

Gustavo Lopes, Homenique Vieira, Lucas Santiago, Rafael Amauri,
Thiago Henriques

Um estudo estatístico sobre ataques cardíacos e seu prognóstico

Belo Horizonte

2021

Resumo

"Estatística Descritiva e Estatística Inferencial" são áreas de grande importância para a Estatística, pois elas ajudam a descrever uma população através de um conjunto de dados amostrais. Este relatório técnico para a disciplina de Estatística e Probabilidade serve como um estudo de tais áreas, fazendo a descrição de dados coletados do site Kaggle sobre o assunto: ataques cardiovasculares e suas previsões.

Palavras-chave: Ataques cardiovasculares, Estatística e Probabilidade, Kaggle, Estatística Descritiva e Estatística Inferencial

Lista de ilustrações

Figura 1 – Histograma das idades	9
Figura 2 – Box Plot do colesterol	10
Figura 3 – Histograma do colesterol	10
Figura 4 – Box Plot do pressão sanguínea em repouso	11
Figura 5 – Histograma da pressão sanguínea repouso	11
Figura 6 – Histograma do pico de batimento cardíaco	12

Lista de tabelas

Tabela 1 – Medidas de tendência central e de variabilidade	13
--	----

Sumário

	Introdução	5
1	RECOLHIMENTO DOS DADOS	6
2	CLASSIFICAÇÃO DAS VARIÁVEIS	7
2.1	Qualitativas Nominais	7
2.2	Qualitativas Ordinais	7
2.3	Quantitativas Discretas	7
2.4	Quantitativas Contínuas	8
3	BOX PLOT E HISTOGRAMA	9
4	MEDIDAS DE TENDÊNCIA CENTRAL E DE VARIABILIDADE	13
5	INTERVALO DE CONFIANÇA PARA PROPORÇÃO DE IN- TERESSE	15
5.1	Proporção Populacional de Interesse	15
5.2	Média Populacional de Interesse	15
	Conclusão	17
	REFERÊNCIAS	18
	APÊNDICES	19
	APÊNDICE A – BANCO DE DADOS	20

Introdução

As doenças cardiovasculares é um conjunto de doenças do coração e dos vasos sanguíneos, incluindo problemas estruturais e coágulos. De acordo com dados distribuídos pela Organização Mundial de Saúde([OMS, 2017](#)), é estimado que no ano de 2016, 17.9 milhões de pessoas morreram por conta de doenças cardiovasculares, representando 31% de todas as mortes em nível global. Além disso, de acordo com a Sociedade Brasileira de Cardiologia, doenças cardiovasculares(DCV), tem sido a principal causa de mortalidade no Brasil desde a década de 1960.

Devido à pandemia ocasionada pelo COVID-19, admite-se que muitos desses casos vão ocorrer com mais frequência, principalmente em pessoas mais velhas devido ao estresse. Uma matéria da CNN Brasil de Janeiro deste ano, comenta dados de uma pesquisa feita no Brasil, afirma que: "o número de mortes por doenças cardiovasculares cresceu até 132% no Brasil durante a pandemia"([REZENDE, 2020](#)).

Sendo assim, por este ser um assunto relevante no contexto atual, foi selecionado um banco de dados, fornecido pelo site Kaggle, uma subsidiária da Google LLC, com foco em Cientistas de Dados e Machine Learning, afim de estudar estatisticamente as variáveis presentes na amostra.

1 Recolhimento dos dados

Como mencionado anteriormente, para este relatório decidimos usar uma base de dados do site Kaggle, mais especificamente, usamos o "Heart Attack Analysis & Prediction Dataset - A dataset for heart attack classification"([RAHMAN, 2021](#)). Estes dados contêm espaço amostral de 303 pessoas, apresentando um total de 14 atributos. Considerando que a última atualização destes dados foi em Março de 2021, este repositório se demonstra perfeito para este estudo.

