

CURSO DE ESTATÍSTICA BÁSICA

MÓDULO 1

CAPÍTULO 1 – Conceitos Fundamentais e Compreensão da Estatística

1- Introdução

HOUAISS e VILLAR (2009) definem Estatística como sendo o ramo da Matemática que trata da coleta, da análise, da interpretação e da apresentação de massas de dados numéricos. Estatística significa, também, um conjunto de dados numéricos.

A palavra estatística é derivada do termo latino “status” (Estado). Esse termo provém do primeiro uso da estatística que tinha como função o registro de dados (nº de habitantes da população, nº de casamentos, etc.) e a elaboração de tabelas e gráficos para descrever resumidamente um determinado país em números.

A palavra estatística também designa o ramo da Matemática que analisa dados estatísticos, abrangendo a estatística descritiva e a inferência estatística.

Entretanto, mais importante do que defini-la é poder identificar sua importância e o aumento de sua utilização, pois a Estatística evoluiu, tornando-se uma ampla e complexa ciência, tirando conclusões sobre o conjunto todo a partir de amostras representativas, o que faz com que as decisões sejam baseadas na incerteza.

Uma boa definição de Estatística é a de ser um conjunto de métodos apropriados para a coleta, para a apresentação (organização, resumo e descrição), para a análise e para a interpretação de dados de observação, tendo como objetivo a compreensão de uma realidade específica para a tomada da decisão. Assim sendo, a Estatística se preocupa com:

- a coleta, a organização, a sintetização e a apresentação de dados.
- a medição da variação nos dados e o levantamento de dados.
- a estimativa dos parâmetros da população e a determinação da precisão das estimativas.

-a aplicação dos testes de hipótese em relação aos parâmetros.

-a análise da relação entre duas ou mais variáveis.

A Estatística é bastante utilizada em diversos ramos da atividade humana, no intuito de realizar pesquisas, colher dados e processá-los, analisar informações, apresentar situações através de gráficos de fácil compreensão. Os meios de comunicação, ao utilizarem gráficos, deixam a leitura mais agradável.

Em reportagens de jornais e revistas, pesquisas de opinião, recenseamentos, em ciências tais como a Geografia, a Economia, a Medicina, etc., são utilizados números para descrever e representar fatos observados. Esses números assim empregados são dados estatísticos.

O IBGE (Instituto Brasileiro de Geografia e Estatística) é considerado um órgão importante e conceituado na área.

2- Panorama Histórico

Historicamente, o desenvolvimento da Estatística pode ser compreendido a partir de dois fenômenos: a necessidade de governos coletarem dados censitários e o desenvolvimento da teoria do cálculo das probabilidades.

As primeiras aplicações da Estatística estavam voltadas para as necessidades do Estado na formulação de políticas públicas, fornecendo dados demográficos e econômicos à administração pública.

Há indícios de que 3000 anos a.C. já se faziam censos na Babilônia, China e Egito e até mesmo o 4º livro do Velho Testamento faz referência à instrução dada a Moisés, para que fizesse um levantamento dos homens de Israel que estivessem aptos para guerrear. Usualmente, essas informações eram utilizadas para a taxaço de impostos ou para o alistamento militar. O Imperador César Augusto, por exemplo, ordenou que se fizesse o Censo de todo o Império Romano.

O primeiro levantamento estatístico de que se tem conhecimento se deve a Heródoto e se refere a um estudo da riqueza da população do Egito, cuja finalidade era averiguar quais eram os recursos humanos e econômicos disponíveis para a construção das pirâmides, isso no ano de 3050 a. C. No ano de 2238 a. C., o Imperador Chinês Yao ordenou a realização de uma Estatística com fins industriais e comerciais. No ano de 1400 a. C., o famoso faraó egípcio Ramsés II ordenou um levantamento das terras do Egito. Existem ainda, outros casos de Estatísticas no período antigo da civilização.

A Estatística está ligada a vários ramos da atividade humana. Seu estudo e aperfeiçoamento também foram impulsionados por essas atividades, como é o caso da Biologia e, em particular, da Genética. Estudos de pesquisadores, como os britânicos Karl Pearson (1857 – 1936), professor de eugenia na Universidade de Londres, e Ronald A. Fisher (1890 – 1962), não foram únicos. Além de deixarem contribuições valiosas como geneticistas, desenvolveram, paralelamente, alguns trabalhos em Estatística.

A Estatística, como todas as ciências, tem suas raízes na história do homem. Fisher refere-se à Estatística como o ramo da Matemática Aplicada dedicada à análise de dados de observação. Independentemente das críticas que essa concepção possa merecer, ela evidencia claramente dois aspectos importantes do método estatístico: o tratamento quantitativo a ser aplicado ao fenômeno e a observação, tomada em seu sentido mais amplo.

Desde a Antiguidade, vários povos já registravam o número de habitantes, de nascimentos, de óbitos, faziam estimativas das riquezas individual e social, cobravam impostos, distribuíam equitativamente terras aos povos, realizavam inquéritos quantitativos por processos, etc. que, atualmente, denominaríamos de “estatísticas”.

Na Idade Média colhiam-se informações, comumente com finalidades tributárias ou bélicas.

A partir do século XVI começaram a aparecer as primeiras análises sistemáticas de fatos sociais, como batizados, casamentos, funerais, originando as primeiras tábuas e tabelas e os primeiros números relativos.

Os fundamentos matemáticos da Estatística iniciaram no século XVII com o desenvolvimento da teoria das probabilidades por Pascal e Fermat, surgida com o estudo dos jogos de azar. O uso de computadores modernos tem permitido a computação de dados estatísticos em larga escala e também tornaram possível novos métodos antes impraticáveis. Segundo alguns autores, o marco inicial da Estatística é atribuído à publicação das “Observações sobre os Sentos de Mortalidade” (1662), de John Graunt.

No século XVIII o estudo de tais fatos foi adquirindo, aos poucos, aspecto verdadeiramente científico. Godofredo Achenwall batizou a nova ciência (ou método) com o nome de Estatística, determinando o seu objetivo e suas relações com as ciências.

A abrangência da Estatística aumentou no início do século XIX para incluir a acumulação e análise de dados de maneira geral. Hoje, a Estatística é largamente aplicada nas ciências naturais, e sociais, inclusive na administração pública e privada

As tabelas tornaram-se mais complexas, surgiram as representações gráficas e o cálculo das probabilidades, e a Estatística deixou de ser simples catalogação de dados numéricos coletivos para se tornar o estudo de como chegar a conclusões sobre o todo (população), partindo da observação de partes desse todo (amostras).

Atualmente, informações numéricas são necessárias para cidadãos e organizações de qualquer natureza, e de qualquer parte do mundo globalizado.

Podem-se sintetizar as preocupações com a Estatística em quatro fases.

Primeira Fase	Pepino, no ano de 758, e Carlos Magno, em 762, realizaram estatísticas sobre as terras que eram de propriedade da Igreja. Essas foram as únicas estatísticas importantes desde a queda do Império Romano.
Segunda Fase	Na Inglaterra, no século XVII, já se analisavam grupos de observações numéricas referentes à saúde pública, nascimentos, mortes e comércio. Destacam-se, nesse período, John Graunt (1620-1674) e William Petty (1623-1687) que procuraram leis quantitativas para traduzir fenômenos sociais e políticos.
Terceira Fase	Também no século XVII, inicia-se o desenvolvimento do Cálculo das Probabilidades que, juntamente com os conhecimentos estatísticos, redimensionou a Estatística. Nessa fase, destacam-se: Fermat (1601-1665), Pascal (1623-1662) e Huygens (1629-1695)
Quarta Fase	No século XIX, inicia-se a última fase do desenvolvimento da Estatística, alargando e interligando os conhecimentos adquiridos nas três fases anteriores. Nesta fase, a Estatística não se limita apenas ao estudo da Demografia e da Economia, como antes; agora, o seu campo de aplicação se estende à análise de dados em Biologia, Medicina, Física, Psicologia, Indústria, Comércio, Meteorologia, Educação etc., e ainda, a domínios aparentemente desligados, como Estrutura de Linguagem e estudo de Formas Literárias. Destacam-se, no período, Ronald Fisher (1890-1962) e Karl Pearson (1857-1936).

3- Método Estatístico

3.1- O método científico

Muitos conhecimentos foram obtidos na Antiguidade por acaso e, outros, por necessidades práticas, sem aplicação de um método.

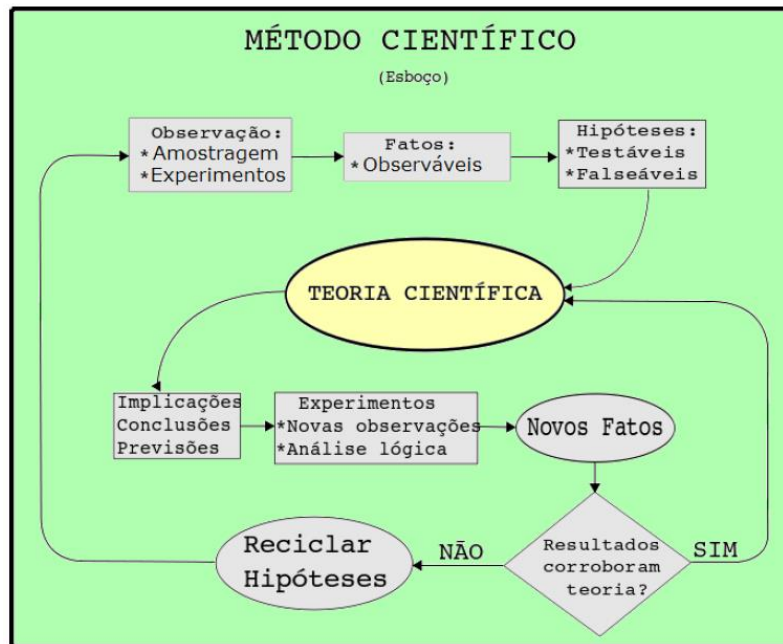
Hoje em dia, quase todo acréscimo de conhecimento resulta da observação e do estudo, embora muito desse conhecimento possa ter sido observado inicialmente por acaso. A verdade é que são desenvolvidos processos científicos para seu estudo e para se adquirir tais conhecimentos. Pode-se afirmar que:

Método é um conjunto de meios dispostos convenientemente para se chegar a um fim que se deseja. Uma definição simples: método é a maneira como se faz alguma coisa.

A metodologia científica tem sua origem no pensamento de Descartes, que foi posteriormente desenvolvido empiricamente pelo físico inglês Isaac Newton. René Descartes propôs chegar à verdade através da dúvida sistemática e da decomposição do problema em pequenas partes, características que definiram a base da pesquisa científica.

O método científico refere-se a um aglomerado de regras de como se deve agir a fim de produzir conhecimento dito científico, seja este um novo conhecimento, ou seja, produto de uma totalidade, evolução ou um aumento da área de incidência de conhecimentos previamente existentes.

Na maioria das disciplinas científicas consiste em juntar evidências empíricas verificáveis, fundamentadas na observação sistemática e controladas, na maioria das vezes, resultantes de experiências ou pesquisa de campo, e analisá-las com o uso da lógica. Para muitos autores o método científico nada mais é do que a lógica aplicada à ciência.



Fonte: Wikipédia, acesso em 05/12/2017.

Dos métodos científicos, convém destacar o método experimental e o estatístico.

3.2- O método experimental

Para que um pesquisador realize e avalie uma pesquisa corretamente, um bom conhecimento de Estatística é essencial, sobretudo para compreender as potencialidades e as limitações das técnicas utilizadas.

O método experimental consiste em manter constantes todas as causas (fatores), menos uma, e variar esta causa de modo que o pesquisador possa descobrir seus efeitos, caso existam. Obs.: É o método preferido no estudo da Física, da Química, etc.

Por exemplo, para fazer café pode-se usar ½ litro de água, 3 colheres de café, um coador, 4 colheres de açúcar. Repetindo esta receita diversas vezes, provavelmente, em todas as vezes acontecerá o mesmo resultado. Contudo, se for alterado algum dos fatores, como aumentar a quantidade de água, o café ficará mais aguado; se aumentar o açúcar, ficará mais doce e assim por diante.

3.3- O método estatístico

Diversas vezes tem-se a necessidade de descobrir fatos em um campo em que o método experimental não se aplica (como nas ciências sociais), já que os vários fatores que afetam o fenômeno em estudo não podem permanecer constantes enquanto se variar a causa que, naquele momento, interessa à pesquisa.

Exemplificando pode-se citar a determinação das causas que definem o preço de uma mercadoria. Para se aplicar o método experimental, precisa-se fazer variar a quantidade da mercadoria e verificar se tal fato iria influenciar no preço da mesma.

No entanto, seria necessário que não houvesse alteração nos outros fatores. Dessa forma, deveria existir, no momento da pesquisa, uma uniformidade dos salários, o gosto dos consumidores deveria permanecer constante, seria necessária a fixação do nível geral dos preços das outras necessidades, etc. Mas isso tudo é impossível. Nesses casos, lança-se mão de outro método, embora mais difícil e menos preciso, denominado método estatístico.

O método estatístico, diante da impossibilidade de manter as causas constantes, admite todas essas causas presentes variando-as, registrando essas variações e procurando determinar, no resultado final, que influências cabem a cada uma delas.

O método estatístico é o mais usado pela Estatística, nele é impossível manter as causas ou fatores constantes, assim sendo deve-se admitir os valores dessas variáveis e analisá-los, procurando determinar qual é a influência que cada fator apresenta no resultado final.

Exemplo: Uma empresa teve uma queda nas vendas no mês de julho. Os estudos indicam que nesse mês aconteceu recesso escolar, aumentou o fluxo de turistas na região, no entanto fez mais frio, o concorrente baixou o preço dele, e o produto dessa empresa perdeu qualidade. Qual desses fatores poderia ter feito as vendas da empresa despencarem?

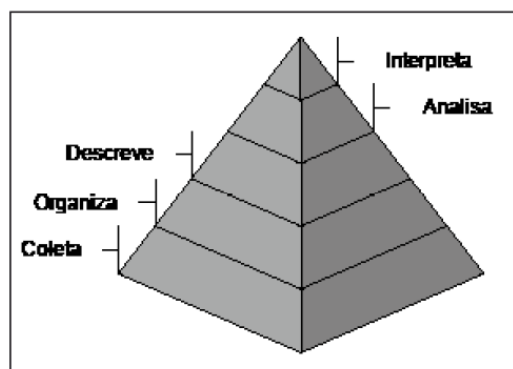
4- Estatística

A Estatística é uma parte da Matemática Aplicada que fornece métodos para a coleta, organização, descrição, análise e interpretação de dados e para o emprego dos mesmos na tomada de decisões.

Em outras palavras, a Estatística é um conjunto de métodos e processos quantitativos que serve para estudar e medir os fenômenos coletivos.

Estatística é a ciência que se preocupa com a coleta, a organização, descrição (apresentação), análise e interpretação de dados experimentais e tem como objetivo fundamental o estudo de uma população. Esse estudo pode ser realizado de duas maneiras: investigando todos os elementos da população ou por amostragem, ou seja, selecionando alguns elementos da população.

De um lado, a Estatística, basicamente, coleta, organiza e descreve os dados e, de outro, analisa e interpreta esses dados.



Estatística: Pirâmide da Definição.

Fonte: MEDEIROS, Carlos Augusto de. **Estatística aplicada à educação**. Brasília: UNB, 2007.

A coleta, a organização e a descrição dos dados estão sob a responsabilidade da Estatística Descritiva, enquanto a análise e a interpretação desses dados ficam por conta da Estatística Indutiva ou Inferencial.

Geralmente, as pessoas se referem ao termo “estatística” no sentido da organização e descrição dos dados (estatística do Ministério da Educação, estatística do Ministério da Saúde, estatística dos acidentes de trânsito, etc.), desconhecendo que o aspecto essencial da Estatística é o de proporcionar métodos inferenciais, que permitam conclusões que transcendam os dados obtidos inicialmente.

Dessa forma, a análise e a interpretação dos dados estatísticos tornam possível o diagnóstico de uma empresa (por exemplo, de uma escola), o conhecimento de seus problemas (condições de funcionamento, produtividade), a formulação de soluções apropriadas e um planejamento objetivo de ação.

Através da análise e interpretação dos dados estatísticos é possível ter o conhecimento de uma realidade, de seus problemas, bem como, a formulação de soluções apropriadas por meio de um planejamento objetivo da ação, excluindo os “achismos” e “casuísmos” comuns.

