

../../../../Imagens/barraunesp-assvisual.png

Prof^a Dr^a Mara Lúcia Martins Lopes

Caio da Silva Pereira

Gabriel Rodrigues Munhoz

Heitor Martins da Silva

Leandro Suzuki Barboza dos Santos

Rafael Gomes de Oliveira

Thiago Rissetti Roquette

../../../../Imagens/SA03x.jpg

Estatística e Probabilidade
Trabalho sobre Apresentação de Dados
Parte 1 - Teórica

Ilha Solteira, São Paulo

Setembro de 2017

Prof^a Dr^a Mara Lúcia Martins Lopes

Caio da Silva Pereira

Gabriel Rodrigues Munhoz

Heitor Martins da Silva

Leandro Suzuki Barboza dos Santos

Rafael Gomes de Oliveira

Thiago Rissetti Roquette

Estatística e Probabilidade

Trabalho sobre Apresentação de Dados

Parte 1 - Teórica

../../../../Imagens/SA03x.jpg

"Estatística é método, ciência e arte."

Universidade Estadual Paulista - Unesp

Faculdade de Engenharia de Ilha Solteira - FEIS

Ilha Solteira, São Paulo

Setembro de 2017

Sumário

| | | |
|------------|---|-----------|
| 0.1 | Introdução | 3 |
| 0.2 | Variáveis | 4 |
| 0.2.1 | Variáveis Qualitativas | 4 |
| 0.2.1.1 | Variáveis Qualitativas Nominais | 4 |
| 0.2.1.2 | Variáveis Qualitativas Ordinais | 4 |
| 0.2.2 | Variáveis Quantitativas | 5 |
| 0.2.2.1 | Variáveis Discretas | 5 |
| 0.2.2.2 | Variáveis Contínuas | 5 |
| 0.3 | Distribuição de Frequência | 6 |
| 0.3.1 | Tabela para Variáveis Qualitativas | 6 |
| 0.3.2 | Tabela para Variáveis Quantitativas | 6 |
| 0.3.2.1 | Tabela para Dados de tipo Quantitativo Discreto | 6 |
| 0.3.2.2 | Tabela para Dados de tipo Quantitativo Contínuo | 7 |
| 0.4 | Gráficos | 8 |
| 0.4.1 | Gráficos para Variáveis Qualitativas | 8 |
| 0.4.1.1 | Gráfico de Barras | 8 |
| 0.4.1.2 | Gráfico de Pizza | 10 |
| 0.4.2 | Gráficos para Variáveis Quantitativas | 11 |
| 0.4.2.1 | Histograma | 11 |
| 0.4.2.2 | Polígono de Frequência | 11 |
| 0.4.2.3 | BoxPlot | 12 |
| 0.5 | Conclusão | 14 |
| | REFERÊNCIAS | 15 |

0.1 Introdução

Estatística pode ser pensada como a ciência de aprendizagem a partir de dados. Logo, temos que apresentar esses dados e organizá-los para que o nosso estudo acerca deles se simplifique e seja correto. Podemos dividir a estatística em dois ramos, o primeiro sendo o descritivo que é composto por toda a parte de coleta, organização, descrição e cálculo dos dados; e a segunda que é o ramo indutivo que é formado pela análise e interpretação, associados a uma margem de incerteza. [1]

Existem diferentes tipos de variáveis e para que tal estudo sobre elas ocorra de forma ordenada há diferentes tipos de distribuição de frequência e gráficos para simplificar a visualização de cada tipo.

Primeiramente iremos diferenciar os tipos de variáveis e em seguida expor todos os principais modelos de tabelas e gráficos. Tudo isso junto compõe os diferentes tipos de apresentações de dados.



0.2 Variáveis

Uma variável é aquilo que se deseja observar para tirar algum tipo de conclusão. Os dados de uma certa pesquisa são compostos por variáveis. Essas, podem ser divididas em:

- Variáveis Qualitativas;
- Variáveis Quantitativas;

0.2.1 Variáveis Qualitativas

As variáveis qualitativas não podem ser expressas numericamente, pois relacionam situações como cor da pele, cor dos olhos, marca de refrigerante, marca de automóvel entre outras. Podem ser divididas em ordinais e nominais.[2]

0.2.1.1 Variáveis Qualitativas Nominais

Não estão relacionadas a ordem, elas são identificadas apenas por nomes, por exemplo, as cores: Vermelho, amarelo, preto, azul, rosa, verde. Além das cores temos as marcas de carros, nome de bebidas, local de nascimento.[3]

