Universidade Federal de Goiás
Instituto de Matemática e Estatística
Programa de Mestrado Profissional em
Matemática em Rede Nacional

Estatística e trânsito: a conscientização por meio de um ensino contextualizado

Rodolpho Mamede de Oliveira

Goiânia

2018

sibi sistema de bibliotecas ufg

PRPG PRÓ-REITORIA DE PÓS-GRADUAÇÃO

UFG

TERMO DE CIÊNCIA E DE AUTORIZAÇÃO PARA DISPONIBILIZAR VERSÕES ELETRÔNICAS DE TESES E DISSERTAÇÕES NA BIBLIOTECA DIGITAL DA UFG

Na qualidade de titular dos direitos de autor, autorizo a Universidade Federal de Goiás (UFG) a disponibilizar, gratuitamente, por meio da Biblioteca Digital de Teses e Dissertações (BDTD/UFG), regulamentada pela Resolução CEPEC nº 832/2007, sem ressarcimento dos direitos autorais, de acordo com a Lei nº 9610/98, o documento conforme permissões assinaladas abaixo, para fins de leitura, impressão e/ou *download*, a título de divulgação da produção científica brasileira, a partir desta data.

1. Identificação do material bibliográfico: **[x] Dissertação** **[ ] Tese**

2. Identificação da Tese ou Dissertação:

Nome completo do autor: Rodolpho Mamede de Oliveira

Título do trabalho: Estatística e trânsito: a conscientização por meio de um ensino contextualizado.

3. Informações de acesso ao documento:

Concorda com a liberação total do documento [x] SIM [ ] NÃO[1]

Havendo concordância com a disponibilização eletrônica, torna-se imprescindível o envio do(s) arquivo(s) em formato digital PDF da tese ou dissertação.

Rodolpho Mamede de Oliveira
Assinatura do(a) autor(a)[2]

Ciente e de acordo:

Tiago Moreira Vargas
Assinatura do(a) orientador(a)[2]

Data: 28 / 08 / 2018

[1] Neste caso o documento será embargado por até um ano a partir da data de defesa. A extensão deste prazo suscita justificativa junto à coordenação do curso. Os dados do documento não serão disponibilizados durante o período de embargo.
Casos de embargo:
- Solicitação de registro de patente;
- Submissão de artigo em revista científica;
- Publicação como capítulo de livro;
- Publicação da dissertação/tese em livro.

[2] A assinatura deve ser escaneada.

Rodolpho Mamede de Oliveira

Estatística e trânsito: a conscientização por meio de um ensino contextualizado

Trabalho de Conclusão de Curso apresentado ao Instituto de Matemática e Estatística da Universidade Federal de Goiás, como parte dos requisitos para obtenção do grau de Mestre em Matemática.
Área de Concentração: Matemática do Ensino Básico
Orientador: Prof. Dr. Tiago Vargas

Goiânia

2018

Ficha de identificação da obra elaborada pelo autor, através do
Programa de Geração Automática do Sistema de Bibliotecas da UFG.

Oliveira, Rodolpho Mamede de
Estatística e trânsito: a conscientização por meio de um ensino contextualizado. [manuscrito] / Rodolpho Mamede de Oliveira. - 2018.
lxxxviii, 88 f.: il.

Orientador: Prof. Dr. Tiago Moreira Vargas.
Dissertação (Mestrado) - Universidade Federal de Goiás, Instituto de Matemática e Estatística (IME), PROFMAT - Programa de Pós graduação em Matemática em Rede Nacional - Sociedade Brasileira de Matemática (RG), Goiânia, 2018.
Bibliografia. Apêndice.
Inclui fotografias, abreviaturas, gráfico, tabelas, lista de figuras, lista de tabelas.

1. Estatística. 2. Ensino contextualizado. 3. Trânsito . 4. Goiânia. I. Vargas, Tiago Moreira, orient. II. Título.

CDU 51

Universidade Federal de Goiás - UFG
Instituto de Matemática e Estatística - IME
Mestrado Profissional em Matemática
em Rede Nacional – PROFMAT/UFG
Campus Samambaia – Caixa Postal 131 – CEP: 74.001-970 – Goiânia-GO.
Fones: (62) 3521-1208 e 3521-1137 www.ime.ufg.br

UFG

PROFMAT

Ata da reunião da banca examinadora da defesa de Trabalho de Conclusão de Curso do aluno Rodolpho Mamede de Oliveira – Aos treze dias do mês de agosto do ano de dois mil e dezoito, às 14:00 horas, reuniram-se os componentes da Banca Examinadora: Prof. Dr. Tiago Moreira Vargas – Orientador, Prof. Dr. Valdivino Vargas Júnior, Profª. Drª. Thaynara Arielly de Lima e o Prof. Dr. Hugo Leonardo da Silva Belisário, para, sob a presidência da primeira, e em sessão pública realizada na sala de aula do IME, procederem a avaliação da defesa intitulada **"Estatística e trânsito: a conscientização por meio de um ensino contextualizado"**, em nível de mestrado, área de concentração Matemática do Ensino Básico, de autoria de Rodolpho Mamede de Oliveira, discente do Programa de Mestrado Profissional em Matemática em Rede Nacional - PROFMAT da Universidade Federal de Goiás. A sessão foi aberta pela presidente da banca, Prof. Dr. Tiago Moreira Vargas, que fez a apresentação formal dos membros da banca. A seguir, a palavra foi concedida ao autor do TCC que, em 30 minutos, procedeu à apresentação de seu trabalho. Terminada a apresentação, cada membro da banca arguiu o examinando, tendo-se adotado o sistema de diálogo sequencial. Terminada a fase de arguição, procedeu-se à avaliação da defesa. Tendo em vista o que consta na Resolução nº. 1403/2016 do Conselho de Ensino, Pesquisa, Extensão e Cultura (CEPEC), que regulamenta os Programas de Pós-Graduação da UFG, e procedidas as correções recomendadas, o Trabalho foi **APROVADO** por unanimidade, considerando-se integralmente cumprido este requisito para fins de obtenção do título de **MESTRE EM MATEMÁTICA**, na área de concentração Matemática do Ensino Básico pela Universidade Federal de Goiás. A conclusão do curso dar-se-á quando da entrega, na secretaria do IME, da versão definitiva do trabalho, com as devidas correções supervisionadas e aprovadas pelo orientador. Cumpridas as formalidades de pauta, às 16:00 horas, a presidência da mesa encerrou a sessão e, para constar, eu, Sóstenes Soares Gomes, secretário do PROFMAT/UFG, lavrei a presente ata que, após lida e aprovada, segue assinada pelos membros da Banca Examinadora em quatro vias de igual teor.

Tiago Moreira Vargas

Prof. Dr. Tiago Moreira Vargas
Presidente – IME/UFG

Valdivino Vargas Júnior

Prof. Dr. Valdivino Vargas Júnior
Membro – IME/UFG

Thaynara Arielly de Lima

Profª. Drª. Thaynara Arielly de Lima
Membro – IME/UFG

Hugo Leonardo da S. Bel

Prof. Dr. Hugo Leonardo da Silva Belisário
Membro - IFG/Goiânia

Rodolpho Mamede de Oliveira é agente municipal de trânsito desde 1998, graduou-se em Matemática pela Universidade Federal de Goiás em 2008, fez pós-graduação em Gestão Pública pela Faculdade Brasileira de Educação e Cultura (FABEC) em 2015 e atua como professor efetivo da rede estadual de ensino desde 2010.

Dedico este trabalho especialmente aos meus filhos Ana Carolina e Rodolpho Filho, a minha esposa, aos meus pais e a todos meus familiares.

Agradecimentos

Agradeço primeiramente a Deus, por me dar forças e disposição para voltar aos estudos após oito anos da graduação. Tudo Dele, por Ele e para Ele!

Agradeço a minha esposa Talitha pela sua sabedoria, por sempre me incentivar e suportar minha ausência e o nervosismo em alguns momentos. À quem honra, honra!

Aos meus amigos de curso, com os quais pude aprender muito dentro e fora da sala de aula.

À Secretaria Municipal de Trânsito, Transportes e Mobilidade(SMT) nas pessoas de Maria Auxiliadora, chefe da divisão de multas e Thyara Sampaio, chefe da divisão de Geoprocessamento, que me deram todo suporte de dados, sem os quais não seria possível a realização deste trabalho.

Ao meu orientador, professor Dr. Tiago Vargas, por todo comprometimento e contribuição no desenvolvimento deste trabalho.

Aos alunos do 3º PROFEN A do Colégio Estadual Robinho Martins de Azevedo pelo envolvimento no projeto.

Aos integrantes da banca de avaliação por suas sugestões e contribuições.

À CAPES, pelo suporte financeiro, que muito contribuiu nesses dois anos de estudo e pesquisa.

Resumo

O presente trabalho traz um tema transversal à educação bastante presente na vida das pessoas: o trânsito. E tem por finalidade propor um ensino contextualizado da Estatística básica focado nesse tema. Não há o que se falar em trânsito sem que esteja presente conceitos dessa importante e útil área do conhecimento. Acreditamos ser uma abordagem inédita e de grande relevância para o contexto sócio educacional na educação básica. Para tanto, foi selecionada uma turma do 3º PROFEN(Programa de Fortalecimento do Ensino Noturno), um novo modelo adotado pela Secretaria Estadual de Educação de Goiás, equivalente ao 2º ano do ensino médio, com 30 alunos, de um colégio estadual no município de Goiânia para se desenvolver um projeto de pesquisa participante, com foco na contextualização do ensino da estatística voltada para o trânsito (mais especificamente sobre as principais infrações cometidas em Goiânia), procurando inserir o tema em sala de aula e familiarizar os alunos com o assunto. Aqui, o leitor encontrará alguns conceitos e definições sobre a legislação de trânsito vigente em nosso país, inerentes à proposta do trabalho, bem como os conceitos da estatística descritiva e problemas de aplicação. Todos os problemas aqui propostos foram elaborados pelo autor com dados reais, coletados junto aos órgãos de trânsito e, também, de maneira empírica *in loco*.

Palavras-chave

Estatística. Ensino Contextualizado. Trânsito. Goiânia

Abstract

The present work brings a cross - cutting theme to education that is very present in people 's lives: traffic. And it has the purpose of proposing a contextualized teaching of Basic Statistics focused on this theme. There is nothing to talk about in transit without the presence of concepts of this important and useful area of knowledge. We believe it to be an unprecedented and highly relevant approach to the socio-educational context in basic education. To this end, a second-year high school class with 30 students from a state college in the city of Goiânia was selected to develop a participatory research project, focusing on the contextualization of traffic-oriented statistics (more specifically on main infractions committed in Goiânia), trying to insert the theme in the classroom and familiarize the students with the subject. Here, the reader will find some concepts and definitions about the current traffic legislation in our country, inherent to the proposal of the work, as well as the concepts of descriptive statistics and application problems. All the problems proposed here were elaborated by the author with real data, collected from the transit agencies and, also, empirically *in loco*.

Keywords

Statistic. Contextualized Teaching. Transit. Goiânia.

Lista de Figuras

Lista de Tabelas

Sumário

1 Introdução

Um dos assuntos mais comentados e discutidos diariamente, em todo Brasil é o trânsito. E esse é um tema muito polêmico e presente no cotidiano das pessoas, além de ser um grande campo para aplicações de conceitos e definições de uma importante e útil área do conhecimento matemático, a Estatística. Diariamente, nos jornais e na TV, são divulgadas informações com o uso de dados estatísticos sobre o trânsito. O presente trabalho tem como proposta a delineação de uma pesquisa participante à cerca das principais infrações de trânsito cometidas na cidade de Goiânia nos anos de 2017 e 2018, trazendo como questão norteadora a importância de se inserir o tema "trânsito" em sala de aula. Acreditamos que trata-se de uma abordagem inédita e que a inserção do tema em sala de aula é de alta relevância pois, além de trazer um ensino de matemática contextualizado em um tema contundente, tem a intenção de levar os alunos à uma reflexão sobre o assunto.

Os temas transversais são de grande importância para ensino e de acordo com os Parâmetros Curriculares Nacionais (PCN's) [1]:

> *A inclusão de questões sociais no currículo escolar não é uma preocupação inédita. Essas temáticas já têm sido discutidas e incorporadas às áreas ligadas às Ciências Sociais e Ciências Naturais, chegando mesmo, em algumas propostas, a constituir novas áreas, como no caso dos temas Meio Ambiente e Saúde.*

É fato que os PCN's compreendem apenas seis temas transversais (Ética, Orientação Sexual, Meio Ambiente, Saúde, Pluralidade Cultural, Trabalho e Consumo), mas os conteúdos de educação para o trânsito, por exemplo, entram em quase todos eles. No entanto, baseado nos próprios dados estatísticos divulgados diariamente em todo tipo de mídia, que trazem índices cada vez mais alarmantes com inúmeras mortes no trânsito, acreditamos que esse tema deveria ser trabalhado não só como um tema transversal, mas sim, como uma disciplina curricular obrigatória.

As demandas do mundo moderno fazem com que a socidedade passe a ter que assimilar novos conhecimentos para lidar com fatos e fenômenos no seu dia a dia. Obviamente, espera-se que a educação seja uma possibilidade de acesso a tais conhecimentos para a população e, de acordo com de acordo com Tomaz e David [3]:

> A matemática vem ganhando espaço nesse cenário e sendo demandada a produzir modelos para descrever ou ajudar a compreender fenômenos nas

> diversas áreas do saber, produzindo conhecimentos novos nessas áreas, ao mesmo tempo que se desenvolve enquanto campo de conhecimento científico

Entendemos que o conhecimento dos conteúdos de maneira geral por si só não favorecem a compreensão de maneira ampla das diversas situações reais vividas pelos alunos, daí a importância de se propor um ensino interdisciplinar e contextualizdo. Com este ensejo, esse projeto será desenvolvido em uma turma do 3º período do ensino médio(3ºPROFEN) com 30 alunos, do Colégio Estadual Robinho Martins de Azevedo situado no Jardim Nova Esperança, no município de Goiânia. Para isso, serão utilizados dados secundários, coletados junto aos órgãos de trânsito e também dados primários, coletados pelo próprio autor, *in loco*, nos meses de Abril e Maio de 2018. Dados esses reais e atuais que poderão dar aos alunos e também ao leitor, uma noção e interpretação da atual situação do trânsito (no tocante às principais infrações) no município de Goiânia. Foi idealizado por ser possível fazer um relação empírica entre a Estatística e uma das profissões do autor que, além de professor de matemática na rede estadual de ensino, é agente municipal de trânsito (figura que atua diariamente na orientação, educação e fiscalização de trânsito) há 19 anos no município de Goiânia. Com isso, uma das propostas aqui é inserir o tema "trânsito" em sala de aula, buscando familiarizar os alunos (e futuros condutores) com o assunto, além de fornecer uma fonte auxiliar a professores de matemática da educação básica com problemas contextualizados e aplicações dos conceitos da estatística descritiva. Professores que tenham por objetivo a ministração deste importante conteúdo matemático e, fundamentalmente, a conscientização dos futuros condutores de veículos.

O objetivo geral aqui é contextualizar a Estatística com ênfase nas principais infrações de trânsito buscando conscientizar os alunos através dos números, levando-os a uma profunda reflexão sobre a atual situação do trânsito em nossa capital. Como objetivos específicos, temos:

- Abordar os conceitos e definições básicas inerentes às infrações de trânsito;
- Despertar o espírito pesquisador nos alunos;
- Elaborar e resolver problemas de estatística contextualizados com dados reais sobre o trânsito;
- Conscientizar os futuros condutores à partir dos problemas propostos;
- Avaliar a viabilidade da inclusão de uma disciplina sobre o trânsito no currículo escolar.

No segundo capítulo, apresentaremos alguns conceitos e definições presentes no Código de Trânsito Brasileiro(CTB) [5], pertinentes aos objetivos do trabalho, onde os alunos terão o contato inicial com a lei que rege as normas de trânsito. A intenção é de que eles associem o que está presente na lei com o que vêem no seu dia-a-dia nas ruas da cidade e até mesmo nas atitudes de seus familiares quando estão na condução de um veículo automotor.

