



**Sociedade Brasileira de Computação**

---

**Associação Brasileira de Estatística**

---

Referenciais de Formação para o Curso  
de Bacharelado em Ciência de Dados

**2023**

## Créditos de elaboração

A Sociedade Brasileira de Computação (SBC) em conjunto com a Associação Brasileira de Estatística (ABE) produziram os referenciais de formação para os cursos de Bacharelado em Ciência de Dados com o propósito de utilizar noções de competências definidas pela força tarefa em Ciência de Dados da Association for Computing Machinery (ACM), em 2021.

Este documento foi elaborado pelos seguintes membros da SBC, em ordem alfabética: Adenilso da Silva Simão, Anderson Luiz Ara Souza, André Carlos Ponce de Leon Ferreira de Carvalho, Angelo Roncalli Alencar Brayner, Duncan Dubugras Alcoba Ruiz, Eduardo Ogasawara, Itana Maria de Souza Gímenes, Marcos Oliveira Prates, Maristela Terto de Holanda, Milene Selbach Silveira, Paulo Canas Rodrigues, Rodrigo Duran, Sílvia Amélia Bim, Taciana Pontual da Rocha Falcao e Viviana Giampaoli.

Como citar este documento:

SOCIEDADE BRASILEIRA DE COMPUTAÇÃO; ASSOCIAÇÃO BRASILEIRA DE ESTATÍSTICA. Referenciais de formação para o curso de Bacharelado em Ciência de Dados. Porto Alegre: Sociedade Brasileira de Computação (SBC), 2023. 40p. DOI 10.5753/sbc.ref.2023.126

# Organização

Diretoria de Educação: Itana Maria de Souza Gimenes

Comissão de Educação:

- Adenilso da Silva Simão, Universidade de São Paulo, Instituto de Ciências Matemáticas e de Computação (ICMC-USP)
- Flávio de Oliveira Silva, Universidade Federal de Uberlândia (UFU) ▪ Maristela Terto de Holanda, Universidade de Brasília (UNB)
- Milene Selbach Silveira, Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS)
- Rodrigo Duran, Instituto Federal de Mato Grosso do Sul (IFMS)
- Ronaldo Celso Messias Correia, Universidade Estadual Paulista (UNESP) ▪ Silvia Amélia Bim, Universidade Tecnológica Federal do Paraná (UTFPR)
- Taciana Pontual da Rocha Falcao, Universidade Federal Rural de Pernambuco (UFRPE)

Elaboração pela SBC

- Angelo Roncalli Alencar Brayner (Coordenador), Departamento de Computação (DC) da Universidade Federal do Ceará (UFC); ▪ André Carlos Ponce de Leon Ferreira de Carvalho, Instituto de Ciências Matemáticas e de Computação (ICMC) da Universidade de São Paulo (USP);
- Duncan Dubugras Alcoba Ruiz, da Escola Politécnica (EP), Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS);
- Eduardo Ogasawara, do Departamento de Computação (DC), Centro Federal de Educação Tecnológica Celso Suckow da Fonseca (CEFET/RJ).

Colaboração da ABE

- Anderson Luiz Ara Souza, Departamento de Estatística da Universidade Federal do Paraná (UFPR);
- Marcos Oliveira Prates, Departamento de Estatística da Universidade Federal de Minas Gerais (UFMG);
- Paulo Canas Rodrigues, Departamento de Estatística da Universidade Federal da Bahia (UFBA);
- Viviana Giampaoli, Departamento de Estatística da Universidade de São Paulo (USP).

# Apresentação

A Sociedade Brasileira de Computação (SBC), por meio de sua Diretoria de Educação e Comissão de Educação, promoveu nos eventos Workshop de Educação em Computação (WEI) 2020 e Curso de Qualidade (CQ) 2021 discussões sobre cursos emergentes na área de Computação no Brasil e no mundo. A partir desses eventos, em julho de 2021, iniciou-se o processo de produção de referenciais de formação para cursos emergentes, a saber: Ciência de Dados, Cibersegurança e Inteligência Artificial. São cursos já existentes em outros países e que também já existem no Brasil, porém ainda não fazem parte das Diretrizes Curriculares Nacionais (DCN) para cursos superiores do MEC. A última DCN para a área de computação é de 2016 e contempla os cursos de: Ciência da Computação, Engenharia de Computação, Licenciatura em Computação, Engenharia de Software e Sistemas de Informação. Ao mesmo tempo, a Associação Brasileira de Estatística (ABE) tem discutido o papel da Ciência de Dados perante a Estatística, em especial nos últimos Simpósios Brasileiros de Probabilidade e Estatística (SINAPE) nos anos de 2016, 2018 e 2022. A última DCN para área de Estatística é de 2008 e contempla os cursos de Bacharelado em Estatística.

Assim, a SBC, em parceria com a ABE, entende que é seu compromisso perante a comunidade, produzir um material especializado, sobre os referidos cursos emergentes, para orientar as Instituições de Ensino Superior (IES) sobre as competências e habilidades requeridas para formação de profissionais desses cursos. Os referenciais estão sendo desenvolvidos por subcomissões das Comissões Especiais da SBC em interação com a Diretoria e Comissão de Educação.

Este documento apresenta os referenciais de formação para o curso de Bacharelado em Ciências de Dados. Esses referenciais foram elaborados pela seguinte comissão:

- Angelo Roncalli Alencar Brayner (coordenador), Departamento de Computação (DC) da Universidade Federal do Ceará (UFC);
- André Carlos Ponce de Leon Ferreira de Carvalho, Instituto de Ciências Matemáticas e de Computação (ICMC) da Universidade de São Paulo (USP);
- Duncan Ruiz, da Escola Politécnica (EP), Pontifícia Católica do Rio Grande do Sul (PUCRS); e,
- Eduardo Ogasawara, do Departamento de Computação (DC), Centro Federal de Educação Tecnológica Celso Suckow da Fonseca (CEFET/RJ).

Os referenciais foram revisados pela Comissão de Educação da SBC e submetidos à consulta pública, coordenada pela SBC, no período de 21/09/22 a 24/11/2022. Em seguida, os referenciais foram revisados para atender as sugestões da comunidade.

Concomitantemente à chamada pública, foi criada uma comissão conjunta entre a SBC e a ABE. Essa comissão foi composta pelos membros da SBC supracitados e os seguintes membros da ABE:

- Anderson Luiz Ara Souza, Departamento de Estatística da Universidade Federal do Paraná (UFPR);
- Marcos Oliveira Prates, Departamento de Estatística da Universidade Federal de Minas Gerais (UFMG);

- Paulo Canas Rodrigues, Departamento de Estatística da Universidade Federal da Bahia (UFBA);
- Viviana Giampaoli, Departamento de Estatística da Universidade de São Paulo (USP).

A comissão conjunta reavaliou o documento consolidado e aprovado pela Comissão de Educação da SBC para propor adequações com relação às competências e os conteúdos da área de Ciência de Dados no âmbito da Estatística.

Finalmente, visando aprimorar o alinhamento entre os grupos de trabalho da Sociedade Brasileira de Computação (SBC) e da Associação Brasileira de Educação (ABE), os presidentes das duas entidades, Raimundo Macêdo e Viviana Giampaoli, realizaram várias reuniões, em que foram discutidos os termos expressos no documento conjunto e os trâmites necessários para a sua aprovação nos respectivos conselhos. Essas reuniões foram realizadas para garantir que as duas organizações estivessem em pleno acordo em relação aos termos e diretrizes contidos no documento conjunto.

Itana Maria de Souza Gimenes  
Diretora de Educação

# Sumário

Organização	03
Apresentação	04
Resumo	07
1 Introdução	08
2 Breve histórico do curso de Bacharelado em Ciência de Dados	09
3 Os benefícios do curso para a Sociedade	11
4 Aspectos relacionados com a formação de um profissional em Ciência de Dados	12
5 Perfil do egresso, Competências e Habilidades	13
6 Eixos de formação, competências e conteúdos	15
II.7 Relação com as Diretrizes Curriculares Nacionais	31
II.8 Estágios, TCC e atividades complementares	34
II.8.1 Estágio Supervisionado	34
II.8.2 Trabalho de Conclusão de Curso	34
II.8.3 Atividades complementares	35
II.9 Metodologia de ensino	36
II.10 Requisitos legais	36
II.11 Agradecimentos	38
Referências	39

## Resumo

Este documento apresenta os referenciais de formação para os cursos de Bacharelado em Ciência de Dados. Esses referenciais foram construídos utilizando a noção de competência, em consonância com as competências definidas pela Força Tarefa em Ciência de Dados da Association for Computing Machinery (ACM) em 2021 e com os Referenciais de Formação para os Cursos de Graduação em Computação 2017 produzidos pela Sociedade Brasileira de Computação (SBC). O perfil do egresso e as competências e habilidades foram concebidos e agrupados em 8 (oito) eixos de formação que são relacionados aos conteúdos necessários para desenvolvimento das respectivas competências. Houve a colaboração da Associação Brasileira de Estatística (ABE) para adequar competências e conteúdos às expectativas desta área, em consonância com as orientações do Special Interest Group (SIG) on Data Science da International Statistical Institute (ISI) e da American Statistical Association. Assim, este documento é assinado pela SBC e ABE. Destaca-se ainda que esses referenciais visam a facilitar a construção de projetos pedagógicos de curso nas Instituições de Ensino Superior (IES) nacionais de acordo com seus objetivos, estratégias e vocações.