Um exemplo mais concreto se vê na observação de ursos, separando-os por sexo, não há uma ordem natural em suas categorias, a ordem na linha das tabelas pode ser qualquer uma, nesse caso se dá por:

Tabela 1 – Distribuição de Dados Qualitativos Nominais

| Sexo | Frequência Absoluta | Frequência Relativa(%) |
|-----------|---------------------|------------------------|
| Feminino | 35 | 36,1 |
| Masculino | 62 | 63,9 |
| Total | 97 | 100,0 |

Fonte: Elaborado pelo autor

0.2.1.2 Variáveis Qualitativas Ordinais

Nem sempre são numéricas, mas obedecem a uma relação de ordem, por exemplo, conceitos como ótimo, bom, regular e ruim, classe social, grau de instrução.[3]

No caso da observação dos ursos, a distribuição de frequência dos ursos segundo o mês de observação representa uma variável qualitativa ordinal, segue a tabela:

Tabela 2 – Exemplo de variável qualitativa ordinal: Distribuição de frequência dos 97 ursos segundo mês de observação.

| Mês de Observação | Frequência Absoluta | Frequência Relativa(%) | Frequência Absoluta Acumulada | Frequência Relativa Acumulada |
|-------------------|---------------------|------------------------|-------------------------------|-------------------------------|
| Abril | 8 | 8,3 | 8 | 8,3 |
| Maio | 6 | 6,2 | 14 | 14,5 |
| Junho | 6 | 6,2 | 20 | 20,7 |
| Julho | 11 | 11,3 | 31 | 32,0 |
| Agosto | 23 | 23,7 | 54 | 55,7 |
| Setembro | 20 | 20,6 | 74 | 76,3 |
| Outubro | 14 | 14,4 | 88 | 90,7 |
| Novembro | 9 | 9,3 | 97 | 100,0 |
| Total | 97 | 100,0 | – | – |

Fonte: www.Google.com.br

0.2.2 Variáveis Quantitativas

As variáveis quantitativas são as características que podem ser medidas em uma escala quantitativa, ou seja, apresentam valores numéricos que fazem sentido. Podem ser contínuas ou discretas.

../../../../Imagens/SA03x.jpg

0.2.2.1 Variáveis Discretas

São características mensuráveis que podem assumir apenas um número finito ou infinito contável de valores e, assim, somente fazem sentido valores inteiros. Geralmente são o resultado de contagens. Variáveis discretas são como: número de filhos, número de bactérias por litro de leite, números de garrafas quebradas em uma fábrica e assim por diante. [2]

0.2.2.2 Variáveis Contínuas

Essas variáveis, diferentemente das discretas, podem possuir valores fracionados, pois, são organizados em intervalos e com isso podem possuir quaisquer número entre os limites inferior e superior. Usualmente devem ser medidas através de algum instrumento. Exemplos: peso (balança), altura (régua), tempo (relógio), pressão arterial, idade. [2]

0.3 Distribuição de Frequência

A distribuição de frequência é um arranjo de valores, organizando os dados para que assim possam ser melhor utilizados. Primeiramente possuímos os dados brutos que foram coletados de uma certa pesquisa, para organizá-los precisamos colocá-los em rol (em ordem crescente) e em seguida transferi-los para uma tabela de frequência. As tabelas de frequência servem de base para as representações gráficas.

Existem dois tipos de tabelas:

- Tabela para Variáveis Qualitativas;
- Tabela para Variáveis Quantitativas;

0.3.1 Tabela para Variáveis Qualitativas

É organizada na maioria das vezes em 3 colunas, sendo a primeira preenchida pelos dados qualitativos, a segunda a frequência absoluta e por último a frequência relativa. [4]

Exemplo:

Tabela 3 – Cor dos olhos de uma certa amostra de população

| Categoria | Frequência Absoluta | Frequência Relativa(%) |
|-----------|---------------------|------------------------|
| Castanhos | 10 | 0,50 |
| Pretos | 7 | 0,35 |
| Azuis | 2 | 0,10 |
| Verdes | 1 | 0,05 |
| Total | 20 | 1,00 |

Fonte: www.Google.com.br

0.3.2 Tabela para Variáveis Quantitativos

Esse tipo de tabela é subdividido em dois tipos: de ordem discreta e de ordem contínua. Ambos possuem frequência absoluta, relativa e a acumulada. O que difere uma da outra é o modo como estão os dados, se são pontuais ou se estão em intervalos.

