

# MAE116 – Noções de Estatística

Farmácia - Noturno - 2º semestre de 2017

Lista de exercícios 2 - Estatística Descritiva II - C A S A (gabarito)

## Exercício 2.

Os dados a seguir correspondem aos recordes de atletas de 10 países na Olimpíada de Los Angeles em 1984 em algumas provas de atletismo.

Mulheres				
País	100m (seg)	400m (seg)	3000m (min)	Maratona (min)
Argentina	11,61	54,50	9,79	178,52
Brasil	11,31	52,80	9,77	168,75
Chile	12,00	54,90	9,37	171,38
Colômbia	11,6	53,26	9,46	165,42
Alemanha	11,01	48,16	8,75	148,53
França	11,15	51,73	8,98	155,27
Portugal	11,81	54,30	8,84	151,20
Canadá	11,00	50,06	8,81	149,50
USA	10,79	50,62	8,50	142,72
Kenya	11,73	52,70	9,20	181,05

Homens				
País	100m (seg)	400m (seg)	3000m (min)	Maratona (min)
Argentina	10,39	46,84	14,04	137,72
Brasil	10,22	45,21	13,62	133,13
Chile	10,34	46,20	13,61	134,03
Colômbia	10,43	46,10	13,49	131,35
Alemanha	10,16	44,50	13,21	132,23
França	10,11	45,28	13,34	132,30
Portugal	10,53	46,70	13,13	128,65
Canadá	10,17	45,68	13,55	131,15
USA	9,93	43,86	13,20	128,22
Kenya	10,46	44,92	13,10	129,75

Figura 3: Recordes de atletas.

- a. Para cada gênero e modalidade, calcule a média, mediana e desvio padrão dos recordes.

**Resposta:** Nas seguintes tabelas apresentamos as medidas para cada gênero:

MUJERES	10m(seg)	400m(seg)	3000m(min)	Maratona(min)
Média	11,40	52,30	9,15	161,23
Mediana	11,45	52,75	9,09	160,34
D.P	0,40	2,16	0,44	13,52

HOMBRES	10m(seg)	400m(seg)	3000m(min)	Maratona(min)
Média	10,27	45,53	13,43	131,85
Mediana	10,28	45,48	13,41	131,79
D.P	0,18	0,96	0,29	2,78

- b. Compare os resultados em a). Em qual modalidade as diferenças quanto ao gênero são maiores? E em qual são menores? **Resposta:** Observando nas tabelas podemos dizer que as diferenças são

# MAE116 – Noções de Estatística

Farmácia - Noturno - 2º semestre de 2017

Lista de exercícios 2 - Estatística Descritiva II - C A S A (gabarito)

maiores nas modalidades de 400m, 3000m e Maratona, sendo as medidas nos homens menores em 400m e Maratona e maior na modalidade de 3000m. Enquanto na modalidade de 100m, as médias e medianas são similares com valores no desvio padrão baixos.

- c. Utilizando gráficos boxplot faça uma comparação do desempenho entre os sexos nas quatro modalidades. Comente.

