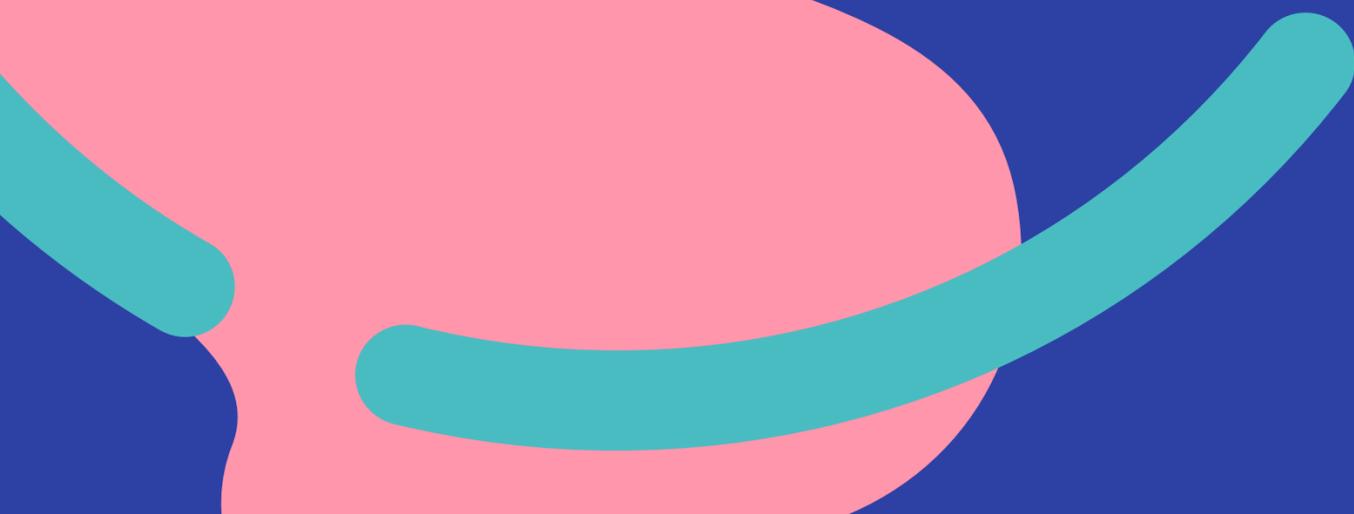


progr{a}m{aria}



ROADMAP PARA AS ÁREAS DEDADOS

Vem cá ver o que você precisa
saber para começar!





CIÊNCIA DE DADOS

Uma das profissões que vamos conhecer por aqui é a Ciência de Dados, uma profissão onde você vai poder trabalhar com machine learning e muitas coisas incríveis. Bora lá?