0.3.2.1 Tabela para Dados de tipo Quantitativo Discreto

Um exemplo de tabela para dados discretos: Número de irmãos em 20 alunos de uma turma.

Tabela 4 – Tabela de Frequência do Número de Irmãos

| Classe | Frequência Absoluta | Frequência Relativa(%) | Frequência Absoluta Acumulada | Frequência Relativa Acumulada |
|--------|---------------------|------------------------|-------------------------------|-------------------------------|
| 0 | 5 | 25 | 5 | 25 |
| 1 | 8 | 40 | 13 | 65 |
| 2 | 5 | 25 | 18 | 90 |
| 3 | 2 | 10 | 20 | 100 |
| Total | 20 | 100 | | |

Fonte: www.Google.com.br

Convém salientar que as colunas referentes às frequências acumuladas só fazem sentido em tabelas de frequências onde a variável em estudo se possa ordenar (no exemplo da tabela de frequências para dados de tipo qualitativo, apresentado anteriormente, não tem sentido considerar as frequências acumuladas).[4]

0.3.2.2 Tabela para Dados de tipo Quantitativo Contínuo

Se os dados são de natureza quantitativa contínua, consideram-se classes na forma de intervalos. Sempre que possível estes intervalos devem ter a mesma amplitude.

Exemplo:

Considerando a seguinte amostra que resultou de observar a variável Altura em 30 alunos de uma turma.

Tabela 5 – Tabela de Frequência da Altura dos Alunos

| Classes | Frequência Absoluta | Frequência Relativa(%) | Frequência Absoluta Acumulada | Frequência Relativa Acumulada |
|----------|---------------------|------------------------|-------------------------------|-------------------------------|
| 131├ 141 | 2 | 7 | 2 | 7 |
| 141├ 151 | 6 | 20 | 8 | 27 |
| 151├ 161 | 7 | 23 | 15 | 50 |
| 161├ 171 | 9 | 30 | 24 | 80 |
| 171├ 181 | 6 | 20 | 30 | 100 |
| Total | 30 | 100 | | |

Fonte: www.Google.com.br

0.4 Gráficos

Gráfico é uma composição de barras, colunas ou qualquer elemento que consiga expressar a valor dos dados de forma visual. Assim, podemos dizer que os gráficos foram criados para que possamos compreender o que se passa com um certo experimento ou pesquisa apenas olhando para os diferentes tamanhos dos elementos ou as linhas que o complementam. Esse estilo de apresentação de dados pode ser dividido em:

- Gráficos para Variáveis Qualitativas;
- Gráficos para Variáveis Quantitativas;

0.4.1 Gráficos para Variáveis Qualitativas

A representação de gráficos de variáveis qualitativas possui várias formas, onde vários são apenas variações do mesmo princípio. Esta representação é feita basicamente na forma de gráficos de barras ou colunas, ou gráficos de setores (*pizza*), com suas diversas variações.[5]

Podem, também, ser utilizados gráficos pictóricos, que, na são como variações dos gráficos de barras. Todos os gráficos também podem ser usados para gerar gráficos de variáveis quantitativas discretas.

