

PROBABILIDADE E ESTATÍSTICA I

Aula 1

Prof. Regina Meyer Branski

Conteúdo

1. **Introdução a Estatística e Probabilidade**
2. **Probabilidade e Distribuições de Probabilidade**
 - ▣ Probabilidade Condicional, Regra de Multiplicação e Regra da Adição
 - ▣ Distribuições de Probabilidade Discretas
 - ▣ Distribuições de Probabilidades Normais
3. **Estatística Descritiva**
 - ▣ Distribuição de Frequência e Gráficos
 - ▣ Medidas de Tendência Central, Variação e Posição
4. **Inferência Estatística:**
 - ▣ Intervalo de Confiança
 - ▣ Teste de Hipótese com uma amostra
 - ▣ Teste de Hipótese com duas amostras
 - ▣ Regressão Linear e Correlação
 - ▣ Testes não paramétricos

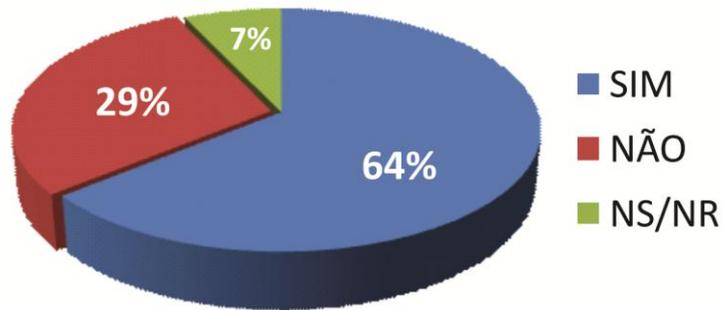
Visão Geral de Estatística e Probabilidade

- Objetivos
 - ▣ Definir estatística
 - ▣ Diferenciar População e Amostra
 - ▣ Diferenciar Parâmetros e Dados Estatísticos
 - ▣ Apresentar Estatística Descritiva, Probabilidade e Estatística Inferencial

O que são dados?

Resultado das observações, contagens, medições ou respostas.

VOCÊ ACHA QUE O USO DE CELULAR NA ESCOLA PODE SER ÚTIL PARA SUA APRENDIZAGEM?



EM QUAIS DISCIPLINAS VOCÊ ACHA QUE É POSSÍVEL USAR O CELULAR?

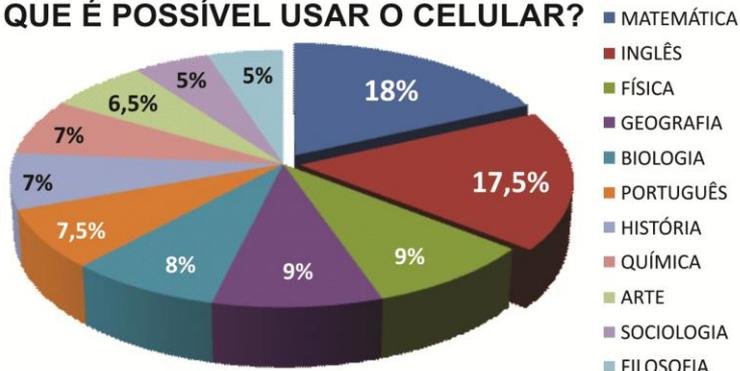
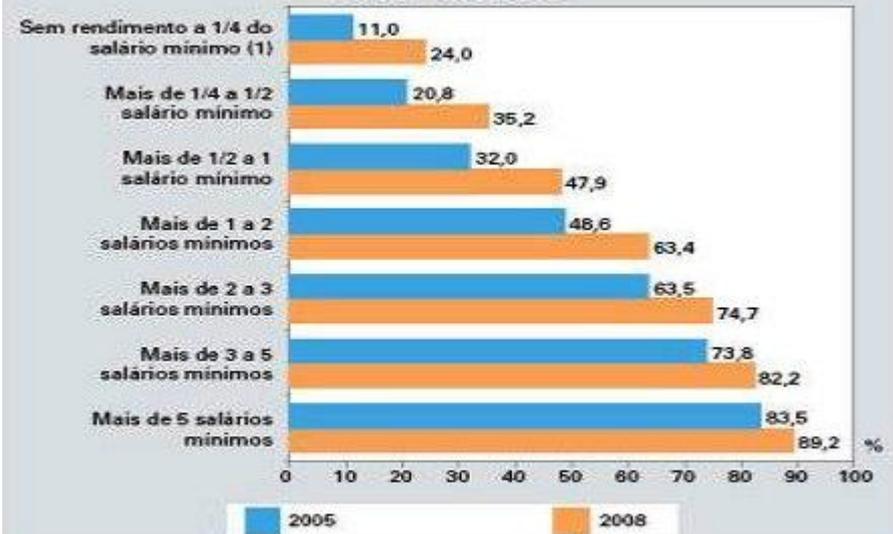


Gráfico 19 - Percentual das pessoas que tinham telefone móvel celular para uso pessoal na população de 10 anos ou mais de idade, segundo as classes de rendimento mensal domiciliar *per capita* Brasil - 2005/2008



Fonte: IBGE, Diretoria de Pesquisas, Coordenação de Trabalho e Rendimento, Pesquisa Nacional por Amostra de Domicílios 2005/2008.
Nota: Exclusivo as pessoas cuja condição na unidade domiciliar era pensionista, empregado doméstico ou parente do empregado doméstico.
(1) Inclusive as pessoas moradoras em unidades domiciliares cujos componentes recebiam somente em benefícios.

