

## Roteiro do Integrante 4: O Investigador (Visualização e Conclusão)

### Introdução:

"Agora que processamos os dados e calculamos as métricas, meu papel é traduzir esses números em evidências visuais e conclusões práticas para a auditoria. Vamos analisar o que o nosso 'motor de grafos' revelou."

---

### Seção 8.1: Visualização Básica da Rede

*(Mostre o primeiro grafo colorido gerado)*

#### Fala Sugerida:

"Começamos com a **Visualização 8.1**, que é o mapa geral do território.

- Aqui, cada **nó (bolinha)** é uma empresa.
- As **cores** não são aleatórias: elas representam as **comunidades** que o algoritmo de Louvain detectou.
- O **tamanho** dos nós é proporcional à importância da empresa (Centralidade de Grau).

O que isso nos diz? Que o mercado não é homogêneo. Existem grupos bem definidos que operam isoladamente, o que pode indicar divisão regional ou especialização, mas também reservas de mercado."

---

### Seção 8.2: Visualização com Peso das Arestas

*(Mostre o código e o gráfico onde as linhas têm espessuras diferentes)*

#### Fala Sugerida:

"Na **Seção 8.2**, refinamos a visualização para focar na intensidade das relações.

No código, vejam a linha: `edge_widths = [w / max_weight * 5 for w in weights]`

O que fizemos aqui foi definir a **espessura das linhas** (arestas) com base no **peso** da relação.

- Uma linha fina significa que as empresas se encontraram em apenas uma licitação.
- Uma **linha grossa** indica que elas participaram de múltiplas licitações juntas.

Visualmente, isso nos permite ignorar o ruído e focar apenas nas conexões fortes e recorrentes, que são os principais indícios de conluio."

---

### 🔥 Seção 8.3: Heatmap de Co-participações

*(Mostre a matriz de calor quadrada)*

#### Fala Sugerida:

"A **Seção 8.3** traz o Heatmap, ou Mapa de Calor.

Construímos uma **Matriz de Adjacência** onde os eixos X e Y são as empresas. O código varre todos os pares e preenche a célula com a contagem de encontros.

**Interpretação:** Quanto mais quente a cor (vermelho/escuro), mais vezes aquelas duas empresas disputaram o mesmo contrato. Em uma auditoria, nós não perderíamos tempo olhando para as áreas claras. Nós iríamos direto nos **quadrados vermelhos**, pois ali estatisticamente existe uma anomalia de comportamento."

---

### 🕒 Seção 9.1: Identificar Cliques

*(Mostre a lista de cliques detectados)*

#### Fala Sugerida:

"Entrando na detecção de padrões suspeitos na **Seção 9.1**, usamos a função `nx.find_cliques(G)`.

Na teoria dos grafos, um **Clique** é um grupo onde **todos os membros estão conectados a todos os outros**.

- Identificamos um clique com 10 empresas.
- Isso significa que existe um grupo fechado onde todas essas 10 empresas se conhecem e competem entre si.

Se descobrirmos que as vitórias se alternam apenas entre os membros desse clique, temos a definição clássica de um cartel operando em rodízio."

---

### 💛 Seção 9.2: Alta Frequência de Co-participação

*(Mostre a tabela de pares suspeitos)*

#### Fala Sugerida:

"Na **Seção 9.2**, fomos cirúrgicos.

O código filtra as arestas com a condição: `if data['weight'] >= 5`. Ou seja, estamos isolando apenas os pares de empresas que participaram de **5 ou mais licitações juntas**.

O resultado gerou uma lista de pares suspeitos, como [Citar exemplo do output, ex: Zeta e Theta]. Isso direciona a fiscalização: ao invés de auditar o setor inteiro, o auditor deve intimar especificamente essas empresas para explicar tal coincidência."

---

### **Seção 9.3: Análise de Vitórias por Comunidade**

*(Mostre o código da query e os gráficos de barras)*

#### **Fala Sugerida:**

"A **Seção 9.3** cruza os dados do Grafo com os dados financeiros.

Usamos uma query Cypher (`query_vencedores`) para extrair quem ganhou cada contrato e somar os valores. Depois, agrupamos isso pelas **Comunidades** que detectamos antes.

Os gráficos de barras nos mostram:

1. **Azul:** Qual comunidade ganha mais contratos (volume).
2. **Laranja:** Qual comunidade leva mais dinheiro (valor).

Se uma comunidade pequena (poucas empresas) tem uma barra laranja gigante, isso indica uma concentração de renda suspeita, sugerindo que aquele grupo pode estar manipulando os preços de contratos de alto valor."

---

### **Seção 9.4: Queries Cypher para Análise**

*(Mostre a lista de queries)*

#### **Fala Sugerida:**

"Na **Seção 9.4**, deixamos preparadas 'perguntas prontas' para o banco de dados Neo4j.

Estas queries permitem ao auditor aprofundar a investigação em tempo real:

- Podemos perguntar o '**Caminho mais curto**' (`shortestPath`) para ver se uma empresa suspeita tem ligação direta com uma empresa idônea.

- Podemos filtrar empresas que **sempre perdem** (possíveis 'laranjas' apenas para fazer número).

Isso transforma o notebook em uma ferramenta de consulta dinâmica."

---



## **Seção 10.1: Conclusões e Relatório Final**

*(Mostre o resumo final gerado pelo código)*

### **Fala Sugerida:**

"Para encerrar, a **Seção 10.1** gera o **Relatório Executivo Automatizado**.

O código compila todas as métricas que calculamos — número de nós, conexões, coeficiente de clustering e os principais 'players' (baseado em centralidade) — e imprime um resumo legível.

**Conclusão do Trabalho:** Saímos de dados brutos e desconexos para um **painel de inteligência**. Conseguimos não apenas visualizar o mercado, mas apontar com precisão matemática onde estão os riscos de fraude.

Isso prova que o uso de Grafos e Ciência de Redes é uma das armas mais poderosas disponíveis hoje para combater a corrupção e aumentar a transparência pública."