

3.3 - Entendendo o Jupyter

00;00;00;27 - 00;00;27;25

Olá pessoal! Na última aula nós fizemos um código em um bloco de notas para somar ingredientes para a compra de um bolo. Na programação, podemos fazer scripts escrevendo códigos sequencialmente em um arquivo, como fizemos na última aula. Contudo, esse arquivo único tem algumas desvantagens, pois podem ficar longos demais, pode ser difícil de encontrar algum erro, etc. Em análise de dados, é mais comum a utilização de notebooks que são como cadernos.

00;00;27;25 - 00;00;53;18

Nos notebooks podemos fazer isso em pequenos blocos de textos de código. Cada bloco é chamado de célula. Nesse curso, nós vamos usar o Jupyter Notebook, que é um ambiente que tem essa abordagem de notebooks que estamos falando. Para facilitar, vamos utilizar o Google Colab, que é um serviço online do Jupyter Notebook e não precisa de configuração ou instalação no computador, e é totalmente gratuito.

00;00;53;26 - 00;01;14;23

Mas se quiserem rodar o Jupyter no pelo computador de vocês, a gente vai deixar aqui alguns materiais complementares de como instalar e executar. Vamos já aprender a abrir o Google Colab pra visualizar como é esse tal de Jupyter Notebook. Bom, vamos lá no Google Drive, na pasta que criamos o curso da PrograMaria, pra ficar todo organizadinho.

00;01;15;00 - 00;01;41;12

Aqui dentro da pasta, a gente clica com o botão direito, vai em “mais”. Bom, aqui na verdade já pode aparecer a opção do Google Colaboratory se você já tiver instalado aí. Se não aparecer, você vai aqui em “mais”. Vai em “conectar mais apps”. Aí nessa barra de pesquisa, a gente coloca “colaboratory”. Aqui já vai aparecer pra gente poder clicar. Nessa primeira opção aqui, a gente clica e vai em “instalar”.

00;01;42;17 - 00;02;20;00

Bom, aqui em “continuar”. A gente seleciona qual que é a conta, no caso, a minha. Vai em “continuar” e vai ser feita a instalação do Colaboratory. Concluído. Bacana. Concluído. Ok. Agora a gente já tem Colaboratory instalado, a gente pode fechar essa aba e, de novo, clicar com o botão direito aqui dentro da pasta. Não apareceu aqui, a gente pode ir em “mais”.

00;02;20;07 - 00;02;43;00

E aqui está o nosso Google Colaboratory. Aí a gente já vai abrir um notebook aqui em branco para a gente começar a programar. Vamos ver: criar e compartilhar. Bom, aqui a gente tem o notebook. Vou aumentar um pouquinho para ficar melhor de a gente visualizar. Perfeito. Nessa barra de ferramentas aqui em cima, a gente tem duas opções: a de código e de texto.

00;02;43;06 - 00;03;07;08

É intuitivo, pois quando a gente quiser adicionar mais uma célula de código, a gente clica em código, mas a gente quiser adicionar um texto, imagem, etc, a gente clica em texto. Vamos começar pela célula de código para reproduzir a primeira coisa que a gente programou, que é o nosso “Olá mundo!” Vamos aqui nessa primeira célula de código e digitar print

00;03;08;19 - 00;03;35;04

Olá mundo! E aí, para executar essa célula a gente aperta ctrl+enter. Viu que mágico? Já apareceu aqui o nosso Olá mundo! A gente pode apertar ctrl+enter para executar a célula. Ou a gente pode ir aqui nesse símbolo aqui de executar, que vai aparecer também. O Jupyter é uma plataforma de código interativo, então conseguimos executar nossos códigos sem precisar do terminal nem nada disso.

00;03;35;16 - 00;04;04;07

E cada célula pode ser executada de forma independente, uma por vez. Mas lembre-se que a ordem da execução das células importa. Vamos supor que eu crie duas células de código. Vamos lá: duas. Na primeira, eu coloque $a=1$ e na segunda eu coloque $a+2$. E aí, se eu executar a segunda célula antes da primeira, vai dar erro.

00;04;05;21 - 00;04;29;05

Por quê? Isso porque a gente ainda não guardou que $a=1$. Então, se eu rodar a primeira primeiro aqui e depois rodar a segunda célula, aí ele vai me dar o resultado de que $a=3$. Legal, né? Agora vamos repetir o nosso código de somar lista de compras, que foi o que a gente fez na última aula.

