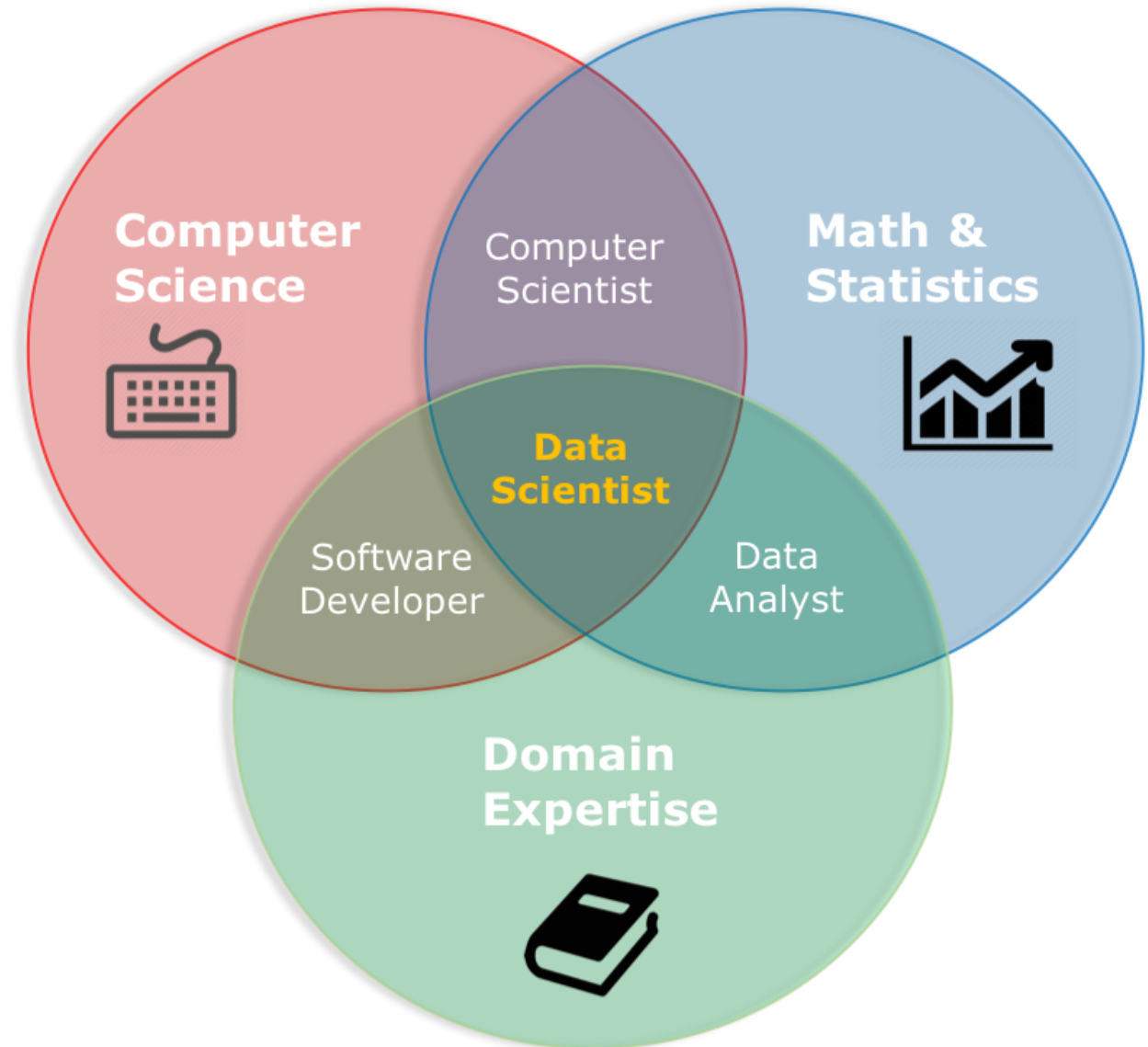


# *Introdução a Ciência de Dados*



# ***Apresentação Pessoal***



**2000 a 2004 – Graduação em Engenharia de Computação no ITA**



**2005 a 2008 – Mestrado em Eng. de Computação e Eletrônica no ITA**



**2009 a 2015 – Doutorado em Eng. de Computação e Eletrônica no ITA**



**2004 a 2010 – Empreendedor, sócio em empresa de base tecnológica**



**2013 a 2017 – Censipam / Ministério da Defesa**



**2016 – Professor de Concursos**



**2017 – Ministério da Economia**



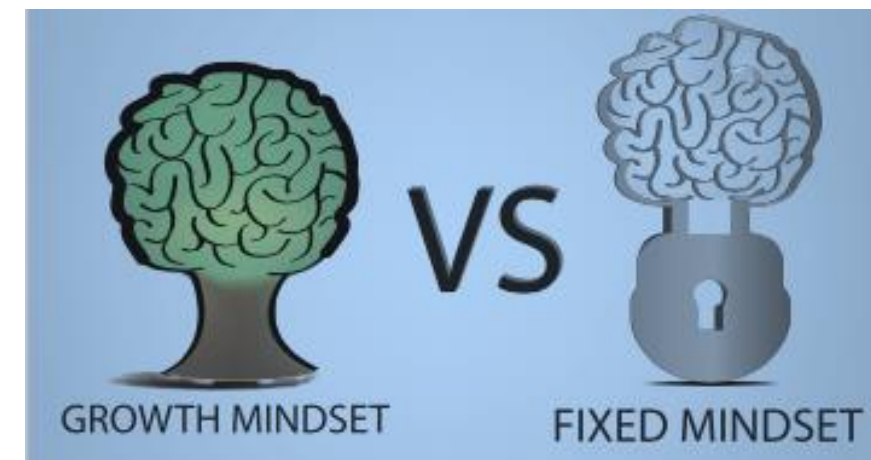
**2020 – Fintech Empréstimo P2P**

# *Conteúdo do Curso*

- Introdução a Ciência de Dados
- Introdução ao Jupyter Notebook / JupyterHub
- Introdução ao Python
- Introdução ao Python Pandas
- Coleta de Dados em Diversos formatos
  - CSV, SQL, API REST, Web Scrapy
- Visualização de Dados
- Conceitos Básicos de Estatística Descritiva
- Estudos de Caso
- Configuração de Ambiente de Desenvolvimento

# *Porque agregar novas habilidades (capacitação)*

- Manter-se atualizado com as tendências
  - Que vão dominar o futuro
- Atitude de crescimento (Growth mindset)
  - Quando acreditamos que nossa
    - ✓ inteligência,
    - ✓ habilidades criativas e
    - ✓ caráter
      - são coisas que podemos melhorar significativamente.