# 1 Introdução

Os referenciais de formação para os cursos de Bacharelado em Ciência da Dados (RF-CD) estão em consonância com as Diretrizes Curriculares Nacionais (DCN2016), homologadas em novembro de 2016, por meio da Resolução N<sup>o</sup> 05 de 16/11/2016 (MEC, 2016). Conforme o Manual de Classificação de Cursos Superiores, esses cursos fazem parte da área 6 – Computação e Tecnologias da Informação e Comunicação (TIC), área detalhada Soluções computacionais para domínios específicos, rótulo 0617C01 – Ciência de Dados (MEC, 2019).

A metodologia adotada para desenvolvimento dos referenciais segue um modelo baseado em competências e a mesma estrutura e princípios adotados para construção dos Referenciais de Formação para os Cursos de Graduação em Computação 2017 produzidos pela Sociedade Brasileira de Computação (SBC) (SBC, 2017). Assim, adotou-se como referência a Taxonomia de Bloom Revisada (Ferraz e Belhot, 2010). Nesta taxonomia, uma competência pode expressar o conhecimento, as habilidades ou as atitudes esperadas do egresso do curso, sob a perspectiva de objetivos de aprendizagem. Além disso, as competências foram articuladas e estruturadas em eixos temáticos de formação (Anastasiou, 2010).

É importante destacar que a SBC é a favor da liberdade do exercício profissional, sendo o conhecimento técnico-científico e social, normalmente adquirido em curso superior de boa qualidade, o principal diferencial de competência profissional. Nesse sentido, a SBC posiciona-se contra a regulamentação da profissão de Cientista de Dados por um conselho profissional nos moldes tradicionais.

## 2 Breve histórico do curso de Bacharelado em Ciência de Dados

Na virada do século 20 para o século 21, a humanidade testemunhou um crescimento exponencial no volume de dados armazenados em mídias digitais. Esse fenômeno foi impulsionado pelo avanço de tecnologias de geração de dados como redes de sensores sem fio, Internet das coisas (IoT), entre outras.

Naquele momento, a Ciência da Computação sentiu-se impelida a prover recursos tecnológicos, em nível de hardware e software, que fossem capazes de garantir o gerenciamento daquele grande volume de dados de forma segura e eficiente, surgindo assim a tecnologia de Big Data. Concomitantemente, para a área de Estatística, e sob a liderança a nível mundial da American Statistical Association, observou-se o enorme potencial do uso de um grande número de métodos estatísticos para extrair informações importantes a partir dos dados existentes, bem como para criar modelo de predição e inferência estatística, tendo sido criadas diretrizes específicas para a área de Ciência de Dados em 2014 amplamente discutidas no Curriculum Guidelines for Undergraduate Programs in Statistical Science e que culminou em um número especial do periódico American Statistician, publicado em novembro de 2015 e intitulado “Statistics and the undergraduate curriculum” e em um evento organizado em fevereiro de 2015 intitulado “The increasing role of data science in undergraduate statistics”.

Esse novo cenário fez com que se percebesse a necessidade de uma nova área da computação, que abarcasse profissionais com conhecimentos sólidos em Computação e Estatística. Assim, surgiu em escala global a Ciência de Dados. Conseqüentemente, várias iniciativas surgiram em diferentes países, como Inglaterra, Alemanha, EUA, Canadá, entre outros, com o intuito de criar cursos de graduação para formar Cientistas de Dados. Por exemplo, já em 2014, a Universidade da Califórnia, em Berkeley, deu início a um curso de graduação em Ciência de Dados, formando sua primeira turma em 2014.

No Brasil, as primeiras iniciativas em formar cientistas de dados começam a surgir em 2020, a partir de duas abordagens distintas, além das atualizações das grades curriculares dos cursos do Bacharelado em Estatística. Uma abordagem é a de adotar uma ênfase em Ciência de Dados dentro dos cursos de Estatística. Essa abordagem foi utilizada, por exemplo, em: (1) Instituto de Matemática da UFRJ, e no Instituto de Matemática, Estatística e Computação Científica da UNICAMP. A segunda abordagem é a de criar cursos com a denominação de Bacharelado em Ciência de Dados ou similares. Essa abordagem foi adotada: (1) pela Universidade Anhembi Morumbi e pela PUC-SP, na criação de bacharelado em Estatística e Ciência de Dados, (2) Na Universidade Municipal de São Caetano do Sul (USCS); (3) pela PUCRS, na criação, em 2020 de bacharelado em Ciência de Dados e Inteligência Artificial, já em seu 3o ano de implantação, (4) pela UFC, com a criação do curso em Ciência de Dados pelos Departamentos de Computação e de Estatística, (5) pela USP, com a criação do curso em Ciência de Dados pelos Departamentos de Ciências de Computação, de Matemática, de Matemática Aplicada e Estatística e de Sistemas de Computação.

Para ambos os casos, encontram-se Programas de Pós-graduação *stricto sensu* com foco em Ciência de Dados. Alguns precursores como o da Estatística do IME-USP de 1970, outros mais recentes como é da área de Ciência da Computação da CAPES, por exemplo, abordagem adotada pelo CEFET/RJ desde 2016. Em termos de formação *latu sensu*, as primeiras especializações em Ciência de Dados foram criadas em 2018 nos Departamentos de Estatística da Universidade Federal da Bahia e da Universidade Federal do Paraná.

Devido ao aumento na criação de cursos autodenominados Ciência de Dados, o INEP resolveu atualizar o CINE (Classificação Internacional Normalizada da Educação) no Brasil. Esta decisão foi tomada após os trabalhos desenvolvidos por um grupo de especialistas, com membros da Sociedade Brasileira de Computação e por meio do Ofício Circular No. 0714111/2021/CGCES/DEED-INEP, classificando cursos, com a denominação Ciência de Dados, na Área 6 (Computação e Tecnologia da Informação e Comunicação), Área Detalhada 0617 (Desenvolvimento de Soluções Computacionais em Domínios Específicos). Assim, cursos de Ciência de Dados podem ter seus Projetos Pedagógicos de Curso (PPC) balizados referenciais curriculares apresentados neste documento.

### **3 Os benefícios do curso para a Sociedade**

A quantidade e complexidade de dados gerados por seres humanos e máquinas cresce a uma velocidade cada vez maior. Esses dados contêm informações relevantes que podem resultar em avanços não apenas em diversas áreas do conhecimento, mas também para oferta de melhores serviços públicos, desenvolvimento social, crescimento econômico e aumento da oferta de empregos. A atuação de profissionais na área de Ciência de Dados pode reduzir riscos não só para os seres humanos, mas também para todo o planeta, assim facilitando a preservação e recuperação do meio ambiente e a ética nas relações humanas. Para que isso ocorra, é importante a formação de recursos humanos em toda a cadeia de análise de dados com sólido conhecimento nas áreas de Computação, Estatística e Matemática. A boa formação desses profissionais permitirá o desenvolvimento de ferramentas computacionais capazes de explorar de uma forma competente e eficiente os dados gerados e apoiar o processo de tomada de decisão nos setores de Ciências, Governo e Empresas.

## **4 Aspectos relacionados com a formação de um profissional em Ciência de Dados**

O Bacharel em Ciência de Dados deve ser capaz de compreender os processos para coleta, gerenciamento e análise eficientes de dados em larga escala, em diferentes formatos e de fontes heterogêneas. Devem ser capazes de definir e executar procedimentos de gestão e uso de dados, bem como desenvolver estratégias para formular e testar hipóteses, interpretar e extrair conhecimentos úteis e significativos, desenvolver e validar algoritmos para analisá-los, interagir com profissionais de outras áreas de conhecimento e trabalhar de forma colaborativa.

