

# CURSO DE CIÊNCIA DE DADOS APLICADA AO PODER JUDICIÁRIO



**FUNDAMENTOS DE ESTATÍSTICA PARA CIÊNCIA DE DADOS**

## FICHA DO CURSO

**Curso:** Fundamentos de Estatística para Ciência de Dados

**Modalidade:** EAD - autoinstrucional

**Carga horária total:** 40 horas

**Carga horária semanal:** 4 horas

**Início da oferta:** -

**Fim da oferta:** -

**Pré-requisito:** Não há pré-requisitos recomendados para o cursista.

**Conteudistas:** Wagner Hugo Bonat

**1. Objetivos**

Habilitar o(a) aluno(a) a utilizar métodos estatísticos tais como, medidas descritivas, gráficas e tabulares. Princípios de probabilidade e inferência estatística.

Ao final do curso o participante deverá demonstrar capacidade de uso e interpretação racional de métodos estatísticos relacionados a estatística descritiva e inferencial.

**2. Programa**

Estatística: contextualização e motivação; Métodos de amostragem; Estatística descritiva; Medidas de associação; Elementos de probabilidade; Inferência estatística; Testes de hipóteses; Comparando dois grupos; Analisando a associação entre duas variáveis categóricas; comparando mais de dois grupos.

**3. Procedimentos didáticos**

A cada módulo (em um total de **10 módulos**) serão desenvolvidas as seguintes atividades pelos cursistas:

1. Acessar o material pré aula para preparar-se para a aula.
2. Assistir as videoaulas programadas para o módulo.
3. Fazer o estudo individual de materiais indicados como: leituras complementares, resolução de exercícios e acesso a vídeos adicionais.
4. Realizar as atividades de avaliação do módulo no formato de quiz.

Como suporte ao Ensino Remoto será utilizada a infraestrutura da plataforma moodle do Centro de Formação e Aperfeiçoamento de Servidores do Poder Judiciário (CEAJUD).

#### 4. Procedimentos metodológicos

Conjunto de vídeo-aulas separadas por módulos de conhecimento acompanhado de avaliações para reforçar o aprendizado.

#### 5. Distribuição da carga horária do cursista

A Tabela 1 a seguir apresenta distribuição da carga horária semanal total do cursista, de 4 horas por semana, nas atividades semanais previstas na vigência do curso.

Tabela 1 - Carga horária semanal do(a) cursista.

Atividade semanal	Carga horária (horas)	Fração (%)
Estudo individual de conteúdo pré-aula	00:50	20,83%
Assistir vídeo-aulas	01:10	29,17%
Estudo individual pós aula	01:00	25,0%
Atividades avaliativas	01:00	25,0%
<b>Total</b>	<b>04:00</b>	<b>100%</b>

#### 6. Conteúdo programado

O curso está dividido em 10 módulos sendo que cada módulo configura uma semana de atividades. O conteúdo programado é apresentado na Tabela 2.

Tabela 2 - Cronograma detalhado do conteúdo das unidades didáticas.

<b>Módulo</b>	<b>Conteúdo programado</b>	<b>Duração</b>
1	<p>A era do Big Data e Analytics</p> <p>O que é Ciência de Dados?</p> <p>Como estruturar problemas com dados</p> <p>Etapas de uma análise de dados</p> <p>Tipos de estudos</p> <p>Estatística e Ciência de dados</p> <p>Conceitos fundamentais</p> <p>O que é estatística? Por que estudar estatística?</p> <p>Variáveis e tipos de variáveis</p> <p>Níveis de medidas e escalas de medida</p> <p>Estatística e métodos computacionais</p>	4 horas
2	<p>Tipos de estudos</p> <p>Métodos de amostragem probabilística</p> <p>Métodos de amostragem não probabilística</p> <p>Comentários adicionais</p>	4 horas
3	<p>Importância da estatística descritiva e exploratória</p> <p>Tipos de variáveis e formas de análise</p> <p>Avaliando a qualidade dos dados</p> <p>Descrivendo uma variável qualitativa</p> <p>Descrivendo uma variável quantitativa</p> <p>Representações gráficas</p> <p>Medidas de posição e posição relativa</p> <p>Medidas de dispersão</p>	4 horas
4	<p>Tabelas de dupla entrada</p> <p>Representações gráficas</p> <p>Medindo associação: duas variáveis qualitativas, qualitativa e quantitativa, e duas quantitativas</p>	4 horas

5	Introdução a probabilidade Distribuições para variáveis discretas e contínuas Distribuição Normal Distribuição amostral	4 horas
6	Estimação pontual e intervalar Intervalo de confiança para proporção Intervalo de confiança para média Cálculo do tamanho da amostra	4 horas
7	Componentes de um teste de hipóteses Teste de significância para média Teste de significância para proporção Decisões e tipos de erros P-valor	4 horas
8	Comparando dois grupos Comparando duas proporções Comparando duas médias Comparando médias com amostras dependentes	4 horas
9	Tabelas de contingência Teste qui-quadrado para independência Medindo a associação em tabelas de contingência	4 horas
10	Relacionamentos lineares O problema de mínimos quadrados O modelo de regressão linear simples Inferência para o coeficiente angular Análise de variância (ANOVA)	4 horas
Encerramento do curso		

## 9 Desempenho no curso

O desempenho no curso será determinado pela nota em atividade avaliativa aplicada aos cursistas. Requer-se, no mínimo, 70% de aproveitamento para obtenção de certificado.

## 10 Referências bibliográficas

- Magalhães, M. N.; Lima, A. C. P. **Noções de Probabilidade e Estatística**. EdUSP, 6ed. 2008.
- Agresti, A.; Finaly, B. **Statistical Methods for the Social Sciences**. Prentice Hall, 4ed, 2009.
- Triola, Mario F. **Elementary statistics using Excel**. Boston: Pearson, 2018.
- Finkelstein, M. **Basic Concepts of Probability and Statistics in the Law**. Springer, 2009.
- Foreman, John W. **Data smart: using data science to transform information into insight**. Indianapolis, IN: John Wiley & Sons, 2014.
- Dalgaard, Peter. **Introductory statistics with R**. New York: Springer, 2008. Print.

## 11 Informações sobre os conteudistas

### Wagner Hugo Bonat

Doutorado em Matemática pela University of Southern Denmark, Dinamarca(2016).  
Professor adjunto da Universidade Federal do Paraná, Brasil.

## 12 Requisitos Técnicos

Computador com acesso à internet. Permissão para instalar programas.

# CURSO DE CIÊNCIA DE DADOS APLICADA AO PODER JUDICIÁRIO