**Resposta:** Na seguinte figura apresenta-se os boxplots dos recordes de homens e mulheres por modalidade.

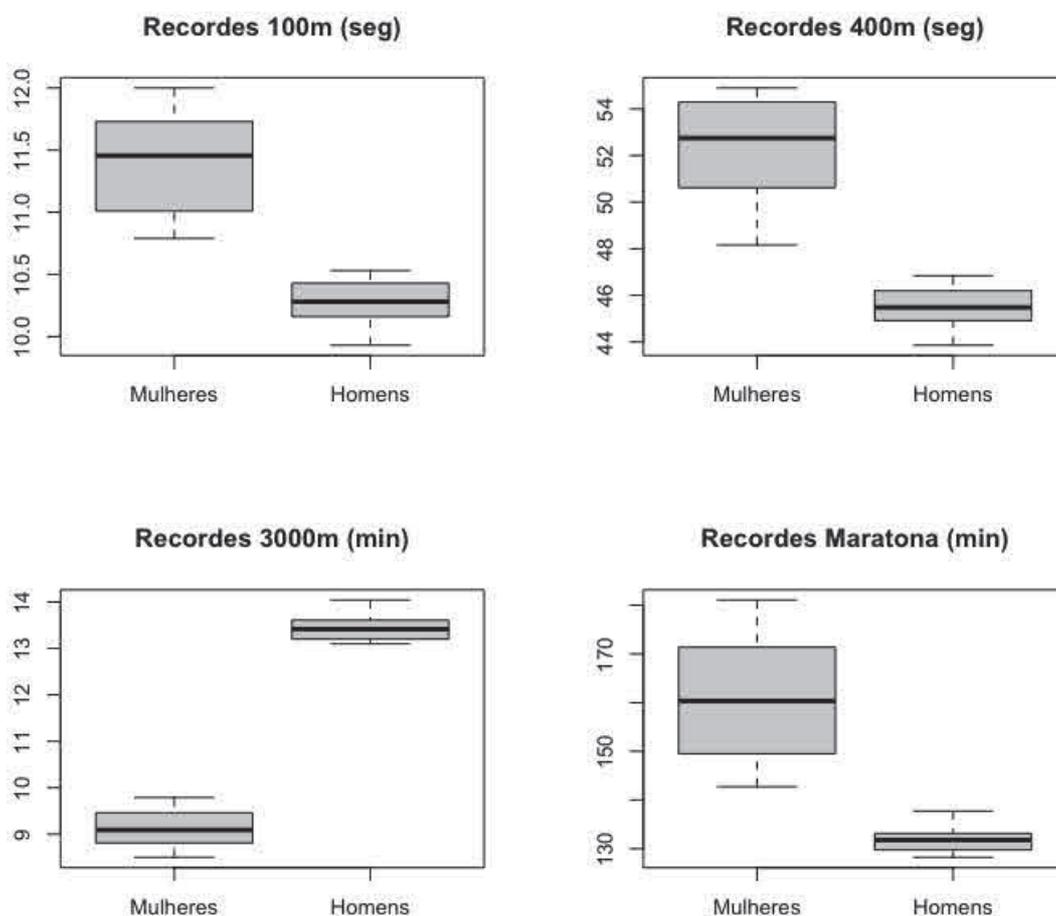


Figura 4: Boxplots por modalidade.

- Recordes 100m : Se observa que o tempo dos homens é menor que o tempos das mulheres, mostrando uma variabilidade maior nas mulheres.
- Recordes 400m : A distribuição dos dados apresenta comportamento similar aos recordes de 100m.
- Recordes 3000m : Observamos um tempo menor nas mulheres que nos homens; além disso a variabilidade nos recordes é baixa.
- Recordes Maratona : Notemos que o tempo dos homens é menor que o tempos das mulheres, também vemos uma variabilidade significativamente maior nas mulheres.

# MAE116 – Noções de Estatística

Farmácia - Noturno - 2º semestre de 2017

Lista de exercícios 2 - Estatística Descritiva II - C A S A (gabarito)

## Exercício 3.

Os dados a seguir representam indivíduos que foram contaminados pelo veneno de um certo tipo de inseto e submetidos a três tipos de tratamentos. As variáveis são:

**Idade:** idade do paciente no momento de admissão, em anos;

**Diag:** tempo, em horas, gasto entre o contato com o inseto e administração do tratamento;

**Recup:** tempo, em horas, entre a administração do tratamento e recuperação;

**Tratam:** tipo do tratamento administrado;

**Coag:** presença de coágulos no momento de admissão.

Pac.No.	Idade	Diag	Recup	Tratam	Coag
1	28	7	3	II	Não
2	15	52	45	I	Não
3	76	30	23	III	Sim
4	15	53	46	I	Sim
5	21	3	2	II	Não
6	11	46	42	I	Não
7	16	55	47	I	Não
8	16	54	47	I	Sim
9	47	13	12	III	Sim
10	18	59	51	II	Não
11	40	20	11	III	Sim
12	24	3	1	II	Não
13	32	9	3	II	Não
14	31	9	3	II	Não
15	10	44	40	I	Sim
16	31	9	3	II	Sim
17	31	10	4	II	Sim
18	46	13	11	III	Sim
19	21	1	2	II	Sim
20	39	17	8	III	Sim
21	15	53	46	I	Sim
22	9	42	39	I	Não
23	75	30	22	III	Sim
24	54	18	16	III	Não
25	35	12	5	II	Sim
26	18	58	50	II	Sim

Figura 5: Dados de indivíduos contaminados.

- a. Classifique cada uma das variáveis;

**Resposta:**

Idade: quantitativa contínua

Diag: quantitativa contínua

Recup: quantitativa contínua

Tratam: qualitativa nominal

Coag: qualitativa nominal

b. Através de representações gráficas adequadas, compare os três tratamentos com relação à idade dos pacientes.