O Bacharel em Ciência de Dados, também denominado Cientista de Dados, possui uma vasta área de atuação. Cientistas de Dados poderão analisar dados de áreas tão diversas como administração, agronomia, ciências florestais, ciências sociais, economia, engenharia, geografia, história, medicina e veterinária, para citar algumas. Sua atuação deve se dar de forma colaborativa, com profissionais de outras áreas do conhecimento, que entendem o significado dos dados por ele analisados.

## 5 Perfil do egresso, Competências e Habilidades

Nesta seção, são apresentados o perfil dos egressos do curso de Bacharelado em Ciência de Dados e suas respectivas competências e habilidades específicas. Esses elementos adicionam características próprias da área de ciência de dados ao perfil, competências e habilidades estabelecidos nos Art. 4º. e Art. 5º. para a área de Computação das DCN2016.

O curso de Bacharelado em Ciência de Dados visa formar profissionais capazes de “pensar com dados”. Esses profissionais devem possuir competência teórica, técnica e metodológica, bem como experiência prática para lidar com as mais variadas situações e domínios de aplicação. Espera-se que os egressos do curso de Bacharelado em Ciência de Dados sejam capazes de:

- I. entender, formular e refinar as questões apropriadas no processo de análise de dados;
- II. obter, modelar e explorar os dados relacionados;
- III. processar os dados e realizar as análises necessárias;
- IV. obter e comunicar o conhecimento relevante da área;
- V. apoiar o desenvolvimento e implantação de soluções com base nos resultados atingidos; e,
- VI. entender e atender aspectos éticos e sociais relacionados à sua atuação.

Levando em consideração a flexibilidade necessária para atender domínios diversificados de aplicação e as vocações institucionais, os egressos dos cursos de bacharelado em ciência de dados devem apresentar as seguintes habilidades e competências (CE):

- CE-I Possuir sólida formação nas áreas de Computação, matemática e estatística que permita a aplicações de conceitos dessas áreas em tarefas de ciência de dados;
- CE-II Utilizar efetivamente técnicas computacionais, matemáticas e estatísticas para, de forma analítica, avaliar a factibilidade e, quando factível, extrair conhecimento dos dados disponíveis para realizar descobertas em diferentes domínios de aplicação, de forma a apoiar o processo de tomada de decisão;
- CE-III Empregar os princípios de Tecnologia de Informação e Comunicação para pesquisar, projetar, implementar e avaliar novas abordagens e técnicas para construção de ferramentas para análise de dados;
- CE-IV Ser capaz de planejar e realizar experimentos computacionais simulados utilizando diferentes infraestruturas que apoiem a gestão e o manuseio eficiente de dados, estruturados e não estruturados, durante todo o ciclo de vida dos dados;
- CE-V Definir e implementar estratégias de gerenciamento de dados para curadoria, coleta, integração, armazenamento, visualização, preservação e disponibilização de dados para futuro processamento;

- CE-VI Gerenciar projetos interdisciplinares que incluam as diversas etapas do ciclo de vida de dados;
- CE-VII Identificar novos desafios, necessidades, oportunidades de negócios e desenvolver soluções inovadoras;
- CE-VIII Investigar, compreender e estruturar as características de domínios de aplicação em diversos contextos que levem em consideração questões ambientais, éticas, sociais, legais e econômicas;
- CE-IX Ser capaz de trabalhar de forma tanto individual como colaborativa, com profissionais da mesma área ou de diferentes áreas;
- CE-X Seguir os princípios de uma ciência de dados justa, transparente, sem viés, respeitando a privacidade e atendendo aos requisitos da Lei Geral de Proteção de Dados (LGPD);
- CE-XI Ter uma visão crítica e criativa na identificação e resolução de problemas contribuindo para o desenvolvimento da área;
- CE-XII Atuar de forma empreendedora, abrangente e cooperativa no atendimento às demandas sociais da região onde atua, no Brasil e no mundo;
- CE-XIII Utilizar racionalmente os recursos disponíveis de forma transdisciplinar; e,
- CE-XIV Atuar em um mundo globalizado do trabalho, buscando o domínio de idiomas estrangeiros, em particular o idioma inglês.

Com uma formação sólida, o profissional formado no curso de bacharelado em ciência de dados está apto a trabalhar em empresas de setores diversos, em órgãos públicos e do terceiro setor, além de IES e em institutos de pesquisa.

## 6 Eixos de formação, competências e conteúdos

As 14 competências e habilidades, acima definidas, foram agrupadas em 8 (oito) eixos de formação. Cada eixo de formação corresponde a uma macro competência que agrupa um conjunto de competências derivadas correlacionadas. Este grupo de competências derivadas levará o egresso a atingir a competência do eixo. Adicionalmente, elas possibilitam o egresso de um Bacharelado em Ciência de Dados a lidar profissionalmente com as várias facetas das atividades de computação. Os eixos de formação traduzem o entendimento de que tal formação deve levar em conta: a capacidade de atuar em todas as fases que envolvem a aplicação da ciência de dados em problemas diversos, desde a concepção de sistemas computacionais até a efetiva implementação de soluções adequadas; a capacidade de se reciclar e buscar novos conhecimentos; e, a capacidade de seguir estudos avançados visando ao desenvolvimento da ciência e da tecnologia. Os eixos de formação são os seguintes:

1. Fundamentos de Matemática, Estatística e Computação para Ciência de Dados
2. Modelagem Estatística e Resolução de Problemas
3. Desenvolvimento de Sistemas
4. Engenharia e Exploração de Dados
5. Big Data
6. Mineração de Dados e Aprendizado de Máquina
7. Aprendizado Contínuo e Autônomo
8. Ciência, Tecnologia, Inovação e Empreendedorismo

Neste documento, cada eixo de formação será apresentado com a seguinte estrutura:

- Código: algarismo indo-arábico que identifica o eixo de formação.
- Título: rótulo que identifica o eixo de formação.
- Descrição: pequeno texto que contextualiza a competência associada ao eixo de formação.
- Competência de eixo: descrição da competência associada ao eixo de formação.
- Competências derivadas: lista de competências, oriundas das competências e habilidades, gerais da área de computação e específicas, definidas para a área de ciência de dados. As competências gerais são indicadas pelo identificador CG e as específicas do curso de Bacharelado em Ciência de Dados, pelo identificador CE. Por sua vez, cada competência derivada é constituída dos seguintes subcampos:
  - Código: é formado pela junção da letra C (inicial da palavra “competência”), do código do eixo (1 a 8) e de um número indo-arábico que ordena sequencialmente a competência derivada no contexto do eixo de formação.
  - Classificação: um dos seis níveis do processo cognitivo da Taxonomia de Bloom Revisada (Ferraz e Belhot, 2010).

○ Conteúdo: lista de conhecimentos que devem ser trabalhados para desenvolver a competência derivada. Cada conteúdo é definido por um título, em geral, oriundo da listagem constante nas seções 3.1 e 3.2 do Parecer CNE/CSE 136/2012 (MEC, 2012), assim como foram utilizados nos referenciais da SBC para a área de computação 2017 (SBC, 2017). Porém, conteúdos próprios da área de ciência de dados e da estatística foram utilizados.

Salienta-se que conteúdo e disciplina (ou componente curricular) não são sinônimos. De fato, a associação entre conteúdos e disciplinas é um dos grandes desafios na elaboração de um projeto pedagógico de curso e sua matriz curricular. Uma disciplina oferecida por uma Instituição de Ensino Superior (IES) em particular poderá abordar mais de um conteúdo elencado nestes referenciais, combinando-os a fim de tratar situações complexas. Ao mesmo tempo, certo conteúdo poderá ser abordado em mais de uma disciplina, evidenciando a sua aplicação em diferentes contextos, possivelmente com diferentes níveis de profundidade. Esta associação de conteúdos e disciplinas dependerá, essencialmente, da estratégia adotada por cada curso para o desenvolvimento de seu projeto pedagógico.

Deve-se enfatizar que uma competência específica pode requerer diferentes conteúdos, dependendo do eixo de formação ao qual ela está associada. Ainda, um conteúdo pode estar presente em mais de um eixo. Por último, um conteúdo pode estar presente em mais de uma competência específica de certo eixo.

Um projeto pedagógico de curso pode usar uma estratégia para implementar sua matriz curricular tal que cada disciplina seja desenhada para oferecer ao egresso uma ou mais competências específicas, no contexto de um ou mais eixos de formação. Assim, cada disciplina deverá abordar (integral ou parcialmente) os conteúdos recomendados para as respectivas competências específicas, de acordo com eixos de formação em questão.

A seguir, cada eixo de formação é detalhado em termos de suas competências derivadas e conteúdos associados.

