

PYTHON PARA ANÁLISE DE DADOS**FICHA DO CURSO****1. Resumo do curso**

Curso: Python para a análise de dados

Modalidade: Ensino à distância

Carga horária total: 40 horas

Carga horária síncrona: 2,5 horas por semana

Início da oferta: Janeiro/2022

Duração do curso: 10 semanas

Carga horária semanal: 4 horas

Carga horária assíncrona: 1,5 horas por semana

Fim da oferta: Abril/2022

Pré-requisito: Informática intermediária (instalar softwares e executar comandos em terminais) e estatística básica (média, desvio padrão, etc).

2. Objetivos

Habilitar o(a) aluno(a) a utilizar a linguagem de programação Python para análise de dados. Realizar tarefas de leitura (entrada), processamento (tratamento e organização), manipulação e análise de dados com estatística descritiva e construção de gráficos simples.

Ao final do curso, o participante terá adquirido conhecimentos básicos sobre programação, será capaz de instalar e configurar seu próprio ambiente para programação e análise de dados com Python, importar e manipular dados em formatos textuais e estruturados, confeccionar e exportar gráficos, e salvar resultados de análises em arquivos.

3. Programa

Introdução ao Python; Funções; Entrada e Saída de Dados (linha de comando, argumentos, arquivos); Estruturas de dados (arrays, tuplas, listas e dicionários); Manipulação de strings; Bibliotecas/módulos padrão e suas funções; Numpy; Pandas; Visualização de Dados.

4. Procedimentos didáticos

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

7. Acessar o material pré aula para preparar-se para a aula.
8. Assistir as aulas programadas para a semana no horário definido.
9. Trabalhar nos tutoriais guiados disponibilizados pelo professor.
10. Fazer o estudo individual de materiais indicados como: leituras complementares, resolução de exercícios e acesso a vídeos adicionais.
11. Participar do fórum do curso contribuindo com tópicos para a discussão ou respondendo e complementando tópicos em aberto relacionados ao conteúdo apresentado.
12. Realizar as atividades de avaliação semanal de desempenho e avaliação do curso.

Como suporte ao Ensino Remoto será utilizada a infraestrutura de AVA (Ambiente Virtual de Aprendizado) que é acessado pelo endereço <https://treinamento.omegadatascience.com.br> para cursistas cadastrados.

Serão observados os seguintes detalhes sobre os procedimentos didáticos no período de vigência do curso, conforme itens a seguir.

4.1 Sistema de comunicação durante aulas ao vivo

Será usado o sistema de comunicação por mensagens instantâneas do serviço de streaming utilizado para transmissão das aulas ao vivo durante a apresentação do professor. Durante a apresentação do professor, a comunicação por texto é com a equipe de monitores do curso. Durante a aula, em momentos de pausa para dúvidas definidos pelo professor, os alunos poderão interagir por voz e vídeo.

4.2 Sistema de comunicação entre as aulas ao vivo

Será usado o recurso de fórum do AVA, permitindo que alunos abram tópicos e contribuam com a discussão ou resolução de dúvidas em tópicos abertos.

4.3 Sistema de comunicação nos plantões de dúvida

Será usado o sistema de comunicação por mensagens instantâneas, voz e vídeo do serviço de streaming utilizado para transmissão dos plantões de dúvidas.

4.4 Materiais didáticos

1. No Moodle terá o guia rápido do curso, onde constarão detalhes das plataformas de ensino, instruções sobre a comunicação com os instrutores, monitores e coordenação, cronograma e programa da disciplina, orientação das atividades remotas assíncrona e síncrona, seus prazos e pesos na determinação de desempenho e participação.
2. Dentro da conta de cada usuário na Sympla estarão disponíveis os links para a transmissão das aulas e dos plantões de dúvidas.
3. No Moodle estarão os links para acesso às transmissões gravadas das aulas. Elas ficarão hospedadas no canal do Youtube designado para o treinamento.
4. Slides (em formato pdf) usados nas aulas bem como eventuais arquivos como scripts, planilhas e arquivos de dados serão disponibilizados no Moodle.
5. No Moodle será usado o recurso de fórum para cursistas abrirem e contribuírem em tópicos, tirarem dúvidas e contribuírem com materiais.
6. No Moodle poderão ser indicados materiais complementares àqueles produzidos pelos professores do curso, como: vídeos que demonstram conceitos e ideias sobre os conteúdos das unidades didáticas e links para outros sítios eletrônicos contendo materiais relacionados aos temas das aulas.
7. No Moodle serão disponibilizadas atividades avaliativas de desempenho e participação bem como formulários de avaliação da qualidade do curso.

4.5 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	00:30	12,5
Participação na aula ao vivo	01:30	37,5
Estudo individual pós aula	00:30	12,5
Participação no plantão de dúvidas	01:00	25,0
Participação no fórum	00:30	12,5

Total	04:00	100
-------	-------	-----

4.6 Modelo de tutoria a distância

A dinâmica deste curso é disponibilizar para o cursista um serviço de autoaprendizagem que consta de vídeo-aulas síncronas com acesso posterior às gravações, disponibilização de materiais didáticos usados nos vídeos e complementares, atividades para a prática do conteúdo.

Além das aulas e respectivos materiais, o cursista terá um encontro remoto síncrono de plantão de dúvidas para interagir com o professor e pares.

