

# PROJETO FINAL

Equipe: Alexandre Matias, Giuliana  
Martins, Matheus Fonseca  
Grupo: Bando de Dados

Disciplina: Introdução a Banco de Dados  
Professor: Luiz Celso Gomes Junior



# O QUE FOI FEITO:

- Bússola Política de gêneros musicais
- Clusterização para criar grupos baseados nos gêneros musicais
- Recomendação de filmes para assistir baseado na similaridade entre usuários e filmes curtidos (e odiados!) em comum.
- Recomendação de amigos através de grafos.

# POR QUE?

- Foram pensados em ideias simulando uma rede social real:
  - ◆ Criação de grupos baseado em interesses comuns de várias pessoas
  - ◆ Recomendação de filmes usando a similaridade entre usuários (gostos parecidos)
  - ◆ Recomendação de amigos com base em “amigos de amigos”

# FERRAMENTAS

- Python como linguagem de programação.
- E demais bibliotecas:
  - ◆ igraph
  - ◆ Numpy
  - ◆ Matplotlib
  - ◆ Scikit-learn (kmeans)

# ESTRATÉGIAS UTILIZADAS

- Criação de clusters a partir usando kmeans.
- Criação de matriz de similaridades entre usuários.
- Criação de um grafo contendo as amizades.

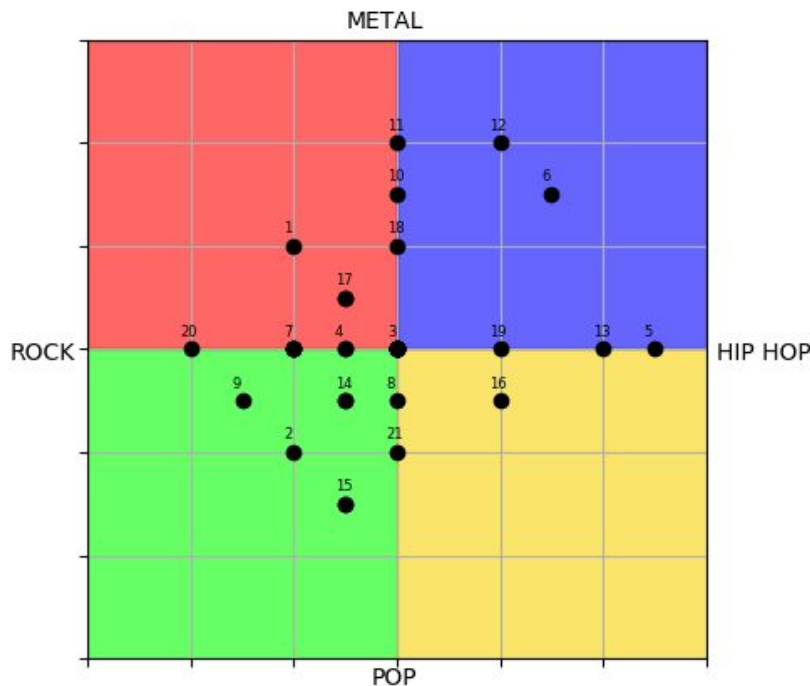
# BÚSSOLA POLÍTICA DE GÊNEROS MÚSICAIS

- Com inspiração nas “bússolas políticas” que ficaram virais nos últimos meses foi criada uma bússola de gêneros musicais;
- A bússola foi implementada levando em consideração os 4 gêneros musicais mais populares e as notas atribuídas por cada pessoa. Os gêneros são:
  - ◆ Heavy metal music
  - ◆ Pop music
  - ◆ Alternative rock
  - ◆ Hip hop music
- Apresentando de maneira mais geral os gêneros foram: Metal, Pop, Rock e Hip hop
- Foi decidido um semi-eixo para cada gênero, no caso Metal foi escolhido o semi-eixo positivo de y, Pop ficou com o semi-eixo negativo de y, Hip hop com o semi-eixo positivo de x e Rock com o semi-eixo negativo de x;