1. EIXO DE FORMAÇÃO: FUNDAMENTOS DE MATEMÁTICA, ESTATÍSTICA E COMPUTAÇÃO PARA CIÊNCIA DE DADOS		
Competência geral esperada para o eixo: entender teorias e princípios básicos das áreas de Computação, Estatística e Matemática para aplicá-los na resolução de problemas de Ciência de Dados, incluindo sistemas de domínios específicos.		
Competências derivadas	Classificação	Conteúdos
<b>C.1.1.</b> Dominar os principais conceitos de programação.	<b>Aplicar</b>	Algoritmos
		Linguagem de Programação
		Estruturas de dados

<b>C.1.2.</b> Compreender e saber usar os principais temas em ciência da computação.	<b>Aplicar</b>	Análise de Algoritmos
		Teoria dos Grafos
		Estruturas de dados
		Lógica Matemática
		Indução Matemática
		Linguagens Formais e Autômatos Finitos
<b>C.1.3.</b> Compreender e saber usar os principais temas de estatística introdutória	<b>Aplicar</b>	Estatística Descritiva
		Experimentos Simulados
		Inferência Estatística
		Probabilidade
		Indução Matemática
		Modelagem Paramétrica
<b>C.1.4.</b> Compreender e saber usar os principais temas de matemática básicos para um curso de Ciência de Dados	<b>Aplicar</b>	Análise Combinatória
		Matemática Discreta
		Cálculo Diferencial e Integral
		Geometria Analítica
		Álgebra Linear e Matricial
		Cálculo Numérico

## 2. EIXO DE FORMAÇÃO: MODELAGEM ESTATÍSTICA E RESOLUÇÃO DE PROBLEMAS

A resolução de problemas por meio da modelagem estatística e computacional é possível com a execução de passos finitos e bem definidos. Nesse sentido, os egressos devem ser “capazes de criar soluções, individualmente ou em equipe, para problemas complexos caracterizados por relações entre domínios de conhecimento e de aplicação”.

**COMPETÊNCIA:** entender, implementar e aplicar técnicas de análise e modelagem estatística para resolver problemas que tenham solução analítica, considerando os limites da Ciência de Dados, o que inclui:

- Identificar os problemas que apresentem soluções estatísticas e algorítmicas viáveis.
- Identificar métodos de coleta de dados;
- Elaborar sínteses numéricas e gráficas de dados;
- Compreender e aplicar técnicas paramétricas e não paramétricas de análise de dados univariados e multivariados;
- Selecionar ou criar algoritmos apropriados para situações particulares.
- Avaliar técnicas de análise e de modelagem estatística
- Implementar a solução usando o paradigma de programação adequado.

Competências derivadas	Classificação	Conteúdos
<b>C.2.1.</b> Reconhecer a importância social da Ciência de Dados	<b>Aplicar</b>	Sustentabilidade
		Ética em Ciência de Dados, Privacidade, Legislação e Direitos Civis
		Projetos de Extensão Comunitários
<b>C.2.2.</b> Avaliar criticamente e comunicar a resolução dos problemas, reconhecendo a importância científica e econômica da Ciência de Dados.	<b>Aplicar</b>	Empreendedorismo
		Fundamentos de Administração
		Metodologia Científica
		Fundamentos de Economia
		Gerenciamento de Projetos
		Comunicação Profissional e Científica
		Propriedade Intelectual
<b>C.2.3.</b> Conhecer técnicas de coleta, síntese e visualização de dados	<b>Aplicar</b>	Conceitos básicos de amostragem
		Estatística Descritiva
		Visualização de Dados
<b>C.2.4.</b> Compreender aspectos fundamentais de temas relacionados e relevantes para a Ciência de Dados	<b>Aplicar</b>	Aprendizado de máquina
		Probabilidade

		Inferência Estatística
		Modelagem Paramétrica
		Engenharia e Exploração de Dados
		Pesquisa Operacional, Otimização e Meta-heurísticas
<b>C.2.5.</b> Ter domínio de técnicas de modelagem estatística para Ciência de Dados	<b>Aplicar</b>	Modelagem Paramétrica
		Estatística Multivariada
		Modelagem Não paramétrica
<b>C.2.6.</b> Ser capaz de identificar soluções estatísticas adequadas para o contexto da aplicação	<b>Avaliar</b>	Modelagem Paramétrica
		Estatística Multivariada
		Previsão de séries temporais
<b>C.2.7.</b> Aplicar temas relevantes para o projeto de soluções eficientes de Ciência de Dados	<b>Aplicar</b>	Sistemas Paralelos e Distribuídos
		Computação de Alto Desempenho
		Segurança de Sistemas Computacionais
		Arquitetura e Organização de Computadores

### 3. EIXO DE FORMAÇÃO: DESENVOLVIMENTO DE SISTEMAS

O desenvolvimento de sistemas computacionais inclui tanto a criação de sistemas quanto a adaptação de sistemas existentes, em especial aqueles que apresentam análise de dados. Deve contemplar o levantamento de requisitos funcionais e não-funcionais, a sua análise, modelagem, projeto, implementação e teste. Em todo o processo de desenvolvimento de sistemas computacionais deve-se empregar teorias, métodos, técnicas e ferramentas para garantia e controle de qualidade do processo e do produto e identificar potenciais vulnerabilidades.

**COMPETÊNCIA:** desenvolver sistemas computacionais que atendam qualidade de processo e de produto, considerando princípios e boas práticas de engenharia de sistemas e engenharia de software, incluindo:

- Identificar, analisar, especificar, validar requisitos.
- Projetar soluções computacionais em harmonia com o ambiente social e físico no seu entorno de aplicação.
- Implementar sistemas computacionais utilizando ambientes de desenvolvimento apropriados.
- Testar e manter sistemas computacionais.
- Identificar e analisar potenciais vulnerabilidades.

Competências derivadas	Classificação	Conteúdos
<b>C.3.1.</b> Resolver problemas usando ambientes de programação	<b>Aplicar</b>	Algoritmos
		Estatística
		Programação
		Banco de Dados
		Interação Humano-Computador
		Processamento Paralelo e Distribuído
<b>C.3.2.</b> Tomar decisões e inovar, com base no conhecimento do funcionamento e das características técnicas de hardware e da infraestrutura de software dos sistemas de computação, consciente dos aspectos éticos, legais e dos impactos ambientais decorrentes	<b>Aplicar</b>	Arquitetura e Organização de Computadores
		Banco de Dados
		Segurança de Sistemas Computacionais
		Computação em Nuvem
		Legislação de Computação, inteligência artificial e de proteção de dados
		Engenharia de Software
		Sustentabilidade
		Ciência de Dados, Ética e Sociedade
<b>C.3.3.</b> Avaliar criticamente projetos de sistemas de computação	<b>Avaliar</b>	Engenharia de Software
		Métodos Quantitativos em Computação

		Avaliação de Desempenho de Sistemas
		Confiabilidade e Segurança de Sistemas
		Sistemas Paralelos e Distribuídos
<b>C.3.4.</b> Ser capaz de realizar trabalho cooperativo e entender os benefícios que este pode produzir	<b>Aplicar</b>	Gerenciamento de Projetos
		Comunicação Profissional
		Ciência de Dados, Ética e Sociedade
		Comportamento Humano nas Organizações
<b>C.3.5.</b> Compreender os fatos essenciais, os conceitos, os princípios e as teorias relacionadas à Ciência da Computação para o desenvolvimento de software e hardware e suas aplicações	<b>Aplicar</b>	Algoritmos
		Matemática Discreta
		Arquitetura e Organização de Computadores
		Teoria da Computação
		Teoria dos Grafos
<b>C.3.6.</b> Identificar e analisar requisitos e especificações para problemas específicos e planejar estratégias para suas soluções	<b>Aplicar</b>	Engenharia de Software
<b>C.3.7.</b> Empregar metodologias que visem garantir critérios de qualidade ao longo de todas as etapas de desenvolvimento de uma solução computacional	<b>Aplicar</b>	Engenharia de Software
<b>C.3.8.</b> Analisar quanto um sistema baseado em computadores atende os critérios definidos para seu uso corrente e futuro (adequabilidade)	<b>Avaliar</b>	Engenharia de Software
<b>C.3.9.</b> Aplicar os princípios de gerência, organização e recuperação da informação de vários tipos, incluindo texto imagem som e vídeo	<b>Aplicar</b>	Estruturas de Dados
		Banco de Dados
		Busca e Recuperação da Informação
		Inteligência Artificial

		Sistemas Multimídia
		Inteligência Computacional
		Processamento de Imagens
<b>C.3.10.</b> Aplicar os princípios de interação humano-computador para avaliar e construir uma grande variedade de produtos incluindo interface do usuário, modelos cognitivos, sistemas de apoio a visualização, escalas, cores, tipos de gráficos, mapas, páginas WEB, sistemas multimídia e sistemas móveis	<b>Criar</b>	Interação Humano-Computador
		Visualização de dados
		Sistemas Multimídia

#### 4. EIXO DE FORMAÇÃO: ENGENHARIA E EXPLORAÇÃO DE DADOS

Na formação de um cientista de dados, a compreensão de como os dados podem ser produzidos, armazenados e gerenciados é de fundamental importância. Nesse sentido, o estudo de sistemas gerenciadores de banco de dados proverá ao egresso do curso de Ciência de Dados a fundamentação teórica e prática de como grandes volumes de dados são manipulados. Para isto, o entendimento do funcionamento de diversos sistemas de bancos de dados dará solidez à formação de cientistas de dados. A governança de dados deve embasar o egresso na habilidade de trabalhar aspectos de qualidade, privacidade e curadoria dos dados. Nessa linha de estudo, a proveniência de dados e LGPD devem estar presentes.

