



WIDaT

2022

UFES

# **DESCOBERTA DE CONHECIMENTO APLICADA À BASE DE DADOS ABERTOS DA ANVISA SOBRE PREÇOS DE MEDICAMENTOS POR MEIO DE ANÁLISE DE REDES DE INFORMAÇÃO**

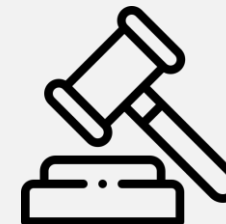
Lucas Vale | Henrique Monteiro Cristóvão



**Desenvolvimento  
da sociedade**



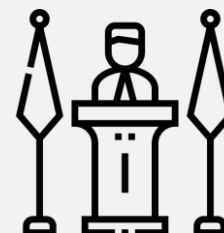
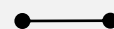
**Acúmulo de  
dados**



**Dados com  
relevância**



**Saúde das  
pessoas**



**Tomada de  
decisão**



CSV

## Preço de Medicamentos no Brasil – Consumidor

Os dados representam a lista de preços de Medicamentos , contemplando o preço Fábrica, ou preço fabricante (PF), que é o preço máximo que pode ser praticado pelas empresas produtoras ou importadoras do produto e pelas empresas distribuidoras. O PF indica o preço máximo permitido para venda a farmácias e drogarias e o Preço Máximo ao Consumidor (PMC) indica o preço teto de venda ao consumidor.

🕒 Data de criação: 31/05/2017 🕒 Última atualização: 31/05/2017 📄 Downloads: 33

[Download](#)[Preview](#)

# KDD

**KNOWLEDGE DISCOVERY IN DATABASE**

Fayyad et al. (1996)

# KDD

Fayyad et al. (1996)

**Seleção de dados**

**Pré-processamento**

**Transformação**

**Mineração de dados**

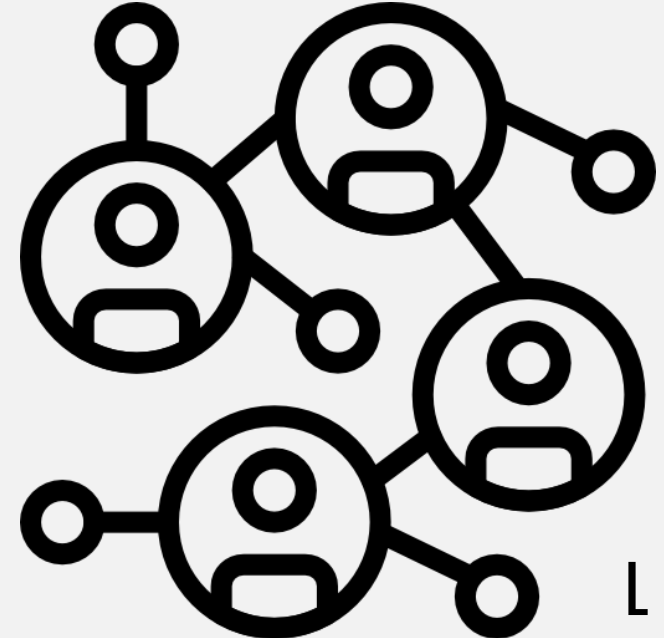
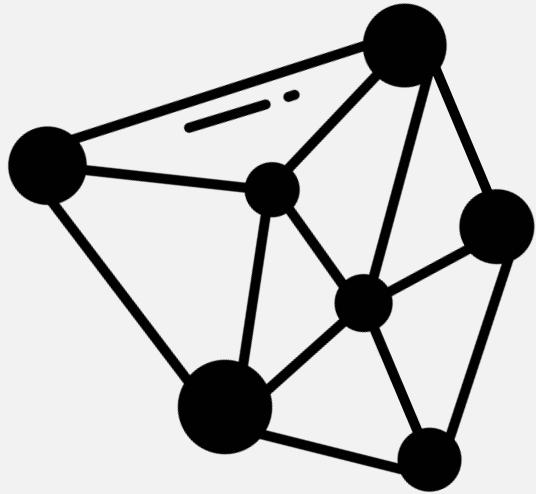
**Interpretação**

**Avaliação e implantação do conhecimento**

# Mineração de dados

Diversidade de metodologias

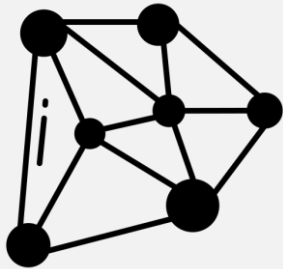
—  
**ANÁLISE DE  
REDES  
COMPLEXAS**



# KDD

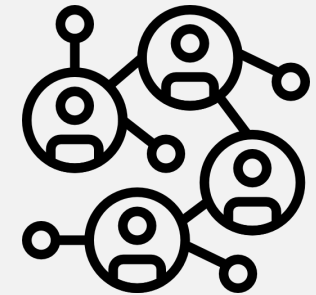
Fayyad et al. (1996)

