

Aula 01



[Rodrigo Moutinho | LinkedIn](#)

Técnico → Não Técnico

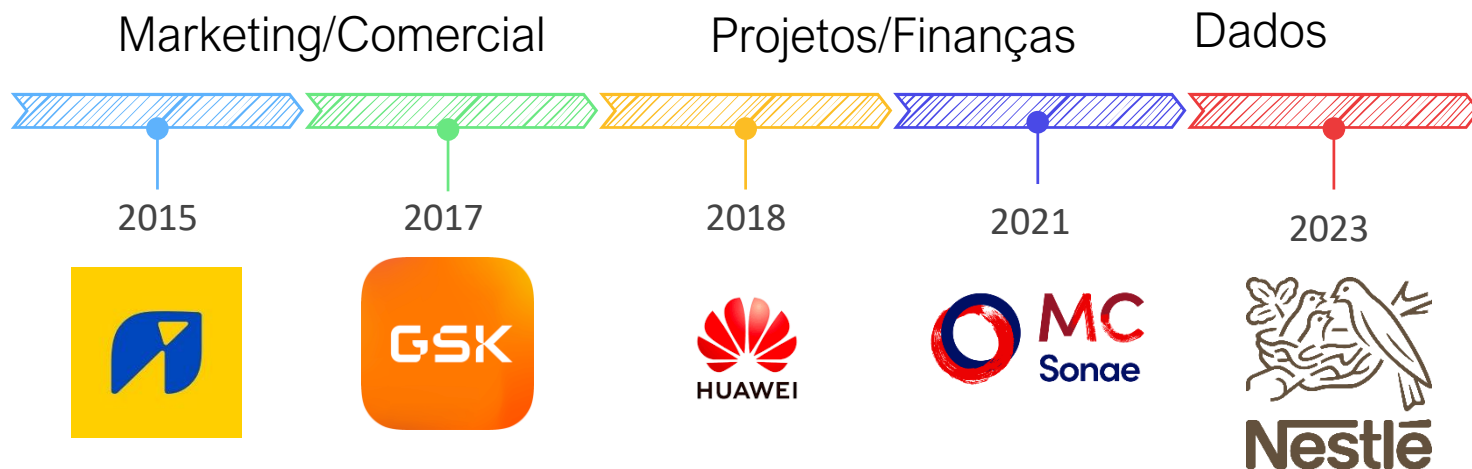
Comunicativo

Curioso

Sintaxe de código

Decorar conceitos

Detalhes



E vocês?

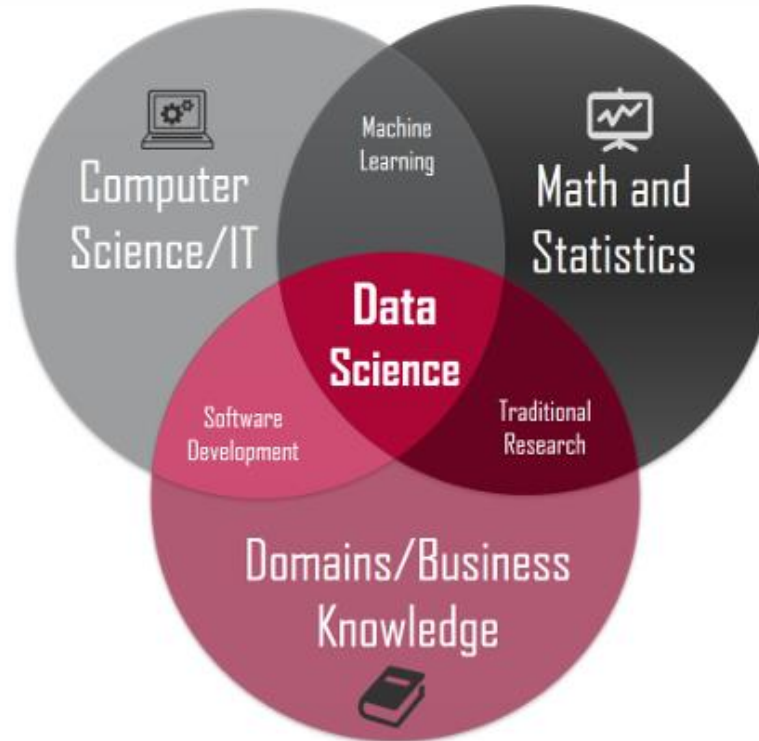
Motivo de estarem fazendo curso de DSBA?
Transição de carreira? Aprender Novas Habilidades?

1 minuto apresentação

Avaliação

Assiduidade e Pontualidade (15%)
Motivação e Participação (35%)
Trabalho Individual (50%)

O que é Data Science?