Com as referências em mão e fazendo a extração dos dados do arquivo .csv, a classificação dos dados pode ser iniciada.

2 Classificação das variáveis

Antes de começar a aprofundar no estudo, inicialmente deve se fazer uma análise dos elementos presentes no conjunto. Ao selecionar uma amostragem, são analisadas informações capazes de explicar e de mostrar as características da população em questão.

Essas características são denominadas de variáveis que podem ser classificadas de diferentes formas.

2.1 Qualitativas Nominais

Variáveis de características não numérica, que nomeia ou rótulo as características por meio de números ou símbolos.

Na amostragem em questão, as variáveis a seguir são classificadas dessa forma:

- a) "sex";
- b) "fbs";
- c) "exng";
- d) "oldpeak";
- e) "cp";

2.2 Qualitativas Ordinais

Variáveis de características não numérica, que mantém uma relação de ordem.

Na amostragem em questão, as variáveis a seguir são classificadas dessa forma:

“Não apresenta variáveis nessa amostragem com essa classificação”

2.3 Quantitativas Discretas

Variáveis que assumem valores inteiros e pontuais pertencentes a um conjunto enumerável.

Na amostragem em questão, as variáveis a seguir são classificadas dessa forma:

- a) "age";
- b) "thalachh";

2.4 Quantitativas Contínuas

Variáveis que assumem valores qualquer valor real em um intervalo, associados a medição.

Na amostragem em questão, as variáveis a seguir são classificadas dessa forma:

- a) "trtbps";
- b) "chol";
- c) "restecg";

3 Box Plot e Histograma

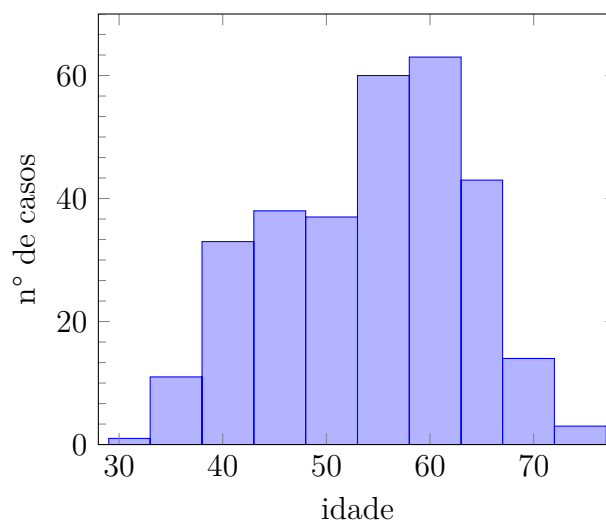
Com as categorias das variáveis colocadas, é possível inserir agora os dados coletados em gráficos, para analisar tais dados em conjunto. O primeiro gráfico utilizado é o chamado diagrama de caixa, também conhecido como Box Plot. o Box Plot é nada mais do que uma ferramenta gráfica utilizada na estatística para visualizar a variação numérica(vista pelo eixo X) dos quartis.

Enquanto isso, o quartil refere-se a qualquer um de três valores que divide o grupo ordenado de dados em quatro partes iguais. Através dele, é possível de forma visual avaliar a dispersão de um conjunto de dados, assim como também a presença de outliers(observação que diferencia um tanto das demais).

Por outro lado, o histograma serve para analisar a distribuição de frequências.

