

UNIDADE 1

1) A direção do CED (Centro de Educação da UFSC), e os departamentos do centro, têm interesse em avaliar se a biblioteca setorial está atendendo adequadamente os alunos de graduação, pós-graduação, pessoal docente e técnico-administrativo do CED. Há preocupação com o acervo em si (atualização, composição, número de cópias disponíveis, adequação às necessidades de cada curso), e com o atendimento aos usuários (número de atendentes, horário, "cortesia"). Faça o planejamento da pesquisa estabelecendo

a) Objetivo geral da pesquisa.

b) Objetivos específicos da pesquisa.

c) Tipo de pesquisa (Levantamento ou Experimento).

d) População da pesquisa.

e) Quais são as variáveis da pesquisa? Por quê? Como serão medidas? Por quê?

f) Como serão coletados os dados: secundários ou primários, censo ou amostragem? Por quê?

g) Escolha do instrumento de pesquisa. Justificativa.

i) Elaboração do instrumento de pesquisa. Justificativa dos itens e opções escolhidas.

a) Objetivo geral da pesquisa. Justificativa.

Avaliar a opinião dos alunos de graduação em do CED no semestre 2009.1 sobre a biblioteca setorial do centro.

b) Objetivos específicos da pesquisa. Justificativa.

Para avaliar a qualidade da biblioteca setorial serão verificados os seguintes aspectos: acervo, estrutura e atendimento. Esses três itens permitirão caracterizar perfeitamente a qualidade da biblioteca setorial do CED.

c) Tipo de pesquisa (Levantamento ou Experimento). Justificativa.

Levantamento. Não há maneira de controlar as variáveis que não são importantes para a pesquisa (o que seria necessário no caso de um experimento).

d) População da pesquisa. Justificativa.

As opiniões de todos os alunos de graduação dos cursos do CED (Pedagogia e Biblioteconomia) regularmente matriculados no semestre 2009.1. São aqueles para quem queremos que os resultados da pesquisa sejam válidos. Os alunos que estiverem com suas matrículas trancadas, ou em estágio (sem frequentar disciplinas), ou que por algum outro motivo (saúde por exemplo) não comparecerem regularmente à UFSC poderão ser inacessíveis aos pesquisadores, ou seu número pode ser tão pequeno que não justifique sua inclusão.

Neste ponto é importante também decidir se algum subgrupo da população será excluído ou não: os calouros, por exemplo, será que já possuem vivência universitária suficiente para opinar sobre a qualidade da biblioteca do CED.

e) Quais são as variáveis da pesquisa? Por quê? Como serão medidas? Por quê?

As variáveis estão relacionadas aos objetivos específicos, permitindo sua caracterização. Todas as variáveis serão medidas através de uma escala de opções, a partir de questões feitas aos alunos.

Objetivo específico Acervo: atualização, conservação, número de cópias, adequação às disciplinas.

Objetivo específico Estrutura: equipamentos de acesso à informação, conforto dos usuários,.

Objetivo específico Atendimento: horário de atendimento, número de atendentes, qualificação dos atendentes, cortesia dos atendentes.

f) Como serão coletados os dados: secundários ou primários, censo ou amostragem? Por quê?

Dados primários. Mesmo que exista algum levantamento prévio neste assunto, o objetivo geral será atingido com maior eficácia se os dados forem atuais (referentes ao semestre 2009.1). Assim é preciso coletá-los.

Amostragem. Entrevistar todos os alunos de Pedagogia e Biblioteconomia pode não ser possível por falta de tempo ou recursos, obrigando a selecionar uma amostra representativa da população.

g) Escolha do instrumento de pesquisa. Justificativa.

Questionário. As informações procuradas somente poderão ser obtidas através de perguntas feitas diretamente aos alunos.

h) Elaboração do instrumento de pesquisa. Justificativa dos itens e opções escolhidas.

Serão feitas perguntas objetivas aos alunos sobre as variáveis descritas no item g. As perguntas serão objetivas para evitar “devaneios” por parte dos respondentes e para facilitar a análise dos dados. Será feita uma pergunta inicial sobre o uso ou não da biblioteca: para os que declararem não utilizar questionar-se-á por quê. As outras questões objetivas (apenas para que os declararem usar a biblioteca setorial) terão opções apresentadas em uma escala de 1 a 5 (sendo 1 o mínimo e 5 o máximo), e a questão final será discursiva para que os alunos possam deixar alguma sugestão ou reclamação.

O questionário a seguir faz parte de um trabalho acadêmico. Você não precisa se identificar. Por favor responda todas as perguntas CORRETA E FRANCAMENTE. Os resultados serão posteriormente divulgados para a comunidade universitária.

1) Você usa regularmente (pelo menos uma vez por semana) a biblioteca setorial do CED?