4.1- Divisões da Estatística

- **Estatística Descritiva** – é aquela que se preocupa com a coleta, organização, classificação, apresentação, interpretação e análise de dados referentes ao fenômeno por

meio de gráficos e tabelas, além de calcular medidas que permitam descrever o fenômeno.

- **Estatística Indutiva (amostral ou inferencial)** – é aquela que, partindo de uma amostra, estabelece hipóteses, tira conclusões sobre a população de origem e formula previsões fundamentando-se na teoria das probabilidades. A estatística indutiva cuida da análise e interpretação dos dados.

O processo de generalização do método indutivo está associado a uma margem de incerteza. Isso se deve ao fato de que a conclusão que se pretende obter para o conjunto de todos os indivíduos analisados quanto a determinadas características comuns baseia-se em uma parcela do total de observações.

5- Fases do Método Estatístico

Quando se pretende empreender um estudo estatístico completo, existem diversas fases do trabalho que devem ser desenvolvidas para se chega aos resultados finais do estudo. Essas etapas ou operações são chamadas fases do trabalho estatístico e são de âmbito da Estatística Descritiva.

As fases do método estatístico são:

- ❖ Definição do problema
- ❖ Planejamento
- ❖ Coleta dos dados
- ❖ Crítica dos dados
- ❖ Apuração dos dados
- ❖ Exposição ou Apresentação dos dados
- ❖ Análise e Interpretação dos resultados

1.

A seguir uma breve descrição sobre cada fase.

Definição do problema – A primeira fase do trabalho estatístico consiste em uma definição ou formulação correta do problema a ser estudado. Além de considerar detidamente o problema objeto do estudo, devem-se examinar outros levantamentos realizados no mesmo campo e análogos, uma vez que parte da informação de que se

necessita pode, muitas vezes, ser encontrada dessa forma. Como o próprio nome já diz essa é a fase em que se define o porquê de usar o método estatístico. Qual será a sua pesquisa? Saber exatamente aquilo que se pretende pesquisar é o mesmo que definir corretamente o problema.

Exemplo: Um fabricante de sabonete, que deseja lançar um produto novo no mercado, poderia estar interessado em um estudo sobre as características dos consumidores atuais. Não havendo estudos semelhantes, ele deverá formular o problema com base em sua própria experiência. Uma lista de fatores relevantes deverá resultar dessa investigação preliminar: número de unidades consumidas por família em cada ano, número médio de pessoas que compõe cada família, número de membros adultos da família, as marcas preferidas e assim por diante. Saber exatamente aquilo que se pretende pesquisar é o mesmo que definir corretamente o problema.

Planejamento – O passo seguinte, após a definição do problema, compreende a fase do planejamento, que consiste em se determinar o procedimento necessário para resolver o problema e, em especial, como levantar informações sobre o assunto objeto do estudo. Elaborar como fazer o levantamento dos dados. Como levantar dados? Que dados deverão ser obtidos? Como se deve obtê-los? Quem pode patrocinar minha pesquisa? É preciso planejar o trabalho a ser realizado, tendo em vista o objetivo que se pretende atingir. Mais especificamente, na fase do planejamento a preocupação maior reside na escolha das perguntas, bem como sua correta formulação, qualquer que seja a modalidade de coleta de dados. É nesta fase que será escolhido o tipo de levantamento a ser utilizado. Sob esse aspecto, pode haver dois tipos de levantamento.

- Levantamento censitário, quando a contagem for completa, abrangendo todo o Universo.
- Levantamento por amostragem, quando a contagem for parcial.

Outros elementos importantes que devem ser tratados nesta fase são o cronograma das atividades, através do qual são fixados os prazos para as várias fases, os custos envolvidos, o exame das informações disponíveis, o delineamento da amostra, a forma como serão escolhidos os dados e assim por diante.

Coleta de dados – O terceiro passo é essencialmente operacional, compreendendo a coleta das informações propriamente ditas. Formalmente, a coleta de dados se refere à obtenção, reunião e registro sistemático de dados, com um objetivo determinado.

Exemplo: Uma empresa pode valer-se de diversas fontes ao utilizar os dados em seus trabalhos estatísticos. Assim, os seus registros contábeis podem conter muitas informações úteis para outros fins, que não meramente o da avaliação do ativo, do passivo e do patrimônio, e a determinação dos lucros e perdas. Os vários departamentos de uma organização, no curso de sua atividade normal, mantêm igualmente registros de natureza estatística. Pode ocorrer, entretanto, que os registros da própria empresa não proporcionem toda a informação necessária para resolver determinado problema. Por exemplo, para saber se uma queda nas vendas está sendo experimentada também por outras empresas do mesmo setor industrial em que opera a empresa não irá recorrer simplesmente a seus registros internos. Poderá haver algum organismo especializado que reúna os dados e os distribua em forma de publicações ou de outra maneira. Assim, os registros internos não serão suficientes nesta situação, necessitando a empresa recorrer a fontes externas para obter as informações desejadas. Se a empresa precisasse saber a opinião dos consumidores sobre algum aspecto particular do seu produto, um exame direto sobre ele seria recomendável.

É possível distinguir dois tipos de fontes externas, as quais dão origem a duas espécies de dados: dados primários e dados secundários.

- Dados primários – Os dados são primários quando são publicados ou comunicados pela própria pessoa ou organização que os tenha recolhido.
- Dados secundários – Os dados são secundários quando são publicados ou comunicados por outra organização.

Um conjunto de dados é primário ou secundário em relação a alguém. As tabelas do Censo Demográfico são fontes primárias. Quando determinado jornal publica estatísticas extraídas de várias fontes e relacionadas com diversos setores industriais, os dados são secundários para quem desejar utilizar-se deles em alguma pesquisa que esteja desenvolvendo. Embora muitas vezes possa ser conveniente recorrer a fontes secundárias, é mais seguro trabalhar com fontes primárias, por várias razões.

- Uma fonte primária oferece, em geral, informação mais detalhada do que uma fonte secundária.
- É mais provável que as definições de termos e de unidades figurem somente nas fontes primárias.
- O uso da fonte secundária traz o risco adicional de erros de transcrição.

- Uma fonte primária poderá vir acompanhada de cópias dos impressos utilizados para coletar as informações, juntamente com o procedimento adotado na pesquisa, a metodologia seguida e o tipo e o tamanho da amostra.

Essas informações proporcionam ao usuário uma ideia do grau de garantia que os dados oferecem. Uma coleta de dados pode ser realizada de duas maneiras, direta ou indiretamente.

Coleta Direta – A coleta é direta quando é obtida diretamente da fonte, como no caso da empresa que realiza uma pesquisa para saber a preferência dos consumidores pela sua marca. Há três tipos de coleta direta.

- Coleta Contínua – a coleta de dados é contínua quando os dados são obtidos ininterruptamente, automaticamente e na vigência de um determinado período: um ano, por exemplo. Os registros de nascimento, de casamento, de óbito podem ser considerados como uma forma de coleta direta contínua.
- Coleta Periódica – a coleta de dados é periódica quando é realizada em períodos curtos, determinados, de tempos em tempos. O recenseamento demográfico, a cada dez anos, e o censo industrial, anualmente, são exemplos de coleta periódica.
- Coleta Ocasional – a coleta de dados é ocasional quando os dados forem colhidos esporadicamente, atendendo a uma conjuntura qualquer ou a uma emergência. A coleta de dados em um surto epidêmico e o registro de pedidos de um determinado artigo que uma grande empresa recebe em um dia de greve são formas de coleta ocasional.

Coleta Indireta – A coleta dos dados é indireta quando inferida a partir dos elementos conseguidos pela coleta direta, ou através do conhecimento de outros fenômenos que, de algum modo, estejam relacionados com o fenômeno em questão. É feita, portanto, por deduções e conjeturas, podendo ser realizada

- por analogia – quando o conhecimento de um fenômeno é induzido a partir de outro que com ele guarda relações de casualidade.
- por proporcionalização – quando o conhecimento de um fato se induz das condições quantitativas de uma parte dele.
- por indícios – quando são escolhidos fenômenos sintomáticos para discutir um aspecto geral da vida social.

- por avaliação – quando, através de informações fidedignas ou estimativas cadastrais, se presume o estado quantitativo de um fenômeno.

Complementando:

A coleta de dados que podem ser de dois tipos:

1- Dados Primários: os dados são obtidos diretamente na fonte originária (coleta direta).

Exemplo: Preferência dos consumidores por um determinado produto.

Métodos de coleta de dados primários: é importante garantir que a coleta de dados primários seja executada de maneira estatisticamente correta, senão os resultados podem ser tendenciosos.

- Observação: O pesquisador não pergunta, observa. Por exemplo: pesquisa de observação para diagnosticar as necessidades de trânsito de uma cidade.
- Levantamento: É o método mais comum de se coletar dados. O instrumento pode ser um questionário estruturado ou um roteiro de itens em que o entrevistado disserta à vontade sobre cada item da pesquisa.

Principais formas de levantamento, resumindo as vantagens e desvantagens:

- Entrevista pessoal: mais flexível e muito caro.
- Telefone: mais barato, penetra em segmentos difíceis, mas é de fácil recusa.
- Questionários (postal, fax ou e-mail): mais lento, média de retorno das respostas muito baixas, mas sem interferência do pesquisador.

2- Dados Secundários: os dados são obtidos de algo já disposto. Provém da coleta direta. Exemplo: Pesquisa sobre a mortalidade infantil, que é feita através de dados colhidos por outras pesquisas.

Obs.: É mais seguro trabalhar com fontes primárias. O uso da fonte secundária traz o grande risco de erros de transcrição.

Crítica dos dados – Após os dados serem obtidos eles são analisados para conferir se há erros, para que não ocorram distorções que interfiram no resultado final. Obtidos os dados, eles devem ser cuidadosamente criticados, à procura de possíveis falhas e imperfeições, a fim de não incorrerem erros grosseiros ou de certo vulto, que possam influir sensivelmente nos resultados. A crítica é externa quando visa às causas dos erros por parte do informante, por distração ou má interpretação das perguntas que lhe foram feitas; é interna quando visa observar os elementos originais dos dados da coleta.

Apuração dos Dados – Antes de começar a analisar os dados, é conveniente que lhes seja dado algum tratamento prévio, a fim de torná-los mais expressivos. Esta etapa, apuração ou sumarização, consiste em resumir os dados, através de sua contagem e agrupamento. É um trabalho de condensação e de tabulação dos dados, que chegam ao analista de forma desorganizada, tornando impossível a tarefa de apreender todo o significado pela simples leitura. Há várias formas de se fazer a apuração, dependendo das necessidades e dos recursos disponíveis do interessado: manual, mecânica, eletromecânica e eletrônica. A apuração é manual quando não se recorre a qualquer máquina para ser realizada. Quando a apuração for feita com o auxílio de máquinas mecânicas, como as de somar e de calcular, ela será mecânica. A apuração eletromecânica é realizada com máquinas diferentes das anteriores pelo fato de suas engrenagens internas serem movidas a energia elétrica. As máquinas eletrônicas efetuam as operações através de impulsos elétricos dispensando qualquer tipo de engrenagem, o que lhes confere uma velocidade infinitamente maior que a das anteriores. Através da apuração tem-se a oportunidade de condensar os dados, de modo a obter um conjunto compacto de números, o qual possibilita distinguir melhor o comportamento do fenômeno na sua totalidade. A apuração nada mais é que a soma e o processamento dos dados obtidos. Contudo, a contrapartida da melhor apreciação dos dados em seu conjunto é a perda correspondente de detalhes, uma vez que se trata de um processo de sintetização.

Apresentação dos Dados – Há duas formas de apresentação dos dados, que não se excluem mutuamente.

- **Apresentação Tabular** – É uma apresentação numérica dos dados. Consiste em dispor os dados em linhas e colunas distribuídas de modo ordenado, seguindo algumas regras práticas adotadas pelos diversos sistemas estatísticos. De maneira mais formal, define-se como tabela a disposição escrita que se obtém, fazendo-se referir uma coleção de dados numéricos a uma determinada ordem de classificação.
- **Apresentação Gráfica** – Constitui uma apresentação geométrica. Embora a apresentação tabular seja de extrema importância, no sentido de facilitar a análise numérica dos dados, não permite ao analista obter uma visão tão rápida, fácil e clara do fenômeno e sua variação como a conseguida através de um gráfico.

Análise e Interpretação dos Resultados – A última fase do trabalho estatístico é a mais importante e também a mais delicada. Nesta etapa, o interesse maior reside em tirar conclusões que auxiliem o pesquisador a resolver seu problema. A análise dos dados estatísticos está ligada essencialmente ao cálculo de medidas, cuja finalidade principal é descrever o fenômeno. Assim, o conjunto de dados a ser analisado pode ser expresso por números-resumos, as estatísticas, que evidenciam características particulares desse conjunto. O significado exato de cada um dos valores obtidos através do cálculo das várias medidas estatísticas disponíveis deve ser bem interpretado. É possível mesmo, nesta fase, arriscar algumas generalizações, as quais envolverão, naturalmente, algum grau de incerteza, porque não se pode estar seguro de que o que foi constatado para aquele conjunto de dados (a amostra) se verificará igualmente para a população. O processo de generalização constitui um campo mais avançado da Estatística Indutiva ou Inferência Estatística. Está ligada essencialmente ao cálculo de medidas e coeficientes, cuja finalidade principal é descrever o fenômeno (estatística descritiva). Na estatística indutiva a interpretação dos dados se fundamenta na teoria da probabilidade.

Obs.:

Estatística Descritiva se preocupa com a coleta, organização, classificação, apresentação, interpretação e análise de dados experimentais.

Estatística Indutiva se preocupa com as hipóteses e conclusões sobre a população.

6- Aplicações da Estatística à Administração

- Uma empresa que está se preparando para lançar um novo produto precisa conhecer as preferências dos consumidores no mercado de interesse. Para isso, pode fazer uma pesquisa de mercado entrevistando um número de residências escolhidas aleatoriamente. Poderá, então, usar os resultados para estimar as preferências de toda a população.
- As técnicas estatísticas são necessárias para separar efeitos de fatores diferentes. É possível que, em uma comunidade, o consumo de sorvete dependa do preço do produto, da renda média local, do número de crianças na comunidade e da temperatura média. Se dispuser de observações de todos os diferentes fatores em

jogo, o analista poderá aplicar a análise da regressão para determinar quais fatores têm os efeitos mais importantes.

- Um auditor deve verificar os livros de uma empresa para se certificar de que os lançamentos refletem efetivamente a situação financeira da companhia. O auditor deve examinar pilhas de documentos originais, como notas de venda, ordens de compra e requisições. Seria um trabalho incalculável consultar todos os documentos originais; em vez disso, o auditor pode verificar uma amostra de documentos escolhidos aleatoriamente e, com base nessa amostra, fazer inferências sobre toda a população.
- Antes de lançar um novo remédio no mercado, é necessário fazer várias experiências para garantir que o produto é seguro e eficaz. O melhor modo de testar um remédio consiste em tomar dois grupos tão semelhantes quanto possível e dar o remédio a um grupo, mas não ao outro, e verificar então se os resultados nos dois grupos são diferentes. O grupo ao qual o remédio foi dado é chamado grupo experimental e o outro é o grupo de controle. Torna-se necessária a análise estatística para determinar se todas as diferenças observadas realmente foram causadas pelo remédio ou poderiam ter sido causadas por outros fatores.
- Se estivermos recebendo um grande embarque de mercadorias de um fornecedor, teremos de certificar-nos de que o produto realmente satisfaz os requisitos de qualidade acordados. Seria muito dispendioso fazer uma verificação de cada item; mas, aqui, mais uma vez, as técnicas estatísticas vêm em nosso auxílio, permitindo-nos fazer inferências sobre a qualidade de todo o lote mediante inspeção de uma amostra de itens escolhidos aleatoriamente.
- Imagine que você trabalha em uma fábrica e deseja reduzir o número de rejeitos de uma determinada matéria-prima. Você pode utilizar a estatística para identificar padrões que levam ao erro, como máquina específica, horário que os erros acontecem, pessoa ou equipe que mais falham, etc. Pode também analisar a equipe que menos comete erros, entender o seu padrão e replicar.
- Agora imagine que você está tendo um alto custo de estoque. Utilize a estatística para determinar quando comprar produtos, levando em conta fatores como sazonalidade, eventos, picos de demanda e comportamento do consumidor.