Abaixo encontra-se um histograma feito com as idades do conjunto amostral analisado:

Figura 1 – Histograma das idades

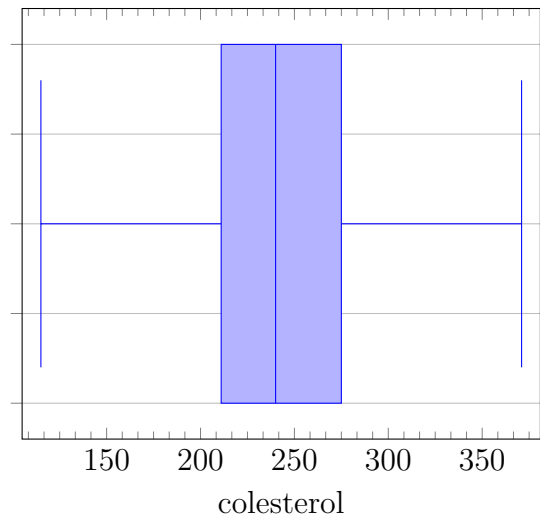


Fonte: Produzido pelos próprios autores

Com o histograma, podemos observar uma ocorrência maior de casos em pessoas com faixa etária entre 50 à 65 anos. A presença de casos para pessoas abaixo de 30 anos e acima dos 70 anos é quase nula.

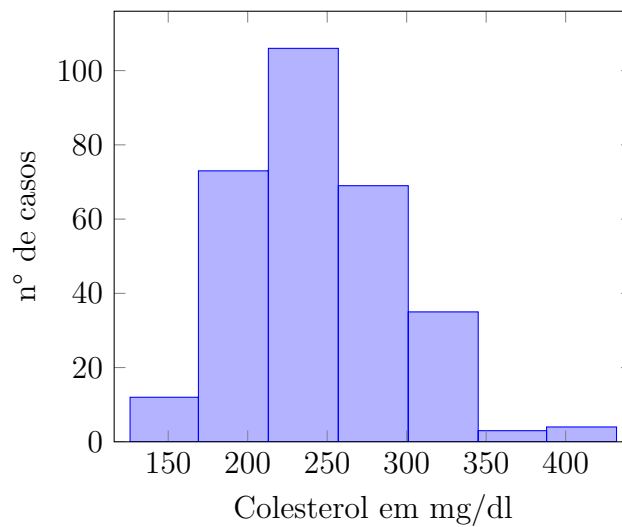
Abaixo encontra-se os gráficos com base no colesterol do conjunto amostral analisado:

Figura 2 – Box Plot do colesterol



Fonte: Produzido pelos próprios autores

Figura 3 – Histograma do colesterol

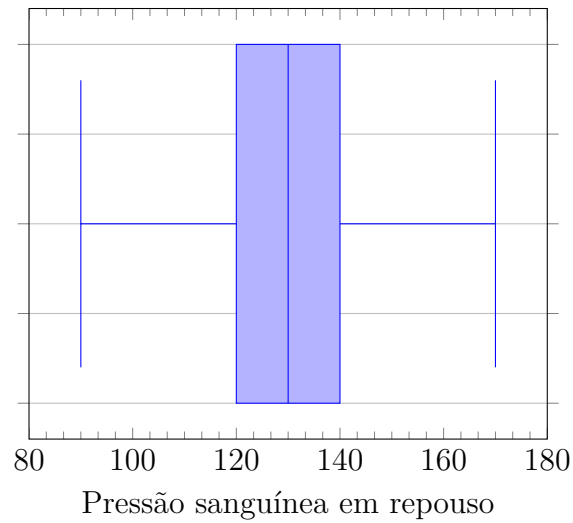


Fonte: Produzido pelos próprios autores

No boxplot, é possível observar que os outliers inferior e superior são 115 e 371, e que cinco valores da amostra ultrapassam esses outliers. Analisando o histograma, a presença da maior parte dos casos se encontra concentrada no intervalo de 169 mg/dl a 301 mg/dl de colesterol.