A partir dos dados posso fazer inferências!!

- “Pessoas que comem grãos integrais nas três refeições diárias têm risco de derrame reduzido em 37%.” (Fonte: *Whole Grains Council*.)
- “Das 1.500 lesões na coluna vertebral em menores nos EUA, 70% são resultado de acidentes em veículos e 68% dos acidentados não usavam cinto de segurança.” (Fonte: *UPI*.)

Cuidado com os dados!!!

REPORTAGEM – Celular: de vilão a mocinho

PESQUISA REVELA O QUE PENSAM OS ALUNOS DO MECCA SOBRE O USO DE CELULAR NA ESCOLA

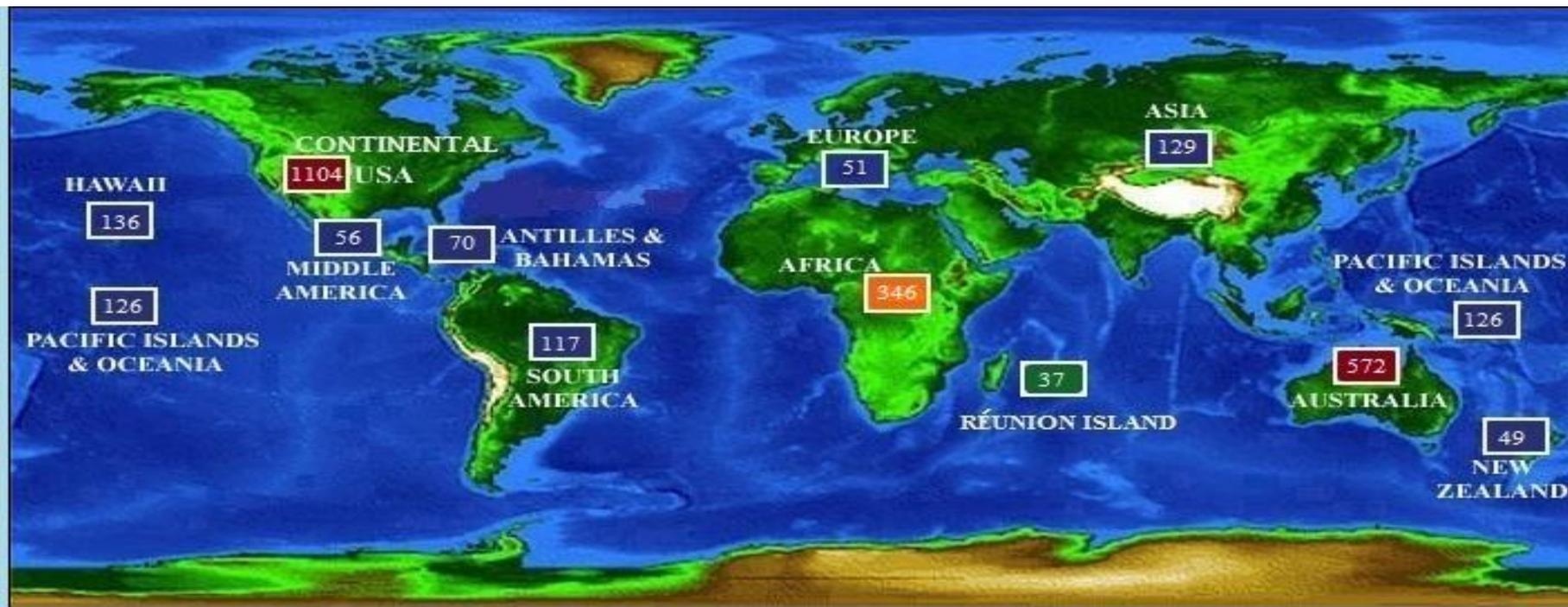
Bruna Alcântara e Jaíne Rodrigues



Nos dias de hoje, encontrar um adolescente que não tenha celular virou raridade. Em uma pesquisa realizada pelo Jornal MeccAtitude com 42 alunos do período noturno (1ª, 2ª e 3ª séries do Ensino Médio), revelou que apenas 10% não possuem aparelho celular.

De todos os 90% que possuem celular, apenas 12% não têm acesso à internet. Porém, dos 88% dos alunos que possuem internet, a maioria não usa o celular com maior prioridade para estudos e trabalhos escolares, e sim para redes sociais e músicas (confira gráficos ao lado).

Cuidado com os dados!!!



Territory	Total Attacks	Fatal Attacks	Last Fatality	Territory	Total Attacks	Fatal Attacks	Last Fatality
Continental USA	1104	35	2012	Antilles & Bahamas	70	16	2013
Australia	572	153	2014	Middle America	56	27	2011
Africa	346	94	2014	New Zealand	49	9	2013
Asia	129	48	2000	Europe	51	27	1974
Hawaii	136	9	2013	Réunion Island	37	17	2013
Pacific/Oceania Islands (w/out Hawaii)	126	49	2011	Unspecified/Open Ocean	21	7	1995
South America	117	26	2013				
				WORLD	2,777	497	2014

500 and above

200 - 499

40 - 199

1 - 39

0

Os dados!