O terceiro capítulo irá abordar a parte matemática do trabalho, com referencial teórico, a relevância da Estatística e os principais conceitos da Estatística Descritiva.

O quarto capítulo traz o foco principal do projeto, onde aplicaremos a matemática de forma bem contextualizada trabalhando os conceitos da estatística descritiva presentes no Capítulo 3. É a parte prática e participante e se subdividirá em duas etapas: na primeira, apresentaremos aos alunos os dados coletados secundariamente, junto aos órgãos de trânsito, momento em que iremos propor os primeiros problemas. Na segunda etapa, a coleta de dados far-se-à *in loco* e os problemas propostos serão elaborados e trabalhados em sala de aula.

No quinto capítulo, faremos a análise dos resultados obtidos quando da elaboração, execução e resolução dos problemas propostos bem como o nível de comprometimento dos alunos e qual o impacto que o projeto causou na turma como um todo. Por fim, faremos uma avaliação da relevância e importância da abordagem empregada neste trabalho, nas considerações finais.

2 Considerações iniciais e algumas definições do CTB

Como a proposta deste trabalho é de uma pesquisa participante com alunos do ensino médio, contextualizando a estatística com enfoque nas principais infrações de trânsito, faz-se necessário apresentar alguns dados sobre a cidade de Goiânia assim como algumas noções preliminares acerca das normas gerais de circulação e conduta definidas no Código de Trânsito Brasileiro (CTB), bem como os artigos referentes às principais infrações de trânsito, no que diz respeito à circulação, estacionamento e parada dos veículos automotores.

2.1 Goiânia, a capital dos veículos

Goiânia, capital do estado de Goiás, foi fundada em 24 de outubro de 1933 e planejada para 50 mil habitantes. No entanto, de acordo com o último censo do Instituto Brasileiro de Geografia e Estatística (IBGE) [2] feito em 2010, Goiânia tinha uma população de 1.302.001 pessoas e foi estimado, para 2017, um total de 1.466.105 pessoas. Tem uma área de aproximadamente $740km^2$ e uma densidade demográfica aproximadamente de 1.981,22 hab/km^2. O gráfico a seguir, mostra a evolução da frota de veículos em Goiânia de 1999 à 2006 de acordo com a Associação Brasileira de Medicina de Tráfego(ABRAMET) [4]:

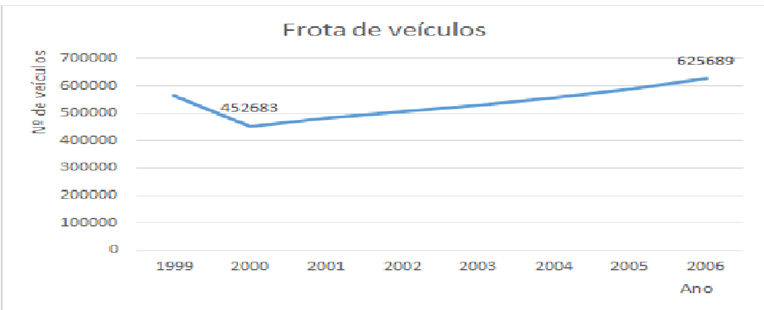


Figura 1: Frota de 1999 à 2006/(*fonte: própria com dados da ABRAMET*)

Com base no gráfico, podemos observar que o aumento da frota de 2000 à 2006 foi cerca de 38%. Nos dias atuais, de acordo com o DETRAN-GO a cidade conta com uma frota de 1.185.875 veículos registrados(atualizado em 12/01/2018), o que corresponde a um aumento de mais de 89% nos últimos 12 anos, isso da uma média de aproximadamente 1,23 habitantes para cada veículo e que, de acordo com estudos e projeções de dois pesquisadores, doutores em matemática pela Pontifícia Universidade Católica de Goiás, chegará a incrível marca de um veículo por habitante em 2022. É a capital com maior índice de veículos por habitante do Brasil. Pode-se observar com esses dados que o número de veículos em circulação é 23 vezes maior que o número de habitantes inicialmente para o qual a cidade fora projetada. Considerando tudo isso, é de se entender porque o trânsito aqui piora à cada ano. Apesar disso, pode-se afirmar que em Goiânia não temos congestionamentos, temos sim, trânsito lento (e quase sempre somente nos horários de pico) diferentemente de cidades como São Paulo e Rio de Janeiro, que chegam a ter de 5km a 10km de congestionamento em horários de pico. Goiânia conta, nos dias atuais, com 312 agentes de trânsito, dos quais menos de 250 atuam diretamente no trabalho nas ruas da cidade (os demais ficam em serviços internos). Isso dá uma média de um agente para 4743 veículos sendo que, de acordo com orientação do Departamento Nacional de Trânsito (DENATRAN), o ideal seria que houvesse um agente para cada 1000 veículos. Assim, seria possível um trabalho mais eficaz e intensivo.

2.2 Alguns conceitos do CTB pertinentes

Até o ano de 1997, estava em vigência o "Código Nacional de Trânsito" o qual regulamentava que o trânsito nas vias urbanas era de responsabilidade exclusiva dos estados brasileiros. Com o advento da Lei 9.503 de setembro de 1997, que instituiu o "Código de Trânsito Brasileiro"(CTB), que passou a vigorar em 23 de janeiro de 1998, houve a municipalização do trânsito e, então, as prefeituras tornaram-se responsáveis pelo planejamento, fiscalização, educação, monitoramento, intervenção e controle do trânsito nas cidades, assumindo as questões relacionadas ao pedestre, à circulação, ao estacionamento, à parada de veículos e à implantação da sinalização, atendendo de forma direta as necessidades da comunidade. Para isso, os municípios criaram órgãos que funcionariam como secretarias, agências ou superintendências de trânsito e ficaria a cargo dos estados as questões referentes aos registros de veículos, documentação e equipamentos de segurança obrigatórios. Em Goiânia, esse órgão é a SMT (Secretaria Municipal de Trânsito, Transporte e Mobilidade) e no estado de Goiás o órgão competente é o DETRAN-GO.

No anexo I do CTB é definida a figura do agente municipal de trânsito: pessoa, civil ou policial militar, credenciada pela autoridade de trânsito para o exercício das atividades de fiscalização, operação, policiamento ostensivo de trânsito ou patrulhamento. Ele é o servidor público, concursado, responsável por orientar e fiscalizar as infrações de trânsito referentes à circulação, estacionamento e parada dos veículos. Vale ressaltar aqui, que é possivel ser firmado convênios entre estados e municípios que autorizam a mútua fiscalização de todas as infrações previstas no código. Hoje, este convênio está vigente em Goiânia.

A seguir, foram selecionadas algumas definições extraídas do CTB (acerca do seu capítulo III), muito pertinentes a um dos objetivos deste trabalho, que serão apresentadas à turma. O objetivo é introduzir os conceitos iniciais da lei levantando um debate e promovendo uma constante interação entre pesquisador e alunos, já buscando levá-los às primeiras reflexões.

CAPÍTULO III: DAS NORMAS GERAIS DE CIRCULAÇÃO E CONDUTA

Art. 29. O trânsito de veículos nas vias terrestres abertas à circulação obedecerá às seguintes normas:
I - a circulação far-se-á pelo lado direito da via, admitindo-se as exceções devidamente

sinalizadas;
II - o condutor deverá guardar distância de segurança lateral e frontal entre o seu e os demais veículos, bem como em relação ao bordo da pista, considerando-se, no momento, a velocidade e as condições do local, da circulação, do veículo e as condições climáticas;
III - quando veículos, transitando por fluxos que se cruzem, se aproximarem de local não sinalizado, terá preferência de passagem:
a) no caso de apenas um fluxo ser proveniente de rodovia, aquele que estiver circulando por ela;
b) no caso de rotatória, aquele que estiver circulando por ela;
c) nos demais casos, o que vier pela direita do condutor;
IV - quando uma pista de rolamento comportar várias faixas de circulação no mesmo sentido, são as da direita destinadas ao deslocamento dos veículos mais lentos e de maior porte, quando não houver faixa especial a eles destinada, e as da esquerda, destinadas à ultrapassagem e ao deslocamento dos veículos de maior velocidade;
V - o trânsito de veículos sobre passeios, calçadas e nos acostamentos, só poderá ocorrer para que se adentre ou se saia dos imóveis ou áreas especiais de estacionamento;
VI - os veículos precedidos de batedores terão prioridade de passagem, respeitadas as demais normas de circulação;
VII - os veículos destinados a socorro de incêndio e salvamento, os de polícia, os de fiscalização e operação de trânsito e as ambulâncias, além de prioridade de trânsito, gozam de livre circulação, estacionamento e parada, quando em serviço de urgência e devidamente identificados por dispositivos regulamentares de alarme sonoro e iluminação vermelha intermitente.
IX - a ultrapassagem de outro veículo em movimento deverá ser feita pela esquerda, obedecida a sinalização regulamentar e as demais normas estabelecidas neste Código, exceto quando o veículo a ser ultrapassado estiver sinalizando o propósito de entrar à esquerda;
O condutor utilizará o pisca-alerta nas seguintes situações:
a) em imobilizações ou situações de emergência;
b) quando a regulamentação da via assim o determinar;
Art. 45. Mesmo que a indicação luminosa do semáforo lhe seja favorável, nenhum condutor pode entrar em uma interseção se houver possibilidade de ser obrigado a imobilizar o veículo na área do cruzamento, obstruindo ou impedindo a passagem do trânsito transversal.

Art. 47. Quando proibido o estacionamento na via, a parada deverá restringir-se ao tempo indispensável para embarque ou desembarque de passageiros, desde que não interrompa ou perturbe o fluxo de veículos ou a locomoção de pedestres.
Parágrafo único. A operação de carga ou descarga será regulamentada pelo órgão ou entidade com circunscrição sobre a via e é considerada estacionamento.

3 A Estatística

É cada vez mais acentuada a utilização da Estatística em qualquer atividade profissional da vida moderna e, nos seus diversos ramos de atuação, as pessoas estão frequentemente expostas a essa área do conhecimento. Isto se deve às múltiplas aplicações que o método estatístico proporciona àqueles que rotineiramente fazem o uso dele. É possível distinguir duas concepções para a palavra Estatística, de acordo com Toledo e Ovalle [7]:

> *a) No plural(estatísticas), indica qualquer coleção consistente de dados numéricos, reunidos com a finalidade de fornecer informações acerca de uma atividade qualquer.*
> *b) No singular, indica a atividade humana especializada ou um corpo de técnicas, ou ainda uma metodologia desenvolvida para a coleta, a classificação, a apresentação, a análise e a interpretação de dados quantitativos e a utilização desses dados para a tomada de decisões.*

Por meio dessa segunda assertiva, é possível imaginar o objeto dos estudos etatísticos, que resida naqueles fenômenos que se referem principalmente a um conjunto muito numeroso de indivíduos, que são semelhantes quanto a, pelo menos uma, característica específica. Para Bussab e Morettin (2013, pg.1) [12] "Estatística é a parte da metodologia da ciência que tem por objetivo a coleta, redução, análise e modelagem de dados."

Há que se ressaltar que a Estatística não é um método mediante o qual se pode provar tudo aquilo que se deseja. Para estabelecer o âmbito dos estudos da disciplina e adotando-se um esquema prático de raciocínio, pode-se dizer que a Estatística compreende dois campos bem amplos, a saber: a Estatística descritiva (que discorreremos mais a seguir por ser o campo que trata este trabalho) e a Estatística indutiva, também chamada de Inferência Estatística. Esta última, tem o papel de generalizar conclusões, inferir propriedades para um todo, com base em uma parte.

Em um sentido mais amplo, a Estatística Descritiva pode ser interpretada como uma área cujo objetivo é a observação de fenômenos de mesma natureza, a coleta de dados numéricos referentes a esses fenômenos, a organização e a classificação desses dados observados e a sua apresentação através de gráficos e tabelas, além do cálculo de coeficientes que permitam descrever resumidamente os fenômenos. Nesse compasso, a estatística, conteúdo que só passou a ser inserido no curriculo em 1997 com a publicação dos Parâmetros Curriculares Nacionais (PCN) para os anos iniciais do Ensino

Fundamental, em 1998 com a publicação dos PCN para os anos finais do Ensino Fundamental e, em 1999 com a publicação dos PCN para o Ensino Médio tem um papel preponderante para a formação crítica do aluno pois, de acordo com os Parâmetros Curriculares Nacionais (PCN's) [1]:

> *A Matemática, por sua universalidade de quantificação e expressão, como linguagem portanto, ocupa uma posição singular. No Ensino Médio, quando nas ciências torna-se essencial uma construção abstrata mais elaborada, os instrumentos matemáticos são especialmente importantes. Mas não é só nesse sentido que a Matemática é fundamental. Possivelmente, não existe nenhuma atividade da vida contemporânea, da música à informática, do comércio à meteorologia, da medicina à cartografia, das engenharias às comunicações, em que a Matemática não compareça de maneira insubstituível para codificar, ordenar, quantificar e interpretar compassos, taxas, dosagens, coordenadas, tensões, freqüências e quantas outras variáveis houver.*

Com isso, entendemos que é totalmente justificado o ensino da Estatística já nas séries iniciais e com o foco na contextualização e na interdisciplinaridade, pois trata-se de um conteúdo rico em situações cotidianas que podem ser traduzidas em gráficos e tabelas facilitando o aprendizado e desenvolvendo uma visão crítica no aluno. Assim, de acordo com Carvalho [8]:

> *A inclusão da Estatística, porém, ao menos no seu aspecto descritivo, justificar-se-ia pelo fato de que a imprensa escrita e televisionada vem utilizando mais e mais tabelas e gráficos estatísicos para transmitir as informações. Além disso, os manuais didáticos de outras disciplinas, notadamente de Geografia, cada vez mais constroem os seus textos baseando-se em índices, além dos gráficos e tabelas.*

3.1 Conceitos estatísticos

Definição 1. *Em Estatística, população (ou universo estatístico) é o conjunto constituido por todos os indivíduos que apresentam pelo menos uma característica comum, cujo comportamento interessa analisar.*

Definição 2. *Uma amostra pode ser definida como um subconjunto, uma parte selecionada da totalidade da população, por meio da qual se faz algum juízo sobre suas características.*

Definição 3. *A Estatística dedica-se ao estudo dos fenômenos de massa, que são resultantes da ocorrência de um grande número de causas, total ou parcialmente desconhecidas, que serão chamadas de "fenômenos estatísticos".*

Definição 4. *Variáveis são características de interesse que são medidas em cada elemento da amostra ou população. Como o nome diz, seus valores variam de elemento para elemento.*

As variáveis podem ter valores numéricos ou não e são classificadas em variáveis qualitativas e quantitativas:
Variáveis Quantitativas: são as características que podem ser medidas em uma escala quantitativa, ou seja, apresentam valores numéricos que fazem sentido. Estas, podem ser contínuas ou discretas.
Variáveis quantitativas discretas: são as 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. Exemplos: velocidade dos veículos aferida pelo radar, número de veículos apreendidos em um mês, número de pessoas fiscalizando o trânsito em uma cidade, etc...
Variáveis quantitativas contínuas: são as características mensuráveis que assumem valores em uma escala contínua (na reta real), para as quais valores fracionais fazem sentido. Usualmente devem ser medidas através de algum instrumento. Exemplos: peso bruto de um veículo, valor registrado no teste do bafômetro, etc...
Variáveis Qualitativas: são as características que não possuem valores quantitativos, mas, ao contrário, são definidas por várias categorias, ou seja, representam uma classificação dos indivíduos. Podem ser nominais ou ordinais.
Variáveis qualitativas nominais: não existe ordenação dentre as categorias. Exemplos: sexo, cor dos olhos, fumante/não fumante, doente/sadio.
Variáveis qualitativas ordinais: existe uma ordenação entre as categorias. Exemplos: escolaridade (1º, 2º, 3º graus), estágio da doença (inicial, intermediário, terminal), mês de observação (janeiro, fevereiro,..., dezembro).