**Resposta:** Primeiramente tem-se que carregar os dados. O boxplot pode ser obtido como segue.

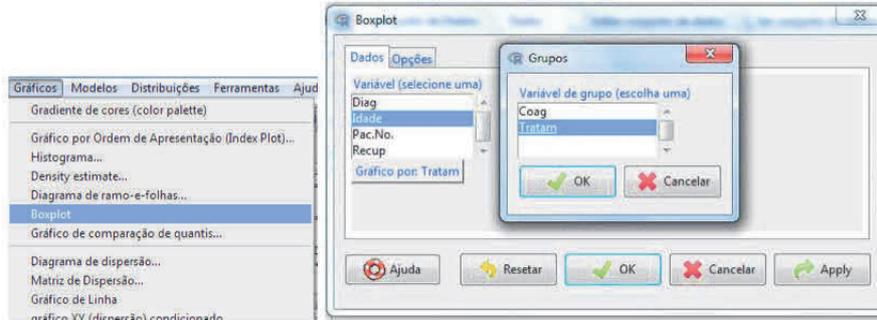


Figura 6

Na Figura 7 apresenta-se o gráfico de boxplot para os três tratamentos com relação à Idade. Nesta figura observa-se que,

- o tratamento III apresentou-se a maior variabilidade,
- as distribuições dos tratamentos parecem ser assimétricas,
- a mediana e demais quartis da variável idade são diferentes nos três tratamentos,
- os pacientes submetidos ao tratamento I são os mais jovens, seguidos dos pacientes submetidos ao tratamento II e em seguida do III,
- em nenhum dos tratamentos há dados discrepantes.

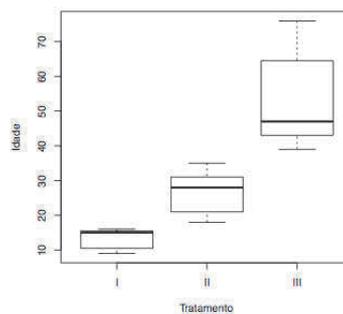


Figura 7: Boxplots da Idade segundo Tratamento.

c. Repita o item (b) para a variável Coag. Suponha que ao invés de trabalhar com a variável Idade, crie-se uma nova variável denominada Etário, assumindo valor 0 se Idade for menor que 29 anos e 1 caso contrário.

**Resposta:** O procedimento para realizar o boxplot é similar ao feito no item (b), neste caso escolhemos a opção Gráfico por Coag. Na Figura 8 apresenta-se o gráfico de boxplot da variável Coag com relação a Idade. Pode-se ver claramente que:

- a presença de coágulos no momento de admissão apresentou maior variabilidade que a não presença,
- nas duas categorias da variável Coag a distribuição da idade parece levemente ser assimétrica,
- a mediana e demais quartis da variável idade são maiores para os pacientes que apresentaram coágulo,
- há uma observação discrepante entre o grupo de pessoas que não tinha coágulo.

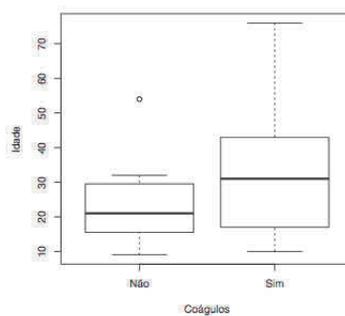


Figura 8: Boxplots da Idade segundo a variável Coag.

Para criar a variável Etário deve-se realizar os seguintes passos:

**Passo1:** Escolher as opções

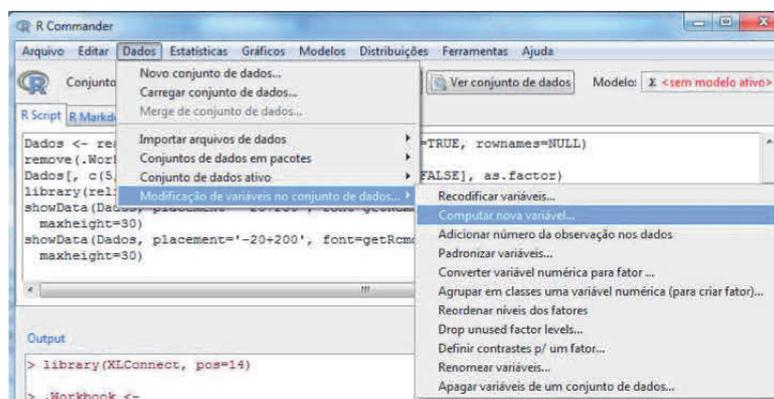


Figura 9

**Passo2:** Escolhemos a variável de interesse (Idade), escrevemos o nome da nova variável (Etário) e indicamos a expressão para computar: `as.numeric(Idade >= 29)`,