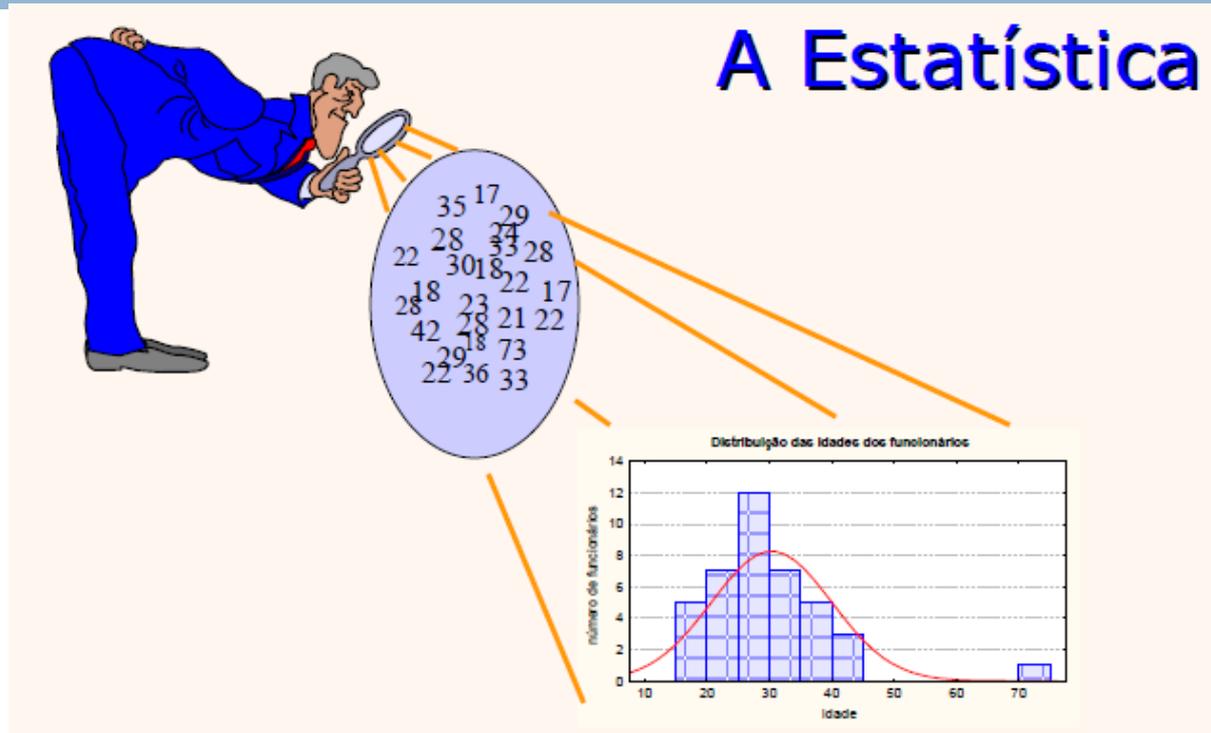
The screenshot shows the IBGE website interface. At the top, there's a navigation bar with tabs for **Indicadores**, **População**, **Economia**, **Geociências**, **Canais**, **Download**, **Pesquisas**, and **Sala de Imprensa**. Below this, there are several columns of links and content:

- Indicadores Sociais:** Intermediação Política, Crianças e Adolescentes, Educação e Trabalho, Mortalidade Infantil, Indicadores Sociais Mínimos, Mercado de Trabalho, Mobilidade Social, População Jovem, Síntese de Indicadores Sociais, Indicadores Culturais, Indicadores Sociodemográficos e de Saúde.
- População:** Censos Demográficos (Censo 2010, Censo 2000), Contagem da População, População Indígena, Estatísticas do Registro Civil, PNAD (Pesquisa Nacional por Amostra de Domicílios, Suplementos, Pesquisas especiais).
- Economia:** Pesquisa de Orçamentos Familiares (2008-2009, 2002-2003), Pesquisa Nacional de Saúde (Percepção do Estado de Saúde, Estilos de Vida e Doenças Crônicas).
- Geociências:** Tábua Abreviadas de Mortalidade, Tábua Completas de Mortalidade, Projeção da População, Estimativas de População, Atlas de Saneamento, Economia Informal Urbana, Assistência Médico-Sanitária, Saneamento Básico, Sindicatos: Indicadores Sociais, Pesquisa de Esporte, Reflexões sobre os Deslocamentos Populacionais no Brasil, Características Étnico-raciais da População.
- Pesquisas:** PeNSE, Pesquisa Nacional de Saúde do Escolar, Avaliação do estado nutricional dos escolares do 9º ano do ensino fundamental.

Below the navigation bar, there are several news and information boxes:

- Acesso à Informação:** IBGE divulgará, em atendimento ao FPE, renda domiciliar per capita 2014. Estabelecido em 25 de fevereiro de 2015 o término do período de transição para o SIRGAS2000. Divulgação da PIMES.
- Transparência Pública:** IPCA e INPC - São Paulo - Transporte, Divulgação da PNAD Continua em 2015, IBGE divulga a Área Territorial Brasileira, IPCA e INPC - Subitem Multa.
- Transparência CPGF:** Destaque anterior.
- Indicadores:** IPCA, INPC, PME, PMC, PIB, SINAPI, PIM, IPP, PMS. PMC (+) ver pesquisa. Gráfico: Variação (%) do Volume de vendas do comércio varejista em relação a igual mês do ano anterior. X-axis: Jul, Ago, Set, Out, Nov, Dez. Y-axis: -1, 0, 1, 2, 3.
- Ciências Estatísticas:** Social media icons for Facebook, Twitter, and YouTube.
- BRASIL EM SÍNTESE:** A graphic with a map of Brazil and a circular arrow.

Estatística



Conjunto de técnicas para coletar, organizar, descrever e interpretar dados para apoio à tomada de decisões

Alguns Conceitos...