# Mineração de dados

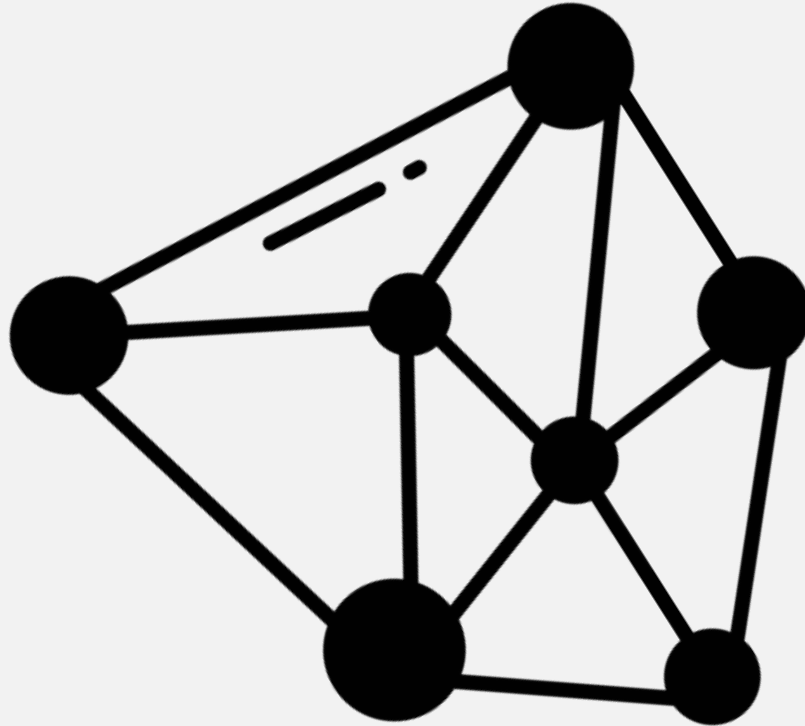


**ANÁLISE DE REDES  
COMPLEXAS**

**REDES SOCIAIS**  
Wasserman et al. (1994)



Métricas | Comunidades  
Projeções bipartidas | Inspeção visual



# REDE COMPLEXA

Estrutura

Topológica

Sofisticada

Barabási (2003)



**Imagina-se que o conjunto de dados escolhido possa, ao ser analisado, revelar padrões ou tendências não evidentes ou despercebidas que, em função da sua forte relevância social, podem trazer algum ganho para os indivíduos diretamente envolvidos com o preço de medicamentos no Brasil.**

**Imagina-se que o conjunto de dados escolhido possa, ao ser analisado, revelar **padrões ou tendências** não evidentes ou despercebidas que, em função da sua forte **relevância social**, podem trazer algum **ganho** para os indivíduos diretamente envolvidos com o preço de medicamentos no Brasil.**

**Investigar e revelar relações entre variáveis da base de dados abertos da ANVISA sobre preços de medicamentos.**

# Abordagem qualitativa & natureza aplicada.

# KDD

**KNOWLEDGE DISCOVERY IN DATABASE**

Fayyad et al. (1996)

# KDD

Fayyad et al. (1996)

**Seleção de dados**

**Pré-processamento**

**Transformação**

**Mineração de dados**

**Interpretação**

**Avaliação e implantação do conhecimento**

# Pré-processamento

## Remoção de ruídos

---

Remoção de colunas irrelevantes

Correção de erros de registro

camelCase

# KDD

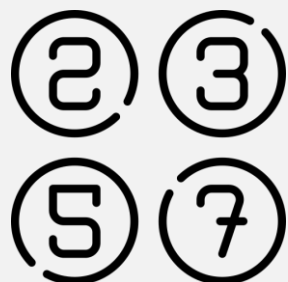
Fayyad et al. (1996)



OpenRefine

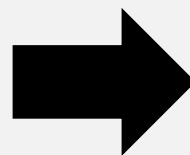
# Transformação

Transcrição



**Valores de Preço**

Quantitativa Contínua



**Categorias de Preço**

Qualitativa Ordinal

**KDD**

Fayyad et al. (1996)



**OpenRefine**



# Transformação

```
graph [
  directed 0

  node [ id {{jsonize(cells.classeTerapeutica.value)}}
  variavel "Classe Terapeutica" agrupamento "grupo A" ]

  node [ id {{jsonize(cells.faixaPrecoFabrica17.value)}}
  variavel "Faixa de Preço de Fabrica" agrupamento "grupo B" ]

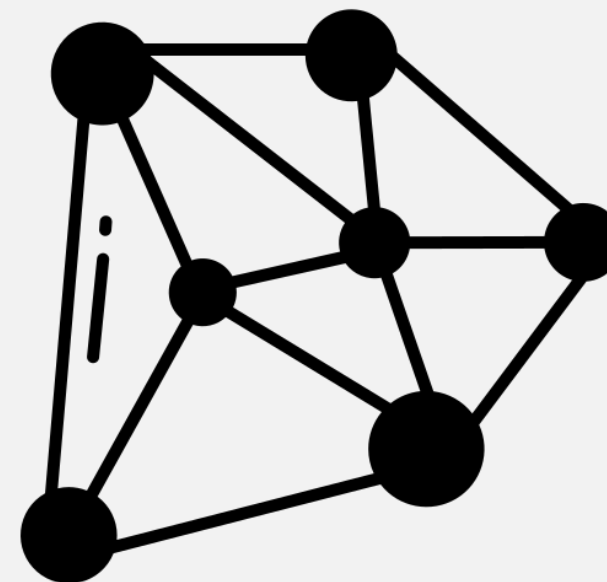
  node [ id {{jsonize(cells.regimeDePreco.value)}}
  variavel "Regime de Preço" agrupamento "grupo B" ]

  edge [ source {{jsonize(cells.classeTerapeutica.value)}}
  target {{jsonize(cells.faixaPrecoFabrica17.value)}}]

  edge [ source {{jsonize(cells.classeTerapeutica.value)}}
  target {{jsonize(cells.regimeDePreco.value)}}]

]
```

Mapeamento  
GML



# KDD

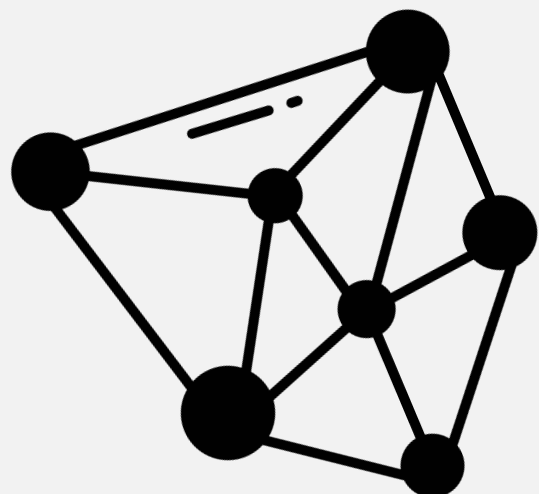
Fayyad et al. (1996)



