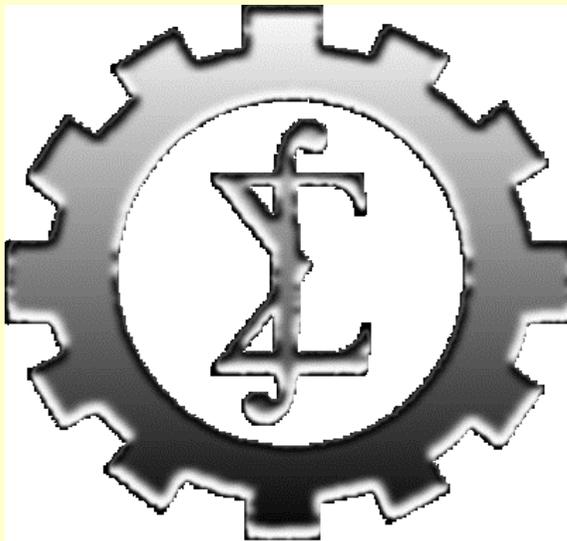


CURSO DE ESTATÍSTICA DESCRITIVA

Adriano Mendonça Souza
Departamento de Estatística
- UFSM -



**“A imaginação é mais importante
que o conhecimento”**

Albert Einstein 

Estatística ?!?



Conjunto de técnicas ou processos, que permitem descrever, analisar e concluir numericamente ou graficamente sobre determinado assunto

Aplicações

- **Indústria (Controle de qualidade)**
- **Economia**
- **Meteorologia**
- **Geologia**
- **Biologia**
- **Ciências Sociais**



Divisões da Estatística

- Geral ou metodológica

- Descritiva
- Indutiva ou inferencial

(amostragem, planejamento de experimentos)

- Estatística aplicada

- Biometria
- Econometria
- Mecânica estatística
- Demografia
- Psicometria
- Sociometria

Conceitos Básicos

População x Amostra

Censo x Amostragem

Dado x Variável

Parâmetros x Estatísticas

| | | | | |
|------------|-------|----------------------|-------|-----------|
| μ | _____ | Média | _____ | \bar{X} |
| σ^2 | _____ | Variância | _____ | S^2 |
| σ | _____ | Desvio-padrão | _____ | S |

Classificação das variáveis

- Qualitativa ou atributos/ Cor, beleza
- Quantitativa discreta/ Número de alunos
- Quantitativa contínua/ Altura, peso

Medições \Rightarrow Var. Contínuas

Contagens \Rightarrow Var. Discretas

Níveis de mensuração das variáveis

- Nominal/Classificação ou símbolos
Moda, frequência
- Ordinal/Relação
Mediana, moda, frequência, estatística por postos
- Intervalar/Distância entre dois números
Estatísticas paramétricas
- Razão/Existe o ponto zero
Todas as estatísticas são aplicáveis

Arredondamento de dados

0, 1, 2, 3, 4 \Rightarrow Conserva

6, 7, 8, 9 ou 5+ \Rightarrow Soma uma unidade

6, 349**2**3 \rightarrow 6,349

6, 342**9**3 \rightarrow 6,343

Par \leftarrow **5** \rightarrow Ímpar
Conserva Soma uma unidade

6, 33**5**003 \rightarrow 6,34

6, 33**5**000 \rightarrow 6,34

6, 32**5**000 \rightarrow 6,32



Fases do método estatístico

São as fases do trabalho estatístico que devem ser desenvolvidas para se chegar aos resultados finais do estudo.

- ✓ Definição do problema;
- ✓ Planejamento da pesquisa;
- ✓ Coleta ou levantamento dos dados;
- ✓ Crítica e digitação dos dados;
- ✓ Organização e representação dos dados;
- ✓ Análise dos dados e interpretação dos resultados;

Representação tabular

Consiste em dispor os dados em linhas e colunas distribuídas de modo ordenado, obedecendo à Resolução n° 886, de 26 de outubro de 1966, do Conselho Nacional de Estatística, editadas pelo Instituto Brasileiro de Geografia e Estatística - IBGE

Representação esquemática

Título: O quê?; Onde?; Quando?

| | |
|-----------|-------|
| Cabeçalho | Total |
|-----------|-------|

← Coluna Indicadora

Total

Fonte

* (*Chamada*)

Nota

} Rodapé

Elementos de uma tabela

Título:

Define o que está sendo estudado

O que? (Assunto a ser representado (Fato));

Onde? (O lugar onde ocorreu o fenômeno (local));

Quando? (A época em que se verificou o fenômeno (tempo)).