Conjunto de Dados

▣ População



▣ Amostra

População e Amostra

- População
 - ▣ Conjunto completo de indivíduos ou objetos de interesse
 - Todos os alunos que já cursaram faculdade no Brasil
 - Todas as sequoias da Califórnia
- Amostra
 - ▣ Conjunto de indivíduos ou objetos selecionados a partir da população
 - Conjunto das 25 alturas coletadas de 25 alunos da escola A

População e Amostra

- ▣ Interesse em identificar **variáveis** específicas
 - idade, cor do cabelo, altura e peso.
- ▣ Valor de Dados: valor para cada variável associada a um elemento da população ou amostra
 - pode ser um número, palavra ou símbolo.
- ▣ Quando João entrou na faculdade tinha 23 anos, cabelo castanho, media 1,80 m e pesava 85 kg.
 - Valores para as quatro variáveis aplicadas a João

Identificar conjunto de dados

- Em uma pesquisa recente nos EUA, 1.708 adultos foram questionados se achavam que o aquecimento global era um problema que requeria ação governamental imediata. 939 adultos disseram sim. Identifique a população e a amostra. Descreva o conjunto de dados. (*Adaptado de: Pew Research Center.*)

Identificar conjunto de dados

- População:
- Amostra:
- Conjunto de dados:

Identificar conjunto de dados

- População: Todos os adultos dos EUA
- Amostra: 1708 adultos
- Conjunto de dados: 939 sim e 769 não



Exercícios

- Identifique População e Amostras
 - Um estudo com 33.043 crianças na Itália foi conduzido para encontrar a ligação entre anormalidades no ritmo cardíaco e síndrome da morte súbita.
 - Uma pesquisa com 1 000 usuários de computador descobriu que 17% planejam comprar o sistema operacional Windows.
 - Pesquisa com 496 estudantes de uma faculdade descobriu que 10% planejam viajar para fora do país durante as férias
 - Uma pesquisa com 546 mulheres descobriu que mais de 56% são o investidor primário em suas residências.

Parâmetros e Estatística

□ Parâmetro

- Número que descreve uma característica **da população**
 - *Idade média da população dos Estados Unidos*

□ Estatística

- Número que descreve uma característica **da amostra**
 - *Idade média da população de três estados*

Identificar Parâmetro e Estatística

- Decida se o valor numérico descreve um parâmetro populacional ou uma estatística amostral.

Uma pesquisa recente de uma amostra de MBAs revelou que o salário médio de um MBA é superior a \$ 82.000 a.a. (*Fonte: The Wall Street Journal.*)

Identificar Parâmetro e Estatística

Uma pesquisa recente de uma amostragem de MBAs revelou que o salário médio de um MBA é superior a \$ 82.000. (*Fonte: The Wall Street Journal.*)

Estatística amostral (a média de \$ 82.000 é baseado em uma parte da população).

Identificar Parâmetro e Estatística

O salário inicial dos 667 MBAs da *University of Chicago Graduate School of Business* aumentou 8,5% desde o ano passado

Identificar Parâmetro e Estatística

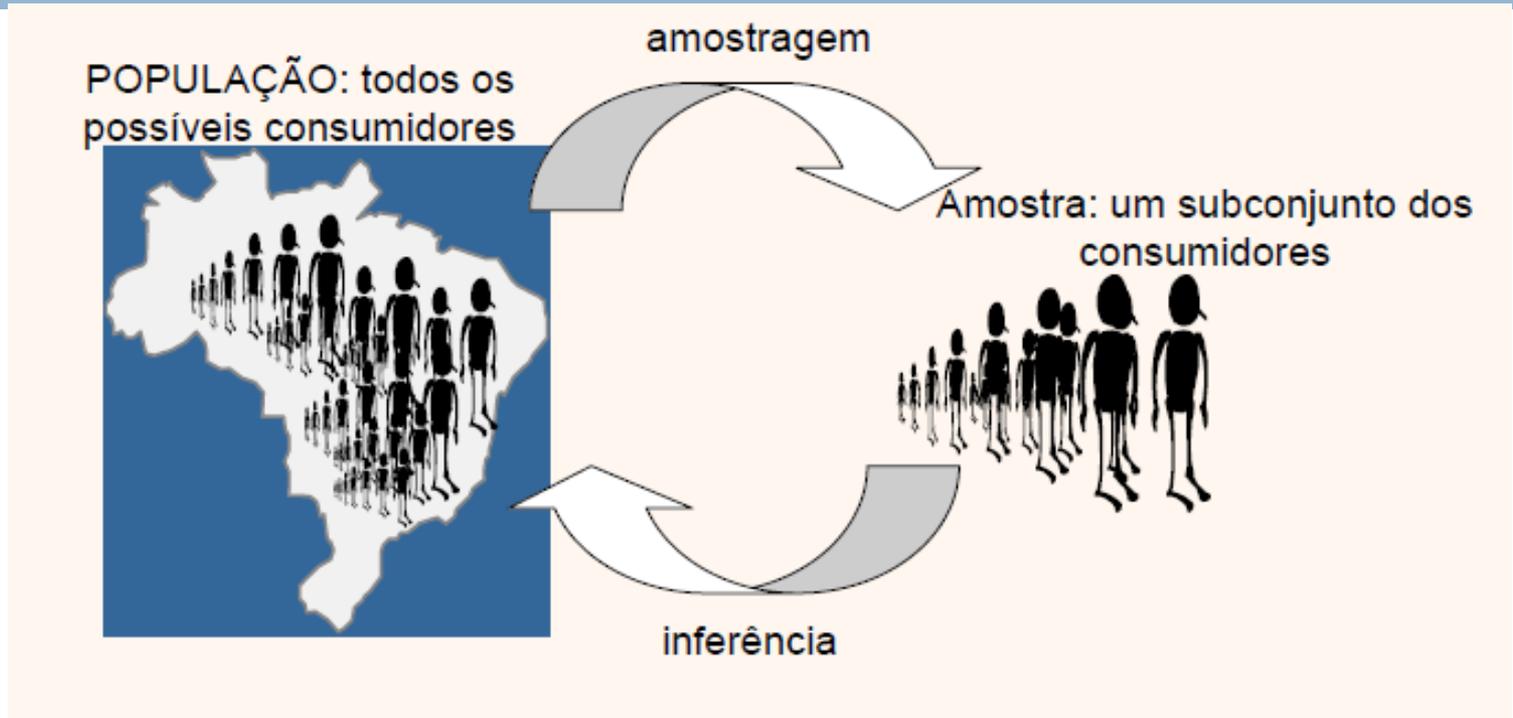
O salário inicial dos 667 MBAs da *University of Chicago Graduate School of Business* aumentou 8,5% desde o ano passado

