Unidade Temática 4: Resumo Numérico de Dados

Medidas de tendência central

O tratamento e interpretação torna-se difícil para quem pretende colocar uma informação como um todo, razão pela qual constuma-se calcular algumas medidas que resumidamente traduzem as importantes caracteristicas das distribuições de frequencias.

Existem quatro grupos de medidas numéricas que resumem um conjunto de dados observados. As medidas de tendência central, medidas de localização ou separatrizes, as medidas de dispersão e as características de forma de distribuição.

Os índices de centro de distribuição ou medidas de tendência central, assim designados em vertude da tendência que todos os dados observados têm em torno desses valores centrais. Este grupo inclui as grandezas médias, a moda e a mediana.

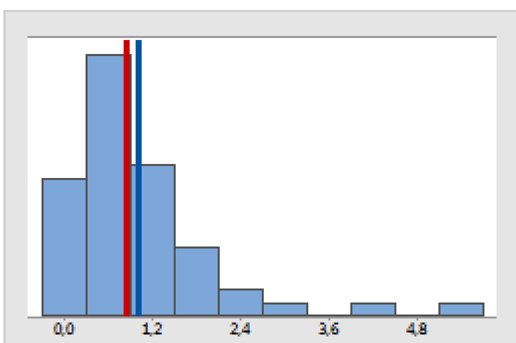
A média Aritmétrica: é uma grandeza ou valor representativo de um conjunto de dados como eles tendem a se localizar em torno do ponto central. A média aritmética é a mais usada nas investigações, e esta pode ser simples ou ponderada quando temos frequências.

Amostra: $\bar{x} = \frac{1}{n}\sum x_i$ simples; $\bar{x} = \frac{1}{n}\sum x_i \times f_i$ ponderada, usado para dados agrupados.

População $\mu = \frac{1}{N}\sum x_i$ simples $\mu = \frac{1}{N}\sum x_i \times f_i$ ponderada, usado para dados agrupados.

Média aparada

A média aparada é a média dos dados sem os 5% maiores e os 5% menores valores. Usa-se a médias aparadas para eliminar o impacto de valores muito maiores ou muito menores da média (valores extremos). Quando os dados contêm outliers, a média aparada pode ser uma medida melhor da tendência central do que a média ariginal.



A linha azul (a 2ª da esquerda para direita) representa a média original, que é fortemente influenciada pelos valores extremos à direita. A linha vermelha (a 1ª da esquerda para direita) representa a média aparada, que se desloca para a esquerda porque foram excluidos os valores extremos nos 5% maiores dados.

CARACTERÍSTICAS DA MÉDIA ARITMÉTICA

- ✓ valor da média é determinado em função de todos os valores da distribuição;
- ✓ O valor da média é influenciado pelos valores extremos da distribuição;
- ✓ A média independe da ordem das classes (todas as classes) igualmente importantes no seu cálculo), não dizendo nada sobre a distribuição das classes;
- ✓ A média tem um valor determinado.

VANTAGENS DA MÉDIA ARITMÉTICA

- ✓ É a mais comum das médias, mais compreensível e a mais reconhecida como média propriamente dita, pois se presta a tratamento algébrico por ter seus termos matematicamente bem definidos;
- ✓ Seu cálculo é simples, para realiza-los são necessários somente valores totais e número de classes ou valores.

DESVANTAGENS

- ✓ Seu valor pode ficar alienado devido à influência de valores extremos por não representar bem o conjunto de valores de uma Série Simples ou da Distribuição de Frequência.

Mediana: é uma separactriz que devide a distribuição de frequência ou conjunto de dados em duas partes iguais. A mediana é o valor que está na posição do meio para "n impar" e é igual a média aritmética dos dois valores centrais para "n par".,

isto é:

Quando n é impar: $mediana = X_p = X_{\left(\frac{n+1}{2}\right)}$

Quando n é par: $mediana = \dfrac{X_{\left(\frac{n}{2}\right)} + X_{\left(\frac{n}{2}+1\right)}}{2}$

Para dados agrupados em classe, a mediana é determinada por aproximação, bastando conhecer a classe mediana, que é localizada calculando as frequências acumuladas.