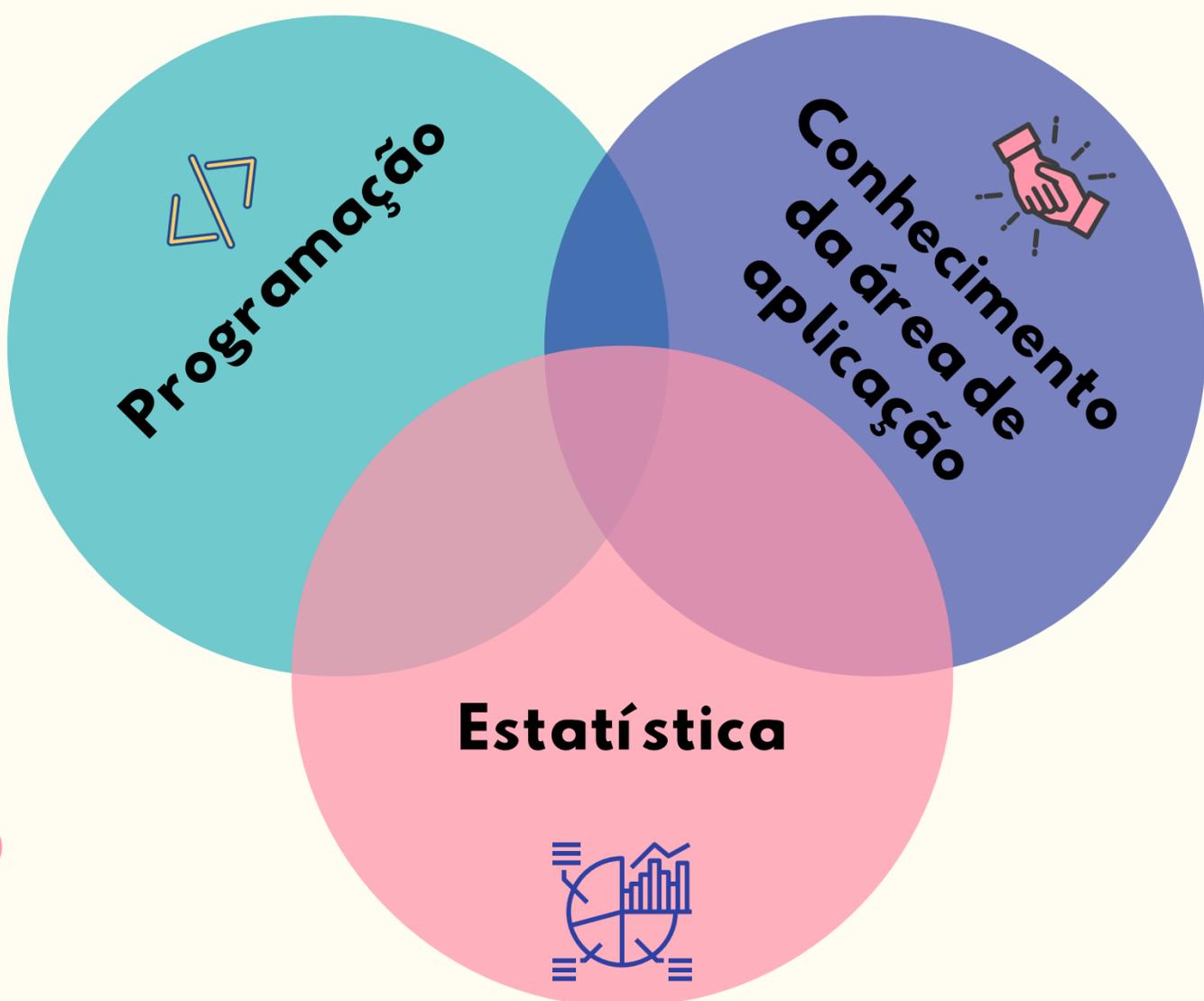


O QUE É CIÊNCIA DE DADOS?

A ciência de dados é uma área multidisciplinar que nos permite transformar dados brutos em conhecimento, principalmente através de modelos.

*

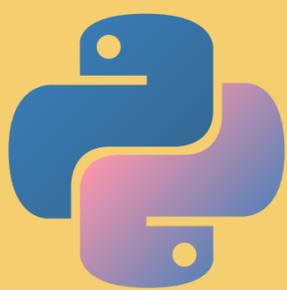
Hard skills necessárias



*Hard skills: conhecimentos técnicos

1. Programação

Vamos passar pela primeira etapa das hard skills: a programação. Quando falamos de ciência de dados, temos duas linguagens-chave:



Python



Python

1. Variáveis
2. Operadores
3. Estrutura básica de dados
4. Loop de repetições
5. Estrutura de condições
6. Métodos de funções
7. Bibliotecas básicas: numpy, math, SciPy.
8. Tratamento de erros

1

2

SQL

1. Introdução a banco de dados
2. Consultas em geral
3. Manipulação de dados
4. Joins
5. Funções

3

Manipulação de dados com Python

1. Biblioteca: Pandas

4

Visualização de dados com Python*

1. Biblioteca: Matplotlib

2. Biblioteca: Seaborn

5

Modelagem de dados com Python*

Biblioteca: Scikit Learn

1. Regressão Linear – Simples

2. Regressão Linear – Múltipla

3. Análise de Resíduos

4. Regularização L1 e L2

5. Modelos Lineares Generalizados

6. Regressão polinomial

7. Naives Bayes

8. Árvores de Decisão

9. KNN

10. Regressão Logística

11. SVM (Máquinas de vetores de suporte)

12. K-means

13. DBSCAN

14. PCA

15. Métricas de avaliação de modelos:

Acurácia, F1 Score, Precisão, Recall, Matriz de Confusão, Area Under the ROC Curve, Log Loss, Mean Squared Error, Mean Absolute Error, Mean Absolute Percentage Error, R2, Kolmogorov Smirnov.

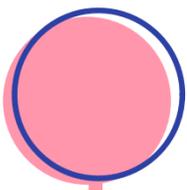
***Estudar estas etapas após o estudo teórico de estatística**

2. Estatística

Vamos passar pela segunda etapa das hard skills: a estatística.

Estudo teórico de estatística:

1. Estatística Descritiva Univariada:
Média, Moda, Mediana.
2. Medidas de Dispersão:
Amplitude, Desvio-Médio,
Variância, Desvio-padrão, Erro-padrão.
3. Medidas de Assimetria e Curtose
4. Medidas de Associação entre
Variáveis Quantitativa:
Diagrama de dispersão, Covariância, Correlação.
5. Probabilidade:
Experimento Aleatório, Espaço amostral,
Eventos, Operações com eventos,
Eventos mutuamente excludentes,
Eventos independentes,
Probabilidade condicional,
Teorema de Bayes.



Estudo teórico de estatística (continuação):

6. Variáveis Aleatórias e Distribuições de probabilidade:

Esperança Matemática, Variância, Covariância, Coeficiente de correlação.

7. Distribuições discretas:

Distribuição de Bernoulli, Distribuição binomial, Distribuição de Poisson.

8. Distribuições contínuas:

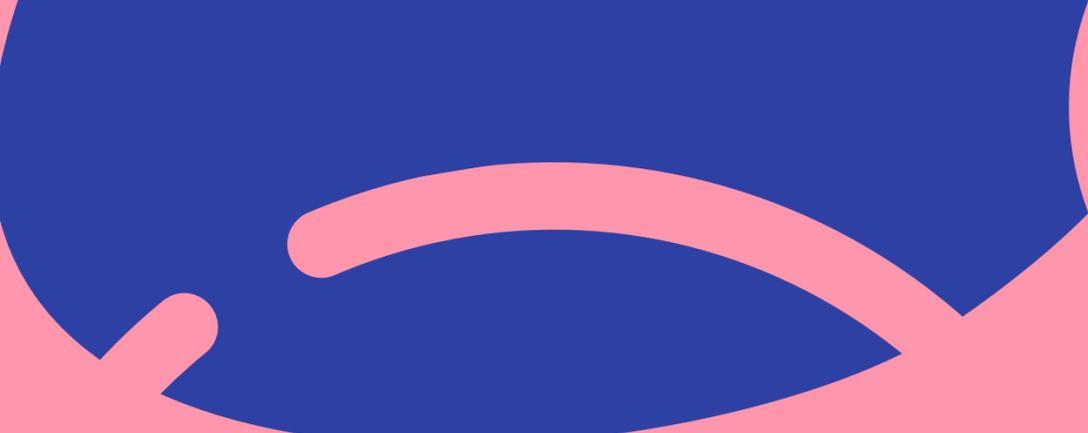
Distribuição uniforme, Distribuição normal, Distribuição Log-normal.