- **Cabeçalho:** É a parte da tabela na qual é designada a natureza do conteúdo de cada coluna.

Corpo:

É a parte da tabela composta por linhas e colunas.

- **Linhas:** é a parte do corpo que contém uma sequência horizontal de informações.
- **Colunas:** é a parte do corpo que contém uma sequência vertical de informações.
- **Coluna Indicadora:** é a coluna que contém as discriminações correspondentes aos valores distribuídos pelas colunas numéricas.
- **Casa ou célula:** é a parte da tabela formada pelo cruzamento de uma linha com uma coluna.

Rodapé:

É o espaço aproveitado em seguida ao fecho da tabela, onde são colocados as notas de natureza informativa (fonte, notas e chamadas)

- **Fonte:** refere-se à entidade que organizou ou forneceu os dados expostos.
- **Notas e Chamadas:** são esclarecimentos contidos na tabela (**nota** - conceituação geral; **chamada** - esclarecer minúcias em relação a uma célula).

Séries estatísticas

Uma série estatística é um conjunto de dados ordenados segundo uma característica comum, sendo apresentadas sob forma de tabela e/ou gráfico.

Ela é classificada de acordo com a variação dos elementos que a compõem:

- **a espécie** (o fato ou fenômeno)
- **o local** (o lugar onde o fenômeno acontece)
- **a época** (fator temporal ou cronológico)

Série Temporal ou Cronológica

É a série cujos dados estão dispostos em correspondência com o **tempo**, varia o tempo e permanece constante o fato e o local.

Produção de Petróleo Bruto no Brasil de 1976 a 1980 (x 1000 m³)

| Anos | Produção |
|------|----------|
| 1976 | 9 702 |
| 1977 | 9 332 |
| 1978 | 9 304 |
| 1979 | 9 608 |
| 1980 | 10 562 |

Fonte: Conjuntura Econômica (fev. 1983)

O nome da série depende do(s) elemento(s) que varia(m).

Série Geográfica ou Territorial

É a série cujos dados estão dispostos em correspondência com o **local**, varia o local e permanece constante a época e o fato.

População Urbana do Brasil em 1980 (x 1000)

| Região | População |
|--------------|-----------|
| Norte | 3 037 |
| Nordeste | 17 568 |
| Sudeste | 42 810 |
| Sul | 11 878 |
| Centro-Oeste | 5 115 |
| Total | 80 408 |

Fonte: Anuário Estatístico (1984)

Pode-se ter uma série específica, geográfica, temporal, mista ou uma distribuição de frequências.

Série Específica ou Qualitativa

É a série cujos dados estão dispostos em correspondência com a **espécie** ou qualidade, varia o fato e permanece constante a época e o local.

População Urbana e Rural do Brasil em 1980 (x 1000)

| Localização | População |
|-------------|-----------|
| Urbana | 80 408 |
| Rural | 38 566 |
| Total | 118 974 |

Fonte: Anuário Estatístico (1984)

Quando em uma série estatística permanecem constantes o fato, o local e a época em que determinado fenômeno ocorreu temos uma **Distribuição de Frequência**

Série Mista ou Composta

A combinação entre duas ou mais séries constituem novas séries denominadas compostas e apresentadas em tabelas de dupla entrada.

População Urbana do Brasil por Região de 1940 a 1980 (x 1000)

| Anos | REGIÕES | | | | |
|------|---------|--------|--------|--------|-------|
| | N | NE | SE | S | CO |
| 1940 | 406 | 3 381 | 7 232 | 1 591 | 271 |
| 1950 | 581 | 4 745 | 10 721 | 2 313 | 424 |
| 1960 | 958 | 7 517 | 17 461 | 4 361 | 1 007 |
| 1970 | 1 624 | 11 753 | 28 965 | 7 303 | 2 437 |
| 1980 | 3 037 | 17 567 | 42 810 | 11 878 | 5 115 |

Fonte: Anuário Estatístico (1984)

Local + Época = Série Geográfica Temporal

O nome da série mista surge de acordo com a combinação de pelo menos dois elementos.