☐ SIM

☐ NÃO

Se você respondeu SIM vá para a pergunta 3, se respondeu NÃO vá para a pergunta 2.

2) Se você não usa a biblioteca setorial do CED indique o motivo PREDOMINANTE para isso (assinale apenas uma alternativa):

☐ desatualização do acervo ☐ horário de atendimento inadequado

☐ número de cópias insuficiente ☐ problemas com o atendimento

☐ desconforto ao utilizar (calor, frio, ventilação, barulho, etc)

☐ usa outra biblioteca (Central, Pública do Estado, FAED, etc)

☐ Outro (indique): _____

Devolva o questionário ao aplicador. Muito obrigado pela colaboração.

3) Em qual período você PREDOMINANTEMENTE frequenta a biblioteca (assinale apenas uma alternativa):

☐ Manhã

☐ Tarde

☐ Noite

4) Responda às questões abaixo dando uma nota de 1 a 5, sendo 1 o valor mínimo (PIOR score) e 5 o valor máximo (MELHOR score), para os seguintes aspectos relacionados à biblioteca do CED (assinale apenas uma alternativa):

a) Atualização do acervo (1) (2) (3) (4) (5)

b) Conservação do acervo (1) (2) (3) (4) (5)

c) Número de cópias por documento do acervo (1) (2) (3) (4) (5)

d) Adequação dos títulos do acervo às disciplinas (1) (2) (3) (4) (5)

f) Equipamentos para auxiliar na busca de informações (1) (2) (3) (4) (5)

g) Número de mesas e postos de estudo (1) (2) (3) (4) (5)

h) Condições ambientais (calor, frio, ventilação) (1) (2) (3) (4) (5)

g) Nível de ruído (1) (2) (3) (4) (5)

h) Horário de atendimento (1) (2) (3) (4) (5)

i) Número de atendentes no período que você frequenta (1) (2) (3) (4) (5)

j) Qualificação dos atendentes (1) (2) (3) (4) (5)

k) Cortesia dos atendentes (1) (2) (3) (4) (5)

5) Aponte o principal ponto positivo e negativo da biblioteca setorial do CED:

POSITIVO: _____

NEGATIVO: _____

UNIDADE 2

1) Construa a distribuição de frequências para a variável opinião sobre design (Design) dos veículos da Toyord. Analisando os resultados, os clientes, de uma forma geral, tem boa impressão sobre o design dos veículos da TOYORD? JUSTIFIQUE.

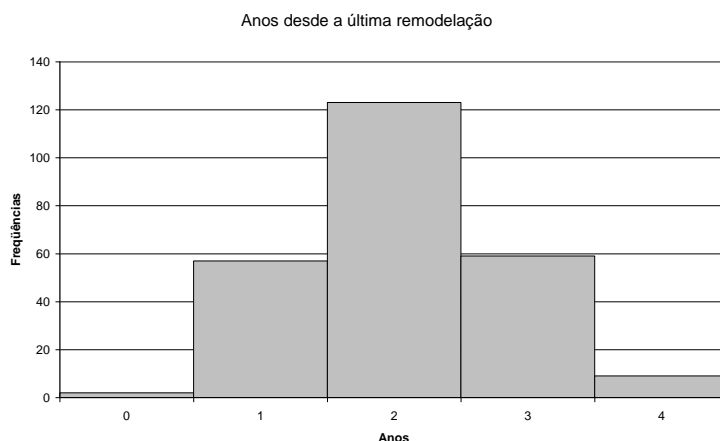
Aqui é necessário consultar o arquivo “Como fazer análise exploratória de dados com o Microsoft Excel®”, disponível no ambiente virtual da disciplina, e realizar a análise propriamente dita no arquivo AmostraToyord.xls, também disponível no ambiente virtual.

A variável opinião sobre design (Design) é QUALITATIVA, e neste caso estamos interessados em realizar uma análise individual para saber se os clientes têm uma boa impressão sobre os veículos da TOYORD. Basta seguir as recomendações do item 1.1 do referido texto, o que resultará na seguinte distribuição de frequências:

| Opinião sobre design | Frequência | Percentual |
|----------------------|------------|------------|
| Adiante dos outros | 59 | 23,60% |
| Atualizados | 160 | 64,00% |
| Ultrapassados | 31 | 12,40% |
| Total | 250 | 100,00% |

A opinião predominante (64%) entre os clientes é que o design dos veículos é atualizado. Se somarmos as opiniões de que o design está adiante dos outros concorrentes obteremos 87,60% de opiniões favoráveis, o que é um valor bastante expressivo, indicando que os clientes realmente têm boa impressão sobre o design dos veículos da TOYORD.