9. Distribuições de probabilidade conjunta:

Distribuição conjunta de variáveis discretas, Distribuição conjunta de variáveis contínuas.

10. Inferência Estatística e Principais Teoremas de Probabilidade:

Teorema de Tchebycheff, Lei dos grandes números, Teorema Central do Limite, Intervalo de confiança, Testes de hipóteses.



3. Conhecimento da área de aplicação

Essa é nossa última etapa de hard skills em ciência de dados. No entanto, ela é uma etapa que depende muito da área de aplicação que escolhermos seguir. Podemos aplicar ciência de dados na saúde, na oferta de crédito, na agricultura e até mesmo no marketing.

Descubra sua área de aplicação favorita

1

Nessa etapa, não há certo ou errado. Pesquise bastante sobre as áreas disponíveis no mercado e que combinam com você e com seus gostos. Não tenha medo de explorar várias áreas. Hoje temos diversos eventos de empresas e/ou comunidades que mostram onde os dados podem ser aplicados. Participe sempre que puder e conheça tudo que há para fazer. O céu é o limite!





Procure pessoas que trabalham na área

É muito importante que você procure pessoas que já estão atuando na sua área de interesse. Chame-as no LinkedIn e peça indicação de conteúdos que possam te ajudar nessa jornada. Não tenha medo de pedir ajuda!

2

3

Encontre comunidades que possam te oferecer apoio

Lembre-se: você não está sozinha. Há várias mulheres que estão passando pelo mesmo que você e várias que já passaram por isso. Vá até comunidades que possam te oferecer suporte e sororidade.

4

O aprendizado é contínuo e deve ser compartilhado

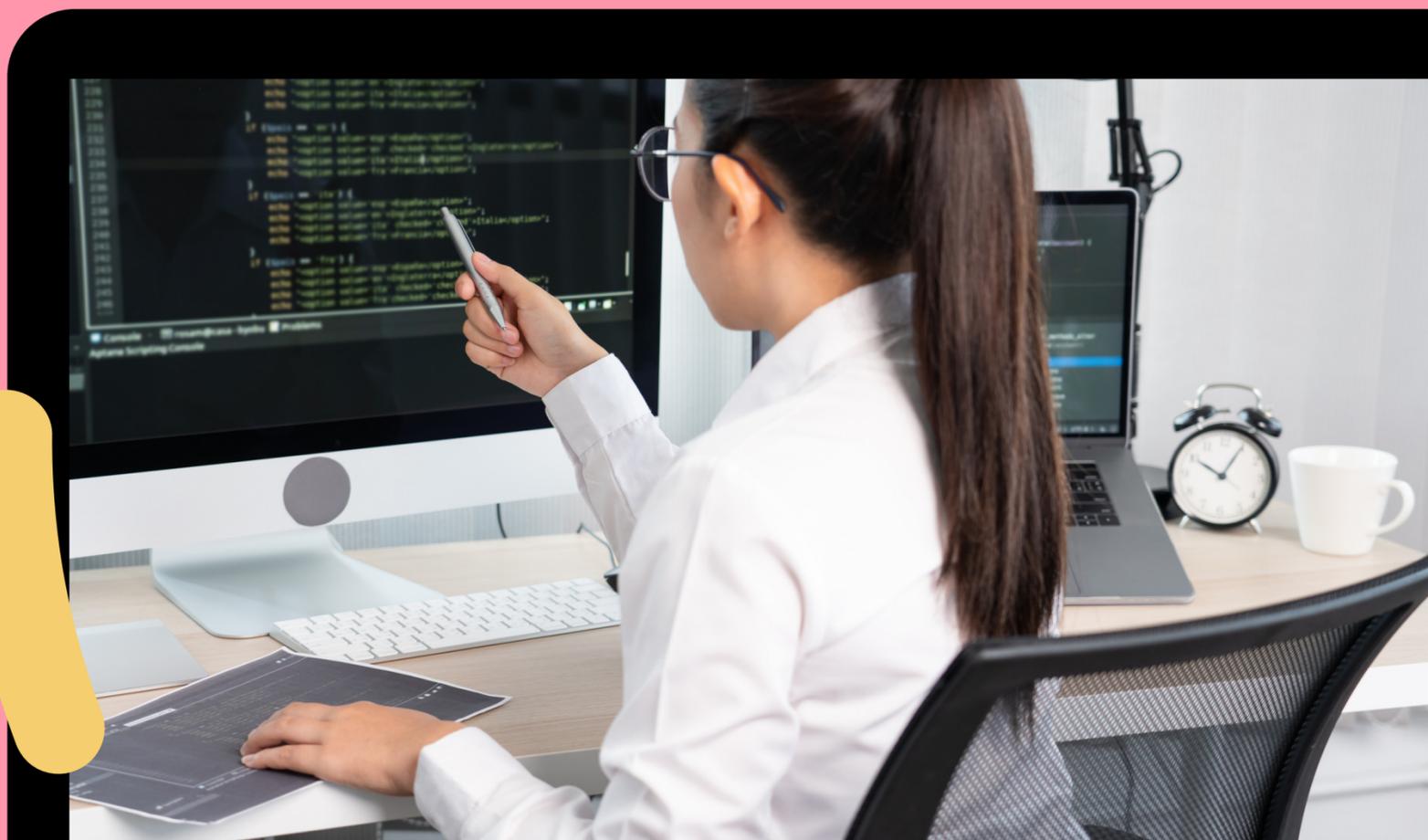
Mesmo depois de estar na área, suas soft skills vão continuar evoluindo. Compartilhe seus aprendizados com mais mulheres - além de ajudá-las, você estará se desenvolvendo também.





