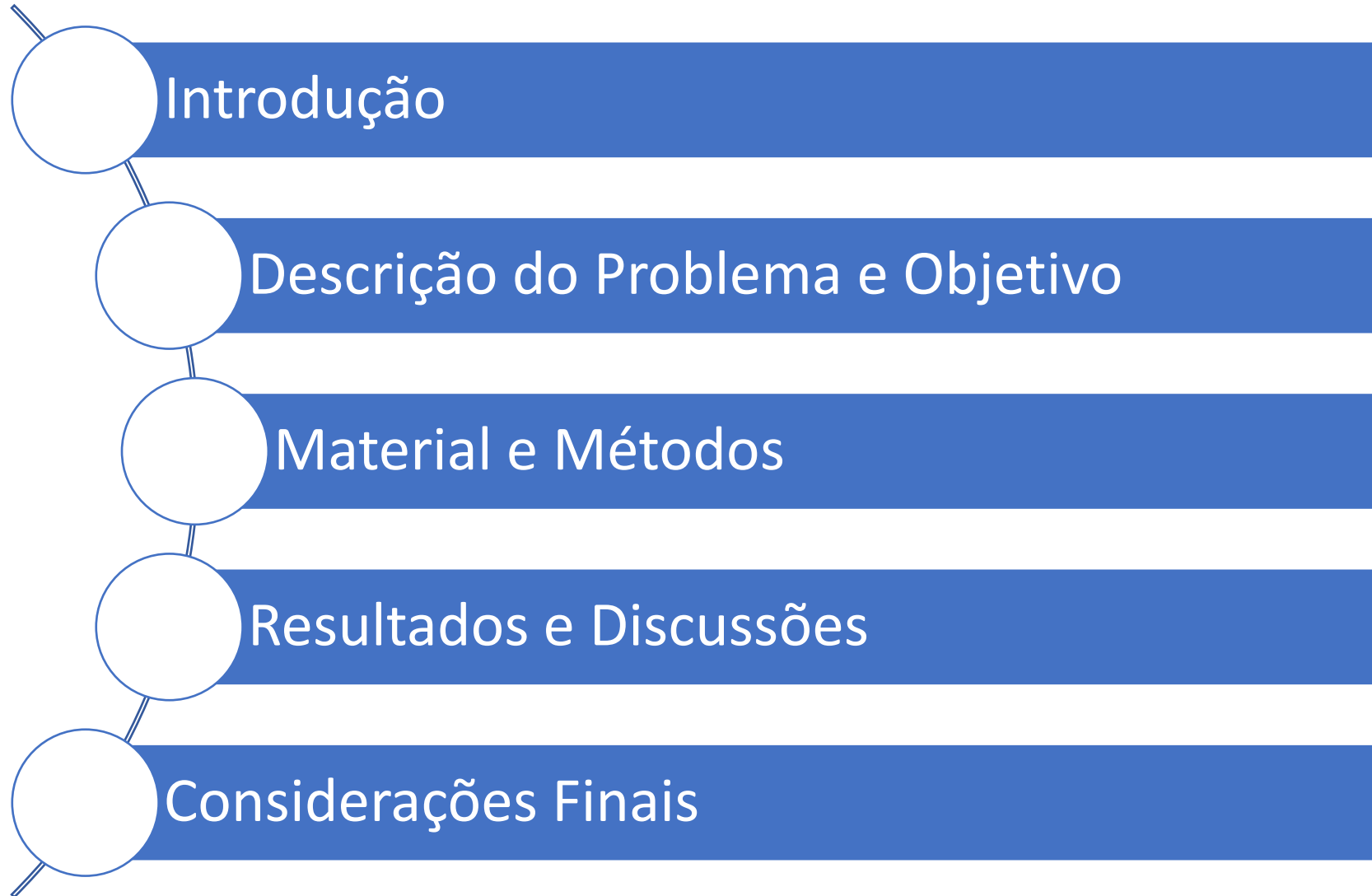
- Possibilita identificar até a última referência societária entre indivíduos do Grupo;

Negativos:

- Não é possível dizer que esse vínculo distante pode trazer algum risco para análise de crédito para todos os outros integrantes do Grupo;

Figura 7. Grupo Econômico – Vínculo Societário – Zoom vínculo com muitas arestas e distante

