# BIG DATA







# Introdução ao Big Data

Tema da Aula: Twitter Streaming API

Prof.: **Dino Magri** 

Data: 05 de Agosto de 2016

#### Coordenação:

Prof. Dr. Adolpho Walter Pimazzi Canton

Profa. Dra. Alessandra de Ávila Montini



#### Coordenação:

Prof. Dr. Adolpho Walter Pimazzi Canton

Profa. Dra. Alessandra de Ávila Montini

#### Contatos:

- E-mail: <u>professor.dinomagri@gmail.com</u>
- Twitter: <a href="https://twitter.com/prof\_dinomagri">https://twitter.com/prof\_dinomagri</a>
- LinkedIn: <a href="http://www.linkedin.com/in/dinomagri">http://www.linkedin.com/in/dinomagri</a>
- Site: <a href="http://www.dinomagri.com">http://www.dinomagri.com</a>

### Currículo

- (2014-Presente) Professor no curso de Extensão, Pós e MBA na Fundação Instituto de Administração (FIA) – <a href="https://www.fia.com.br">www.fia.com.br</a>
- (2013-Presente) Pesquisa e Desenvolvimento no Laboratório de Arquitetura e Redes de Computadores (LARC) na Universidade de São Paulo – <a href="www.larc.usp.br">www.larc.usp.br</a>
- (2013) Professor no MBA em Desenvolvimento de Inovações Tecnológicas para WEB na IMED
   Passo Fundo RS www.imed.edu.br
- (2012) Bacharel em Ciência da Computação pela Universidade do Estado de Santa Catarina
   (UDESC) <a href="https://www.cct.udesc.br">www.cct.udesc.br</a>
- (2009/2010) Pesquisador e Desenvolvedor no Centro de Computação Gráfica Guimarães –
   Portugal <a href="https://www.ccg.pt">www.ccg.pt</a>
- Lattes: http://lattes.cnpq.br/5673884504184733



#### Material das aulas

- Material das aulas:
  - <a href="http://urls.dinomagri.com/posmba-turma2/">http://urls.dinomagri.com/posmba-turma2/</a>
  - Senha: fia2016
- Faça o download do arquivo 2016-08-05-py-aula5.zip
- Salve na Área de Trabalho (Desktop)
- Depois que finalizar o download, acesse a pasta Área de trabalho e descompacte o arquivo 2016-08-05-pyaula5.zip.

### Conteúdo da Aula

- Objetivo
- Facebook API
- Exercícios



### Conteúdo da Aula

- Objetivo
- Facebook API
- Exercícios



# Objetivo

 Objetivo dessa aula é introduzir os conceitos básicos da Graph API do Facebook e como podemos acessá-la utilizando Python.

### Conteúdo da Aula

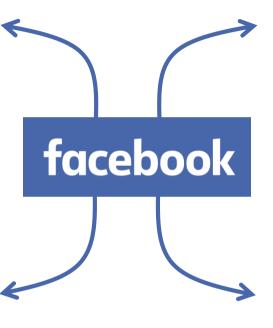
- Objetivo
- Facebook API
- Exercícios



#### Facebook

- Facebook é uma rede social fundada em Fevereiro de 2004 por Mark Zuckerberg, Dustin Moskovitz, Chris Hughes e Eduardo Saverin.
- Nessa rede social podemos se conectar com amigos, familiares, colegas, seguir empresas, compartilhar informações, fotos, vídeos, realizar chamadas, entre outros.
- A primeira API do Facebook foi lançada em Agosto de 2006.

Permite conectar pessoas ao redor do mundo!



Criar páginas, compartilhar imagens, mensagens, vídeos.

