

# Mentoria Semana 1

*Aula 1 - Introdução ao mundo do desenvolvimento e  
Aula 2 - Introdução à lógica*



```
column-8 w-col-stack hero-section
content-left">
  <div class="hero-container
heading-container">
  <div class="heading-tagl
container">
    <h3 class="heading-17 h
subtitle">gama experience-/h3
    <span class="text-span-
<div class="hero-section wf-
section" style="top: 1px 1
important">
  <div class="columns-5 w-row
hero-section-container">
    <div class="hero-section-
columns">
      <div class="column-8 w-col
col-6 w-col-stack hero-section
content-left">
        <div class="hero-container
heading-container">
          <div class="heading-tagl
container">
            <h3 class="heading-17 h
subtitle">gama experience-/h3
            <span class="text-span-
<div class="hero-section wf-
section" style="top: 1px 1
important">
          <div class="columns-5 w-row
hero-section-container">
            <div class="hero-section-
columns">
              <div class="column-8 w-col
col-6 w-col-stack hero-section
content-left">
```



# Caio Fernandes

Professor de curso técnico de informática,  
Especialista de Cultura e Governança de dados  
no Grupo Boticário, e Jogador de Rugby

[linkedin.com/caio--fernandes--santos/](https://linkedin.com/caio--fernandes--santos/)





# Apresentação

*Conhecendo as expectativas*





# Trilha preparatória

Carga horária de estudo: 26h (8h assíncrono + 18h ao vivo)

- Introdução ao mundo do desenvolvimento
- Lógica
- Metodologias Ágeis
- GIT
- Programação
- Fundamentos de BI

# Introdução ao mundo do desenvolvimento



```
<div class="column-8 w-col-6 w-col-stack hero-section content-left">
  <div class="hero-container heading-container">
    <div class="heading-tag container">
      <h3 class="heading-17 h-subtitle">gama experience</h3>
      <span class="text-span">
        <div class="hero-section wf-section" style="top: 1px !important">
          <div class="columns-5 w-row hero-section-container">
            <div class="hero-section-columns">
              <div class="column-8 w-col-6 w-col-stack hero-section content-left">
                <div class="hero-container heading-container">
                  <div class="heading-tag container">
                    <h3 class="heading-17 h-subtitle">gama experience</h3>
                    <span class="text-span">
                      <div class="hero-section wf-section" style="top: 1px !important">
                        <div class="columns-5 w-row hero-section-container">
                          <div class="hero-section-columns">
                            <div class="column-8 w-col-6 w-col-stack hero-section content-left">
```

# Papéis envolvidos na cadeia de valor de dados



No varejo, o percentual de consumidores brasileiros que pretendem adotar um comportamento de compra omnicanal, mesclando as compras entre lojas online e lojas físicas, cresceu de 29% em 2019 para 49% em 2021.<sup>1</sup> Seguindo um estudo do BCG, empresas que usam dados nas áreas de marketing, vendas, supply chain, P&D e produção, podem aumentar o EBITDA até 20% a 30%.<sup>2</sup>

## SEPHORA

Com uma estratégia focada em omnicanalidade, digitalização e dados, a Sephora cresceu de forma exponencial nos últimos anos. Pelo quarto ano em seguida, a empresa se posiciona no primeiro lugar do "Retail Personalization Index 2021".<sup>3</sup>

### Exemplos de uso de dados

- Rastreamento e **análise de qualquer engajamento do cliente com a marca** e integração dos perfis online e off-line, como input para modelos de IA/ML, gerando uma experiência altamente personalizada
- Lançamento do Visual Artist já em 2016, produto mobile e in-store que usa realidade aumentada para **experimentar make-up virtualmente**
- **Chatbots** para atendimento e agendamento de makeovers nas lojas físicas
- Criação de **aplicativos de voz no Google Assistant** para agendar serviços e colecionar dados sobre preferências do consumidor
- Uso de **IoT nas lojas físicas** para identificar clientes para oferecer promoções diárias e personalizadas, e carregar o carrinho de compras e a lista de desejos