**COMPETÊNCIA:** desenvolver e implementar projetos lógico e físico de banco de dados, independente do modelo de dados, garantindo-se o nível de privacidade de dados esperado, bem como identificar gargalos e soluções para melhorar o acesso a bancos de dados, quanto ao tempo de resposta e aumento da vazão do sistema quanto ao volume de dados providos por unidade de tempo. Torna-se necessário, então, o desenvolvimento das seguintes competências específicas:

- Utilizar conceitos, métodos e ferramentas de modelagem conceitual e física de bancos de dados, visualização e privacidade de dados, com o objetivo de garantir o cumprimento dos requisitos especificados quanto à qualidade, desempenho e nível de privacidade dos dados.
- Interagir com pessoas de diferentes perfis, de diversas áreas do conhecimento, incluindo especialistas do domínio da aplicação, estatístico e, engenheiros de dados.
- Realizar ações de prospecção na busca de soluções mais eficazes, incluindo novas tecnologias, produtos e serviços.
- Adequar-se rapidamente às mudanças tecnológicas e aos novos ambientes de trabalho.

Competências derivadas	Classificação	Conteúdos
<b>C.4.1.</b> Compreender especificação de requisitos de dados e construir modelos conceituais de dados	<b>Aplicar</b>	Banco de Dados
		Sistemas de Dados
		Integração de Dados
<b>C.4.2.</b> Avaliar performance de sistemas de dados	<b>Aplicar</b>	Ajuste de Performance de Sistemas de Bancos de Dados
		Sistemas de Dados
		Estrutura de Dados
		Sistemas Paralelos e Distribuídos
<b>C.4.3.</b> Conhecer técnicas de visualização de dados	<b>Aplicar</b>	Visualização de dados
		Estatística Descritiva
		Privacidade de dados
<b>C.4.4.</b> Ter domínio de técnicas de privacidade de dado de acordo com a LGPD	<b>Aplicar</b>	Privacidade de dados
<b>C.4.5.</b> Ser capaz de identificar soluções tecnológicas adequadas para o domínio da aplicação	<b>Avaliar</b>	Banco de Dados
		Ajuste de Performance de Sistemas de Dados
		Integração de Dados
		Visualização de dados
		Privacidade de dados
		Segurança de Sistemas Computacionais
		Sistemas Paralelos e Distribuídos

## 5. EIXO DE FORMAÇÃO: BIG DATA

A implantação de soluções de Big Data compreende a instalação dos sistemas computacionais (desenvolvidos ou adquiridos) no ambiente alvo, podendo envolver a integração de sistemas computacionais, a adequação de infraestrutura, garantia das regras de negócio (requisitos funcionais) e das regras sistêmicas (desempenho, contingência, confiabilidade, segurança), e das políticas internas e externas legais.

**COMPETÊNCIA:** Implantar soluções de Big Data, considerando:

- Planejar e executar o processo de implantação de sistemas computacionais baseados no paradigma de programação e gerência de dados em larga escala.
- Identificar os problemas que demandem soluções algorítmicas escaláveis.
- Selecionar ou criar algoritmos escaláveis que abordem armazenamento de dados, computação de alto desempenho e teoria da complexidade
- Prover capacitação das pessoas envolvidas (técnicos e usuários), considerando a documentação e a operacionalização do sistema computacional.
- Garantir a consistência da implementação com as normas legais e éticas da comunidade envolvida.

Competências derivadas	Classificação	Conteúdos
<b>C.5.1.</b> Tomar decisões e inovar, com base no conhecimento do funcionamento e das características técnicas de hardware e da infraestrutura de software dos sistemas de computação consciente dos aspectos éticos, legais e dos impactos ambientais decorrentes	<b>Aplicar</b>	Arquitetura e Organização de Computadores
		Redes de Computadores
		Sistemas Operacionais
		Segurança de Sistemas Computacionais
		Banco de Dados
		Sistemas Paralelos e Distribuídos
		Privacidade de Dados
		Legislação de Computação, inteligência artificial e de proteção de dados
		Sustentabilidade
<b>C.5.2.</b> Compreender e explicar as dimensões quantitativas de um problema	<b>Aplicar</b>	Estatística
		Métodos Quantitativos em Computação

<b>C.5.3.</b> Conceber e preparar soluções computacionais de larga escala	<b>Aplicar</b>	Arquitetura e Organização de Computadores
		Processamento Paralelo e Distribuído
		Computação em Nuvem
<b>C.5.4.</b> Planejar soluções de gerência e armazenamento de dados distribuídos	<b>Aplicar</b>	Banco de Dados
		Sistemas de Dados
		Integração de Dados
<b>C.5.5.</b> Preparar e apresentar trabalhos e problemas técnicos e suas soluções para audiências diversas, em formatos apropriados (oral e escrito)	<b>Aplicar</b>	Visualização de dados
		Comunicação Profissional
<b>C.5.6.</b> Avaliar criticamente projetos de sistemas de computação	<b>Avaliar</b>	Engenharia de Software
		Gerência de Projetos
		Avaliação de Desempenho de Sistemas
<b>C.5.7.</b> Ser capaz de trabalhar de forma colaborativa e entender os benefícios que este pode produzir	<b>Aplicar</b>	Fundamentos de Administração
		Comportamento Humano nas Organizações
		Comunicação Profissional
		Ciência de Dados, Ética e Sociedade
<b>C.5.8.</b> Conceber soluções computacionais a partir de decisões visando o equilíbrio de todos os fatores	<b>Criar</b>	Engenharia de Software
		Simulação de Sistemas
		Privacidade de Dados
		Sustentabilidade
		Empreendedorismo

C.5.9. Empregar metodologias que visem garantir critérios de qualidade ao longo de todas as etapas de desenvolvimento de uma solução computacional	Aplicar	Engenharia de Software
		Gerência de Projetos
		Teste, Verificação e Validação de Software

## 6. EIXO DE FORMAÇÃO: MINERAÇÃO DE DADOS E APRENDIZADO DE MÁQUINA

Um dos aspectos centrais na formação de bons cientistas de dados é como saber usar e desenvolver técnicas para extrair conhecimento útil de um conjunto de dados e como criar modelos que capturam as principais características presentes nos dados, sejam os modelos descritivos, preditivos ou prescritivos. Para isso, deverá conhecer e selecionar as diferentes técnicas e algoritmos utilizados nos processos de mineração de dados e de aprendizado de máquina para um conjunto de dados.

**COMPETÊNCIA:** Analisar, compreender, aplicar e desenvolver técnicas e algoritmos para mineração de dados e aprendizado de máquina.