OpenRefine

L | H

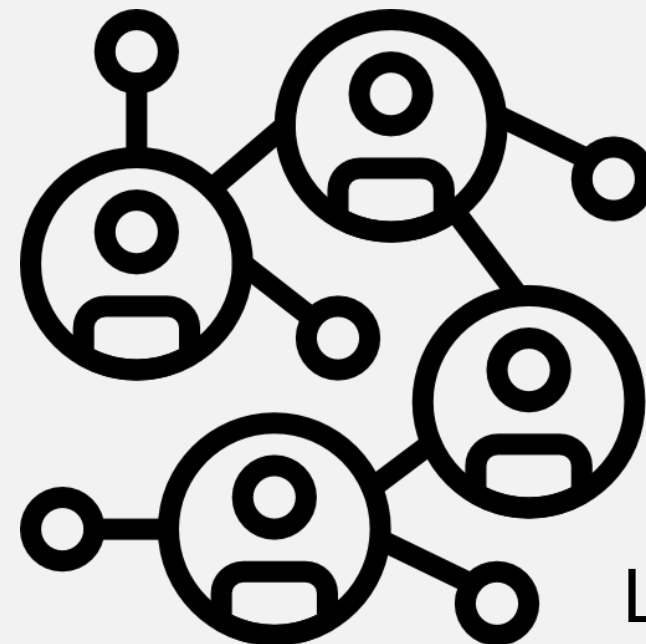
# Mineração de dados



**ANÁLISE DE REDES  
COMPLEXAS**

+

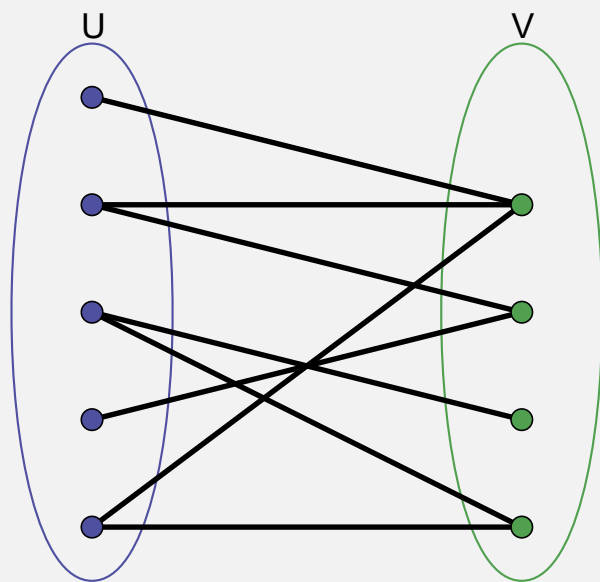
**CRIAÇÃO DE DASHBOARD**



**KDD**

Fayyad et al. (1996)

# Redes



Laboratório & Tarja

Classe Terapêutica & Faixa de Preço Fábrica

Classe terap., Faixa de Preço F. & Regime de Preço

---

Yifan Hu e Fruchterman Reingold

## KDD

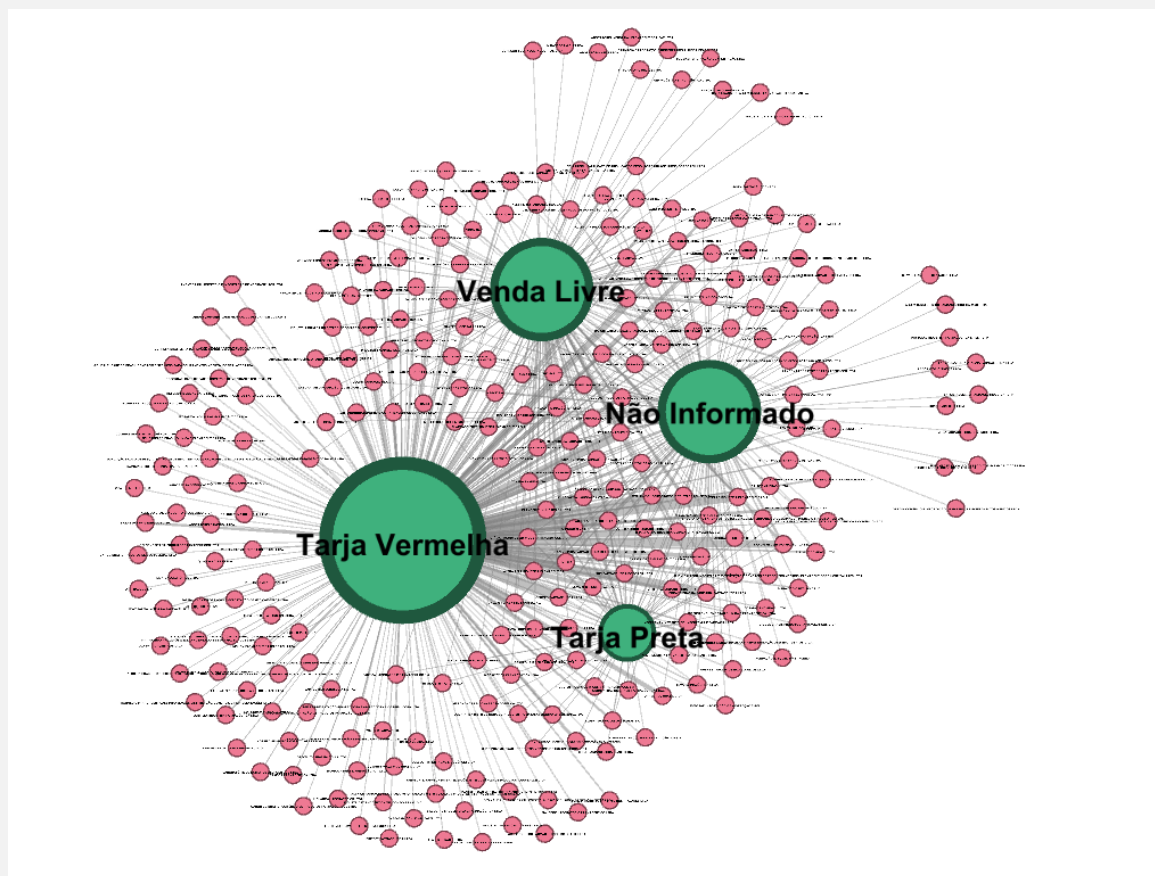
Fayyad et al. (1996)