Parâmetro populacional (porcentagem de 8,5% é baseada nos salários iniciais de todos os 667 graduados).

Exercícios

- Distinguir Parâmetro e Estatística
 - Média dos salários anuais para 35 dos 1200 contadores de uma empresa é \$ 68.000
 - Sessenta e dois dos 97 passageiros a bordo da aeronave sobreviveram à sua explosão.
 - Categoria de interesse para 12% de todas as revistas foi esportes
 - Pesquisa com 1503 adultos nos EUA, 53% disseram que usam tanto a linha fixa quanto o telefone celular

Estatística

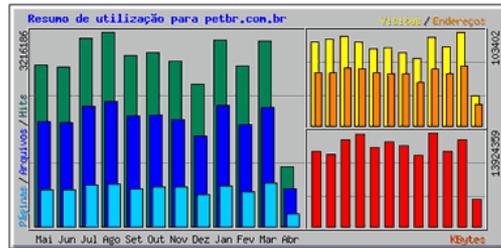


Essência da análise estatística é tirar conclusões sobre uma população a partir de uma amostra

Ramos da Estatística

Estatística Descritiva

Envolve informações, números e gráficos de forma resumida



Mês	Média diária				Total Mensal					
	Hits	Arquivos	Páginas	Visitas	Endereços	KBytes	Visitas	Páginas	Arquivos	Hits
Abr 2007	89853	56865	19098	3063	23434	4046054	33699	210087	625518	988393
Mar 2007	99102	63822	23112	3335	65687	12869748	103402	716476	1978508	3072184
Fev 2007	94530	60398	20393	3114	57549	11083680	87200	571008	1691169	2646863
Jan 2007	99216	64522	21731	3163	62513	13924359	98079	673690	2000200	3075710
Dez 2006	75709	48482	17189	2412	47670	10545345	74793	532867	1502961	2346994
Nov 2006	91047	58659	21738	2691	57379	12063974	80745	652157	1759785	2731428
Out 2006	92668	59781	21106	2786	57687	12593577	86392	654296	1853229	2872732
Sep 2006	94363	61149	20842	2846	58156	11779954	85384	625271	1834485	2830897
Ago 2006	103747	66770	22840	2989	63111	13691140	92660	708067	2069874	3216186
Jul 2006	100306	64277	22447	3198	64321	12859668	99162	695875	1992594	3109486
Jun 2006	88076	57336	20422	3172	58046	10777765	95170	612662	1720095	2642292
Mai 2006	86031	55837	19705	2985	58207	11183023	92548	610871	1730961	2666989
Totais					137418287	1029234	7263326	20759379	32200154	

Estatística Inferencial

Envolve tirar conclusões sobre a população a partir da amostra



E a Probabilidade?

- Probabilidade é o veículo da estatística



Probabilidade

Estatística



E a Probabilidade?

- Probabilidade é o veículo da estatística
- Probabilidade: indaga a chance de ocorrer algo específico se você conhece a população
- Estatística: extrai uma amostra, descreve a amostra e faz inferências sobre a população
- Em um conjunto com 20 fichas verdes, 20 vermelhas e 20 azuis
 - ▣ Probabilidade: se eu tirar uma ficha, qual a chance dela ser verde?
 - ▣ Estatística: sem conhecer o conjunto de fichas, extrai uma amostra e com base nela faço conjecturas sobre o conjunto

Estatística Descritiva X Estatística Inferencial

- Uma amostra de homens de 48 anos foi estudada por 18 anos. Para homens não casados, aproximadamente 70% estavam vivos aos 65 anos. Para os casados, 90% estavam vivos aos 65 anos. *(Fonte: The Journal of Family Issues)*

Estatística Descritiva X Estatística Inferencial

Estatística descritiva envolve afirmações como “Para homens não casados, aproximadamente 70% estavam vivos aos 65 anos” e “Para homens casados, 90% estavam vivos aos 65 anos”.

Estatística Descritiva X Estatística Inferencial

Estatística descritiva envolve afirmações como “Para homens não casados, aproximadamente 70% estavam vivos aos 65 anos” e “Para homens casados, 90% estavam vivos aos 65 anos”.

Uma inferência possível tirada desse estudo é que ser casado está associado com uma vida mais longa para homens.