A Tabela 1 traz alguns dados coletados da turma selecionada onde este trabalho foi desenvolvido, com a variável quantitativa "idade" e as variáveis qualitativas "sexo" e se tem "habilitação", com intuito de traçar o seu perfil:

Nº de chamada	Sexo	Idade	Habilitação
1	M	63	SIM
2	M	19	NÃO
3	M	19	NÃO
4	F	19	NÃO
5	F	16	NÃO
6	F	16	NÃO
7	F	18	NÃO
8	F	16	NÃO
9	F	16	NÃO
10	M	28	NÃO
11	F	22	NÃO
12	F	32	NÃO
13	F	19	NÃO
14	M	20	NÃO
15	F	25	NÃO
16	M	34	NÃO
17	F	17	NÃO
18	M	18	NÃO
19	M	17	NÃO
20	F	23	NÃO
21	M	16	NÃO
22	M	18	NÃO
23	F	22	NÃO
24	M	48	SIM
25	F	16	NÃO
26	F	16	NÃO
27	M	18	NÃO
28	F	16	NÃO
29	M	17	NÃO
30	M	17	NÃO

Tabela 1: Perfil da turma

3.2 Distribuição de frequências

Quando se estuda uma variável, o maior interesse do pesquisador é conhecer o seu comportamento, analisando a ocorrênica de suas possíveis realizações. Uma distribuição de frequência é uma tabela que contém um resumo dos dados obtido em uma amostra. Cada entrada da tabela, contém a frequência dos dados em um determinado intervalo, ou em um grupo.

Definição 5. *Chamamos de dados brutos, dados originais obtidos a partir da coleta, e que ainda não se encontram prontos para análise.*

Exemplo 1. *A seguir estão relacionadas as idades dos 30 alunos da turma do 3º Profen A, que foram coletadas pela ordem de chamada:*
63,19,19,19,16,16,18,16,16,28,22,32,19,20,25,34,17,18,17,23,16,18,22,48,16,16,18,16,17,17

Definição 6. *O rol é uma lista em que os valores estão dispostos em uma determianda ordem, crescente ou decrescente.*

Exemplo 2. *Os dados brutos do exemplo 1 organizados em um rol, ficam da seguinte forma:*
16,16,16,16,16,16,16,16,17,17,17,17,18,18,18,18,19,19,19,19,20,22,22,23,25,28,32,34,42,63.

3.3 Tabela de Frequências

As tabelas de frequências são representações nas quais os valores se apresentam em correspondência com suas repetições, evitando-se assim que eles apareçam mais de uma vez na tabela(como ocorre com o rol) e podem representar tanto valores individuais(tabulados) como valores agrupados (em classes). As tabelas de frequências facilitam a interpretação dos dados de uma pesquisa, visto que as mesmas agrupam observações com as mesmas características, em um mesmo conjunto.

3.3.1 Elementos de uma distribuição de frequências

Uma tabela de frequências é construída a partir de alguns termos próprios. São eles:
Frequência absoluta(f_i)*:* É o número de observações correspondentes à uma classe ou valor.
Frequência acumulada*(Fac):* É a soma da frequência absoluta de uma classe ou valor com as frequências absolutas das classes ou valores anteriores. Toda vez que se deseja saber quantas observações existem até uma determinada classe ou valor individual, basta observar o valor da frequência acumulada.
Frequência relativa(f_r)*:* Representa a proporção de observações de um valor individual ou de uma classe, em relação ao número total de observações, ou seja, é um

número relativo. O cálculo da f_r é dado pelo quociente entre a frequência absoluta (f_i) da classe e o número total de observações. Assim, temos que:

$$f_r = \frac{f_i}{N} \tag{1}$$

onde N é o número total de observações.

Se desejarmos obter esse valor em termos percentuais, basta multiplicarmos esse quociente por 100. É importante observar que a soma de todas as f_r de uma tabela de frequência é sempre igual a 1(se estiver representada na forma decimal) ou igual a 100%(se estiver representada na forma percentual).

Amplitude total*(At)*: É a diferença entre o maior e o menor valor observado da variável em estudo.

Exemplo 3. *Por exemplo, na Tabela 1 a maior idade foi 63 e a menor foi 16, logo a amplitude total do conjunto de valores é dada por:* ***At*** *= 63 - 16 = 47.*

Classe: Classe de frequência é cada um dos grupos de valores em que se subdivida a amplitude total do conjunto de valores observados da variável. A distribuição deve conter um número adequado de classes para não comprometer a pesquisa e a sua interpretação. Para determinar o número de classes há diversos métodos. Um deles, é a regra de **Sturges**, que estabelece que o número de classes é igual a:

$$k = 1 + 3,3.log_{10}N,$$

em que k é o número de classes e N é o número total de observações.

Exemplo 4. *Se o número de observações for 25, temos que* $N = 25$*, e* $k = 1 + 3,3.log_{10}25 \approx 5,613202$. *Fazendo o arredondamento, consideramos* $k = 6$.

Truman L. Kelley [6], sugere os seguintes números de classes, com base no número total de observações(para efeito de representação gráfica):

N	5	10	25	50	100	200	500	1000
k	2	4	6	8	10	12	15	15

Limites de classes: São os valores extremos de cada classe em uma tabela de distribuição de frequências com dados agrupados. O valor à esquerda do símbolo $\vdash$ é denominado limite inferior (é incluido na classe) e, o valor à direita, limite superior (não é incluido na classe).

Exemplo 5. *Se uma determinada classe tem o intervalo* $15 \vdash 18$*, significa que a classe compreende valores de 15, inclusive, até 18, exclusive.*

Amplitude do Intervalo de Classe: Ou, simplesmente intervalo de classe, é definido como a diferença entre seus limites superior e inferior.

Exemplo 6. *No Exemplo 5, a amplitude do intervalo é 18 - 15 = 3.*

Ponto Médio de Classes(x_j): Ponto médio ou valor médio de classe é o valor que a representa, para efeito de cálculo de certas medidas. É um valor importante, por ser equidistante dos limites de classe. Para se obter o ponto médio de uma classe, basta acrescentar ao seu limite inferior a metade da amplitude do intervalo de classe. No caso de uma distribuição com dados agrupados, o ponto médio de cada classe é calculado pela média aritmética entre os seus limites.

3.3.2 Tabela de frequências de dados tabulados individualmente

Utilizando os dados da Tabela 1, foi construída a Tabela 2 de frequência de dados tabulados, ou seja, uma tabela onde os valores da variável aparecem individualmente, da variável idade.

3.3.3 Distribuição de Frequências em dados agrupados

Muitas vezes os dados coletados são muito extensos e dificultam a análise da variável. Por isso, há vantagem em resumir os dados originais em uma distribuição onde os valores observados não mais aparecerão individualmente, mas sim, agrupados em classes. A Tabela 3 abaixo, mostra a distribuição de frequênicas com dados agrupados da variável idade com base na Tabela 1.

Idade	Freq. absoluta(f_i)	Freq. acumulada(Fac)	Freq. relativa(f_r)	Percentual(%)
16	8	8	0,266	26,6%
17	4	12	0,133	13,3%
18	4	16	0,133	13,3%
19	4	20	0,133	13,3%
20	1	21	0,033	3,33%
22	2	23	0,066	6,66%
23	1	24	0,033	3,33%
25	1	25	0,033	3,33%
28	1	26	0,033	3,33%
32	1	27	0,033	3,33%
34	1	28	0,033	3,33%
48	1	29	0,033	3,33%
63	1	30	0,033	3,33%
Total	30	-	1	100%

Tabela 2: Tabela de distribuição de frequência da variável idade

Classes de idades	Frequência absoluta(f_i)
16$\vdash$ 24	24
24$\vdash$ 32	2
32$\vdash$ 40	2
40$\vdash$ 48	0
48$\vdash$ 56	1
56$\vdash$ 64	1
Total	30

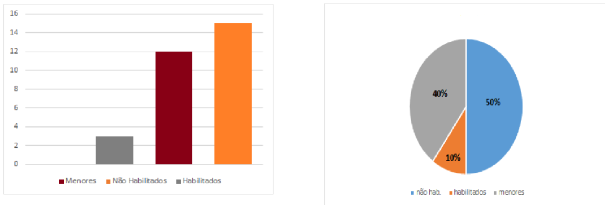
Tabela 3: Tabela de distribuição de frequência com dados agrupados

O símbolo $\vdash$ indica inclusão na classe do valor situado à sua esquerda e exclusão do valor situado à sua direita. Assim, na Tabela 3 a primeira classe ($16 \vdash 24$) agrega os valores de 16, inclusive, até 24, exclusive.

3.4 Gráficos Estatísticos

A representação gráfica é um complemento importante da apresentação dos dados estatísticos feitos em uma tabela. Ela facilita a visualização e interpretação dos resultados de uma amostra. Os principais tipos de gráficos são: gráfico de barras, de barras agrupadas, de colunas, de linhas, de setores circulares e o histograma. Para representar variáveis qualitativas, os gráficos mais utilizados são os gráficos de colunas e de setores circulares(gráfico de pizza).

Exemplo 7. *A Figura 2 mostra o gráfico de colunas e de setores para análise da variável qualitativa "habilitação", de acordo com a Tabela 1:*



(a) Gráfico de colunas (b) Gráfico de setores

Figura 2: Gráfico de Colunas e Setores para Variável qualitativa

3.4.1 Gráfico de barras

Tem por finalidade comparar grandezas, por meio de retângulos de igual largura e comprimentos proporcionais às respectivas grandezas.

Exemplo 8. *A Figura 3 traz o gráfico de barras com as dez infrações mais cometidas em Goiânia no ano de 2017.*

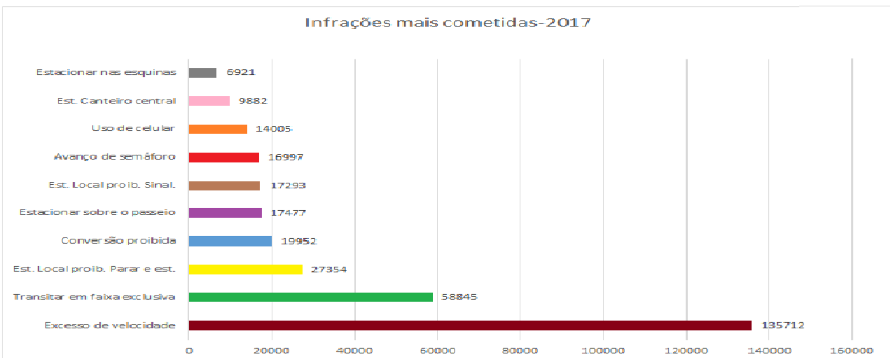


Figura 3: Gráfico de barras: Infrações mais cometidas em 2017/*(fonte: própria com dados da SMT)*

Exemplo 9. *A Figura 4 mostra o número de veículos autuados por excesso de velocidade no período de Janeiro a Outubro de 2017 em Goiânia:*

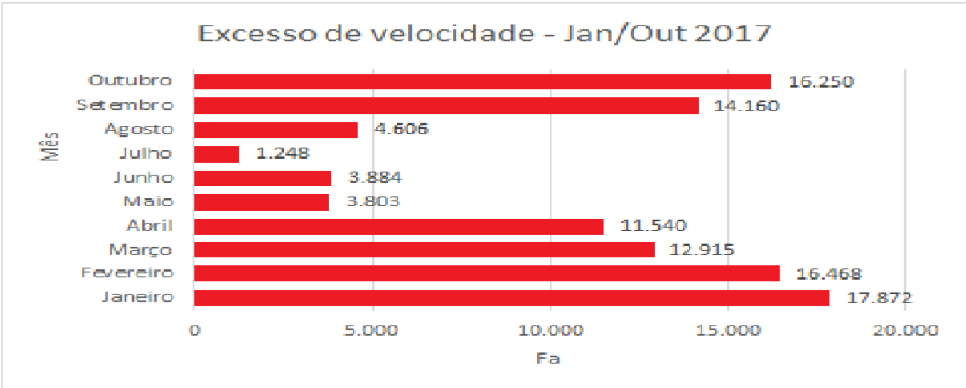


Figura 4: Gráfico de barras: Excesso de velocidade/*(fonte: própria com dados da SMT)*

3.4.2 Gráfico de barras agrupadas

É um gráfico que para cada valor da variável aparece um grupo de barras. Tem a vantagem de permitir comparar diferentes grupo de dados para os mesmo valores de variável e a desvantagem de não poder ser utilizado para variáveis que apresentam muitas modalidades.

Exemplo 10. *A Figura 5 mostra o gráfico de barras agrupadas com o número de multas aplicadas manualmente e eletrônicamente por avanço de semáforo no primeiro trimestre de 2018 em Goiânia:*

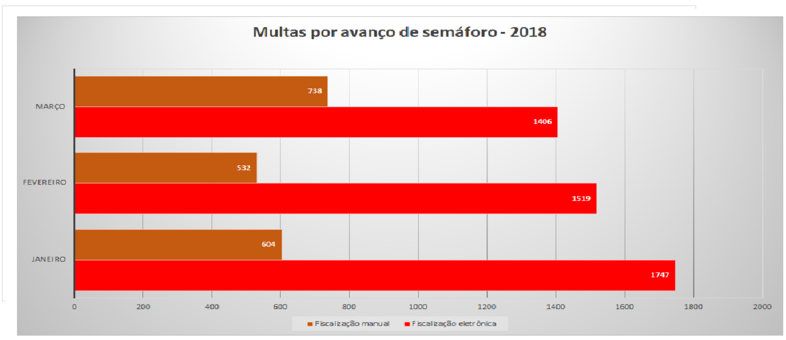


Figura 5: Gráfico de barras agrupadas/*(fonte: própria com dados da SMT)*

Exemplo 11. *A Figura 6 traz o número de acidentes (com ou sem vítimas) nos anos de 2016 e 2017 em Goiânia:*

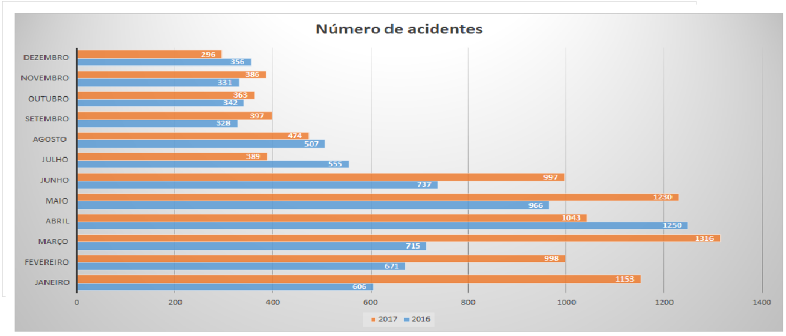


Figura 6: Gráfico de barras agrupadas/*(fonte: própria com dados da SMT)*

3.4.3 Gráfico de colunas

Os gráficos de colunas (ou barras verticais) têm a mesma finalidade dos gráficos de barras horizontais, sendo, no entanto, preferíveis quando se quer uma representação, análise e interpretação de dados relacionados com séries de tempo como, por exemplo, o número de veículos caracterizados como roubados/furtados no estado de Goiás nos anos de 2014 a 2018.