Mineração de dados

Gephi

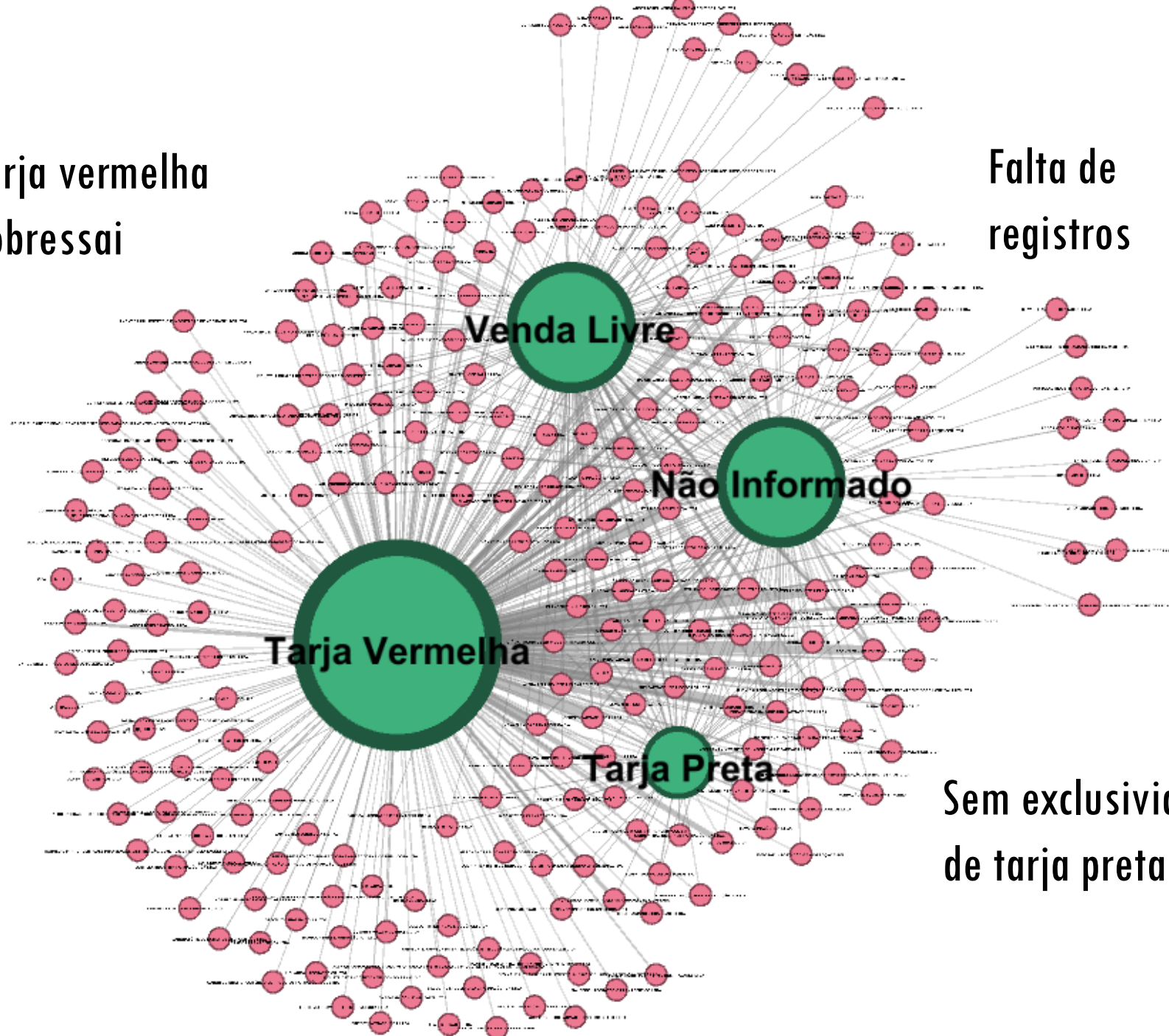
L | H

## Laboratório & Tarja



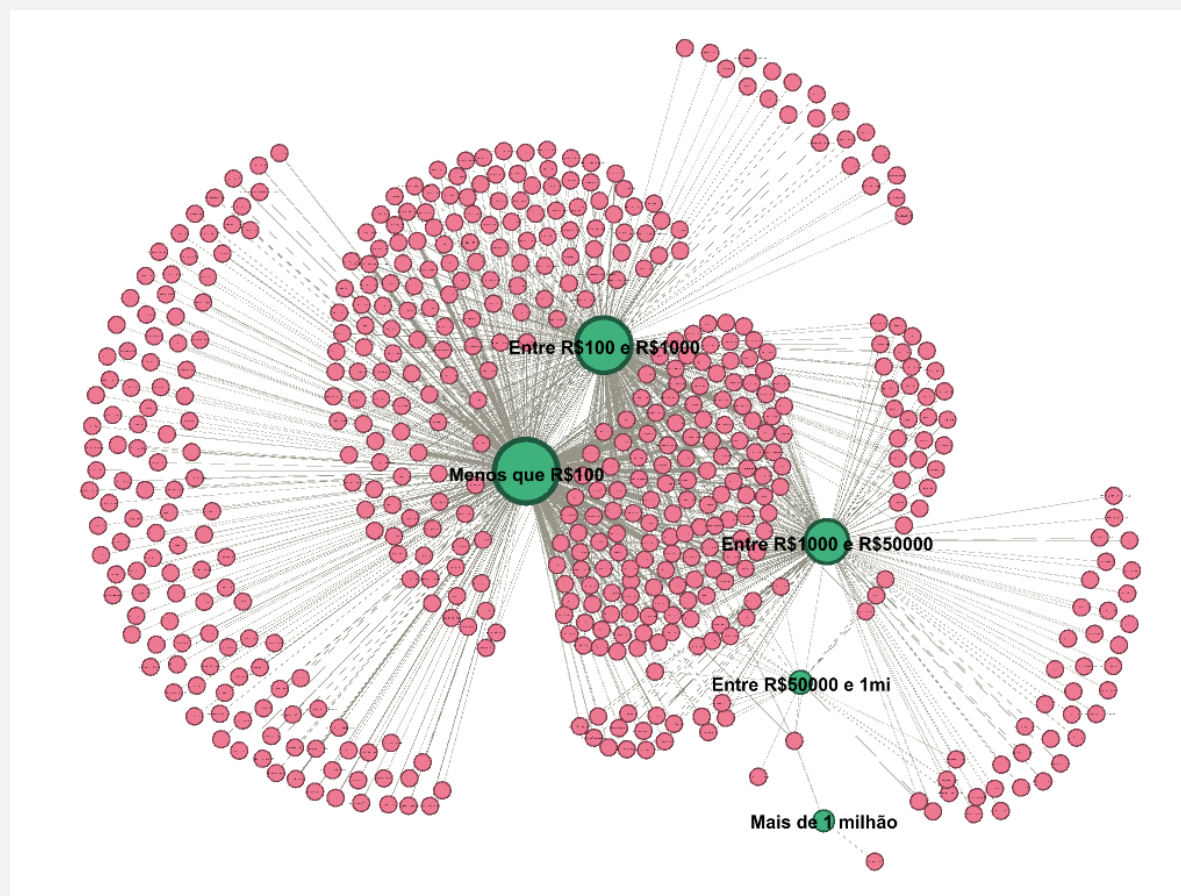
Tarja vermelha  
sobressai

Falta de  
registros

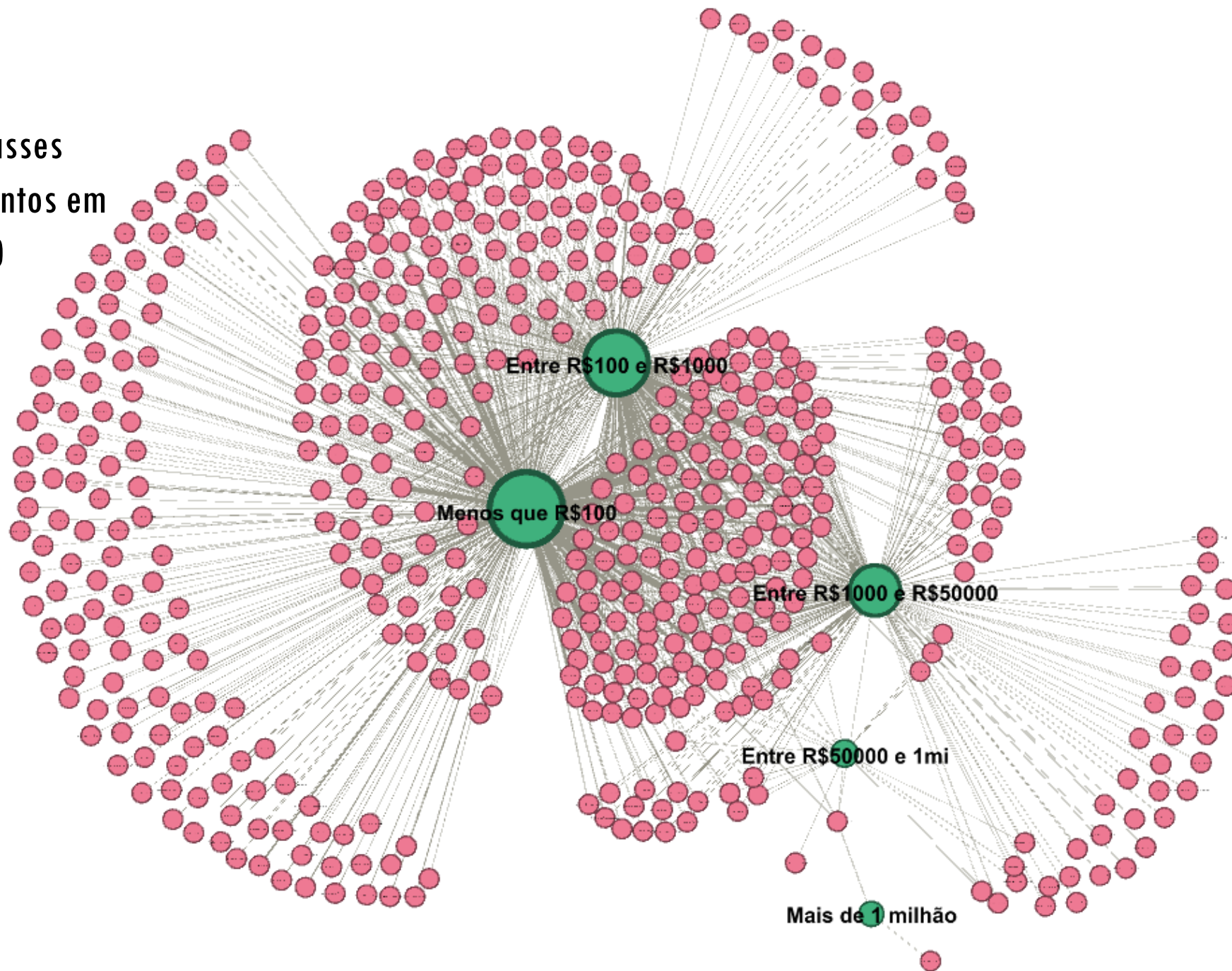