Representação gráfica

Gráfico é toda a forma de representação das séries estatísticas que seja baseada no desenho.

os gráficos, devem apresentar um **título** que informe a espécie, o lugar, o tempo do fenômeno representado e a **fonte** de onde foram coletados os dados expostos.

Classificação dos gráficos analíticos

✧ Pontos

✧ Linhas

✧ Superfície

📁 Barras (Simples, Sobrepostos, Justapostos)

📁 Colunas (Simples, sobrepostos, Justapostos)

📁 Setores

Um gráfico deve ser:

atraente

simples

claro

verdadeiro

preciso

Alguns Gráficos

Produção de Petróleo Bruto no Brasil de 1976 a 1980 (x 1000 m³)

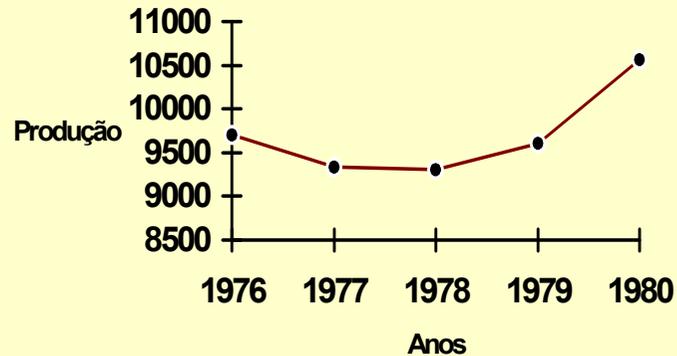


Gráfico de linha

Fonte: Conjuntura Econômica (fev. 1983)

População Urbana do Brasil em 1980 (x 1000)

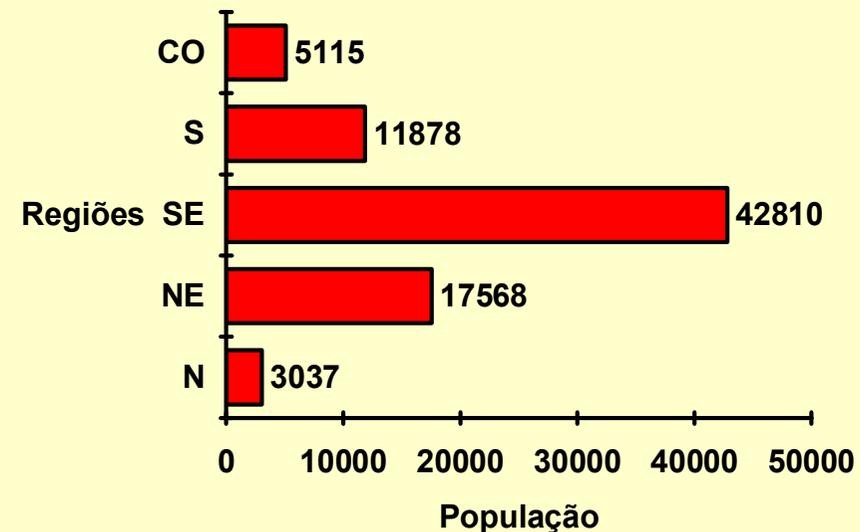
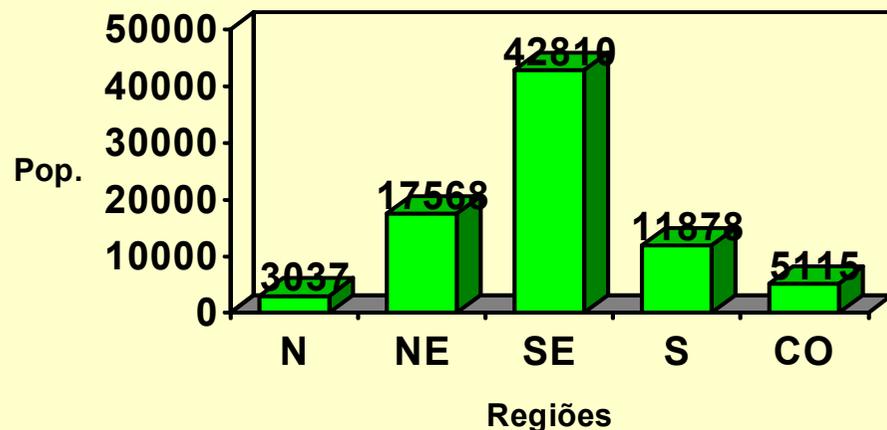