Exemplo 12. *A Figura 7 mostra o gráfico de colunas que ilustra as dez infrações mais cometidas em 2017 na cidade de Goiânia:*

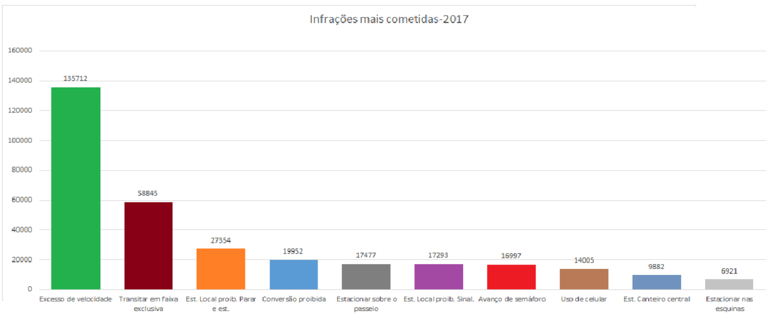


Figura 7: Gráfico de Colunas/*(fonte: própria com dados da SMT)*

Exemplo 13. *A Figura 8 ilustra as dez infrações mais cometidas em Janeiro de 2017 em Goiânia:*

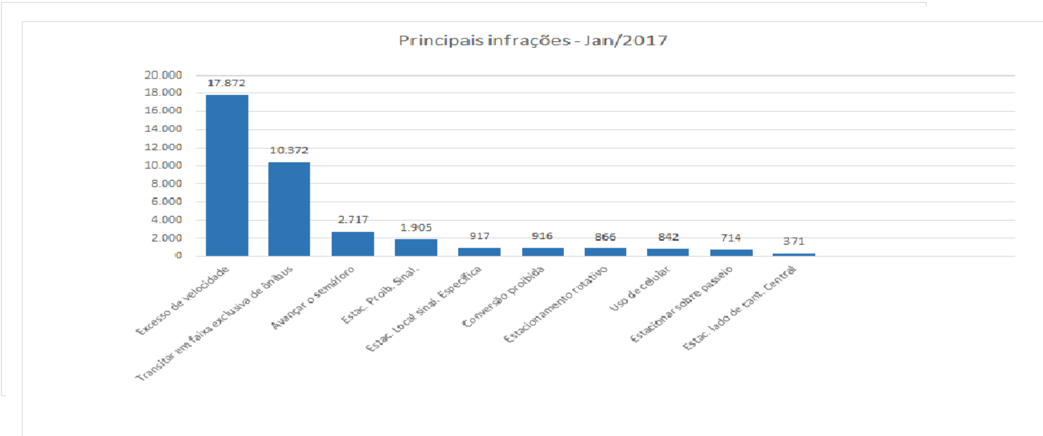


Figura 8: Gráfico de Colunas/*(fonte: própria com dados da SMT)*

3.4.4 Gráfico de linhas

Os gráficos de linhas são frequentemente usados para a representação de séries de tempo (quando um dos fatores for o tempo), isto porque quando a série cobre um grande número de períodos de tempo, a representação dos valores através de colunas pode conduzir a uma excessiva concentração de dados. As linhas são mais eficientes que as colunas quando existem intensas flutuações nas séries ou quando há necessidade de se representarem várias séries em um mesmo gráfico.

Exemplo 14. *A seguir, a Figura 9 traz o gráfico de linhas, referente ao número de acidentes por ano, de 1999 à 2013, em Goiás:*

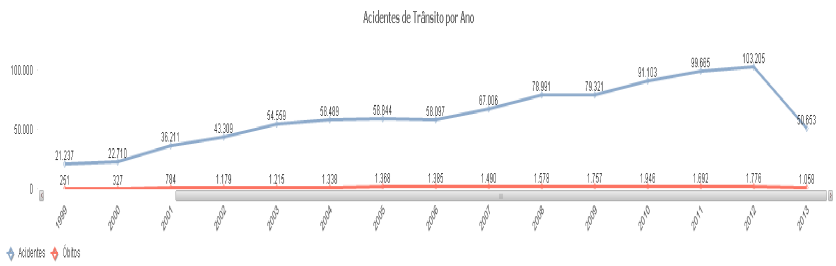


Figura 9: Gráfico de linhas/*(fonte: Detran-Go)*

Exemplo 15. *A Figura 10 traz o gráfico de linhas com o nùmero de autuações por avanço de semáforo no período de Janeiro a Outubro de 2017:*

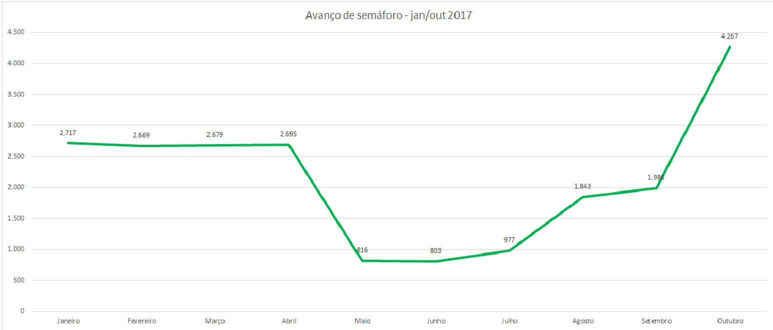


Figura 10: Gráfico de linhas/*(fonte: própria com dados da SMT)*

3.4.5 Gráfico de setores circulares

Os gráficos de setores (ou setogramas),também chamados de gráfico de pizza, são usados para representar valores absolutos ou porcentagens complementares. Para construir um gráfico de setores, partimos do pressuposto que o total de valores analisados, corresponde a 360º e cada classe analisada (setor) poderá ser expressa, em graus, por meio de uma regra de três simples. Se já dispomos de uma tabela de frequências, utilizamos os percentuais obtidos a partir da frequência relativa ***Fr***.

Exemplo 16. *Vamos construir o gráfico de setores circulares para a variável "habilitação" com base na Tabela 1:*

Cálculo do setor correspondente ao nº de menores:

30 - 360^o

12 - x^o

$x = \frac{12 \times 360}{25} = 144^o$ *que corresponde à 40% do circulo.*

Cálculo do setor correspondente ao nº de habilitados:

30 - 360^o

3 - x^o

$x = \frac{3 \times 360}{30} = 36^o$ *que corresponde à 10% do circulo.*

Cálculo do setor correspondente ao nº de não habilitados:

30 - 360^o

15 - x^o

$x = \frac{15 \times 360}{30} = 180^o$ *que corresponde à 50% do circulo.*

Após esses cálculos, com auxílio de um transferidor, fazemos a marcação dos ângu-

los correspondentes. Os valores percentuais são obtidos a partir da frequência relativa Fr. A Figura 11 ilustra o gráfico de setores:

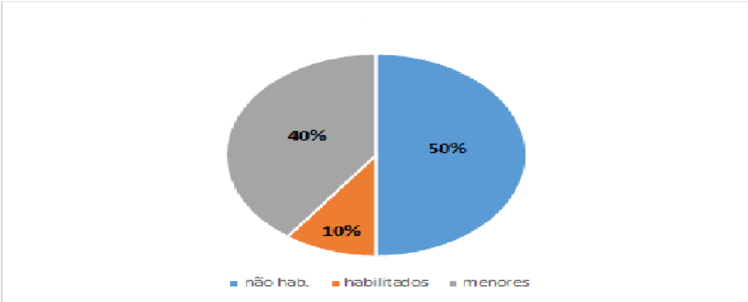


Figura 11: Gráfico de Setores

Exemplo 17. *A Figura 12 traz o gráfico de setores circulares com os números de multas aplicadas pelas principais infrações de veículos em trânsito, no período de Janeiro a Outubro de 2017:*

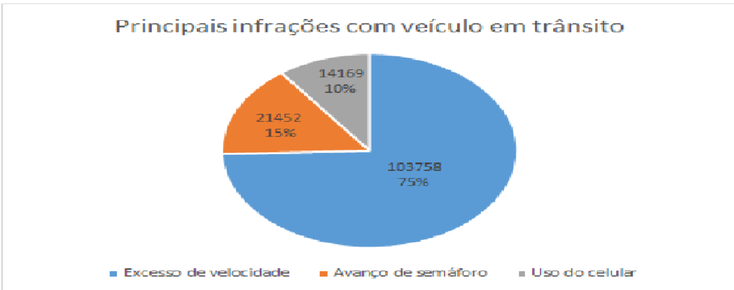


Figura 12: Gráfico de Setores/*(fonte: própria com dados da SMT)*

3.4.6 Histograma

O histograma é um gráfico formado por um conjunto de retângulos justapostos, de forma que a área de cada retângulo seja proporcional à frequência da classe que ele representa. É o tipo de gráfico mais usado para representar uma tabela de frequência com dados agrupados.

Exemplo 18. *A Figura 13, mostra o histograma que relaciona cada classe de idades com a frequência absoluta* (f_i)*, referentes à Tabela 1:*

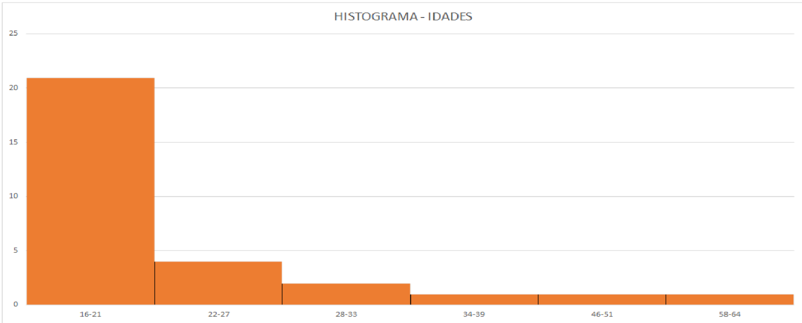


Figura 13: Histograma

3.4.7 Polígono de frequências

Unindo por linhas retas os pontos médios das bases superiores dos retangulos do histograma, obtém-se outra representação dos dados, denominada Polígono de Frequências.

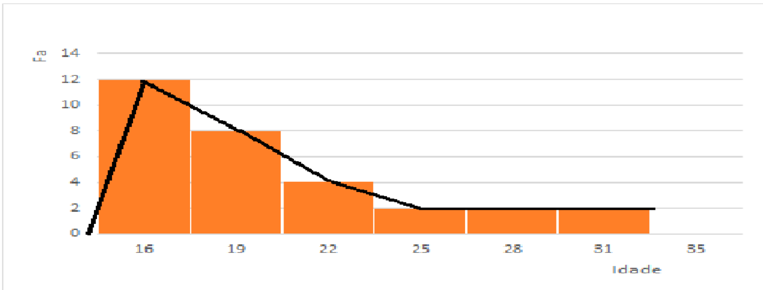


Figura 14: Polígono de frequência

3.5 Medidas de posição

Muitas vezes há a necessidade de condensar mais os dados de uma pesquisa, apresentando um ou mais valores que sejam representativos, da série toda. Porém, se usarmos somente um valor, obtemos uma redução drástica dos dados. Assim, de acordo com Marconi e Lakatos [9]:

> *As medidas de posição, também chamadas parâmetros de posição ou medidas de tendência central, constituem-se em um dos procedimentos para a redução dos dados , expressando valores que se encontram situados entre os extremos*

de uma série ou distribuição. Referem-se a dados não tabulados e a dados tabulados

Usualmente, emprega-se as chamadas medidas de posição central: **média**, **mediana** e a **moda**.

3.5.1 Média

A média é a medida de tendência central mais usada para descrever resumidamente uma distribuição e sua ideia geral é substituir uma lista de valores por um único valor, que preserve uma certa característica dessa lista. Existem vários tipos de médias: média aritmética, média geométrica, média harmônica, média quadrática, média cúbica e média biquadrada. Neste trabalho, abordaremos apenas as principais médias, a saber: aritmética, geométrica e harmônica.

Média Aritmética(A)

A média aritmética de um conjunto de valores pode ser de dois tipos: simples ou ponderada.

Definição 7. *Dada uma lista de n valores $x_1, x_2, ..., x_n$, a média aritmética **A** é dada pelo quociente entre a soma dos valores dessa lista e o número total de valores e tem a caracteristica de preservar a soma desses valores. Simbolicamente, temos:*

$$A = \frac{x_1 + x_2 + ... + x_n}{n} = \frac{\sum_{i=1}^{n} x_i}{n} \tag{2}$$

em que x_i é valor genérico da observação, x_n é o n-ésimo valor observado e n é número de observações.

Exemplo 19. *Com base na Tabela 4, vamos calcular a média aritmética do número de acidentes no município de Goiânia em 2016:* ***Solução:***
De acordo com a tabela, temos que:

$$A = \frac{606+671+715+1250+966+737+55+507+328+342+331+356}{12} = \frac{7364}{12} = 613,6$$

Portanto, a média mensal foi de 613,6 acidentes em 2016.

Note que, a soma dos valores da tabela é 7364 e se substituirmos cada valor da tabela pela média, teremos como resultado a mesma soma (7364), preservando, assim, a característica da soma.

mês	nº de acidentes
janeiro	606
fevereiro	671
março	715
abril	1250
maio	966
junho	737
julho	555
agosto	507
setembro	328
outubro	342
novembro	331
dezembro	356

Tabela 4: Acidentes em 2016/*(fonte: SMT)*

Definição 8. *A média aritmética ponderada (ou simplesmente média ponderada) é utilizada quando os valores do conjunto tiverem pesos diferentes e é dada pelo quociente entre o produto dos valores da variável pelos seus respectivos pesos e a soma desses pesos. Simbolicamente, temos:*

$$A = \frac{x_1p_1 + x_2p_2 + ... + x_np_n}{p_1 + p_2 + ...p_n} \tag{3}$$

Exemplo 20. *De acordo com o CTB, as infrações são classificadas e pontuadas quanto à sua natureza: leve (3 pontos), média (4 pontos), grave (5 pontos) e gravíssima (7 pontos). Um certo condutor cometeu, no período de 12 meses, 15 infrações, sendo elas: 2 gravíssimas, 3 graves, 4 médias e 6 leves. Vamos calcular a média ponderada desses pontos considerando, como pesos, os pontos referentes à cada tipo de infração:*
Solução: *Aplicando a fórmula 3, temos:*

$$A = \frac{2\times7+3\times5+4\times4+6\times3}{7+5+4+3} = \frac{14+15+16+18}{19} = \frac{63}{19} = 3,31$$

Logo, a média de pontos por infração desse condutor é de 3,31.

Normalmente, para o cálculo da média ponderada recorre-se a uma tabela de dados tabulados, o que possibilita maior rapidez de operação e organização dos valores.

Exemplo 21. *A quantidade de autos de infração lavrados por vinte agentes de trânsito em um dia de serviço foram os seguintes, já em forma de rol,:*

8,8,8,8,9,9,10,10,10,10,10,10,12,12,12,13,13,15,15,15

Calcular a média de autos por agente para esse dia.

Nº de autos(x_i)	Frequência absoluta(f_i)	$x_i.f_i$
8	4	32
9	2	18
10	6	60
12	3	36
13	2	26
15	3	45
	$\sum_{i=1}^{6} fi = 20 = n$	$\sum_{i=1}^{6} x_i.f_i = 217$

Tabela 5: Média aritmética para dados tabulados

Solução:

$$A = \frac{\sum_{i=1}^{k} x_i.f_i}{n} = \frac{\sum_{i=1}^{6} x_i.f_i}{20} = \frac{217}{20} = 10,85.$$

Logo, a média de autos por agente nesse dia foi de 10,85.

Se, por algum motivo, não se tiver acesso aos dados de uma amostra, mas apenas a sua tabela de freqüências ou ao seu histograma, não será possível calcular exatamente os valores da sua média. Assim, o melhor que se pode fazer é calculá-las aproximadamente. Neste caso, a tabela requer mais uma coluna, necessária para dispor os pontos médios das classes, como na Tabela 4 do Exemplo 19:

Exemplo 22. *A quantidade de autos de infração lavrados por vinte agentes de trânsito em uma semana de trabalho estão representados na Tabela 6. Determinar a média de autos de infração por agente nessa semana.*

Classes	Frequência absoluta(f_i)	Ponto médio(x_j)	$f_i.x_j$
$0 \vdash 10$	3	5	15
$10 \vdash 20$	9	15	135
$20 \vdash 30$	5	25	125
$30 \vdash 40$	2	35	70
$40 \vdash 50$	1	45	45
	$\sum_{i=1}^{5} fi = 20 = n$		$\sum_{i=1}^{5} f_i.x_j = 390$

Tabela 6: Média aritmética para dados agrupados

$$A = \frac{\sum_{i=1}^{k} f_i.x_j}{n} = \frac{\sum_{i=1}^{5} f_i.x_j}{20} = \frac{390}{20} = 19,5.$$

Portanto, a média de autos por agente, nessa semana, foi de 19,5.

Temos duas principais propriedades da média aritmética, que agrupamos na proposição a seguir.