Exercícios

- Estudo recente, voluntários que dormiram 8 horas em uma noite eram três vezes mais capazes de responder corretamente às questões de um teste de matemática em relação àqueles que não tiveram horas de sono suficientes.
- a) identifique a amostra
- b) qual a população
- c) que parte representa o ramo descritivo da estatística
- d) faça uma inferência com base nos resultados

Vimos até agora!

- Definir Estatística
- Diferenciar População e Amostra
- Diferenciar Parâmetros e Dados Estatísticos
- Diferenciar Estatística Descritiva e Inferencial

Classificação dos Dados

- Dados Qualitativas
- Dados Quantitativas

Dados Qualitativas

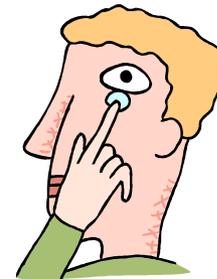
- Consistem em atributos, qualidades ou entradas não numéricas



Graduação



Local de
Nascimento



Cor dos Olhos

Dados Quantitativas

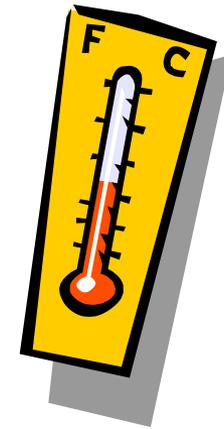
- Medidas ou contagens numéricas.