Gráfico de barras

Fonte: Anuário Estatístico (1984)

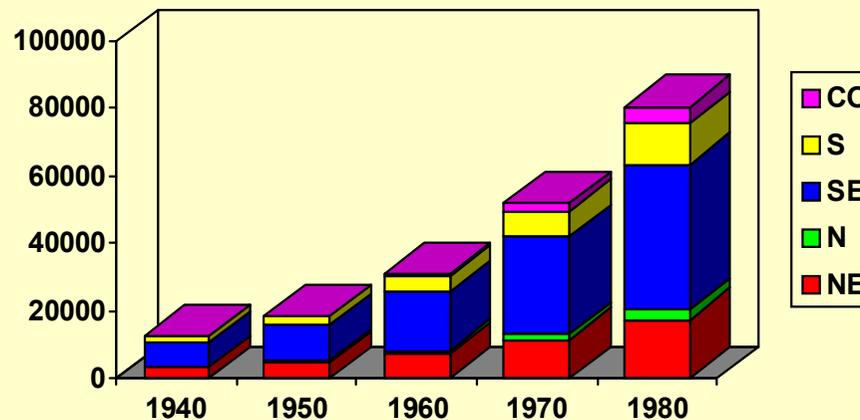
População Urbana do Brasil em 1980 (x 1000)



Fonte: Anuário Estatístico (1984)

Gráfico de colunas

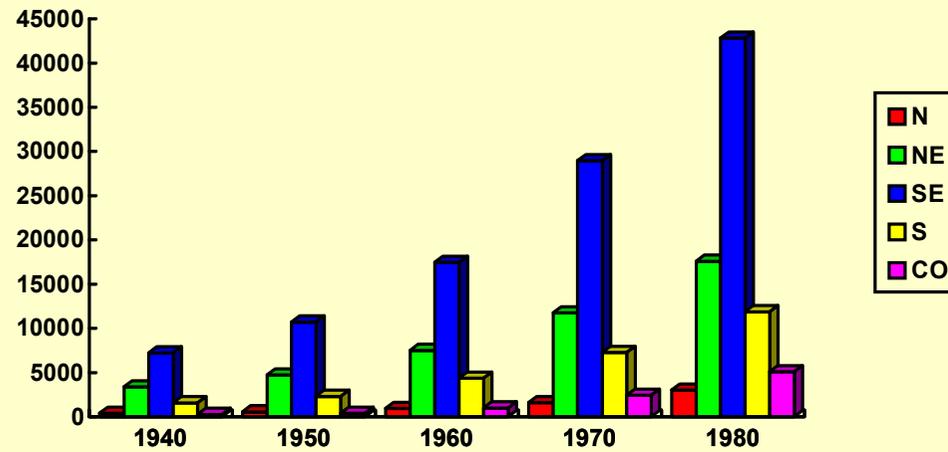
População Urbana do Brasil por Região de 1940 a 1980 (x 1000)



Fonte: Anuário Estatístico (1984)

Gráfico de colunas
sobrepostas

População Urbana do Brasil por Região de 1940 a 1980 (x 1000)

