

eu ainda lembro do tempo que para trabalhar com data Science você tinha que gastar uma parte do seu tempo codificando ou realizando processos complexos manualmente eu ainda lembro do meu processo de entrevista de emprego no Google quando eu acabei trabalhando lá como analytical Lead que eu tive que fazer todo o reverse Engineering de um processo de ab testing Durante a entrevista de emprego tudo isso na mão e hoje para fazer um processo de ab testing eu tenho ferramentas que tornam esse processo muito mais simples do que mexer numa planilha isso porque com a evolução das ciências de dados várias e ferramentas foram evoluindo para tornar o processo mais fácil e mais prático Além de muito mais rápido e é isso aí que a gente vai ver no vídeo de hoje as ferramentas mais importantes de data Science que hoje essas ferramentas fazem parte da vida da maioria dos cientistas de dados eu vou separar as ferramentas por grupos para que a gente possa entender melhor primeiramente eu vou primeiro trabalhar com as novidades as novas ferramentas e depois eu vou trabalhar com as mais comuns e para começar na nossa lista eu vou trazer uma ferramenta inusitada que é bastante utilizada por agências principalmente na área de predictive Analytics Você entendeu o poder da essa ferramenta ela faz o trabalho que muitas vezes levaria horas e horas de codificação em poucos minutos e com uma acuracidade incrível a ferramenta se chama aqu em uma única ferramenta você consegue preparar limpar e trabalhar os dados realizar forecasting e fazer análises criar Ed haw analyses utilizar nlp para conversar com os dados criar relatórios utilizando Inteligência Artificial generativa além de todas as integrações que a ferramenta possui é uma ferramenta sensacional principalmente se você é um consultor de data science ou se você trabalha com data Science Analytics para agências ou para Marketing a próxima ferramenta da nossa lista é o Big ML ML de machine learning essa ferramenta é outro recurso das ciências de dados de alto nível que fornece aos usuários um ambiente completamente interativo e baseado em Cloud e ela é ideal para você processar aplicativos de machine learning você pode criar uma conta gratuita ou Premium dependendo das suas necessidades Mas você já pode ir lá agora e utilizar a ferramenta com essa ferramenta você pode criar soluções empresariais de Mach learning automaticamente você pode pegar padrões de dados preditivos e

transformá-los em aplicativos práticos e inteligentes que você vai estar utilizando pra sua empresa e aí todo o processo de data Science Analytics dependendo da empresa ou Dependendo do que você for fazer você pode fazer através dessa ferramenta a próxima ferramenta na nossa lista é o data Robot que é uma plataforma de machine learning que ela é projetada para as empresas implementarem modelos de machine learning dentro de processos específicos Nos quais elas vão estar trabalhando o objetivo da ferramenta é democratizar utilização dos ciências de dados como é que eles fazem isso pensa que você tem técnicas avançadas de ciências de dados que você precisa de um grande trabalho grandes especialistas para trabalhar naquilo a ferramenta reúne tudo isso de uma maneira fácil ou seja projetos complexos que você precisaria de um cientista de dados um time grande de ciências de dados para estar desenvolvendo Você vai precisar de um ou dois cientistas de dados para estar trabalhando naquele projeto com as ferramentas dentro da própria plataforma as ferramentas que eles possuem facilitam o processo de implementação de criação e de manutenção de sistemas preditivos ele automatiza o processo de machine learning desde a preparação dos dados até a seleção dos modelos além de todo o ajuste de parâmetros tudo isso você faz na ferramenta isso permite que os usuários criem modelos avançados de altíssima qualidade rapidamente e sem precisar ter um profundo conhecimento de ciências de dados a próxima ferramenta da nossa lista é chamada de Nime ela é projetada para criar fluxos de trabalho de ciências de dados de uma forma visual permitindo análise modelagem e a manipulação de dados de uma maneira muito fácil sem exigir conhecimento de programação e ela cuida de processos que vão desde a limpeza de dados até implementação desses modelos de machine learning pros trabalhos que você estiver realizando ela faz previsão de ch de clientes detecção de frauds previsão de demanda e muitas outras atividades automaticamente Além disso ela possui nodes para mineração de texto e processamento de nlp processamento de linguagem natural ele também é bastante utilizado na área de mercado de trading para tendências de mercado bom e saindo dessas ferramentas incríveis que não é todo mundo que conhece vamos para as ferramentas mais comuns utilizadas no dia a dia de um cientista de dados eu vou começar pelas