Competências derivadas	Classificação	Conteúdos
C.6.1. Avaliar e tratar a qualidade dos dados	Aplicar	Estatística Descritiva
		Banco de Dados
		Visualização de Dados
		Pré-processamento de Dados
C.6.2. Compreender e explicar as dimensões quantitativas de um problema	Aplicar	Estatística Descritiva
		Métodos Quantitativos em Computação
		Visualização de Dados
C.6.3. Planejar experimentos simulados	Projetar	Simulação computacional
		Fundamentos de Programação
		Visualização de Dados
C.6.4. Avaliar necessidade e aplicar técnicas de pré-processamento de dados	Aplicar	Estatística Descritiva
		Fundamentos de Programação
		Introdução a Ciência de Dados

		Visualização de Dados
		Modelagem Não paramétrica
		Mineração de Dados Aprendizado de Máquina
		Álgebra Linear e Matricial
<b>C.6.5.</b> Identificar e analisar requisitos e especificações para problemas específicos e planejar estratégias para suas soluções	<b>Criar</b>	Introdução à Ciência de Dados
		Inferência Estatística
		Projeto de Sistemas Computacionais
<b>C.6.6.</b> Especificar, projetar, implementar, manter e avaliar sistemas de computação, empregando teorias, práticas e ferramentas adequadas	<b>Criar</b>	Projeto de Sistemas Computacionais
		Aprendizado de Máquina
		Modelagem Paramétrica
		Modelagem Não paramétrica
		Previsão de séries temporais
		Avaliação de Desempenho de Sistemas
		Banco de Dados
		Processamento Paralelo e Distribuído
		Computação em Nuvem
		Redes Neurais
		Internet das Coisas (IoT)
<b>C.6.7.</b> Conceber soluções eticamente e socialmente responsáveis	<b>Aplicar</b>	Sustentabilidade
		Ciência de Dados, Ética e Sociedade
		Projetos Comunitários
<b>C.6.8.</b> Aplicar os princípios de gerência, organização e recuperação da informação de vários tipos, incluindo texto, imagem, som e vídeo	<b>Aplicar</b>	Banco de Dados
		Recuperação da Informação
		Sistemas Multimídia

		Processamento da Língua Natural
		Processamento de Imagens
		Processamento Paralelo e Distribuído

## 7. EIXO DE FORMAÇÃO: APRENDIZADO CONTÍNUO E AUTÔNOMO

A teoria e prática da Ciência de Dados estão em permanente evolução, levando a: i. ao surgimento de novos instrumentos (processos, métodos e ferramentas) que visam à melhoria do gerenciamento e da qualidade dos dados, além das informações e conhecimentos extraídos deles; ii, ao aperfeiçoamento de instrumentos existentes; iii. ao surgimento de novas tecnologias de infraestrutura computacional; e, iv. à expansão dos domínios de aplicação da ciência de dados.

Este eixo de formação em especial agrupa competências orientadas ao desenvolvimento pessoal (habilidades e atitudes), em vez da assimilação de conteúdos tradicionais. Por exemplo, é possível apresentar o conteúdo "auto-regulação da aprendizagem" em forma de palestras extra-curriculares ou de aula em alguma disciplina da matriz curricular. Porém, essa habilidade será plenamente desenvolvida se for estimulada transversalmente à apresentação de conteúdos técnicos. Os docentes devem ser estimulados a adotarem abordagens pedagógicas para promover as competências deste eixo de formação, tais como (mas não exclusivamente), aprendizagem colaborativa, aprendizagem baseada em projetos (project-based learning – PBL), aprendizagem ativa (active learning), ensino híbrido (blended), entre outras.

**COMPETÊNCIA:** aprender contínua e autonomamente sobre métodos, instrumentos, tecnologias de ciência de dados e de seus domínios de aplicação, abrangendo:

- Desenvolver estudos para manter-se atualizado sobre a evolução da ciência de dados, além de desenvolver pesquisas que contribuam para essa evolução. Essa atividade de estudo inclui o acompanhamento de publicações científicas e comerciais em periódicos e eventos especializados, bem como a participação em cursos de treinamento, extensão, especialização, mestrado e doutorado.
- Avaliar novos instrumentos e métodos, bem como a viabilidade de seu emprego no seu contexto de trabalho e, eventualmente, propor novos instrumentos e métodos a partir da sua própria experiência.
- Avaliar novas tecnologias computacionais e a sua adequação à ciência de dados.
- Adequar-se rapidamente às mudanças tecnológicas e aos novos ambientes de trabalho (CG-IX).

Competências derivadas	Classificação	Conteúdos
<b>C.7.1.</b> Gerir a sua própria aprendizagem e desenvolvimento, incluindo a gestão de tempo e competências organizacionais	<b>Criar</b>	Gestão de Tempo
		Gestão de Carreira Profissional
		Autorregulação da Aprendizagem
<b>C.7.2.</b> Ser capaz de realizar trabalho cooperativo e entender os benefícios que este pode produzir	<b>Aplicar</b>	Aplicável a todos os conteúdos, utilizando práticas pedagógicas colaborativas
<b>C.7.3.</b> Compreender os fatos essenciais, os conceitos, os princípios e as teorias relacionadas à Ciência de Dados e suas aplicações	<b>Avaliar</b>	Algoritmos
		Aprendizado de Máquina
		Teoria da Computação
		Modelagem Paramétrica
		Estatística Multivariada
		Modelagem Não paramétrica
		Mineração de Dados
<b>C.7.4.</b> Analisar quanto um sistema baseado em computadores atende os critérios definidos para seu uso corrente e futuro (adequabilidade)	<b>Avaliar</b>	Métodos Quantitativos em Computação
		Modelagem de Sistemas
		Banco de Dados
		Simulação de Sistemas

## 8. EIXO DE FORMAÇÃO: CIÊNCIA, TECNOLOGIA E INOVAÇÃO

Um curso de Bacharelado em Ciência de Dados deve fornecer aos seus egressos uma base teórica sólida que os permita desenvolver estudos avançados e prepará-los para grandes desafios.

A inovação em ciência de dados exige conhecimentos científicos e tecnológicos que vão além dos necessários para suas aplicações tradicionais. Além disso, a formação do egresso deve levar em conta a cultura das pessoas envolvidas, as oportunidades do mercado e as necessidades da sociedade.

**COMPETÊNCIA:** Desenvolver estudos avançados visando o desenvolvimento científico e tecnológico da ciência de dados e a criação de soluções computacionais inovadoras para problemas em qualquer domínio de conhecimento, abrangendo:

- Entender os fundamentos científicos e tecnológicos da computação, em profundidade, relacionados à ciência de dados.
- Dominar ferramentas matemáticas e estatísticas necessárias para a pesquisa e desenvolvimento de produtos e processos em ciência de dados.
- Adaptar-se a novos domínios de aplicação, que envolvam contextos particulares e novas tecnologias.
- Realizar ações inovadoras na busca de soluções mais eficazes, incluindo novos produtos e processos
- Adequar-se rapidamente às mudanças tecnológicas e aos novos ambientes de trabalho.

Competências derivadas	Classificação	Conteúdos
<b>C.8.1.</b> Identificar problemas que tenham solução algorítmica	<b>Aplicar</b>	Algoritmos
		Teoria da Computação
		Complexidade de Algoritmos
		Simulação Computacional
<b>C.8.2.</b> Tomar decisões e inovar, com base no conhecimento do funcionamento e das características técnicas de Ciência de Dados, consciente dos aspectos éticos, legais e dos impactos ambientais decorrentes	<b>Aplicar</b>	Privacidade de Dados
		Legislação de Computação, inteligência artificial e de proteção de dados
		Sustentabilidade
		Ciência de Dados, Ética e Sociedade
<b>C.8.3.</b> Conceber soluções computacionais capazes de trazer inovação a um produto ou processo	<b>Aplicar</b>	Empreendedorismo
		Interação Humano-Computador
		Metodologia Científica
		Gerência de Projeto
<b>C.8.4.</b> Identificar e analisar requisitos e especificações para problemas específicos e planejar estratégias para suas soluções de Ciência de Dados	<b>Aplicar</b>	Engenharia de Software
		Teste, Verificação e Validação de Software

## II.7. Relação com as Diretrizes Curriculares Nacionais

Esta seção apresenta em quais eixos de formação estão mapeadas cada uma das competências e habilidades gerais da área de Computação, conforme DCN2016 e específicas definidas neste documento.