Proposição 1. *Seja $x_1, \ldots, x_n$ uma lista de valores. Seja $\bar{x}$ a média aritmética destes valores. Então,*
(a) A soma algébrica dos desvios de um conjunto de números tomados em relação à média aritmética é zero;
(b) Multiplicando-se cada elemento de um conjunto de números por um valor constante e arbitrário, a média fica multiplicada por essa constante.
Demonstração: *Na demonstração do item (a), considerando dados brutos, temos:*

$$\sum_{i=1}^{n} d_i = \sum(x_i - \overline{x}) = \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} \overline{x}$$

Como $\overline{x}$ é uma constante, para um dado conjunto de valores, temos que $\sum_{i=1}^{n} \overline{x} = n\overline{x}$. Más, $\overline{x} = \dfrac{\sum_{i=1}^{n} \overline{x}}{n}$ e, assim, $n\overline{x} = n\dfrac{\sum_{i=1}^{n} x_i}{n} = \sum_{i=1}^{n} x_i$. Então,

$$\sum_{i=1}^{n} d_i = \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} \overline{x} = \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} x_i = 0.$$

Na demonstração do item (b), considere c uma constante arbitrária. Façamos $Y = cX$. Devemos ter:

$$\overline{y} = \frac{\sum_{i=1}^{n} y_i}{n} = \frac{\sum_{i=1}^{n} cx_i}{n} = c\frac{\sum_{i=1}^{n} x_i}{n} = c\overline{x} \Rightarrow \overline{y} = c\overline{x}.$$

Média Geométrica (G)

A média geométrica de n valores positivos é definida, genericamente, como a raiz n-ésima do produto de todos esses n valores e tem a característica de preservar o produto desses valores. É muito utilizada na matemática financeira sendo a mais indicada quando estamos analisando situações envolvendo aumentos sucessivos e uma de suas aplicações é nos problemas em que se deseja calcular as taxas médias.

Definição 9. *Dados os valores $x_1, x_2, ..., x_n$, a média geométrica desses valores é dada por:*

$$G = \sqrt[n]{x_1 \times x_2 \times ... \times x_n} \tag{4}$$

Exemplo 23. *Vamos calcular a média geométrica entre os valores dos pontos assinalados na CNH de acordo com a natureza da infração.*
Já sabemos que esses valores são 3,4,5 e 7. Então, temos que:

$$G = \sqrt[4]{3 \times 4 \times 5 \times 7} = \sqrt[4]{420} \cong 4,527$$

$\therefore$ *a média geométrica entre os números 3,4,5 e 7 é aproximadamente 4,527.*

Note que, o produto dos valores é 420 e se substituirmos cada valor pela média geométrica, teremos como resultado o mesmo produto(420), preservando, assim, a característica do produto.

Exemplo 24. *A tabela abaixo mostra o número de veículos registrados pelo DETRAN-GO nos anos de 2014 a 2017. Qual a taxa média anual de aumento dessa frota?*

Ano	nº de veículos
2014	3.428.168
2015	3.585.854
2016	3.686.343
2017	3.798.186

Tabela 7: Frota de veículos em Goiás/*(fonte: DETRAN-GO)*

Solução: *Da tabela, obtemos as taxas de aumento anuais:*

De 2014 a 2015 o aumento foi de 4,6%, de 2015 a 2016 de 2,8% e de 2016 a 2017 de 3,03%. Suponha uma quantia inicial x. Se a frota aumenta 4,6% no primeiro ano,

temos: $x + 0,046 \cdot x = 1{,}046 \cdot x$. Assim $1{,}046 \cdot x$ é quantia sobre a qual incidirá o aumento do próximo ano, 2,8%. Assim temos: $1,046 \cdot x + 0,028 \cdot 1,046 \cdot x = 1,046 \cdot x \cdot 1,028 = 1,075288 \cdot x$. Analogamente, 1,075288.x é a quantia a qual incidirá o aumento do próximo ano, 3,03%. Temos então:
1,07288x + 0,0303.1,07288.x = 1,07288.x.1,0303 = 1,1054.x, ou ainda:

$$1{,}046.x.1{,}028.1{,}0303 = 1{,}1054.x.$$

Portanto, o aumento de 2014 a 2017 foi de 10,54%. A taxa média i, é aquela que deve ser aplicada igualmente em cada mes a fim de produzir o mesmo aumento no período, de 10,54%. Desse modo, para encontrá-la, basta efetuarmos os mesmos cálculos substituindo as taxas mensais 4,6% , 2,8% e 3,03% por i%. Supondo a mesma quantidade inicial x, se ela aumenta i% no primeiro mês, teremos $x + i \cdot x = x \cdot (1 + i)$. Se essa nova quantia aumenta i%, teremos $x.(1+i)+i.x.(1+i) = x.(1+i).(1+i) = x.(1+i)^2$ e, se essa última quantia aumenta i% , então teremos: $x.(1+i)^2 + i.x.(1+i)^2 = x.(1+i)^2(1+i) = x.(1+i)^3$ que, como mostrado acima, tem que ser igual a 1,1054.x. Assim, teremos:

$$\begin{gathered} x.(1+i)^3 = 1,046.x.1,028.1,0303 \\ (1+i)^3 = 1,046.1,028.1,0303 \\ 1+i = \sqrt[3]{1,046.1,028.1,0303} \cong 1,034 \\ i \cong 1,034 - 1 \\ i \cong 0,034 = 3,4\% \end{gathered}$$

Portanto, a taxa média é de 3,4%.

Nota: Observe que na terceira igualdade, temos a média geométrica dos três aumentos sucessivos, com isso, conclui-se que a taxa média é a média geométrica das taxas acrescidas de uma unidade, de cujo resultado se subtrai um. Um importante resultado que demonstraremos a seguir é a chamada desigualdade das médias (aritmética e geométrica).

Teorema 1. *(desigualdade das médias) Se $x_1, x_2, ..., x_n$ são números reais positivos e A e G, são respectivamente suas médias aritmética e geométrica, então $A \geq G$. Além do mais, a igualdade ocorre se, e somente se, $x_1 = x_2 = ... = x_n$.*
Demonstração: *Devemos provar que $\dfrac{x_1 + x_2 + \ldots + x_n}{n} \geq \sqrt[n]{x_1 \cdot x_2 \cdot \ldots \cdot x_n}$ ocorrendo a igualdade se, e só se, $x_1 = x_2 = \ldots = x_n$.*
*Primeiro, vamos provar para **n = 2:**(caso inicial)*

$$A-G=\frac{x_1+x_2}{2}-\sqrt{x_1.x_2}=\frac{x_1+x_2-2.\sqrt{x_1.x_2}}{2}=\frac{(\sqrt{x_1}-\sqrt{x_2})^2}{2}\geq 0 \Rightarrow A\geq G$$

Além disso, a igualdade só ocorre quando $x_1 = x_2$.

Agora, provaremos quando $\boldsymbol{n = 2^k}$*:*

Usaremos indução sobre k, $k \geq 1$ *e consideremos* $n = 2^k$ *a quantidade dos números em que aplicaremos nossas médias. O caso k = 1 já foi provado.*

Hipótese de Indução: Suponha que a desigualdade seja verdade para $n = 2^k$*. Vamos ver o que ocorre quando* $n = 2^{k+1}$.

Para 2^{k+1} *, temos uma lista de 2n números. Então, temos que:*

$$\begin{aligned}
A &= \frac{x_1+x_2+\ldots+x_n+x_{n+1}+x_{n+2}\ldots+x_{2n}}{2n}\\
&= \frac{\dfrac{x_1+x_2+\ldots+x_n}{n}+\dfrac{x_{n+1}+x_{n+2}\ldots+x_{2n}}{n}}{2}\\
&\geq \frac{\sqrt[n]{x_1\cdot x_2\cdot\ldots\cdot x_n}+\sqrt[n]{x_{n+1}\cdot x_{n+2}\cdot\ldots\cdot x_{2n}}}{2}\\
&\geq \sqrt{\sqrt[n]{x_1\cdot x_2\cdot\ldots\cdot x_n}.\sqrt[n]{x_{n+1}\cdot x_{n+2}\cdot\ldots\cdot x_{2n}}}\\
&= \sqrt[2n]{x_1\cdot x_2\cdot\ldots\cdot x_n\cdot x_{n+1}\cdot x_{n+2}\cdot\ldots\cdot x_{2n}}=G
\end{aligned}$$

Note que, na primeira desigualdade usamos a hipótese de indução e, na segunda desigualdade, usamos o caso inicial.

Além disso, a igualdade só ocorre quando

$$x_1=x_2=\ldots=x_n, x_{n+1}=x_{n+2}=\ldots=x_{2n} \text{ e } x_1\cdot x_2\cdot\ldots\cdot x_n=x_{n+1}\cdot x_{n+2}\cdot\ldots\cdot x_{2n}$$

ou seja, quando

$$x_1=x_2=\ldots=x_n=x_{n+1}=x_{n+2}=\ldots=x_{2n}$$

Logo, a propriedade é válida também para $2n = 2^{k+1}$ *e, portanto a desigualdade das médias vale para todo n da forma* 2^k.

Caso geral: *A prova consiste em completar a lista com valores iguais à media aritmética dos n números, até o número de valores tornar-se uma potência de 2.*

Ilustramos para n = 3, quando completamos a lista x_1, x_2, x_3 *com um quarto valor igual a* $A = \dfrac{x_1+x_2+x_3}{3}$.

$$\begin{aligned}
A = \frac{x_1+x_2+x_3+A}{4} &\geq \sqrt[4]{x_1.x_2.x_3.A} \Rightarrow\\
A^4 &\geq x_1.x_2.x_3.A \Rightarrow
\end{aligned}$$

$$A^3 \geq x_1.x_2.x_3 \Rightarrow$$
$$A \geq \sqrt[3]{x_1.x_2.x_3} = G.$$

Além disso, a igualdade só vale quando $x_1 = x_2 = x_3$.
De maneira análoga, provamos para qualquer valor de n, o que completa a prova no caso geral.

■

Média Harmônica (H)

A média harmônica está relacionada ao cálculo matemático em situações envolvendo as grandezas inversamente proporcionais. A média harmônica de um conjunto de valores x_i é o inverso da média aritmética dos inversos.

Definição 10. *Dados os valores $x_1, \ldots, x_n$, definimos a média harmônica desses valores como a quantidade*

$$H = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + ... + \frac{1}{x_n}} \tag{5}$$

Exemplo 25. *Como no exemplo 23, vamos calcular agora, a média harmôncia entre os valores dos pontos assinalados na CNH:*

$$H = \frac{4}{\frac{1}{3}+\frac{1}{4}+\frac{1}{5}+\frac{1}{7}} = \frac{4}{\frac{389}{420}} \cong 4,318$$

$\therefore$ *a média harmônica entre os números 3,4,5 e 7 é aproximadamente 4,318.*

Uma importante observação sobre essas médias é a seguinte relação entre elas:
Se $x_1, x_2, x_3, ..., x_n$ são números positivos e A, G e H são suas médias aritmética, geométrica e harmônica, respectivamente, então $A \geq G \geq H$. Além disso, duas quaisquer dessas médias são iguais se, e somente se, os números da lista são todos iguais, $x_1 = x_2 = x_3 = ... = x_n$

3.5.2 Moda (Mo)

A moda (também chamada de norma, valor típico ou valor dominante) é uma outra medida de centralidade. É o valor que ocorre com maior frequência ou o valor mais comum em um conjunto de dados. Os valores da varável dispostos em uma tabela de frequências podem apresentar-se individualmente ou agrupados em classes. No

primeiro caso, a determinação da moda é imediata, bastando, para isso, consultar a tabela, localizando o valor que apresenta a maior frequência. Esse valor será a moda do conjunto. Assim, tomando como exemplo a Tabela 1, a moda para a variável idade é 18, pois é o valor que aparece com maior frequência (11 vezes). Logo, $Mo = 18$.
Tratando-se de uma tabela de frequências com valores agrupados em classes, o procedimento não é imediato, sendo disponíveis alguns métodos de cálculos distintos. Neste trabalho, apresentaremos dois desses métodos:
Moda Bruta: O método mais rudimentar de cálculo da moda em tabelas de frequências com valores agrupados em classes consiste em tomar o ponto médio da classe modal. Esse valor recebe o nome de moda bruta. Examinando os dados da Tabela 6, por exemplo, podemos dizer que a segunda classe é a classe modal e a moda bruta será seu ponto médio: $Mo = 15$.
O método que apresentaremos a seguir é mais elaborado e baseia-se não apenas na frequência da classe modal, mas tambem nas frequências das classes adjacentes.
Método de King: O método de King baseia-se na influência das frequências das classes adjacentes sobre a classe modal. Considerando essas frequências, está se admitindo implicitamente que a moda se desloca dentro do intervalo de classe para um determinado ponto (valor), de tal sorte que as distâncias desse ponto aos limites de classe sejam inversamente proporcionais às frequências das respectivas classes adjacentes. O cálculo da moda pelo método de King é dado por:

$$Mo = l + c.(\frac{f_{post}}{f_{ant} + f_{post}}), \tag{6}$$

em que l é o limite inferior da classe modal, c é a amplitude do intervalo de classe, f_{ant} é a frequência da classe anterior à classe modal, e f_{post} é a frequência da classe posterior à classe modal.

Exemplo 26. *Calcular, pelo método de King, a moda dos valores constantes na Tabela 8.*
Solução:
A classe modal é a segunda: $10 \vdash 20$. *A moda, segundo a fórmula de King, será:*

$$Mo = l + c.\frac{f_{post}}{f_{ant} + f_{post}} = 10 + 10.\frac{5}{3+5} = 10 + \frac{50}{8} = 16,25$$

A moda bruta desses valores é ligeiramente menor: Mo = 16.

Classes	Frequência absoluta(f_i)
$0 \vdash 10$	3
$10 \vdash 20$	9
$20 \vdash 30$	5
$30 \vdash 40$	2
$40 \vdash 50$	1
	n = 20

Tabela 8: Moda para dados agrupados

3.5.3 Mediana(Md)

A mediana é a terceira medida de tendência central e pode ser definida como o valor que divide uma série ordenada(em forma de rol) de tal forma que metade dos itens sejam iguais ou menores do que ela e que a outra metade dos itens sejam maiores ou iguais do que ela. Em termos mais simples, a mediana pode ser o valor do meio de um conjunto de dados. A vantagem da mediana em relação à média é que a mediana pode dar uma ideia melhor de um valor típico porque não é tão distorcida por valores extremamente altos ou baixos. Assim, considere $x_1, x_2, ...x_n$, os n valores de uma variável X ordenados em ordem crescente. Vamos denotar a menor observação por $x_{(1)}$, a segunda por $x_{(2)}$ e assim por diante, obtendo-se

$$x_{(1)} \leq x_{(2)} \leq ... \leq x_{(n-1)} \leq x_{(n)}. \tag{7}$$

Com esta notação, a mediana da variável X pode ser definida como:

$$Md(X) = \begin{cases} x_{(\frac{n+1}{2})}, \text{ se } n \text{ ímpar;} \\ \frac{x_{(\frac{n}{2})} + x_{(\frac{n}{2}+1)}}{2}, \text{ se } n \text{ par.} \end{cases} \tag{8}$$

As observações ordenadas como em (8) são chamadas *estatísticas de ordem.*

Exemplo 27. *A seguir, estão os dados brutos com as idades dos alunos da turma selecionada. Vamos determinar a mediana(ou idade mediana)da turma:*
63,19,19,19,16,16,18,16,16,28,22,32,19,20,25,34,17,18,17,23,16,18,22,48,16,16,18,16,17,17
Solução:
Primeiro, vamos ordenar em forma de rol:

16,16,16,16,16,16,16,16,17,17,17,17,18,18,18,18,19,19,19,19,20,22,22,23,25,28,32,34,42,63.
Como o conjunto desses dados é um número par($n = 30$), temos que a mediana é:

$$Md = \frac{18+18}{2} = 18$$

Portanto, a mediana(ou idade mediana) da turma é 18.

Quando os valores da variável estiverem já tabulados, o procedimento a ser adotado será praticamente idêntico ao anterior. Primeiro, deve-se verificar se o número de observações é par ou ímpar, em seguida, analisamos a ordem da série com o auxilio da coluna das frequências acumuladas.

Exemplo 28. *Vamos determinar a mediana do conjunto de valores indicados na Tabela 9:*

Nº de autos(x_i)	Frequência absoluta(f_i)	Fac
8	4	4
9	2	6
10	6	12
12	3	15
13	2	17
15	3	20
	n = 20	

Tabela 9: Mediana para dados tabulados

Solução:
Como o número de observações é par($n = 20$) e, de acordo com a definição de mediana, devemos ter dois elementos centrais, o décimo e o décimo primeiro. Analisando a coluna das frequências acumuladas, verificamos que o valor mediano está na 3º classe. Como a mediana é a média aritmética entre os elementos centrais da série, segue que

$$Md = \frac{10+10}{2} = 10.$$

Logo, a mediana é 10.