Definição: Data science é o estudo dos dados para extrair insights significativos para negócios. É uma abordagem multidisciplinar que combina práticas das áreas de matemática, estatística, inteligência artificial e engenharia da computação para analisar grandes quantidades de dados. (Amazon Web Services)

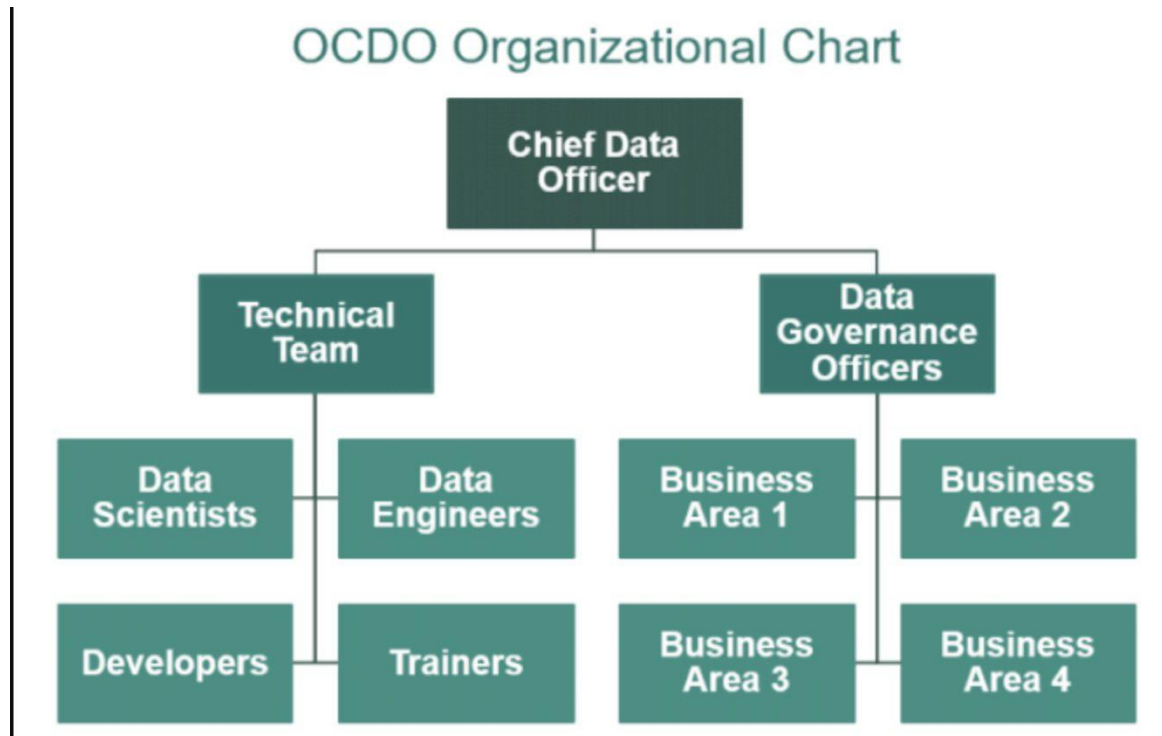
Dados: O novo petróleo



Profissionais de dados

Data Engineer	Data Scientist	Data Analyst
		
Construir e otimizar a infraestrutura de dados que irá permitir que o Cientista de Dados e o Analista de Dados rodem suas análises.	Utiliza estatística, aprendizado de máquina e algoritmos matemáticos para realizar previsões, identificar padrões e responder perguntas de negócio.	Compreende questões de negócio, consolida e elabora dashboards e relatórios para comunicar resultados.
Skills Programação, Bancos de Dados, Infraestrutura Cloud, Big Data.	Skills Programação, Bancos de Dados, Estatística, Matemática	Skills Comunicação, UX, Conhecimentos de Negócio, Bancos de Dados
Techs SQL, Python, Cloud (Azure, AWS, Google)	Techs SQL, Python, R, Cloud	Techs SQL, Power BI, Tableau, Qlik View, Excel

Estrutura Departamento de Dados (Tradicional)



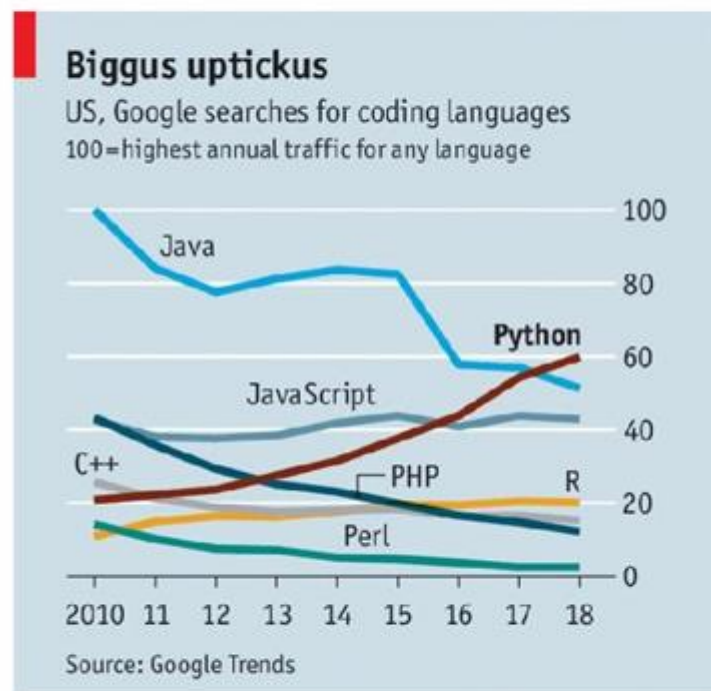


Porquê Python?