- Muito além de planilhas e tabelas a estatística é a principal fonte para tomada de decisões assertivas, analisando cenários e relacionando os mais variados fatores internos e externos da organização. Com isso, acaba por impulsionar os resultados da empresa.

CAPÍTULO 2 – População e Amostra

1- Variáveis

As variáveis nos estudos estatísticos são os valores que assumem determinadas características dentro de uma pesquisa e podem ser classificadas em qualitativas ou quantitativas, ou seja, as variáveis podem ter valores numéricos ou não numéricos.

Variável é, convencionalmente, o conjunto de resultados possíveis de um fenômeno.

As variáveis qualitativas não podem ser expressas numericamente, pois relacionam situações como a cor da pele, cor dos olhos, marca de refrigerante, marca de automóvel, preferência musical entre outras. Elas podem ser divididas em ordinais e nominais. As variáveis qualitativas ordinais, apesar de não serem numéricas, obedecem a uma relação de ordem, por exemplo: conceitos como ótimo, bom, regular e ruim, classe social, grau de instrução, etc. Já as variáveis qualitativas nominais não estão relacionadas à ordem, elas são identificadas apenas por nomes, por exemplo, as cores: vermelho, amarelo, preto, azul, rosa, verde, etc. Também como exemplo de nominais temos as marcas de carros, nome de bebidas, local de nascimento entre outros. Resumindo, uma variável é qualitativa quando seus valores são expressos por atributos.

No caso das variáveis quantitativas usamos a representação numérica. Elas podem ser classificadas em discretas e contínuas. As variáveis quantitativas discretas acontecem relacionadas a situações limitadas, por exemplo: número de revistas vendidas, quantidade de consultas médicas, número de filhos de um casal. No caso das variáveis quantitativas contínuas, a abrangência pertence a um intervalo que se caracteriza por infinitos valores, como exemplo podemos citar: o peso de um produto, altura dos alunos de uma escola, velocidade de objetos, entre outras situações.

Uma variável originalmente quantitativa pode ser coletada de forma qualitativa. Por exemplo, a variável idade, medida em anos completos, é quantitativa (contínua);

mas, se for informada apenas a faixa etária (0 a 5 anos, 6 a 10 anos, etc...), é qualitativa (ordinal). Outro exemplo é o peso dos lutadores de boxe, uma variável quantitativa (contínua) se trabalharmos com o valor obtido na balança, mas qualitativa (ordinal) se o classificarmos nas categorias do boxe (peso-pena, peso-leve, peso-pesado, etc.).

Outro ponto importante é que nem sempre uma variável representada por números é quantitativa. O número do telefone de uma pessoa, o número da casa, o número de sua identidade. Às vezes o sexo do indivíduo é registrado na planilha de dados como 1 se macho e 2 se fêmea, por exemplo. Isto não significa que a variável sexo passou a ser quantitativa.

Em síntese: As características das populações são chamadas de variáveis, que podem ser divididas em:

Qualitativas:

- nominal (sexo, cor dos olhos...)
- ordinal (classe social, grau de instrução...)

Quantitativas:

- contínua (peso, altura...)
- discreta (número de filhos, número de carros...)

VARIÁVEL QUALITATIVA: <i>variável não numérica, mas que apresenta qualidade ou atributo do indivíduo ou objeto pesquisado.</i>	NOMINAL: <i>é a variável para a qual não há nenhuma ordenação possível.</i>	Exemplos: <i>Gênero (masculino ou feminino), Religião (católico, evangélico, espírita,), Nacionalidade (brasileiro ou estrangeiro) etc.</i>
	ORDINAL: <i>é a variável para a qual há ordenação em seus resultados.</i>	Exemplos: <i>Nível de instrução (ensino fundamental, ensino médio, ensino superior ou pós-graduação), Classe social (classe baixa, classe média, classe alta) etc.</i>
VARIÁVEL QUANTITATIVA: <i>variável que é expressa numericamente.</i>	DISCRETA: <i>é a variável que assume determinado valor.</i>	Exemplos: <i>Número de funcionários em uma empresa, Número de passageiros por voo (avião); número de pessoas inscritas no INSS, tempo de serviço em uma empresa, idades dos funcionários de uma empresa, etc.</i>
	CONTÍNUA: <i>é a variável que assume determinado intervalo de valor, resultante da coleta de dados na variável discreta.</i>	Exemplos: <i>Faixa salarial dos funcionários de uma empresa (de 1001 a 1500 reais, De 1501 a 2000 reais, ...); Tempo de serviço em uma empresa (até 2 anos, de 2,1 a 4 anos, de 4,1 a 6 anos,...); estatura (até 1,60 m, de 1,61 a 1,65 m, de 1,66 a 1,70 m,...) etc.</i>

Fonte: Pandolfi et al.

2- População e Amostra

Entende-se por população o conjunto de objetos, itens ou eventos com alguma característica ou propriedade comum mensurável, ordenável ou comparável de acordo com os limites propósitos e objetivos do estudo. Sendo assim, formam-se um conjunto

de elementos que formam o universo que é passível de serem observados, sob as mesmas condições.

Ao conjunto de entes portadores de, pelo menos, uma característica comum denominamos população estatística ou universo estatístico.

Na maioria das vezes, por impossibilidade ou inviabilidade econômica ou temporal, limitamos as observações referentes a uma determinada pesquisa a apenas uma parte da população. A essa parte proveniente da população em estudo determinamos amostra.

Uma amostra é um subconjunto finito de uma população.

Pode-se classificar a população em finita e infinita.

População Finita: é aquela que se consegue enumerar todos os elementos que a formam. Refere-se a um universo limitado em uma dada unidade de tempo. Exemplificando pode-se dizer que a quantidade de automóveis produzidos por uma fábrica em um mês, a população de uma cidade e o número de alunos de uma sala de aula são exemplos de uma população finita.

População Infinita: é aquela cujos elementos não podem ser contados. Refere-se a um universo não delimitado. Os resultados (cara ou coroa) obtidos em sucessivos lances de uma moeda, o conjunto dos números inteiros, reais ou naturais são exemplos de populações infinitas.

Como em qualquer estudo estatístico tem-se em mente pesquisar uma ou mais características dos elementos de alguma população. Essa característica deve estar perfeitamente definida e isso se dá quando, considerado um elemento qualquer, podemos afirmar, sem ambiguidade, se esse elemento pertence ou não à população.

3- Amostragem

A amostragem – inspeção de parte da população – é mais econômica mais rápida (reduz o número de dados), incomoda menos (consulta menor número de pessoas) e

pode até apresentar resultados mais confiáveis: é que poucas consultas permitem contar com pessoal especializado, mais bem qualificado.

Grande parte das pesquisas científica ou de resoluções de problemas de engenharia é feita por amostragem, ou seja, observamos apenas um subconjunto de elementos da população. Observações:

1. a população é grande ou infinita.
2. as observações ou mensurações têm alto custo.
3. as medidas exigem testes destrutivos.
4. há necessidade de rapidez, etc.

Em geral, o uso de amostragem leva à redução de custos e tempo. Mas a amostragem precisa ser feita com critérios, pois pretendemos ter amostras que permitam, a partir de uma análise estatística apropriada, obter conclusões satisfatórias sobre toda a população. Existem várias técnicas de amostragem, cada uma tem vantagens e desvantagens, e a escolha deverá ser feita pelo pesquisador de acordo aos objetivos propostos pela pesquisa. Os principais modelos de amostragem probabilística são: a amostragem aleatória simples, a amostragem estratificada proporcional e amostragem sistemática. Os principais modelos de amostragem não probabilística são: a amostragem acidental, a amostragem de voluntários, a amostragem por escolhas racionais, a amostragem por julgamento do especialista.

Exemplos para selecionarmos amostras:

Se quiséssemos fazer um estudo sobre as preferências musicais entre os 500 alunos de uma escola, como escolheríamos uma amostra? Para correr menos riscos de chegar a conclusões erradas, devemos estabelecer um número mínimo de elementos para compor a amostra, que não deve ser inferior a 10% do total de elementos da população. Assim, nossa amostra deve ser de, pelo menos, 50 alunos.

- **Amostragem aleatória (ao acaso):** todos os elementos da população têm igual possibilidade de serem selecionados para constituir a amostra. No nosso exemplo, poderíamos atribuir um número a cada aluno da escola e escolher ao acaso 50 números para obter os 50 alunos que constituem a amostra.
- **Amostragem sistemática:** os elementos são selecionados para a amostra por um sistema preestabelecido. No caso dos alunos, considerando-os ainda numerados, poderíamos escolher os números segundo uma lei de formação: 2,4,6,8 e, assim, sucessivamente, até obtermos os 50 alunos.

- **Amostragem estratificada:** é utilizada quando a população está dividida em estratos ou grupos diferenciados. No exemplo dos alunos da escola podemos considerar cada série de escolaridade como um estrato e assim escolher, em cada uma, um determinado número de alunos, isto é, uma amostra dentro de cada estrato. A escolha das amostras poderá ser feita por um dos dois processos anteriores. O número de elementos das amostras de cada estrato não precisa ser igual, mas deve respeitar as mesmas proporções. Se escolhermos uma amostra de 1ª série com muitos alunos e uma de 2ª série com poucos, correremos o risco de ter resultados que não são representativos da população.

Exercício Resolvido

Numa escola, os professores de Educação Física resolveram fazer um estudo sobre o “peso” dos alunos da 1ª série do Ensino Médio. Sabendo que há 400 alunos nessa série, selecione uma amostra pelo:

- a) método aleatório.
- b) método sistemático.
- c) método estratificado.

Resolução:

- a) Para construir uma amostra pelo método da amostragem aleatória, podemos:
 - Fazer uma lista com os 400 alunos objeto do estudo estatístico, numerada de 1 a 400.
 - Sortear ao acaso 40 números (10% de 400). Os 40 alunos sorteados serão os componentes da amostra.
- b) Para a amostra pelo método da amostragem sistemática, podemos proceder da seguinte forma:
 - Elaborar a lista como no exemplo anterior.
 - Sortear um número de 1 a 10; se o número sorteado for 5, por exemplo, formaremos nossa amostra por: 5,15,25,35,45,55,65,75,85,95,105,115, 125, 135, até obter os 40 alunos.
- c) Para a amostragem estratificada é preciso considerar o número de alunos por classe de 1ª série e, então, sortear os alunos que participarão da pesquisa em quantidades proporcionais ao número de alunos de cada classe. Suponhamos que a quantidade de alunos por classe seja a seguinte:

Classe	Alunos
A	35
B	40
C	45
D	30
E	42
F	36
G	45
H	42
I	45
J	40

Sabemos que nossa amostra deve ter 40 alunos tomados de 10 classes. Para a constituição da amostra precisamos, em primeiro lugar, calcular a porcentagem de alunos de cada classe em relação à população de 400 alunos.

Classe	Alunos	%
A	35	8,75
B	40	10
C	45	11,25
D	30	7,5
E	42	10,5
F	36	9
G	45	11,25
H	42	10,5
I	45	11,5
J	40	10
Total	400	100

A primeira série A, com 35 alunos, tem 8,75% dos elementos da população, logo, na amostra, 8,75% dos elementos deverão ser dessa classe. O mesmo deve ser utilizado para as demais classes. Como precisamos de uma amostra de 40 alunos, calculamos: A: 8,75% de 40 é $3,5 \cong 4$ alunos

B: 10% de 40 é 4

C: 11,25% de 40 é $4,5 \cong 4$

D: 7,5% de 40 é 3

E: 10,5% de 40 é $4,2 \cong 4$

F: 9% de 40 é $3,6 \cong 4$

G: 11,25% de 40 é $4,5 \cong 4$

H: 10,5% de 40 é $4,2 \cong 4$

I: 11,5% de 40 é $4,6 \cong 5$

J: 10% de 40 é 4

Observe que esses dados foram arredondados para o inteiro mais próximo porque indicam números de alunos. Podemos organizar uma tabela.

Classe	Alunos	%	Amostra
A	35	8,75	4
B	40	10	4
C	45	11,25	4
D	30	7,5	3
E	42	10,5	4
F	36	9	4
G	45	11,25	4
H	42	10,5	4
I	45	11,5	5
J	40	10	4
Total	400	100	40

Obs.: Na última coluna da tabela está representada a quantidade de alunos de cada classe que comporá a amostra. A escolha desses alunos poderá ser feita por um dos processos anteriores.

4- CENSO

É o exame completo de toda uma população.

Quanto maior for a amostra, mais precisas e confiáveis serão as induções feitas sobre a população. Assim, resultados mais perfeitos serão obtidos pelo censo. No entanto, muitas vezes o emprego de amostras, com certo rigor técnico, pode levar a resultados mais confiáveis ou até mesmo melhores daqueles que seriam obtidos mediante um censo.

As razões para se recorrer a amostras são: menor custo e tempo para levantar dados; melhor investigação dos elementos observados.

CAPÍTULO 3 – Séries Estatísticas

1- Tabelas

A apresentação tabular é uma apresentação numérica que consiste em dispor os dados em linhas e colunas distribuídos de modo ordenado, segundo algumas regras práticas definidas pelo Conselho Nacional de Estatística e pelo IBGE.

As tabelas apresentam a vantagem de conseguir expor, sinteticamente em um só local, os resultados sobre determinado assunto, de forma a se obter uma visão global mais rápida do que se pretende analisar.

A integração de valores contidos nas tabelas favorece a utilização de representações gráficas, geralmente, uma maneira mais útil e elegante de demonstrar as características que serão analisadas.

Tabela é um quadro que resume um conjunto de observações.

Uma tabela compõe-se de

- **corpo** – conjunto de linhas e colunas que contêm informações sobre a variável em estudo.
- **cabeçalho** – parte superior da tabela que especifica o conteúdo das colunas.
- **coluna indicadora** – parte da tabela que especifica o conteúdo das linhas.
- **linhas** – retas imaginárias que facilitam a leitura, no sentido horizontal, de dados que se inscrevem nos seus cruzamentos com as colunas.
- **casa ou célula** – espaço destinado a um só número.
- **título** – conjunto de informações, as mais completas possíveis, respondendo às perguntas: O quê?, Quando?, Onde?, localizado no topo da tabela.

Exemplo:

Tabela 11 – Evolução da renda dos Empreendedores Brasileiros				
Faixa de Renda	2000(%)	2000(%)	2000(%)	2000(%)
Menos de 3 SM	30	39	43	53
De 3 a 6 SM	30	31	34	22
Mais de 6 a 9 SM	14	12	11	8
Mais de 9 a 15 SM	12	12	9	6
Mais de 15 SM	1	3	3	1
Não sabe Recusou	4	6	2	2
Total	100	100	100	100
Fonte: Pesquisas GEM 2000 – 2003				

Título	Valor em dólares dos principais produtos que o Brasil vende à Argentina	
Cabeçalho	Produto	Valor em dólares (em bilhões)
Coluna Indicadora	Automóveis	606
	Veículos de carga	541
	Autopeças	531
	Motores	264
	Minério	248
	Tratores	130

Fonte: Época 25 de janeiro de 1999



Devem-se considerar ainda os elementos complementares da tabela: fonte, chamadas e notas, colocadas, de preferência, no seu rodapé.

- **Fonte** – indicação da entidade responsável pelo fornecimento dos dados.
- **Notas** – são usadas para conceituação ou esclarecimento em geral.
- **Chamadas** – são usadas para esclarecer certas minúcias em relação a casas, linhas e colunas.

De acordo com a Resolução 886 da Fundação IBGE, nas casas ou células devemos colocar

- ✓ **um traço horizontal (—)** quando o valor é zero, não só quanto à natureza das coisas, como quanto ao resultado do inquérito.
- ✓ **três pontos (...)** quando não temos os dados.
- ✓ **um ponto de interrogação (?)** quando temos dúvida quanto à exatidão de determinado valor.
- ✓ **zero (0)** quando o valor é muito pequeno para ser expresso pela unidade utilizada. Se os valores são expressos em numerais decimais, precisamos acrescentar à parte decimal um número correspondente de zeros (0,0; 0,00; 0,000; ...)

Recomendações para a construção de tabelas

Na construção de tabelas, os dados são apresentados em colunas verticais e linhas horizontais, conforme a classificação dos resultados da pesquisa.