$$Me = Li + \frac{\frac{n}{2} - Fa}{f_{me}} \times a$$ Onde

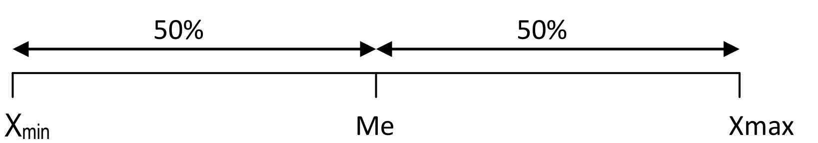
Li – Limite inferior da classe mediana

f_{me} - irequência da classe mediana

Fa – Frequência acumulada da classe anterior a classe mediana

a – Amplitude da classe mediana

REPRESENTAÇÃO GEOMÉTRICA DA MEDIANA



Moda

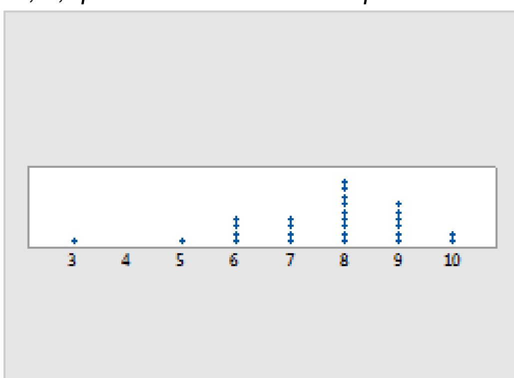
O termo Moda ou Norma (valor dominante de um conjunto) foi introduzido por **Karl Pearson** em 1895 e originou-se talvez na maneira de se falar “tal coisa está na moda”, “tal coisa é mais freqüente”.

Portanto, define-se Moda ao valor da variável que ocorre com mais frequência. Para casos em que o número das observações não é muito elevado esta medida pode ser lida no rol.

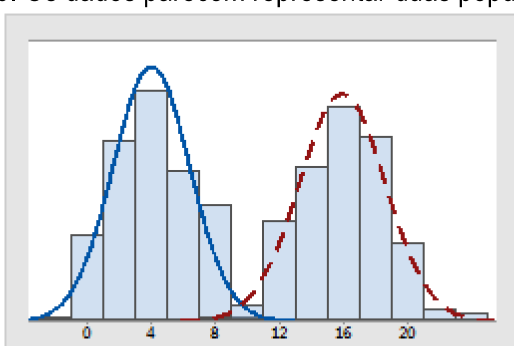
A moda pode não existir, e mesmo que exista, pode não ser única. Assim, um conjunto de valores ou uma Distribuição de Freqüência pode ter várias modas e é dito ser “multimodal”. Se tem duas modas, bimodal e assim por diante. De acordo com as frequências de cada valor observado, é possível

- não ter a moda (amodal),
- ter uma moda (unimodal),
- duas modas (bimodal)
- três modas (trimodal),
- muitas modas (plurimodal ou simplesmente, multimodal).

Unimodal: *Existe somente um modo, 8, que ocorre com mais frequência.*



Bimodal: Existem dois modos, 4 e 16. Os dados parecem representar duas populações diferentes.



Exemplos 1:

a) X={2, 2, 5, 7, 9, 9, 9, 10, 10, 11, 12, 18} → **Mo=9**;

b) X={3, 5, 8, 10, 12, 15, 16} → **<u>não</u> tem moda**;

c) X={ 2, 3, 4, 4, 5, 5, 7, 7, 7, 9} → **é bimodal: Mo_1=4** e **Mo_2=7**.

Nas Distribuições de Freqüência em classe, não temos o valor que mais se repete e sim a classe que possui a maior frequência. Nestas condições, é dita **CLASSE MODAL** (classe de maior freqüência).

Quando temos dados agrupados em classe, a moda é calculada por aproximação, começando por localizar a classe modal (aquela classe que tiver maoir frequencia absoluta), se necessário deve –se conhecer a moda bruta que é o ponto médio da classe que contém a moda. As formulas de King e de Czuber são alternativas para calcular o valor da moda com aproximação.

Formula de King

$$Mo = Li + \frac{fp}{fa + fp} \times a$$

A formula de Czuber é uma das alternativas para calcular o valor da moda com aproximação.

