

Projeto Pedagógico

Especialização em Ciência
de Dados aplicada a
Políticas Públicas

Sumário

Introdução.....	03
Objetivo do curso.....	03
Competências a serem desenvolvidas.....	04
Público-Alvo.....	04
Carga horária.....	05
Processo seletivo.....	05
Vagas.....	05
Requisitos para ingresso no curso.....	06
Metodologia.....	06
Estrutura curricular.....	08
Ementas da disciplinas.....	09

Introdução

Este documento sistematiza as concepções básicas e fundamentos teóricos do Curso “Especialização em Ciência de Dados aplicada a Políticas Públicas”, regido pela Resolução no 22, de 27 de outubro de 2017, que aprova o Regulamento dos Cursos de Pós-Graduação *Lato Sensu* da Escola Nacional de Administração Pública - Enap. Especifica ainda as finalidades e a organização curricular integrando conteúdos e práticas no campo da Ciência de Dados. A proposta curricular tem como eixo norteador a relação direta entre a Ciência de Dados e o aumento da capacidade de entendimento dos problemas organizacionais e a tomada de decisões baseada em evidências.

Com um olhar voltado para a prática, o curso, ao longo dos seus três eixos temáticos: Ciência da Computação, Matemática/Estatística e Políticas Públicas, desenvolverá habilidades de pensamento crítico não só de manipulação e integração de dados e construção de modelos, mas principalmente de experimentação, identificação, formulação de problemas relevantes no âmbito das políticas públicas e de tomada de decisão. Para isso, combina modalidades de ensino, estratégias e instrumentos didático-pedagógicos que levarão os estudantes a analisar, compreender, distinguir e avaliar os conceitos e preceitos da administração pública e do seu papel profissional, para (re)elaborar seu conhecimento sobre o contexto complexo da ação governamental e decidir sobre como agir e interagir em situações concretas.

Objetivo do curso

Desenvolver e aprimorar a capacidade e competência profissional de servidores públicos federais que possibilitem aplicar os conhecimentos de Ciências de Dados na análise e avaliação de Políticas Públicas no âmbito do governo brasileiro em suas diferentes esferas.

Ao final do curso os alunos devem ser capazes de:

- Entender o processo de elaboração de políticas públicas;
- Conhecer técnicas de avaliação de políticas públicas;
- Desenvolver habilidades básicas de programação;
- Desenvolver proficiência na análise estatística de dados com estatística descritiva e inferencial;
- Criar e avaliar modelos baseados em dados;
- Desenvolver análises estatísticas em linguagem de programação apropriada;
- Desenvolver habilidades de coleta, manipulação, limpeza e integração de dados;
- Aplicar conceitos e métodos de ciência de dados para tratar problemas no contexto de políticas públicas e comunicar suas soluções efetivamente.

Competências a serem desenvolvidas

- Pensamento crítico e analítico;
- Mentalidade de crescimento (growth mindset);
- Comunicação por meio da apresentação de dados;
- Capacidade de experimentação e tomada de decisão baseado em evidências;
- Capacidade de identificar, formular e enunciar problemas relevantes; e
- Empatia e capacidade de entender as necessidades das personas.

Público-Alvo

O curso tem como público-alvo servidores ou empregados públicos federais que:

- atuem em área diretamente relacionada à análise e avaliação de políticas públicas do Governo Federal ou em áreas/unidades que

trabalham diretamente com coleta, manipulação, limpeza e integração de dados para tratar problemas no contexto de políticas públicas;

- tenham necessidade de desenvolver competências requeridas à modelagem e produção de dados usando protótipos, algoritmos, modelos preditivos e análises personalizadas.

E ainda servidores públicos federais ocupantes de cargo efetivo e empregados públicos concursados em exercício na Controladoria-Geral da União e demais servidores públicos federais ocupantes de cargo efetivo, empregados públicos federais concursados e militares das Forças Armadas.

Carga horária

A carga horária total do curso será de **375 horas**, a ser cursada em aproximadamente 13 meses. A ENAP poderá oferecer até 40% da carga horária total do curso a distância.

A carga horária semanal de 9 horas está prevista para ser realizadas às **segundas-feiras** no período da tarde – **das 14h30 às 17h30** e **terças-feiras** no período da manhã – **das 9h às 12h** e no período da tarde – **das 14h30 às 1730**.