1. **Fácil de aprender e compreender.** A curva de aprendizagem é relativamente baixa sendo perfeito para programadores de todos os níveis
2. **Open source.** A utilização de Python é gratuita, bem como as suas bibliotecas.
3. Beneficia de um extenso **apoio da comunidade.**
4. Bibliotecas focadas na limpeza e processamento de dados, bem como em machine learning.



Porquê Python?



Economist.com

Python is not perfect. Other languages have more processing efficiency and specialized capabilities. C and C++ are “lower-level” options which give the user more control over what is happening within a computer’s processor. Java is popular for building large, complex applications. JavaScript is the language of choice for applications accessed via a web browser. Countless others have evolved for various purposes. But Python’s killer features—**simple syntax that makes its code easy to learn and share**, and its **huge array of third-party packages**—make it a good general-purpose language. Its versatility is shown by its range of users and uses. The Central Intelligence Agency has employed it for hacking, Pixar for producing films, Google for crawling web pages and Spotify for recommending songs.”

<https://www.economist.com/graphic-detail/2018/07/26/python-is-becoming-the-worlds-most-popular-coding-language>

Breve História do Python



Guido van Rossum



...In December 1989, I was looking for a "hobby" programming project that would keep me occupied during the week around Christmas. My office ... would be closed, but I had a home computer, and not much else on my hands. I decided to write an interpreter for the new scripting language I had been thinking about lately: a descendant of [ABC](#) that would appeal to [Unix/C hackers](#). I chose Python as a working title for the project, being in a slightly irreverent mood (and a big fan of [Monty Python's Flying Circus](#)).

Google Colab for Python



PySeek

<https://colab.research.google.com/>



Google Colab Tutorial for Beginners | Get Started with Google Colab

270K views • 3 years ago



Doga Ozgon

In this tutorial you will discover all the basics you need to know to get started with Google Colab. You don't need any setup to get ...

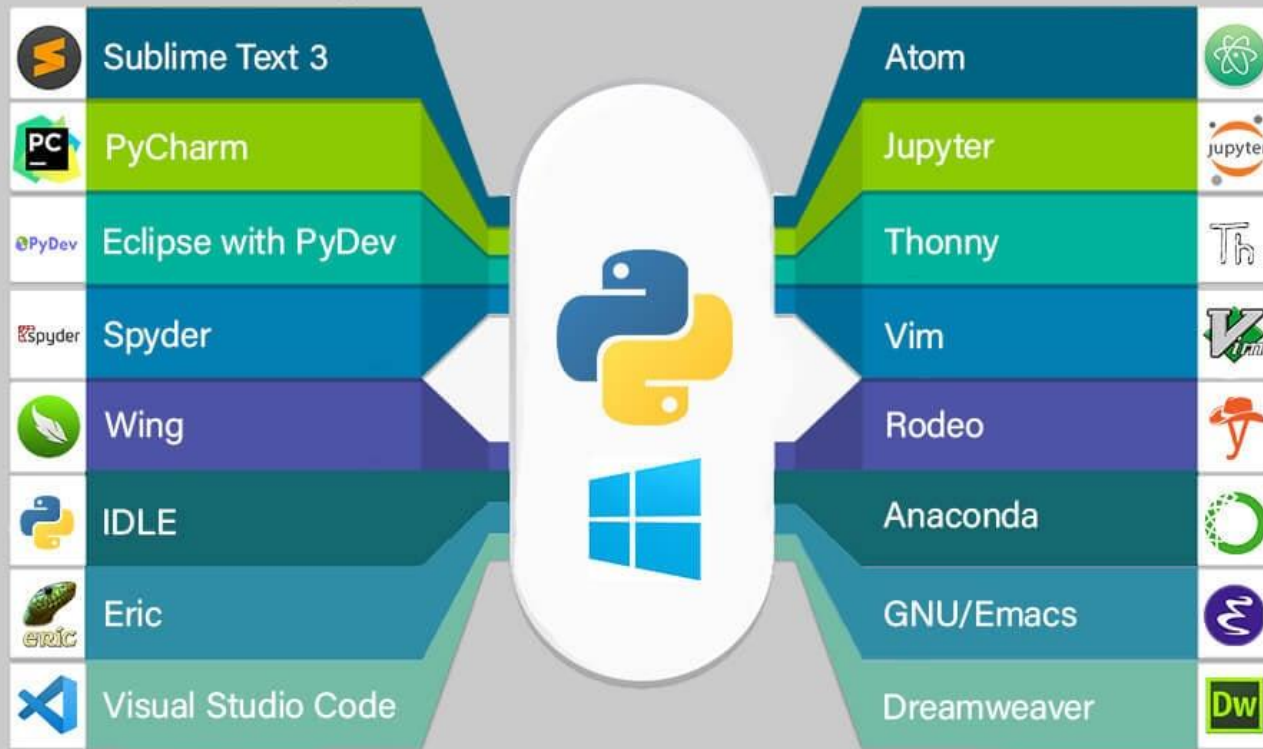


3 moments COMMAND + ENTER | SHIFT + ENTER | Graphics Processing Unit Tensor Processing Unit