Neste processo considera-se a influência, sobre a classe modal, das freqüências das classes adjacentes, anterior e posterior, assim como da freqüência da classe modal. O ponto correspondente à moda divide o intervalo em partes proporcionais às diferenças respectivas da freqüência da classe modal, com as freqüências adjacentes.

$$Mo = L_i + \frac{fmo - fa}{2 \times fmo - (fa + fp)} \times a$$ Onde

Li - Limite inferior da classe modal

f_{mo} - Frequência da classe modal

fa – Frequência anterior a classe modal

fp – Frequência posterior a classe modal e

a – Amplitude da classe modal

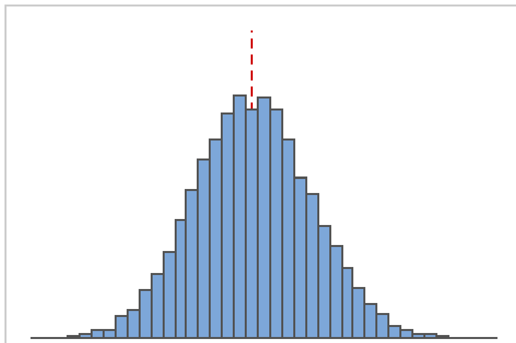
Nota: São vantagens e desvantagens de cada uma das medidas de tendência central:

- ✓ A moda é uma medida que requer apenas o conhecimento da frequência absoluta e pode ser utilizada para qualquer tipo de variáveis, tanto qualitativas, quanto quantitativas.
- ✓ A mediana é uma medida que exige uma ordenação de categorias, da mais baixa a mais alta (ou *vice*-versa), assim ela só pode ser obtida para variáveis qualitativas ordinais ou para as quantitativas, jamais para variáveis qualitativas nominais. Além disso, a mediana não é influenciada por valores extremos.
- ✓ A média aritmética trabalha com todos os elementos do conjunto de dados, enquanto a mediana utiliza apenas um ou dois valores. No entanto a média sofre influência de valores extremos (muito alto ou baixo).
- ✓ A média é uma medida que pode ser calculada apenas para variáveis quantitativas. E embora a média seja um valor mais fácil de entender, tem o defeito de nos induzir ao erro se a nossa amostra tiver valores muito extremos.

Assim, no caso das variáveis quantitativas, quando o valor da Mediana é muito diferente da Média, é aconselhável considerar sempre a Mediana como valor de referência mais importante.

O centro dos dados é a área onde a maioria dos valores em um conjunto de dados se agrupam. A tendência central pode ser descrita por diversas estatística, como média, média aparada, mediana ou modo. Conhecer a tendência central dos dados é um primeiro passo importante para entender os dados.

Gráficos como histogramas, boxplots e dotplots são úteis para visualizar a tendência central dos dados e pode ajudar a decidir qual estatística de tendência central é mais apropriada para um determinado conjunto de dados.



Exemplo 2: Os dados que se seguem representam as percentagens do valor total que 17 credores do Millennium Bim retornaram ao banco durante um ano.

32.2 , 29.5 , 29.9 , 32.4 , 30.5 , 30.1 , 32.1 , 35.2 , 10.0 , 20.6 , 28.6 , 30.5 , 38.0 , 33.0 , 29.4 , 37.1 , 28.6

Determine a média, moda e mediana dos 17 credores.

Resolução

Média. $\bar{x} = \dfrac{\sum x_i}{n} = \dfrac{507.7}{17} = 29.86$

Moda: Para localizar o valor que está na moda vamos escolher o valor que ocorre mais vezes

10.0 20.6 **28.6** **28.6** 29.4 29.5 29.9 30.1 **30.5** **30.5** 32.1 32.2 32.4 33 35.2 37.1 38

Do rol temos dois volores que ocorrem mais vezes (duas vezes cada), portanto, temos duas modas :

$Mo_1 = 28.6$, $Mo_2 = 30.5$, assim, diremos que temos uma distribuição bimodal.

Mediana: Para localizar o valor mediano, primeiro, vamos ordenar os dados em uma ordem crescente. Como o n é impar (n=17), a posição (P) em que se encontra o valor mediano será: $P(me) = \dfrac{n+1}{2} = \dfrac{17+1}{2} = 9$, portanto, o elemento mediano está na posição 9, vejamos:

1ª	2ª	3ª	4ª	5ª	6ª	7ª	8ª	**9ª**	10ª	11ª	12ª	13ª	14ª	15ª	16ª	17ª
10.0	20.6	28.6	28.6	29.4	29.5	29.9	30.1	**30.5**	30.5	32.1	32.2	32.4	33	35.2	37.1	38

Assim, a mediana é 30,5. Isto significa que 50% das pessoas pagaram menos de 30,5% do valor que receberam quando pediram o emprestimo e outras 50% das pessoas devolveram igual ou maior que 30,5% do total.