- ✓ A tabela deve ser simples, pois tabelas simples são mais claras e objetivas. Desse modo, é conveniente que grandes volumes de informação sejam descritos em várias tabelas, em vez de em uma só.
- ✓ A tabela deve ser autoexplicativa, ou seja, sua explicação deve estar desvinculada do texto.
- ✓ Nenhuma casa da tabela deve ficar em branco, contendo sempre um número ou um sinal.
- ✓ Se houver duas ou mais tabelas em um texto, deverão receber um número, que será referido no texto.
- ✓ As colunas externas de uma tabela não devem ser fechadas.
- ✓ Nas partes superior e inferior, as tabelas devem ser fechadas por linhas horizontais. O emprego de linhas verticais para a separação de colunas no corpo da tabela é opcional.

- ✓ É conveniente que sejam evitados os arredondamentos. Quando for necessário o arredondamento dos números que compõem a tabela deve ser efetuado segundo critérios de minimização de erros (com isso procura-se evitar o acúmulo de erros de arredondamentos decorrentes do processo de aproximação).
- ✓ Deverá ser mantida uniformidade quanto ao número de casas decimais.
- ✓ Os totais e subtotais devem ser destacados.
- ✓ A tabela deve ser maior no sentido vertical que no horizontal. Entretanto, se uma tabela apresentar muitas linhas e poucas colunas (estreita demais), convém separá-la em uma maior quantidade de colunas. Nesse caso, as colunas deverão ser separadas por linhas duplas.

2- Séries Estatísticas

Denomina-se série estatística toda tabela que apresenta a distribuição de um conjunto de dados estatísticos em função da época, do local ou da espécie (fenômeno).

Em uma série estatística observamos a existência de três elementos ou fatores: o **tempo**, o **espaço** e a **espécie**. Conforme varie um dos elementos da série, podemos classifica-la em histórica (ou temporal), geográfica e específica.

As séries estatísticas consistem na apresentação das informações (variáveis estatísticas) em formas de tabelas, objetivando sintetizar os dados estatísticos observados e tornando-os mais compreensivos. Uma tabela e mesmo um gráfico deve apresentar o cabeçalho, o corpo e o rodapé.

O título deve conter o suficiente para compreender os seguintes elementos fundamentais:

- Fato: fenômeno observado.
- Espaço geográfico: local ou região onde o fato ocorreu.
- Época: refere-se à data ou ao tempo em que o fato foi observado.

O corpo é reservado para o registro dos dados.

O rodapé é reservado para a identificação da fonte de dados.

2.1- Série histórica, temporal, evolutiva ou cronológica

É a série estatística cujos dados são observados segundo a época de ocorrência. O tempo é variável; o fato e o local são fixos. Descreve os valores da variável, em determinado local, discriminados segundo intervalos de tempo variáveis.

Exemplo:

**Produção Agrícola na PB
2004-2009**

Ano	Produção (R\$ 10 ⁶)
2004	4,5
2005	5,3
2006	4,9
2007	5,1
2008	6,8
2009	7,1

Fonte: Dados fictícios

2.2- Série geográfica, espacial, territorial ou de localização

É a série estatística cujos dados são observados segundo a localidade de ocorrência. O local varia; o tempo e o fato são fixos. Descreve os valores da variável, em determinado instante, discriminados segundo regiões.

Exemplo:

**% do PIB destinado a Educação
por País — 2009**

País	% do PIB
EUA	2,5
Brasil	1,0
Japão	7,1
México	0,8

Fonte: Dados fictícios

2.3- Série específica ou categórica

É a série estatística cujos dados são agrupados segundo a modalidade de ocorrência. Fato variável, tempo e local fixos. Descreve os valores da variável, em determinado tempo e local, discriminados segundo especificações ou categorias.

Exemplo:

PIB por setor econômico, Brasil – 2009.

Setor	PIB (US\$ 10 ⁹)
Primário	12,5
Secundário	5,8
Terciário	279,1

Fonte: Dados fictícios

2.4- Série mista ou conjugada – Tabela de dupla entrada

É uma combinação de duas ou mais dos 3 tipos de séries anteriores. Muitas vezes temos necessidade de apresentar, em uma única tabela, a variação de valores de mais de uma variável, isto é, fazer uma conjugação de duas ou mais séries.

Conjugando duas séries em uma única tabela, obtemos uma tabela de dupla entrada. Em uma tabela desse tipo ficam criadas duas ordens de classificação: uma horizontal (linha) e uma vertical (coluna).

Exemplo:

Produção Agrícola, por região, segundo os principais produtos — 2009 (valores em milhões de ton.)

Região	Produtos				Total
	Grãos	Frutas	Legumes	Outros	
Norte	1,5	45,2	5,4	0,5	52,6
Nordeste	2,4	120,0	1,2	4,5	128,1
Centro-Oeste	10,5	10,5	4,2	1,5	26,7
Sudeste	50,7	100,1	10,2	5,9	166,9
Sul	12,5	75,0	15,2	2,5	105,2

Fonte: Dados fictícios

3- Dados Absolutos e Dados Relativos

No estudo sobre formas de apresentação de dados estatísticos, principalmente quando se planeja proceder a uma análise exploratória ou descritiva para uma variável, será apontado que em muitos casos pode ser mais conveniente a apresentação de resultados na forma relativa ao invés de na forma absoluta.

3.1- Dados Absolutos

Os dados estatísticos resultantes da coleta direta da fonte, sem outra manipulação senão a contagem ou medida, são chamados **dados absolutos**. Por

exemplo, se dissermos “uma sala possui 20 alunos”, esses dados são considerados absolutos.

A leitura dos dados absolutos é sempre enfadonha e inexpressiva, embora esses dados traduzam um resultado exato e fiel, não têm a virtude de ressaltar de imediato as suas conclusões numéricas. Daí o uso imprescindível que faz a Estatística dos dados relativos.

3.2- Dados Relativos

Dados relativos: resultado de comparações por quocientes (razões) que se estabelecem entre dados absolutos, e têm por finalidade realçar ou facilitar as comparações entre quantidades.

Traduzem-se os dados relativos, em geral, por meio de **porcentagens, índices, coeficientes e taxas.**

3.2.1- As porcentagens

O emprego da porcentagem é de grande valia quando queremos destacar a participação da parte no todo. Formalmente, PORCENTAGEM é uma RAZÃO cujo CONSEQUENTE (DENOMINADOR) é 100. Se p =porcentagem, P é o valor da Parte e B é o valor do todo (ou Base), então pela definição:

$$\frac{p}{100} = \frac{P}{B}$$

Exemplos:

Consideremos a série abaixo.

MATRÍCULAS NAS ECOLAS DA CIDADE A - 2016

CATEGORIA	NÚMERO DE ALUNOS
Ensino Fundamental	19286
Ensino Médio	1681
Ensino Superior	234
Total	21201

Dados Fictícios

Calculemos as porcentagens dos alunos de cada nível.

Ensino Fundamental $\rightarrow 19286 \times 100 : 21201 = 90,96 = 91\%$

Ensino Médio $\rightarrow 1681 \times 100 : 21201 = 7,92 = 7,9\%$

Ensino Superior $\rightarrow 234 \times 100 : 21201 = 1,10 = 1,1\%$

Com esses dados, podemos formar uma nova coluna na série em estudo.

MATRÍCULAS NAS ESCOLAS DA CIDADE A - 2016

CATEGORIA	NÚMERO DE ALUNOS	%
Ensino Fundamental	19286	91,0
Ensino Médio	1681	7,9
Ensino Superior	234	1,1
Total	21201	100,0

Dados Fictícios

Os valores dessa nova coluna nos dizem que, de cada 100 alunos da cidade A, 91 estão matriculados no Ensino Fundamental; 8, aproximadamente, no Ensino Médio e 1 no Ensino Superior. O emprego da porcentagem é de grande valia quando é nosso intuito destacar a participação da parte no todo.

Consideremos, agora, a série.

MATRÍCULAS NAS ESCOLAS DAS CIDADES A E B – 2016

CATEGORIA	NÚMERO DE ALUNOS	
	CIDADE A	CIDADE B
Ensino Fundamental	19286	38660
Ensino Médio	1681	3399
Ensino Superior	234	424
Total	21201	42483

Dados Fictícios

Qual das cidades tem, comparativamente, maior número de alunos em cada nível? Como o número total de alunos é diferente nas duas cidades, não é fácil concluir a respeito usando os dados absolutos. Porém, usando as porcentagens, tal tarefa fica bastante facilitada. Assim, acrescentando, na tabela anterior, as colunas correspondentes às porcentagens, obtemos

MATRÍCULAS NAS ESCOLAS DAS CIDADES A E B – 2016

CATEGORIA	CIDADE A		CIDADE B	
	NÚMERO DE ALUNOS	%	NÚMERO DE ALUNOS	%
Ensino Fundamental	19286	91,0	38660	91,0
Ensino Médio	1681	7,9	3399	8,0
Ensino Superior	234	1,1	424	1,0
Total	21201	100,0	42483	100,0

Dados Fictícios

o que nos permite dizer que, comparativamente, contam, praticamente, com o mesmo número de alunos em cada nível.

Observações:

- Do mesmo modo que tomamos 100 para base de comparação, também podemos tomar outro número qualquer, entre os quais destacamos o número 1. É claro que, supondo o total igual a 1, os dados relativos das parcelas serão menores que 1.
- Em geral, quando usamos 100 para base, os dados são arredondados até a primeira casa decimal, e quando tomamos 1 por base, são arredondados até a terceira casa decimal.

3.2.2- Os índices. Índices econômicos.

Os **índices** são razões entre duas grandezas tais que uma não inclui a outra.

São exemplos de índices:

$$\text{índice cefálico} = \frac{\text{diâmetro transverso do crânio}}{\text{diâmetro longitudinal do crânio}} \times 100$$

$$\text{quociente intelectual} = \frac{\text{idade mental}}{\text{idade cronológica}} \times 100$$

$$\text{densidade demográfica} = \frac{\text{população}}{\text{superfície}}$$

Índices econômicos:

$$\text{produção per capita} = \frac{\text{valor total da produção}}{\text{população}}$$

$$\text{consumo per capita} = \frac{\text{consumo do bem}}{\text{população}}$$

$$\text{renda per capita} = \frac{\text{renda}}{\text{população}}$$

$$\text{renda per capita} = \frac{\text{receita}}{\text{população}}$$

3.2.3- Os coeficientes

Os **coeficientes** são razões entre o número de ocorrências e o número total (número de ocorrências e número de não ocorrências).

São exemplos de coeficientes:

$$\text{coeficiente de natalidade} = \frac{\text{número de nascimento}}{\text{população total}}$$

$$\text{coeficiente de mortalidade} = \frac{\text{número de óbitos}}{\text{população total}}$$

Coeficientes educacionais:

$$\text{coeficiente de evasão escolar} = \frac{\text{número de alunos evadidos}}{\text{número inicial de matrículas}}$$

$$\text{coeficiente de aproveitamento escolar} = \frac{\text{número de alunos aprovados}}{\text{número final de matrículas}}$$

$$\text{coeficiente de recuperação escolar} = \frac{\text{número de alunos recuperados}}{\text{número de alunos em recuperação}}$$

3.2.4- As taxas

As taxas são coeficientes multiplicados por uma potência de 10 (10, 100, 1000, etc.) para tornar o resultado mais inteligível.

São exemplos de taxas:

Taxa de mortalidade = coeficiente de mortalidade x 1000

Taxa de natalidade = coeficiente de natalidade x 1000

Taxa de evasão escolar = coeficiente de evasão escolar x 100

Exemplo:

O Estado A apresentou 733986 matrículas na 1ª série, no início do ano de 2017, e 683816 no fim do ano. O Estado B apresentou, respectivamente, 436127 e 412457 matrículas. Qual é o Estado que apresentou maior evasão escolar?

$$A \rightarrow \text{taxa de evasão escolar} = \frac{733986 - 683816}{733986} \times 100 = 0,0683 \times 100 = 6,83 = 6,8\%$$

$$B \rightarrow \text{taxa de evasão escolar} = \frac{436127 - 412457}{436127} \times 100 = 0,0542 \times 100 = 5,42 = 5,4\%$$

Resposta: O Estado que apresentou maior evasão escolar foi o Estado A.

CAPÍTULO 4 – Gráficos Estatísticos

1- Gráfico Estatístico

O gráfico estatístico é uma forma de apresentação dos dados estatísticos, cujo objetivo é o de produzir, no investigador ou no público em geral, uma impressão mais rápida e viva do fenômeno em estudo, já que os gráficos falam mais rápido à compreensão que as séries.

Para tornar possível uma representação gráfica, estabelece-se uma correspondência entre os termos da série e determinada figura geométrica, de tal modo que cada elemento da série seja representado por uma figura proporcional. A representação gráfica de um fenômeno deve obedecer a certos requisitos fundamentais, para ser realmente útil.

- ❖ **Simplicidade** – o gráfico deve ser destituído de detalhes de importância secundária, assim como de traços dispensáveis que possam levar o observador a uma análise demorada ou com erros.
- ❖ **Clareza** – o gráfico deve permitir uma correta interpretação dos valores representativos do fenômeno em estudo.

- ❖ **Veracidade** – o gráfico deve comunicar a verdade sobre o fenômeno em estudo.

Os gráficos constituem uma forma clara e objetiva de apresentar dados estatísticos. A intenção é a de proporcionar aos leitores em geral a compreensão e a veracidade dos fatos. De acordo com a característica da informação precisamos escolher o gráfico correto. Os mais usuais são: gráfico de segmentos ou linhas, gráfico de barras ou colunas e gráfico de setores.

2- Diagramas

Os diagramas são gráficos geométricos de, no máximo, duas dimensões; para sua construção, em geral, fazemos uso do sistema cartesiano.

Dentre os principais diagramas, destacam-se:

2.1- Gráfico de segmento ou gráfico em linha

Este tipo de gráfico emprega a linha poligonal para representar a série estatística. Constitui uma aplicação do processo de representação das funções num sistema de coordenadas cartesianas. Nesse sistema são utilizadas duas retas perpendiculares; as retas são os eixos coordenados e o ponto de interseção, a origem. O eixo horizontal denomina-se eixo das abscissas (ou eixo x) e o eixo vertical, eixo das ordenadas (ou eixo y).

Exemplo:

Para tornar bem clara a explanação, considere a seguinte série.

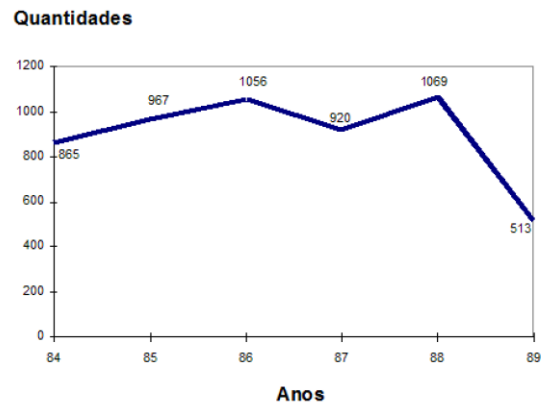
**Produção de veículos de autopropulsão
Brasil - 1984-89**

Anos	Quantidades (x 1000)
1984	865
1985	967
1986	1056
1987	920
1988	1069
1989	513

Fonte: ANFAVEA

Vamos tomar os anos como abscissas e as quantidades como ordenadas. Desse modo, um ano dado (x) e a respectiva quantidade (y) formam um par ordenado (x,y),

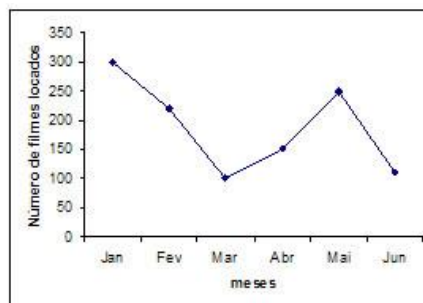
que pode ser representado num sistema cartesiano. Determinados, graficamente, todos os pontos da série, usando as coordenadas, ligam-se todos esses pontos, dois a dois, por segmentos de reta, formando a poligonal, que é o gráfico em linha ou de segmentos correspondente à série em estudo.



Exemplo:

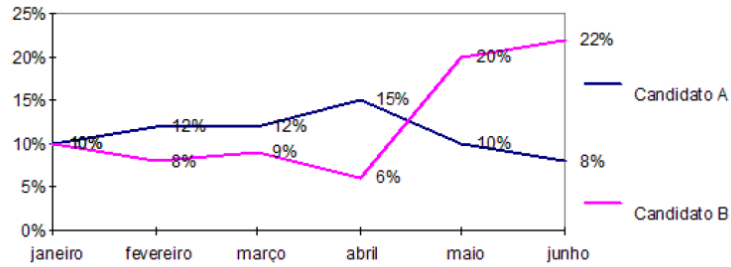
Uma locadora de filmes em DVD registrou o número de locações no 1º semestre do ano de 2012. Os dados foram expressos em um gráfico de segmentos.