ferramentas de visualização existem cinco ferramentas de visualização mais comuns e de modo algum você precisa conhecer todas elas dessa forma as ferramentas de visualização mais importantes são minitab Power bi tablô matplotlib Lib e looker cada uma dessas ferramentas vai ser mais interessante para específicos trabalhos que você estiver realizando por exemplo por exemplo ler é muito interessante se você trabalha na área de digital marketing Market Science por exemplo porque ela já tem várias implementações Integradas a ferramenta todas as ferramentas elas fazem trabalhos muito parecidos as diferenças são em algumas das funcionalidades que você tem nelas e aí qualquer uma dessas ferramentas que você aprender Tá bom se você quiser aprender apenas uma e usar essa como parâmetro para todas as outras que você precisar mudar ou mudar de empresa e utilizar Eu recomendo ir pro tablô ou pro Power bi se você trabalha na área de digital marketing ou trabalha na parte de agências Eu recomendo end Aria ir pro ler mas existe uma outra ferramenta que pouca gente fala que é o D3 DS o D3 DS é uma biblioteca JavaScript usada para criar visualizações de dados dinâmicas além de interativas elas permitem que os usuários da ferramenta vinculem os dados a projetos Dom ela permite vincular os dados a esses elementos possibilitando a criação de visualizações interativas através delas Mas além dela tem uma outra ferramenta extremamente importante que 95% de cientistas de dados usam ela tá presente em todas as empresas eu uso pelo menos uma vez por semana mas pouca gente fala que é o Excel Além disso ele possui várias funcionalidades avançadas que podem ser utilizadas para ciências de dados ele tem várias ferramentas para limpar e preparar os dados funções estatísticas funções matemáticas que ajuda você fazer análises descritivas e inferenciais além de facilitar muito a análise exploratória de dados ou o eda Além disso você utilizando o solw no Excel você consegue realizar problemas comp complexos de decisão como maximização de lucros ou minimização de custos através de técnicas de programação linear e não linear Além disso você pode rodar Python no Excel e vincular ele a outros projetos de data Science que você tá fazendo utilizando Python além da Ampla Gama de funções estatísticas como a nova desvio padrão e várias outras coisas que você pode estar realizando na ferramenta dito isso vamos agora pras ferramentas Python

quando a gente fala de machine learning e nlp as ferramentas mais comuns são o kit learn tensor Flow P torch Mat plot Lib num Pie e pandas cada uma dessas bibliotecas tem funcionalidades específicas e áreas de aplicação específicas por exemplo py kit learn é uma biblioteca de machine learning Python que fornece ferramentas simples e eficientes para mineração de dados ou análise de dados já o p torch é uma ferramenta de Deep learning uma biblioteca que você utiliza para Deep learning Mat plot Lib é uma ferramenta de plotagem 2D que permite que você Gere gráficos através do do Python e o npai é uma ferramenta uma biblioteca básica bastante avançada na verdade de computação científica eu disse básica porque ela é fundamental Ou seja no mesmo projeto você pode carregar várias bibliotecas diferentes que vão estar executando funções diferentes dentro daquele seu arquivo Resumindo cada uma dessas ferramentas desempenha um papel crucial na área de ciências de dados pyc learn focado em machine learning tor Floor e pytorch são especializados em Deep learning Mat plot Lib trabalhando com a visualização de dados não em computação numérica e pandas na análise de dados e manipulação de dados saindo do Python existem outras ferramentas que dependendo da sua função empresa que você estiver trabalhando que você pode estar usando também entre elas a gente tem o sas o rup e o apach Spark o sas já é uma plataforma integrada Tenta colocar tudo dentro deles e o que que eles te fornecem uma ampla Gama de ferramentas e soluções para análise de dados ela é utilizada geralmente quando você tá gerenciando grandes volumes de dados e você quer fazer análises sofisticadas a partir dele já o hadoop é um software que permite o processamento de grandes quantidades de dados através de um cluster de computadores usando modelos de programação simples já o apach spar oferece processamento em tempo real e em Bat e depois dessas ferramentas não podiam faltar as ferramentas de narrative ai a mais comum delas como todo mundo sabe é o chat de PT que tem sido cada vez mais utilizado para ciências de dados seja diretamente seja através das integrações Ele oferece a possibilidade de você gerar código Python e executar esse código além de gerar relatórios de análises bastante complexo e ele vem equipado com uma variedade de plugins que você pode utilizar para matemática estatística experimentação pesquisa automação revisão de

código documentação e diversas outras coisas que são tarefas comuns de data Science e o mais importante é que existe uma tendência muito grande não só da adoção do chats de PT e outras ferramentas de narrative ai como tá evoluindo muito que essas ferramentas podem fazer e por último a ferramenta que eu não citei aqui antes depois eu vou fazer um vídeo sobre ela e eu uso de vez em quando já usei bastante alguns projetos que é o Google bigquery que tem ficado muito mais avançado você inclusive inclusive pode fazer consultas em nlp dentro do próprio bigquery hoje E aí nos próximos vídeos eu vou falar um pouco mais sobre ele e aí trazer um pouco mais dessas outras ferramentas para vocês também legal Muito obrigado quem participou do vídeo de hoje não deixem de se cadastrar aqui se inscrever dá o seu like qualquer dúvida que você tiver em relação a ferramentas ou outras áreas de ciências de dados análise de dados marketing Science que é a minha especialidade perguntem aqui eu vou estar respondendo para vocês um abraço e até a próxima