Permite compartilhar interesses, trocar mensagens públicas e privadas

Controle de privacidade: é possível controlar quem pode acessar sua página



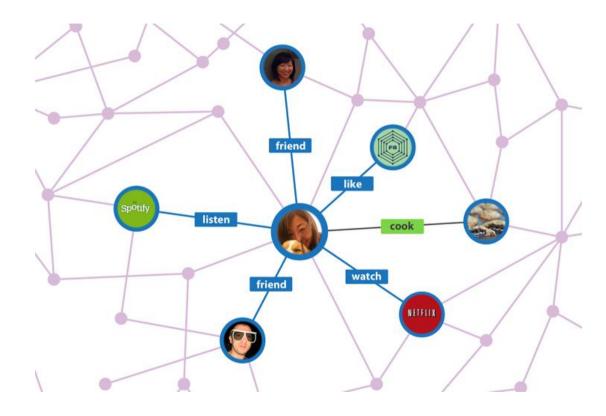
#### Facebook

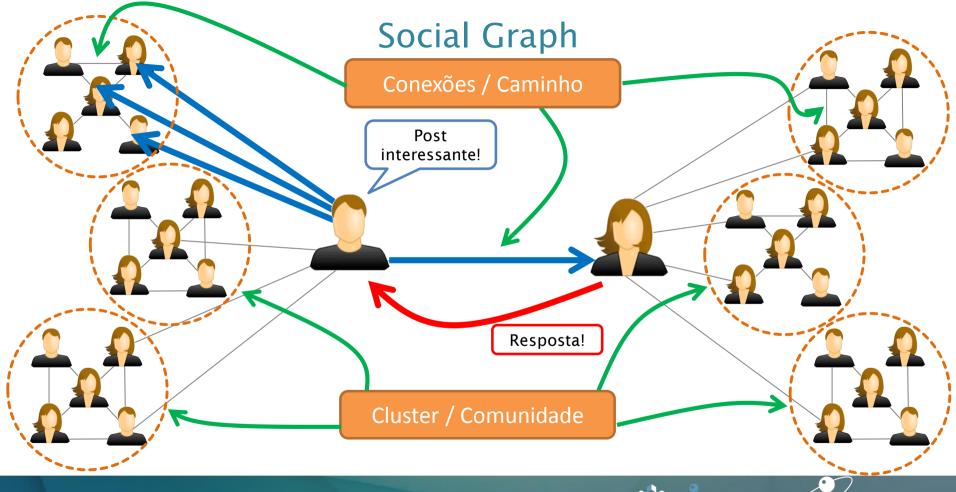
Quantidade de usuários ativos no Facebook ao longo dos anos



- Desde 2012, o Facebook substituiu a API REST pela Graph API.
- A Graph API recebeu esse nome com base na ideia de um <u>Social</u>
   <u>Graph</u>, que nada mais é do que uma representação das informações formada por:
  - Nós basicamente "coisas" como um usuário, uma foto, uma página, um comentário.
  - Bordas as conexões entre essas "coisas", como fotos de uma página ou os comentários de uma Foto.
  - Campos Informações sobre essas "coisas", como o aniversário de uma pessoa ou o nome de uma Página.

Basicamente uma
 Social Graph
 descreve um mapa
 de usuários, suas
 conexões e objetos.





- A Graph API do Facebook é a chave para Social Graph
- Acesso via URL HTTP(S) para qualquer objeto na Social Graph.
- Todos os dados disponíveis são retornados como objetos JSON.
- Acesso aos objetos:
  - Usuários, Páginas, Eventos, Grupos, Aplicações, Mensagens de
     Status, Fotos, Álbuns, Imagens do perfil, vídeos, notas e check-in.

- Os módulos que iremos utilizar são:
  - facebook-sdk
  - requests

- Para instalar, abra o terminal/CMD e digite:
  - pip install facebook-sdk
  - pip install requests

- Para acessar a Graph API do Facebook, existem duas possibilidades:
  - Criar um aplicativo
  - Utilizar alguma versão existente da Graph API
- Iremos utiliza-las para realizar algumas ações, como escrever ou recuperar informações do Facebook.