Exemplo 3: Para os dados da tabela abaixo, dertemine a media , mediana e a moda.

x_i	f_i	$x_i f_i$
5	1	5 x 1 = 5
6	5	6 x 5 = 30
7	6	7 x 6 = 42
8	5	8 x 5 = 40
9	3	9 x 3 = 27
$\sum$	20	144

- $\bar{x} = \frac{\sum x_i f_i}{n} = \frac{144}{20} \Rightarrow \bar{x} = 7{,}2$ Mediana = 7 Moda = 7

Exemplo 4: A partir da tabela abaixo, determine a média, moda e mediana da distribuição.

Classes	xi	fi	xi*fi	Fi
10 - 20	15	10	150	10
20 - 30	25	20	500	30
30 - 40	35	35	1225	65
40 - 50	45	**40**	1800	**105**
50 - 60	55	25	1375	130
60 - 70	65	15	975	145
70 - 80	75	5	375	150
Total	**-**	**150**	**6400**	**-**

Resolução

Média: $\bar{x} = \frac{\sum x_i \times f_i}{n} = \frac{6400}{150} = 42{,}67$

Moda: $Mo = Li + \frac{fp}{fa + fp} \times a = 40 + \frac{25}{(25+35)} \times 10 = 44{,}17$

Mediana: $Me = Li + \frac{\frac{n}{2} - Fa}{f_{me}} \times a = 40 + \frac{75-65}{40} \times 10 = 42{,}5$

ESTATISTICA BÁSICA - 2019

FICHA DE EXERCICIOS 4: *Medidas de Tendencia central*

1. Responda se cada uma das proposições a seguir é verdadeira ou falsa. Se a proposição for falsa, corrija a palavra sublinhada para que se torne verdadeira.
   a) Metade dos valores de uma variável quantitativa são sempre menores que a <u>média</u>.
   b) A <u>mediana</u> é a melhor medida de tendencia central para variáveis qualitativa nominal.

2. Quando seus dados forem qualitativos nominais, podemos usar como medida de tendência central: a média? a mediana? a moda?
3. Toda distribuição tem média, mediana e moda?
4. Determine a posição mediana para n=47 e n=64.

5. Para os seguintes conjuntos de dados, determine os valores da média aritmética, mediana e moda.

a) 12,15, 16, 15, 12, 15, 15, 5, 7, 14　　　b) 2, 6, 3, 6, 3, 3, 4

6. As idades dos 11 alunos de uma turma de matemática são respectivamente iguais a:
   11; 11; 11; 12; 12; 13; 13; 13; 13; 15; 16. A moda e a mediana desses 11 valores correspondem a:
   <u>a) 16, 12</u>　<u>b) 12, 11</u>　<u>c) 15, 12</u>　<u>d) 13, 13</u>　<u>e) 11, 13</u>

7. Observando a distribuição abaixo, podemos dizer que a mediana é:
   17, 12, 9, 23, 14, 6, 3, 18, 42, 25, 18, 12, 34, 5, 17, 20, 7, 8, 21, 13, 31, 24, 9
   <u>a) 13,5</u>　<u>b) 17</u>　<u>c) 14,5</u>　<u>d) 15,51</u>　<u>e) 14</u>

8. Nos conjuntos de valores abaixo:
   A = {3, 5, 6, 8, 9, 10, 10, 10, 11, 12, 17}

   B = {4, 5, 7, 10, 11, 13, 15}

   C = {2, 3, 4, 5, 5, 5, 5, 6, 7, 8, 8, 8, 8, 9, 10, 11}

Em relação à moda podemos dizer que:	Então:
I. A é unimodal e a moda é 10;	a) Estas afirmações estão todas corretas;
II. B é unimodal e a moda é 10;	b) Estas afirmações estão todas erradas;
III. C é bimodal e as modas são 5 e 8	c) I e II estão corretas
	d) I e III estão corretas;
	e) II e III estão corretas.