## NORDSTROM

A Nordstrom é uma das empresas mais digitalizadas no segmento de varejo, oferecendo uma experiência omnicanal, que é fundada no amplo uso de dados.<sup>4</sup> Nos EUA, ela lidera o ranking das empresas mais queridas no segmento de varejo de roupas há cinco anos consecutivos, utilizando uma estratégia muito simples: Criar melhores experiências para os clientes através da coleta e do uso de dados.<sup>5</sup>

### Exemplos de uso de dados

- **Rastreamento de pins no Pinterest** para identificar produtos populares e em seguida o uso dessas informações para promover os mesmos produtos nas lojas físicas
- Uso de **IoT e sensores nas lojas físicas** para identificar quem entra, qual parte da loja visita e quanto tempo fica
- **Integração do estoque da loja online e física** para que os clientes saibam onde encontrar o produto procurado
- Analisa um grande volume de **dados das mídias sociais** e incentiva o uso do cartão de crédito da Nordstrom e do programa de fidelidade para criar perfis de clientes com o objetivo da melhora dos produtos e do marketing



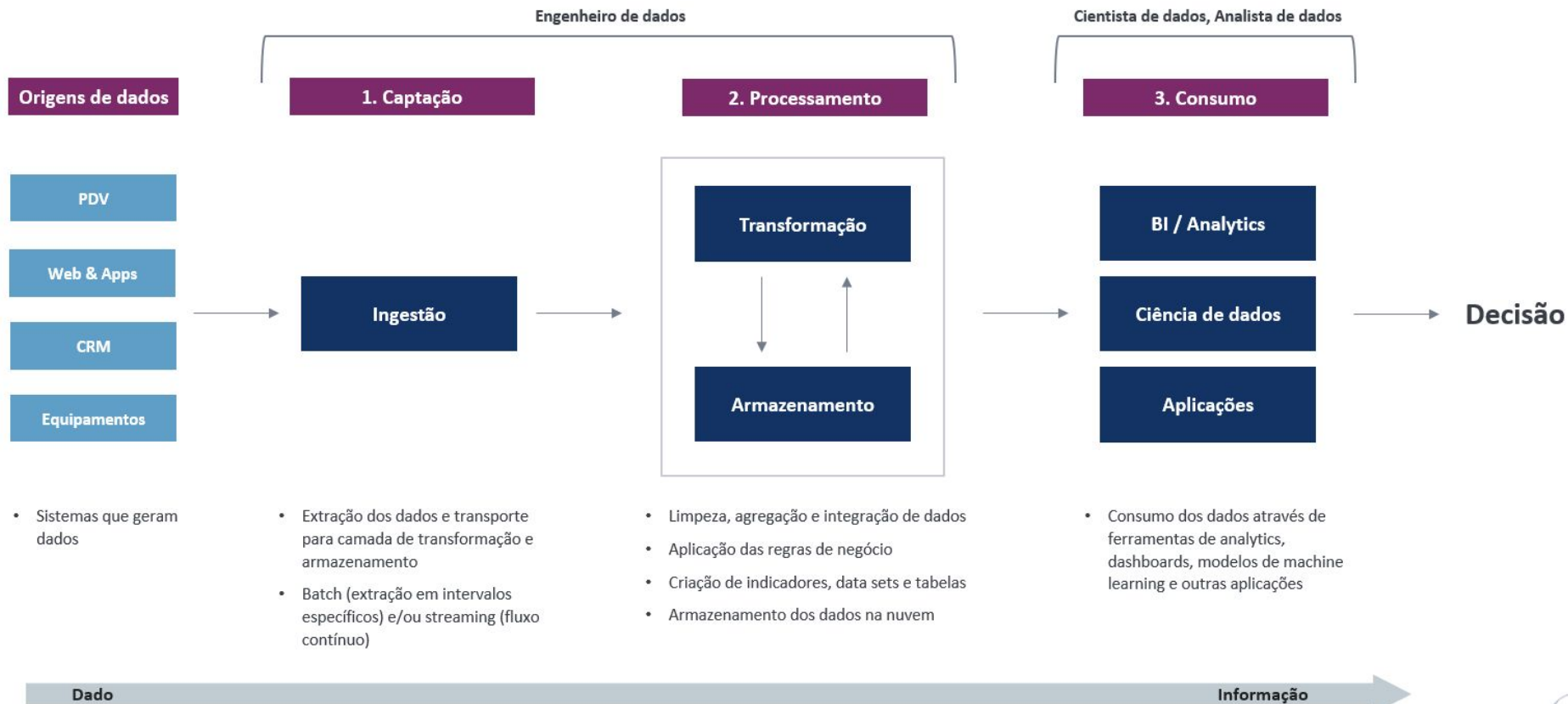
# Dados revolucionam a experiência de compra e do cliente