# BÚSSOLA POLÍTICA DE GÊNEROS MÚSICAIS

- Para fazer a bússola foi necessário atribuir uma coordenada no plano cartesiano para cada pessoa;
- As coordenadas foram feitas inicialmente todas em (0, 0) e eram atualizadas baseadas nas notas atribuídas a cada gênero. A atualização das coordenadas foi feita da seguinte maneira:
  - ◆ Nota 5 em algum gênero, soma 2 no sentido atribuído ao gênero;
  - ◆ Nota 4 em algum gênero, soma 1 no sentido atribuído ao gênero;
  - ◆ Nota 3 em algum gênero é considerada neutra por isso não altera a coordenada;
  - ◆ Nota 2 em algum gênero, soma 1 no sentido contrário ao gênero;
  - ◆ Nota 1 em algum gênero, soma 2 no sentido contrário ao gênero;
- Após a implementação chegou-se a seguinte bússola:

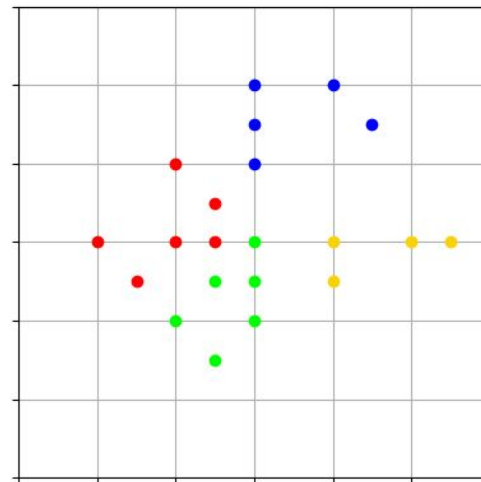
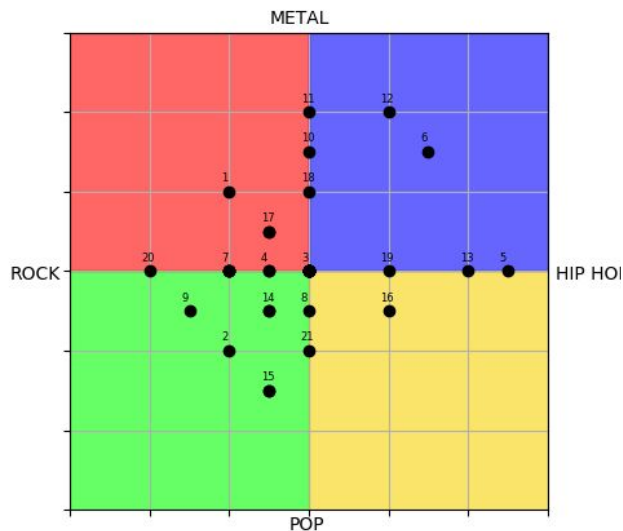
# BÚSSOLA POLÍTICA DE GÊNEROS MÚSICAIS



- É fácil notar que não há os 44 números que seriam de cada uma das pessoas da turma;
- Isso se deve pelo fato de haver uma grande quantidade de pessoas nos mesmos pontos;
- (-2, 2) - DI1802alexandrematias - 1
- (-2, -2) - DI1802annabucharles - 2
- (0, 0) - DI1802cassiomorales - 3
- (-1, 0) - DI1802damonoliniski - 4
- (5, 0) - DI1802danielvieira - 5
- (3, 3) - DI1802francoisrueckert - 6
- (-2, 0) - DI1802gabrielguerrero - 7
- (0, -1) - DI1802heronfernandez - 8
- (-3, -1) - DI1802joaojorge - 9
- (0, 3) - DI1802lincolnbatista - 10
- (0, 4) - DI1802lucasrech - 11
- (2, 4) - DI1802lucasribeiro - 12
- (4, 0) - DI1802lucastakahashi - 13
- (-1, -1) - DI1802matheuscosta - 14
- (-1, -3) - matheusgoncalves - 15
- (2, -1) - DI1802nortonjunior - 16
- (-1, 1) - DI1802tasipasin - 17
- (0, 2) - DI1802tiagoplakitca - 18
- (2, 0) - DI1802victormenuzzo - 19
- (-4, 0) - DI1802vivianebonfim - 20
- (0, -2) - DI1802giulianasilva - 21



# CLUSTERIZAÇÃO PARA CRIAR GRUPOS DE AMIGOS BASEADOS NOS GÊNEROS MÚSICAIS



# CLUSTERIZAÇÃO PARA CRIAR GRUPOS DE AMIGOS BASEADOS NOS GÊNEROS MÚSICAIS

Grupo 1:

['Francois Bruno Rueckert', 'Lincoln Batista', 'Lucas Coradin Rech', 'Lucas Felipe Ribeiro', 'Tiago Colli Plakitca']

Grupo 2:

['Alexandre Herrero Matias', 'Damon Nicholas Lopes Oliniski', 'Gabriel Carrico Guerrero', 'João Pedro Gonçalves Barreiro Passos Jorge', 'Tasi Guilhen Pasin', 'Viviane Lima Bonfim']

Grupo 3:

['Daniel Eduardo Vieira', 'Lucas Shoiti Takahashi', 'Norton Wianey Pezzi Cortabitarte Junior', 'Victor Antonio Menuzzo']

Grupo 4:

['Anna Carolina Flumignan Bucharles', 'Cassio Henrique Garcia Morales', 'Heron Gomes Fernandez', 'Matheus Bigarelli Dantas Da Costa', 'Matheus Ernesto Silva Gonçalves', 'Giuliana Linda Maravilhosa']

## **RECOMENDAÇÃO DE FILMES PARA ASSISTIR BASEADO NA SIMILARIDADE ENTRE USUÁRIOS E FILMES CURTIDOS (E ODIADOS!)**

- Distância de Jaccard para aproximar dois usuários utilizando filmes que ambos gostaram ou odiaram.
  - ◆ Armazenamento em matriz usuario x usuario.
- Recomendação é baseada no usuário que é o mais similar a você
  - ◆ Claro, levando em consideração que a similaridade entre você e você mesmo é 100% :)
- Filmes recomendados são filmes que você não assistiu.
  - ◆ E apenas filmes que a pessoa mais próxima de você gostou.
  - ◆ Nada de recomendar filmes odiados!

# EXEMPLOS

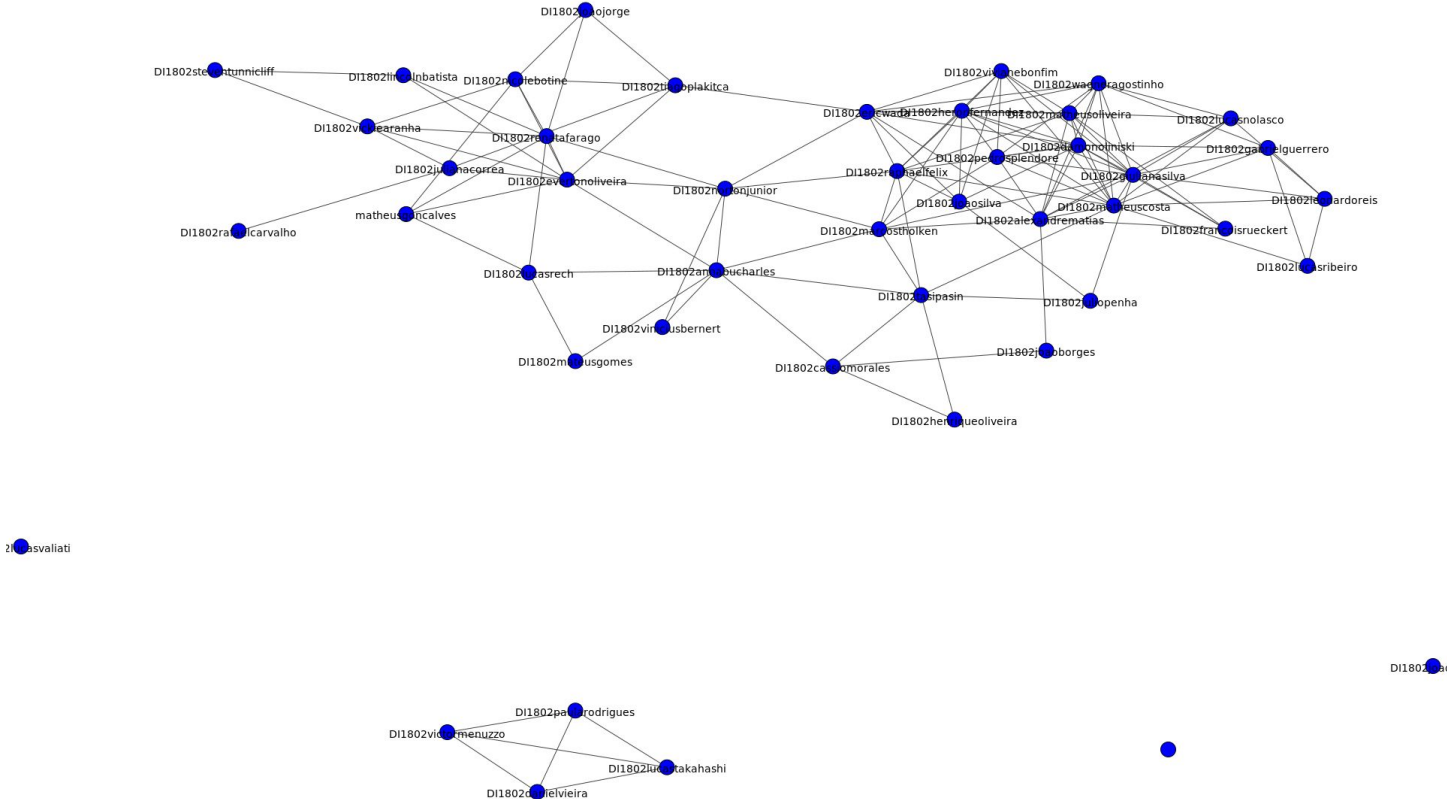
```
Bem vindo ao menu.  
Digite seu número de usuário e lhe recomendaremos alguns filmes!  
4  
Aqui está os filmes recomendados, [('Damon Nicholas Lopes Oliniski',)]:  
["[('Scarface',)]", "[('The Matrix',)]", "[('V for Vendetta',)]", "[('WALL·E',)]"]  
Obrigado por usar nosso programa! Até logo!
```

```
Bem vindo ao menu.  
Digite seu número de usuário e lhe recomendaremos alguns filmes!  
8  
Me desculpe, [('Francois Bruno Rueckert',)], mas não há filmes para te  
recomendar no momento. Assista mais filmes!
```