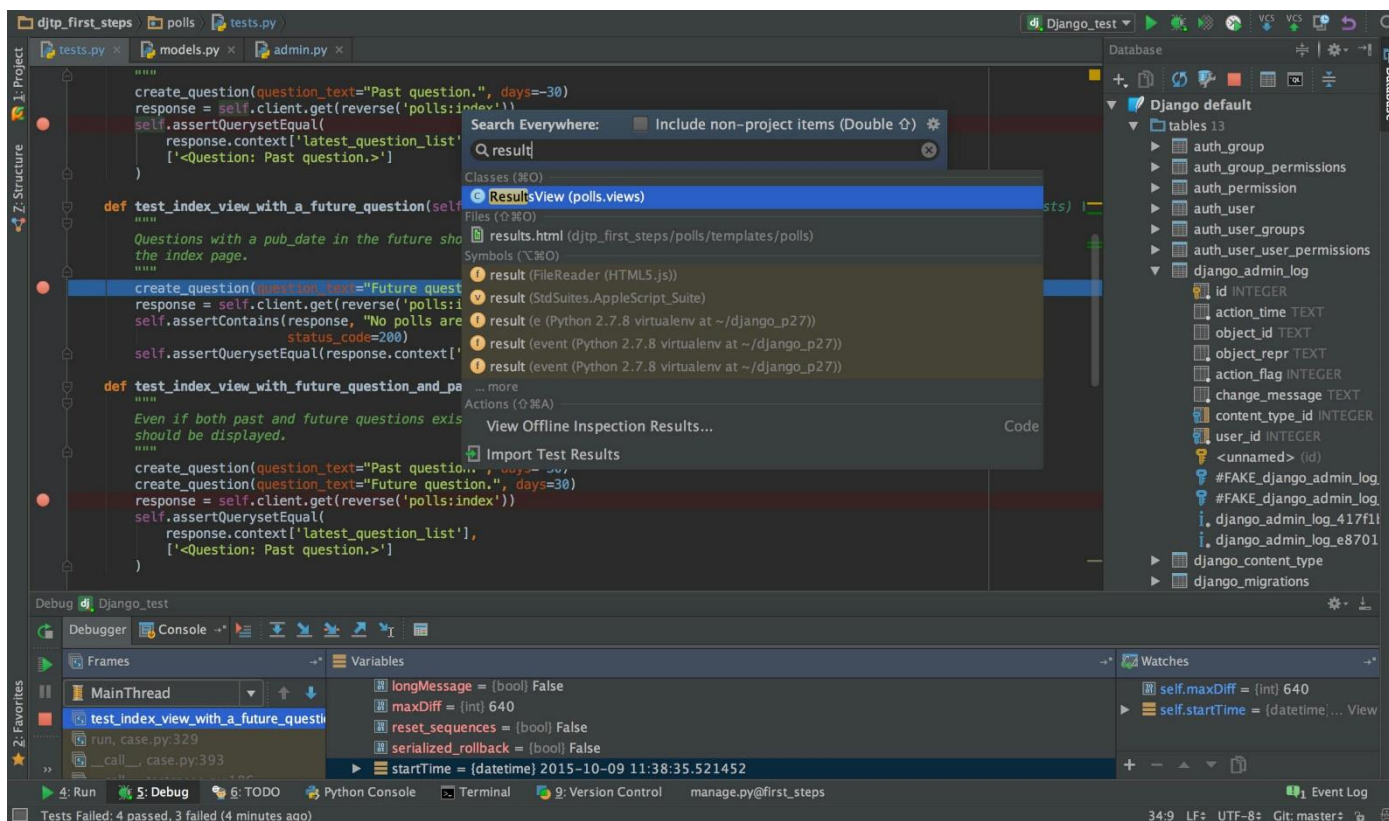


[Google Colab Tutorial for Beginners | Get Started with Google Colab](#)

Python IDE for Windows

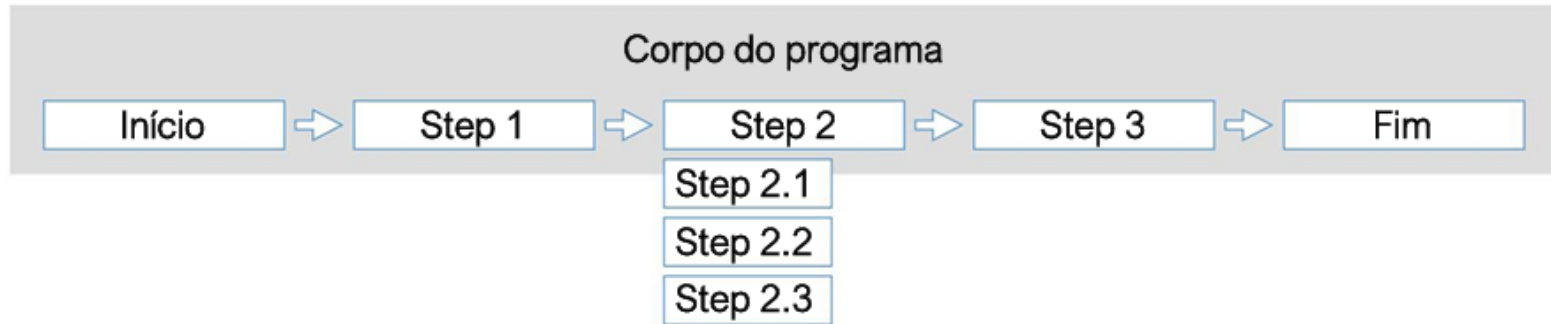


Como é no mercado de trabalho ?