	Arquiteto de dados	Engenheiro de dados	Engenheiro de ML Ops	Cientista de dados	Desenvolvedor de BI	Data Product Owner
Resumo	<ul style="list-style-type: none"><li>• Cria o "blue print" do ambiente de dados</li></ul>	<ul style="list-style-type: none"><li>• Disponibiliza dados para a organização inteira</li></ul>	<ul style="list-style-type: none"><li>• Coloca e monitora modelos de IA no ambiente produtivo</li></ul>	<ul style="list-style-type: none"><li>• Cria modelos de inteligência artificial, olhando para o futuro</li></ul>	<ul style="list-style-type: none"><li>• Transforma dados e disponibiliza dashboards</li></ul>	<ul style="list-style-type: none"><li>• Especifica a construção e conceito do dado</li></ul>
Skills	<ul style="list-style-type: none"><li>• Plataformas cloud</li><li>• Tecnologias de processamento e integração de dados</li></ul>	<ul style="list-style-type: none"><li>• Programação</li><li>• Modelagem de dados</li><li>• Sistemas distribuídos</li></ul>	<ul style="list-style-type: none"><li>• Programação</li><li>• DevOps</li></ul>	<ul style="list-style-type: none"><li>• Estatística e matemática</li><li>• Programação</li><li>• Análise de dados</li><li>• Comunicação</li></ul>	<ul style="list-style-type: none"><li>• Modelagem e bancos de dados</li><li>• Visualização de dados e UX</li><li>• Ferramentas de BI</li></ul>	<ul style="list-style-type: none"><li>• Levantamento e documentação de requisitos;</li><li>• Metodologias ágeis</li></ul>
Responsabilidades	<ul style="list-style-type: none"><li>• Define como dados serão armazenados, integrados e consumidos</li><li>• Escolhe as ferramentas certas</li></ul>	<ul style="list-style-type: none"><li>• Limpa, transforma e disponibiliza dados no datalake e data warehouse</li></ul>	<ul style="list-style-type: none"><li>• Otimiza modelos de machine learning (ML)</li><li>• Coloca modelos de ML em produção</li><li>• Cria monitoramento e automação dos modelos</li></ul>	<ul style="list-style-type: none"><li>• Analisa dados</li><li>• Cria modelos de IA</li><li>• Apresenta resultados</li></ul>	<ul style="list-style-type: none"><li>• Cria dashboards</li><li>• Garante atualização e funcionamento dos dashboards</li></ul>	<ul style="list-style-type: none"><li>• Estrutura soluções técnicas e as transcreve em um backlog de histórias para desenvolvimento</li></ul>
Tecnologias	<ul style="list-style-type: none"><li>• AWS, GCP, Azure</li><li>• SQL, NoSQL</li><li>• Docker, Kubernetes</li></ul>	<ul style="list-style-type: none"><li>• Python, SQL, NoSQL, Spark, Kafka</li></ul>	<ul style="list-style-type: none"><li>• AWS, GCP</li><li>• Kubeflow, Tensorflow</li></ul>	<ul style="list-style-type: none"><li>• Python, R</li></ul>	<ul style="list-style-type: none"><li>• SQL</li><li>• Tableau, PowerBI, Qlik, SAS</li></ul>	<ul style="list-style-type: none"><li>• SQL, GCP, AWS, SAP DS, Qlik, Power BI, Tableau, Looker</li></ul>