Mês	Número de filmes locados
Janeiro	300
Fevereiro	220
Março	100
Abril	150
Mai	250
Junho	110



Obs.: Quando é representada em um mesmo sistema de eixos a variação de dois ou mais fenômenos, a título de comparação, obtém-se um gráfico em linhas comparativo ou poligonal comparativo.

Pesquisa Eleitoral, na cidade X, no período de janeiro a junho de 2017



2.2- Gráfico em colunas ou barras

É a representação de uma série por meio de retângulos, dispostos verticalmente (em colunas) ou horizontalmente (em barras).

Quando dispostos em colunas, os retângulos têm a mesma base e as alturas são proporcionais aos respectivos dados.

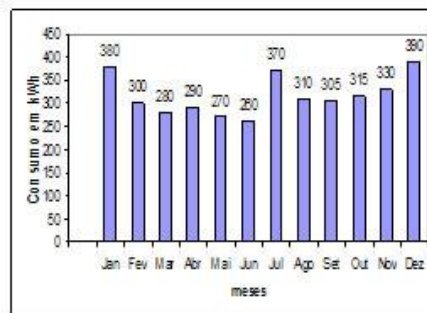
Quando dispostos em barras, os retângulos têm a mesma altura e os comprimentos são proporcionais aos respectivos dados.

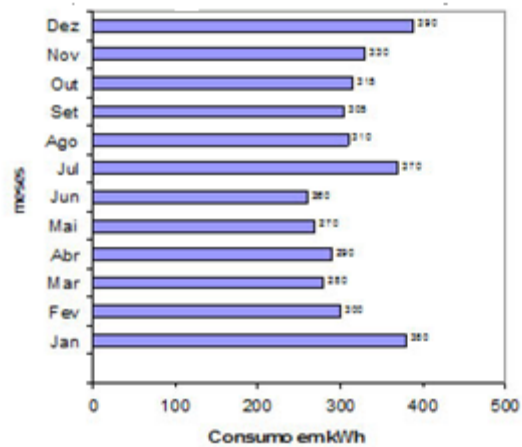
Dessa forma assegura-se a proporcionalidade entre as áreas dos retângulos e os dados estatísticos.

Exemplo:

Consumo de energia elétrica no decorrer do ano de 2015 de uma família.

Mês	Consumo em kWh
Janeiro	380
Fevereiro	300
Março	280
Abril	290
Mai	270
Jun	260
Jul	370
Agosto	310
Set	305
Out	315
Nov	330
Dezembro	390



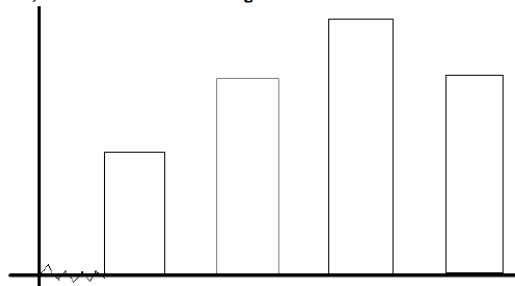


Notas:

- Sempre que os dados a serem inscritos são extensos, deve-se dar preferência aos gráficos em barras (séries geográficas e específicas).
- A ordem a ser observada é a cronológica, se a série for histórica, e a decrescente, se for geográfica ou categórica.
- A distância entre as colunas (ou barras), por questões estéticas, não deverá ser menor que a metade nem maior que os dois terços da largura (ou da altura) dos retângulos.
- A distância entre duas barras deve ser constante.

IMPORTANTE:

As larguras dos retângulos, bem como as distâncias entre eles, devem manter uma regularidade.



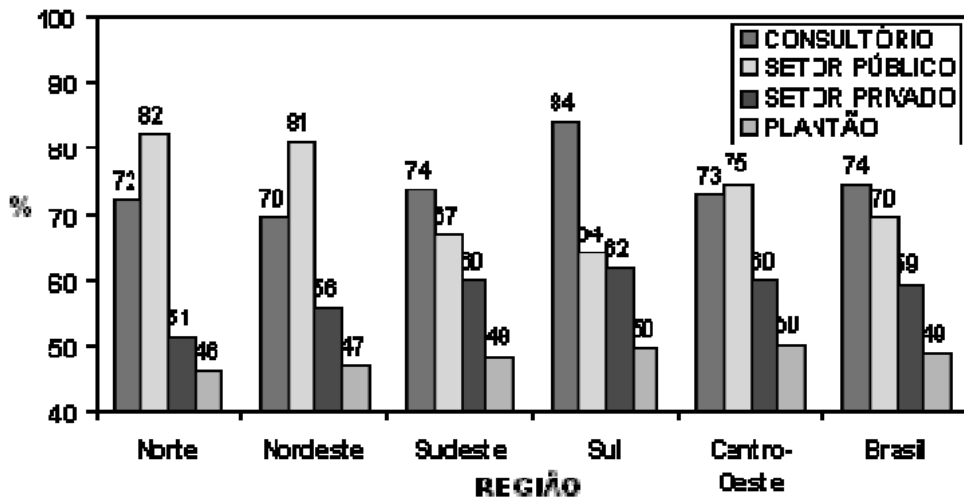
Observação:

Os gráficos em colunas ou em barras podem ainda apresentar as seguintes variações.

I- Gráfico em Colunas ou Barras Múltiplas: Geralmente empregado quando queremos representar, simultaneamente, dois ou mais fenômenos, com propósito de comparação.

Exemplos:

a) Em 1995, quase três quartos dos médicos brasileiros atuavam em consultório próprio (74%), mais de dois terços atuavam em instituições do setor público (70%), mais da metade atuavam no setor privado (59%) e quase metade faziam plantão (49%).

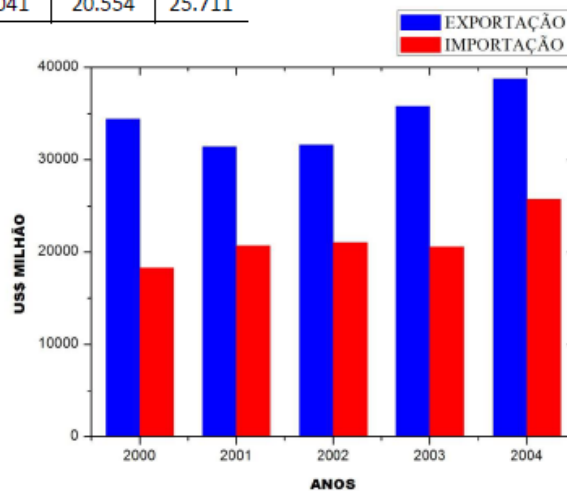


b)

BALANÇA COMERCIAL DO BRASIL – 2000-04

ESPECIFICAÇÕES	VALOR (US\$ 1.000.000)				
	2000	2001	2002	2003	2004
Exportação	34.383	31.414	31.620	35.793	38.783
Importação	18.263	20.661	21.041	20.554	25.711

FONTE: Dados Fictícios.



II- Gráfico em Colunas ou Barras Compostas: É um gráfico constituído por um único retângulo base, subdividido em várias porções, cujos comprimentos são proporcionais às partes no qual se divide o todo.

Constrói-se um retângulo de comprimento qualquer escolhido para representar todo o fato, e faz-se uma regra de três para se calcular o comprimento representativo de cada parte que o compõe.

Exemplo:

Alunos da Escola Y , em 2017

Série	Quantidade
Primeira	400
Segunda	300
Terceira	200
Quarta	100
Total	1000

Será construído o retângulo base com 20 cm.

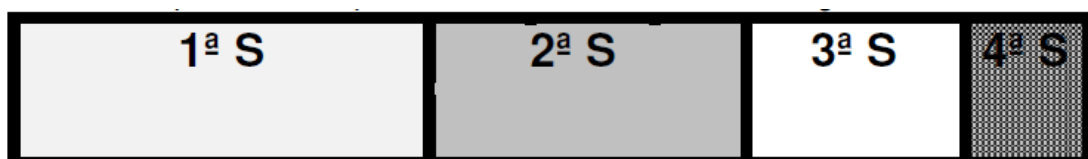
Por exemplo, para a 1ª série, tem-se:

$$1000 \text{ ----- } 20 \text{ cm}$$

$$400 \text{ ----- } x$$

$$x = \frac{400.20}{1000} = 8 \text{ cm}$$

Depois de determinados os dados para as demais séries, pode-se elaborar o gráfico a seguir.



2.3- Gráfico em setores

Este gráfico é construído com base em um círculo, e é empregado sempre que se pretenda ressaltar a participação do dado no total. O total é representado pelo círculo, que fica dividido em tantos setores quantas são as partes.

Os setores são tais que suas áreas são, respectivamente, proporcionais aos dados da série. Cada setor é obtido por meio de uma regra de três simples e direta, lembrando que o total da série corresponde a 360° .

Exemplos:

a)

Rebanho Suíno do Sudeste do Brasil 2005

Estados	Quantidade (mil Cabeças)
Minas Gerais	3.363,7
Espirito Santo	430,4
Rio de Janeiro	308,5
São Paulo	2.035,9
Total	6.138,5

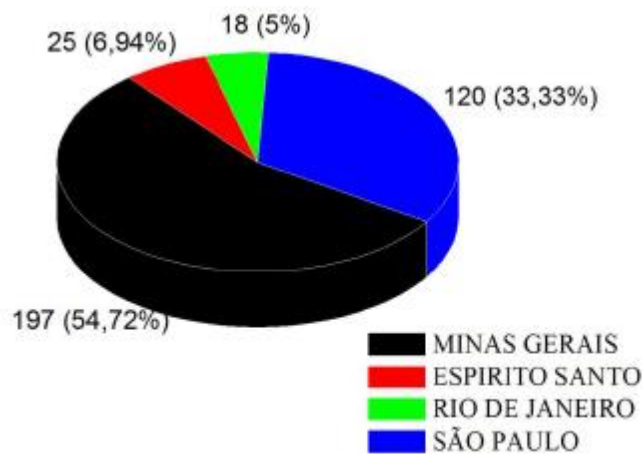
Fonte: Dados Fictícios

$$\begin{aligned}
 &6.138,5 - 360^\circ \\
 &3.363,7 - x_1 \\
 &x_1 = 197,2 \cong 197^\circ
 \end{aligned}$$

$$\begin{aligned}
 &6.138,5 - 360^\circ \\
 &430,4 - x_2 \\
 &x_2 = 25,2 \cong 25^\circ
 \end{aligned}$$

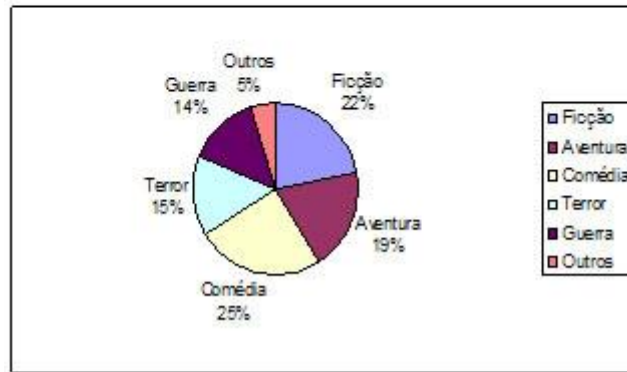
$$\begin{aligned}
 &6.138,5 - 360^\circ \\
 &308,5 - x_3 \\
 &x_3 = 18,0 \cong 18^\circ
 \end{aligned}$$

$$\begin{aligned}
 &6.138,5 - 360^\circ \\
 &2.035,9 - x_4 \\
 &x_4 = 119,3 \cong 120^\circ
 \end{aligned}$$



b) O gráfico a seguir mostra a preferência dos clientes de uma locadora quanto ao gênero dos filmes locados durante a semana.

Gênero	Frequência	
	Absoluta	Relativa
Ficção	88	22%
Aventura	76	19%
Comédia	100	25%
Terror	60	15%
Guerra	56	14%
Outros	20	5%
	400	100%

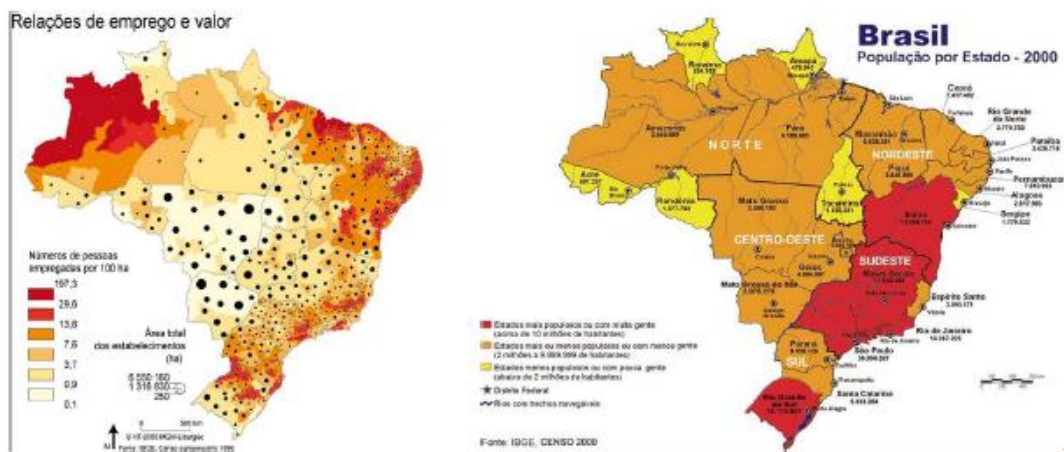


2.4- Cartograma

O cartograma é a representação sobre uma carta geográfica.

Este gráfico é utilizado quando se tem o objetivo de que os dados estatísticos figurem diretamente relacionados com áreas geográficas ou políticas.

Exemplo:



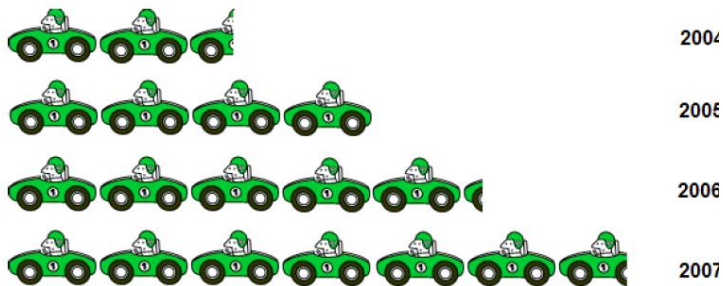
2.5- Pictograma

O pictograma constitui um dos processos gráficos que melhor fala ao público, pela sua forma ao mesmo tempo atraente e sugestiva. A representação gráfica consta de figuras.


Exemplos:



Produção de Veículos da Empresa X, no período de 2004 a 2007



Legenda

 = 10 000 unidades

Fonte: Dados Fictícios

Regras fundamentais para a construção de um pictograma.

- Os símbolos devem explicar-se por si próprios.
- As quantidades maiores são indicadas por meio de um número maior de símbolos, mas não por um símbolo maior.
- Os símbolos comparam quantidades aproximadas, não detalhes minuciosos.
- Os gráficos pictóricos só devem ser usados para comparações, nunca para afirmações detalhadas ou isoladas;

2.6- Gráfico Polar

É o tipo de gráfico ideal para representar séries temporais cíclicas, ou seja, toda série que apresenta uma determinada periodicidade. Como, por exemplo, a variação da precipitação pluviométrica ao longo do ano ou da temperatura ao longo do dia; a arrecadação de um determinado estabelecimento durante uma semana; o consumo de

energia elétrica durante o mês ou o ano; o número de passageiros de uma linha de ônibus ao longo da semana, etc.

O gráfico polar faz uso do sistema de coordenadas polares.

Como construir um gráfico polar:

- Traça-se uma circunferência de raio arbitrário (preferencialmente, a um raio de comprimento proporcional a média dos valores da série).
- Constrói-se uma semirreta (de preferência horizontal) partindo do ponto O (polo) e com uma escala (eixo polar).
- Divide-se a circunferência em tantos arcos quantas forem as unidades temporais.
- Traçam-se semirretas a partir do ponto O (polo) passando pelo ponto de divisão.
- Marcam-se os valores correspondentes da variável, iniciando pela semirreta horizontal (eixo polar).
- Ligam-se os pontos encontrados com segmentos de reta.
- Para fechar o polígono obtido, emprega-se uma linha interrompida.