00;04;30;13 - 00;04;54;25

Eu tenho aqui aberto, aliás, vou abrir agora a soma ingredientes que foi o que a gente fez na última aula. Vou abrir com um txt, com um bloco de notas que foi onde a gente editou, para ficar mais fácil. Beleza. Vou voltar aqui no meu notebook e vou criar uma célula de texto só pra a gente poder ver que acontece.

00;04;56;07 - 00;05;20;00

Veja bem, a estrutura de célula de texto, ela é diferente da que temos para código, com várias opções, inclusive de anexar alguma imagem, foto. Pra aprendemos melhor, vamos reproduzir o script da soma ingredientes que fizemos na última aula aqui nesse notebook e já vamos deixar aqui aberto pra gente poder só copiar e colar uma na outra. Mas para ficar organizado, eu vou colocar nessa célula de texto aqui bonitinho um “soma ingredientes” que funciona como um título para o nosso código.

00;05;20;02 - 00;05;46;21

Ficou bem grande, bonitinho. E já utilizando aquela ideia de caderno, de anotações, a gente pode escrever alguma coisa que vai nos auxiliar. Eu escrevi aqui soma ingredientes, mas se a gente estiver mexendo com algum outro código, você pode colocar aqui o que é esse notebook pra te auxiliar e ficar mais organizado.

00;05;47;17 - 00;06;04;27

Bom, a gente já tem aqui o nosso bloco de notas aberto e vamos fazer exatamente na ordem que está no bloco de notas. Primeiro vamos fazer as importações das bibliotecas, que foi o que a gente fez aqui no início. Eu falei no plural, mas a gente tem só uma biblioteca que a gente importou, que foi essa sys.

00;06;06;23 - 00;06;33;16

Bom, veja bem, aqui pra poder colocar um comentário, eu vou usar esse jogo da velha e aí eu vou colocar "importações". A gente não precisa comentar tudo, mas é legal deixar alguns comentários porque a gente não sabe quem vai ou quando alguém vai pegar o nosso código depois. Às vezes pode ser que a gente mesmo pegue o nosso código e se não tiver nenhum comentário, a gente fica assim "nossa, o que foi que eu escrevi aqui?"

00;06;33;28 - 00;07;07;16

Então é importante a gente deixar algum comentário. Então vamos lá copiar do bloco de notas essa importação. Import. Vou colocar aqui, vou executar essa célula. A gente já fez a nossa importação. É possível a gente colocar uma célula para cada ingrediente, mas pra deixar tudo organizado, vamos colocar tudo que é declarado de variável nessa mesma célula. Então vamos colocar aqui um comentário "declaração de variáveis".

00;07;07;16 - 00;07;37;26

Aqui a gente copia tudo que está aqui e coloca aqui. Vou tirar esses comentários que a gente colocou depois porque na última aula a gente estava só conferindo se o ingrediente estava na lista ou não. Bom, perfeito. Nessa célula a gente tem, então, a declaração das variáveis de preços dos ingredientes. Vamos executar ela e vamos pra uma próxima célula de códigos.

00;07;38;05 - 00;08;04;22

Bom, gente, aqui já declaramos as variáveis e vocês podem ter notado que aqui, do lado de cada célula está aparecendo um número: seis, sete... Se a gente for aqui no primeiro: dois, quatro. Esses números indicam a ordem que a gente está executando as células. Por exemplo, se eu executar essa de importação agora de novo, vai aparecer o número oito. Veja bem, apareceu o número oito.

00;08;04;24 - 00;08;28;08

Se eu executar essa de baixo, vai aparecer o número nove e assim vai... Então se eu executar dez e voltar e executar essa declaração de variáveis de novo, vai ser a onze, que vai ser a ordem que a gente está executando as células. E esse outro número que está aparecendo aqui é só o número de linhas. Então, aqui nessa célula, um aqui é a primeira linha, o dois a segunda linha e assim vai...

00;08;28;20 - 00;08;53;04

É muito legal pra gente poder se orientar, porque às vezes em algumas células a gente pode colocar uma função que seja muito longa e a gente consegue ver qual linha. E até mesmo quando aparece algum erro, no próprio erro pode indicar "olha, erro em tal linha" e a gente consegue vir aqui e ver qual é. Bom, voltando: no caso de função, não podemos iniciar a função em uma célula e terminar em outra, porque a função é um bloco fechado.

00;08;53;13 - 00;09;29;03

Ou seja, quando a gente clicar em executar, precisamos executar a função inteira. Se a gente dividir em duas células, não vai ter executado a função toda, beleza? Ótimo, executamos essa linha de declaração. Vamos pra função. Vamos copiar a função todinha aqui e colar lá na nossa célula. Aqui está ela. Vamos executar e vamos para mais uma célula de código.