2) A variável anos de remodelação dos veículos (na percepção do cliente) está representada no histograma a seguir:



Fonte: adaptada pelo autor de Microsoft ®

O departamento de marketing alega que precisa de mais orçamento para “convencer” os clientes que os veículos da TOYORD têm design avançado, pois eles creem que a maioria dos clientes acha que eles foram remodelados “há vários anos atrás”. Os dados confirmam a crença do departamento de marketing? JUSTIFIQUE!

A definição de “há vários anos atrás” pode ser considerada bastante subjetiva, mas poderíamos concordar com 4 ou 5 anos no mínimo, especialmente no mercado automobilístico. A maioria esmagadora dos clientes (182, ou 72,8% do total) considera que os veículos foram remodelados há, no máximo 2 anos, não configurando o “há vários anos atrás”. Portanto, a crença do departamento de marketing NÃO é confirmada pelos dados.

3) Construa a distribuição de frequências agrupada em classes para a variável quilometragem média mensal percorrida com o veículo. Você considera que os clientes da Toyord usam bastante o veículo ou não? JUSTIFIQUE!

Vamos encontrar o intervalo do conjunto $\text{Intervalo} = 754 - 286 \text{ (máximo - mínimo)} = 468$

Número conveniente de classes = $\sqrt{250} = 15,81$ Talvez 15 sejam muitas classes, podemos arbitrariamente escolher 10 classes.

Amplitude das classes = $468/10 = 46,8$. Teremos um valor fracionário, o que pode dificultar a futura visualização da tabela, podemos, também arbitrariamente, escolher amplitude igual a 47.

Limites das classes: vamos começar pelo mínimo (286):

286|--333 333|--380 380|--427 427|--474 474|--521 521|--568
568|--615 615|--662 662|--709 709|--756

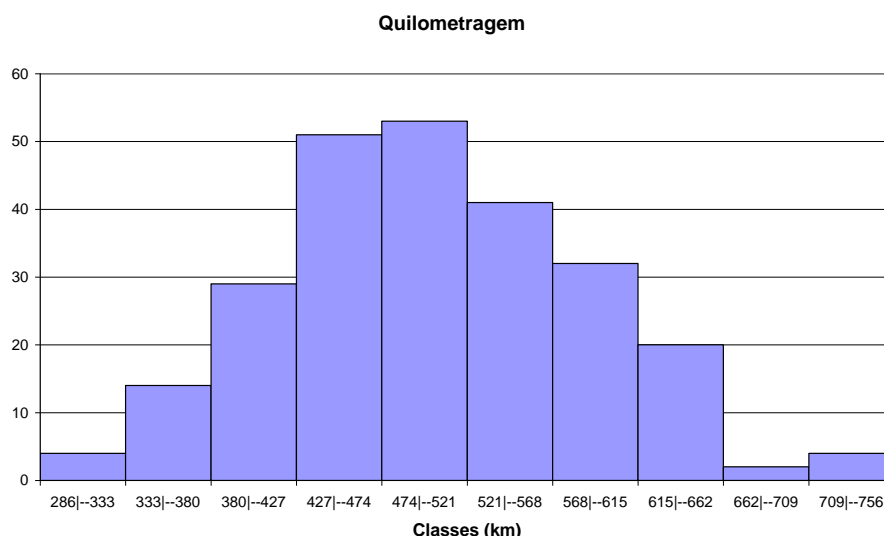
Vamos então obter a distribuição agrupada em classes para a variável Quilometragem, usando a função CONT.SE, exatamente como descrito na seção 2.1 do arquivo “Como fazer análise exploratória de dados com o Microsoft Excel®”. Tomando os cuidados na utilização da função CONT.SE chegaremos ao resultado:

| Classes | Frequências | Percentuais | Pontos médios |
|-----------|-------------|-------------|---------------|
| 286 --333 | 4 | 1,60 | 309,5 |
| 333 --380 | 14 | 5,60 | 356,5 |
| 380 --427 | 29 | 11,60 | 403,5 |
| 427 --474 | 51 | 20,40 | 450,5 |
| 474 --521 | 53 | 21,20 | 497,5 |
| 521 --568 | 41 | 16,40 | 544,5 |
| 568 --615 | 32 | 12,80 | 591,5 |
| 615 --662 | 20 | 8,00 | 638,5 |
| 662 --709 | 2 | 0,80 | 685,5 |
| 709 --756 | 4 | 1,60 | 732,5 |
| Total | 250 | 100,00 | - |

A resolução é um tanto subjetiva, pois a concepção do que significa usar “bastante” o veículo pode variar de pessoa para pessoa. Aqui adotamos apenas uma das possíveis opções.