Sem exclusividade  
de tarja preta

## Classe Terapêutica & Faixa de Preço Fábrica

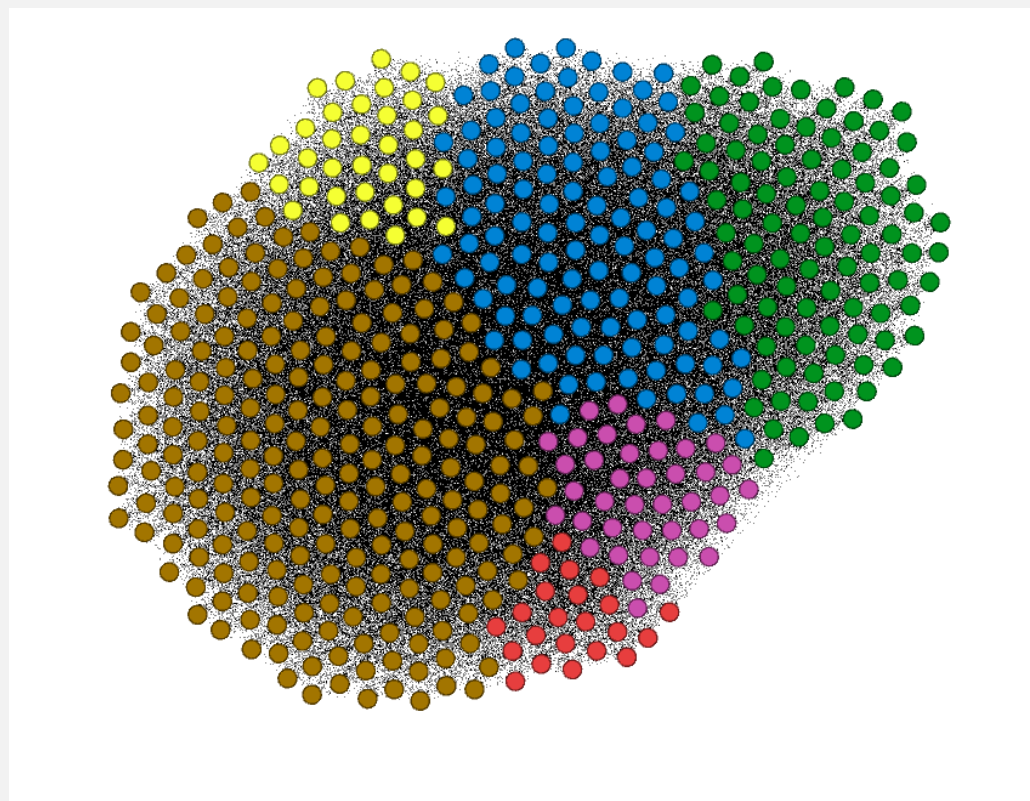


Maioria das classes  
com medicamentos em  
torno de R\$100

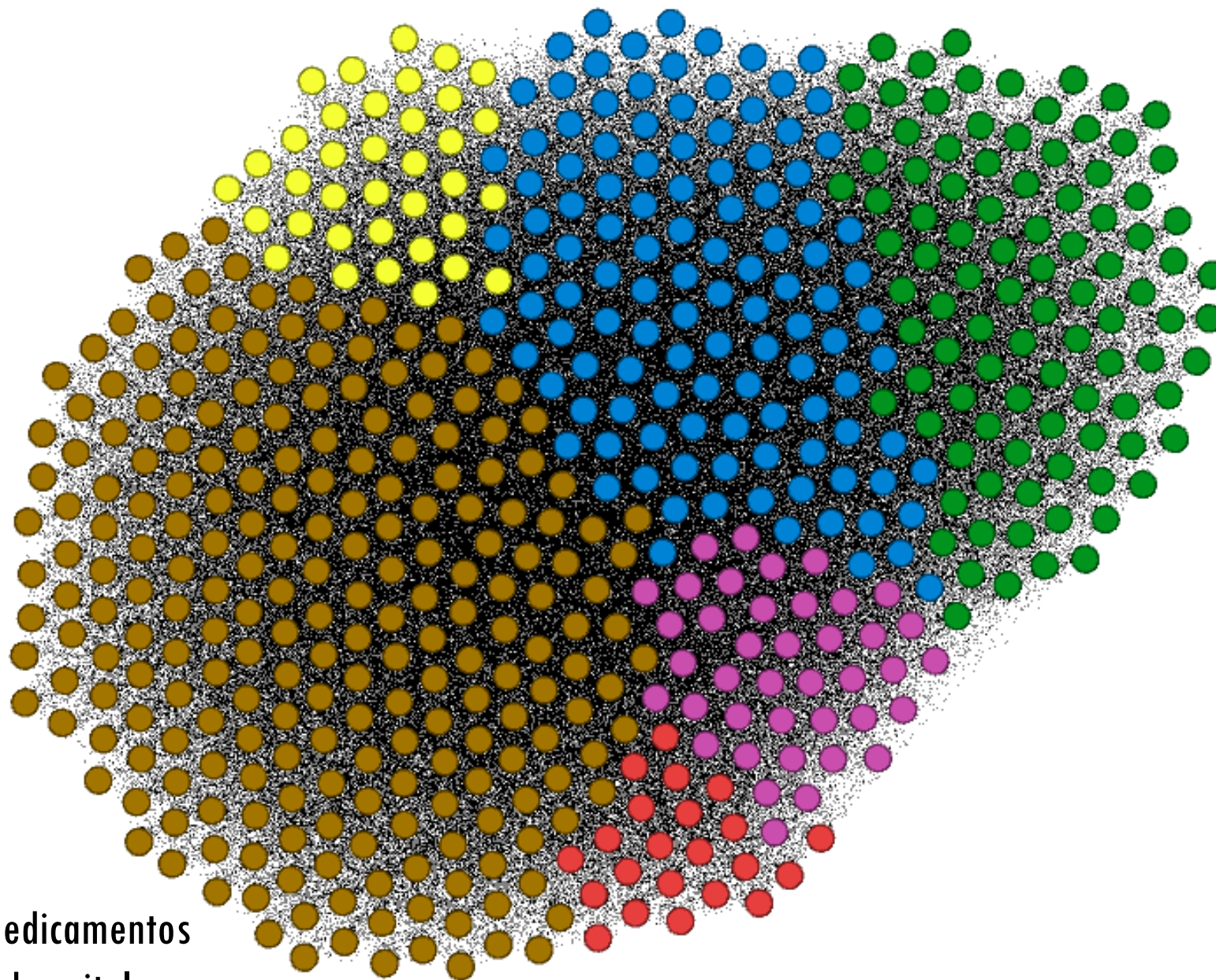