Mediana a partir do histograma

Quando os valores da variável estiverem apresentados por um histograma, a mediana será o valor de x(eixo das abscissas) que corresponde à perpendicular que divide o histograma em duas partes que apresentam áreas iguais.

Exemplo 29. *No gráfico abaixo, o histograma apresenta-se dividido pela linha tracejada OP em duas partes exatamente iguais quanto às suas áreas.*

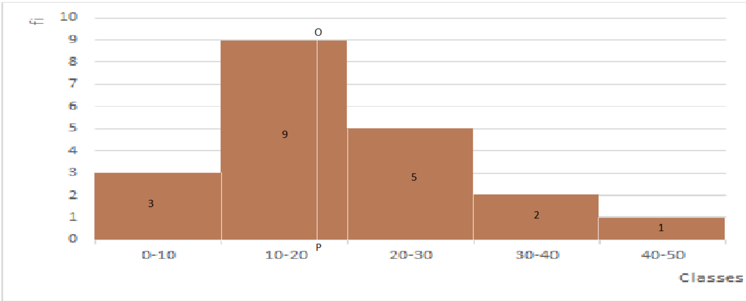


Figura 15: Mediana a partir do histograma

A mediana será, então, a abscissa correspondente à linha OP, a qual divide o histograma em duas partes iguais. Em virtude de as áreas dos retângulos do histograma corresponderem às frequências das respectivas classes, OP será tal que a soma das áreas situadas à sua esquerda ou à sua direita será sempre igual à metade da frequência total, ou seja, igual ao elemento mediano. Examinando o gráfico, podemos ver que

$$n = 3 + 9 + 5 + 2 + 1 = 20$$

e que

$$E_{Md} = \frac{n}{2} = 10.$$

Dessa forma, podemos afirmar que a linha OP deverá dividir o histograma em duas partes, sendo que a área de cada uma é igual a 10. Observando o gráfico e, sabendo que $E_{Md} = 10$, *vemos que para chegar à linha OP, pela esquerda, faltam 7 unidades, pois 3 + 7 = 10, que é o elemento mediano. Isso faz com que tenhamos que calcular proporcionalmente a frequência dessa classe, cuja área e 9. Para isso, devemos ter* $\frac{7}{9}$ *valores da classe mediana e, uma vez que a amplitude do intervalo de classe é igual a 10, para chegarmos à mediana devemos ter:*

$$\frac{7}{9}.10 = 7,7...$$

Assim, Md = 10 + 7,7 = 17,7.

Uma outra maneira de calcular a mediana quando os valores da variável estiverem agrupados em classes, é por meio do polígono de frequências acumuladas, o qual gera a seguinte fórmula para o cálculo:

$$Md = l + c.\frac{E_{Md} - F_{ant}}{f_{Md}}, \tag{9}$$

em que $E_{Md} = \dfrac{N}{2}$ é elemento mediano, l é o limite inferior da classe mediana, c é a amplitude do intervalo de classe, f_{Md} é a frequência simples da classe mediana e F_{ant} é a frequência acumulada até a classe anterior à classe mediana.

Exemplo 30. *Calcular o número mediano de autos de infração por agente de acordo com a Tabela 10.*

Solução:

Classes	Frequência absoluta(f_i)	Fac
$0 \vdash 10$	3	3
$10 \vdash 20$	9	12
$20 \vdash 30$	5	17
$30 \vdash 40$	2	19
$40 \vdash 50$	1	20
	n = 20	

Tabela 10: Mediana para dados agrupados

De acordo com a tabela, temos que:

$$E_{Md} = 10,\ l = 10,\ c = 10,\ f_{Md} = 9,\ F_{ant} = 3$$

Com o uso da fórmula 6, segue que:

$$Md = l + c.\frac{E_{Md} - F_{ant}}{f_{Md}} = 10 + 10\frac{10-3}{9} = 10 + \frac{70}{9} \cong 17,7.$$

Por conseguinte, o número mediano de autos é aproximadamente 17,7.

3.5.4 Quartis, Decis e Centis

Há uma série de medidas de posição semelhantes na sua concepção à mediana, embora não sejam medidas de tendência central. Como foi definido antes, a mediana divide a distribuição em duas partes iguais quanto ao número de elementos de cada parte. Já os Quartis, permitem dividir a distribuição em quatro partes iguais quanto ao número de elementos de cada uma; os Decis, em dez partes e os Centis (ou Percentis) em cem partes iguais. Simbolicamente, temos:

$Q_i = quartis$ com $i = 1, 2, 3.$
$D_i = decis$ com $i = 1, 2, 3, ..., 9.$
$C_i = centis$ com $i = 1, 2, 3, ..., 99.$

Dessa forma, para dividir uma série ordenada de valores em quatro partes iguais, precisamos de três separatrizes(quartis); para dividi-la em dez, precisaremos de nove separatrizes(decis); e em cem, precisaremos de noventa e nove separatrizes(centis).

Quartis - Q_i

Primeiro Quartil - Q_1

Dado um conjunto, em ordem crescente de valores, o primeiro quartil, Q_1, é o valor que divide o conjunto em duas partes tais que um quarto ou vinte e cinco por cento dos valores sejam menores do que ele e três quartos ou setenta e cinco por cento dos restantes sejam maiores. O elemento que indica a ordem ou posição do primeiro quartil é determinado, para dados agrupados, pela seguinte expressão:

$$E_{Q_1} = \frac{n}{4}$$

em que n é o número de valores do conjunto, ou número de observações.

Segundo Quartil - Q_2

Dado um conjunto, em ordem crescente de valores, o segundo quartil ou mediana, é o valor que o divide em duas partes iguais quanto ao número de elementos, isto é, cinquenta por cento dos valores do conjunto são menores, e cinquenta por cento são maiores do que ele.

$$E_{Q_2} = \frac{2n}{4} = \frac{n}{2}.$$

Terceiro Quartil - Q_3

Dado um conjunto, em ordem crescente de valores, o terceiro quartil é o valor que divide o conjunto em duas partes tais que setenta e cinco por cento ou três quartos dos valores sejam menores e vinte e cinco por cento ou um quarto sejam maiores do que ele. O elemento que indica a ordem em que n encontra o terceiro quartil é calculado como segue:

$$E_{Q_3} = \frac{3n}{4.}$$

Genericamente, para determinar a ordem ou posição do quartil a ser calculado, usaremos a seguinte expressão:

$$E_{Q_i} = \frac{in}{4}.$$

em que i é o número do quartil a ser calculado e n é o número de observações.

Decis - D_i

A definição dos decis obedece ao mesmo pricípio da dos quartis, com a modificação da porcentagem de valores que ficam aquém e além do decil que se pretenda calcular.

Primeiro Decil - D_1

O primeiro decil de um conjunto em ordem crescente de valores é o valor que divide um conjunto em duas partes tais que dez por cento ou um décimo dos valores sejam menores e nove décimos ou noventa por cento sejam maiores do que ele. O elemento que indica a posição do segundo decil é calculado pela expressão:

$$E_{D_1} = \frac{n}{10}.$$

Segundo Decil - D_2

Trata-se do valor que divide o conjunto em duas partes tais que, vinte por cento ou dois décimos dos valores sejam menores e oitenta por cento ou oito décimos dos valores sejam maiores. Para saber a ordem do segundo decil, usamos a expressão:

$$E_{D_2} = \frac{2n}{10}.$$

Entre todos os decis, o de maior interesse é o quinto decil, o qual divide o conjunto de valores em duas partes tais que, metade dos valores sejam menores e a outra metade, maiores do que ele. Assim sendo, o quinto decil é igual ao segundo quartil, que por sua vez é igual à mediana. O elemento que indica a ordem do quinto decil é igual ao elemento mediano, ou seja:

$$E_{D_5} = \frac{5n}{10} = \frac{n}{2} = \frac{2n}{4}.$$

Ante ao exposto, podemos afirmar, então que:

$$Md = D_5 = Q_2.$$

De forma geral, para calcular os decis, recorremos à seguinte expressão que define a ordem em que o decil se encontra:

$$E_{D_i} = \frac{in}{10}.$$

em que n é o número de valores observados e i é o número que identifica o decil a ser calculado.

Centis ou Percentis - C_i

No caso dos centis(ou percentis), cada parte em que foram subdivididos os valores do conjunto, através dos noventa e nove centis, contará com um centésimo ou um por cento dos valores do conjunto. O elemento que definirá a ordem do centil, em uma distribuição de frequências de valores tabulados agrupados em classes, será encontrado pelo emprego da expressão:

$$E_{C_i} = \frac{in}{100}.$$

em que i é o número identificador do centil e n é o número total de observações. Note que os centis englobam todos os decis e quartis.

Exemplo 31. *O Vigésimo Centil é igual ao segundo decil, porque:*

$$E_{C_{20}} = \frac{20 \times n}{100} = 0,2n = E_{D_2} = \frac{2n}{10} = 0,2n.$$

Exemplo 32. *O vigésimo quinto centil é igual ao primeiro quartil, pois:*

$$E_{C_{25}} = \frac{25n}{100} = \frac{n}{4} = 0,25n$$

$$E_{Q_1} = \frac{1n}{4} = 0,25n$$

$$\therefore E_{C_{25}} = E_{Q_1}.$$

3.6 Medidas de Dispersão

Os fenômenos analisados estatisticamente caracterizam-se tanto pela sua semelhança quanto pela sua variabilidade. Não tem sentido algum se calcular a média de um conjunto de dados onde não haja variação desses elementos. No entanto, se a variabilidade desses dados for muito grande, sua média terá um grau de confiabilidade tão pequeno que será inútil calculá-la. Caracterizar um conjunto de valores apenas através de uma média, por exemplo, é descreve-lo inadequadamente, uma vez que os dados diferem entre si, em maior ou menor grau. Há situações em que as medidas de tendência central, como a média, a moda e a mediana, não são as mais adequadas para a análise de uma amostra de valores. Nesses casos, é necessário utilizar as chamadas medidas de dispersão que, segundo Martins (2008, p.52) [11] "São medidas estatísticas utilizadas para avaliar o grau de variabilidade, ou dispersão, dos valores em torno da média".

Assim, suponhamos que se deseja comparar a produtividade de dois agentes de trânsito, com base na quantidade de autos de infrações lavrados de segunda a sexta-feira, num período de 6 horas diárias:

	Segunda	Terça	Quarta	Quinta	Sexta
Agente A	5	6	6	7	6
Agente B	4	7	10	6	8

Tabela 11: Tabela produtividade

De acordo com a Tabela 11, verificamos que a produtividade média do agente A é de 6 autos lavrados diariamente, enquanto que a do agente B é de 7 autos. Consequentemente, baseados nestes únicos resultados($\overline{x}_A = 6$ e $\overline{x}_B = 7$), diríamos que a performance de B é melhor do que a de A. Mas, se analisarmos corretamente, perceberemos que a produção de A varia apenas de 5 a 7, ao passo que a de B varia de 4 a 10, o que mostra que a produção de A é mais uniforme. Ocorre, por outro lado, que um alto grau de uniformidade ou pequena dispersão costuma ser considerado em outras áreas e/ou profissões, como algo de qualidade desejável em um processo produtivo pois, qualquer produção em série seria antieconômica se houvesse muita variabilidade nos materiais ou peças fabricadas, por exemplo.

Para avaliar o grau de variabilidade ou dispersão dos valores de um conjunto de números, lançaremos mão das estatísticas denominadas medidas de dispersão(variância, desvio-padrão e coeficiente de variação), que nos proporciona um conhecimento mais completo do fenômeno a ser analisado, permitindo estabelecer comparações entre fenômenos de mesma natureza e mostrando até que ponto os valores se distribuem acima ou abaixo da tendência central(média aritmética).

3.6.1 Variância (S^2)

A variância é uma medida de dispersão que tem como objetivo a avaliação de um conjunto de dados, analisando o quanto eles estão dispersos em relação à média aritmética.

Definição 11. *Dados n elementos $x_1, x_2, ..., x_n$, a variância é a medida de dispersão definida como a média do quadrado dos desvios dos elementos em relação à média aritmética desses elementos. Ou seja, a variância é dada por:*

$$S^2 = \frac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{n} \tag{10}$$

onde $(x_i - \overline{x})$ é o desvio de cada elemento em relação à média e $\overline{x}$ = média aritmética dos n elementos.

Quando a variância representar uma descrição da amostra e não da população, caso mais frequente em estatística, o denominador das expressão 8 será igual a $n-1$ em vez de n. O motivo é pelo fato de que, utilizando do divisor $n-1$, obtém-se uma estimativa não viesada do parâmetro da população. Dessa forma, o cálculo da variância de uma amostra de dados é feito por:

$$S^2 = \frac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{n-1}. \tag{11}$$

Exemplo 33. *No dia 19/3/2018, no período matutino, foi realizada uma fiscalização com o uso do radar móvel na Alameda Ricardo Paranhos no setor Marista, onde a velocidade máxima permitida é de 40km/h. À seguir, temos as velocidades constatadas de dez veículos em apenas um minuto de fiscalização (Fonte: divisão de radar da SMT):*

43 - 55 - 51 - 55 - 36 - 42 - 46 - 49 - 35 - 45

Calcule a variância para esse conjunto de valores:
Solução:
Primeiro, calculamos a média aritmética para esses valores:

$$\overline{x} = \frac{43+55+51+55+36+42+46+49+35+45}{10} = 45,7$$

Como os dados são de uma amostra, usamos a fórmula 11:

$$S^2 = \frac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{n-1} = \frac{(43-45,7)^2 + ... + (45-45,7)^2}{10-1} = \frac{442,91}{9} = 49,21$$

Logo, a variância é $S^2 = 49,21$.

Quando os valores vierem dispostos em uma tabela de frequências de dados agrupados, o cálculo da variância será dado por:

$$\frac{\sum_{i=1}^{k}(x_i - \overline{x})^2 f_i}{n} \tag{12}$$

Exemplo 34. *A tabela abaixo traz as velocidades de vinte veículos aferidas pelo radar móvel descrito no Exemplo 30. Calcule a variância dessas velocidades.*

Velocidade	f_i	Ponto médio(x_j)	$f_i.x_j$	$(x_j - \overline{x})^2 f_i$
24 ⊢ 30	2	27	54	584,82
30 ⊢ 36	1	33	33	122,1
36 ⊢ 42	6	39	234	156,06
42 ⊢ 48	3	45	135	2,43
48 ⊢ 54	5	51	255	238,05
54 ⊢ 60	3	57	171	449,23
	$\sum fi = 20$		$\sum f_i.x_j = 882$	$\sum (x_j - \overline{x})^2 f_i = 1602,69$

Tabela 12: Variância para dados agrupados

Solução: *Inicialmente, note que:*

$$\overline{x} = \frac{\sum f_i.x_j}{20} = \frac{882}{20} = 44,1.$$

Usando a fórmula 10, temos que:

$$\frac{\sum_{i=1}^{k}(x_i - \overline{x})^2 f_i}{n} = \frac{\sum_{i=1}^{7}(x_i - 44,1)^2 f_i}{n} = \frac{1602,69}{20} = 80,13.$$

Logo, a variância é $S^2 = 80,13$.

3.6.2 Desvio-padrão(S)

O desvio-padrão é a medida de dispersão mais usada e é capaz de identificar o "erro" em um conjunto de dados, caso quiséssemos substituir um dos valores coletados pela média aritmética, informando o quão "confiável" é esse valor. Quanto mais próximo de 0 for o desvio padrão, mais homogêneos são os dados, ou seja, quanto maior for a variabilidade entre os dados, maior será o desvio padrão. Como a unidade de variância é o quadrado da unidade dos dados, o desvio-padrão serve para compatibilizar essas unidades pois, o cálculo do desvio padrão é feito a partir da raiz quadrada positiva da variância. Em termos simbólicos, temos que:

$$S = \sqrt{S^2}. \tag{13}$$

No caso do Exemplo 31, temos que o desvio-padrão é:

$$S = \sqrt{S^2} = \sqrt{80,13} = 8,95.$$

Portanto, o desvio-padrão é $S = 8,95$.