Muitas seguradoras consideram que se uma pessoa percorre mais de 500 km por mês deve pagar mais pelo seguro, pois tem mais probabilidade de sofrer acidentes, problemas mecânicos ou mesmo ter o veículo furtado: usar bastante o veículo seria usar mais de 500 km. No nosso caso, há um limite de classe igual a 521, podemos usá-lo como referência: observe que apenas 39,60% dos clientes usam o veículo por 521 km ou mais, sendo assim, podemos dizer que os clientes da TOYORD NÃO usam bastante o veículo.

Podemos construir um histograma a partir da tabela acima, usando os procedimentos descritos no item 2.2.1 do texto “Como fazer Análise Exploratória de Dados com o Microsoft Excel”, disponível no ambiente virtual, chegaremos ao seguinte gráfico:



Visualmente percebemos que o valor 500 km está quase no centro do conjunto de dados.

4) Construa a tabela de contingências para modelo por opinião sobre concessionárias (concessionária). Com base nela avalie se os clientes de todos os modelos estão satisfeitos com os serviços prestados. JUSTIFIQUE a sua resposta.

Aqui é necessário consultar o arquivo “Como fazer análise exploratória de dados com o Microsoft Excel®”, disponível no ambiente virtual da disciplina, e realizar a análise propriamente dita no arquivo AmostraToyord.xls, também disponível no ambiente virtual.

As variáveis Modelo e opinião sobre concessionárias (Concessionária) são QUALITATIVAS, e neste caso estamos interessados em realizar uma análise conjunta para avaliar se os clientes de todos os modelos estão satisfeitos. Devemos usar os procedimentos descritos na seção 1.2 do arquivo citado acima, incluindo os percentuais, que nos levará à tabela abaixo

| Modelo | Opinião sobre concessionárias | | | | | Total geral |
|--------------|-------------------------------|----------------|----------------------|----------------------|--------------|-------------|
| | Bastante satisfatória | Insatisfatória | Muito insatisfatória | Não causou impressão | Satisfatória | |
| Chiconaultla | 1 | 33 | 20 | 22 | 5 | 81 |
| | 1,23% | 40,74% | 24,69% | 27,16% | 6,17% | 100,00% |
| | 2,63% | 57,89% | 68,97% | 29,33% | 10,00% | 32,53% |
| | 0,40% | 13,25% | 8,03% | 8,84% | 2,01% | 32,53% |
| DeltaForce3 | 6 | 13 | 6 | 19 | 12 | 56 |
| | 10,71% | 23,21% | 10,71% | 33,93% | 21,43% | 100,00% |
| | 15,79% | 22,81% | 20,69% | 25,33% | 24,00% | 22,49% |
| | 2,41% | 5,22% | 2,41% | 7,63% | 4,82% | 22,49% |
| LuxuriousCar | 13 | 1 | | 4 | 11 | 29 |
| | 44,83% | 3,45% | 0,00% | 13,79% | 37,93% | 100,00% |
| | 34,21% | 1,75% | 0,00% | 5,33% | 22,00% | 11,65% |
| | 5,22% | 0,40% | 0,00% | 1,61% | 4,42% | 11,65% |
| SpaceShuttle | 12 | 4 | | 14 | 12 | 42 |
| | 28,57% | 9,52% | 0,00% | 33,33% | 28,57% | 100,00% |
| | 31,58% | 7,02% | 0,00% | 18,67% | 24,00% | 16,87% |
| | 4,82% | 1,61% | 0,00% | 5,62% | 4,82% | 16,87% |
| Valentiniana | 6 | 6 | 3 | 16 | 10 | 41 |
| | 14,63% | 14,63% | 7,32% | 39,02% | 24,39% | 100,00% |
| | 15,79% | 10,53% | 10,34% | 21,33% | 20,00% | 16,47% |
| | 2,41% | 2,41% | 1,20% | 6,43% | 4,02% | 16,47% |
| Total | 38 | 57 | 29 | 75 | 50 | 249 |
| | 15,26% | 22,89% | 11,65% | 30,12% | 20,08% | 100,00% |
| | 100,00% | 100,00% | 100,00% | 100,00% | 100,00% | 100,00% |
| | 15,26% | 22,89% | 11,65% | 30,12% | 20,08% | 100,00% |

Torna-se imperativo fazer a análise através de algum percentual, pois os modelos têm frequências (vendas) diferentes, o que poderia dificultar a análise com base nas frequências. Mas qual percentual? Observe que a questão quer obter informações por modelo, cujos valores estão nas linhas, portanto, devemos usar os percentuais das linhas, para cada modelo e compará-los com o percentual da linha total.

Observe que o total da tabela não é 250 porque há um dado perdido de modelo (uma pessoa não declarou o modelo comprado, e tal valor foi removido).