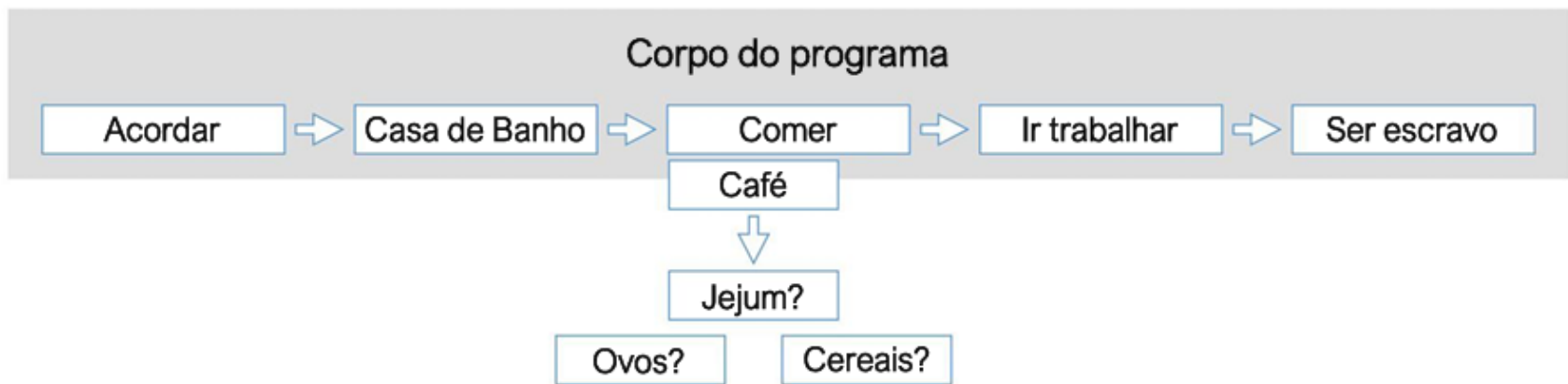
Programas de computador

Um programa corresponde a uma **sequência de instruções** que especifica como proceder a um determinado cálculo:



Programas de computador

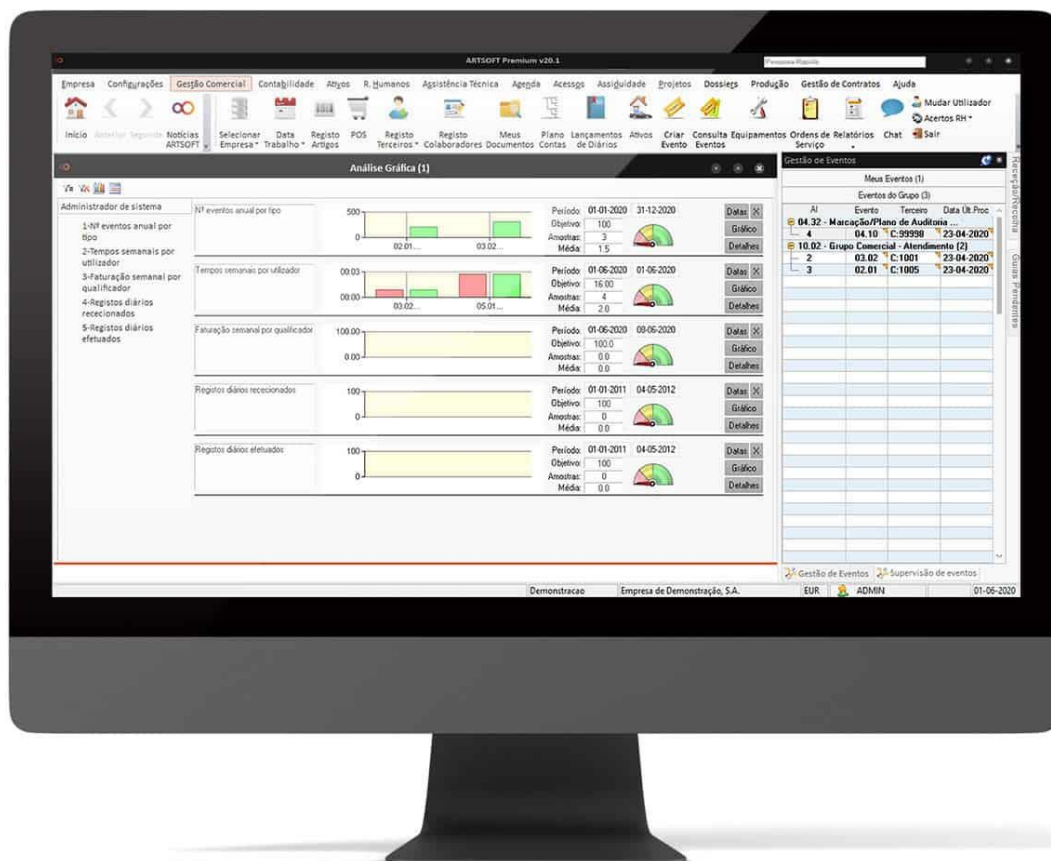
Um programa corresponde a uma **sequência de instruções** que especifica como proceder a um determinada cálculo:

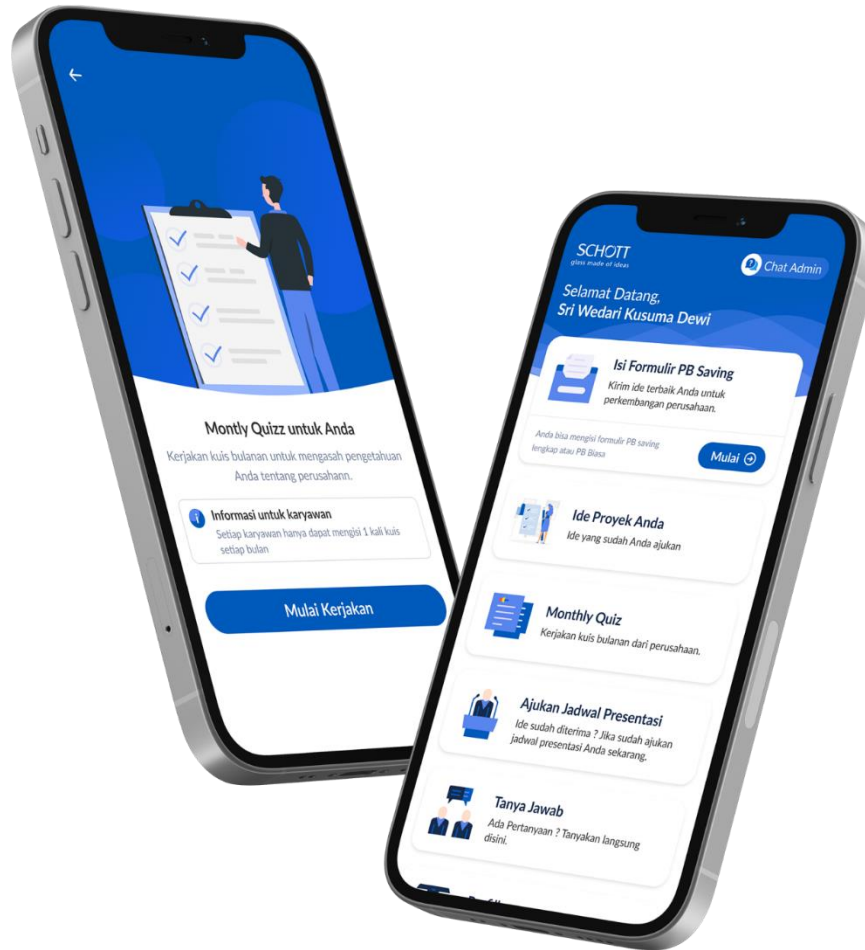


Programas de computador

Um programa corresponde a uma **sequência de instruções** que especifica como proceder a um determinada cálculo:

Input	Obter dados definidos por nós, de um ficheiro ou outro dispositivo.
Output	Apresentar dados no ecrã ou enviar dados para um ficheiro ou outro dispositivo.
Matemática	Efetuar operações matemáticas básicas como a adição e a multiplicação.
Execução condicional	Verificar determinadas condições e executar a sequência de instruções adequada.
Repetição	Realizar uma ação repetidamente, normalmente com algumas variações.





