Treinamento CIS - 1º Período

1. **Conteúdos do Período**
   1. Conceitos básicos de Data Science
   2. Python para DS
   3. Manipulação de base de dados
   4. Visualização de dados
   5. Pandas, numpy, seaborn e matplotlib
   6. Normalização de parâmetros
   7. Tipos de dados
   8. Pré-processamento de dados
   9. Redução de dimensionalidade
   10. Clustering
   11. PCA
   12. Datasets de Treinamento x Validação
   13. Regressão Linear
   14. Regressão Logística
2. **Conteúdos de Apoio**
   1. [Python Básico ao Intermediário](https://youtu.be/rfscVS0vtbw) - Vídeo que ensina todo o básico de Python.
   2. [Fundamentos de Python para Análise de Dados](https://www.datascienceacademy.com.br/course?courseid=python-fundamentos) - Curso DataScience Academy, Módulos recomendados - 5, 8 e 9
   3. [Cursos do Kaggle](https://www.kaggle.com/learn/overview) - Cursos recomendados para o primeiro período - Python, Data Visualization, Pandas e Data Cleaning
   4. [Playlist - Introdução à Data Science, Visualização de Dados e Machine Learning](https://www.youtube.com/playlist?list=PLMdYygf53DP7YZiFUtGTWJJlvynRyrna-) - Playlist brasileira que explica dos conceitos básicos até as primeiras implementações dos conceitos citados.
   5. [Python DataScience HandBook](https://github.com/jakevdp/PythonDataScienceHandbook/blob/8a34a4f653bdbdc01415a94dc20d4e9b97438965/notebooks/Index.ipynb) - Livro sobre Numpy, Pandas, MatplotLib, relativamente curto e direto.
   6. [Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow](https://drive.google.com/file/d/1S4j5ivHtVArowR5TEFFQocjQDx4T07bG/view?usp=sharing) - Livro completo: Para o primeiro período recomenda-se capítulos 1, 2, 8 e 9
   7. [Data Analysis with Dr Mike Pound](https://www.youtube.com/playlist?list=PLzH6n4zXuckpfMu_4Ff8E7Z1behQks5ba) - Playlist com diversos conceitos de análise de dados avançada.
   8. [Tipos de Dados](https://towardsdatascience.com/data-types-in-statistics-347e152e8bee) - Artigo com os tipos de dados em modelos estatísticos.
   9. [Linear Regression StatQuest](https://www.youtube.com/playlist?list=PLblh5JKOoLUIzaEkCLIUxQFjPIlapw8nU) - Playlist sobre regressão linear
   10. [Logistic Regression StatQuest](https://www.youtube.com/playlist?list=PLblh5JKOoLUKxzEP5HA2d-Li7IJkHfXSe) - Playlist sobre regressão logística
   11. [PCA StatQuest](https://youtu.be/FgakZw6K1QQ) - Vídeo explicando detalhadamente PCA
   12. [Machine Learning Sentdex](https://www.youtube.com/playlist?list=PLQVvvaa0QuDfKTOs3Keq_kaG2P55YRn5v) - Vídeos iniciais da playlist sobre Machine Learning.
3. **Tarefas**
   1. Procurar um dataset para fazer os trabalhos
   2. Fazer uma análise no dataset utilizando as ferramentas aprendidas no período
   3. No meio do período, haverá uma reunião com o monitor para consolidação das informações
   4. Demonstrar como pré processar e representar os diferentes tipos de dados
   5. Demonstrar a utilização de técnicas como one-hot encoding, redução de dimensionalidade e PCA, clustering;
   6. Aplicar regressão nos dados - dividir o dataset em treinamento e validação e aplicar regressão para prever um resultado. Testar e visualizar os resultados;
   7. As atividades descritas são apenas sugestões e podem ser modificadas;
   8. A entrega é individual e deverá ser colocada no seu GitHub pessoal