MAE116 – Noções de Estatística  
Farmácia - Noturno - 2º semestre de 2017  
Lista de exercícios 2 - Estatística Descritiva II - C A S A (gabarito)

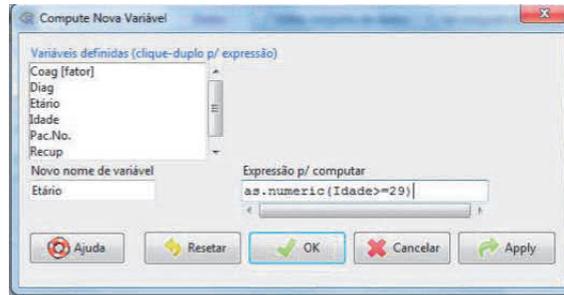


Figura 10

Clicamos em ver conjunto de dados, e vemos que na última coluna da base de dados encontra-se a variável Etário. Note que todas as idades que são menores a 29 anos estão codificadas com o número zero, e as maiores com o número 1.

	Pac.No.	Idade	Diag	Recup	Tratam	Coag	Etário
1	1	28	7	3	II	Não	0
2	2	15	52	45	I	Não	0
3	3	76	30	23	III	Sim	1
4	4	15	53	46	I	Sim	0
5	5	21	3	2	II	Não	0
6	6	11	46	42	I	Não	0
7	7	16	55	47	I	Não	0
8	8	16	54	47	I	Sim	0
9	9	47	13	12	III	Sim	1
10	10	18	59	51	II	Não	0
11	11	40	20	11	III	Sim	1
12	12	24	3	1	II	Não	0
13	13	32	9	3	II	Não	1
14	14	31	9	3	II	Não	1
15	15	10	44	40	I	Sim	0
16	16	31	9	3	II	Sim	1
17	17	31	10	4	II	Sim	1
18	18	46	13	11	III	Sim	1
19	19	21	1	2	II	Sim	0
20	20	39	17	8	III	Sim	1
21	21	15	53	46	I	Sim	0
22	22	9	42	39	I	Não	0
23	23	75	30	22	III	Sim	1
24	24	54	18	16	III	Não	1
25	25	35	12	5	II	Sim	1
26	26	18	58	50	II	Sim	0

Figura 11

- d. Construa um boxplot para a variável Recup, para cada grupo de Etário. Com base nos gráficos, você diria que o tempo de recuperação se modifica com a faixa etária? Justifique sua resposta. Uma nova variável denominada Cura é criada: Cura será rápida se Recup for menor ou igual a 10, será normal se entre 10 e 40 (inclusive) e será lenta para Recup acima de 40.

**Resposta:** Em Gráficos - Boxplot, selecionamos a variável de interesse (Recup) e clicamos em gráfico por grupos (Etário),



Figura 12

Na Figura 13 apresenta-se o gráfico boxplot da variável Recup segundo a variável Etário. Nesta figura pode-se observar que o tempo de recuperação dos pacientes é bem diferente para os grupos etários, sendo maior e mais disperso para o grupo dos mais jovens.

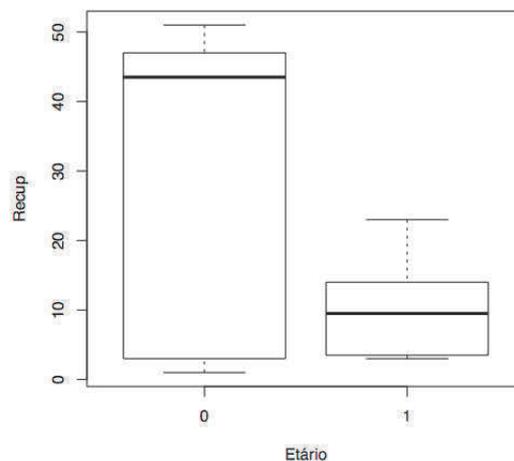


Figura 13: Boxplots da variável Recup segundo a variável Etário.