# O que iremos fazer?

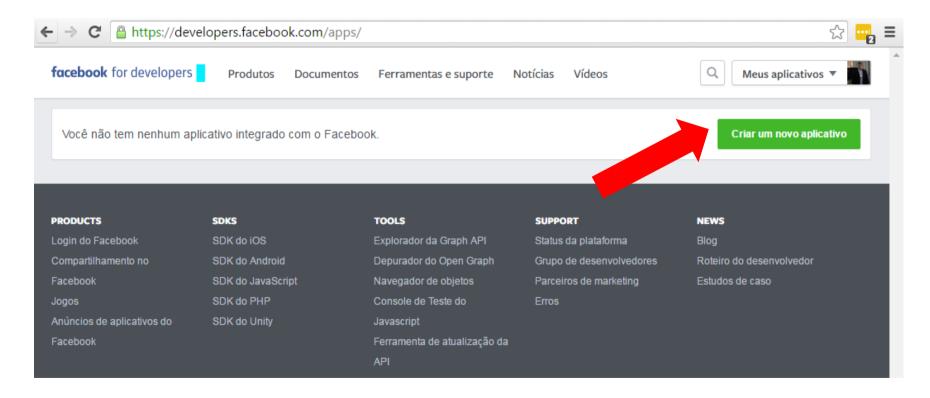
- 1. Realizar uma postagem simples
- 2. Postar links, descrição e imagem
- 3. Comentar e curtir uma publicação realizada
- 4. Postar em uma página específica
- 5. Postar uma foto
- 6. Recuperar posts



- Ter um conta no Facebook
- Aceitar o termos de desenvolvimento (Passo 1) e ter acesso ao ambiente (<a href="https://developers.facebook.com/">https://developers.facebook.com/</a>)
- Criar um aplicativo Facebook (Passo2a) ou utilizar a própria Graph API (Passo2b) para garantir acesso a informação
- Pegar o token de acesso com as devidas permissões para o aplicativo que foi criado (Passo3).

- Passo 1: Aceitar os termos de desenvolvimento e ter acesso ao ambiente
  - Realize o login no Facebook
  - Acesse o endereço: <a href="https://developers.facebook.com/">https://developers.facebook.com/</a>

- Passo 2a: Criar um aplicativo Facebook para garantir o acesso a Graph API
  - Realize o login no Facebook
  - Acesse o endereço: https://developers.facebook.com/apps/







# Add a New App

Select a platform to get started









If you're developing on another platform or want to skip this step for now, use the basic setup.