Idade

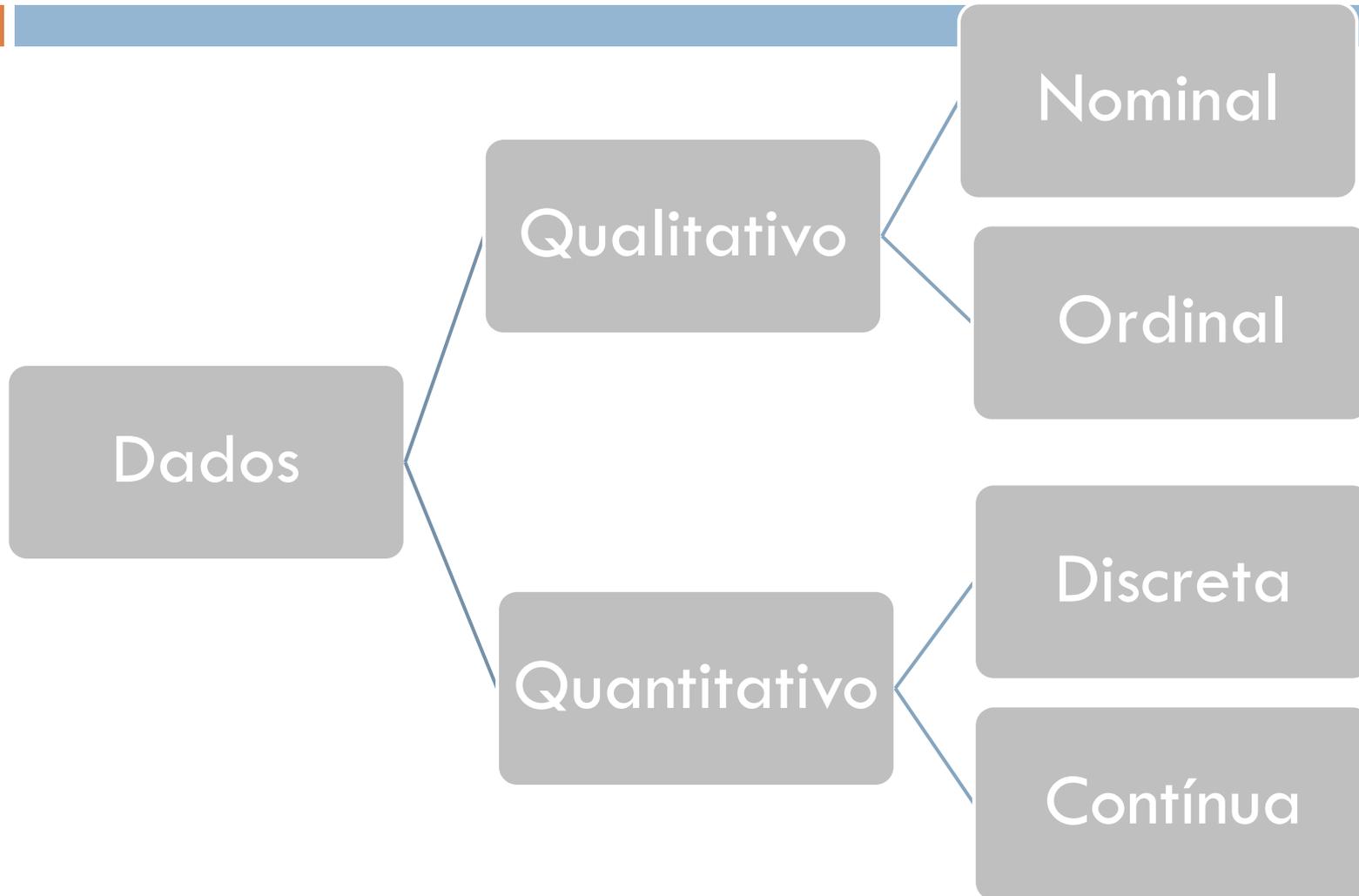


Peso



Temperatura

Tipos de datos



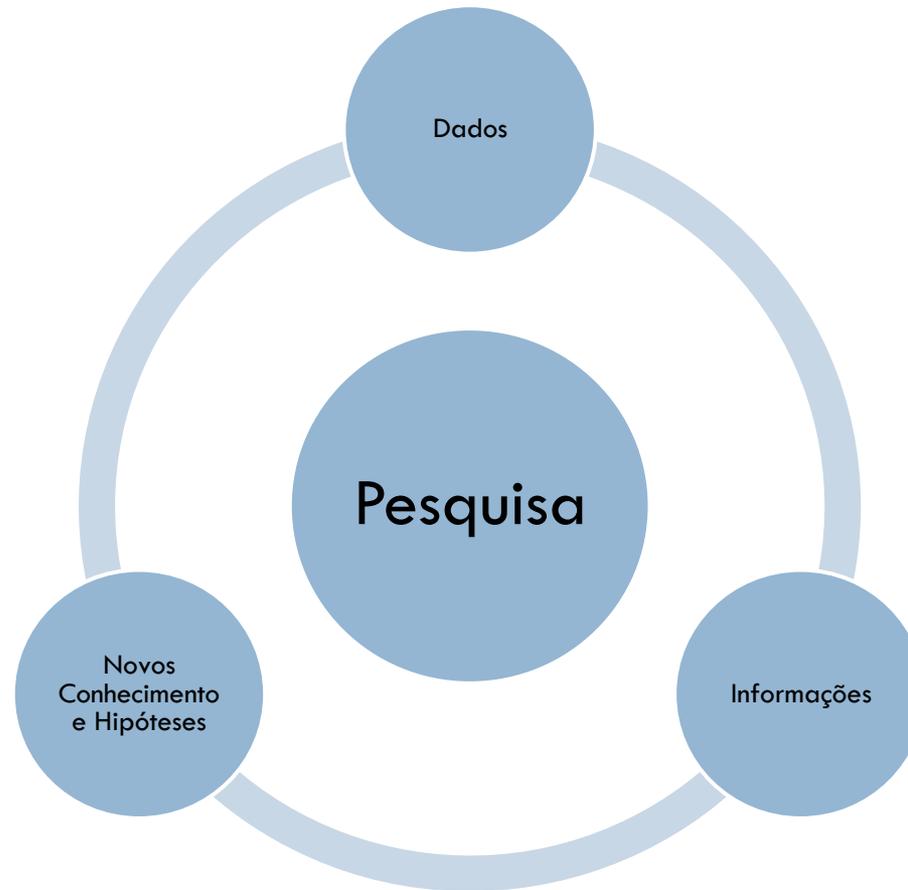
Tipos de Dados

- Amostra de 4 clientes de salão de beleza
- Coletada cor de cabelo, cidade de origem e nível de satisfação com o serviço executado pelo salão.
- Variáveis qualitativas (atributos)
 - ▣ Loiro, castanho, preto, castanho
 - ▣ Guarujá, São Vicente, Santos e São Paulo
 - ▣ Muito satisfeito, satisfeito, pouco satisfeito, insatisfeito

Tipos de Dados

- Número de ligações que um vendedor faz em um único dia
 - ▣ 0, 1, 2, 3.....(pode ser representado como pontos na reta numérica)
- Notas dos juízes em uma competição de patinação
 - ▣ 9,9 - 8,8 - 8,7 - 9,8
- Tempo diário em horas que vendedor passa fazendo ligações
 - ▣ Pode ser qualquer número entre 0 e 24

Estatística e Pesquisa



Planejamento

- Planejar Estudo Estatístico
- Apresentar técnicas para coleta de dados
- Planejar um experimento
- Discutir técnicas de amostra

Planejamento de Estudo Estatístico

1. Identificar variáveis de interesse e a população do estudo
 - ▣ Tempo de recuperação de pacientes com certa doença;
 - ▣ Renda total das famílias brasileiras
2. Definir sistema de coleta e medição dos dados
 - ▣ Inclui método de amostragem, tamanho da amostra etc.
3. Coletar amostra
 - ▣ Selecionar o que será incluído na amostra e coletar os dados. A amostra deve ser representativa da população.
4. Revisar o processo de amostragem após a coleta

Métodos de Coleta de Dados

- Experimental
- Observacional
- Pesquisa
- Simulação

Métodos de Coleta de Dados

□ Estudo Experimental

- O pesquisador controla ou modifica o ambiente e observa o efeito sobre a variável de estudo
- Experimento no qual diabéticos tomam extrato de canela diariamente, enquanto um grupo de controle não. Depois de 40 dias, os diabéticos que tomaram canela reduziram o risco de doenças do coração, enquanto o grupo de controle não apresentou mudanças

Fonte: Diabetes Care

Métodos de Coleta de Dados

□ Estudo Observacional

- ▣ Pesquisador observa o processo sem interferir em nada
- ▣ Pesquisadores observaram e registraram o comportamento oral de crianças acima de três anos de idade com objetos não comestíveis

Fonte: Pediatric Magazine

Métodos de Coleta de Dados

□ Pesquisa

- ▣ Investigação de uma ou mais características de uma população
- ▣ Estudo conduzido em uma amostra de mulheres médicas para determinar se a razão primária na escolha de suas carreiras é estabilidade financeira