../../../../Imagens/SA03x.jpg

0.4.1.1 Gráfico de Barras

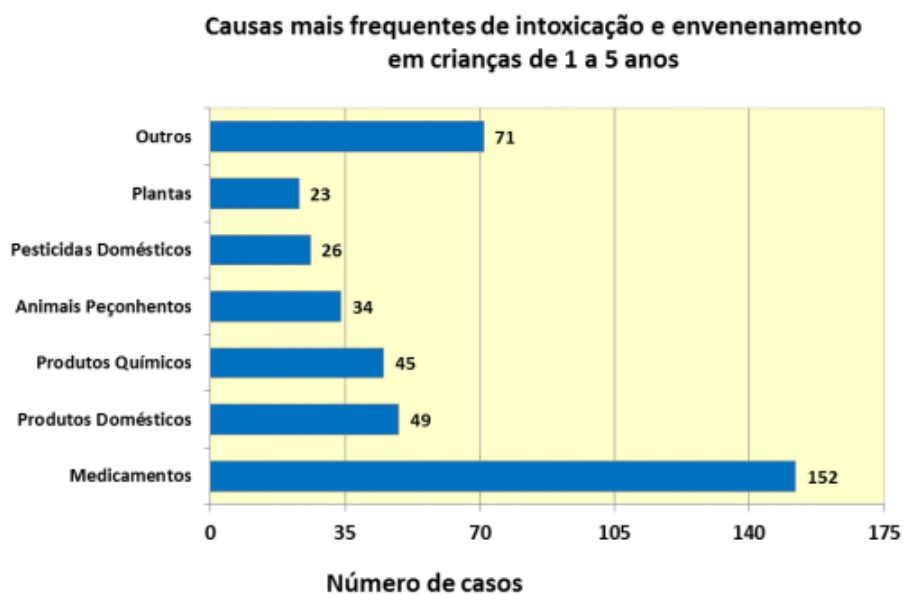
Este tipo de gráfico possui como vantagem a sua forma simples, que permite visualizar as diferenças entre os dados com maior facilidade e rapidez. Tem a finalidade de comparar grandezas por meio de retângulos de igual largura e alturas proporcionais às respectivas grandezas.

Neste tipo de gráfico, cada barra representa a intensidade ou frequência de uma categoria ou atributo. Os espaços existentes entre as barras devem ser iguais. É preferível a este quando as legendas das categorias forem mais curtas, por maior facilidade de posicionamento e formatação.[5]

Na área geográfica este gráfico é muito utilizado, seja para representar fenômenos geográficos ou as diferentes etnias de um determinado local.

Gráfico de Barras Horizontais:

Exemplo: Causas mais frequentes de intoxicação e envenenamento em crianças de 1 a 5, anos em valores absolutos



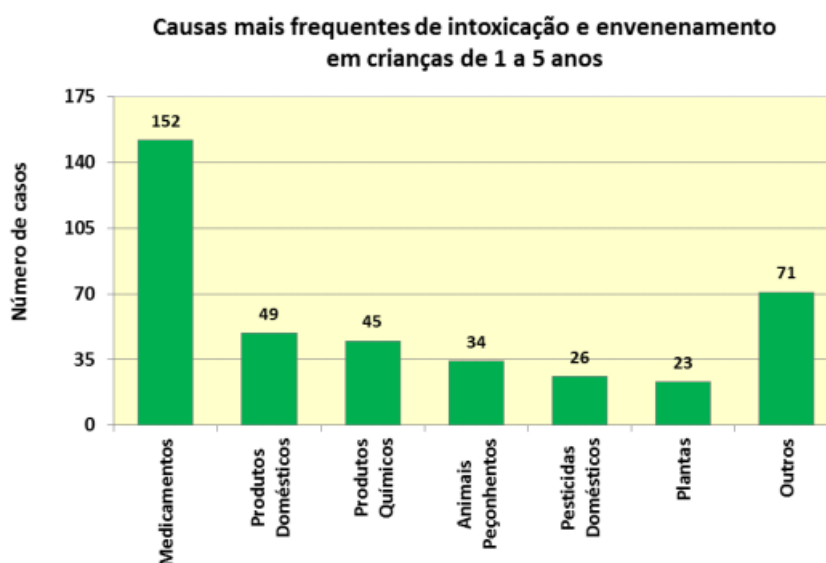
Fonte: www.Google.com.br

Gráfico de Colunas:

Este tipo de gráfico possui a mesma finalidade do gráfico de barras, entretanto, neste as barras são posicionadas na vertical, de maneira a formar colunas.

Exemplo:

Causas mais frequentes de intoxicação e envenenamento em crianças de 1 a 5, anos em valores absolutos.