#### Crie um novo ID do Aplicativo

Comece integrando o Facebook ao seu aplicativo ou site

#### Nome de exibição

O nome do seu aplicativo ou site

#### Email de contato

Usado para comunicados importantes sobre o seu aplicativo

#### Categoria

Escolha uma categoria v

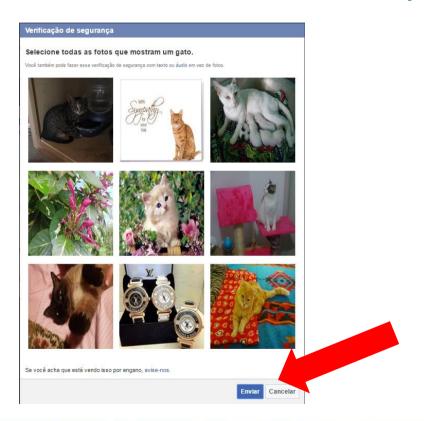


Ao prosseguir, você concorda com as Políticas da Plataforma do Facebook

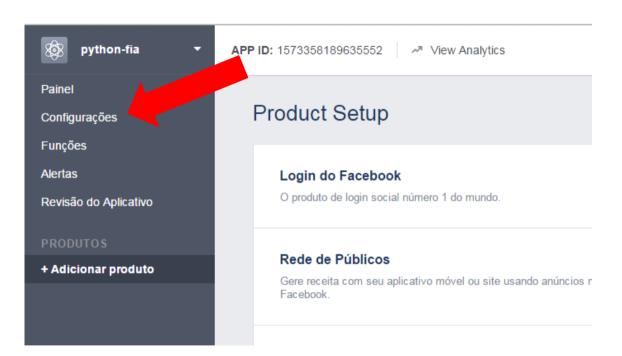
Cancelar

Crie um ID do Aplicativo





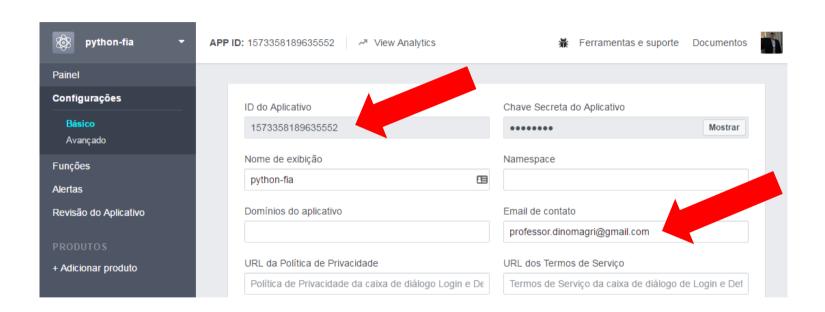
Após clicar em Criar um ID do Aplicativo, pode ser necessário realizar alguma checagem de segurança. Como por exemplo selecionar as imagens que contêm determinados animais (imagem ao lado).



Após será redirecionado para a página da Aplicação criada. No caso, o app chamado python-fia.

Clique em **Configurações**.





Certifique-se que o e-mail esteja preenchido.

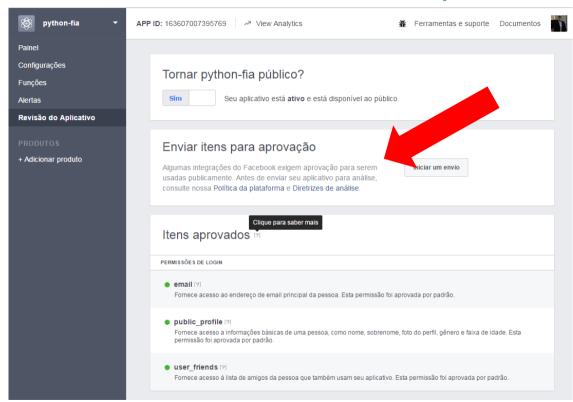












Note que quando torna público um aplicativo, automaticamente teremos os seguintes itens aprovados:

- email
- public\_profile
- user friends



- Passo 2b: Outra forma de interagir com a Graph API é utilizar a Graph API Explorer.
  - Realize o login no Facebook
  - Acesse o endereço: https://developers.facebook.com/tools/explorer

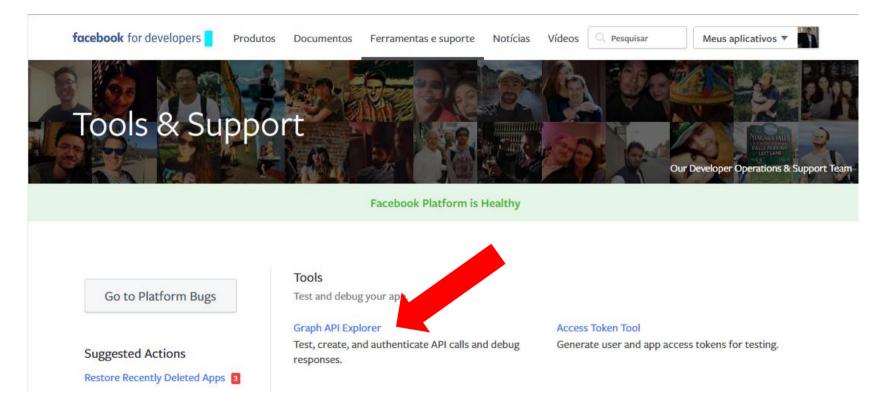
- Passo 3: Criar os tokens de acesso para o aplicativo criado.
  - Realize o login no Facebook
  - Acesse: <a href="https://developers.facebook.com/tools/explorer">https://developers.facebook.com/tools/explorer</a>
  - Clique em Graph API Explorer e selecione a aplicação criada.
  - Também é possível ir pelo menu da aplicação, clicando em
     Ferramentas e Suporte -> Graph API Explorer.