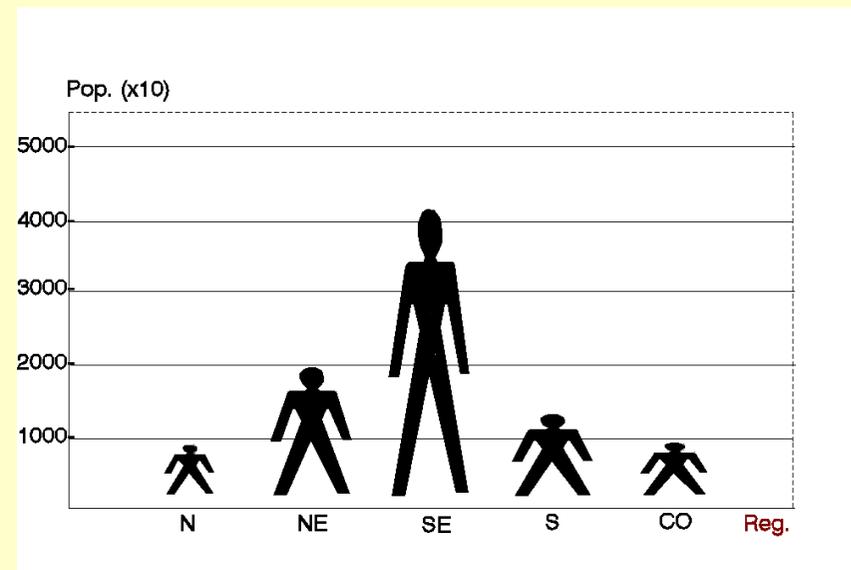


Fonte: Anuário Estatístico (1984)

Colunas justapostas

Gráfico Pictorial

População Urbana do Brasil em 1980 (x 1000)



Fonte: Anuário Estatístico (1984)

Distribuição de Frequência

Uma distribuição de frequência é uma tabela que reúne o conjunto de dados, conforme as frequências ou as repartições de seus valores, podendo ser discreta ou contínua.

É uma série estatística onde permanecem constantes o
fato, o local e a época
em que o fenômeno ocorreu.

Distribuição de Freqüência Intervalar (Contínua)

Altura em centímetros de 160 alunos do
Curso de Administração da UFSM - 1990

| Altura (cm) | | X_i | f_i |
|-------------|---------|-------|-------|
| 150 | --- 158 | 153.5 | 18 |
| 158 | --- 166 | 161.5 | 25 |
| 166 | --- 174 | 169.5 | 20 |
| 174 | --- 182 | 177.5 | 52 |
| 182 | --- 190 | 185.5 | 30 |
| 190 | --- 198 | 193.5 | 15 |
| Σ | | ---- | 160 |

Fonte: Departamento de Estatística (1990)

Distribuição de Frequência Pontual (Discreta)

Notas do Aluno "X" na Disciplina de Estatística segundo
critérios de avaliação do DE da UFSM - 1990

| X_i | f_i |
|----------|-------|
| 6.3 | 2 |
| 8.4 | 3 |
| 5.3 | 2 |
| 9.5 | 3 |
| 6.5 | 5 |
| Σ | 15 |

Fonte: Departamento de Estatística (1990)

Alguns conceitos de uma Distribuição de freqüência

Dados brutos: Valores originais conforme foram coletados;

Rol: Dados brutos, ordenados de forma crescente ou decrescente;

Amplitude total: Diferença entre o maior e o menor valor do rol;

$$H = X_{\text{máx}} - X_{\text{min}}$$

Classe: São intervalos que subdividem a amplitude total, número de classes determinado pela fórmula de Sturges ($4 \leq K \leq 12$);

$$K = 1 + 3,3 \log n$$

Limites de classe: São os limites extremos de cada classe;

L_i → é o menor valor das classe considerada;

L_s → é o maior valor da classe considera;

Amplitude de Classe: É a diferença entre o limite L_i e o L_s da classe e determina a amplitude das classes de uma distribuição de freqüência;

$h = L_s - L_i$ → Quando a DF já existe

$h = H/K$ → Para se construir uma DF;