Fonte: Resultados originais da pesquisa



- **Contribuições:**

- A implementação do algoritmo mostrou-se eficaz na identificação de GEs, fornecendo insights valiosos para a gestão de risco em IFs;
- Automatização pode acelerar a análise de grandes volumes de dados, possibilitando decisões rápidas e precisas;
- Identificação de interdependências financeiras que poderiam passar despercebidas, reduzindo a exposição ao risco;

- **Limitações:**

- Dependência de dados públicos e precisão variável na identificação de influência de sócios.
- Ou seja, o algoritmo não consegue reconhecer a relevância de sócios minoritários no risco geral do grupo.

- **Pesquisas Futuras:**

- Exploração de variáveis adicionais, como dados alternativos e redes sociais corporativas, para aumentar a robustez e precisão na identificação de GEs;
- Integração de dados alternativos, transações financeiras, vínculos de procuração, etc. ;
- Uso de aprendizado de máquina para prever a formação de novos grupos e analisar variáveis qualitativas;

- **Conclusões:**

- A automatização da análise de GEs pode representar um avanço significativo na gestão de risco de crédito, proporcionando às IFs ferramentas para uma gestão mais eficiente;
- Modelos de grafos são um ponto de partida promissor, mas devem ser complementados por análises adicionais para decisões de crédito mais seguras e informadas.

REFERÊNCIAS

- Banco Central do Brasil [BACEN]. 2017. Resolução nº 4557, de 23 de fevereiro. Dispõe sobre a estrutura de gerenciamento de riscos e apetite por riscos nas instituições financeiras. Disponível em: <https://normativos.bcb.gov.br/Lists/Normativos/Attachments/50344/Res_4557_v1_O.pdf>. Acesso em: 10 jun. 2024.
- Base dos Dados. 2023. Quadro Societário CNPJ. Disponível em: <<https://basedosdados.org/dataset/e43f0d5b-43cf-4bfb-8d90-c38a4e0d7c4f?table=81272674-f522-4e43-a70b-05bf46f0a163>>. Acesso em: 10 mar. 2024.
- Camargos, M.A. 2012. A inadimplência em um programa de crédito de uma instituição financeira pública de minas gerais: uma análise utilizando regressão logística. REGE - Revista de Gestão. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1809227616303204>>. Acesso em: 10 mar. 2024.
- Gonçalves, Reinaldo. 1991. Grupos econômicos: uma análise conceitual e teórica. Disponível em: <<https://periodicos.fgv.br/rbe/article/view/534>>. Acesso em: 10 mar. 2024.
- Google BigQuery. Serviço de armazenamento de dados. 2024. Disponível em: <<https://cloud.google.com/bigquery>>. Acesso em: 10 jun. 2024.
- Governo do Brasil. Cadastro Nacional da Pessoa Jurídica – CNPJ. 2024. Disponível em: <<https://dados.gov.br/dados/conjuntos-dados/cadastro-nacional-da-pessoa-juridica---cnpj>>. Acesso em: 24 mar. 2024.
- Instituto de Pesquisa Econômica Aplicada [IPEA]. 2006. Risco de Crédito: Desenvolvimento de modelo Credit Scoring para a gestão da inadimplência de uma instituição de microcrédito. Disponível em: <https://www.ipea.gov.br/ipeacaixa/premio2006/docs/trabpremiados/IpeaCaixa2006_Profissional_02lugar_tema03.pdf>. Acesso em: 30 jun. 2024.
- Jassé, P. 2020. Gestão Do Risco De Crédito Bancário: Estudo Empírico. Disponível em: <<https://comum.rcaap.pt/handle/10400.26/31674>>. Acesso em: 30 jun. 2024.
- Gestel, T. V; Baesens, B. 2008. Credit Risk Management: Basic Concepts financial risk componentes, rating analysis, models, economic and regulatory capital. Disponível em: <https://www.academia.edu/37069057/Credit_Risk_Management_Basic_Concepts>. Acesso em: 01 abr. 2024.
- Wilson, R. J. 1996. Introduction to Graph Theory. 4. ed. Harlow: Pearson Education.

MBA
USP
ESALQ

Obrigado!

Aluno: Nome do Aluno

Orientador: Prof. Marcos dos Santos