Devemos também definir o que consideramos “clientes satisfeitos”: uma boa escolha é considerar a soma das opiniões “positivas” sobre as concessionárias, no caso “Bastante satisfatória” e “Satisfatória”. Com esses aspectos em mente podemos avaliar a satisfação por modelo.

No total geral observamos equilíbrio entre os percentuais de clientes satisfeitos (35,34%) e insatisfeitos (34,54%), além de 30,12% que não se impressionaram com o serviço. Tal situação “inspira cuidados”, pouco mais de um terço dos clientes está satisfeito com o atendimento das

concessionárias.

- há um problema muito sério com o serviço prestado pelas concessionárias aos proprietários do veículo Chiconaultla, pois apenas 7,41% estão satisfeitos, assombrosos 65,43% estão insatisfeitos (30% acima da opinião geral) e preocupantes 27,16% dizem que o serviço não causou impressão.

- no DeltaForce3 há grande equilíbrio entre satisfeitos (32,14%), insatisfeitos (33,93%) e os sem impressão (33,93%), contudo, a situação também é ruim pois nem um terço dos proprietários está feliz com o serviço prestado pelas concessionárias (situação parecida com a opinião geral).

- o modelo LuxuriousCar é, de longe, o veículo cujos proprietários estão mais satisfeitos com as concessionárias (82,76%), sendo que há apenas 3,45% insatisfeitos (e 0% de opinião Muito insatisfatória).

- o modelo SpaceShuttle tem a maioria (57,14%) de seus clientes satisfeitos com as concessionárias (contra 35,34% da opinião geral), mas saiu-se pior do que o LuxuriousCar.

- já a Valentiniana ficou em posição entre o DeltaForce3 e o SpaceShuttle, havendo grande equilíbrio entre satisfeitos (39,02%) e os indiferentes (39,02%), indicando que cerca de 60% dos clientes ou está insatisfeita ou não se impressionou com as concessionárias.

Obviamente, no geral, o serviço das concessionárias precisa ser muito melhorado, pois há apenas 34,34% de satisfeitos. Mas, no modelo Chiconaultla o problema é gritante, com o maior percentual de insatisfeitos de todos. Talvez a solução fosse utilizar para os outros modelos os procedimentos usados no LuxuriousCar, uma vez que seus proprietários são os mais satisfeitos (82,76%, percentual mais de 40% acima do total geral de satisfeitos).

5) Os executivos da Toyord creem que seus clientes mais abastados são mais críticos, tendem a ser mais insatisfeitos com seus veículos. Para verificar se isso é verdade construíram a tabela a seguir. Com base nela, a crença dos executivos é verificada (calcule os percentuais que achar necessários)? JUSTIFIQUE!

| Renda | Opinião geral sobre o veículo | | | | |
|----------------|-------------------------------|--------------|--------------------|------------|-------|
| | Bastante satisfeito | Insatisfeito | Muito insatisfeito | Satisfeito | Total |
| 1,795 -18,645 | 0 | 16 | 78 | 4 | 98 |
| 18,645 -35,495 | 7 | 47 | 26 | 22 | 102 |
| 35,495 -52,345 | 15 | 4 | 5 | 14 | 38 |
| 52,345 -69,195 | 7 | 0 | 1 | 1 | 9 |
| 69,195 -86,045 | 3 | 0 | 0 | 0 | 3 |
| Total geral | 32 | 67 | 110 | 41 | 250 |

Fonte: adaptado pelo autor de Microsoft ®

Novamente, torna-se imperativo a utilização de percentuais para permitir a comparação entre as faixas de renda, já que as frequências são diferentes. O ideal seria usar percentuais por linha, possibilitando a comparação das opiniões por faixa de renda. Calculando estes percentuais vamos chegar à tabela a seguir:

| Renda | Opinião geral sobre o veículo | | | | |
|----------------|-------------------------------|--------------|--------------------|------------|--------|
| | Bastante satisfeito | Insatisfeito | Muito insatisfeito | Satisfeito | Total |
| 1,795 -18,645 | 0,00 | 16,33 | 79,59 | 4,08 | 100,00 |
| 18,645 -35,495 | 6,86 | 46,08 | 25,49 | 21,57 | 100,00 |
| 35,495 -52,345 | 39,47 | 10,53 | 13,16 | 36,84 | 100,00 |
| 52,345 -69,195 | 77,78 | 0,00 | 11,11 | 11,11 | 100,00 |
| 69,195 -86,045 | 100,00 | 0,00 | 0,00 | 0,00 | 100,00 |
| Total geral | 12,80 | 26,80 | 44,00 | 16,40 | 100,00 |

Se considerarmos clientes mais abastados os das 3 últimas classes (mínimo de R\$ 14730) observamos o oposto do que era esperado: as opiniões positivas (satisfeito e bastante satisfeito) representam pelo menos 76% dos clientes, bem acima da opinião geral (que chega a 29,20%). É na menor classe (até 18,645 salários mínimos – R\$ 7737) que se encontram os mais insatisfeitos, nesta classe nada menos do que 79,59% dos clientes declaram-se muito insatisfeitos com seus veículos, o que muito preocupante...