# Growth Mindset

- "Anyone who has never made a mistake has never tried anything new."
  - Albert Einstein
- "Vision is not enough. It must be combined with venture. It is not enough to stare up the steps; we must also step up the stairs."
  - Vaclav Havel
- "A man who has committed a mistake and doesn't correct it is committing another mistake."
  - Confucius



# *Growth Mindset*

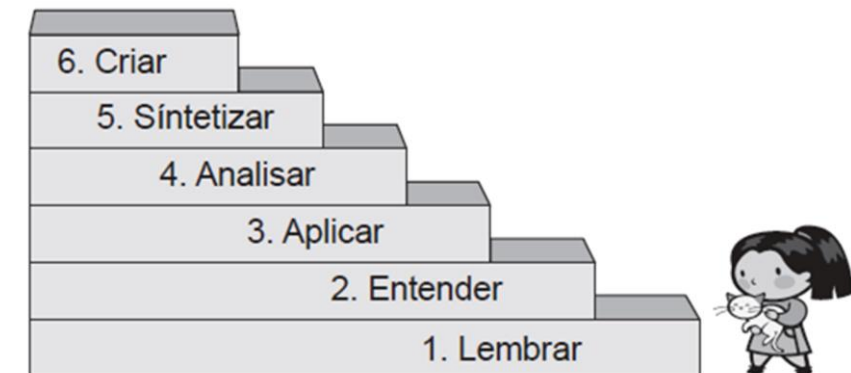
- "Challenges are what make life interesting. Overcoming them is what makes life meaningful."
  - Joshua Marine
- "It's kind of fun to do the impossible."
  - Walt Disney
- "The problem human beings face is not that we aim too high and fail, but that we aim too low and succeed."
  - Michelangelo





# *Capacitando-se em TI com eficiência*

- Recomendações pedagógicas
  - Meta: tornar-se independente/autossuficiente em TI.
    - ✓ Descobrir a solução para os seus próprios problemas sozinho. Lendo e buscando alternativas.
- Não existe gene da TI.
  - Essa habilidade pode ser aprendida e melhorada.
- Inglês
  - Vocabulário
    - ✓ aprender língua / linguagem de programação
  - Aprender a se expressar
    - ✓ buscar aprender os termos exatos



# *Capacitando-se em TI com eficiência*

- Trade-off

- Capacitação prévia X Busca de informação sob demanda.