No AVA haverá durante todo período de vigência do curso, fórum separados por encontro no qual os alunos poderão abrir e contribuir em tópicos diversos, como por exemplo, dúvidas em relação ao conteúdo, solicitar e/ou recomendar materiais complementares, contribuir com tópicos abertos pelos pares, apresentar resultados de atividades desenvolvidas, entre muitas outras possibilidades. Será gerado um FAQ (frequently asked questions) com as questões mais relevantes ou populares do fórum.

Em cada unidade didática, o cursista poderá avaliar seu aprendizado por meio de quizzes sobre os tópicos cobertos no encontro imediatamente anterior.

E por fim, toda unidade didática terá uma enquete anônima com questões voltadas à avaliação do curso e espaço de texto livre para expressar sua opinião.

5 Conteúdo programado

O curso está dividido em unidades didáticas (UD), sendo uma UD para cada semana de atividades. Na sequência, estão listadas as 10 UD's, com respectivo conteúdo planejado, na Tabela 2.

Tabela 2 - Cronograma detalhado do conteúdo das unidades didáticas.

UD	Conteúdo previsto	Período
1	Ambientação ao Python Instalação e configuração do ambiente Python IDEs para trabalhar com Python Acesso à documentação Variáveis e seus tipos básicos	31/01 - 04/02

Comandos de atribuição e operações básicas (soma, divisão, resto, ...)	
Realização de entrada e saída com dispositivos padrão (teclado e tela)	
2 Expressões lógicas: comparações de valores, operadores lógicos e estruturas de controle	
Estruturas de repetição: laços “para” (<i>for</i>) e “enquanto” (<i>while</i>).	07/02 - 11/02
Funções das bibliotecas padrão (pow, cos, ceil, sqrt, ...)	
3 Criação de funções: parâmetros e retornos	
Estruturas de dados simples padrão: arrays, tuplas, listas	14/02 - 18/02
Iteração e acesso a itens nas estruturas de dados padrão	
4 Estruturas de dados padrão do tipo chave-valor: dicionários	
Criação de matrizes utilizando as estruturas padrão.	21/02 - 25/02
Argumentos de linha de comando (argv)	
5 Manipulação de cadeias de caracteres (<i>strings</i>)	
Entradas e saída em arquivos não estruturados	07/03 - 11/03
Leitura de arquivos estruturados (csv, outros arquivos delimitados)	
6 Introdução ao numpy	
Instalação do numpy no ambiente de programação	
Vetores e matrizes em numpy	14/03 - 18/03
Inicialização e acesso de itens do numpy	
Importação de csv para numpy	
7 Operações lógicas em vetores numpy	
Ordenação de vetores	
Estatística em numpy: média, desvio padrão, máximos, mínimos, ...	21/03 - 25/03
Introdução ao matplotlib	
Instalação do matplotlib	
Criando gráficos cartesianos com numpy e matplotlib.	
Criando e visualizando histogramas com numpy e matplotlib	
8 Introdução ao Pandas	
Instalação do Pandas	28/03 - 01/04
Estruturas de dados do Pandas: Series e Dataframes	
Inicialização e acesso de itens do Pandas	

Carga de dados de arquivos csv Visualização de dados com auxílio do matplotlib	
9	Adição, remoção e alteração de colunas. Lidando com datas Rolling: médias móveis, desvio padrão, médias móveis 04/04 - 08/04 exponenciais, ... Dados categóricos
10	Introdução ao Seaborn Instalação do Seaborn Criando gráficos scatter e categóricos Criando e visualizando regressões
Encerramento do curso	

6 Avaliação de desempenho e participação

6.1 Desempenho no curso

O desempenho no curso será determinado pela nota média das 7 maiores notas dos 10 quizzes semanais aplicadas aos cursistas.

6.2 Participação no curso

A frequência de participação no curso será determinada pela realização dos quizzes semanais bem como pelo preenchimento das enquetes semanais de avaliação do curso ou questões do material pré-aula/pós-aula. Receberão certificados de participação aqueles com 70% ou mais de aproveitamento nestas atividades.

7 Referências bibliográficas

1. Chen, Daniel Y. **Pandas for everyone: Python data analysis**. Boston: Addison-Wesley, 2018.
2. McKinney, Wes. **Python for data analysis: data wrangling with pandas, NumPy, and IPython**. Sebastopol, CA: O'Reilly Media, Inc, 2018.

3. Haslwanter, Thomas. **An introduction to statistics with Python: with applications in the life sciences**. Switzerland: Springer, 2016.
4. Dale, Kyran. **Data visualization with Python and JavaScript: scrape, clean, explore & transform your data**. Sebastopol, CA: O'Reilly Media, 2016.
5. Matthes, Eric. **Python crash course: a hands-on, project-based introduction to programming**. San Francisco, CA: No Starch Press, 2019.
6. Zumstein, Felix. **Python for Excel**. O'Reilly Media, Inc, 2021.
7. Lutz, Mark. **Programming Python**. Sebastopol, CA: O'Reilly Media, 2010.

8 Informações sobre o professor

Abaixo estão as informações dos membros da equipe envolvidos diretamente com os cursistas.

Tabela 3 – Informações dos membros da equipe.

Função	Nome	Contato
Instrutores do curso	Andre Ricardo Abed Grégio	gregio@inf.ufpr.br
	Paulo Ricardo Lisboa de Almeida	paulo@inf.ufpr.br
Monitor do curso	A definir	A definir
Coordenador	Walmes Marques Zeviani	walmes@ufpr.br