A variável Cura é uma recodificação da variável Recup, assim que escolhemos as opções, Escolhemos a variável que queremos recodificar (Recup), o nome da variável a criar (Cura) e definimos a recodificação, Visualizamos os dados,

# MAE116 – Noções de Estatística

Farmácia - Noturno - 2º semestre de 2017

Lista de exercícios 2 - Estatística Descritiva II - C A S A (gabarito)

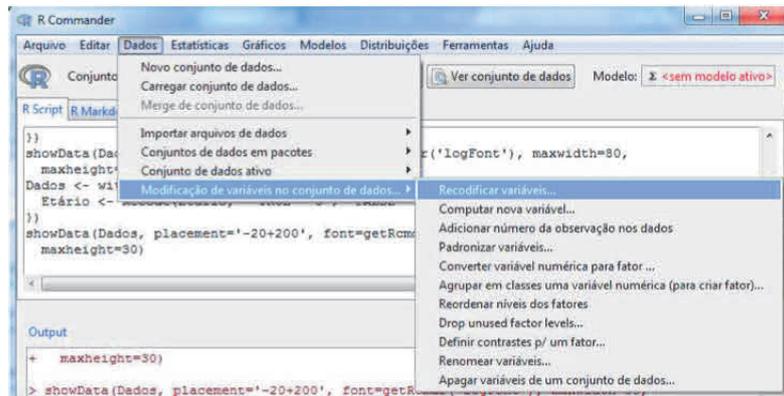


Figura 14

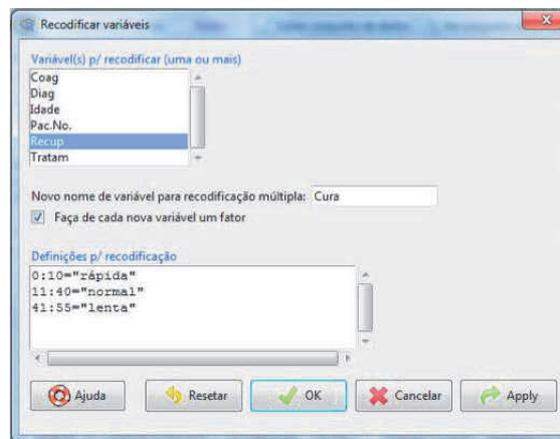


Figura 15

	Pac.No.	Idade	Diag	Recup	Tratam	Coag	Cura
1	1	28	7	3	II	Não	rápida
2	2	15	52	45	I	Não	lenta
3	3	76	30	23	III	Sim	normal
4	4	15	53	46	I	Sim	lenta
5	5	21	3	2	II	Não	rápida
6	6	11	46	42	I	Não	lenta
7	7	16	55	47	I	Não	lenta
8	8	16	54	47	I	Sim	lenta
9	9	47	13	12	III	Sim	normal
10	10	18	59	51	II	Não	lenta
11	11	40	20	11	III	Sim	normal
12	12	24	3	1	II	Não	rápida
13	13	32	9	3	II	Não	rápida
14	14	31	9	3	II	Não	rápida
15	15	10	44	40	I	Sim	normal
16	16	31	9	3	II	Sim	rápida
17	17	31	10	4	II	Sim	rápida
18	18	46	13	11	III	Sim	normal
19	19	21	1	2	II	Sim	rápida
20	20	39	17	8	III	Sim	rápida
21	21	15	53	46	I	Sim	lenta
22	22	9	42	39	I	Não	normal
23	23	75	30	22	III	Sim	normal
24	24	54	18	16	III	Não	normal
25	25	35	12	5	II	Sim	rápida
26	26	18	58	50	II	Sim	lenta

Figura 16

e. Verifique graficamente se, para cada uma das categorias de Cura, os pacientes apresentam diferenças com relação a tempo entre o contato com o inseto e a administração do tratamento.

**Resposta:** Em Gráficos - Boxplot, selecionamos a variável Diag, e em Gráficos por grupo escolhemos Cura,

Na Figura 18 apresenta-se o gráfico boxplot da variável Diag segundo a variável Cura. Nesta figura



Figura 17

se observam muitas diferenças entre os indivíduos:

- quanto mais tempo o paciente demora para receber o tratamento, mais lenta é a sua cura,
- o grupo que curou-se lentamente apresentou variabilidade menor,
- não há dados discrepantes em nenhum dos grupos.