UNIDADE 3

1) A variável anos de remodelação dos veículos (na percepção do cliente) está representada na distribuição de frequências expressa no quadro a seguir:

| Anos de remodelação | Frequências |
|---------------------|-------------|
| 0 | 2 |
| 1 | 57 |
| 2 | 123 |
| 3 | 59 |
| 4 | 9 |
| Total | 250 |

Fonte: elaborado pelo autor.

a) Calcule a média, mediana, moda e quartis da variável anos de remodelação.

b) A direção da Toyord acredita que se uma parcela considerável dos clientes perceber que seus modelos são atualizados (foram remodelados há no máximo 2 anos) o design e o marketing dos veículos estão coerentes. Com base nos resultados da letra a, os dados mostram isso? Justifique.

c) Calcule o intervalo, desvio padrão e coeficiente de variação percentual da variável anos de remodelação.

d) Com base nos resultados dos itens a e c, você considera que os dados estão fortemente concentrados em torno da média? JUSTIFIQUE.

A resolução deste exercício começa semelhante ao 2º exemplo da Unidade 4 do Livro, para calcular medidas a partir de uma tabela de frequências.

a) Para calcular as medidas de posição precisamos:

- Média

Precisamos multiplicar a coluna de valores (anos de remodelação, nosso x_i) pela das frequências f_i , somar os resultados, e dividi-los por 250, que é o número de elementos do conjunto (n). Observe que a variável discreta pode assumir 5 valores diferentes, logo $k = 5$. No quadro abaixo podemos observar o resultado:

| Anos de remodelação x_i | Frequências f_i | $f_i \times x_i$ |
|---------------------------|-------------------|------------------|
| 0 | 2 | 0 |
| 1 | 57 | 57 |
| 2 | 123 | 246 |
| 3 | 59 | 177 |
| 4 | 9 | 36 |
| Total | 250 | 516 |

Agora podemos calcular a média:

$$\bar{x} = \frac{\sum_{i=1}^k (x_i \times f_i)}{n} = \frac{\sum_{i=1}^5 (x_i \times f_i)}{250} = \frac{516}{250} = 2,064 \text{ anos de remodelação.}$$

- Moda

Basta encontrar o valor que ocorre com maior frequência, rapidamente encontramos o valor 2, que possui a maior frequência, logo $\text{Moda} = 2$.

- Mediana

Precisamos obter a posição da mediana.

Posição mediana = $(n + 1)/2 = (250 + 1)/2 = 125,5^a$. Esta posição, obviamente não existe, precisamos calcular a média entre os elementos que estão na 125ª e 126ª posições respectivamente. Temos que acrescentar uma coluna com as frequências acumuladas, o que é relativamente simples basta somar a frequência de um valor com as de todos os anteriores:

| Anos de remodelação x_i | Frequências f_i | Freq. Acumulada |
|---------------------------|-------------------|-----------------|
| 0 | 2 | 2 |
| 1 | 57 | 59 |
| 2 | 123 | 182 |
| 3 | 59 | 241 |
| 4 | 9 | 250 |
| Total | 250 | 2 |

Até o valor 0 temos duas posições (1ª e 2ª). O valor 1 ocupa 57 posições (da 3ª à 59ª), e o valor 2 nada menos do que 123 posições (da 60ª à 182ª): portanto os valores que estão na 125ª e 126ª posições são iguais a 2. Logo, $Md = (2+2)/2 = 2$.

- Quartis

Precisamos obter as posições dos quartis.

Posição do quartil inferior = $(n + 1)/4 = (250 + 1)/4 = 62,75^a$.

Posição do quartil superior = $[3 \times (n+1)]/4 = [3 \times (250 + 1)]/4 = 188,25^a$.