ANÁLISE DE DADOS

A segunda profissão que vamos conhecer por aqui é a análise de dados, área na qual você vai poder trabalhar com visualizações de dados, trazendo muito conhecimento para compartilhar. Vem com a gente!

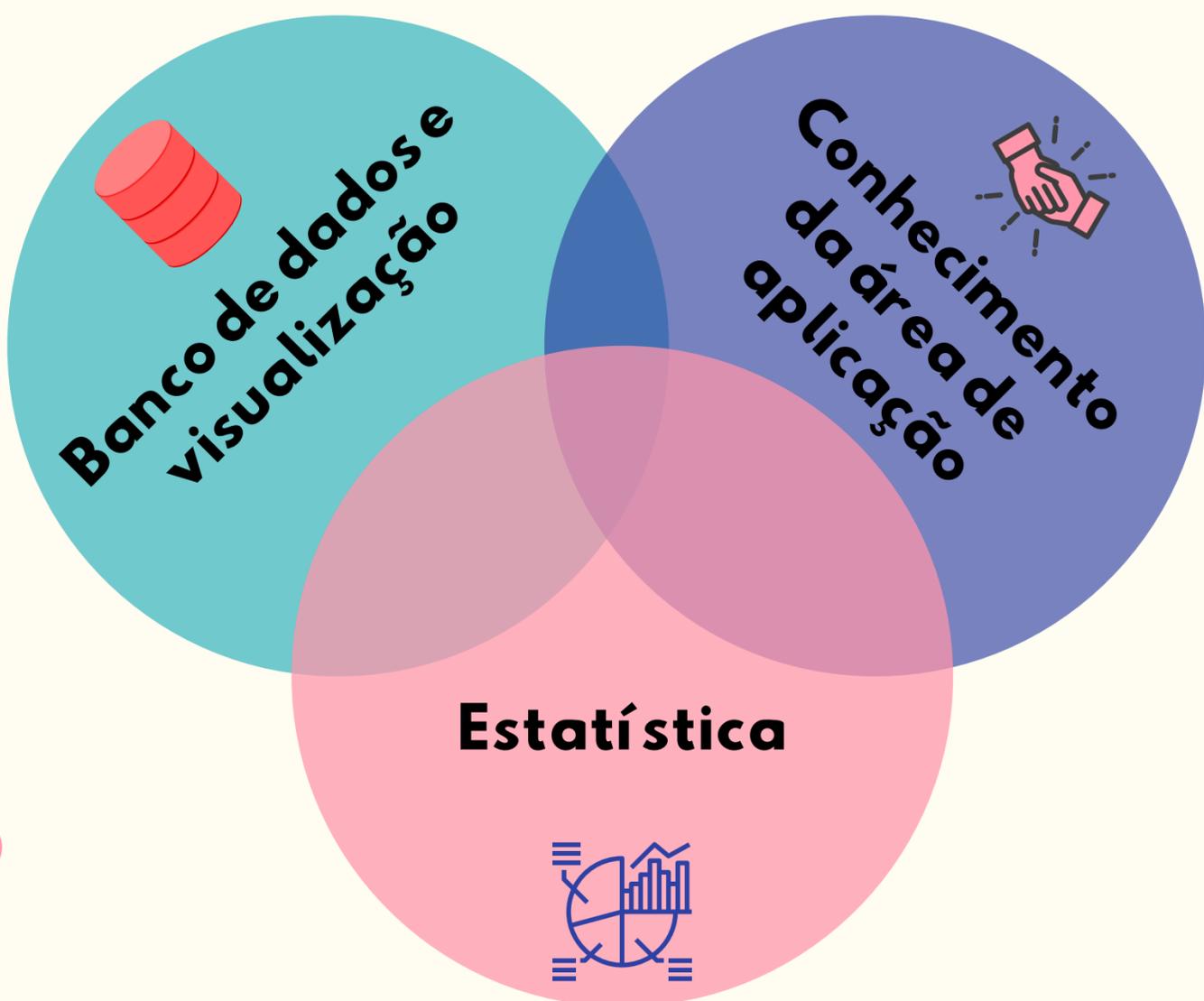




O QUE É ANÁLISE DE DADOS?

A análise de dados também é uma área multidisciplinar que nos permite transformar dados brutos em conhecimento, mas com o foco em visualização dos dados.

* **Hard skills** **necessárias**



***Hard skills**: conhecimentos técnicos

1. Banco de dados e visualização

Vamos passar pela primeira etapa das hard skills: a parte de banco de dados e visualização. Quando falamos de análise de dados, focamos sempre em ferramentas de coleta, manipulação e visualização de dados.