#### Passo 3: Criar tokens de acesso





#### Passo 3: Criar tokens de acesso



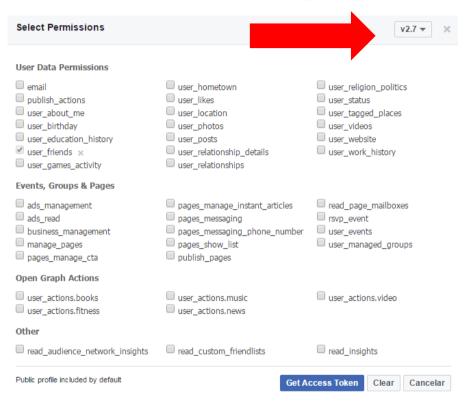


35

#### Passo 3: Criar tokens de acesso



#### Passo 3: Criar tokens de acesso



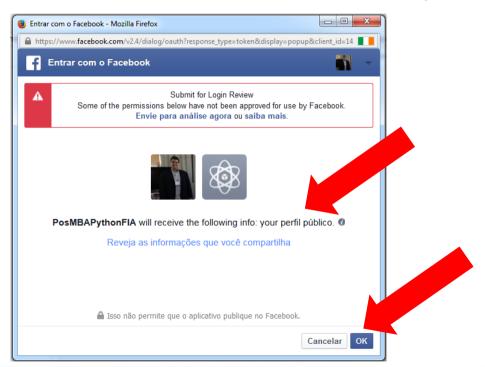
Para gerar o token de acesso precisamos definir quais dados queremos ter acesso.

Para os desafios que iremos resolver, precisamos habilitar:

- publish\_pages
- publish\_actions
- user friends
- user\_posts
- user status
- manage\_pages

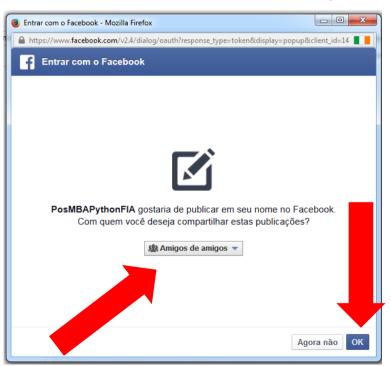


Passo 3: Criar os tokens de acesso para o aplicativo criado ou Graph API.



Toda vez que selecionarmos os elementos que queremos compartilhar, temos que dar permissão de acesso para a aplicação que criamos.

Passo 3: Criar os tokens de acesso para o aplicativo criado.



- É importante verificar com quem deseja compartilhar as informações.
- Também será necessário permitir que a publicação feita pelo aplicativo seja pública ou Amigos de amigos.



- Copie o Token de acesso para utilizarmos em nossos códigos!
- Caso seja necessário recuperar alguma outra informação e o aplicativo não tenha as permissões necessárias, temos que gerar novamente o Token de acesso.