- ✓ Afilar o machado

- *“Dê-me seis horas para derrubar uma árvore e passarei as quatro primeiras afiando o machado.” (Abraham Lincoln)*

- Como eu programo:

- Interativamente e iterativamente,

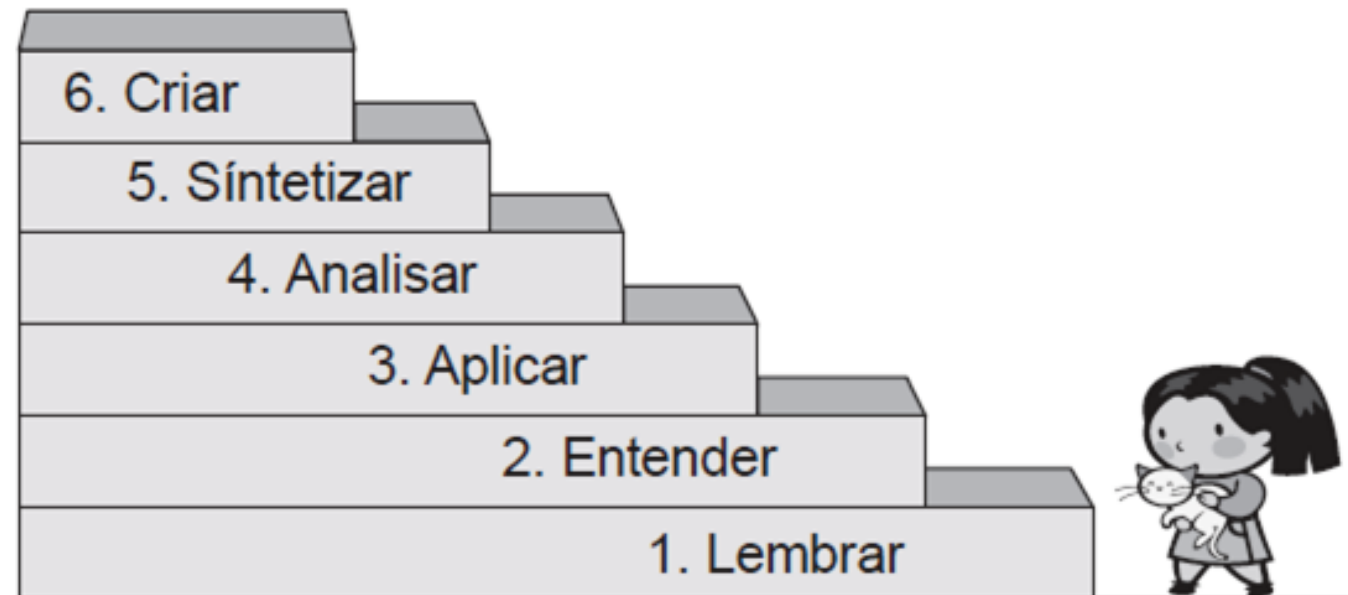
- sem focar em memorizar sintaxe/nomes, e

- sabendo perguntar/pesquisar.



# Metodologia de Aula

- Conteúdo associado a experiências e emoções
  - <https://podcasts.apple.com/br/podcast/tea-for-teaching/id1293488412>
- Taxonomia de Bloom
  - para adquirir uma nova habilidade pertencente ao próximo nível,
    - ✓ deve-se ter adquirido a habilidade do nível anterior.



