

Cronograma 2025 (Prof. Evandro)
(Para cronograma completo, considere as aulas do Prof. Ildeberto)

Aulas Presenciais: Sexta-feira: 17:00-18:40, LEIA 1
Atendimento: Sexta-feira: 14:00-16:00, Bloco C2, sala 12
Acesso ao STOA: <https://edisciplinas.usp.br/>
Disciplina: Nome completo: RAD2607 - Ciência de Dados I (2024)
Nome breve: RAD2607-2024

1. Cronograma 2024 (Prof. Evandro, apenas)

Data	Semana	Conteúdo	Vídeo ou tutorial	Entregas
28-fev	1	Introdução à Ciência de Dados		
7-mar	2	R e RStudio, estruturas de dados no R	https://www.youtube.com/watch?v=ZFF7q5nNOHw	
14-mar	3	Transformação de dados e análise descritiva		L01: Resumo de dados
21-mar	4	Análise descritiva	https://blog.proffernandamacieli.com.br/estatistica-r/	
28-mar	5	Análise multivariada de dados		L02: Análise Descritiva
4-abr	6	Avaliação 1 – Análise Descritiva 1		A1
11-abr	7	Análise multivariada de dados		
18-abr	8	Semana Santa. Não haverá aula		
25-abr	9	Correlação linear de Pearson		L03: Análise Multivariada
2-mai	10	Recesso Dia do Trabalho. Não haverá aula		
9-mai	11	Análise de Comp. Princ. e Análise Fatorial		
16-mai	12	Análise de Agrupamentos (Cluster)		L04: Comp. Principais
23-mai	13	Avaliação 2 – Análise Multivariada		A2
30-mai	14	Modelos de Regressão		
6-jun	15	Modelos de Regressão e Classificação		
13-jun	16	Documentos no R: Quarto Markdown	https://rmarkdown.rstudio.com/	
20-jun	17	Recesso (Corpus Christi) - Não haverá aula		L05: Regressão e Classificação
27-jun	18	Avaliação 3 – Machine learning		A3

2. Critérios para avaliação (Prof. Evandro)

A avaliação (Prof. Evandro) será feita através de

- **Três Avaliações** (A1, A2 e A3);
- **05 Listas** de exercícios (L01, L02, L03, L04 e L05).

A média final (MF_Prof. Evandro) de cada aluno será determinada pelo critério a seguir:

$$\text{Critério: } MF_Evandro = 0,5 A + 0,5 L \quad (1)$$

Sendo A e L valores entre zero e dez, dados por:

A = Média aritmética obtida em 3 Avaliações = $(A1 + A2 + A3) / 3$;

L = Média aritmética obtida em **05 Listas de Exercícios** = $(L01 + L02 + L03 + L04 + L05) / 5$

2.1 Listas de exercícios (L):

- As respostas aos itens de cada lista devem ser entregues no e-disciplinas. Para obter nota máxima, além de entregar as respostas corretas a lista deve ser entregue na semana agendada (SA). É permitido entregas com atraso, porém, a nota será ajustada dependendo da semana da entrega (SE).

Considere: **SA** como sendo a **Semana Agendada** para entrega. Ex., para a L01 temos SA = 3.

SE a **Semana da Entrega** realizada. Ou seja, a semana na qual o aluno fez a entrega.

As datas de cada semana estão definidas no quadro acima, por ex. a semana 3 vai de 14 a 20 mar.

- A nota máxima de uma lista ($Nota_{max}$) é dada por:

$$Nota_{max} = 10 \quad \text{Se a SE for até a SA} \quad (2.a)$$

$$10 * [1 - (SE - SA)/18] \quad \text{Se a SE for a partir da SA} \quad (2.b)$$

O resultado será arredondado para uma casa decimal.

Observação: No cronograma, uma semana de atividades com o Prof. Evandro, tem o início numa sexta-feira e o final na quinta-feira seguinte. Por exemplo, Semana = 3 será do dia 14/março a 20/março. Cada lista está agendada para o primeiro dia de uma determinada semana. Na prática, a lista pode ser entregue em qualquer dia da SA para que nenhuma nota seja descontada por atraso.