RELAÇÃO DE COMPETÊNCIAS DOS REFERENCIAIS DE FORMAÇÃO COM AS COMPETÊNCIAS DESCRITAS NAS DCNS	
Competências e habilidades gerais dos egressos dos Cursos de Bacharelado e Licenciatura	Competências dos Referenciais de Formação
1. Identificar problemas que tenham solução algorítmica	C.1.1, C.8.1
2. Conhecer os limites da computação	C.1.2, C.3.5
3. Resolver problemas usando ambientes de programação	C.1.1, C.3.1
4. Tomar decisões e inovar, com base no conhecimento do funcionamento e das características técnicas de hardware e da infraestrutura de software dos sistemas de computação consciente dos aspectos éticos, legais e dos impactos ambientais decorrentes	C.3.2, C.4.2, C.4.5, C.5.1, C.5.3, C.8.2
5. Compreender e explicar as dimensões quantitativas de um problema	C.4.5, C.5.2, C.6.2, C.8.3
6. Gerir a sua própria aprendizagem e desenvolvimento, incluindo a gestão de tempo e competências organizacionais	C.7.1
7. Preparar e apresentar seus trabalhos e problemas técnicos e suas soluções para audiências diversas, em formatos apropriados (oral e escrito)	C.4.3, C.5.5
8. Avaliar criticamente projetos de sistemas de computação	C.2.2, C.3.3, C.5.6
9. Adequar-se rapidamente às mudanças tecnológicas e aos novos ambientes de trabalho	Eixo 4, C.5.3, C.5.4
10. Ler textos técnicos na língua inglesa	Eixo 1, C.7.1
11. Empreender e exercer liderança, coordenação e supervisão na sua área de atuação profissional	C.3.4, C.5.7, C.7.2
12. Ser capaz de realizar trabalho cooperativo e entender a força que dele pode ser derivada	C.3.4, C.5.7, C.7.2

Competências e habilidades dos egressos dos Cursos de Bacharelado em Ciência de Dados	Competências dos Referenciais de Formação
1. Possuir sólida formação nas áreas de Computação, matemática e estatística que permita a aplicações de conceitos dessas áreas em tarefas de ciência de dados;	C.1.2, C.1.3, C.1.4, C.3.5, C.7.3
2. Utilizar efetivamente técnicas computacionais, matemáticas e estatísticas para, de forma analítica, avaliar a factibilidade e, quando factível, extrair conhecimento dos dados disponíveis para realizar descobertas em diferentes domínios de aplicação, de forma a apoiar o processo de tomada de decisão;	C.1.3, C.1.4, C.2.3, C.2.4, C.2.5, C.2.6, C.6.1, C.6.4, C.6.5
3. Empregar os princípios de Tecnologia de Informação e Comunicação para pesquisar, projetar, implementar e avaliar novas abordagens e técnicas para construção de ferramentas para análise de dados;	C.2.4, C.4.1, C.4.5, C.5.3, C.5.4, C.5.8, C.6.5, C.6.8
4. Ser capaz de planejar e realizar experimentos computacionais simulados utilizando diferentes infraestruturas que apoiem a gestão e o manuseio eficiente de dados, estruturados e não estruturados, durante todo o ciclo de vida dos dados;	C.3.6, C.4.2, C.6.3, C.7.4, C.8.4
5. Definir e implementar estratégias de gerenciamento de dados para curadoria, coleta, integração, armazenamento, visualização, preservação e disponibilização de dados para futuro processamento;	C.2.3, C.4.3, C.4.5, C.6.1, C.6.4, C.6.6
6. Gerenciar projetos interdisciplinares que incluam as diversas etapas do ciclo de vida de dados;	C.5.8
7. Identificar novos desafios, necessidades, oportunidades de negócios e desenvolver soluções inovadoras;	C.2.2, C.2.6, C.2.7, C.3.7, C.3.10, C.5.8, C.5.9
8. Investigar, compreender e estruturar as características de domínios de aplicação em diversos contextos que levem em consideração questões ambientais, éticas, sociais, legais e econômicas;	C.2.1, C.3.8, C.3.9, C.6.7, C.7.4
9. Ser capaz de trabalhar de forma tanto individual como colaborativa, com profissionais da mesma área ou de diferentes áreas;	C.7.2
10. Seguir os princípios de uma ciência de dados justa, transparente, sem viés, respeitando a privacidade e atendendo aos requisitos da Lei Geral de Proteção de Dados (LGPD);	C.4.4, C.6.7

11. Ter uma visão crítica e criativa na identificação e resolução de problemas contribuindo para o desenvolvimento da área;	C.7.4, C.8.3
12. Atuar de forma empreendedora, abrangente e cooperativa no atendimento às demandas sociais da região onde atua, no Brasil e no mundo;	C.2.1, C.3.9, C.3.10
13. Utilizar racionalmente os recursos disponíveis de forma transdisciplinar; e,	C.4.5, C.5.4
14. Atuar em um mundo globalizado do trabalho, buscando o domínio de idiomas estrangeiros, em particular o idioma inglês.	EIXO 1, C.3.4, C.5.7

## **II.8 Estágios, TCC e atividades complementares**

De acordo com esses referenciais, cabe às IES estabelecerem a obrigatoriedade ou não do Estágio Supervisionado ou do Trabalho de Curso e a definição dos respectivos regulamentos.

### **II.8.1 Estágio Supervisionado**

O estágio supervisionado é um ato escolar educativo desenvolvido no ambiente de trabalho. A atividade não cria vínculo empregatício e seu objetivo é preparar o estudante para o mercado de trabalho, por meio do desenvolvimento de competências inerentes à atividade profissional e da contextualização curricular. O estágio, bem como seu regulamento, deve fazer parte do projeto pedagógico do curso.

O estágio supervisionado pode ser obrigatório ou não. No primeiro caso, a sua carga horária é requisito para aprovação no curso e obtenção do diploma, enquanto no segundo caso trata-se de atividade opcional com carga horária acrescida à carga horária regular e obrigatórias. Compete às IES, por meio dos Núcleos Docentes Estruturantes (NDE), determinar a obrigatoriedade ou não do estágio supervisionado para os cursos de bacharelado em Ciência de Dados.

De acordo com as DCN2016, os cursos podem envolver a formação para a realização de: (i) atividades orientadas à realização de processos; e (ii) atividades orientadas à transformação de processos, com desenvolvimento de novas tecnologias. Nos cursos orientados à realização, ou reprodução, de processos, há forte recomendação de estágio para os alunos. Por meio de estágio, os estudantes podem conhecer previamente o ambiente onde são desenvolvidas as atividades de trabalho para as quais eles estão sendo preparados, como forma de iniciação à profissionalização.

Como o estágio se trata de uma atividade em que o estudante replica conhecimentos para atividades de realização de processos, é recomendável que o NDE do curso defina no regulamento de estágio o momento, no que concerne períodos, ou uma formação mínima, em se tratando de conclusão de componentes curriculares, para que o estudante possa iniciar o estágio supervisionado.

### **II.8.2 Trabalho de Conclusão de Curso**

O Trabalho de Conclusão de Curso (TCC) é uma atividade acadêmica curricular cujo objetivo é desenvolver e verificar as habilidades cognitivas de compreensão, aplicação, análise, avaliação e criação acerca dos conhecimentos científicos, técnicos e culturais produzidos ao longo do curso.

Compete às IES, por meio dos Núcleos Docentes Estruturantes (NDE), determinar a obrigatoriedade ou não do TCC para os cursos de bacharelado em Ciência de Dados.

De acordo com as DCN2016, os cursos podem envolver a formação para a realização de: (i) atividades orientadas à realização de processos; e (ii) atividades orientadas à transformação de processos, com desenvolvimento de novas tecnologias. Nos cursos orientados à transformação de processos, ou criação de novas tecnologias, há forte recomendação de realização de TCC para os alunos.

Por meio de TCC, os estudantes devem aplicar conhecimentos de vanguarda na produção de aplicações científicas, tecnológicas ou de inovações.

Compete às IES regulamentar as normas e os procedimentos no seu regimento e dar conhecimento ao aluno acerca do TCC. O regulamento deve especificar os critérios, procedimentos e mecanismos de avaliação, assim como as diretrizes e técnicas relacionadas à sua elaboração. Deverá constar no regulamento quando o estudante iniciará o TCC.

### **II.8.3 Atividades complementares**

De acordo com as DCN2016, as Atividades Complementares são componentes curriculares enriquecedores e implementadores do próprio perfil do formando e deverão possibilitar o desenvolvimento de habilidades, conhecimentos, competências e atitudes do aluno, inclusive as adquiridas fora do ambiente acadêmico, que serão reconhecidas mediante processo de avaliação.

As Atividades Complementares podem incluir atividades desenvolvidas na própria Instituição ou em outras instituições e variados ambientes sociais, técnico-científicos ou profissionais de formação profissional, incluindo:

- experiências de trabalho;
- estágios não obrigatórios;
- extensão universitária
- iniciação científica;
- participação em eventos técnico-científicos;
- publicações científicas;
- programas de monitoria e tutoria;
- disciplinas de outras áreas;
- representação discente em comissões e comitês;
- participação em empresas juniores;
- incubadoras de empresas;
- atividades de empreendedorismo e inovação.