9. O Paulinho, o Toninho e o Pedrinho são três atacantes de uma equipa de futebol. Nesta época, o Paulinho e o Pedrinho já fizeram cinco jogos e o Toninho quatro. O número de chutes à balisa, nos jogos realizados foi o seguinte:

Paulinho	7	8	3	10	7
Toninho	5	5	0	4	4
Pedrinho	9	8	10	5	0

   a) Qual deles fez a melhor média?
   b) Qual é a Moda, para cada um dos jogadores?
   c) Qual é a Mediana, de cada um dos jogadores?

10. As notas de estatística de uma turma estão na tabela seguinte.

Nota	1	2	3	4	5	6	7	8	9
Nº de alunos	2	6	9	12	14	9	5	4	1

- Calcule os valores da média, da mediana e da moda dessas notas.

11. A tabela seguinte apresenta a produção de café, em milhões de toneladas, na região DELTA.

ANO	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002
Ton	12	15	18	22	17	14	18	23	29	32	34

Calcule o valor da produção média, da moda e da produção mediana.

12. A tabela abaixo informa o número de pessoas atendidas de urgência no HPS de certa cidade no período de 22 dias.

Nº DE ATENDIMENTOS	0	1	2	3	4
Nº DE DIAS	4	6	9	2	1

- Qual a média, a moda e a mediana dos atendimentos?

13. Uma empresa possui dois serventes recebendo salários de $600,00 cada um, quatro escriturários recebendo $2.000,00 cada um, um chefe de escritório com salário de $3.500,00 e três técnicos recebendo $6.000,00 cada um. Determine o salário médio, mediano (mediana) e salário modal (moda).

14. Em um posto de controle rodoviário, doze motoristas multados por excesso de velocidade estavam dirigindo a
8 11 14 6 8 10 20 11 13 18 9 15 quilômetros por hora acima do limite regular de velocidade.

a) Em média, em quantos quilômetros por hora esses motoristas estavam excedendo o limite?
b) Se um motorista que excedia o limite em menos de 15 quilômetros por hora foi multado em $ 150,00 e os outros foram multados em $ 220,00, determine a média das multas que esses motoristas tiveram de pagar.

15. A tabela abaixo representa os salários pagos a 100 operários da empresa XYZ.

Nº. de salários mínimos	0 – 2	3 – 5	6 – 8	9 – 11	12 – 14
Nº. de operários	40	30	10	15	5

Determine:

a) o salários médio;
b) o intervalo que contém o valor da mediana e o intervalo que contém a moda.
c) a moda ***Bruta*** desta distribuição.
d) a moda (usando a formula de ***King*** e a formula de ***Czuber***) e a mediana desta distribuição.

16. A tabela a baixo mostra as alturas de 50 atletas:

Altura (Cm)	145 - 150	150 - 155	155 - 160	160 - 170	170 - 175	175 - 180
Frequência	3	6	16	16	6	3

Determine a média, mediana e a moda.

17. Uma distribuidora de refrigerantes fez um levantamento sobre o consumo semanal (em litros) por pessoa, em jan/2002, em uma cidade do litoral, obtendo a tabela abaixo:

CONSUMO	0,5 ├── 1,0	1,0 ├── 1,5	1,5 ├── 2,0	2,0 ├── 2,5	2,5 ├── 3,0
Nº DE PESSOAS	10	25	9	7	6

a) Determine e interprete o consumo médio.
b) Qual o percentual de pessoas que consomem menos de 1 litro por semana?
c) Determine os intervalos que contém o consumo modal e o consumo mediano.
d) a moda ***Bruta*** desta distribuição.
e) Determine a moda (usando a formula de ***King*** e a formula de ***Czuber***) e a mediana desta distribuição

18. João deseja calcular a média das notas que tirou em cada uma das quatro matérias a seguir. Calcule a média ponderada de suas notas, sendo que as duas primeiras provas valem 2 pontos e as outras duas valem 3 pontos:

Inglês	
1ª prova	6,5
2ª prova	7,8
3ª prova	8,0
4ª prova	7,1

Português	
1ª prova	7,5
2ª prova	6,9
3ª prova	7,0
4ª prova	8,2

19. Joanita, deseja calcular a média das notas que tirou em cada uma das quatro matérias a seguir. Calcule a média ponderada de suas notas, sendo que a primeira prova vale 3 pontos, a segunda vale 2 pontos, a terceira vale 4 pontos e quarta vale 5 pontos:

História	
1ª prova	5,4
2ª prova	8,3
3ª prova	7,9
4ª prova	7,0

Matemática	
1ª prova	8,5
2ª prova	9,2
3ª prova	9,6
4ª prova	10,0

20. No conjunto de dados abaixo, calcular a média aritmética e média aparada, com m = 2
    90, 100, 330, 350, 400, 520, 610, 730, 800, 1500, 1700, comente o resultado entre as médias.