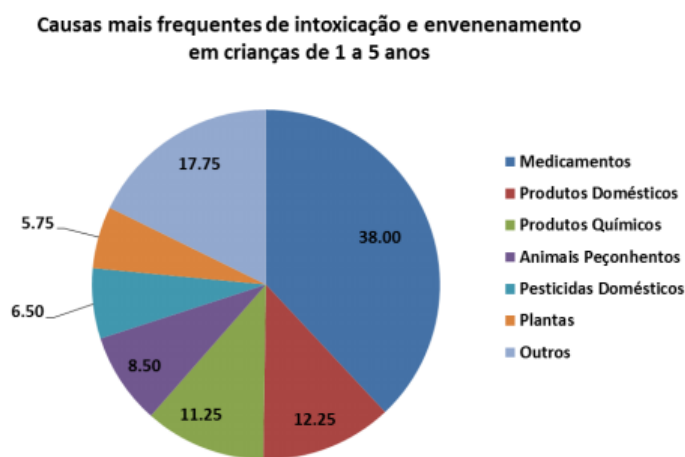
Fonte: www.Google.com.br

0.4.1.2 Gráfico de Pizza

Para construir este gráfico tem-se que encontrar a frequência relativa a princípio de cada setor a ser analisado.[6] Para maior precisão nesse gráfico o ângulo central deve ser calculado, para que assim a disposição de cada categoria não fique de forma confusa, para encontrar tal ângulo multiplica-se a frequência relativa (%) por 360° assim encontrasse o ângulo exato que deve ser desenhado.

Gráfico de Setores:

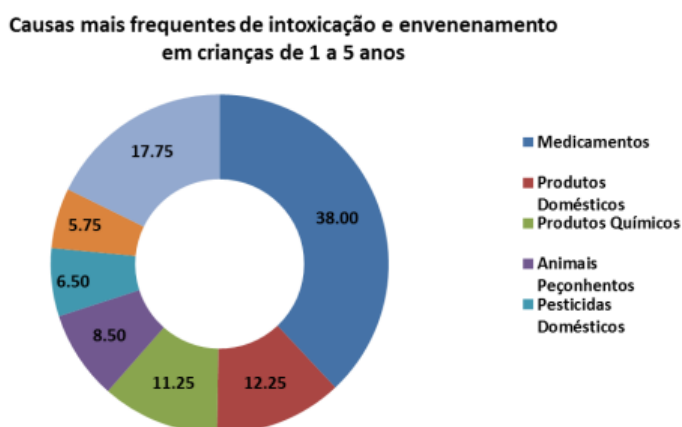
Este tipo de gráfico é a representação dos dados estatísticos em um círculo através de setores. De modo que as áreas são proporcionais aos valores da série. Utilizado principalmente para verificação de percentuais de cada valor da série com o total.[7]



Fonte: www.Google.com.br

Gráfico de Rosca:

Uma das variações do gráfico de pizza é o gráfico de rosca.



Fonte: www.Google.com.br

0.4.2 Gráficos para Variáveis Quantitativas

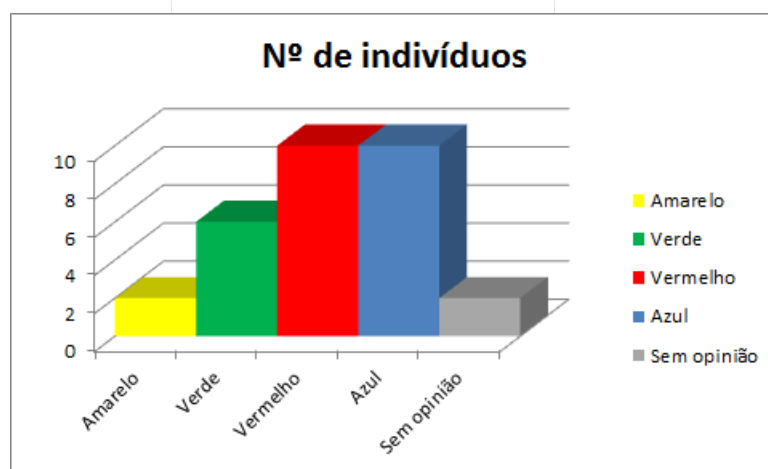
Os gráficos para variáveis quantitativas representam as tabelas de distribuição contínua e discreta de forma mais visual. Eles são os mais utilizados quando estamos trabalhando com dados estatísticos, pois o histograma e o boxplot expressam os dados de forma organizada e que facilita nosso entendimento.[1]

0.4.2.1 Histograma

O gráfico de histograma remete a dados quantitativos contínuos. Esse tipo de gráfico é utilizado para nos mostrar a frequência dos dados. Existem alguns subtipos desse gráfico e um deles é relacionado com a frequência relativa, fazendo com que a área de cada retângulo no gráfico represente essa frequência, assim a área total é igual a 1. Quando os intervalos possuem mesma amplitude sua altura é definida pela frequência de cada dado.[8]

Outro subtipo desse gráfico é o chamado usual, quando os dados representam a frequência absoluta. Além desse, existe o de frequência cumulativa que apresenta os dados acumulados da frequência, o que nos traz dados interessantes para encontrarmos a moda.

