

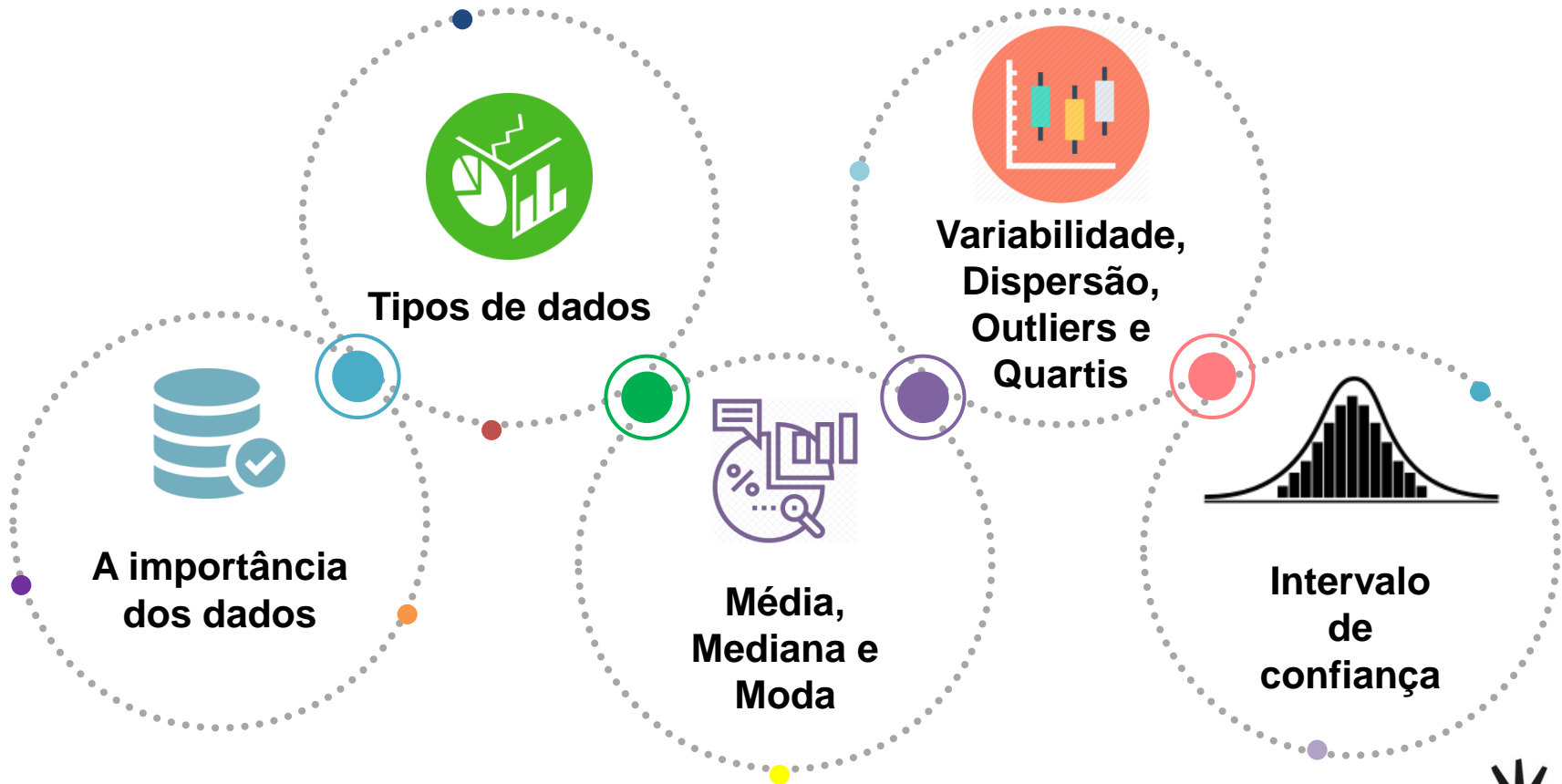
Parte 1



Conceitos Básicos de Estatística

**Grupo de Estudos
Data Science**

Agenda



Tipos de dados

Quantitativos - Quantifica ou mede

Discretos:

Assumem valores em um conjunto especificado de números.

Contínuos:

Assumem valores em um intervalo contínuo de números.

Qualitativos - Característica ou qualidade

Nominal:

Característica que não possui ordem.

Ordinal:

Característica que possui uma ordem de grandeza.

Tipos de dados

Quantitativos



Discretos:

- Quantidade de pessoas na sala?
- Quantidade de dias no mês?

Contínuos:

- Qual sua altura?
- Qual a distância daqui até sua casa?

Qualitativos



Nominal:

- Qual seu Estado? (SP, MG, RJ, Outros)
- Qual gênero você se identifica?

Ordinal:

- Avaliação curso (Ruim a Ótimo).
- Qual seu nível de escolaridade?