Alguns exemplos de possíveis entregas:

(a) A Lista L01 está agendada para a semana 3 (veja no cronograma que 14-mar está indicado como semana 3), ou seja, para a L01 temos SA = 3. Se a entrega da L01 for realizada até a semana 3 então a nota máxima possível é 10. Mas se a entrega for realizada, por exemplo, na semana 10 então SE = 10 e assim: $Nota_{max} = 10 * [1 - (10 - 3)/18] = 6,1$.

(b) Considerando a Lista 01: se a entrega da L01 for realizada no dia 20-mar, então a SE = 3 (veja que 20-mar é o último dia da semana 3). Assim, $Nota_{max} = 10$

(c) Ainda com a Lista L01: se a data de entrega for 04/jul, então SE se dará na última semana, ou seja SE = 18 e $Nota_{max} = 10 * [1 - (18 - 3)/18] = 1,7$.

(d) Ainda com a Lista L01: se a data de entrega for 21/mar, então SE = 3 e $Nota_{max} = 10 * [1 - (4 - 3)/18] = 9,4$.

As entregas serão consideradas até o dia 04/julho/2025. Neste caso sendo a semana 18.

2.2 Avaliações (A):

- As avaliações serão feitas através de trabalho prático que envolve análise de dados.

- Os itens a serem desenvolvidos serão disponibilizados em arquivo pdf no site da disciplina no dia da avaliação, no horário de 17h.

- No dia de Avaliação não haverá aula. Cada aluno deverá acessar o arquivo pdf e realizar o trabalho prático, utilizando os arquivos que forem disponibilizados, como base de dados e as instruções do arquivo pdf da avaliação.

- O prazo máximo para a entrega da avaliação é 24h, sendo que a entrega deve ser feita em arquivo pdf ou html com as respostas aos itens solicitados identificando o item e a resposta. Entregas de arquivos em qualquer outro formato não serão consideradas (não entregue os scripts).

- A nota máxima em cada Avaliação é 10,0 (dez).

- Entregas em atraso:

Entregas após 24h, até 48h terão a nota máxima igual a 6,0. (3.a)

Entregas após 48h terão a nota máxima igual a 2,0. (3.b)

2.2.1 Avaliação Substitutiva (Asub):

- Segue os moldes das Avaliações, porém tem o conteúdo de "Produção de Documentos no R".

- A nota da Avaliação Substitutiva irá substituir a nota de avaliações (A), na composição da Média Final (MF) descrita na expressão (1). Ou seja, quem fizer a Avaliação Substitutiva terá a média final (MF) dada por:

$$MF_Evandro = 0,5 A + 0,5 L \quad \text{Se } A \geq A_{sub} \quad (4.a)$$

$$0,5 A_{sub} + 0,5 L \quad \text{Se } A < A_{sub} \quad (4.b)$$

3. Controle de frequência:

O controle de frequência será feito por chamada na sala de aula, com a frequência apresentada em planilha Excel a cada aula, ou então pela marcação da frequência pelo aluno, no sistema e-disciplinas. No caso dos dias de Avaliação a presença será dada pela entrega da avaliação.

4. Bibliografia

DE OLIVEIRA, P. F.; GUERRA, S; MCDONNELL, R. *Ciência de Dados com R: Introdução*, Brasília: Editora IBPAD, 2018.

FÁVERO, Luiz P.; BELFIORE, Patrícia; DA SILVA, Fabiana L.; CHAN, Betty L. *Análise de Dados: Modelagem multivariada para tomada de decisões*, Rio de Janeiro: Elsevier, 2009.

JOHNSON, Richard A.; WICHERN, Dean W. *Applied Multivariate Statistical Analysis*, 5th Ed. London, 2002.

Tutoriais online

1. The Python Tutorial. Disponível em: <https://docs.python.org/3/tutorial/>
2. Aprenda a programar. Disponível em: <https://wiki.python.org.br/AprendaProgramar>
3. R Markdown: <https://rmarkdown.rstudio.com/>
4. Quarto: <https://quarto.org/>
5. Estatística básica: <https://blog.proffernandamacieli.com.br/estatistica-r/>