Excel

1. Funções de análises
2. Gráficos
3. Manipulação de dados

1

2

SQL

1. Introdução a banco de dados
2. Consultas em geral
3. Manipulação de dados
4. Joins
5. Funções

3

Ferramentas de visualização de dados

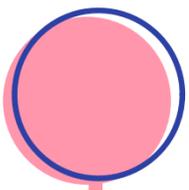
1. Tableau
2. PowerBI
3. Qlik Sense
4. Google Data Studio

2. Estatística

Vamos passar pela segunda etapa das hard skills: a estatística. **Este tópico é idêntico ao tópico de estatística de ciência de dados.**

Estudo teórico de estatística:

1. Estatística Descritiva Univariada:
Média, Moda, Mediana.
2. Medidas de Dispersão:
Amplitude, Desvio-Médio,
Variância, Desvio-padrão, Erro-padrão.
3. Medidas de Assimetria e Curtose
4. Medidas de Associação entre
Variáveis Quantitativa:
Diagrama de dispersão, Covariância, Correlação.
5. Probabilidade:
Experimento Aleatório, Espaço amostral,
Eventos, Operações com eventos,
Eventos mutuamente excludentes,
Eventos independentes,
Probabilidade condicional,
Teorema de Bayes.



Estudo teórico de estatística (continuação):

6. Variáveis Aleatórias e Distribuições de probabilidade:

Esperança Matemática, Variância, Covariância, Coeficiente de correlação.

7. Distribuições discretas:

Distribuição de Bernoulli, Distribuição binomial, Distribuição de Poisson.

8. Distribuições contínuas:

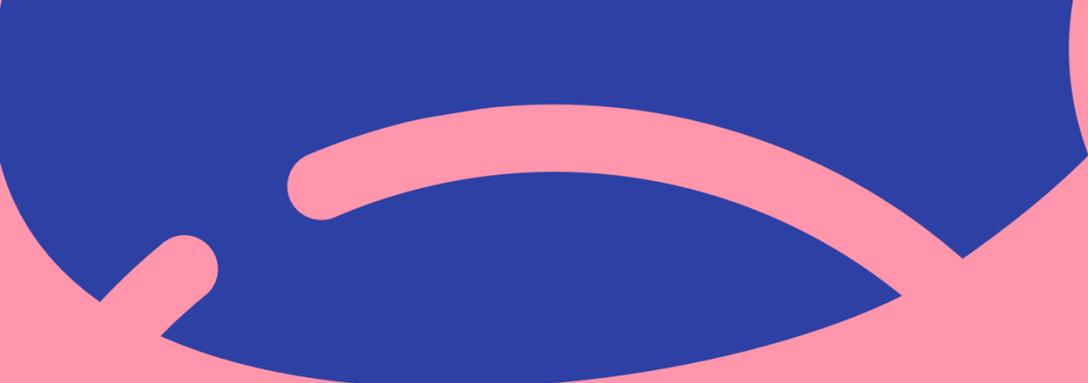
Distribuição uniforme, Distribuição normal, Distribuição Log-normal.

9. Distribuições de probabilidade conjunta:

Distribuição conjunta de variáveis discretas, Distribuição conjunta de variáveis contínuas.

10. Inferência Estatística e Principais Teoremas de Probabilidade:

Teorema de Tchebycheff, Lei dos grandes números, Teorema Central do Limite, Intervalo de confiança, Testes de hipóteses.



3. Conhecimento da área de aplicação

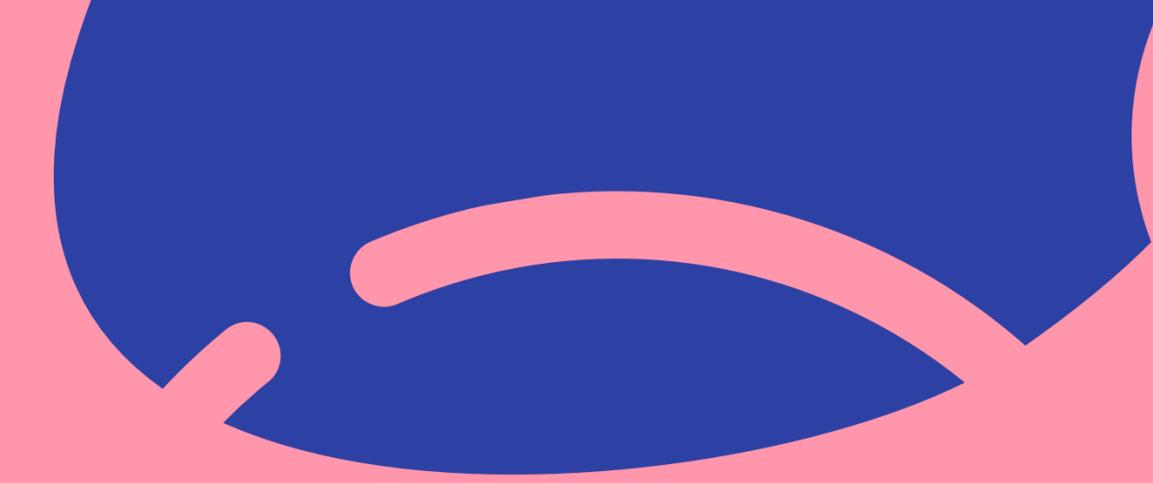
Essa é nossa última etapa de hard skills em análise de dados. Assim como na ciência de dados, ela é uma etapa que depende muito da área de aplicação que escolhemos seguir.

Descubra sua área de aplicação favorita

1

Nessa etapa, não há certo ou errado. Pesquise bastante sobre as áreas disponíveis no mercado e que combinam com você e com seus gostos. Não tenha medo de explorar várias áreas. Hoje, temos diversos eventos de empresas e/ou comunidades que mostram onde os dados podem ser aplicados. Participe sempre que puder e conheça tudo que há para fazer. O céu é o limite!





Procure pessoas que trabalham na área

É muito importante que você procure pessoas que já estão atuando na sua área de interesse. Chame-as no LinkedIn e peça indicação de conteúdos que possam te ajudar nessa jornada. Não tenha medo de pedir ajuda!