Tipos de dados – Exemplo Dados Pyladies

ID	Estado Origem	Idade	Escolaridade	Trabalha como Programadora	Renda Mensal
1	SP	36	4	S	3737,52
2	SP	25	2	N	400,00
3	MG	34	3	S	2366,14
4	RJ	23	3	S	2841,29
5	SP	31	4	N	800
6	SP	34	5	S	3433,02
7	SP	39	5	S	2752,74
8	PE	24	3	S	3682,33
9	RJ	29	3	S	2359,28
10	SP	27	3	S	2119,15
11	SP	30	3	S	3326,79
12	SP	25	4	S	2684,05
13	SP	23	2	S	3507,84
14	SP	16	1	N	0
15	SP	36	4	N	800

Legenda Escolaridade

- | | |
|---|------------------------|
| 1 | Ensino Medio Completo |
| 2 | Graduanda |
| 3 | Graduação Completa |
| 4 | Pós graduanda |
| 5 | Pós graduação completa |

Tipos de dados – Exemplo Dados Pyladies

ID	Estado Origem	Idade	Escolaridade	Trabalha como Programadora	Renda Mensal
1	SP	36	4	S	3737,52
2	SP	25	2	N	400,00
3	MG	34	3	S	2366,14
4	RJ	31	4	N	2841,29
5	SP	31	4	N	800
6	SP	34	5	S	3433,02
7	SP	39	5	S	2752,74
8	PE	24	3	S	3682,33
9	RJ	29	3	S	2359,28
10	SP	27	3	S	2119,15
11	SP	30	3	S	3326,79
12	SP	25	4	S	2684,05
13	SP	23	2	S	3507,84
14	SP	16	1	N	0
15	SP	36	4	N	800

Dados Qualitativos

Legenda Escolaridade

- | | |
|---|------------------------|
| 1 | Ensino Medio Completo |
| 2 | Graduanda |
| 3 | Graduação Completa |
| 4 | Pós graduanda |
| 5 | Pós graduação completa |

Tipos de dados – Exemplo Dados Pyladies

ID	Estado Origem	Idade	Escolaridade	Trabalha como Programadora	Renda Mensal
1	SP	36	4	S	3737,52
2	SP	25	2	N	400,00
3	MG	34	3	S	2366,14
4	Nominal	23	Ordinal	Nominal	2841,29
5	SP	31	4	N	800
6	SP	34	5	S	3433,02
7	SP	39	5	S	2752,74
8	PE	24	3	S	3682,33
9	RJ	29	3	S	2359,28
10	SP	27	3	S	2119,15
11	SP	30	3	S	3326,79
12	SP	25	4	S	2684,05
13	SP	23	2	S	3507,84
14	SP	16	1	N	0
15	SP	36	4	N	800

Legenda Escolaridade

- | | |
|---|------------------------|
| 1 | Ensino Medio Completo |
| 2 | Graduanda |
| 3 | Graduação Completa |
| 4 | Pós graduanda |
| 5 | Pós graduação completa |

Tipos de dados – Exemplo Dados Pyladies

ID	Estado Origem	Idade	Escolaridade	Trabalha como Programadora	Renda Mensal	
1	SP	36	4	S	3737,52	
2	SP	25	2	N	400,00	
3	MG	34	3	S	2366,14	
4	RJ	23	Dados Quantitativos			1,29
5	SP	31				00
6	SP	34	5	S	3433,02	
7	SP	39	5	S	2752,74	
8	PE	24	3	S	3682,33	
9	RJ	29	3	S	2359,28	
10	SP	27	3	S	2119,15	
11	SP	30	3	S	3326,79	
12	SP	25	4	S	2684,05	
13	SP	23	2	S	3507,84	
14	SP	16	1	N	0	
15	SP	36	4	N	800	

Legenda Escolaridade

- | | |
|---|------------------------|
| 1 | Ensino Medio Completo |
| 2 | Graduanda |
| 3 | Graduação Completa |
| 4 | Pós graduanda |
| 5 | Pós graduação completa |

Tipos de dados – Exemplo Dados Pyladies

ID	Estado Origem	Idade	Escolaridade	Trabalha como Programadora	Renda Mensal
1	SP	36	4	S	3737,52
2	SP	25	2	N	400,00
3	MG	34	3	S	2366,14
4	RJ	Discreto	3	S	Contínuo
5	SP	31	4	N	800
6	SP	34	5	S	3433,02
7	SP	39	5	S	2752,74
8	PE	24	3	S	3682,33
9	RJ	29	3	S	2359,28
10	SP	27	3	S	2119,15
11	SP	30	3	S	3326,79
12	SP	25	4	S	2684,05
13	SP	23	2	S	3507,84
14	SP	16	1	N	0
15	SP	36	4	N	800