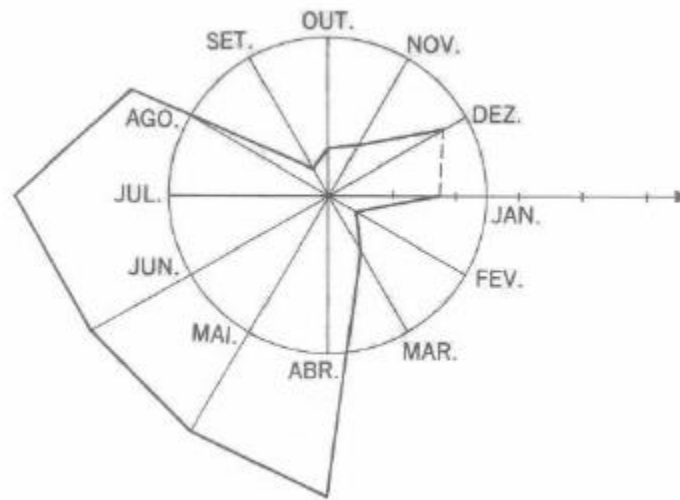
Exemplo:

Precipitação Pluviométrica da Cidade X - 2017

Meses	Precipitação (mm)
Janeiro	174,8
Fevereiro	36,9
Março	83,9
Abril	462,7
Maiο	418,1
Junho	418,4
Julho	538,7
Agosto	323,8
Setembro	39,7
Outubro	66,1
Novembro	83,3
Dezembro	201,2

Fonte: Desconhecida

Precipitação Pluviométrica da Cidade X - 2017



Fonte: Desconhecida

CAPÍTULO 5 – Distribuição de Frequência

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

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

Exemplos:

- a) Distribuição de Frequência 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)

b) 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)

Resumindo:

DISTRIBUIÇÃO DE FREQUÊNCIA	
■	Constitui-se no tipo de tabela mais importante para a estatística descritiva.
■	Distribuição de frequência é constituída por uma tabela resumida, na qual os dados são organizados em grupos de classe ou categorias convenientemente estabelecidas e numericamente ordenadas. Constitui-se num tipo de apresentação que condensa uma coleção de dados conforme as frequências ou repetições de seus valores.
■	As distribuições de frequências são séries heterógrafas, isto é, séries na qual o fenômeno ou fato apresenta graduações ou subdivisões. Embora fixo, o fenômeno varia de intensidade.
■	A construção da distribuição de frequência depende do tipo de dados com os quais se está lidando: contínuos ou discretos .

1-Tabela Primitiva e Rol

Ao fazer a coleta de dados numéricos, certamente eles estarão desorganizados (ou seja, não estarão em ordem crescente nem em ordem decrescente); dados assim formam a **tabela primitiva**.

Considere termos feito uma coleta de dados relativos às estaturas de quarenta alunos, que compõem uma amostra dos alunos de um colégio A, resultando a seguinte tabela de valores.

Estaturas de 40 alunos do Colégio A

166	160	161	150	162	160	165	167	164	160
162	161	168	163	156	173	160	155	164	168
155	152	163	160	155	155	169	151	170	164
154	161	156	172	153	157	156	158	158	161

A esse tipo de tabela, cujos elementos não foram numericamente organizados, denomina-se tabela primitiva. **Dados brutos** são aqueles que não foram numericamente organizados, ou seja, estão na forma com que foram coletados.

Partindo desses dados, é difícil averiguar em torno de que valor tendem a se concentrar as estaturas, qual a menor ou qual a maior estatura ou, ainda, quantos alunos se acham abaixo ou acima de uma dada estatura.

A maneira mais simples de organizar os dados é através de certa ordenação (crescente ou decrescente). A tabela obtida após a ordenação dos dados recebe o nome de **rol**.

Estaturas de 40 alunos do Colégio A

150	154	155	157	160	161	162	164	166	169
151	155	156	158	160	161	162	164	167	170
152	155	156	158	160	161	163	164	168	172
153	155	156	160	160	161	163	165	168	173

2- Distribuição de Frequência

No exemplo trabalhado, a variável estatura, será observada e estudada mais facilmente quando os valores estiverem ordenados em uma coluna e sendo colocado, ao lado de cada valor, o número de vezes que aparece repetido.

Denomina-se frequência o número de alunos que fica relacionado a um determinado valor da variável. A tabela, assim obtida, recebe o nome de **distribuição de frequência**.

ESTAT. (cm)	FREQ.	ESTAT. (cm)	FREQ.	ESTAT. (cm)	FREQ.
150	1	158	2	167	1
151	1	160	5	168	2
152	1	161	4	169	1
153	1	162	2	170	1
154	1	163	2	172	1
155	4	164	3	173	1
156	3	165	1	Total	40
157	1	166	1		

No entanto, o processo apresentado é ainda inconveniente, exigindo muito espaço, mesmo quando o número de valores da variável é de tamanho razoável.

Sendo possível, a solução mais aceitável, pela própria natureza da variável contínua, é o agrupamento dos valores em vários intervalos.

Quando os valores da variável estão sendo agrupados em intervalos, sendo que, em Estatística prefere-se denominar os intervalos de **classes**.

Chamando de frequência de uma classe o número de valores da variável pertencente à classe, os dados podem ser dispostos como na tabela a seguir denominada **distribuição de frequência com intervalo de classe**.

**ESTATURAS DE 40 ALUNOS
DO COLÉGIO A**

ESTATURAS (cm)	FREQÜÊNCIA
150 – 154	4
154 – 158	9
158 – 162	11
162 – 166	8
166 – 170	5
170 – 174	3
Total	40

Ao se agrupar os valores da variável em classes, ganha-se em simplicidade, mas perde-se em pormenores.

O que se pretende com a construção desta nova é realçar o que há de essencial nos dados e, também, tornar possível o uso de técnicas analíticas para sua total descrição, até porque a Estatística tem por finalidade específica analisar o conjunto de valores, desinteressando-se por casos isolados.

Obs.:

- Como o objetivo, desde o início, é a obtenção de uma distribuição de frequência com intervalo de classe, torna-se necessário fazer uma tabulação, com cada traço correspondendo a um valor.
- Quando os dados estão organizados em uma distribuição de frequência, são usualmente denominados **dados agrupados**.

ESTATURAS (cm)	TABULAÇÃO	FREQÜÊNCIAS
150 † 154	<input type="checkbox"/>	4
154 † 158	<input checked="" type="checkbox"/> <input type="checkbox"/>	9
158 † 162	<input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	11
162 † 166	<input checked="" type="checkbox"/> <input type="checkbox"/>	8
166 † 170	<input checked="" type="checkbox"/>	5
170 † 174	<input type="checkbox"/>	3
Total		40

3- Elementos de Uma Distribuição de Frequência

3.1- Classe

Classe de frequência ou, simplesmente, classes são intervalos de variação da variável.

As classes são representadas simbolicamente por i , sendo $i = 1, 2, 3, \dots, k$ (onde k é de o número total de classes da distribuição). O número total de valores da variável é simbolizado por n .

Assim, no exemplo, o intervalo 154 † 158 define a segunda classe ($i = 2$), o intervalo 166 † 170 define a quinta classe ($i = 5$) e assim por diante. Como a distribuição tem seis classes, logo $k = 6$. a variável x assume 40 valores, logo $n = 40$.

3.2- Limites de Classe

Denominam-se limites de classe os extremos de cada classe.

O menor número é o **limite inferior da classe** (l_i) e o maior número, o **limite superior da classe** (L_i). Na segunda classe, por exemplo, tem-se: $l_2 = 154$ e $L_2 = 158$.

Obs.:

De acordo com o IBGE as classes devem ser escritas como *desta quantidade até menor que aquela*, usando para isso o símbolo \vdash . Assim, $l_i \vdash L_i$ significa inclusão de l_i e exclusão de L_i . O indivíduo com estatura 158 cm estaria na terceira classe ($i = 3$) e não na segunda.

3.3- Amplitude de um Intervalo de Classe

Amplitude de um intervalo de classe ou, simplesmente, **intervalo de classe** é a medida do intervalo que define a classe.

O intervalo de classe i é simbolizado por h_i e é obtido pela diferença entre seus limites. $h_i = L_i - l_i$.

No exemplo empregado, o tamanho do intervalo da segunda classe (h_2) vale:

$$h_2 = L_2 - l_2 = 158 - 154 \rightarrow h_2 = 4 \text{ cm.}$$

Todas as outras classes do exemplo também têm intervalo de 4 cm, pois esse é o intervalo entre cada um dos limites inferiores e os limites superiores correspondentes.

3.4- Amplitude Total da Distribuição

Amplitude total da distribuição (AT) é a diferença entre o limite superior da última classe (**limite superior máximo**) e o limite inferior da primeira classe (**limite inferior mínimo**): $AT = L (\text{máx.}) - l (\text{mín.})$

No exemplo utilizado, tem-se: $AT = 174 - 150 \rightarrow AT = 24 \text{ cm.}$

Obs.:

É evidente que, se as classes possuem o mesmo intervalo, verificamos a

relação: $\frac{AT}{h_i} = k \rightarrow \frac{24}{4} = 6.$

3.5- Amplitude Amostral

Amplitude amostral (AA) é a diferença entre o valor máximo e o valor mínimo da amostra: $AA = x(\text{máx.}) - x(\text{mín.})$

No exemplo utilizado, tem-se: $AA = 173 - 150 = 23 \rightarrow AA = 23 \text{ cm.}$

3.6- Ponto Médio de Uma Classe

Ponto médio de uma classe (x_i) é, como o próprio nome indica, o ponto que divide o intervalo de classe em duas partes iguais.

Para se obter o ponto médio de uma classe, calcula-se a semissoma dos limites da classe (média aritmética).

$$x_i = \frac{l_i + L_i}{2}$$

Assim, o ponto médio da segunda classe, no exemplo citado anteriormente, é

$$x_i = \frac{l_i + L_i}{2} \rightarrow x_2 = \frac{l_2 + L_2}{2} \rightarrow x_2 = \frac{154 + 158}{2} = 156 \rightarrow x_2 = 156 \text{ cm}$$

Obs.: O ponto médio de uma classe é o valor que a representa.

3.6- Frequência Simples ou Absoluta

Frequência simples ou **frequência absoluta** ou, simplesmente, **frequência** de uma classe ou de um valor individual é o número de observações correspondentes a essa classe ou a esse valor.

A frequência simples é simbolizada por f_i (lê-se: f índice i ou frequência da classe i). Assim, no exemplo utilizado anteriormente, tem-se:

$$f_1 = 4, f_2 = 9, f_3 = 11, f_4 = 8, f_5 = 5 \text{ e } f_6 = 3.$$

A soma de todas as frequências é representada pelo símbolo de somatório.

$$\sum_{i=1}^k f_i$$

É evidente que

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

Para a distribuição em estudo, tem-se

$$\sum_{i=1}^6 f_i = 40$$

Não havendo possibilidade de engano, usa-se:

$$\sum f_i = 40$$

Pode-se, agora, dar à distribuição de frequência das estaturas dos quarenta alunos do colégio A, a seguinte representação tabular técnica.

i	ESTATURAS (cm)	f _i
1	150 – 154	4
2	154 – 158	9
3	158 – 162	11
4	162 – 166	8
5	166 – 170	5
6	170 – 174	3
		$\Sigma f_i = 40$

4- Número de Classes. Intervalos de Classe.

A primeira preocupação que se tem, na construção de uma distribuição de frequência, é a determinação do número de classes e, conseqüentemente, da amplitude e dos limites dos intervalos de classe.

Para a determinação do número de classes de uma distribuição pode-se lançar mão da **Regra de Sturges**, que permite calcular o número de classes em função da variável: $i \cong 1 + 3,3 \cdot \log n$, sabendo que **i** é o número de classe e **n** é o número total de dados.

Essa regra possibilitou a obtenção da tabela abaixo.

n	i
3 H 5	3
6 H 11	4
12 H 22	5
23 H 46	6
47 H 90	7
91 H 181	8
182 H 362	9
...	...

Além da Regra de Sturges, existem outras fórmulas empíricas que pretendem resolver o problema da determinação do número de classes que deve ter a distribuição.

No entanto, a verdade é que essas fórmulas não levam a uma decisão final; esta vai depender de um julgamento pessoal, que deve estar ligado à natureza dos dados, da unidade empregada para expressá-los, etc.

Decidido o número de classes que deve ter a distribuição, compete agora resolver o problema da determinação da amplitude do intervalo de classe, o que se consegue dividindo a amplitude pelo número de classes: $h = \frac{AT}{i}$.

Quando o resultado não é exato, deve-se arredondá-lo para mais.

Outro problema que surge é a escolha dos limites dos intervalos, os quais deverão ser tais que forneçam, na medida do possível, para os pontos médios, números que facilitem os cálculos – **números naturais**.

No exemplo utilizado, tem-se: para $n = 40 \rightarrow i = 6$.

Logo: $h = \frac{173-150}{6} = \frac{23}{6} = 3,8 = 4$, isto é, seis classes de intervalos iguais a 4.

5- Tipos de Frequências

Frequências simples ou absolutas (f_i) são os valores que realmente representam o número de dados de cada classe.

$$\sum f_i = n$$

Frequências relativas (f_r) são os valores das razões entre as frequências simples e a frequência total.

$$fr_i = \frac{f_i}{\sum f_i}$$

Logo, no exemplo utilizado, a frequência da terceira classe é:

$$fr_3 = \frac{f_3}{\sum f_i} \rightarrow fr_3 = \frac{11}{40} = 0,275 \rightarrow fr_3 = 0,275$$

Evidentemente:

$$fr_i = 1 \text{ ou } 100\%$$

Obs.: O propósito das frequências relativas é o de permitir a análise ou facilitar as comparações.

Frequência acumulada (F_i) é o total das frequências de todos os valores inferiores ao limite superior do intervalo de uma dada classe.

$$F_k = f_1 + f_2 + \dots + f_k$$

ou

$$F_k = \sum f_i \quad (i = 1, 2, \dots, k)$$

Assim, no exemplo apresentado, anteriormente, a frequência acumulada correspondente à terceira classe é:

$F_3 = \sum_{i=1}^3 f_i = f_1 + f_2 + f_3 \rightarrow F_3 = 4 + 9 + 11 \rightarrow F_3 = 24$, o que significa existirem 24 alunos com estatura inferior a 162 cm (limite superior do intervalo da terceira classe).

Frequência acumulada relativa (Fr_i) de uma classe é a frequência acumulada da classe, dividida pela frequência total da distribuição.

$$Fr_i = \frac{F_i}{\sum f_i}$$

Assim, para a terceira classe, temos:

$$Fr_3 = \frac{F_3}{\sum f_i} \rightarrow Fr_3 = \frac{24}{40} = 0,600 \rightarrow Fr_3 = 0,600$$

Resumindo : Tabela com as frequências estudadas

i	Estaturas (cm)	f_i	x_i	fr_i	F_i	Fr_i
1	150 --154	4	152	0,100	4	0,100
2	154 --158	9	156	0,225	13	0,325
3	158 --162	11	160	0,275	24	0,600
4	162 --166	8	164	0,200	32	0,800
5	166 --170	5	168	0,125	37	0,925
6	170 --174	3	172	0,075	40	1,000
Total		40		1,000		

Observação:

O conhecimento dos vários tipos de frequência ajuda para que se possa responder a muitas questões com relativa facilidade, como as seguintes:

- Quantos alunos tem estatura entre 154 cm, inclusive, e 158 cm? Esses são os valores da variável que formam a segunda classe. Como $f_2 = 9$, a resposta é 9 alunos.
- Qual é a porcentagem de alunos cujas estaturas são inferiores a 154 cm? Esses valores são os que formam a primeira classe. Como $fr_1 = 0,100$, obtém-se a resposta multiplicando a frequência relativa por 100: $0,100 \times 100 = 10$. Logo, a porcentagem de alunos é 10%.
- Quantos alunos têm estatura abaixo de 162 cm? É evidente que as estaturas consideradas são aquelas que formam as classes de ordem 1, 2 e 3. Assim, o número de alunos é dado por:

$$f_1 + f_2 + f_3 = \sum_{i=1}^3 f_i = F_3 = 24$$

Portanto, 24 alunos têm estatura abaixo de 162 cm.