Exemplo:



Fonte: www.Google.com.br

0.4.2.2 Polígono de Frequência

Um polígono de frequência é um gráfico que se realiza através da união dos pontos mais altos das colunas num histograma de frequência. Eles são feitos a partir da marca de classe que coincide com o ponto médio de cada coluna do histograma (meio de cada amplitude de classe). Quando são representadas as frequências acumuladas de uma tabela de dados agrupados, obtém-se um histograma de frequências acumuladas, que permite dispor em diagrama o seu polígono correspondente.[9]

Geralmente, os polígonos de frequência são usados quando se pretende mostrar mais de uma distribuição ou a classificação cruzada de uma variável quantitativa contínua com uma qualitativa ou quantitativa discreta num mesmo gráfico. O ponto que tiver mais altura num polígono de frequência representa a maior frequência, ao passo que a área abaixo da curva inclui a totalidade dos dados existentes.

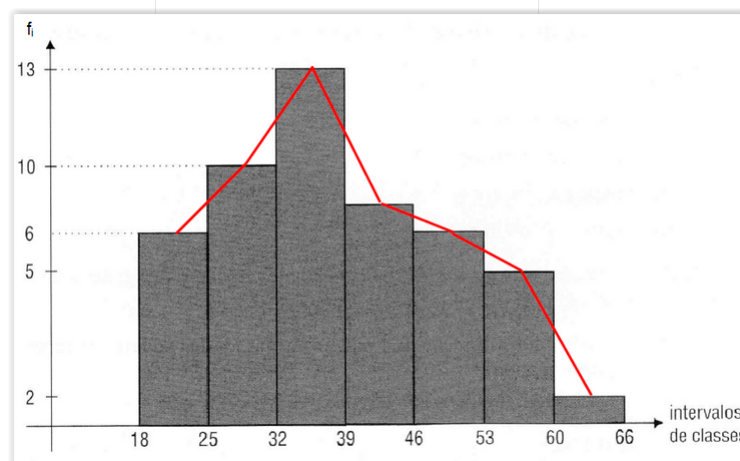
Exemplo:

Considerando as idades de 50 funcionários de uma empresa, agrupados conforme a tabela a seguir:

Tabela 6 – Tabela das idades dos funcionários

| Classes | Intervalos | Frequência (%) |
|---------|------------|----------------|
| 1 | 18– 25 | 6 |
| 2 | 25– 32 | 10 |
| 3 | 32– 39 | 13 |
| 4 | 39– 46 | 8 |
| 5 | 46– 53 | 6 |
| 6 | 53– 60 | 5 |
| 7 | 60– 66 | 2 |
| Total | | 50 |

Fonte: www.professorguru.com.br



Fonte: www.professorguru.com.br

0.4.2.3 BoxPlot

O gráfico BoxPlot é o mais completo dentre todos, pois, em seu ambiente conseguimos notar muitos dados que o histograma não apresenta. Entre esses dados notamos: 1º e 3º quartil, a mediana e os limites inferiores e superiores (esses dois últimos, no entanto, são apresentados no histograma também).

O boxplot tem uma reta (whisker) que estende-se verticalmente ou horizontalmente a partir da caixa, indicando a variabilidade fora do quartil superior e do quartil inferior. Os valores atípicos ou outliers podem ser plotados como pontos individuais. Os espaços entre as diferentes partes da caixa indicam o grau de dispersão, a obliquidade nos dados e os outliers.[1]

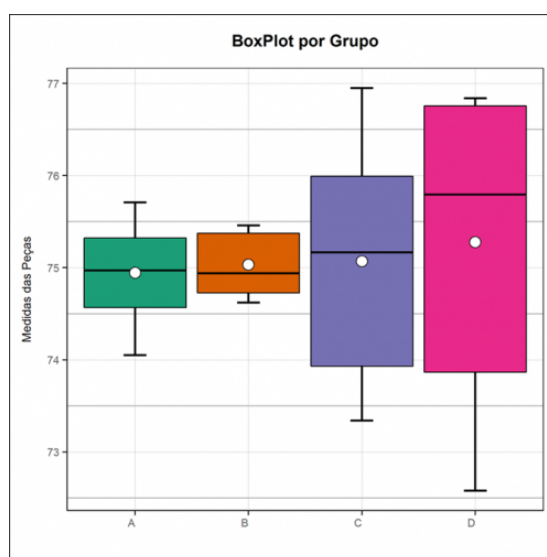
Exemplo:

Uma indústria produz uma peça automotiva cujo valor de referência é 75cm. Após verificar peças fora dos padrões, enviaram duas equipes de trabalhadores (A e B) para um treinamento. Para verificar a eficiência do treinamento, foram selecionadas 10 peças produzidas pelas equipes A e B e 10 peças produzidas pelas equipes C e D que não participaram do treinamento.