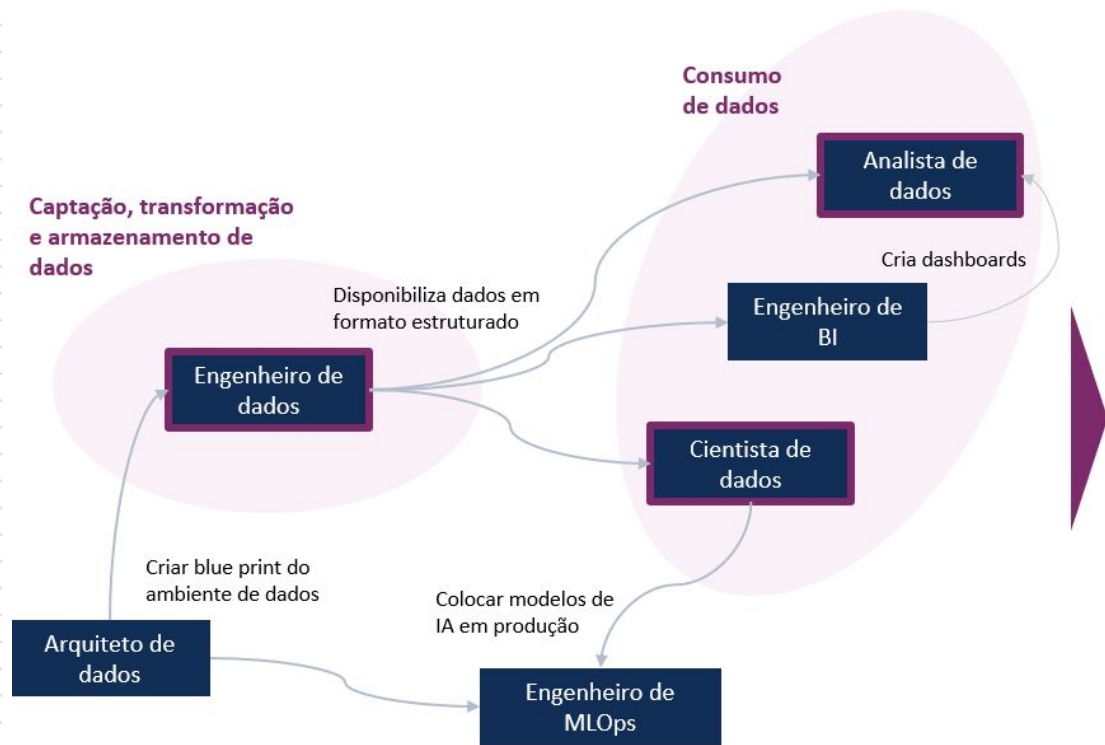
Quando se fala sobre times de dados, geralmente o **engenheiro de dados, cientista de dados e analista de dados** representam a maior quantidade de pessoas

# Processos necessários até o dado se tornar uma informação que possa ser utilizada





# Como os papéis de dados interagem e quais suas dependências?



## Engenheiro de dados

- Disponibiliza dados no BigQuery para consumo através de SQL, Dashboards, IA e APIs
- Complemento do cientista de dados

## Cientista de dados

- Usa dados para fazer análises profundas e criar modelos de IA
- Precisa trabalhar do lado do engenheiro de dados, que gera os insumos, e do engenheiro de MLOps, que coloca os modelos de IA em produção
- Usa dados estruturados e não estruturados (p.ex. textos, imagens, ...)