- Quantos alunos têm estatura não-inferior a 158 cm? O número é dado por:

$$\sum_{i=3}^6 f_i = f_3 + f_4 + f_5 + f_6 = 11 + 8 + 5 + 3 = 27$$

Ou então:

$$\sum_{i=1}^6 f_i - F_2 = n - F_2 = 40 - 13 = 27$$

6- Distribuição de Frequência sem Intervalos de Classe

Quando se trata de variável discreta de variação relativamente pequena, cada valor pode ser tomado como um intervalo de classe (intervalo degenerado) e, nesse caso, a distribuição é chamada distribuição sem intervalos de classe, tomando a seguinte forma:

x_1	f_1
x_1	f_1
x_2	f_2
\vdots	\vdots
x_n	f_n
	$\sum f_i = n$

Exemplo:

Seja X a variável “número de cômodos das casas ocupadas por vinte famílias entrevistadas”.

i	x_i	f_i
1	2	4
2	3	7
3	4	5
4	5	2
5	6	1
6	7	1
		$\Sigma = 20$

Completada com os vários tipos de frequência, tem-se:

i	x_i	f_i	fr_i	F_i	Fr_i
1	2	4	0,20	14	0,20
2	3	7	0,35	11	0,55
3	4	5	0,25	16	0,80
4	5	2	0,10	18	0,90
5	6	1	0,05	19	0,95
6	7	1	0,05	20	1,00
		$\Sigma = 20$	$\Sigma = 1,00$		

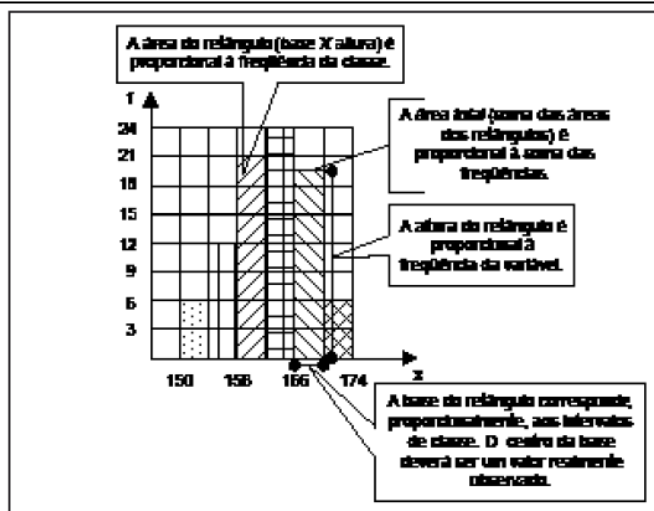
7- Representação Gráfica de Uma Distribuição

Uma distribuição de frequência pode ser representada graficamente pelo **histograma**, pelo **polígono de frequência** e pelo **polígono de frequência acumulada (Ogiva de Galton)**.

Qualquer um dos gráficos mencionados pode ser construído utilizando o primeiro quadrante do sistema de eixos coordenados cartesianos ortogonais. Na linha horizontal (eixo das abscissas) colocam-se os valores da variável; e na linha vertical (eixo das ordenadas), as frequências.

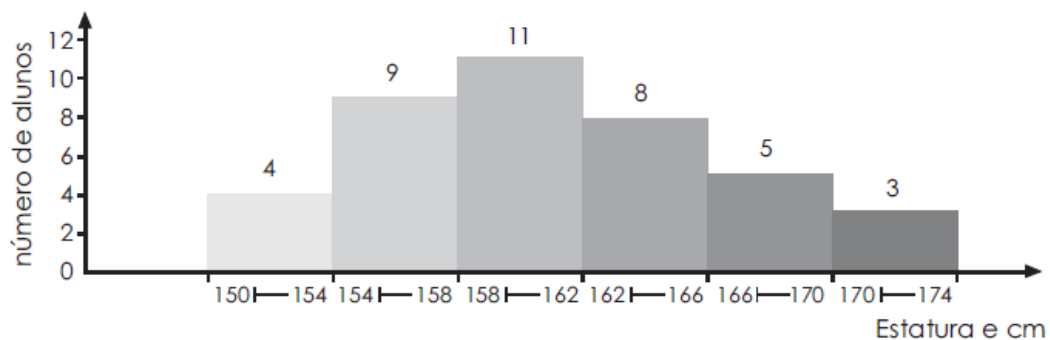
7.1- Histograma

O histograma é formado por um conjunto de retângulos justapostos, cujas bases se localizam sobre o eixo horizontal, de tal modo que seus pontos médios coincidam com os pontos médios dos intervalos de classe.



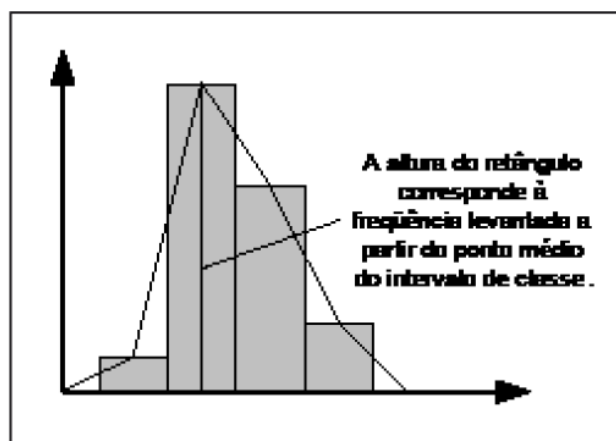
Observações:

- ✓ As larguras dos retângulos são iguais às amplitudes dos intervalos de classe.
- ✓ As alturas dos retângulos devem ser proporcionais às frequências das classes, sendo a amplitude dos intervalos igual. Isso permite que se tomem as alturas numericamente iguais às frequências.
- ✓ O histograma goza da seguinte propriedade: a área de um histograma é proporcional à soma das frequências.
- ✓ Quando se usam as frequências relativas obtém-se um gráfico de área unitária.
- ✓ Quando se pretende comparar duas distribuições, o ideal é fazê-lo pelo histograma de frequências relativas.

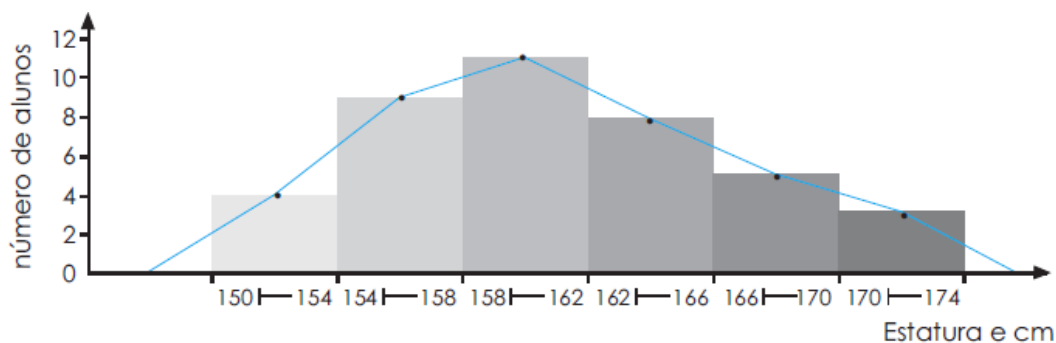


7.2- Polígono de Frequência

O polígono de frequência é um gráfico em linha, sendo as frequências marcadas sobre perpendiculares ao eixo horizontal, levantadas pelos pontos médios dos intervalos de classe.



Obs.: Para realmente se obter um polígono (linha fechada), deve-se completar a figura, ligando os extremos da linha obtida aos pontos médios da classe anterior à primeira e da posterior à última, da distribuição.

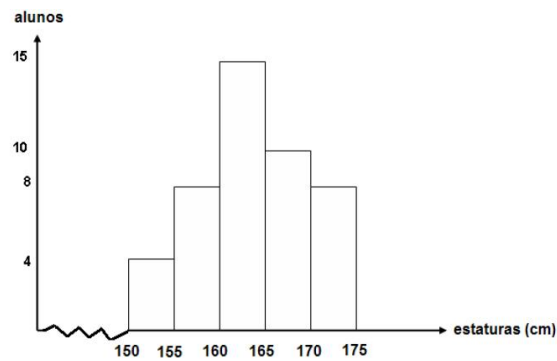


Exemplos:

1) Construir o histograma associado à tabela abaixo, que representa as alturas (em cm) dos alunos da turma A, do Colégio Progresso, em 2018.

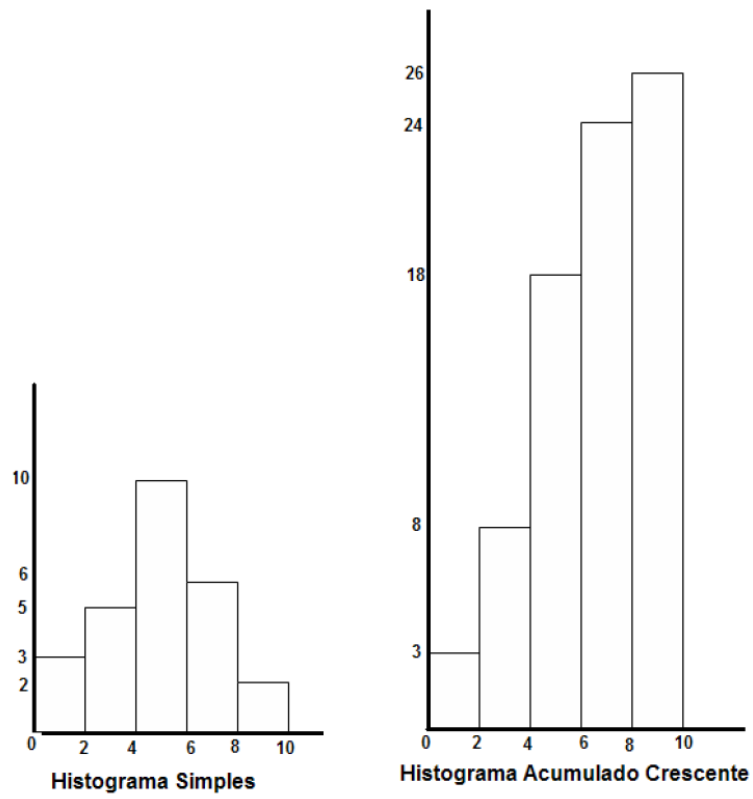
Estaturas (cm)	Alunos (F_i)
150 ----- 155	4
155 ----- 160	8
160 ----- 165	15
165 ----- 170	10
170 ----- 175	8
Total	45

Resposta: Histograma de Frequências Simples

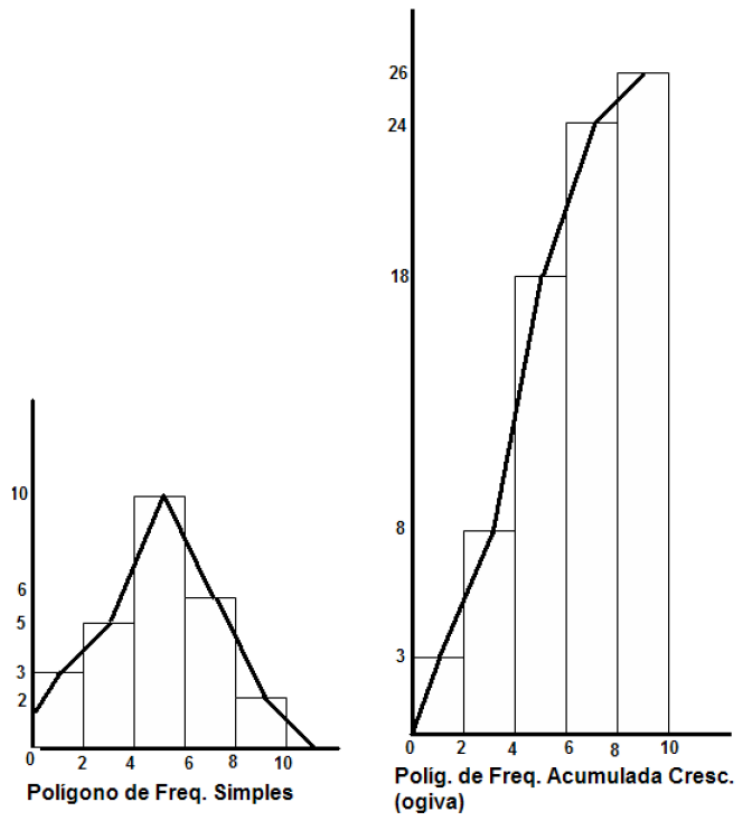


2) A seguir observem como ficam os histogramas para as frequências simples e acumuladas crescentes, referentes às médias dos alunos da turma A, do Colégio União, em julho de 2018.

Notas	Freq. Simples (F_i)	Freq. Acum. Cr.
0 ----- 2	3	3
2 ----- 4	5	8
4 ----- 6	10	18
6 ----- 8	6	24
8 ----- 10	2	26
Σ	26	



3) Observem o polígono de frequência que representa o exemplo anterior.



7.2- Polígono de Frequência Acumulada

O polígono de frequência acumulada é traçado marcando-se as frequências acumuladas sobre perpendiculares ao eixo horizontal, levantadas nos pontos correspondentes aos limites superiores dos intervalos de classe.

Exemplo: Um exercício completo

Análise do desempenho dos alunos do Professor Paulo.

1ª etapa: Levantamento dos dados brutos (Tabela primitiva)

Notas dos alunos do professor Paulo											
5	7	7	2	0	0	3	9	8	4	8	4
1	7	9	6	7	7	1	4	0	2	1	1
3	9	7	5	6	4	9	8	6	5	4	0
8	9	3	2	9	6	8	7	4	5	4	8
3	2	8	8	0	5	3	5	1	5	9	0
9	9	3	9	8	8	7	5	8	7	0	2
7	7	1	7	7	1	7	0	6	3	2	0
2	7	8	6	2	1	6	7	4	6	9	6
5	1	7	9	2	5	9	1	8	5	2	8
7	3	0	7	8	8	6	9	7	4	8	3
5	2	5	1	8	8	8	7	4	0	3	6
2	9	8	4	8	5	8	6	5	8	6	4
2	1	1	0	3	9	0	3	8	1	2	9
1	7	4	9	0	3	8	1	2	9	7	7

2ª etapa: Construção de rol

Notas dos alunos do professor Paulo											
0	1	1	2	3	4	5	6	7	8	8	9
0	1	1	2	3	5	5	7	7	8	8	9
0	1	2	2	4	5	6	7	7	8	8	9
0	1	2	3	4	5	6	7	7	8	8	9
0	1	2	3	4	5	6	7	7	8	8	9
0	1	2	3	4	5	6	7	7	8	8	9
0	1	2	3	4	5	6	7	7	8	8	9
0	1	2	3	4	5	6	7	7	8	8	9
0	1	2	3	4	5	6	7	7	8	8	9
0	1	2	3	4	5	6	7	7	8	8	9
0	1	2	3	4	5	6	7	7	8	8	9
0	1	2	3	4	5	6	7	7	8	8	9
0	1	2	3	4	5	6	7	7	8	8	9
0	1	2	3	4	5	6	7	7	8	8	9
0	1	2	3	4	5	6	7	7	8	8	9
0	1	2	3	4	5	6	7	7	8	8	9
0	1	2	3	4	5	6	7	7	8	8	9
0	1	2	3	4	5	6	7	7	8	8	9
0	1	2	3	4	5	6	7	7	8	8	9

3ª etapa: Construção da tabela de frequência

Notas dos alunos do professor Paulo	
Notas	Freqüência
0	14
1	16
2	15
3	13
4	13
5	15
6	13
7	24
8	26
9	19
10	0
Total 168	

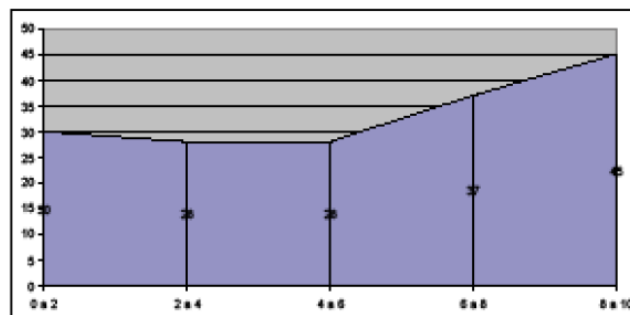
4ª etapa: Construção da tabela de frequência com intervalos de classe

$$\text{amplitude total} = \text{nota maior} - \text{nota menor} = 9 - 0 = 9$$

$$\text{Intervalo de classes} = \frac{\text{amplitude total}}{\text{N}^\circ \text{ de classes}} = \frac{9}{5} = 1,8 = 2$$

Notas dos alunos do professor Paulo	
Notas	Freqüência
0 a 2	30
2 a 4	28
4 a 6	28
6 a 8	37
8 a 10	45
Total 168	

5ª etapa: Representação gráfica (polígono de frequência)

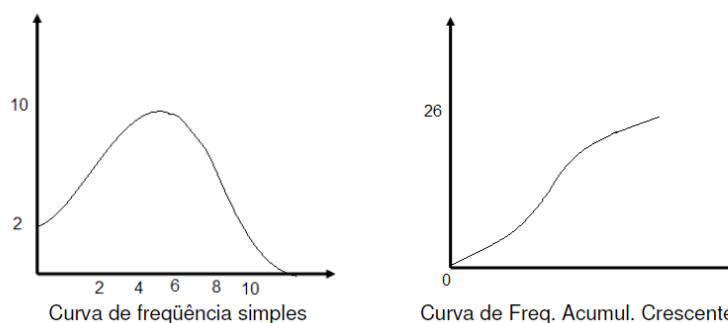
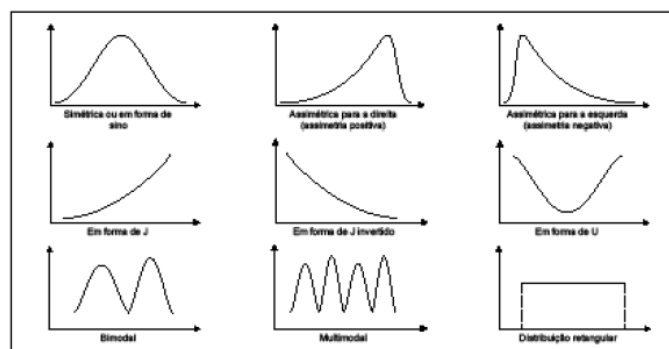


Obs.: Concluindo o estudo, o polígono de frequência parece demonstrar que o resultado do trabalho do professor Paulo é satisfatório, pois há mais alunos acima do intervalo 4 a 6 do que abaixo dele.