Duas classes com  
medicamentos  
acima de 1 mi

## Classe terap., Faixa de Preço F. & Regime de Preço





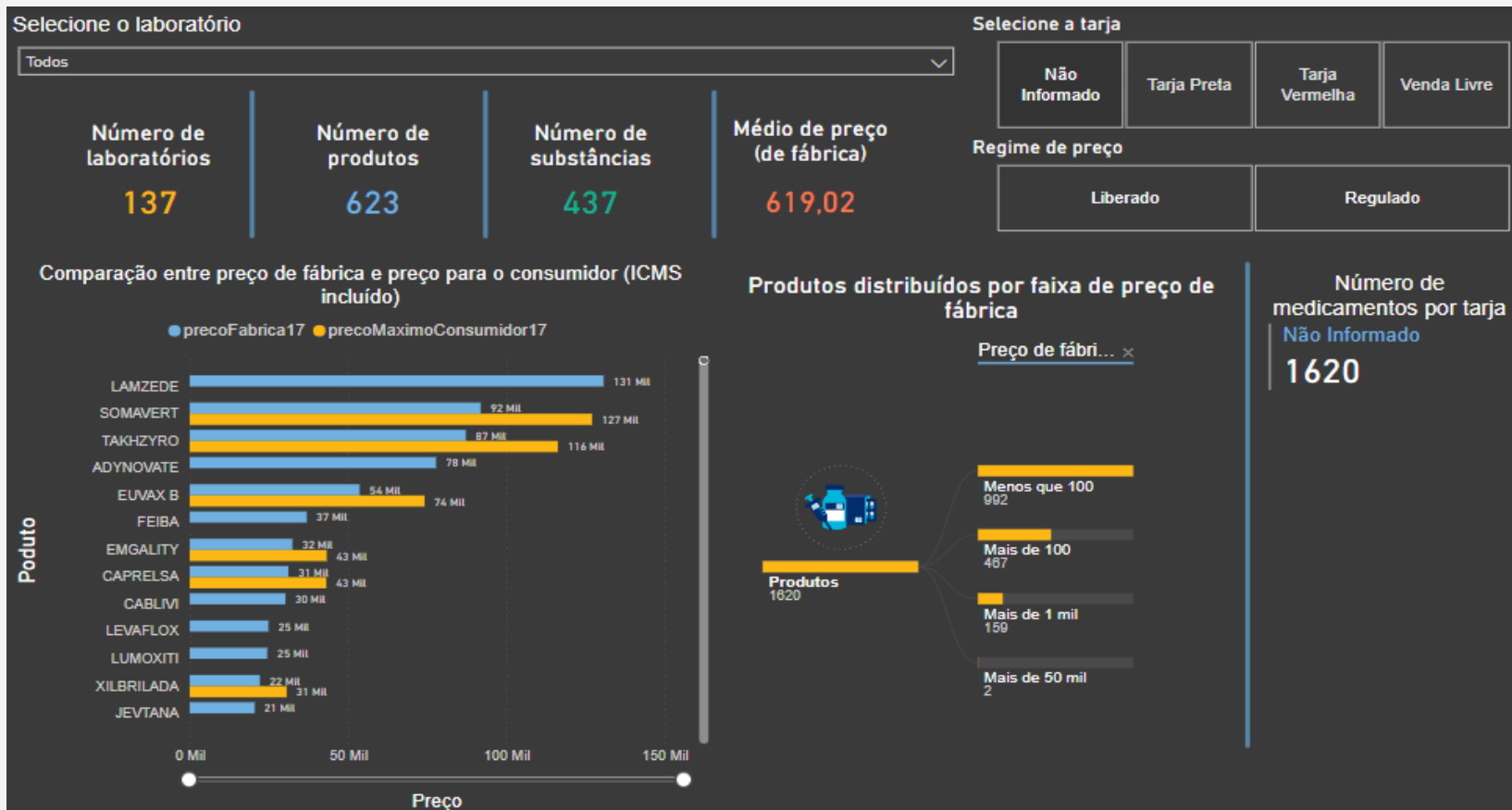


Proporção de  
medicamentos de  
venda livre 6x maior

8x mais medicamentos  
fitoterápicos (tipo de  
produto)

O dobro de medicamentos  
com restrição hospitalar

# Dashboard



Selecione o laboratório

Todos

Número de laboratórios

137

Número de produtos

623

Número de substâncias

437

Médio de preço (de fábrica)