Legenda Escolaridade

- | | |
|---|------------------------|
| 1 | Ensino Medio Completo |
| 2 | Graduanda |
| 3 | Graduação Completa |
| 4 | Pós graduanda |
| 5 | Pós graduação completa |

Tipos de dados – Exemplo Dados Pyladies

ID	Estado Origem	Idade	Escolaridade	Trabalha como Programadora	Renda Mensal
1	SP	36	4	S	3737,52
2	SP	25	2	N	400,00
3	MG	34	3	S	2366,14
4	RJ	Discreto	3	S	Contínuo
5	SP	31	4	N	800
6	SP	34	5	S	3433,02
7	SP	39	5	S	2752,74
8	PE	24	3	S	3682,33
9	RJ	29	3	S	2359,28
10	SP	27	3	S	2119,15
11	SP	30	3	S	3326,79
12	SP	25	4	S	2684,05
13	SP	23	2	S	3507,84
14	SP	16	1	N	0
15	SP	36	4	N	800

Legenda Escolaridade

- | | |
|---|------------------------|
| 1 | Ensino Medio Completo |
| 2 | Graduanda |
| 3 | Graduação Completa |
| 4 | Pós graduanda |
| 5 | Pós graduação completa |

**O que podemos tirar de
informações desses dados?**



Média

A média nos tras informação sobre o valor médio de determinada informação.

Qual o renda mensal média das meninas que observamos?

Para respondermos essa pergunta precisamos:

1º passo) Somar a renda mensal de todas as meninas;

2º passo) Dividir o valor obtido pelo total de meninas.

Soma Renda Mensal	34810,15
-------------------	----------

Total de Meninas	15
------------------	----

Renda Média Mensal	$34810,15/15 = 2320,68$
--------------------	-------------------------

Ou seja, em média a Renda Mensal é de R\$2320,68.

Média

Mas será que a renda mensal média varia com relação às demais características?

Para respondermos essa pergunta precisamos:

1º passo) Somar a renda mensal de todas as meninas considerando a característica em questão;

2º passo) Dividir o valor obtido pelo total de meninas que possuem aquela característica.

Característica: Trabalhar ou não com programação!

Trabalha como Programadora	Soma Renda Mensal	Quantidade Meninas	Renda Mensal Média
S	32810,15	11	2982,74
N	2000	4	500,00

Média

- ✓ Observamos que o fato de trabalhar ou não com programação difere bastante a Renda Mensal Média das meninas observadas;
- ✓ Ponto de atenção: Dentre as 4 meninas que não trabalham com programação uma delas não tem nenhuma renda mensal!

Proposta:

Observar a Renda Mensal Média por Estado de Origem e Escolaridade! :)

Isso indica que a média geral pode ser um número não tão expressivo, mas se observarmos pontos extremos (outliers) ou trabalharmos com outras informações, podemos entender mais adequadamente os dados.

Mediana

Vimos que a média pode não ser a medida mais adequada em alguns casos, agora vamos avaliar a Mediana que nos traz uma medida de tendência central dos dados!

Qual o Mediana quando observamos a idade das meninas?

Para respondermos essa pergunta precisamos:

1º passo) Ordenar os dados do menor para o maior valor;

2º passo) Selecionar o valor mediano dos dados.

Idade														
16	23	23	24	25	25	27	29	30	31	34	34	36	36	39

Observamos que a Mediana não é influenciada pelo valor baixo de idade.

Mediana

A mediana também é interessante quando trabalhamos com dados qualitativos!

Como entendemos qual a tendência central da Escolaridade, por exemplo?

Para respondermos essa pergunta precisamos:

1º passo) Contar a frequência de cada uma das classes;

2º passo) Calcular a frequência acumulada;

3º passo) Selecionar o valor mediano dos dados.

Escolaridade	Frequência	Frequência Acumulada
1 - Ensino Medio Completo	1	1
2 – Graduanda	2	3
3 - Graduação Completa	6	9
4 - Pós graduanda	4	13
5 - Pós graduação completa	2	15

Moda

A Moda é aquele elemento que mais se repete na distribuição dos dados.

Como entendemos qual a Moda dos Estados de Origem? Ou seja, de qual Estado a maioria das meninas vem?

Para respondermos essa pergunta precisamos:

1º passo) Contar a frequência de cada um dos Estados;

2º passo) Selecionar o estado que tem maior frequência.

Estado Origem	Frequência
SP	11
MG	1
RJ	2
PE	1

Média, Mediana e Moda

Mas como essas medidas se relacionam?

Em amostras bem distribuídas a Média, a Mediana e a Moda são iguais ou possuem valores próximos!

Se observarmos a renda mensal das meninas que trabalham com programação temos que a média e a mediana são muito próximas mesmo.

Média	R\$ 2982,74
Mediana	R\$ 2841,29

Se observarmos a idade das meninas, também obtemos valores de média e mediana próximos!

Média	Aprox. 29 anos
Mediana	29 anos