- A biblioteca facebook-sdk foi projetada para ter acesso a Graph API do Facebook.
- É possível autenticar-se, recuperar conexões, criar, atualizar e/ou deletar objetos.
- Documentação da Biblioteca:
  - https://facebook-sdk.readthedocs.org/en/latest/

- 1. Realizar uma postagem simples
- 2. Realizar uma postagem com links, descrição e imagem
- 3. Comentar e curtir uma publicação realizada
- 4. Postar em uma página específica
- 5. Postar uma foto
- 6. Recuperar posts



### 1. Realizar uma postagem simples

- 2. Realizar uma postagem com links, descrição e imagem
- 3. Comentar e curtir uma publicação realizada
- 4. Postar em uma página específica
- 5. Postar uma foto
- 6. Recuperar posts

# Realizar uma postagem simples

- 1. Temos que importar o módulo
- 2. Salvar o token de acesso
- 3. Definir o acesso através da classe GraphAPI
- 4. Utilizar o método put\_wall\_post

Abra o notebook "aula6-parte1-post.ipynb"



- 1. Realizar uma postagem simples
- 2. Realizar uma postagem com links, descrição e imagem
- 3. Comentar e curtir uma publicação realizada
- 4. Postar em uma página específica
- 5. Postar uma foto
- 6. Recuperar posts



# Realizar uma postagem completa, comentar e curtir!

- 1. Temos que importar o módulo
- 2. Salvar o token de acesso.
- 3. Definir o acesso através da classe GraphAPI
- 4. Definir o nome, link, descrição e URL da imagem
- 5. Utilizar os métodos put\_wall\_post, put\_comment, put\_like
- 6. Deletar um objeto

Abra o notebook "aula6-parte2comentario.ipynb"



- 1. Realizar uma postagem simples
- 2. Realizar uma postagem com links, descrição e imagem
- 3. Comentar e curtir uma publicação realizada
- 4. Postar em uma página específica
- 5. Postar uma foto
- 6. Recuperar posts



# Postar em uma página específica

 Iremos realizar uma postagem na página da disciplina no Facebook:

https://www.facebook.com/aulaspythonfia

- 1. Recuperar o ID da página
- 2. Recuperar o token de acesso da página
- 3. Definir o acesso através da classe GraphAPI
- 4. Realizar a publicação

Abra o notebook "aula6-parte3-pagina.ipynb"

- 1. Realizar uma postagem simples
- 2. Realizar uma postagem com links, descrição e imagem
- 3. Comentar e curtir uma publicação realizada
- 4. Postar em uma página específica
- 5. Postar uma foto
- 6. Recuperar posts



#### Postar uma foto

- 1. Temos que importar o módulo
- 2. Salvar o token de acesso
- 3. Carregar a foto
- 4. Utilizar o método put\_photo
- 5. Criar um álbum e adicionar uma foto

Abra o notebook "aula6-parte4-foto.ipynb"



- 1. Realizar uma postagem simples
- 2. Realizar uma postagem com links, descrição e imagem
- 3. Comentar e curtir uma publicação realizada
- 4. Postar em uma página específica
- 5. Postar uma foto
- 6. Recuperar posts



- Até agora, utilizamos uma aplicação específica criada para realizarmos ações de escrita no Facebook (comentar, curtir, postar, entre outros).
- Porém, o Facebook permite explorar a Graph API, sem ter a necessidade de criar aplicações.
- A Graph API Explorer é uma interface simples para autenticação, para realizar requisições para API, e verificar as saídas.

- Não é necessário criar uma aplicação, apenas o token de acesso.
- Vamos utiliza-la para recuperar informações.
- Acesse: <a href="https://developers.facebook.com/tools/explorer">https://developers.facebook.com/tools/explorer</a>