21. Nos quatro primeiros dias úteis de uma semana o gerente de uma agência bancária atendeu 19, 15, 17 e 21 clientes. No quinto dia útil dessa semana esse gerente atendeu n clientes. Se a média do número diário de clientes atendidos por esse gerente nos cinco dias úteis dessa semana foi 19, a mediana foi de quanto?

22. Considere um grupo formado por cinco amigos com idade de 13, 13, 14, 14 e 15 anos. O que acontece com a média de idade desse grupo, se um sexto amigo com 16 anos juntar-se ao grupo?
    a) permanecerá a mesma
    b) diminuiu 1 ano
    c) aumenta 12 anos
    d) aumenta mais de 1 ano
    e) umenta menos de 1 ano

23. A média aritmética entre 50 números é igual a 38. Dois números são retirados: o número 55 e o 21. Calcule a média aritmética dos números que restaram.
a) 32 b) 38 c) 34 d) 45 e) 24

24. A média aritmética das idades de 10 alunos de uma determinada turma é igual a 15 anos. Se dois alunos, um com 12 anos e outro com 18 anos, saírem dessa turma, a média aritmética das idades dos 8 alunos restantes será igual a : a) 13 b) 14 c) 15 d) 16 e) 17

25. A média aritmética das idades de 30 alunos da turma A é 20 anos e a média aritmética das idades de uma outra turma b com 20 alunos é 18 anos. Então a média aritmética das idades dos alunos das duas turmas é:
a) 18 b) 20 c) 21 d) 19,20 e) 17

26. A média aritmética dos pesos de um grupo com 20 pessoas é de 32 kg e a média aritmética de um outro grupo com 80 pessoas é 70 kg. Então a média aritmética dos pesos das pessoas dos dois grupo é:
a) 62,4 b) 51 c) 46,5 d) 41 e) 38

27. A média de idade de um grupo de 30 pessoas, participantes de uma reunião, é de 40 anos. Após a chegada de um novo convidado com 50 anos de idade, a média passa a ser de aproximadamente.
a) 41 anos b) 40,8 anos c) 40,3 anos d) 40 anos e) 30,5 anos

20 Para testar o raciocínio lógico dos estudantes. Um colégio aplicou uma mesma prova para os alunos do 2° e do 3° anos do ensino médio. A nota média da prova, considerando o total dos 150 alunos, foi de 7,80. Sabendo que a nota média dos alunos do 2° ano foi 7,4 e que a nota média dos alunos do 3° anos foi 8,4, o número de alunos do 2° ano desse colégio é: a) 70 b)75 c) 80 d) 85 e) 90

28. João fez 3 pontos a mais que Paulo, e o Paulo fez 6 pontos a mais que Luis. A média aritmética entre os três foi de 6 pontos, quantos pontos fez Paulo? a) 10 b) 7 c) 1 d) 12 e) 18

29. Uma fábrica tem 600 empregados dos quais 430 são homens. A média salarial dos homens é de 4 salários mínimos e a média salarial das mulheres é de 2,5 salários mínimos. Qual é a média salarial dos empregados dessa fábrica?

30. Numa empresa, dez operários têm salários de $ 2.000,00 mensais; doze têm salário de $ 1.000,00 mensais; e oito operários têm salário de $ 1.400,00 mensais. Qual é o salário médio desses operários?

31. Se um aluno já fez dois trabalhos e obteve 8,5 e 5,0, qual deve ser a nota do terceiro trabalho para que a média aritmética dos três seja 7,0?

32. Qual é a média de idade de um grupo em que há 6 pessoas de 14 anos, 9 pessoas de 20 e 5 pessoas de 16 anos?

33. Num determinado país a população feminina representa 51% da população total. Sabendo-se que a idade média (média aritmética das idades) da população feminina é de 38 anos e a da masculina é de 36 anos. Qual a idade média da população?
a) 37,02 anos b) 37,00 anos c) 37,20 anos d) 36,60 anos e) 37,05 anos