619,02

Selecione a tarja

Não Informado

Tarja Preta

Tarja Vermelha

Venda Livre

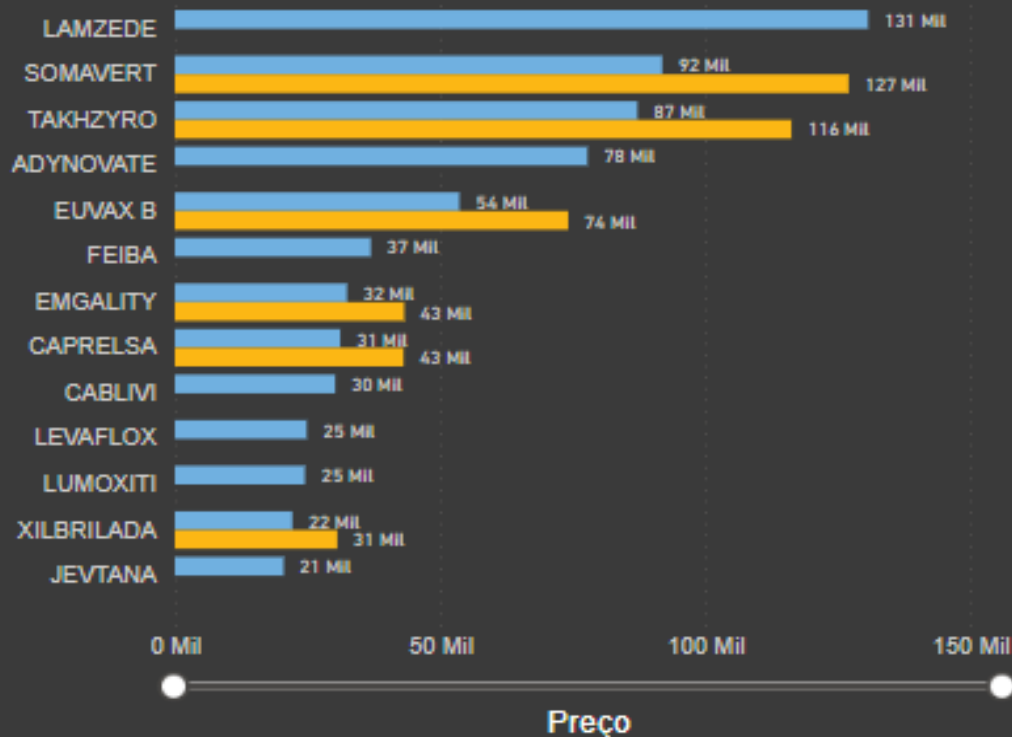
Regime de preço

Liberado

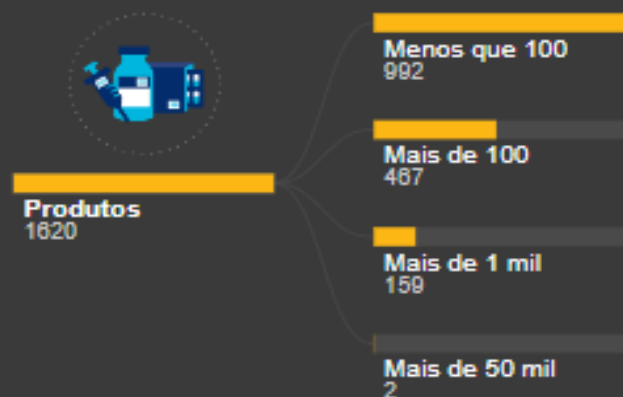
Regulado

Comparação entre preço de fábrica e preço para o consumidor (ICMS incluído)

● precoFabrica17 ● precoMaximoConsumidor17



Produtos distribuídos por faixa de preço de fábrica



Número de medicamentos por tarja

Não Informado

1620



Os medicamentos de

## ~~Tarja Vermelha~~

são os mais produzidos pela indústria farmacêutica brasileira



## Hipertensão

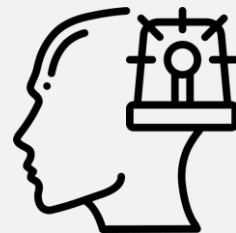
24% dos adultos

Alta mortalidade no Brasil



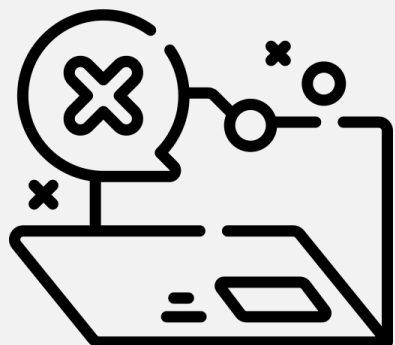
## Diabetes

7% da população



## Ansiedade

Pais mais ansioso do mundo (OMS)



## Muitos registros

incorretos ou ausentes



## Menor assertividade

nas métricas e predições



## Erros humanos

Na alimentação da base de dados



## Má gestão

de dados públicos é  
potencialmente prejudicial



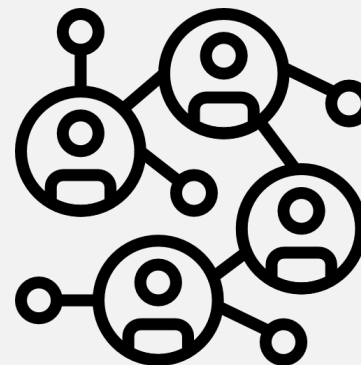
Apenas duas classes terap possuem medicamentos com valor acima de

# 1 Milhão de Reais

**M5X** - TODOS OS  
OUTROS FÁRMACOS  
COM AÇÃO MÚSCULO-  
ESQUELÉTICA

**S1X1** - OUTROS  
PRODUTOS  
OFTALMOLÓGICOS  
SISTÊMICOS

Novartis Biociências S.A.



Relações

# Intracluster

da rede monopartida precisam ser mais investigadas

## Considerações Finais

**Objetivos parcialmente alcançados:**  
algumas relações descobertas

---

**Continuação do esforço de trabalho**

WIDaT

2022

UFES

## Referências

BARABÁSI, Albert-László. **Linked: The new science of networks**. 2003.

FAYYAD, Usama; PIATETSKY-SHAPIRO, Gregory; SMYTH, Padhraic. The KDD process for extracting useful knowledge from volumes of data. **Communications of the ACM**, v. 39, n. 11, p. 27-34, 1996.

METZ, Jean et al. **Redes complexas: conceitos e aplicações**. 2007.

WASSERMAN, Stanley et al. **Social network analysis: Methods and applications**. 1994.