## Analista de dados

- Consome dados estruturados (tabelas, dashboards) para análise e apresentação

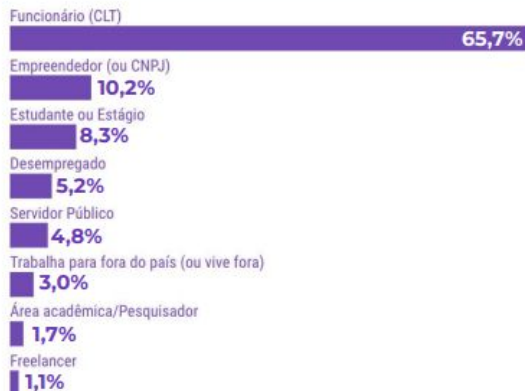
Não se trata de um modelo de cliente-fornecedor, mas de uma interação em qual principalmente engenheiros e cientistas de dados **precisam trabalhar muito bem alinhados e próximos**, para reduzir a quantidade de hand-overs e assim **garantir velocidade na geração de novas informações e nas decisões de negócio**.

# Cargos e Salários



## Perfil Demográfico

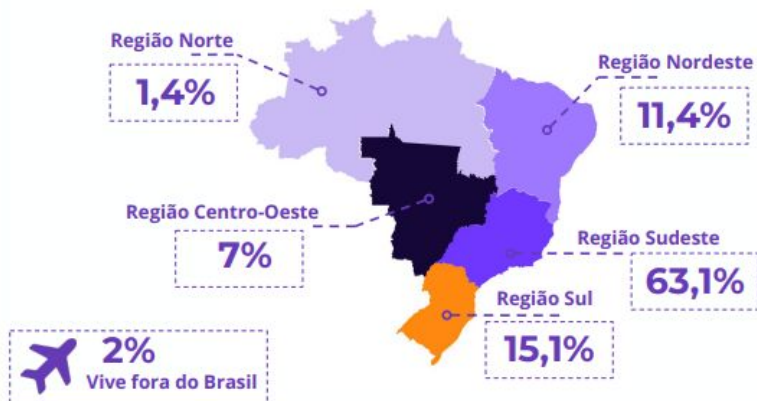
A maioria dos participantes tem um emprego na área



Homens ainda são a grande maioria dos profissionais de dados



Forte concentração de profissionais no Sudeste e Sul do Brasil



# Cargos e Salários



## Formação



Dentre os participantes da pesquisa, **apenas 2,3% não possuem graduação formal**, além disso, mais de 48% possuem algum tipo de pós-graduação.

### A área de dados possui diferentes perfis de profissionais

42,5%		Computação / TI / Eng. Software Sistemas de Informação
23,5%		Área de Engenharia
14,7%		Economia / Administração Contabilidade / Finanças
8,3%		Estatística/ Matemática Matemática Computacional
2,9%		Marketing / Publicidade Comunicação / Jornalismo
2,8%		Física / Química
1,8%		Ciências Biológicas / Farmácia Medicina / Área da Saúde
3,5%		Outras

# Cargos e Salários

## Carreira e experiência

Tempo de experiência na área de dados	Júnior	Pleno	Sênior	Gestor
Sem experiência	7%	4%	5%	5%
Menos de 1 ano	44%	9%	4%	5%
de 1 a 2 anos	33%	21%	7%	12%
de 2 a 3 anos	12%	37%	16%	15%
de 4 a 5 anos	3%	17%	28%	19%
de 6 a 10 anos	1%	8%	20%	19%
Mais de 10 anos	0%	3%	19%	25%