2

3

Encontre comunidades que possam te oferecer apoio

Lembre-se: você não está sozinha. Há várias mulheres que estão passando pelo mesmo que você e várias que já passaram por isso. Vá até comunidades que possam te oferecer suporte e sororidade.

4

O aprendizado é contínuo e deve ser compartilhado

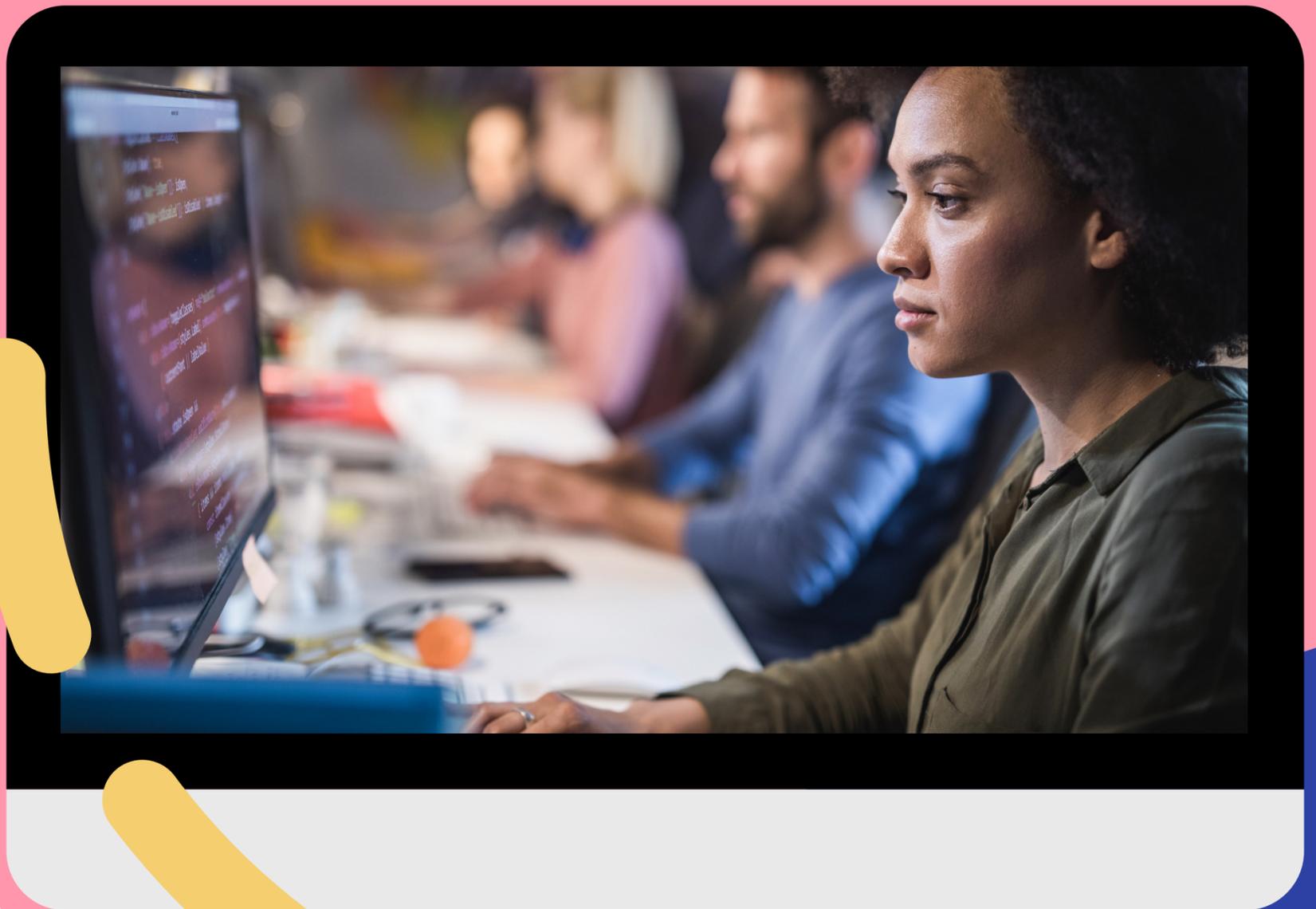
Mesmo depois de estar na área, suas soft skills vão continuar evoluindo. Compartilhe seus aprendizados com mais mulheres - além de ajudá-las, você estará se desenvolvendo também.





ENGENHARIA DE DADOS

A última profissão que vamos conhecer por aqui é a engenharia de dados, área onde você vai poder ajudar a construir uma infraestrutura robusta para que os dados sejam acessíveis. Vamos lá?