## II.9 Metodologia de ensino

De acordo com as DCN2016, a metodologia de ensino deve ser centrada no aluno como sujeito da aprendizagem e apoiada no professor como facilitador do processo de ensino-aprendizagem. O trabalho extraclasse deve ser empregado de forma que o aluno aprenda a resolver problemas e seja estimulado a aprender a aprender, tornando-se independente e criativo. O professor deve apresentar as aplicações dos conteúdos teóricos, ser um mediador, estimular a competição, a comunicação, provocar a realização de trabalho em equipe, motivar os alunos para os estudos e orientar o raciocínio e desenvolver as capacidades de comunicação e de negociação. Quando aplicável, deve-se empregar metodologias ativas, de forma que o aluno passe mais tempo em atividades nas quais seja protagonista no processo de ensino e aprendizagem.

Considerando o cenário atual da educação, numa sociedade com forte viés tecnológico, o projeto pedagógico deve prever o emprego de metodologias de ensino e aprendizagem que promovam a explicitação das relações entre os conteúdos abordados e as competências previstas para o egresso dos cursos. Deve ainda considerar a inserção de novos paradigmas educacionais que abordam práticas curriculares e metodologias de ensino inovadoras. Novas mídias e plataformas tecnológicas de ensino e aprendizagem devem ser inseridas no meio acadêmico, levando a uma reflexão sobre as práticas pedagógicas utilizadas e a efetividade destas no processo de ensino e aprendizagem.

As metodologias de ensino e paradigmas atualmente empregadas são constantemente abordadas nos eventos do CSBC (ex. Curso de Qualidade, WEI, SECOMU, EDUCOMP e JAI) e eventos temáticos organizados pelas Comissões Especiais.

## II.10 Requisitos legais

Esses referenciais foram elaborados em consonância com as DCN2016. Espera-se que eles auxiliem as instituições e coordenações de curso na elaboração e revisão de seus Projetos Pedagógicos de Curso (PPC). Destaque-se fortemente que esses referenciais não são um PPC para os cursos de Bacharelado em Ciência de Dados.

Na elaboração de um PPC, as instituições, coordenações e Núcleos Docentes Estruturantes (NDEs) podem usar este documento como guia para decidir quanto a eixos, competências e conteúdos de formação para seus cursos, mas precisam considerar diversas questões, tais como: o contexto regional onde atua, suas estratégias de formação, sua capacidade de formação, a estratégia institucional para formação de egressos, entre vários outros aspectos.

Deve-se observar também que outros requisitos legais e normativos são exigidos pelas instituições para elaboração de seus PPC. Tais requisitos legais e normativos devem ser cuidadosamente analisados e interpretados pelos NDE antes de se atualizar ou propor um novo PPC. É necessário destacar também a necessidade de adequar o PPC para a modalidade do curso, haja vista que os cursos à distância exigem o cumprimento de requisitos legais e normativos específicos para esse tipo de modalidade, além daqueles exigidos para os cursos presenciais.

O MEC é responsável por organizar o sistema de ensino no país e definir as políticas e diretrizes educacionais nacionais. Isso envolve dar publicidade e zelar pelo cumprimento das leis, decretos, portarias, pareceres, resoluções, e demais instrumentos legais e normativos federais para o ensino, inclusive superior. Em consonância com a esfera federal, as Secretarias de Educação, estaduais, municipais e do Distrito Federal são responsáveis por organizar o sistema de ensino e definir as políticas e diretrizes educacionais; porém, dentro dos seus respectivos domínios: estadual, municipal e distrital. Normalmente, as Instituições de Ensino Superior (IES) elaboram os seus próprios regulamentos e regimentos internos, bem como Planos de Desenvolvimento Institucional e outros instrumentos. Assim, é relevante também observar as questões internas das IES; porém, respeitando a hierarquia legal.

É de grande importância o acompanhamento da evolução legal e normativa pertinentes à educação superior, no país, estados, municípios e na própria IES, haja vista que leis, decretos e outros instrumentos estão sempre sujeitos a mudanças.

## II.11 Agradecimentos

A viabilização deste trabalho se deve ao suporte de vários parceiros. A diretoria da Associação Brasileira de Estatística (ABE) colaborou na definição de conteúdos de Estatística em um curso de Ciência de Dados. A Sociedade Brasileira de Computação (SBC) colocou sua infraestrutura à disposição, para nos auxiliar no que fosse necessário, A Diretoria de Marketing da SBC foi fundamental para a realização da consulta pública, para que este documento fosse discutido pela sociedade em geral e, particular, pela comunidade de computação.

## Referências

ACM/IEEE (2013). Computer Science Curricula 2013: Curriculum Guidelines for Undergraduate Degree Programs in Computer Science. Final Report. ACM, New York, NY, USA. 2013. Disponível em: <http://dx.doi.org/10.1145/2534860>. Último acesso em: 27/03/2017.

ACM (2021) Computing Competencies for Undergraduate Data Science Curricula. Final Report. ACM, New York, NY, USA. 2021, Disponível em: <https://www.acm.org/education/curricularecommendations>. Último acesso em: 17/02/2023.

American Statistical Association Undergraduate Guidelines Workgroup. 2014. 2014 curriculum guidelines for undergraduate programs in statistical science. Alexandria, VA: American Statistical Association. <https://www.amstat.org/docs/default-source/amstat-documents/edu-guidelines2014-11-15.pdf>. Último acesso em 09/05/2023.

Ferraz, A. P. C. M., Belhot, R. V. (2010). Taxonomia de Bloom: revisão teórica e apresentação das adequações do instrumento para definição de objetivos instrucionais. Gest. Prod., São Carlos, 17(2), 421-431.

Le Boterf, G. (2006) Avaliar a competência de um profissional: três dimensões a explorar. Reflexão RH, 1.1: 60-3.

Sacristán, J. G.; Gómez, Á. I. P.; Rodríguez, J. B. M.; Santomé, J. T.; Rasco, F. A.; Méndez, J. M. Á. (2016). Educar por competências: O que há de novo?. Artmed Editora.

SBC (1991). Currículo de Referência da SBC para cursos de Graduação Plena em Computação (CR91). <http://homepages.dcc.ufmg.br/~bigonha/Cr/cr91.html>. Último acesso em: 27/03/2017.

SBC (1996). Currículo de Referência da SBC para cursos de Graduação Plena em Computação (CR96). <http://homepages.dcc.ufmg.br/~bigonha/Cr/cr.html>. Último acesso em: 27/03/2017.

SBC (1999). Currículo de Referência da SBC para cursos de Graduação em Computação (CR99). <http://lad.dsc.ufcg.edu.br/ec/cr99.pdf>. Último acesso em: 27/03/2017.

SBC (2003). Currículo de Referência da SBC para cursos de Graduação em Computação e Informática (CR03). <https://pt.slideshare.net/derciolr/currculo-de-referencia-sbc>. Último acesso em: 27/03/2017.

SBC (2017). Referenciais de Formação para os Cursos de Graduação em Computação 2017. Disponível em: <https://www.sbc.org.br/educacao/referenciais-de-formacao-2017>. Último acesso em: 17/02/2023.

Scallon, G. (2015) Avaliação da aprendizagem numa abordagem por competências. Curitiba: PUCPress.

MEC (1999). Diretrizes Curriculares Nacionais de cursos da área de Computação e Informática (DCN99). Disponível em: <http://www.inf.ufrgs.br/ecp/docs/diretriz.pdf>. Último acesso em: 27/03/2017.

MEC (2008). Diretrizes Curriculares Nacionais para o curso de graduação em Estatística. Disponível em [http://portal.mec.gov.br/cne/arquivos/pdf/2008/pces214\\_08.pdf](http://portal.mec.gov.br/cne/arquivos/pdf/2008/pces214_08.pdf). Último acesso em 1 de novembro de 2022.

MEC (2012). Proposta de Diretrizes Curriculares Nacionais para os cursos de graduação em Computação. Disponível em: [http://portal.mec.gov.br/index.php?option=com\\_docman&view=download&alias=11205-pces136-11-pdf&category\\_slug=julho-2012-pdf&Itemid=30192](http://portal.mec.gov.br/index.php?option=com_docman&view=download&alias=11205-pces136-11-pdf&category_slug=julho-2012-pdf&Itemid=30192). Parecer CNE/CES nº 136/2012, aprovado em 8 de março de 2012. Último acesso em: 27/03/2017.

MEC (2016). Diretrizes Curriculares Nacionais para os Cursos de Graduação em Computação (DCN2016). Disponível em: [http://portal.mec.gov.br/index.php?option=com\\_docman&view=download&alias=52101-rces005-16-pdf&category\\_slug=novembro-2016-pdf&Itemid=30192](http://portal.mec.gov.br/index.php?option=com_docman&view=download&alias=52101-rces005-16-pdf&category_slug=novembro-2016-pdf&Itemid=30192). Resolução CNE/CES nº 5, de 16 de novembro de 2016. Último acesso: 27/03/2017.