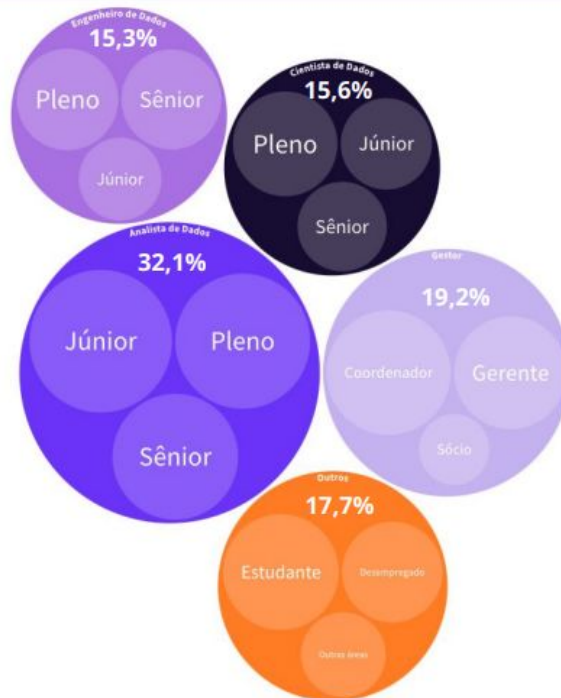
**Nível Júnior:** Entre 1 e 3 anos de experiência em sua maioria.

**Nível Pleno:** Entre 2 e 5 anos de experiência em sua maioria.

**Nível Sênior:** Entre 3 e 10 anos+ de experiência em sua maioria.

**Gestores:** Existem diferentes níveis de senioridade a depender da função exercida, mas a maioria dos entrevistados têm acima de 5 anos de experiência.

As áreas de atuação mais frequentes são DS, DE e DA





# Cargos e Salários

## Salários x Senioridade

Salários dos profissionais de dados no final de 2021 x Senioridade

Faixa salarial	Junior	Pleno	Senior	Gestor
Acima de R\$ 20 mil				
de R\$ 16 mil a R\$ 20 mil				
de R\$ 12 mil a 16 mil				
de R\$ 8 mil a 12 mil				
de R\$ 6 mil a 8 mil				
de R\$ 4 mil a 6 mil				
de R\$ 3 mil a 4 mil				
de R\$ 2 mil a 3 mil				
Até R\$ 2 mil				

\*Obs: As estimativas salariais são aproximações. No momento da pesquisa não coletamos o salário bruto dos participantes, portanto as aproximações foram feitas com base em faixas salariais (presumindo uma distribuição uniforme dentro das faixas).

**Nível Júnior:** Estimamos que 50% dos participantes tem o salário mensal entre R\$ 3 mil e R\$ 6 mil, 25% recebem acima disso.

**Nível Pleno:** Estimamos que 50% dos participantes tem o salário mensal entre R\$ 5 mil e R\$ 10 mil, 25% recebem acima disso.

**Nível Sênior:** Estimamos que 50% dos participantes tem o salário mensal entre R\$ 8 mil e R\$ 12 mil, 25% recebem acima disso.

**Gestores:** Existem diferentes níveis de senioridade a depender da função exercida, mas estimamos que mais de 50% dos participantes tem salário mensal superior a R\$ 12 mil..

Q3 = 75%  
Mediana = 50%  
Q2 = 25%

# Obrigado (a)!



```
<div class="column-8 w-col-6 w-col-stack hero-section content-left">  
  <div class="hero-container heading-container">  
    <div class="heading-tagline container">  
      <h3 class="heading-17 h-subtitle">gama experience</h3>  
      <span class="text-span">  
<div class="hero-section w-section" style="top: 1px !important">  
  <div class="columns-5 w-row hero-section-container">  
    <div class="hero-section-columns">  
      <div class="column-8 w-col-6 w-col-stack hero-section content-left">  
        <div class="hero-container heading-container">  
          <div class="heading-tagline container">  
            <h3 class="heading-17 h-subtitle">gama experience</h3>  
            <span class="text-span">  
<div class="hero-section w-section" style="top: 1px !important">  
  <div class="columns-5 w-row hero-section-container">  
    <div class="hero-section-columns">  
      <div class="column-8 w-col-6 w-col-stack hero-section content-left">
```