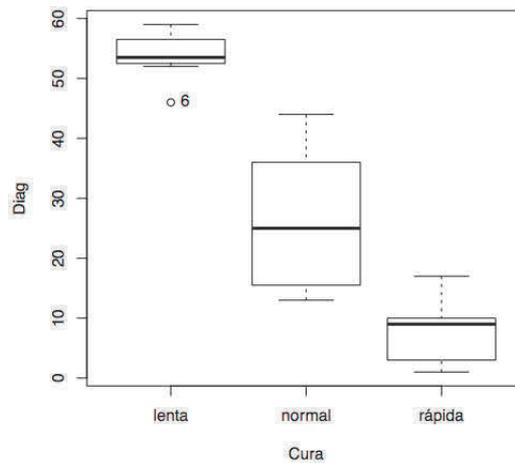


Figura 18: Boxplots da variável *Diag* segundo a variável *Cura*.

f. Crie uma tabela de dupla entrada, contendo Etário nas linhas e Coag nas colunas. Com base em tal tabela, você diria que Coag e Etário estão associados? Por quê?

**Resposta:** Primeiro temos que converter a variável Etário para fator, Em dados - Modificação de variáveis no conjunto de dados - converter variável numérica em fator - Escolhemos a variável Etário - Use números.

A tabela de dupla entrada é obtida utilizando as seguintes opções,

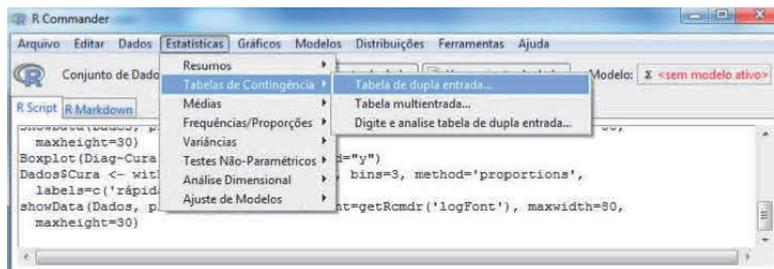


Figura 19

Escolhemos as linhas e as colunas

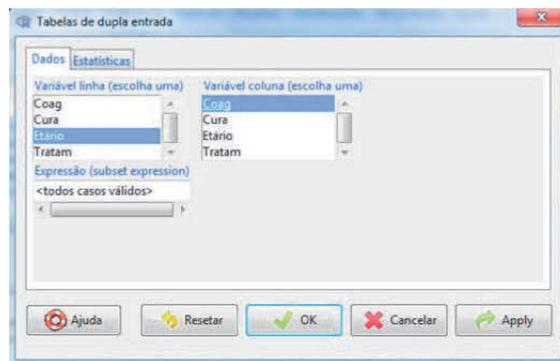


Figura 20

A saída é como segue

Frequency table:

	Coag	
Etário	Não	Sim
0	8	6
1	3	9

Row percentages:

	Coag			
Etário	Não	Sim	Total	Count
0	57.1	42.9	100	14
1	25.0	75.0	100	12

# MAE116 – Noções de Estatística

Farmácia - Noturno - 2º semestre de 2017

Lista de exercícios 2 - Estatística Descritiva II - C A S A (gabarito)

---

-----  
Frequency table:

Coag

	Etário	Não	Sim
0	8		6
1	3		9

Total percentages:

	Não	Sim	Total
0	30.8	23.1	53.8
1	11.5	34.6	46.2
Total	42.3	57.7	100.0

Segundo os resultados acima, podemos concluir que as variáveis Etário e Coag estão associadas, pois as distribuições de frequências de Coag para as diferentes categorias de Etário diferem uma da outra. Parece haver maior porcentagem de pacientes com coágulos entre as pessoas mais velhas.

# MAE116 – Noções de Estatística

Farmácia - Noturno - 2º semestre de 2017

Lista de exercícios 2 - Estatística Descritiva II - C A S A (gabarito)

---

g. Qual é a proporção de indivíduos que apresentam coágulos no momento de admissão?

**Resposta:** De acordo com os resultados da tabela de frequência acima (item (f)), podemos observar que dos indivíduos que foram contaminados pelo veneno e submetidos aos três tipos de tratamentos, 57,7% apresentou coágulos no momento da admissão.

# MAE116 – Noções de Estatística

Farmácia - Noturno - 2º semestre de 2017

Lista de exercícios 2 - Estatística Descritiva II - C A S A (gabarito)

---

- h. Crie uma tabela de dupla entrada com Tratam nas linhas e Cura nas colunas. Com base em tal tabela, você diria que a rapidez da cura depende do tipo de tratamento considerado? Justifique.