# Metodologia do Curso

- Aulas teóricas, acompanhadas de
  - exercícios aplicados e
  - exercícios extra/opcionais.
- Grupos de 4 alunos
  - Ajudar e ser ajudado pelos próprios colegas (*peer instruction*);
  - Consulte a lista de grupos/alunos no canal da aula no slack;
  - Conectados numa sala de vídeo conferência;
    - ✓ Desenvolvendo os exercícios individualmente e tirando dúvidas entre si;
    - ✓ O monitor/professor entra na sala para sanar a dúvida registrada no github.
- **Não desconcentrar** nos momentos da aula teórica
  - Atenção/Foco/Celular

# *Horário das Aulas*

- 9h00 as 17h00
  - 09h00 as 12h00: Aula Teórica;
    - ✓ Presença será registrada às 9h00 todos os dias no Canal \_presença do Slack.
  - 14h00 as 17h00: Aula Prática (Resolução de Problemas com Python).
    - ✓ Live coding e dúvidas personalizadas (nos grupos) com os monitores.
- Monitores estarão a disposição para tirar dúvidas.
  - Registrem as dúvidas no github. O atendimento será por ordem FIFO (first-in, first-out).
- Tente não se atrasar para não perder o conteúdo
  - Existe uma cadeia de dependência entre os conteúdos.
    - ✓ Perdendo um você pode não entender os próximos

# *Metodologia do Curso*

- Jupyter hub/nbgrader,
  - Ambiente na nuvem para resolução de exercícios e correção automática.
- Projeto final
  - Tema livre;
  - Formato de entrega: um caderno Jupyter.
    - ✓ Com código e explicações, contando uma história.

# *Avaliação: Atividades Práticas e Individuais*

- Atividades simples e de propósito específico
  - (\*)Exceto a última: atividade de aplicação
- Cronograma e Pontuação

Atividade	Data da Entrega	Pontos
Atividade 1	14/10	1
Atividade 2	15/10	1
Atividade 3	16/10	1
Atividade 4	19/10	1
Atividade 5	20/10	1
Atividade 6	21/10	1
Atividade 7	22/10	1
Atividade 8	23/10	1
Projeto	25/10	6

# *Quando você pensar que não vale a pena experimentar*

- Se você falha ao planejar, está planejando falhar.
  - *Benjamin Franklin*
- Se enxerguei mais longe, foi porque me apoiei em ombros de gigantes.
  - *Bernard of Chartres / Isaac Newton*
- Dê-me seis horas para derrubar uma árvore e passarei as quatro primeiras afiando o machado.
  - *Abraham Lincoln*
- A definição de insanidade é fazer a mesma coisa repetidamente e esperar resultados diferentes.
  - *Rita Mae Brown / Albert Einstein*

## ***Bibliografia do Curso***

- **McKinney, W. (2018). Python for data analysis: Data wrangling with pandas, NumPy, and IPython.**
- Knafllic, Cole Nussbaumer. Hoboken, New Jersey : John Wiley & Sons, Inc., [2015].
- WHEELAN, C. Estatística, o que é, para que serve, como funciona. Zahar, 2016.



## *Bibliografia Complementar*

- DOWNEY, Allen B. Think Python: How to Think Like a Computer Scientist.  
<http://greenteapress.com/thinkpython/html/index.html>
- Analytics Edge, do MIT na plataforma EdX:  
<https://www.edx.org/course/the-analytics-edge>
- TUFTE, Edward R. 2006. Beautiful evidence. Cheshire, Conn: Graphics Press.

## *Bibliografia Complementar*

- KURT, Will. 2019. Bayesian statistics the fun way: understanding statistics and probability with Star Wars, LEGO, and rubber ducks.
- KAHNEMAN, D. Rápido e Devagar - Duas Formas de Pensar. Objetiva, 1a. ed. 2012.
- SILVER, N. Sinal e Ruído. Intrínseca, 1ª ed. 2013.
- TALEB, N. N. Iludido pelo Acaso – A influência oculta da sorte nos mercados e na vida. Record, 2004.
- TALEB, N. N. A lógica do cisne negro. Best Seller, 15ª ed. 2008.
- DARRELL, H. Como mentir com estatística. Intrínseca, 2016.

# *Conteúdos Recomendados*

- Datahackers – Comunidade de Cientistas de Dados (em português)
  - <https://datahackers.com.br/newsletter>
  - <https://datahackers.com.br/podcast>
  - <https://datahackers.com.br/slack>
- Toward Data Science Podcast
  - <https://towardsdatascience.com/tagged/podcast>
- Conteúdo de jornalismo de dados:
  - <http://meneg.at/>
  - <https://pudding.cool/>

# *Conteúdos Recomendados*

- Effective Programming For Data Science
  - <https://towardsdatascience.com/february-edition-effective-programming-for-data-science-b6c122b35a87>
- Podcast com os autores do livro Freakonomics
  - <https://freakonomics.com/>
- Podcast sobre Lean Startup
  - <https://leanstartup.co/podcasts/>