Processo seletivo

Para ingresso no curso, as pessoas que se candidataram deverão ser aprovadas em processo seletivo, dentro do número de vagas ofertadas, regulamentado por edital.

Vagas

Serão ofertadas 40 vagas, sendo 30 reservadas para a Controladoria-Geral da União e 10 para os demais órgãos e entidades do governo federal. Serão chamados para ocupar as quarenta vagas do curso os(as) candidatos(as) classificados(as), dentro do número de vagas, em processo seletivo, nos termos definidos no edital do processo seletivo do curso.

Requisitos para ingresso no curso

São requisitos para ingresso no curso:

- Ser graduado em qualquer curso de nível superior reconhecido pelo Ministério da Educação – MEC;
- Pertencer ao público-alvo a que se destina o processo seletivo, conforme Edital;
- Não estar cursando outro programa de Pós-Graduação *Lato Sensu* ou *Stricto Sensu* da Enap;
- Não ter cursado outro programa de Pós-Graduação *Lato Sensu* ofertado pela Enap nos dois anos anteriores à data da matrícula;
- Cumprir com o disposto do Edital;
- Ser aprovado em processo seletivo conduzido pela Enap;
- Apresentar os documentos necessários para a efetivação da matrícula, conforme Edital.

Metodologia

O curso aborda as bases teóricas na área de Ciência de Dados e aplica conceitos de vieses cognitivos nos três eixos temáticos que sustentam o curso: Ciência da Computação, Matemática/Estatística e Políticas Públicas, incentivando a discussão de experiências e a prática na manipulação e análise de bases de dados de Políticas Públicas de Infraestrutura, de Saúde e Sociais.

O curso será um espaço de aprendizagem coletiva, integrando a teoria e a prática. A participação ativa dos estudantes será demandada nas aulas online e presenciais, bem como no ambiente virtual de aprendizagem. As atividades envolverão aulas expositivas dialogadas; flash cards de revisão teórica; atividades práticas elaboradas na forma de *Jupyter Notebooks*; discussões em grupos; participação no ciclo de palestras Conhecimentos Conexos; laboratório de casos com utilização de bases de dados robustas, onde serão incentivadas as discussões de experiências e a prática na manipulação e análise de bases de dados de Políticas Públicas de Infraestrutura, de Saúde e Sociais.

A expectativa é que os participantes sejam incentivados à reflexão, visando correlacionar os conceitos e proposições abordados nas disciplinas com situações concretas e a entender os problemas organizacionais e tomar decisões com base em evidências.

A cada disciplina serão realizadas avaliações de aprendizagem que abrangerão a aplicação de conceitos, técnicas e reflexões realizadas durante as aulas.

As avaliações formativas serão realizadas por meio de trabalhos em grupo e individuais, testes práticos, dentre outras. As somativas incluem o Trabalho de Conclusão de Curso (TCC), que tem o foco na aplicação prática como um diferencial metodológico do curso.



Estrutura curricular

DISCIPLINAS		CH
Análise de Políticas Públicas e Processo Decisório	(online)	18h
Desenho e Avaliação Ex-Ante de Políticas Públicas	(online)	21h
Políticas Públicas baseadas em evidências	(online)	18h
Monitoramento e Avaliação Ex-Post de Políticas Públicas	(online)	21h
Introdução a Programação em Python	(online)	30h
Estatística Descritiva	(online)	24h
Introdução a Banco de Dados e Linguagem SQL	(online)	24h
Estatística Inferencial	(online)	36h
Introdução a Ciência de Dados	(online)	42h
Design Sprint e Lean Inception	(online)	18h
Laboratório de Casos - Políticas Públicas de Saúde	(presencial)	09h
Machine Learning	(online)	39h
Laboratório de Casos - Políticas Públicas Sociais	(presencial)	12h
Sistemas de Informação Geográfica (SIG)	(presencial)	30h
Laboratório de Casos - Políticas Públicas de Infraestrutura	(presencial)	09h
Diálogos Metodológicos	(online)	12h
Disciplina Integradora Aplicada	(online)	12h
Conhecimentos Conexos	(online)	Optativa
Avaliação de Impacto de Programas e Políticas Sociais	(online)	Optativa
Monitoria em Ciência da Computação	(online)	Optativa
Monitoria em Matemática e Estatística	(online)	Optativa
CARGA HORÁRIA TOTAL		375h

Ementas das disciplinas

D1. Análise de Políticas Públicas e Processo Decisório

Dimensões de análise das políticas públicas: tipos de políticas públicas, atores de políticas públicas, fases do processo de elaboração de políticas públicas (formação da agenda, formulação de alternativas, tomada de decisão, implementação, avaliação, extinção). Conceitos, Métodos de Estudo e Avaliação. Formulação e implementação das políticas públicas.