Ambas as posições não existem, precisamos obter as médias dos valores que estão na 62ª e 63ª posições para calcular o quartil inferior, e dos valores que estão na 188ª e 189ª posições para chegar ao quartil superior. Observando as frequências acumuladas podemos verificar que da 60ª à 182ª posição os valores são iguais a 2, incluindo as 62ª e 63ª posições, o que nos leva a concluir que o quartil inferior vale $(2+2/2) = 2$. Da 183ª à 241ª posições os valores são iguais a 3, incluindo as 188ª e 189ª posições, o que nos leva a concluir que o quartil superior vale $(3+3/2) = 3$.

b) Observe que a mediana vale 2, indicando que 50% dos clientes consideram que os veículos foram atualizados há no máximo 2 anos atrás. A média vale 2,064 anos, muito próxima do valor da mediana. A moda também vale 2, o que é um indicador importante, pois a variável é discreta e somente pode assumir valores de 0 a 4. Não obstante, o quartil superior vale 3, indicando que 25% dos clientes acreditam que os veículos foram modernizados há 3 anos ou mais. Os executivos da Toyord deveriam ficar um pouco apreensivos pois a mediana indica que apenas metade dos clientes considera que os veículos foram atualizados há no máximo 2 anos.

c) Para calcular as medidas de dispersão

- Intervalo: basta observar os valores máximo e mínimo na tabela de frequências, respectivamente 4 e 0 anos; assim, o intervalo vale $4 - 0 = 4$ anos, ou expresso pelos limites $[0, 4]$ anos.

- Desvio padrão: é preciso usar a fórmula do desvio padrão para o caso em que os dados estão em uma tabela de frequências (veja Unidade 4, página 109).

$$s = \sqrt{\frac{\sum_{i=1}^k (x_i^2 \times f_i) - \frac{\left(\sum_{i=1}^k x_i \times f_i\right)^2}{n}}{n-1}} \quad (\text{amostra})$$

Precisamos encontrar alguns somatórios e o valor de n . Pela tabela sabemos que $n = 250$. O resultado $\sum_{i=1}^k (x_i \times f_i)$ já foi encontrado na letra a. Mas resta o resultado $\sum_{i=1}^k (x_i^2 \times f_i)$. Podemos

estender a tabela usada na letra a:

| Anos de remodelação x_i | Frequências f_i | $f_i \times x_i$ | x_i^2 | $f_i \times x_i^2$ |
|---------------------------|-------------------|------------------|---------|--------------------|
| 0 | 2 | 0 | 0 | 0 |
| 1 | 57 | 57 | 1 | 57 |
| 2 | 123 | 246 | 4 | 492 |
| 3 | 59 | 177 | 9 | 531 |
| 4 | 9 | 36 | 16 | 144 |
| Total | 250 | 516 | - | 1224 |

Agora basta substituir na equação:

$$s = \sqrt{\frac{\sum_{i=1}^k (x_i^2 \times f_i) - \frac{\left(\sum_{i=1}^k x_i \times f_i\right)^2}{n}}{n-1}} = \sqrt{\frac{\sum_{i=1}^5 (x_i^2 \times f_i) - \frac{\left(\sum_{i=1}^5 x_i \times f_i\right)^2}{n}}{n-1}} = \sqrt{\frac{1224 - \frac{(516)^2}{250}}{250-1}} = 0,8 \text{ anos}$$

Então o desvio padrão foi de 0,8 anos.

- Coeficiente de variação percentual: a expressão desta medida está na Unidade 4, página 110 do livro texto.

$$\text{c.v.\%} = \frac{s}{\bar{x}} \times 100\% \quad \text{Basta substituir os valores da média (calculada na letra a) e do desvio}$$

padrão para obter o coeficiente de variação percentual:

$$\text{c.v.\%} = \frac{s}{\bar{x}} \times 100\% = \frac{0,8}{2,024} \times 100 = 38,76\% \quad \text{Então o desvio padrão representa 38,76\% da}$$

média.

d) Para avaliar se os dados estão concentrados em torno da média devemos examinar o resultado do coeficiente de variação percentual. Não há uma regra que defina o valor até o qual se consideram os dados “concentrados” em torno da média. No presente caso o cv% vale 38,76%, desvio padrão representa pouco mais de um terço da média, podemos afirmar que os dados não estão fortemente concentrados, para que isso ocorresse o cv% deveria ser bem menor, algo em torno de 5%, o que não é o caso.

2) Na questão 5 das atividades de aprendizagem da Unidade 2 foi dito: “os executivos da Toyord creem que seus clientes mais abastados são mais críticos, tendem a ser mais insatisfeitos com seus veículos”. Naquela questão foi construída uma distribuição de frequências conjunta, relacionando a renda agrupada em classes com a opinião geral dos clientes sobre seus veículos, para verificar se os executivos estavam certos. Agora, analise a renda dos clientes (variável quantitativa) em função da opinião geral dos clientes (através do Microsoft Excel ®), calculando medidas de síntese de renda em função das opiniões.

a) Com base nos resultados os executivos estão certos? JUSTIFIQUE!

b) Compare com as conclusões que você obteve na questão 5 da Unidade 2.