Tabela 7 – Tamanho das peças produzidas pelas 4 equipes

| Equipe A | Equipe B | Equipe C | Equipe D (%) |
|----------|----------|----------|--------------|
| 75,27 | 74,94 | 75,93 | 75,98 |
| 75,33 | 75,25 | 76,95 | 75,61 |
| 74,58 | 75,44 | 75,47 | 74,20 |
| 75,01 | 74,62 | 73,60 | 76,44 |
| 75,71 | 75,35 | 74,85 | 76,84 |
| 74,93 | 74,75 | 73,34 | 76,75 |
| 74,72 | 74,65 | 74,04 | 76,78 |
| 74,53 | 74,94 | 75,00 | 74,74 |
| 75,32 | 74,92 | 76,18 | 72,58 |
| 74,05 | 75,46 | 75,33 | 72,86 |

Fonte: www.portalaction.com.br



Fonte: www.portalaction.com.br

0.5 Conclusão

Estatística é uma ciência exata que visa fornecer subsídios ao analista para coletar, organizar, resumir, analisar e apresentar dados.[10] Para que esses dados sejam de fácil entendimento e visualização foram criadas: Tabelas de frequência e diferentes tipos de gráficos. E tudo isso dividido conforme o tipo de variável com a qual se está trabalhando seja ela qualitativa, nominal ou ordinal, ou quantitativa, discreta ou contínua.

Tal ciência é importante método de gestão nos dias atuais, e é utilizada na maioria das vezes para aumentar a credibilidade de argumentos, pois, nada mais são que pesquisas feitas para entendermos como uma amostra ou população está sendo afetada.[11]

Estatística é método, ciência e arte. É método quando, na Física, na Biologia, na Medicina ou na Pedagogia, aplica-se a populações específicas, isto é, serve a uma ciência particular, da qual se torna instrumento. É ciência quando, graças às suas teorias, estuda grandes conjuntos, independentemente da natureza destes, sendo autônoma e universal. Finalmente, é arte na construção de modelos para representar a realidade.[12]

../../../../Imagens/SA03x.jpg

Referências

- [1] Wikipedia, “Wikipedia.” <https://pt.wikipedia.org>.
- [2] Ufpr, “Tipos de variáveis.” <http://leg.ufpr.br/~silvia/CE055/node8.html>.
- [3] M. Educação, “Variáveis na estatística.” <http://mundoeducacao.bol.uol.com.br/matematica/variaveis-na-estatistica.htm>.
- [4] Wikiciencias, “Tabela de frequência.” http://wikiciencias.casadasciencias.org/wiki/index.php/Tabela_de_frequ%C3%Aancias.
- [5] Posgraduando, “Variáveis qualitativas apresentação.” <http://posgraduando.com/as-variaveis-quantitativas-e-qualitativas-e-os-testes-estatisticos/>.
- [6] Ufscar, “Apresentação de dados qualitativos.” http://www.ufscar.br/jcfogo/Bioestat/arquivos/Bioestat2_Variaveis_Qualitativas.pdf.
- [7] Ufpb, “Gráficos de variáveis qualitativas.” <http://www.de.ufpb.br/~luiz/AED/Aula5.pdf>.

- [8] PortalAction, “Histograma.” <http://www.portalaction.com.br/estatistica-basica/16-histograma>.
- [9] Conceito.de, “Polígono de frequência.” <http://conceito.de/poligono-de-frequencia>.
- [10] S. Matematica, “Estatística.” <http://www.somatematica.com.br/estat/basica/paginal.php>.
- [11] Administradores, “Estatística.” <http://www.administradores.com.br/artigos/tecnologia/a-importancia-da-estatistica/67496/>.
- [12] P. Educação, “Estatística.” <https://www.portaleducacao.com.br/conteudo/artigos/educacao/a-importancia-da-estatistica/30442>.