Ponto médio de classe: É a média aritmética dos limites de classe, é o valor representativo da classe;

$$X_i = (L_i + L_s)/2$$

Frequência absoluta: é o número de observações que aparece em uma classe ou valor individual;

$$f_i = n = \sum_{i=1}^n f_i = f_1 + f_2 + \dots + f_n$$

Frequência Relativa: É o quociente entre a frequência absoluta da i-ésima classe com o somatório das frequências;

$$fr_i = \frac{f_i}{\sum_{i=1}^n f_i} \qquad \sum_{i=1}^n fr_i = 1$$

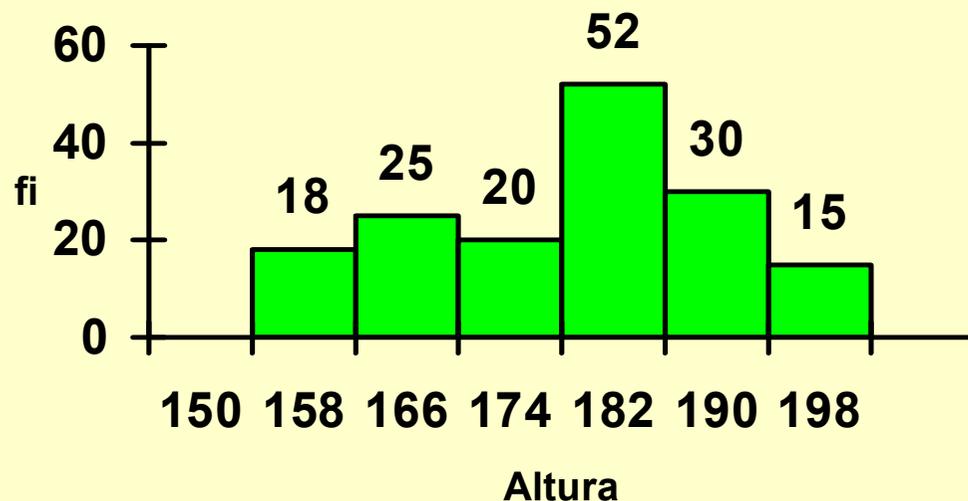
Frequência Acumulada: É o somatório da frequência absoluta da i-ésima classe com a frequência absoluta das classes anteriores;