3.6.3 Coeficiente de variação(CV)

Os estudos estatísticos estão relacionados às situações que envolvem estratégias e planejamentos, coleta e organização de dados, análise e interpretação clara e objetiva dos dados observados. Para comparação de dois ou mais conjuntos de dados, a estatística utiliza o desvio padrão, desde que esses dados estejam na mesma unidade de medida. Caso os conjuntos de dados sejam medidos em grandezas diferentes (unidades de medida diferentes), a comparação será feita utilizando o coeficiente de variação.

O coeficiente de variação é usado para analisar a dispersão em termos relativos a seu valor médio quando duas ou mais séries de valores apresentam unidades de medida diferentes. Dessa forma, podemos dizer que o coeficiente de variação é uma forma de expressar a variabilidade dos dados excluindo a influência da ordem de grandeza da variável. O cálculo do coeficiente de variação é feito por meio da expressão:

$$CV = \frac{S}{\overline{x}} \tag{14}$$

em que S é o desvio padrão e $\overline{x}$ é a média aritmética.

Exemplo 35. *A fiscalização de excesso de velocidade por radar móvel na Alameda Ricardo Paranhos(onde o máximo permitido é 40 km/h) no período das 9:00h às 9:10h do dia 19/3/2018, registrou as velocidades de 57 veículos dos quais, 23 foram multados por excesso de velocidade. A média das velocidades desses 23 veículos foi de 53,4 km/h com desvio-padrão de 5,24. Observou-se também que a média de veículos fiscalizados por minuto foi de 5,7, com desvio-padrão de 2,16. Em qual das análises houve maior grau de dispersão?(Fonte: Divisão de radar da SMT)*

Solução:

Para avaliar o grau de dispersão em cada situação, devemos calcular o coeficiente de variação para cada caso.

1º caso) CV das velocidades dos veículos multados:

$\overline{x} = 53,4$ *e* $S = 5,24$

$$CV = \frac{S}{\overline{x}} = \frac{5,24}{53,4} = 0,0981 = 9,81\%.$$

2º caso) CV do número de veículos fiscalizados po minuto:

$\overline{x} = 5,7$ *e* $S = 2,16$

$$CV = \frac{S}{\overline{x}} = \frac{2,16}{5,7} = 0,379 = 37,9\%.$$

Comparando os dois coeficientes de variação, podemos concluir que houve um maior grau de dispersão para os dados do 2º caso (número de veículos fiscalizados por minuto) ou ainda que os dados do 1º caso(velocidades dos veículos multados) são mais homogêneos.

3.6.4 Box-plot

Conforme Magalhaes e Lima (2013, pg.20) [10] "Uma representação gráfica envolvendo quartis é o *box-plot* ou gráfico de caixa, que permite visualizar diversos aspectos da distribuição dos dados, tais como posição, variabilidade, assimetria e mesmo a ocorrência de valores atípicos.

Para construção de um *box-plot*, precisamos usar informações dos dados brutos de uma tabela. Definimos, então, um retângulo("caixa")em que a aresta inferior coincide com o primeiro quartil e a superior, com o terceiro quartil. A mediana é representada por um traço no interior do retângulo. Segmentos de reta, denominados *bigodes*, são incluídos no *box-plot*, partindo dos primeiro e terceiro quartis e terminando em valores definidos a seguir. O intervalo $[Q_1;Q_3]$ contém 50% das observações centrais e dá uma ideia de quão dispersos são os valores observados. A amplitude desse intervalo, $IQ = Q_3 - Q_1$, recebe o nome de *intervalo interquartil.* Para conjuntos de dados "bem comportados", a maior parte das observações se situa no intervalo $[Q_1 - 1,5IQ; Q_3 + 1,5IQ]$, sendo que dados fora desse intervalo são potenciais valores *atípicos* ou *discrepantes*, pois representariam um padrão distinto do esperado para a maior parte das unidades experimentais. Por essa razão, os limites desse intervalo serão denominados *pontos de corte.* Os segmentos partindo dos primeiro e terceiro quartis são limitados pelos valores mínimo e máximo dentro do intervalo acima descrito. Valores abaixo ou acima desses limites são representados por asteriscos e denominados *valores extremos* também chamados de *outliers.*

A Figura 16 ilustra os elementos de um *box-plot.*

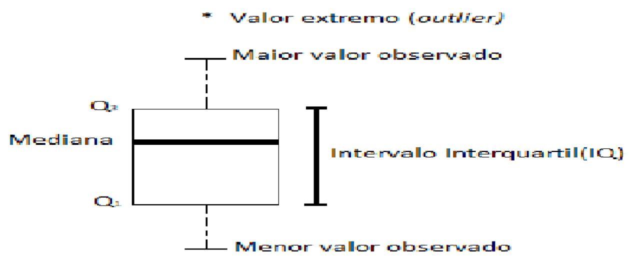


Figura 16: Elementos de um box-plot/*(fonte: própria)*

Exemplo 36. *: A Tabela 13 mostra os números de acidentes ocorridos em Goiânia divididos por regiões, nos anos de 2015 a 2018.*

Região	2015	2016	2017	2018
Sul	5488	4083	2659	488
Centro	5218	4009	2453	455
Sudeste	2814	2185	1038	130
Norte	1893	1571	804	98
Oeste	1878	1451	804	79
Leste	1484	1159	603	72
Noroeste	1429	1089	562	35

Tabela 13: Acidentes por região

Vamos construir o box-plot para análise desses dados(mostraremos os cálculos para o ano de 2015):

De acordo com a tabela, os dados de 2015 , já ordenados, são:

1429 - 1484 - 1878 - 1893 - 2814 - 5218 - 5488

Como o conjunto de dados é ímpar(n = 7), temos:
Menor valor = 1429
$Q_1 = \frac{1484+1878}{2} = 1681$
Md = 1893
$Q_3 = \frac{2814+5218}{2} = 4016$
IQ = 4016 - 1681 = 2335
Maior valor = 5488
Ponto de corte inferior : $Q_1 - 1,5IQ = 1681 - 1,5.2335 = -1821,5$
Ponto de corte superior: $Q_3 + 1,5IQ = 4016 + 1,5.2335 = 7518,5$

O box-plot é apresentado na Figura 17:

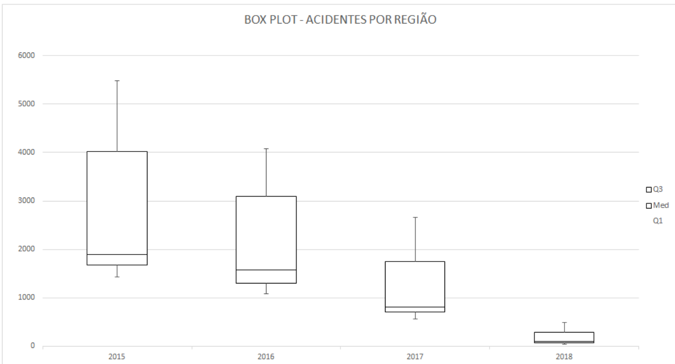


Figura 17: Box-plot/*(fonte: própria com dados da SMT)*

Interpretação do box-plot

O retângulo contém 50% dos valores do conjunto de dados. A posição da linha mediana no retângulo informa sobre a assimetria da distribuição. Uma distribuição simétrica teria a mediana no centro do retângulo. Se a mediana é próxima de Q_1, então, os dados são positivamente assimétricos. Se a mediana é próxima de Q_3 os dados são negativamente assimétricos. O comprimento das linhas fora do retângulo (algumas vezes chamadas de *whisquers*) informam sobre a cauda da distribuição. No exemplo dado, podemos observar que a distribuição é positivamente assimétrica e que os números médios de acidentes são diferentes e decrescentes de 2015 à 2018. Nota-se também que a variabilidade dos acidentes foi decrescendo ao passar dos anos e que em nenhum ano, tivemos valores atípicos(*outlier*). Concluímos, com o auxílio gráfico do *box-plot*, então, que o número de acidentes vem diminuindo em Goiânia.

Exemplo 37. *Apresentamos agora, o número de multas aplicadas por estacionamento sobre o passeio público, nos doze meses de 2017.*

714-765-787-794-937-990-1151-1353-1494-1560-2036-2545

Como o conjunto de dados é par(n = 12), temos:
Menor valor = 714
$Q_1 = 790,5$
Md = 1070,5
$Q_3 = 1527$
IQ = 1527 - 790,5 = 736,5
Maior valor = 2545
Ponto de corte inferior : $Q_1 - 1,5IQ = 790,5 - 1104,75 = -314,25$
Ponto de corte superior: $Q_3 + 1,5IQ = 1527 + 1104,75 = 2631,75$

O box-plot é apresentado na Figura 18:

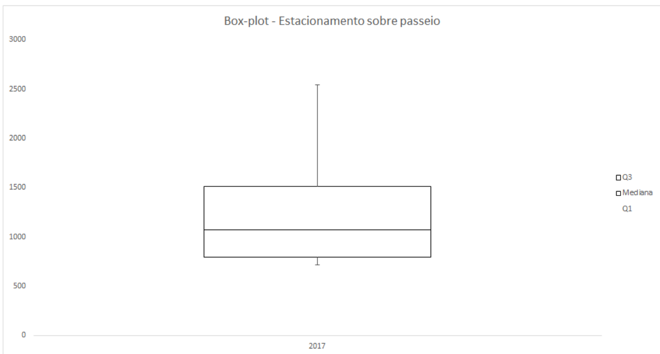


Figura 18: Box-plot/*(fonte: própria com dados da SMT)*

Interpretação do box-plot

Neste exemplo, podemos observar que a distribuição é positivamente assimétrica. Há maior concentração de dados nos valores mais baixos, com isso, a cauda mais longa da distribuição fica acima, indicando a ocorrência de valores altos com baixa frequência. Nota-se também que houve moderada variabilidade no número de multas aplicadas mensalmente e que não há a presença de valores discrepantes(*outlier*).

3.7 Medidas de Assimetria

As medidas de assimetria proporcionam, juntamente com as medidas de posição e de dispersão, a compreensão completa da distribuição de frequência estudada. Assimetria, como o próprio nome insinua, significa desvio ou afastamento da simetria. Em outras palavras, assimetria é o grau de deformação de uma curva de frequências. Conhecer o tipo e intensidade da assimetria de um conjunto de dados pode trazer informações úteis ao analista. Por exemplo, caso a distribuição tenha uma forte assimetria positiva, sabe-se que apesar da alta concentração de dados em valores mais baixos, a média sofrerá influência da cauda à direita deslocando-se em sua direção. Nesse caso, haverá mais observações abaixo da média do que acima dela. O inverso acontece se a assimetria for negativa. Quanto ao grau de de deformação ou assimetria, podemos ter três tipos de curvas de frequência, as quais descreveremos a seguir.

3.7.1 Curva Simétrica ou Distribuição Simétrica

Uma distribuição de frequências simétrica apresenta como característica principal o fato de as três medidas de centralidade mais usadas, a saber: média aritmética, mediana e moda, serem iguais. Em termos gráficos, a curva de frequências apresentará as duas caudas com a mesma configuração, conforme mostra a Figura 19:

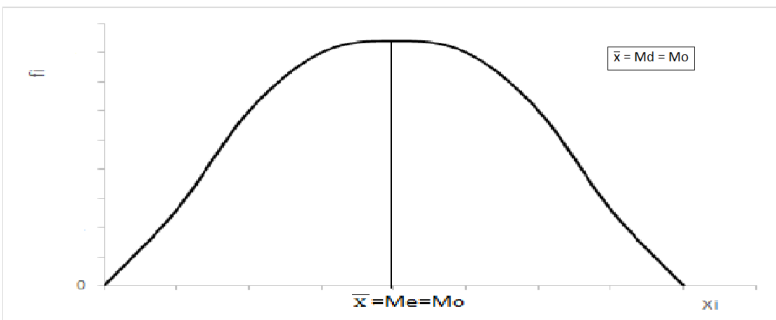


Figura 19: Curva de frequências simétrica, sem deformação/*(fonte: própria)*

3.7.2 Curva ou Distribuição de Frequências Assimétrica Positiva ou Desviada(Deformada) à Direita

Toda distribuição deformada é sempre assimétrica. No entanto, essa assimetria pode ser na cauda esquerda ou na direita da curva de frequências.
Uma distribuição com deformação positiva se apresenta com uma cauda mais alongada à direita da ordenada máxima(correspondente à moda) do que à esquerda. Nas distribuições assimétricas à direita, há uma predominância de valores superiores à moda. Os valores concentram-se na extremidade inferior da escala. Então, temos : $Mo < Md < \overline{x}$. A Figura 20 ilustra esse tipo de gráfico.

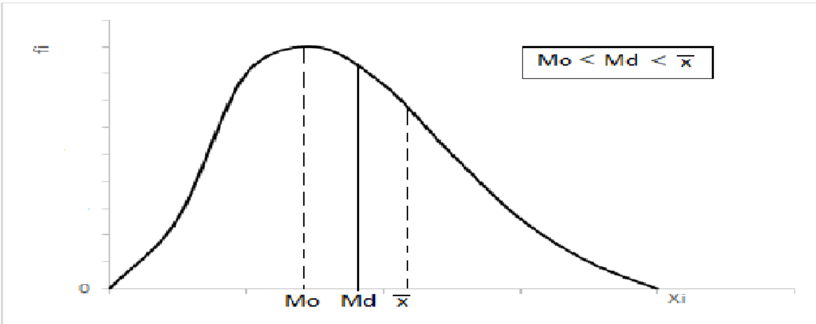


Figura 20: Curva de frequências assimétrica à direita (deformação positiva)/*(fonte: própria)*

Esse tipo de distribuição é bastante comum em administração e economia: variáveis como preços, PIB, salários, etc., possuem, em geral, este comportamento.

3.7.3 Curva ou distribuição de Frequência Assimétrica Negativa ou Desviada (Deformada) à Esquerda

Nas distribuições assimétricas negativas, predominam valores inferiores à moda. Neste caso, a curva apresenta uma cauda mais longe à esquerda do que à direita da ordenada máxima. Nas distribuições com assimetria negativa, a média aritmética é menor que a mediana, e esta, é menor que a moda, conforme pode ser observado na Figura 21.

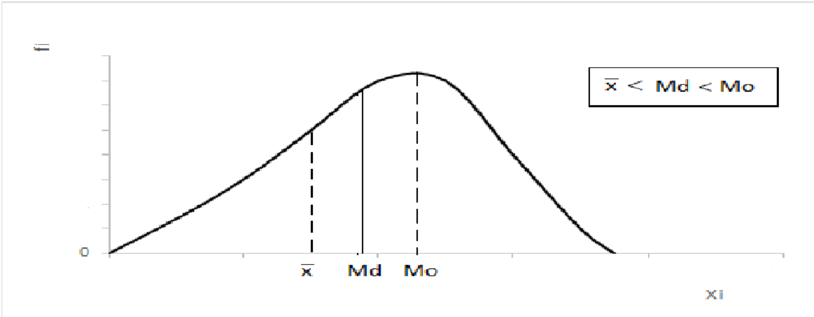


Figura 21: Curva de frequências assimétrica à esquerda (deformação negativa)/*(fonte: própria)*

3.7.4 Principais medidas de Assimetria

Para avaliar o grau de assimetria de uma distribuição, são propostas algumas medidas de assimetria, que veremos a seguir.

Método de Comparação entre Medidas de Tendência Central

É o método mais rudimentar. Não permite estabelecer até que ponto a curva analisada se desvia da simetria. A comparação é simples:

$\overline{x} > Mo \Longrightarrow$ Assimetria positiva
$\overline{x} = Mo \Longrightarrow$ Simetria
$\overline{x} < Mo \Longrightarrow$ Assimetria negativa

Coeficiente (Índice) de Pearson

Sugerida por Karl Pearson, usada muito frequentemente para avaliar o grau de assimetria de uma distribuição e se calcula mediante as expressões:
a) Primeiro coeficiente de Assimetria de Pearson(e_1)

$$e_1 = \frac{\overline{x} - Mo}{s} \tag{15}$$

b) Segundo Coeficiente de Assimetria de Pearson(e_2)

$$e_2 = \frac{3(\overline{x} - Md)}{s} \tag{16}$$

Como essas relações são aproximadas e não exatas, somente quando a distribuição for simétrica elas se equivalerão. Quando os valores de e obtidos através de (16) e (17) variam entre ± 1, então se diz que a assimetria é fraca e o fenômeno em estudo é considerado não muito assimétrico.