00;09;29;03 - 00;09;58;24

Agora a gente vai pra... Qual é a próxima parte? Vamos dar uma olhada aqui. Ah, ótimo, mais uma função que é a função do if name. Vamos copiar toda também e vamos colocar na nossa célula de código. Ótimo, vamos executar. Opa, deu erro. Bom, isso porque quando executamos o Jupyter, não executamos usando o terminal. Então a gente não está chamando o script como fizemos na última aula.

00;09;59;08 - 00;10;22;12

Ou seja, a gente não precisa desse if main, a gente nem precisa mais da biblioteca sys. O objetivo de utilizar a biblioteca era para conseguirmos pegar todos os ingredientes que são enviados via terminal. Se a gente não tem terminal mais, para que a gente vai usar essa biblioteca? Bom, então vamos retirar esse if aqui e vamos apagar, lembrando de voltar a indentação. Como que a gente pode fazer voltar essa indentação de uma forma muito fácil?

00;10;22;23 - 00;10;46;18

A gente pode selecionar tudo e colocar o shift + tab. Apertando o shift + tab, vai voltar. Cada vez que você apertar, é uma indentação que volta. Como a gente tinha duas, a gente aperta duas vezes shift + tab e aí já tem a indentação corretamente. Bom, não vamos mais usar a biblioteca sys, então podemos retirar essas partes aqui que tem a chamada da biblioteca de todos os itens.

00;10;46;28 - 00;11;14;26

Agora nós podemos só escrever true ou false na frente de cada variável pra dizer se tem ou não tem um item. Por exemplo, aqui nesse "tem cenoura", a gente pode apagar isso aqui tudo e colocar true. Isso para todas as variáveis. Vamos lá... Bom, e aí a gente tem o nosso código assim e quando a gente clicar para executar, já vai aparecer a soma bonitinha.

00;11;15;01 - 00;11;40;00

Vamos lá executar. Aqui a gente tem total dos ingredientes: 54,5. Fácil, não é? Bem mais fácil do que mexer no script, abrir terminal e fazer orações para rodar e tudo mais. E é por isso que o Jupyter é mais utilizado para análise de dados, pela facilidade de compreensão e correção do código, podendo descrever toda uma história de análise aqui dentro desse notebook.

00;11;40;05 - 00;12;17;00

Se em uma nova célula, a gente mudar apenas a variável do preço cenoura... Assim... Vamos colocar aqui e copiar lá o preço cenoura, levar lá pra baixo nessa nova célula de código e mudar esse preço aqui pra um preço cenoura: oito. Aí, se a gente executar essa célula e depois voltarmos na célula anterior que faz a soma e executar de novo, nosso resultado que era 54,5, se a gente executar foi pra 58. Viu que o total da compra mudou?

00;12;17;11 - 00;12;40;11

Os valores ficam na memória do programa. Por isso é preciso tomar cuidado com a ordem da execução das células. Ah, não vamos esquecer de nomear o nosso notebook. Aqui em cima a gente pode alterar o nome do arquivo. Vamos colocar “soma ingredientes”, porque a gente é muito criativo, soma ingredientes. Vejam que esse notebook está com a extensão de ipynb.

00;12;40;11 - 00;13;00;08

Isso é de Python notebook. Toda vez que a gente for fazer uma análise de dados no notebook, no ambiente do Jupyter, a gente usa essa extensão. Se a gente for executar no terminal, a gente usa a extensão .py, que é de script. Beleza? A gente não conseguiria usar o script aqui, por exemplo. Legal demais.

00;13;00;08 - 00;13;24;16

Estamos empolgadas para continuar. Então, nessa aula nós aprendemos o que é Jupyter Notebook, Google Colab e refizemos a soma ingredientes em um notebook. Já na próxima aula nós vamos refazer as análises do módulo dois, mas agora utilizando o Colab. Bom, gente, uma coisa que é especial do Google Colab: a gente tem um ambiente de execução, onde a gente vai colocando todo o nosso código e tudo mais.

00;13;24;27 - 00;13;56;06

Se a gente passar um bom tempo sem mexer, deixar a página aqui aberta, pode ser que quando a gente voltar, esse ambiente de execução tenha desconectado. E aí vai aparecer uma mensagem “Olha, você foi desconectada” ou “você está por aí?”. Já apareceu comigo: “você está por aí?”, tipo Netflix. Não se preocupem. Quando acontecer isso, vocês só vão aqui em ambiente de execução e executar tudo, que aí vai executar todas as células e vai parar lá na última que você executou. A gente só faz assim e automaticamente vai executar tudo, ok?