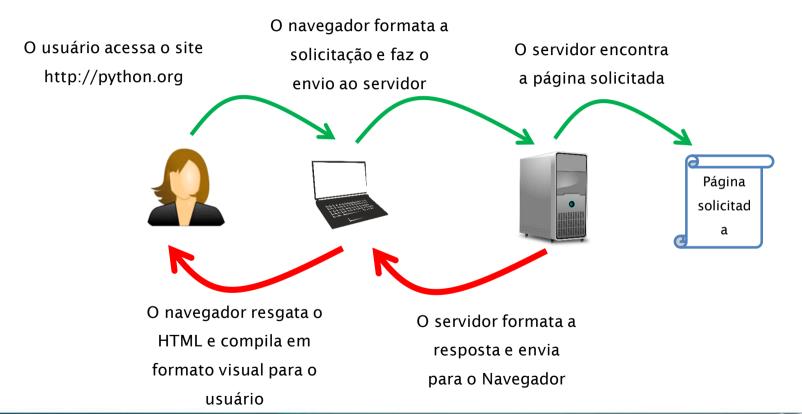
- Podemos utilizar a versão 2.3 para recuperar:
  - /{user-id}/statuses Retorna apenas a atualização de status postado pelo usuário em seu próprio perfil
  - /{user-id}/posts retorna os posts criados pelo usuário no seu próprio mural ou nos dos amigos, e pode incluir qualquer conteúdo como links compartilhados, check-ins, fotos e atualização de status.

- Iremos utilizar a versão 2.3 para recuperar:
  - /{user-id}/home Retorna o fluxo de todos os posts criados pelo usuário e seus amigos. O que normalmente se encontra no Feed de Notícias.
  - /{user-id}/feed inclui tudo que você vê no seu perfil (links compartilhados, check-ins, fotos, atualização de status, além de incluir posts criados por amigos no perfil do usuário.

 Primeiro iremos recuperar o que vemos no perfil do usuário:

- Iremos utilizar os módulos
  - facebook, json e requests

# Requisições HTTP





#### Um parênteses sobre o módulo requests:

 Ele é utilizado para fazer requisições HTTP, será útil para que possamos requisitar as outras páginas de conteúdo do Facebook.

```
>>> r = requests.get('http://python.org')
>>> r.status_code # 200 a requisição foi realizada
>>> r.encoding
>>> r.text # Imprime o texto recuperado através da requisição
>>> 'Python' in r.text
True
```

- 1. Temos que importar o módulo.
- 2. Gerar o token de acesso (read stream)
- 3. Definir o acesso através da classe GraphAPI
- 4. Entender como os dados são retornados
- 5. Salvar em um arquivo para posterior análise

Abra o notebook "aula6-parte5-recuperar.ipynb"

#### Conteúdo da Aula

- Objetivo
- Facebook API
- Exercícios



#### Exercícios

- Exercício 1 Crie um álbum de fotos chamado Cursos.
   Depois adicione as 4 imagens que estão disponíveis na pasta da disciplina.
- Importante A descrição (name) de cada imagem deve ser o nome da imagem, sem a extensão (.jpg ou .png).
- Utilize a biblioteca embutida chamada os para visualizar todos os arquivos da pasta atual através do método os.listdir('.')

#### Exercícios

Exercício 2 – Altere o programa da aula6-parte5recuperar ('me/feed') e imprima, além do histórico,
o nome, tipo e a hora da criação.

- story
- name
- type
- created\_time

#### Exercícios

- Exercício 3 Utilizando o feed de notícias, recuperar os objetos JSON e salve-os em arquivo. Depois carregue no DataFrame do pandas.
  - Faça uma análise exploratória nos dados e defina quais informações seriam relevantes.

Facebook API

63

# Referências Bibliográficas

Mastering pandas – Femi Anthony – Packt Publishing, 2015.

Data Science from Scratch – Joel Grus – O'Reilly, 2015.

Python for Data Analysis – Wes McKinney – USA: O'Reilly, 2013.

- Referência da Graph API
  - https://developers.facebook.com/docs/graph-api

# Referências Bibliográficas

Python for kids – A playful Introduction to programming –
 Jason R. Briggs – San Francisco – CA: No Starch Press, 2013.

 Python Cookbook – David Beazley & Brian K. Jones – O'Reilly, 3th Edition, 2013.

 As referências de links utilizados podem ser visualizados em <a href="http://urls.dinomagri.com/refs">http://urls.dinomagri.com/refs</a>