D2. Desenho e Avaliação Ex-Ante de Políticas Públicas

O papel da avaliação ex-ante das políticas públicas. Diagnóstico do problema e modelo lógico: identificação do problema; relação de sobreposição e complementaridade entre as ações estatais; avaliação de desenho de política pública e definição de indicadores. Desenho e suporte da política pública: objetivo e ações; fundamentação; público-alvo; atores envolvidos e seleção de beneficiários. Análise SWOT para entendimento dos contextos internos e externos. Estratégia de construção da confiança e do suporte da implementação e aprovação da política pública.

D3. Políticas Públicas baseadas em evidências

O que são evidências. Como se distingue evidência de opinião. Como se produzem evidências. Como se prospectam evidências. O papel da meta-avaliação em particular das revisões sistemáticas. Como as evidências instruem as diversas fases do ciclo da política pública. O que é política pública baseada em evidências. Como orientar a prática da política pública baseada em evidências. Quais os limites e dificuldades dessa abordagem.

D4. Monitoramento e Avaliação Ex-Post de Políticas Públicas

Breve histórico da avaliação de programas e políticas públicas. Programas e políticas como intervenções na realidade. Tipologias de avaliação e técnicas

de análise (avaliação normativa e pesquisa avaliativa). Pesquisa avaliativa (avaliação diagnóstica, avaliação de implementação, avaliação de efeitos (impacto). Projeto de avaliação (desenho, gestão e disseminação). problemas e desafios da avaliação de programas e políticas. Estudos de caso de avaliação de políticas públicas.

D5. Introdução a Programação em Python

Algoritmos. Operação de atribuição. Tipos e variáveis. Desvios condicionais. Comandos de seleção múltipla. laços de repetição. Listas, arrays multidimensionais e dicionários. Operadores matemáticos. Expressões e funções. Depuração de programas e exceções. Algoritmos gulosos, divisão e conquista. Busca sequencial e busca binária. Manipulação de arquivos de texto e arquivos JSON. Ferramenta de log em arquivo texto. Funções de array em nível de elemento (element-wise array functions).

D6. Estatística Descritiva

Distribuição de frequências e histograma. Principais medidas de posição: média, mediana, moda, quartis e box plots. Principais medidas de dispersão: amplitude, variância e desvios médio e padrão. Medidas de assimetria e curtose. Medidas envolvendo duas variáveis: covariância e correlação. Análise de dados anômalos ou discrepantes. Aplicações com dados públicos.

D7. Introdução a Banco de Dados e Linguagem SQL

Banco de Dados: Histórico e Características. Projeto de Banco de Dados: Conceitos, Dependência Funcional, Restrições de Integridade e Formas Normais. Modelo Entidade-Relacionamento (ER). Linguagem SQL. Conceitos de transações e propriedades ACID. Mapeamento Objeto Relacional (Object Relational Mapping - ORM). Bancos de dados NoSQL.

D8. Estatística Inferencial

Definição e tipos de amostragem. Distribuições da média amostral e do desvio padrão amostral. Estimativas pontual e intervalar. Intervalo de confiança para a média, para a proporção e para a variância. Hipóteses estatísticas. Teste de Hipóteses. Teste para n médias (ANOVA). Teste para n Proporções (Teste do Qui-Quadrado).

D9. Introdução a Ciência de Dados

Introdução ao Pandas (Python): Estruturas de dados do Pandas e Numpy. Importação, exportação, manipulação e limpeza de dados. Filtro e Agregação de dados. Análise de Dados Exploratória. Análise de Dados Exploratória Automatizada. Story telling with data. Web scrapping. Regressão Linear. Causalidade em Ciência de Dados.