**Resposta:** Como no item (f), vamos ao menu “Estatísticas”, escolhemos “Tabelas de contingência” e depois “Tabela de dupla entrada”. Logo escolhemos as variáveis Cura nas linhas e Tratam nas colunas. A saída é como segue,

Frequency table:

```
          Tratam
Cura I IIIII
lenta 6 2 0
normal 2 0 6
rápida 0 9 1
```

Column percentages:

```
          I   II  III Total
lenta 75.0 18.2 0.0 30.8
normal 25.0 0.0 85.7 30.8
rápida 0.0 81.8 14.3 38.5
Total 100.0 100.0 100.0 100.0
```

Com base nestes resultados podemos observar que as distribuições de frequências da cura são bem diferentes nos três tratamentos. Por exemplo note que, 75% dos indivíduos que foram submetidos ao tratamento I tiveram cura lenta enquanto que 81.8% dos pacientes submetidos ao tratamento II tiveram cura rápida. O tratamento mais eficaz parece ser o II e o menos eficaz o I. O tratamento III poderia ser considerado como o intermediário.

# MAE116 – Noções de Estatística

Farmácia - Noturno - 2º semestre de 2017

Lista de exercícios 2 - Estatística Descritiva II - C A S A (gabarito)

---

- i. Dentre os pacientes submetidos ao tratamento do tipo I, qual é a proporção com cura rápida?

**Resposta:** A proporção de pacientes submetidos ao tratamento tipo I que tiveram cura rápida é nula, pois neste tipo de tratamento nenhum paciente apresentou cura rápida.

# MAE116 – Noções de Estatística

Farmácia - Noturno - 2º semestre de 2017

Lista de exercícios 2 - Estatística Descritiva II - C A S A (gabarito)

## Exercício 4.

Os números registrados de nascidos vivos nos Estados Unidos para cada mês no período de janeiro de 1991 a dezembro de 1992 são apresentados abaixo.

Mês/1991	Número(milhares)	Mês/1992	Número(milhares)
Janeiro	325	Janeiro	334
Fevereiro	312	Fevereiro	304
Março	346	Março	360
Abril	340	Abril	330
Mai	355	Mai	361
Junho	342	Junho	333
Julho	358	Julho	352
Agosto	346	Agosto	350
Setembro	365	Setembro	357
Outubro	355	Outubro	345
Novembro	324	Novembro	332
Dezembro	342	Dezembro	325

Figura 21: Registro de nascidos.

- a. Qual é a variável de estudo? Classifique-a.

**Resposta:** A variável de estudo é o número de nascidos vivos nos Estados Unidos para cada mês no período descrito. Essa é uma variável do tipo quantitativa discreta.

- b. Construa um gráfico de linha que exiba o número registrado de nascidos vivos no tempo.

**Resposta:** Na seguinte figura, a linha grossa corresponde ao período de 1991, e o resto ao período 1992.

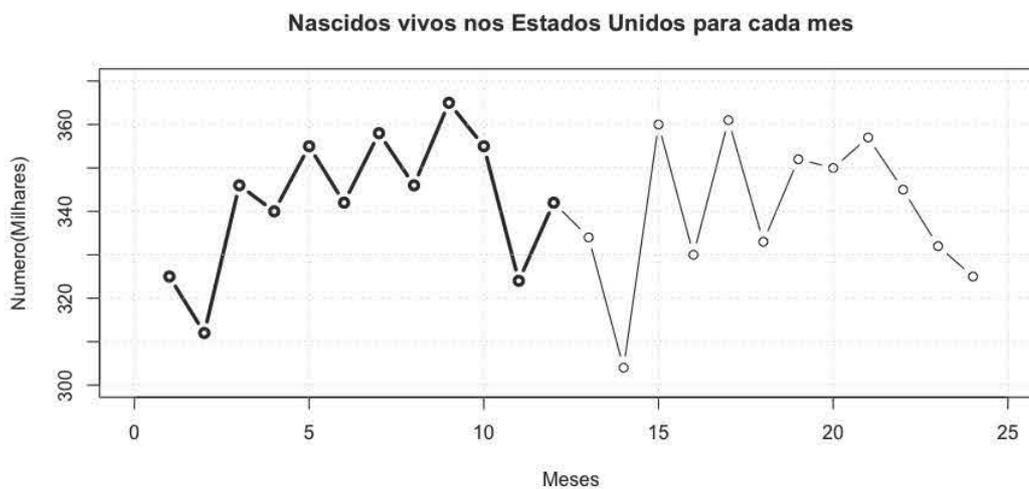


Figura 22: Registro de nascidos vivos nos anos 1991-1992.

- c. Com base nesse período de dois anos, você acha que o número de nascidos vivos segue um padrão sazonal nos Estados Unidos?