Temos uma situação com duas variáveis: opinião geral (qualitativa ordinal) e renda (quantitativa contínua). Podemos usar os procedimentos descritos na seção 2.3.2 do arquivo “Como fazer Análise Exploratória de Dados usando o Microsoft Excel”, disponível no ambiente virtual, para calcular as medidas de síntese de renda em função das opiniões. Lembre-se que é possível apenas obter as medidas mínimo, máximo, média e desvio padrão (e o coeficiente de variação percentual a partir destes), e o resultado será:

| <i>Opinião</i> | <i>Renda</i> | |
|----------------------------|----------------------|---------|
| <i>Bastante satisfeito</i> | <i>Mínimo</i> | 20,220 |
| | <i>Máximo</i> | 86,015 |
| | <i>Média</i> | 46,791 |
| | <i>Desvio padrão</i> | 16,870 |
| | <i>cv%</i> | 36,055% |
| <i>Satisfeito</i> | <i>Mínimo</i> | 13,055 |
| | <i>Máximo</i> | 58,405 |
| | <i>Média</i> | 32,291 |
| | <i>Desvio padrão</i> | 10,175 |
| | <i>cv%</i> | 31,510% |
| <i>Insatisfeito</i> | <i>Mínimo</i> | 10,820 |
| | <i>Máximo</i> | 43,575 |
| | <i>Média</i> | 24,701 |
| | <i>Desvio padrão</i> | 7,014 |
| | <i>cv%</i> | 28,397% |
| <i>Muito insatisfeito</i> | <i>Mínimo</i> | 1,795 |
| | <i>Máximo</i> | 65,390 |
| | <i>Média</i> | 16,267 |
| | <i>Desvio padrão</i> | 9,728 |
| | <i>cv%</i> | 59,803% |
| <i>TOTAL</i> | <i>Mínimo</i> | 1,795 |
| | <i>Máximo</i> | 86,015 |
| | <i>Média</i> | 25,062 |
| | <i>Desvio padrão</i> | 14,492 |
| | <i>cv%</i> | 57,825% |

a) A afirmação dos executivos pode ser traduzida da seguinte forma: quanto pior a opinião sobre o veículo, maiores devem ser as rendas dos clientes. Analisando as medidas de síntese podemos verificar se isso é verdade.

A tabela acima desmente a afirmação cabalmente. Os que apresentam a opinião “Bastante satisfeito” possuem os maiores valores de mínimo (20,220 salários mínimos), máximo (86,015 salários mínimos, que é o máximo do conjunto todo), e média (46,791 salários mínimos). À medida que a opinião se torna pior as médias vão caindo progressivamente, até chegar ao caso de “Muito insatisfeito”, onde vale 16,267 salários mínimos (30 salários mínimos abaixo da média da opinião “Bastante satisfeito”).

b) As conclusões obtidas na letra a coincidem com as da questão 5 da Unidade 2. Lá, a análise feita acima foi realizada construindo tabelas agrupadas em classe para a renda, em função de cada opinião: encontramos maior número de clientes nas classes mais elevadas de renda para as opiniões mais positivas sobre os veículos.

Além disso, podemos concluir que HÁ relação entre as variáveis opinião sobre o veículo e renda dos clientes, pelos seguintes motivos:

- a média total de renda (considerando os 250 clientes da amostra) vale 25,062 salários mínimos; as médias de renda para as opiniões “Bastante satisfeito”, “Satisfeito” e “Muito insatisfeito” distanciam-se muito da total (21 salários mínimos acima, 7 salários mínimos acima e 9 salários mínimos abaixo, respectivamente).

- o mínimo do conjunto total vale 1,795 salários mínimos, e é verificado na opinião “Muito insatisfeito”, os mínimos das outras opiniões aumentam progressivamente à medida que as respostas tornam-se mais positivas sobre os veículos.

- o máximo do conjunto total vale 86,015 salários mínimos, e é verificado na opinião “Bastante satisfeito”, os máximos das outras opiniões diminuem progressivamente à medida que as respostas

tornam-se mais negativas sobre os veículos, com exceção da opinião “Muito insatisfeito” cujo máximo de renda (65,390 salários mínimos) é superior aos máximos das opiniões “Satisfeito” e “Insatisfeito”;

- no que tange à dispersão, avaliando os coeficientes de variação percentual (pois as médias de renda nas opiniões são diferentes) podemos ver que o desvio padrão total da renda representa cerca de 57% da média total;

- somente a opinião “Muito insatisfeito” apresenta dispersão da renda compatível ($cv\% = 59,803\%$), o que provavelmente se deve a valores discrepantes como o máximo igual a 65,390 salários mínimos;

- para todas as outras opiniões a dispersão é menor;

- na opinião “Bastante satisfeito” encontramos o segundo maior $cv\%$ (36,055%), provavelmente porque o seu máximo é o máximo total da renda (86,015 salários mínimos).