# IMPLEMENTAÇÃO

- Views de associação
  - ◆ Associar cada Filme com um número arbitrário
  - ◆ Associar cada Usuário com número arbitrário
- Motivos: Facilitar busca e referência ao longo do programa.
- Criação de Matrizes
  - ◆ Usuario x Filme para armazenar as notas dos usuários por filme.
  - ◆ Usuario x Usuario para armazenar a similaridade entre usuários, utilizando Distância de Jaccard
- Busca pelo usuário mais similar, dado um usuário de interesse.
  - ◆ Através de um menu interativo.
- Agrupamento dos filmes recomendados em listas

# RECOMENDAÇÃO DE AMIGOS COM GRAFOS



# RECOMENDAÇÃO DE AMIGOS COM GRAFOS

- Biblioteca utilizada: python-igraph
- O grafo foi feito com base na tabela “Conhece” normalizada
- Recomenda amigos baseado nas amizades de outras pessoas

```
Digite a id da pessoa que deseja sugerir amigos: DI1802alexandrematias
São amigos de DI1802damonoliniski :
DI1802gabrielguerrero
DI1802heronfernandez
DI1802joaosilva
DI1802raphaelfelix
São amigos de DI1802ericwada :
DI1802heronfernandez
DI1802joaosilva
DI1802nortonjunior
DI1802raphaelfelix
DI1802ttiagoplakitca
DI1802vivianebonfim
São amigos de DI1802francoisrueckert :
São amigos de DI1802giulianasilva :
1234
DI1802gabrielguerrero
DI1802heronfernandez
DI1802juliopenha
DI1802leonardoreis
DI1802vivianebonfim
```

# RELAÇÃO COM A DISCIPLINA

- Consultas e criação de views foram essenciais para o desenvolvimento do trabalho
- Foram analisados dados de todos os tipos para obter as relações obtidas
- Informações implícitas foram extraídas com a análise



# CONCLUSÕES

O trabalho e a disciplina de forma geral nos instigou a aprender mais sobre banco de dados e análise de dados.

Uma parte interessante foram os diversos modos descobertos de tratamento e visualização de dados. O trabalho também nos instigou a buscar maior conhecimento em Python, uma linguagem nunca utilizada pelos integrantes do grupo.