Exemplo 38. *Vamos usar a tabela do Exemplo 18 para calcular o tipo de assimetria pelo método de comparação e pelos coeficientes de Pearson, para aquela situação:*
Solução: *Já temos a média aritmética dos dados:* $\overline{x} = 10,85$.
Analisando a tabela, vemos facilmente que tanto a moda quanto a mediana são iguais a 10 (Mo = 10 e Md = 10). Então, temos que:

$$\overline{x} > Mo$$

Logo, pelo método da comparação, a distribuição tem assimetria positiva.
Calculemos, agora, pelo primeiro coeficiente de Pearson:
Primeiro, vamos calcular o desvio-padrão:

$$S = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{n-1}} = \sqrt{\frac{\sum_{i=1}^{20}(x_i - 10,85)^2}{19}} = 2,39$$

Então, temos:

$$e_1 = \frac{\overline{x} - Mo}{s} = \frac{10,85 - 10}{2,39} = 0,35$$

Por fim, vamos calcular pelo segundo coeficiente de Pearson:

$$e_2 = \frac{3(\overline{x} - Md)}{s} = \frac{3(10,85 - 10)}{2,39} = 1,06$$

Concluímos, assim, que a distribuição tem uma fraca assimetria positiva.

Exemplo 39. *A Tabela 14 mostra os números de atropelamentos fora da faixa de pedestres ocorridos em Goiânia no primeiro semestre de 2017:*

Vamos calcular os coeficientes de Pearson:

Solução: *Da tabela, temos:*

Mês	f_i
Janeiro	39
Fevereiro	19
Março	43
Abril	60
Maio	60
Junho	40

Tabela 14: Atropelamentos - 1º semestre/2017

$\overline{x} = 43,5$
Md = 41,5
Mo = 60

Cáculo do desvio-padrão:

$$S = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{n-1}} = \sqrt{\frac{\sum_{i=1}^{6}(x_i - 43,5)^2}{5}} = 15,34$$

Segue que:

$$e_1 = \frac{\overline{x} - Mo}{s} = \frac{43,5 - 60}{15,34} = -1,07$$

e que

$$e_2 = \frac{3(\overline{x} - Md)}{s} = \frac{3(43,5 - 41,5)}{15,34} = 0,39$$

A relação (16) tem o incoveniente de requerer a determinação prévia da moda. Quando a distribuição em questão não apresentar forte assimetria, deve-se dar preferência à relação (17)

Coeficiente Quartil de Assimetria (e_Q)

É frequentemente usado e para seu cálculo, usa-se os três quartis. Trat-se de uma medida muito útil quando não for possível o emprego do desvio-padrão como medida de dispersão, mas apenas alguma medida que dependa dos quartis. O coeficiente quartil de assimetria é definido pela relação:

$$e_Q = \frac{Q_3 - 2Md + Q_1}{Q_3 - Q_1} \tag{17}$$

Esse coeficiente de assimetria assume valores entre os limites -1 e +1, ou seja:

$$-1 \leq e_Q \leq 1$$

Exemplo 40. *Calcular o coeficiente quartil de assimetria para a distribuição do Exemplo 34.*
Solução: *Os dados são*

714-765-787-794-937-990-1151-1353-1494-1560-2036-2545

Primeiro, vamos calcular os quartis:

$$Q_1 = \frac{787+794}{2} = 790,5$$
$$Q_2 = Md = \frac{990+1151}{2} = 1070,5$$
$$Q_3 = \frac{1494+1560}{2} = 1527$$

Substituindo na relação, temos que:

$$e_Q = \frac{Q_3 - 2Md + Q_1}{Q_3 - Q_1} = \frac{1527 - 2.1070,5 + 790,5}{1527 - 790,5} = 0,22$$

4 Metodologia

Nesta seção, apresentamos a metodologia desenvolvida neste trabalho, parte fundamental para nosso objetivo principal. Ressaltamos aqui, que a proposta ficou um pouco prejudicada pela quantidade de aulas que tivemos pois, devido aos feriados prolongados e o processo eleitoral no colégio, dispusemos apenas de dez aulas para trabalhar o conteúdo inicial de Estatística. Além do que, na grade curricular prevista, o foco principal para o bimestre era Trigonometria e Geometria Espacial, a Estatística seria mais aprofundada no semestre seguinte.

4.1 Desenvolvimento das aulas

Inicialmente, foi aplicado um questionário(apêndice) com intuito de traçar o perfil da turma, com perguntas como idade, se o aluno tinha habilitação e qual sua visão sobre o trânsito de Goiânia. Nesta aula, esclarecemos aos alunos o nosso objetivo, informamos da nossa experiência no trânsito e enfatizamos que eles não só fariam parte, mas que sua participação seria fundamental para o desenvolvimento deste projeto.

Nas aulas 2 e 3, fizemos uso de datashow, com objetivo de familiarizar a turma com o tema, onde apresentamos os principais gráficos estatísticos. Buscamos uma constante interação com os alunos, promovendo debates, expondo alguns pontos específicos sobre as normas gerais de circulação e conduta presentes no CTB e ilustrando as principais infrações de trânsito cometidas em nossa capital.

Para a aula 4, introduzimos os conceitos básicos da estatística descritiva e trabalhamos a construção de uma tabela de frequências para a variável *idade*, com os dados obtidos no questionário aplicado. Em seguida, de posse dos dados referentes ao número de multas aplicadas naquela semana, construimos uma tabela de dados agrupados.

Para as aulas 5 e 6, foi preparada uma lista de atividades sobre gráficos estatísticos. Nesta lista, foram trabalhados gráficos gerados com dados sobre acidentes de trânsito e número de multas, coletados junto aos órgãos de trânsito e também *in loco*. Para este momento, formamos grupos de três e quatro alunos, promovendo uma contínua interação entre alunos-alunos e professor-alunos, onde procuramos dar uma visão interpretativa dos principais gráficos estatísticos e chamar a atenção para o alto índice de acidentes com vítimas e o baixo número de multas aplicadas.

Na aula 7, conceituamos as medidas de centralidade dando ênfase à média aritmética. Nesta aula, foi calculada a média de idade da turma, a idade mediana e a idade modal. Resolvemos, também, exercícios de medida de centralidade com os dados de autuações feitas pela nossa equipe de fiscalização de trânsito naquele dia.
Para as aulas 8 e 9, foi elaborada uma lista de atividades sobre as medidas de tendência central. Momentos em que trabalhamos dados sobre multas aplicadas, acidentes com vitima, atropelamentos e mortos no trânsito de Goiânia, aplicando a mesma dinânica nas aulas 5 e 6.

Na última aula fizemos um *feedback* sobre tudo que foi falado, apresentado e desenvolvido com a turma, enfatizando que nossa proposta era de trazer um modelo novo de ensino de Matemática com um tema bastante importante e presente no cotidiano de cada um, que poderia muito bem ser trabalhado em outras disciplinas como Física, Química e Biologia. Frizamos, também, que o número de aulas que dispusemos para a nossa proposta fora pequeno dada tamanha importância que o tema exige e que, independentemente do tema ou assunto, o raciocínio para a resolução de problemas envolvendo gráficos, tabelas e medidas de centralidade seria análogo aos empregados em sala de aula, nas atividades propostas.

Com os dados obtidos no questionário, construimos juntamente com os alunos, o Gráfico 22. Salientamos que este gráfico é um elemento estatístico real em que eles foram os protagonistas, os pesquisados. Foi produzido obedecendo as fases de uma trabalho estatístico, a saber: a coleta de dados(questionário), o tratamento das informações(tabulação)e o produto final(gráfico), o qual facilita a visualização e interpretação dos resultados de uma amostra.

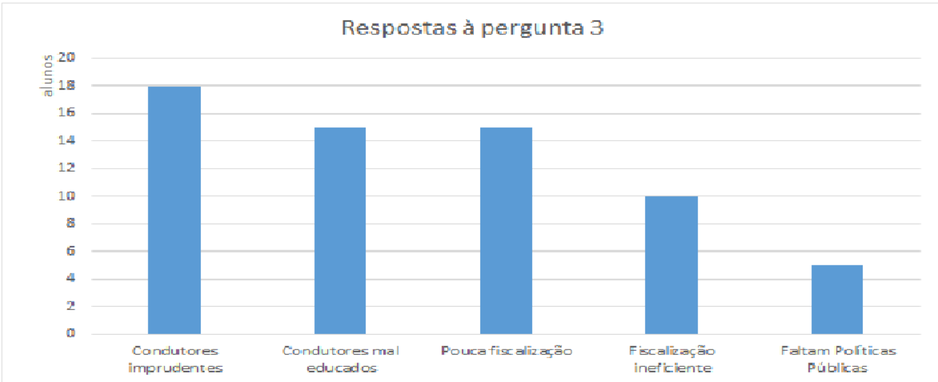


Figura 22: Gráfico de colunas para a pergunta 3

4.2 Avaliação dos resultados

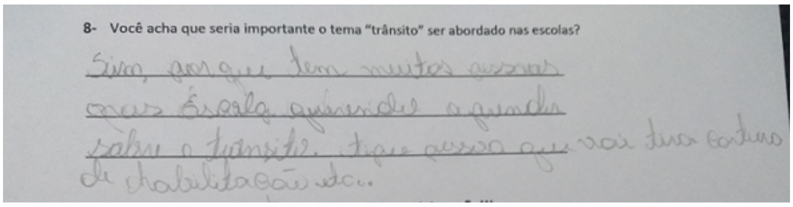
O ensino no período noturno é, historicamente, prejudicado. Fatores como trabalho, cansaço, defasagem de ensino e tempo reduzido das aulas contribuem para isso. Apesar disso, avaliamos muito positivamente essa experiência em trazer uma forma diferenciada e inovadora para o ensino da Estatística. Aulas repetitivas e massantes, sempre com uso de quadro-giz e livro didático são reclamações constantes dos alunos. Acreditamos que a contextualização da Matemática é um processo sociocultural e que o seu ensino não deve se restringir apenas à exercícios e problemas prontos de livros didáticos. Para Tomaz e David [3]:

> No que diz respeito às tentativas de desenvolver o ensino da Matemática na perspectiva da contextualização e da interdisciplinaridade, nem sempre elas tem sido avaliadas como bem sucedidas porque muitas vezes os esforços de contextualização acabam resultando como artificiais, como naqueles livros didáticos em que o contexto das situações serve apenas como ponto de

partida para obtenção dos dados numérico que vão ser usados nas operações matemáticas.

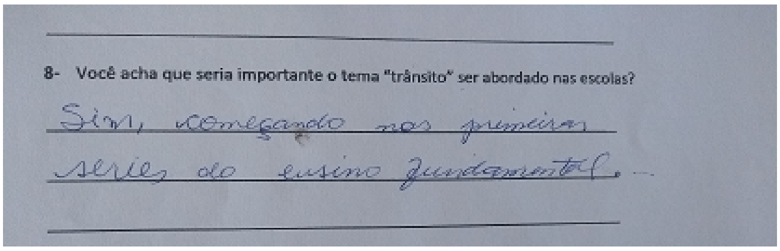
Diferentemente disso, buscamos aqui, trazer uma abordagem sólida e consistente a qual revelou um grande interesse da grande maioria dos alunos. Normalmente, o interesse pela Matemática, qualquer que seja o conteúdo, é mínimo. Somente aqueles alunos que já tem afinidade com a disciplina esboçam o interesse, o que é natural. Contudo, a nossa metodolgia surtiu grande efeito e até mesmo aqueles alunos que tem um certa aversão à disciplina se enganjaram nos exercicios propostos, por se interessarem pelo tema e não considerarem a aula "chata". A exposição com uso de datashow sobre um pouco de legislação de trânsito deixou toda a turma bastante interessada e à todo tempo perguntas sobre a fiscalização no trânsito era feitas. Frizamos que, apesar de nossa abordagem ser a Estatística no trânsito, o raciocínio para a solução de quaisquer problemas de Estatística eram sempre análogos, diferenciando-se apenas nos diversos contextos em que eram abordados e que, a escolha pelo trânsito se fazia necessária por poder proporcionar à eles um ensino concreto com dados reais, que a nossa condição e experiência favorecia. O contato com dados reais, como a grande quantidade de acidentes e mortos no trânsito, trouxe à baila a necessidade de se investir em políticas públicas voltadas para a educação no trânsito. Foram inúmeros os questionamentos à respeito da lei e da fiscalização de trânsito, onde pudemos observar, realmente, o quão importante é a abordagem desse tema em sala de aula.

À seguir, ilustramos com algumas respostas de alunos, a necessidade de se ministrar aulas que abordem o tema trânsito em sala de aula, as quais corroboram o nosso pensamento:


8- Você acha que seria importante o tema "trânsito" ser abordado nas escolas?

Sim, porque tem muitas pessoas nas Escola [illegible] sobre o trânsito. tipo pessoa que vai tirar carteira de habilitação etc..

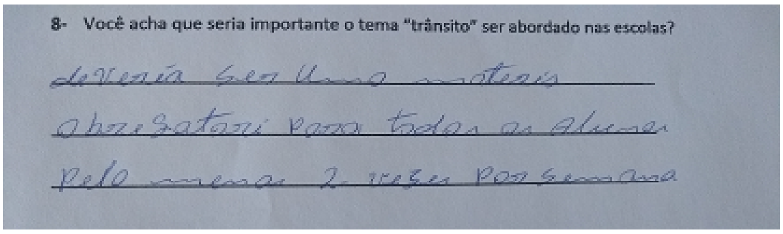
Figura 23: Resposta de uma aluna de 16 anos



8- Você acha que seria importante o tema "trânsito" ser abordado nas escolas?

Sim, começando nas primeiras series do ensino fundamental.

Figura 24: Resposta de um aluno de 48 anos

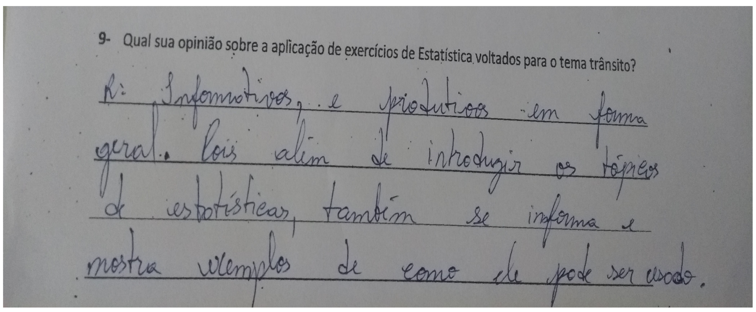


8- Você acha que seria importante o tema "trânsito" ser abordado nas escolas?

deveria ser uma materia obrigatori para todos os alunos pelo menos 2 vezes por semana

Figura 25: Resposta de um aluno de 63 anos

Na resolução das duas listas aplicadas, em grupo, pudemos notar o interesse até mesmo daqueles alunos que em algumas aulas nem mesmo abriam seu caderno e ficavam totalmente alheios à aula. Era notória a dificuldade de alguns alunos na resolução dos problemas propostos, devido a grande defasagem de ensino mas, com uma simples explicação aliada ao interesse que eles demostravam, todas as dúvidas eram sanadas e o aprendizado ficava evidente. Neste momento, percebemos que o que é preciso para despertar o interesse no aluno é a motivação. O cansaço e a falta de interesse foram superados pela nossa proposta, que os motivou desde a primeira aula, por se tratar de um tema de tamanha importância e presente no dia-a-dia de cada um deles. As figuras abaixo trazem opiniões de alunos que coadunam com a nossa:



9- Qual sua opinião sobre a aplicação de exercícios de Estatística voltados para o tema trânsito?

R: Informativos, e produtivos em forma geral. Pois além de introduzir os tópicos de estatísticas, também se informa e mostra exemplos de como ele pode ser usado.

Figura 26: Resposta de um aluno de 20 anos