8- Curva de Frequência

8.1- A Curva de Frequência. Curva Polida

A tendência da análise de populações cada vez mais amplas é de que a linha poligonal se torne uma curva. Essa curva recebe a denominação de **curva de frequência**. Enquanto o polígono de frequência apresenta a imagem real do fenômeno estudado, a curva de frequência exibe a imagem tendencial.



Após o traçado de um polígono de frequência, é desejável, muitas vezes, que se lhe faça um **polimento**, de modo a mostrar o que seria tal polígono com um número maior de dados. Esse procedimento não mostra uma certeza absoluta de que a curva obtida – **curva polida** – seja tal qual a curva resultante de um grande número de dados. No entanto, pode-se afirmar que ela assemelha-se mais à curva de frequência do que o polígono de frequência obtido de uma amostra limitada.

O polimento, geometricamente, corresponde à eliminação dos vértices da linha poligonal. Consegue-se isso com o emprego de uma fórmula bastante simples, a qual, a partir das frequências reais, nos fornece novas frequências – **frequências calculadas** –

que se localizarão, como no polígono de frequências, nos pontos médios. A fórmula para se determinar a frequência calculada (fc_i) é

$$fc_i = \frac{f_i + 2f_{i-1} + f_{i+1}}{4}$$

fc_i é a frequência calculada da classe considerada.

f_i é a frequência simples da classe considerada.

f_{i-1} é a frequência simples da classe anterior à classe considerada.

f_{i+1} é a frequência simples da classe posterior à classe considerada.

Exemplo:

i	ESTATURAS (cm)	f_i
1	150 - 154	4
2	154 - 158	9
3	158 - 162	11
4	162 - 166	8
5	166 - 170	5
6	170 - 174	3
		$\Sigma f_i = 40$

$$fc_1 = \frac{0 + 2 \times 4 + 9}{4} = \frac{17}{4} = 4,25$$

$$fc_2 = \frac{4 + 2 \times 9 + 11}{4} = \frac{33}{4} = 8,25$$

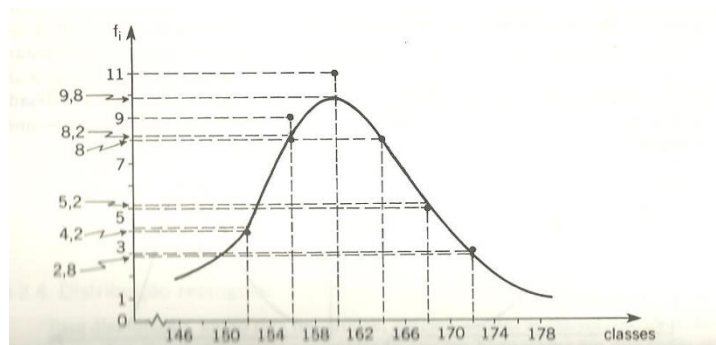
$$fc_3 = \frac{9 + 2 \times 11 + 8}{4} = \frac{39}{4} = 9,75$$

$$fc_4 = \frac{11 + 2 \times 8 + 5}{4} = \frac{32}{4} = 8$$

$$fc_5 = \frac{8 + 2 \times 5 + 3}{4} = \frac{21}{4} = 5,25$$

$$fc_6 = \frac{5 + 2 \times 3 + 0}{4} = \frac{11}{4} = 2,75$$

i	ESTATURAS (cm)	f_i	fc_i
1	150 † 154	4	4,2
2	154 † 158	9	8,2
3	158 † 162	11	9,8
4	162 † 166	8	8,0
5	166 † 170	5	5,2
6	170 † 174	3	2,8
		$\Sigma = 40$	



8.2- As Formas das Curvas de Frequência

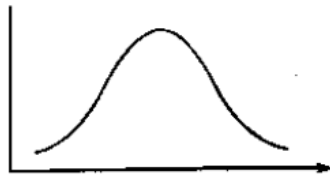
As curvas de frequência assumem as seguintes formas características.

8.2.1- Curvas em Forma de Sino

As **curvas em forma de sino** caracterizam-se pelo fato de apresentarem um valor máximo na região central.

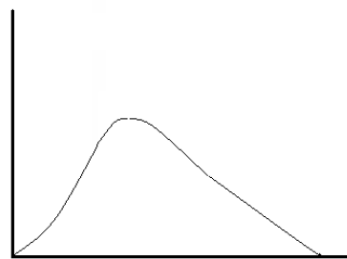
São muitos os fenômenos que oferecem distribuições em forma de sino: a estatura de adultos, o peso de adultos, a inteligência medida em testes mentais, os preços relativos. Tais curvas podem ser simétricas ou assimétricas.

- Curva Simétrica – Essa curva caracteriza-se por apresentar valor máximo no ponto central e os pontos equidistantes desse ponto terem a mesma frequência.

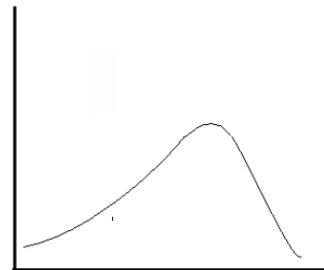


Curva em Sino Simétrica

- Curva Assimétrica – Na prática, não são encontradas distribuições perfeitamente simétricas. As distribuições obtidas de medições reais são mais ou menos assimétricas, em relação à frequência máxima. Assim, as curvas correspondentes a tais distribuições apresentam a **cauda** de um lado da ordenada máxima mais longa do que o outro. Se a cauda mais alongada fica à direita, a curva é chamada **assimétrica positiva** ou **enviesada à direita**. Se a curva se alonga à esquerda, a curva é chamada **assimétrica negativa** ou **enviesada à esquerda**.



Curva assimétrica positiva

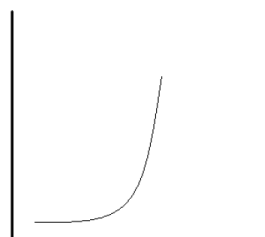


Curva assimétrica negativa

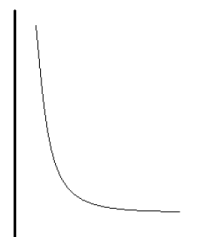
8.2.2- Curvas em Forma de Jota

As curvas em forma de jota são relativas a distribuições extremamente assimétricas, caracterizadas por apresentarem o ponto de ordenada máxima em uma das extremidades.

São curvas comuns aos fenômenos econômicos e financeiros: distribuição de vencimentos ou rendas pessoais.

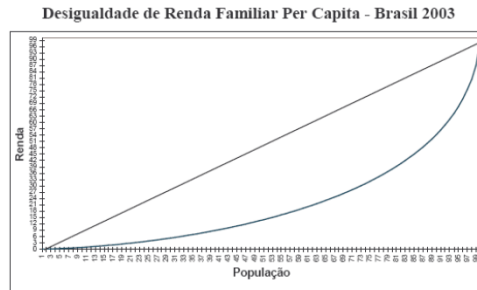


Curva em Jota



Curva em Jota Invertido

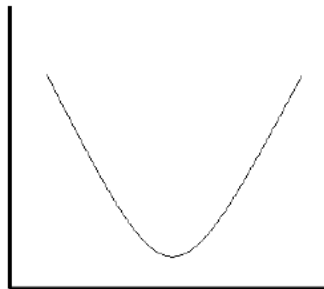
Obs.: Um importante exemplo das curvas em J, na área econômica, é a Curva de Lorenz, que é um gráfico utilizado para representar distribuição de rendas ou de riquezas, etc. A seguir um exemplo.



8.2.3- Curvas em Forma de U

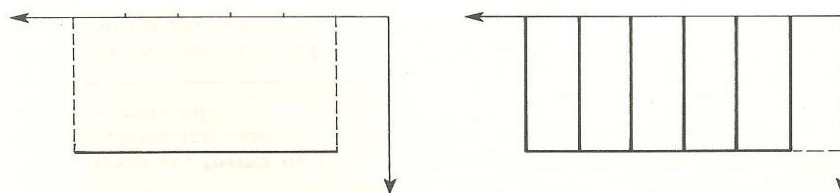
As curvas em forma de U são caracterizadas por apresentarem ordenadas máximas em ambas as extremidades.

Como exemplo de distribuição que dá origem a esse tipo de curva pode-se citar a de mortalidade por idade.



8.2.4- Distribuição Retangular

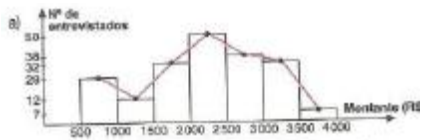
Essa distribuição, muito rara na verdade, apresenta todas as classes com a mesma frequência. Tal distribuição seria representada por um histograma em que todas as colunas teriam a mesma altura ou por um polígono de frequência reduzido a um segmento de reta horizontal.



Exemplo:

A tabela abaixo mostra o montante de pagamentos efetuados em um banco durante um dia.

Montante (R\$)	f
500 -- 1000	28
1000 -- 1500	12
1500 -- 2000	32
2000 -- 2500	50
2500 -- 3000	38
3000 -- 3500	32
3500 -- 4000	7

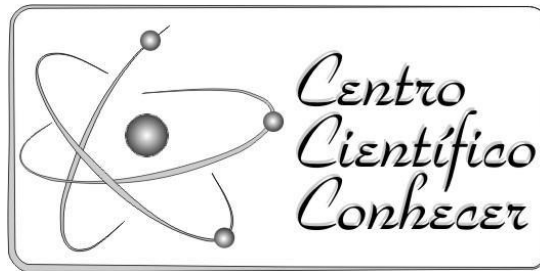


b)

Fr. acum. (fa)	Fr. acum. relativa (far)%
28	14,07
40	20,10
72	36,18
122	61,31
180	80,40
192	96,48
199	100

c) 72

d) ~ 96,48%



Avaliação

Esta avaliação corresponde a 100% da nota do primeiro módulo.

Cursista: _____

1ª Questão

Assinale a alternativa que expressa corretamente o que é população para a Estatística.

- A) População é a quantidade de habitantes de um país.
- B) População é uma parte da amostra estatística.
- C) População é uma reunião de pessoas populares.
- D) População é um conjunto de elementos portadores de, pelo menos, uma característica comum.
- E) População é o conjunto de elementos que responderão a um questionário para fornecer informações relevantes para uma pesquisa.

2ª Questão

Assinale a alternativa que mostra a definição correta de amostra.

- A) Amostra é o conjunto de dados obtidos em uma pesquisa.
- B) Amostra é um brinde que ganhamos em um evento.
- C) Amostra é o conjunto dos resultados obtidos com uma pesquisa.
- D) Amostra é uma tabela de dados obtida com uma pesquisa.
- E) Amostra é um subconjunto finito de uma população

3ª Questão

Dentro do universo estatístico "alunos de uma escola de ensino fundamental", as variáveis "cor dos cabelos" e "número de irmão" podem ser classificadas, respectivamente, como

- A) 1ª. variável quantitativa discreta; 2ª. variável qualitativa.
- B) 1ª. variável quantitativa contínua; 2ª. variável qualitativa.
- C) 1ª. variável qualitativa; 2ª. variável quantitativa contínua.
- D) 1ª. variável qualitativa; 2ª. variável quantitativa discreta.
- E) 1ª. variável quantitativa discreta; 2ª. variável qualitativa contínua.

4ª Questão

Escreva um conceito para Estatística.

5ª Questão

Para os tipos de variáveis, indicar se são discretas (D) ou contínuas (C).

- a) () Peso do conteúdo de um pacote de feijão.
- b) () Diâmetro de um rolamento.
- c) () Número de artigos defeituosos produzidos por uma determinada máquina.
- d) () Número de indivíduos de uma determinada região que recebem seguro desemprego.
- e) () Número de pessoas de uma residência.
- f) () Número médio diário de clientes potenciais visitados por um vendedor na última semana.
- g) () A altura (h) de um indivíduo.
- h) () Temperaturas registradas a cada hora.
- i) () Vida média de lâmpadas.
- j) () Comprimento de 100 parafusos.
- k) () O número de livros de uma biblioteca.

6ª Questão

Seja x a variável “número de cômodos das casas ocupadas por vinte famílias entrevistadas”.

i	x_i	f_i
1	2	4
2	3	7
3	4	5
4	5	2
5	6	1
6	7	1
		$\Sigma = 20$

Complete a tabela com os vários tipos de frequência.

i	x_i	f_i	fr_i	F_i	Fr_i
1	2	4			
2	3	7			
3	4	5			
4	5	2			
5	6	1			
6	7	1			
		$\Sigma = 20$	$\Sigma = 1,00$		

7ª Questão

Assinale a alternativa Correta. As Séries Estatísticas classificam-se em

- A) Geográficas, Históricas e Específicas.
- B) Locais, Históricas e Específicas.
- C) Mundiais, Históricas e Singulares.
- D) Geográficas, Históricas e Classificadas.
- E) Geográficas, Históricas e Singulares.

8ª Questão

Complete o quadro abaixo, classificando as variáveis.

UNIVERSO	VARIÁVEL
Alunos de uma escola	Cor de Cabelos: Variável qualitativa
Casais residentes em uma cidade	Número de filhos: Variável quantitativa discreta
As jogadas de um dado	O ponto obtido em cada jogada:
Peças produzidas por certa máquina	Número de peças produzidas por hora:
Peças produzidas por certa máquina	Diâmetro externo:
Casais residentes em uma cidade	Sexo dos filhos:

9ª Questão

Na construção de um gráfico de setores, para uma população de 500 pessoas, o evento que tiver uma frequência igual a 200, será representado por um setor de quantos graus?

- A)72 B)200 C)120 D)144 E)80

10ª Questão

Dado o rol de medidas das alturas (dadas em cm) de uma amostra de 100 indivíduos de uma faculdade,

151	152	154	155	158	159	159	160	161	161
161	162	163	163	163	164	165	165	165	166
166	166	166	167	167	167	167	167	168	168
168	168	168	168	168	168	168	168	169	169
169	169	169	169	169	170	170	170	170	170
170	170	171	171	171	171	172	172	172	173
173	173	174	174	174	175	175	175	175	176
176	176	176	177	177	177	177	178	178	178
179	179	180	180	180	180	181	181	181	182
182	182	183	184	185	186	187	188	190	190

Calcule.

- a) a amplitude amostral.
- b) o número de classes.
- c) a amplitude de classes.
- d) os limites de classes.
- e) as frequências absolutas da classes.
- f) as frequências relativas.
- g) os pontos médios da classes.
- h) as frequências acumuladas.