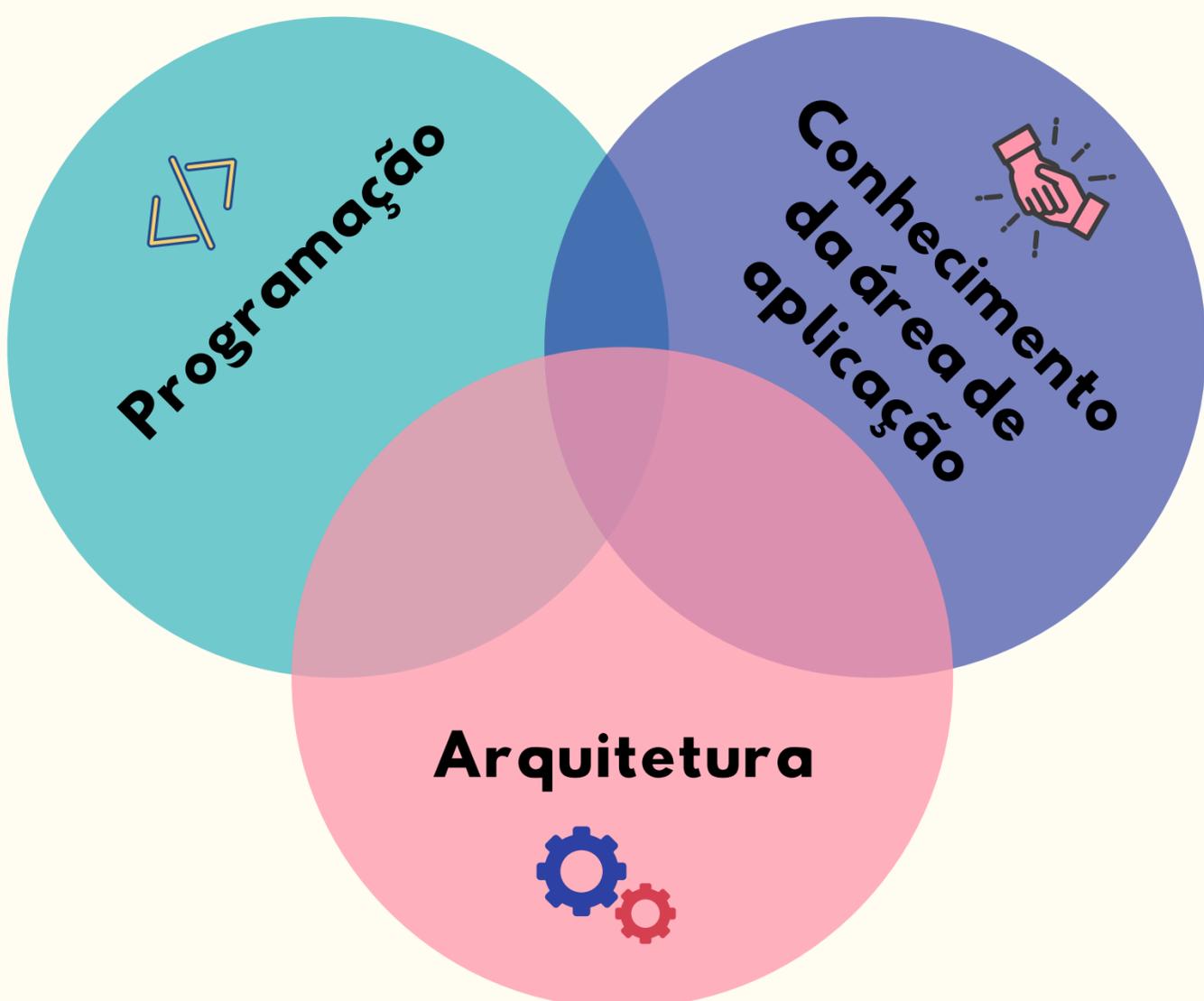




O QUE É ENGENHARIA DE DADOS?

A engenharia de dados é uma área multidisciplinar onde você arquiteta e/ou integra sistemas que coletam, gerenciam e convertem dados brutos em informações.

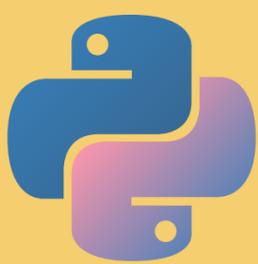
* **Hard skills** **necessárias**



***Hard skills**: conhecimentos técnicos

1. Programação

Vamos passar pela primeira etapa das hard skills: a programação. Quando falamos de engenharia de dados, temos quatro pontos-chave:



Python



Big Data

Python

1. Variáveis
2. Operadores
3. Estrutura básica de dados
4. Loop de repetições
5. Estrutura de condições
6. Métodos de Funções
7. Bibliotecas básicas: numpy, math, SciPy
8. Tratamento de erros

1

2

SQL

1. Introdução a banco de dados
2. Consultas em geral
3. Manipulação de dados
4. Joins
5. Funções



3

NoSQL

1. Manipulação de Dados
 - a) Não Estruturados
 - b) Semi-Estruturados
2. MongoDB

4

Big Data

1. Ecosystema Hadoop
2. Pig e Flume
3. Processamento com Spark

5

ETL

1. Transformação de Dados
2. Criação de Data Warehouse
3. Exportação de Dados

2. Arquitetura

Vamos passar pela segunda etapa das hard skills: a arquitetura.

Cloud Computing

1

1. Introdução a Cloud
2. Modelos de Cloud
3. Nuvem pública
4. Cloud Builders
5. Segurança de dados
6. Banco de dados
7. Virtualização de infraestrutura
8. Soluções Multicloud
9. Orquestração de Serviços e Kubernetes