Data Dashboard

Auto date range

📅

This Week

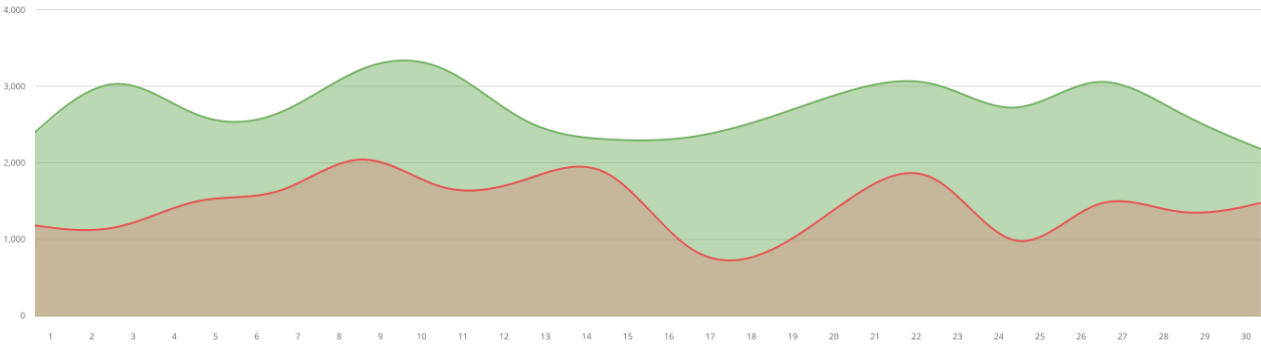
Services

All

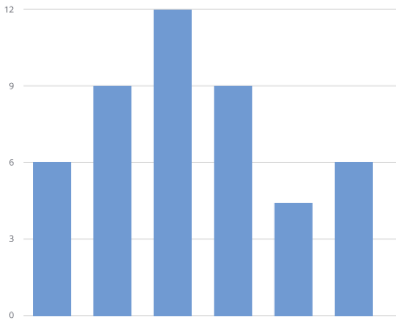
Posts

All

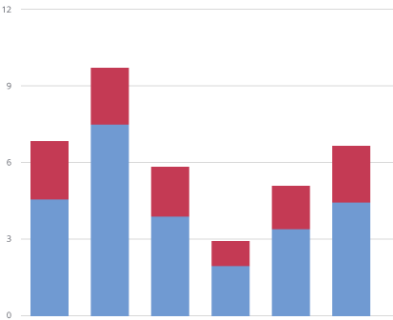
Earnings Before Interest, Taxes, Depreciation, and Amortization (EBITDA)



Net Profit Margin



Debt-to-Equity Ratio



Revenue

\$24.5

↑ 13%

vs previous 7 days

Avg Profit Margin

9.5%

↑ 1

vs previous 7 days

Return On Investment (ROI)

19.1%

8%

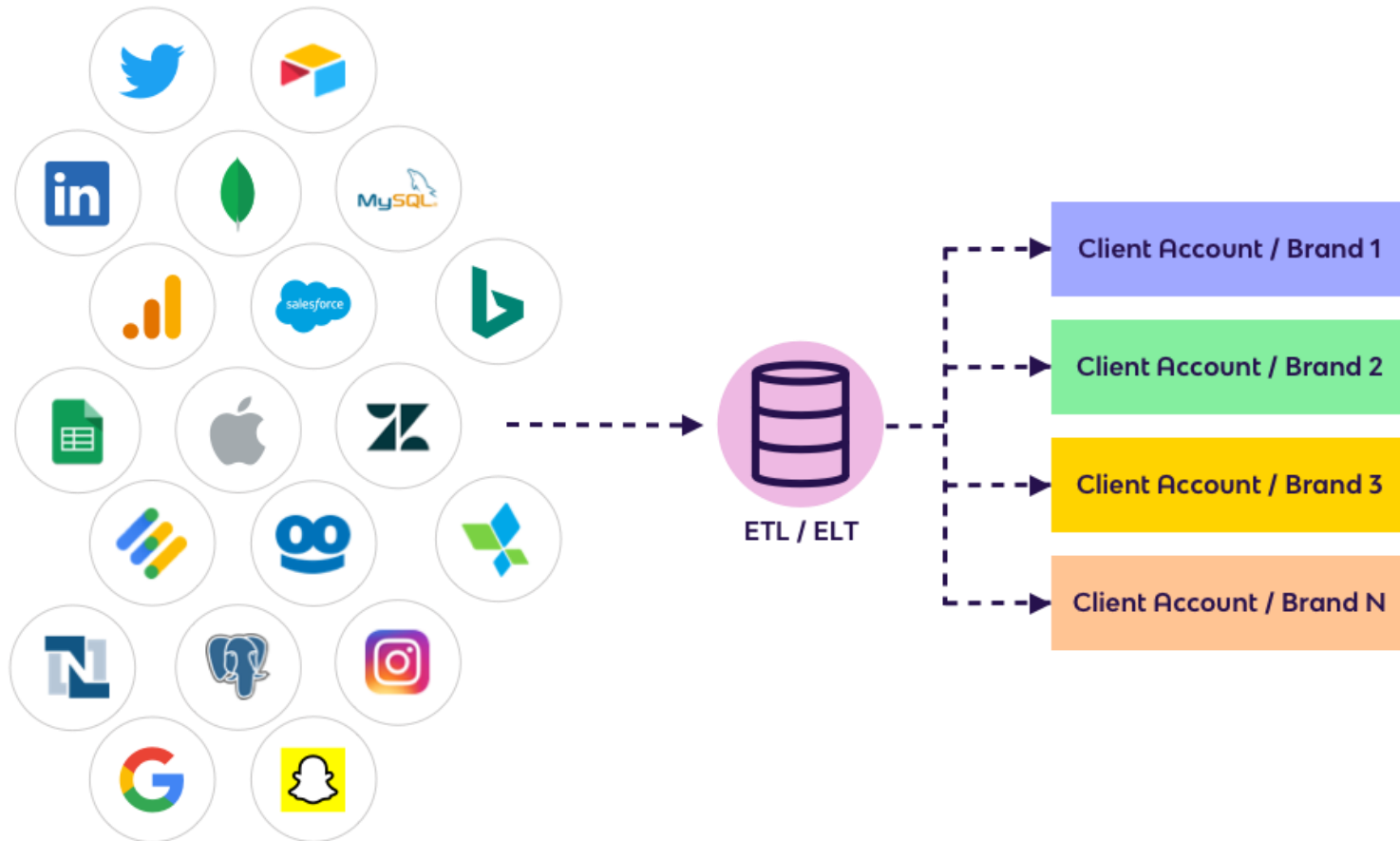
vs previous 7 days

CLV

\$2,176

↑ 2.3%

vs previous 7 days





Python Tutorial

Google Colab for Python



PySeek

[Aula 01](#)