D10. Design Sprint e Lean Inception

Introdução ao Lean Inception. Visão do produto. Objetivos do produto. Personas. Jornada de usuários. Funcionalidades. Nivelamento das Funcionalidades. Sequenciamento das Funcionalidades. Canvas MVP. Introdução ao Design Sprint. Dia 1 - Entender/Definir. Dia 2 – Divergir. Dia 3 – Decidir. Dia 4 – Prototipar. Dia 5 – Validar.

D11. Laboratório de Casos - Políticas Públicas de Saúde

Laboratório sobre atendimentos da saúde básica e atenção familiar. Laboratório sobre a pandemia do COVID-19. Laboratório sobre políticas de vacinação.

D12. Machine Learning

Introdução ao aprendizado de máquina. Aprendizado supervisionado e Não

supervisionado. Introdução a framework de programação de aprendizado de máquina. Regressão linear. Overfitting, complexidade, treinamento, validação, dados de teste. Problemas de classificação, fronteiras de decisão, métodos de vizinhos mais próximos. Árvore de decisão e Random Forest. Redes Neurais.

D13. Laboratório de Casos - Políticas Públicas Sociais

Laboratório sobre pobreza e políticas de redistribuição de renda. Laboratório sobre o mercado de trabalho e empregos. Laboratório sobre criminalidade e justiça com foco no público jovem.

D14. Sistemas de Informação Geográfica (SIG)

Representação computacional do espaço. Conceitos: Espaço, Escala, Modelo, Dependência Espacial. Tipos de Dados Geográficos. Estruturas de Dados em SIG. Conceitos básicos de cartografia para Sistemas de Informação Geográfica. Interoperabilidade de Dados Geográficos. Padrões Abertos e Especificações - OGC. Modelagem Numérica de Terreno. Introdução à Geoestatística. Álgebra de Mapas. Inferência Geográfica e Suporte à Decisão. Estudos de caso: Saúde Coletiva. Estudos Populacionais

D15. Laboratório de Casos - Políticas Públicas de Infraestrutura

Laboratório sobre Demografia, habitação, déficit habitacional. Laboratório sobre o uso e racionamento dos recursos hídricos. Laboratório sobre produção e distribuição de energia elétrica.

D16. Diálogos Metodológicos

Os tipos de conhecimento: uma breve história da Ciência. Métodos científicos: indutivo, dedutivo, hipotético-dedutivo e dialético. Fatos, leis e

teorias. Tipos de produção e publicação científica. O projeto de pesquisa: o processo de planejamento do trabalho científico (tema, problema e possíveis respostas; justificativa, objetivos. Elaboração do aporte teórico. Regras da redação científica -citações e referências). Cronograma de projeto. Metodologias de pesquisa: métodos, tipos e natureza. Métodos qualitativos e quantitativos: características, técnicas de coleta, tratamento, análise e apresentação de dados. Estrutura do projeto capstone. Orientação da forma do pré-projeto de pesquisa.

D17. Disciplina Integradora Aplicada

Trabalhos em grupo com temas específicos escolhidos pelos professores das disciplinas recentemente ministradas para estimular o aluno a articular conceitos e evidenciar o desenvolvimento da capacidade de aplicação do conhecimento adquirido ao longo do curso, conectando teoria e prática.

D18. Conhecimentos Conexos

Palestras, com conteúdo de curta duração, com temas específicos escolhidos pelos professores das disciplinas recentemente ministradas.

D19. Avaliação de Impacto de Programas e Políticas Sociais- EaD

Avaliação de impacto. Avaliação aleatorizada. Teoria da Mudança. Realização de avaliação aleatorizada. Tamanho amostral e poder estatístico. Medição. Aspectos operacionais. Desafios após o desenho. Uso de evidência. Avaliação do início ao fim.

D20. Monitoria em Ciência da Computação

Comandos de terminal. Caminhos absolutos e relativos. Variáveis de ambiente. Versionamento de código, git e Github. Tabela verdade dos

operadores AND, OR e NOT e sentenças SE/ENTÃO. Revisão de conceitos de cartografia.

D21. Monitoria em Matemática e Probabilidade

Regra de três e conceitos elementares de matrizes, conjuntos e funções. Conceitos elementares de probabilidade. Introdução a álgebra linear (matrizes, sistemas lineares). Espaços vetoriais e transformações lineares. Ortogonalidade. Tipos de distribuições (Uniforme, Gaussiana). Noções intuitivas de cálculo diferencial.



Diretoria de Educação Executiva
Coordenação-Geral de Especialização e MBA