# MAE116 – Noções de Estatística

Farmácia - Noturno - 2º semestre de 2017

Lista de exercícios 2 - Estatística Descritiva II - C A S A (gabarito)

---

**Resposta:** Segundo a estrutura do gráfico de linha, podemos dizer que o número de nascidos vivos segue um padrão sazonal, porque os períodos de 1991 e 1992 apresentam no primeiro mês uma queda seguido de um aumento para logo tornar-se estável nos meses intermediários, até uma nova queda nos meses finais.

# MAE116 – Noções de Estatística

Farmácia - Noturno - 2º semestre de 2017

Lista de exercícios 2 - Estatística Descritiva II - C A S A (gabarito)

## Exercício 5.

Na tabela seguinte são listados os gastos com os cuidados com a saúde per capita em 1989 para 23 das 24 nações que constituem a Organização para Cooperação Econômica e desenvolvimento.

Nação	Gasto(US\$)	Nação	Gasto(US\$)	Nação	Gasto(US\$)	Nação	Gasto(US\$)
Austrària	1032	Finlândia	1067	Itália	1050	Portugal	464
Áustria	1093	França	1274	Japão	1035	Espanha	644
Bélgica	980	Alemanha	1232	Luxemburgo	1193	Suécia	1361
Inglaterra	836	Grécia	371	Países Baixos	1135	Suíça	1376
Canadá	1683	Islândia	1353	Nova Zelândia	820	U.S.A.	2354
Dinamarca	912	Irlanda	658	Noruega	1234		

Figura 23: Gastos de saúde.

- a. Ordene esses países de acordo com os gastos per capita com a saúde.

**Resposta:** A continuação apresentamos os países ordenados na ordem decrescente segundo os gastos com os cuidados com a saúde per capita.

```
[1,] "U.S.A"  
[2,] "Canada"  
[3,] "Suica"  
[4,] "Suecia"  
[5,] "Islandia"  
[6,] "Franca"  
[7,] "Noruega"  
[8,] "Alemanha"  
[9,] "Luxemburgo"  
[10,] "Países_Baixos"  
[11,] "Austria"  
[12,] "Finlandia"  
[13,] "Italia"  
[14,] "Japao"  
[15,] "Austraria"  
[16,] "Belgica"  
[17,] "Dinamarca"  
[18,] "Inglaterra"  
[19,] "Nova_Zelandia"  
[20,] "Irlanda"  
[21,] "Espanha"  
[22,] "Portugal"  
[23,] "Grecia"
```

- b. Calcule a média, mediana e desvio padrão dos gastos per capita.

# MAE116 – Noções de Estatística

Farmácia - Noturno - 2º semestre de 2017

Lista de exercícios 2 - Estatística Descritiva II - C A S A (gabarito)

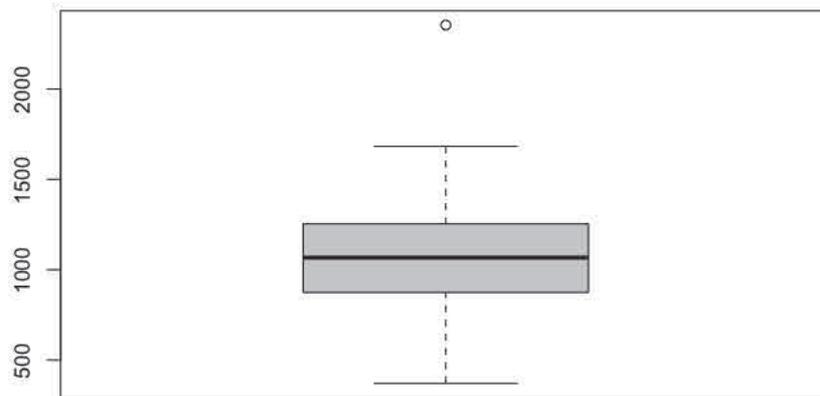
---

**Resposta:** As medidas dos gastos são as seguintes:

- Média:  $\bar{X} = 1093,78$ .
- Mediana:  $X_{Med} = 1067$ .
- Desvio padrão:  $s = 412,58$ .

c. Construa o gráfico boxplot. Comente.

**Resposta:**



**Figura 24:** *Boxplot de gastos dos países .*

Notemos que neste gráfico tem-se um ponto excepcional correspondente a os gastos de "U.S.A", concluindo que esse país é o que mais gasta na saúde per capita em comparação dos outros países.