Métodos de Coleta de Dados

□ Simulação

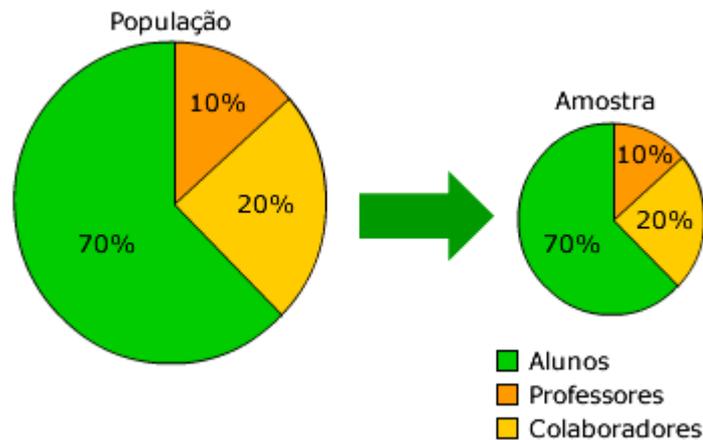
- ▣ Usa modelos matemáticos ou físico para reproduzir as condições de uma situação ou processo
- ▣ Empresas automobilísticas fazem simulações com bonecos para estudar os efeitos de acidentes em pessoas

Métodos de Coleta de Dados

- Estudo do efeito da mudança dos padrões de voo no número de acidentes com aviões
- Um estudo sobre o efeito da aveia para abaixar a pressão sanguínea
- Um estudo sobre como alunos da quarta série montam um quebra-cabeça
- Um estudo dizendo se os residentes nos EUA aprovam seu presidente ou não.

Método de Amostragem

- Processo de seleção de itens ou eventos que serão a amostra.
 - ▣ Devem representar com precisão a população



Métodos de Amostragem

- Como garantir que um método produzirá bons dados?
 - ▣ Amostras tendenciosas
 - Seleccionadas por julgamento, conveniência etc
 - Dados diferem sistematicamente da população

 - ▣ Amostras aleatórias
 - Todos os membros da população tem chances iguais de serem seleccionados
 - Dados representam a população

Amostra Aleatória Simples

- Toda amostra possível do mesmo tamanho tem a mesma chance de ser selecionadas
- Aleatório = Chances Iguais
 - e não ao acaso, sem padrão
- Ideia é que o próximo resultado não é previsível
- Por que criar uma amostra aleatória simples?
 - Para ser imparcial. Procedimentos estatísticos só valem para amostras aleatórias

Como gerar uma Amostra Aleatória Simples?

- Número aleatórios podem ser gerados por uma tabela, calculadoras, softwares.
- Numere toda a população
- Membros da população correspondentes aos números aleatórios gerados farão parte da amostra

Amostra Aleatória Simples

Há 731 estudantes que se inscreveram no curso de estatística em sua faculdade. Você deseja formar uma amostra de oito estudantes para responder às questões de uma pesquisa. Selecione os estudantes que pertencerão à amostra aleatória simples

Amostra Aleatória Simples

1. Dê números de 1 até 731 para cada estudante de estatística
2. Na tabela de números aleatórios, escolha um espaço inicial aleatoriamente (suponha que comece na terceira linha da segunda coluna)

Tabela de Números aleatórios

92630 78240 19267 95457 53497 23894 37708 79862
79445 78735 71549 44843 26104 67318 00701 34986
59654 71966 27386 50004 05358 94031 29281 18544
31524 49587 76612 39789 13537 48086 59483 60680
06348 76938 90379 51392 55887 71015 09209 79157

3. 719 – 662 – 738 – 650 – 004 – 053 – 589 – 403 – 129

Outras Técnicas de Amostragem

- Amostra Estratificada
- Amostra por Agrupamento
- Amostra Sistemática

Outras Técnicas de Amostragem

□ Amostra Estratificada



□ Amostra por Agrupamento



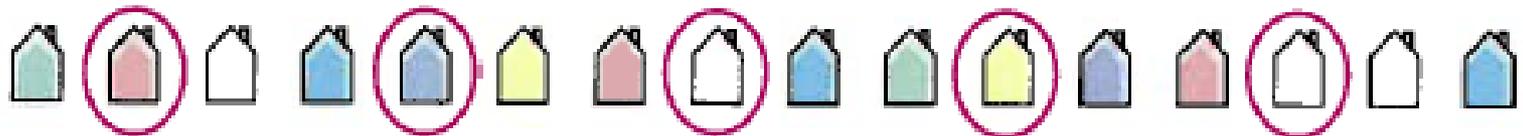
Outras Técnicas de Amostragem

□ Amostra Sistemática



Amostra Sistemática

- Escolha um valor inicial aleatoriamente. Então, separe todo x^o membro da população
- Na cidade de Santos, dar um número diferente para cada casa, escolher um número inicial aleatório e, então, escolher toda 100^o casa.
- Resultados próximos à amostragem aleatória



Amostra Estratificada

- Divida a população em grupos (estrato) e selecione uma amostra aleatória de cada grupo
 - ▣ Para coletar uma amostragem estratificada do número de pessoas que vivem em casas na cidade de Santos, divida as casas em níveis socioeconômicos e selecione casas aleatoriamente em cada nível

Exercício

- Você está realizando um estudo para determinar a opinião dos estudantes em sua escola sobre a pesquisa de células-tronco. Identifique a técnica de amostragem.
 - ▣ Seleciona uma classe aleatoriamente e questiona cada aluno da classe
 - ▣ Divide a população de estudantes com relação às graduações, seleciona aleatoriamente e questiona alguns alunos de cada curso
 - ▣ Designa um número para cada aluno e gera números aleatoriamente. Então, questiona os estudante que corresponde à cada número aleatório.

Visão Geral de Estatística e Probabilidade

- Definir estatística
- Diferenciar População e Amostra
- Diferenciar Parâmetros e Dados Estatísticos
- Apresentar Estatística Descritiva, Probabilidade e Estatística Inferencial
- Planejar um experimento