$$F_n = \sum_{i=1}^n f_i = n$$

Gráficos da distribuição de frequência

Histograma

Altura em centímetros de 160 alunos do Curso de Administração da UFSM - 1990

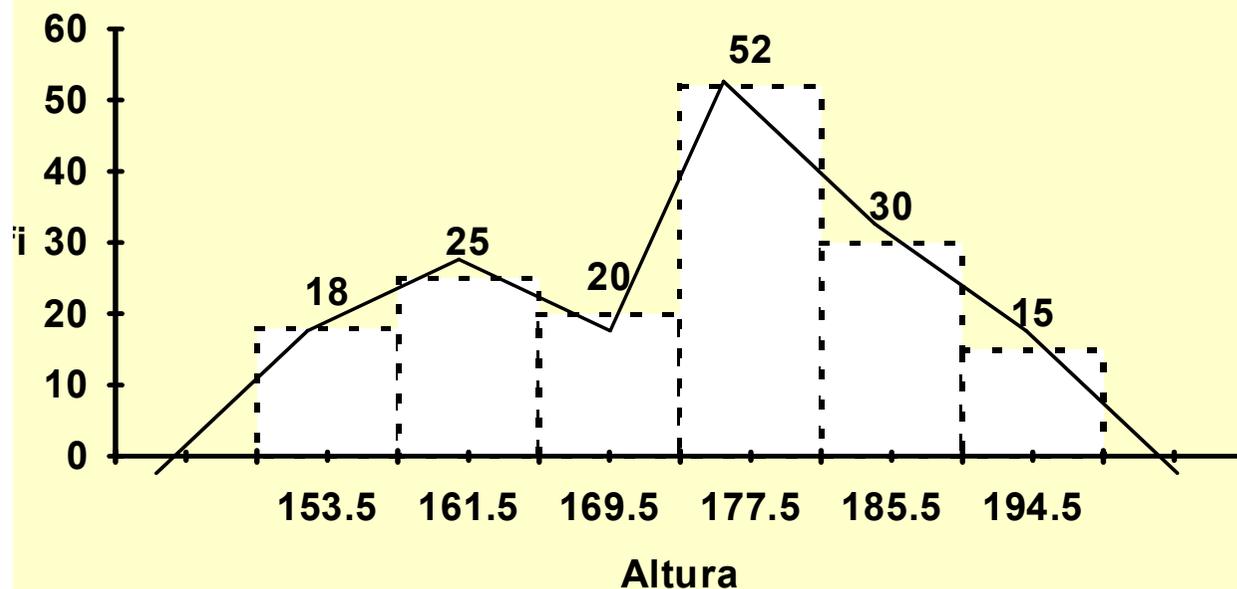


Fonte: Departamento de Estatística (1990)

É um gráfico de colunas justapostas cujas alturas são proporcionais às frequências absolutas e as bases correspondem ao intervalo de classe da distribuição.

Polígono de Frequência

Altura em centímetros de 160 alunos do Curso de Administração da UFSM - 1990

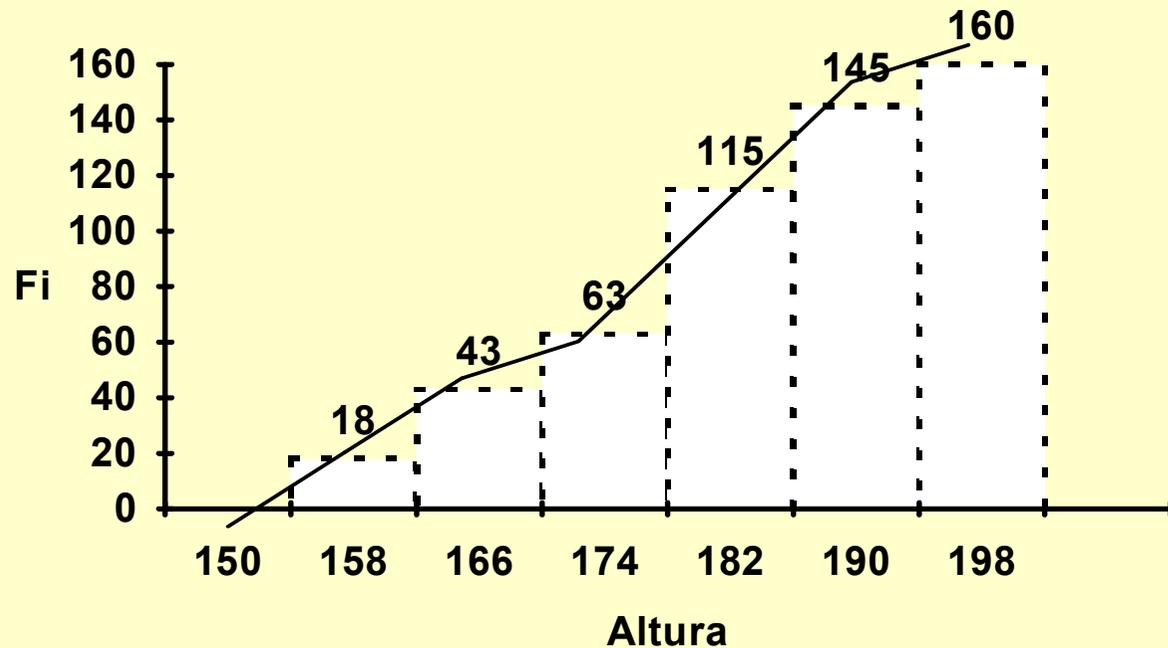


Fonte: Departamento de Estatística (1990)

**É um gráfico de linha
cujos vértices são
proporcionais às
frequências absolutas e
correspondem aos pontos
médios das classes de
distribuição.**

Ogiva

Altura em centímetros de 160 alunos do Curso de Administração da UFSM - 1990



Fonte: Departamento de Estatística (1990)

É um gráfico de linha cujos vértices são proporcionais às frequências acumuladas e correspondem aos limites inferiores das classes da distribuição.

Desafio !!!



Procure em seu local de trabalho dados que possam ser representados através de uma distribuição de frequência e de um gráfico que você julgue útil para o desempenho do seu trabalho.