Google Colab for Python



PySeek

[Aula 01](#)

Próximas Aulas

1. Data Science Fundamentals

- › Data Science Definition /Applications
- › Data Mining main branches/techniques
- › Python Fundamentals – Key concepts
- › Data Understanding
- › Data loading and cleansing, Plotting, Feature extraction

Aula 1: Iniciando Data Science Fundamentals

- Introdução
- O que é Data Science
- Profissionais de dados
- Python
- Google Colab

Aula 2: Variáveis, Operadores e Estrutura

- Data Mining
- Variáveis
- Operadores
- Estrutura de Dados

Próximas Aulas

1. Data Science Fundamentals

- › Data Science Definition /Applications
- › Data Mining main branches/techniques
- › Python Fundamentals – Key concepts
- › Data Understanding
- › Data loading and cleansing, Plotting, Feature extraction

Aula 3: Iniciando Data Science Fundamentals

- Dados x Conhecimento x Informações
- Market Basket Analysis
- Python (Flow Control, Loop, IF etc)
- Exercício 1

Aula 4: Variáveis, Operadores e Estrutura

- Artificial Intelligence x Machine Learning
- Python (Funções, Lambda)
- Exercício 2

Próximas Aulas

1. Data Science Fundamentals

- › Data Science Definition /Applications
- › Data Mining main branches/techniques
- › Python Fundamentals – Key concepts
- › Data Understanding
- › Data loading and cleansing, Plotting, Feature extraction

Aula 5: Pandas I

- Series e Dataframes
- Prática

Aula 6: Pandas II

- Join/Merge
- Prática
- Exercício 3

Próximas Aulas

1. Data Science Fundamentals

- › Data Science Definition /Applications
- › Data Mining main branches/techniques
- › Python Fundamentals – Key concepts
- › Data Understanding
- › Data loading and cleansing, Plotting, Feature extraction

Aula 7: Data Viz

- Matplotlib x Seaborn
- Data Storytelling
- Prática

Aula 8: Git (Version Control)

- Prática
- Exercício 4

Extra

2. Databases Fundamentals

- › Database concept
- › Datawarehouse concept
- › Relational and non-relational
- › Major providers
- › Architecture, Models, Definition
- › Types of Data Models
- › Fact and Dimension tables, Keys
- › Star and Snowflake schema
- › SQL Fundamentals, Key concepts

3. Exploratory Data Analysis

- › Types of data
- › Descriptive Statistics
- › Dirty & missing Data
- › Cleaning Data Techniques and tools
- › Visual Exploration

Extra

4. Analysing & Visualizing Data.

- › Introduction to Information Visualization
- › The role of visualisation in data science
- › Good/bad examples
- › Plotting with python
- › Visualisations (and dashboards) for different audiences
- › Building a dashboard with python

5. Business Analytics

- › Business Analytics
- › Introdução ao Google Analytics – overview do software
- › Diferentes soluções/ferramentas de Web Analytics para além do Google Analytics
- › Perspetivas de métricas e dimensões (KPIs técnicos, quantitativos e qualitativos)
- › Objetivos e KPI's (O que são? Para que servem? Exemplos de utilização)
- › Account e Tracking Code – o que é e onde encontrar, como criar e instalar
- › Análise dos vários painéis (público-alvo, aquisição, comportamento, conversões e real-time)
- › Relatórios personalizados
- › Dashboards (com Data Studio)
- › Funis de conversão: O que são? Qual a sua importância? Otimização de um funil de conversão
- › Perceber para onde fogem as visitas? Porquê? Como evitar?
- › Segmentos avançados. Qual a utilidade? Como se criam?
- › UTMS – o que são? Para que servem?

Extra

6. Big Data Technologies

- › Big Data Ecosystem
- › Introduction to Big Data
- › Defining a Big Data Cluster
- › Distributed Query SQL Query Engine
- › Distributed Data Processing
- › Distributed Processing with Spark
- › Data Pipelines
- › Data Integration Tools
- › Workflow Concepts
- › Building a Data Pipeline

7. Quantitative & Statistical Analysis and Modeling

- › Quantitative Data Definition
- › Types of Quantitative Data
- › Collection Methods
- › Analysis Methods
- › Steps to conduct Quantitative Data Analysis
- › Quantitative Data Examples
- › Advantages & Disadvantages of Quantitative Data
- › Differences between Quantitative and Qualitative Data

Extra

8. Machine Learning Models

- › Introduction to Machine Learning
- › ModelsSupervised Learning:
- › RegressionSupervised Learning: Classification
- › Machine Learning: Advanced Concepts
- › Unsupervised Learning

9. Introduction to Networks

- › Introduction to Neural
- › NetworksDifferent Types of Neural
- › NetworksSome Selected Examples