2

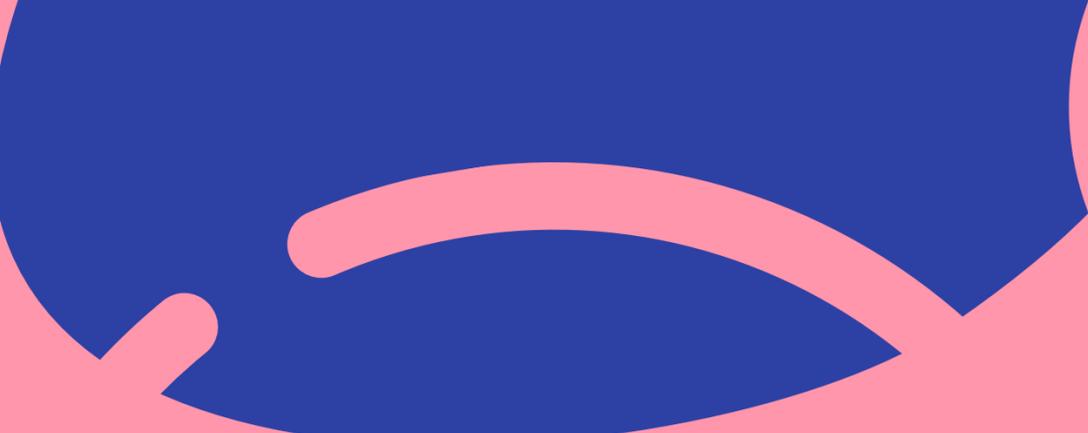
GitHub

1. Arquitetura do GitHub
2. Comandos

3

APIs

1. Desenvolvimento de APIs
2. Modelagem de APIs



3. Conhecimento da área de aplicação

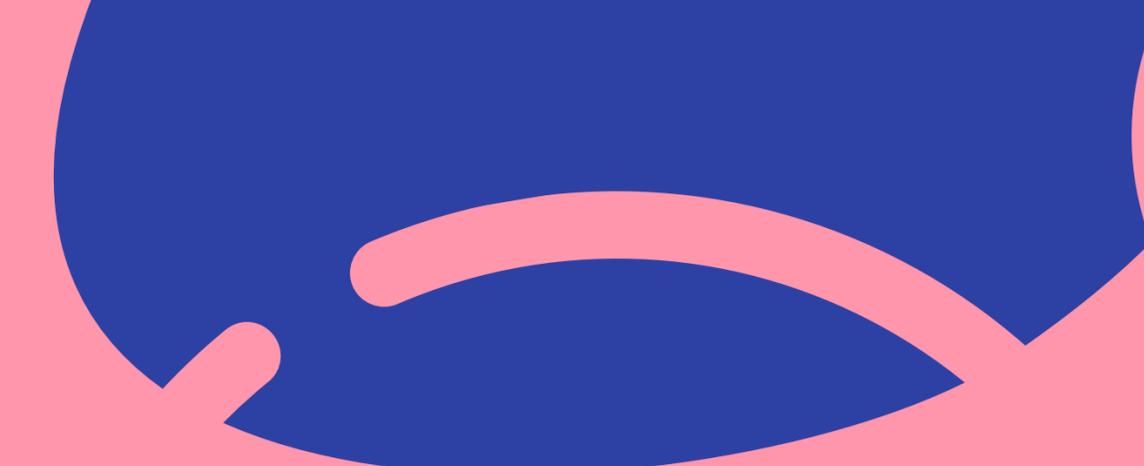
Essa é nossa última etapa de hard skills em engenharia de dados. No entanto, assim como nas outras duas profissões que vimos, ela é uma etapa que depende muito da área de aplicação que escolhemos seguir.

Descubra sua área de aplicação favorita

1

Nessa etapa, não há certo ou errado. Pesquise bastante sobre as áreas disponíveis no mercado e que combinam com você e com seus gostos. Não tenha medo de explorar várias áreas. Hoje, temos diversos eventos de empresas e/ou comunidades que mostram onde os dados podem ser aplicados. Participe sempre que puder e conheça tudo que há para fazer. O céu é o limite!





Procure pessoas que trabalham na área

É muito importante que você procure pessoas que já estão atuando na sua área de interesse. Chame-as no LinkedIn e peça indicação de conteúdos que possam te ajudar nessa jornada. Não tenha medo de pedir ajuda!

2

3

Encontre comunidades que possam te oferecer apoio

Lembre-se: você não está sozinha. Há várias mulheres que estão passando pelo mesmo que você e várias que já passaram por isso. Vá até comunidades que possam te oferecer suporte e sororidade.

4

O aprendizado é contínuo e deve ser compartilhado

Mesmo depois de estar na área, suas soft skills vão continuar evoluindo. Compartilhe seus aprendizados com mais mulheres - além de ajudá-las, você estará se desenvolvendo também.





DICAS FINAIS

Parabéns por ter chegado até aqui! Vamos às últimas dicas para que você continue nessa linda jornada!

E AS SOFT SKILLS?

Não são só as hard skills que nos fazem ter sucesso no mundo de dados. Precisamos também ter atributos não técnicos que nos ajudam a interagir com outras pessoas e com os dados propriamente ditos.



Alguns fatores relevantes

1 ano

Tempo médio de dedicação para ingressar na área

5 mil reais

Salário médio da área

64,9%

É a porcentagem de equipes de tech onde as mulheres representam no máximo 20% do time. Tech precisa de você!



Referências para estudar

- [Playlist completa de Estatística básica da UFPR](#)
- [Estatística e Probabilidade \(Khan Academy\)](#).
- [Estatística \(Univesp\)](#).
- [Curso de Python](#)
- [Python para Machine Learning](#)
- [Matemática para Machine Learning](#)
- [Algoritmos de Machine Learning](#)
- [Tutoriais para Python e SQL \(W3schools\)](#).
- [Cursos da Kaggle](#)
- [PlayList de SQL \(Curso\)](#).
- [Playlist Excel](#)
- [Google Data Studio](#)
- [Tableau](#)
- [Power BI](#)
- [MongoDB](#)
- [Spark](#)
- [Hadoop](#)
- [Big Data](#)
- [APIs](#)
- [GitHub](#)
- [Azure](#)



progr{a{m}aria