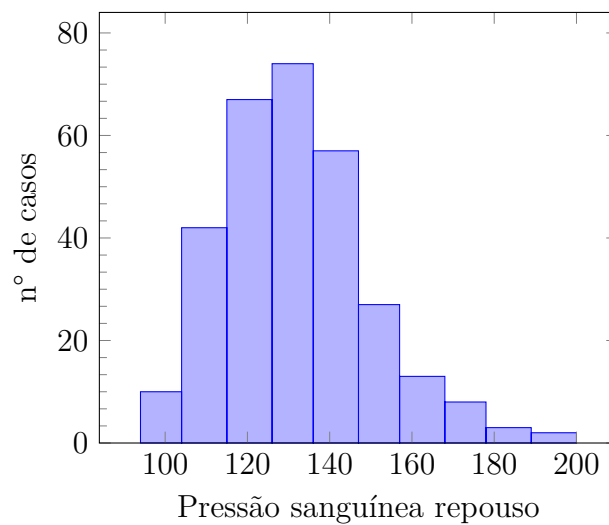
Abaixo encontra-se os gráficos feitos com as pressões sanguíneas em repouso do conjunto amostral analisado:

Figura 4 – Box Plot do pressão sanguínea em repouso



Fonte: Produzido pelos próprios autores

Figura 5 – Histograma da pressão sanguínea repouso

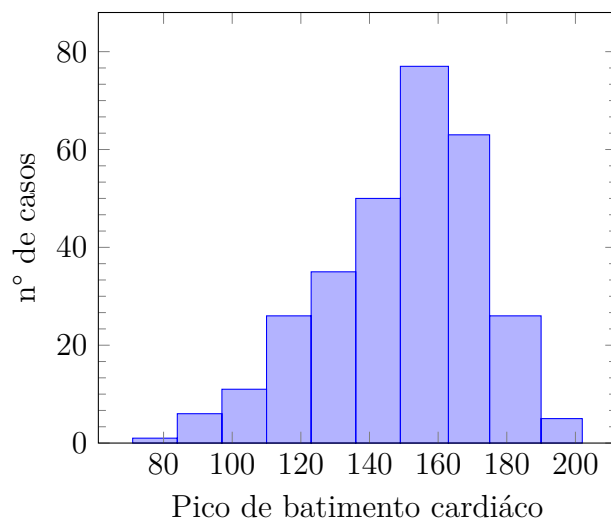


Fonte: Produzido pelos próprios autores

No boxplot, é possível observar que os outliers inferior e superior são 90 e 170, e que treze valores da amostra ultrapassam esses outliers, mais especificamente o limite superior. Analisando o histograma, a presença da maior parte dos casos concentra no intervalo de 110 mm/Hg a 140 mm/Hg equivalente a pressão sanguínea em repouso.

Abaixo encontra-se um histograma feito com os picos de batimento cardíaco do conjunto amostral analisado:

Figura 6 – Histograma do pico de batimento cardíaco



Fonte: Produzido pelos próprios autores

Analisando o histograma, é possível observar uma curva crescente do intervalo de 80 a 160 do pico de batimento cardíaco. Além disso, é possível observar um número maior de casos, onde o batimento cardiovascular teve seu pico em 150 a 160 bpm.

É importante ressaltar que todos os gráficos feitos estão relacionados com as variáveis quantitativas da amostra, já que é uma forma mais fácil de analisar as mesmas e pode nos oferecer mais informações a respeito do conjunto como um todo.

4 Medidas de tendência central e de variabilidade

Tendência central ou centralidade, refere-se a propensão de dados quantitativos acumularem em proximidade de um valor central. Em outras palavras, a partir destas medidas é possível descobrir um número que ocupa a posição central em um conjunto de valores.

Eis aqui as medidas de tendência central mais utilizadas e que também foram utilizadas nesse estudo:

- a) Média: Também chamado de média aritmética:, é a soma de todos os elementos de um conjunto;
- b) Mediana: Valor em um conjunto de dados que divide o grupo ao meio;
- c) Moda: Valor que ocorre com a maior frequência em um grupo de dados;

Abaixo pode ser encontrado uma tabela feita para parear os dados quantitativos do banco de dados, e encontrar as medidas de tendência central. Ela também apresenta o desvio padrão, importante parâmetro estatístico para encontrar o grau de variação de um conjunto de elementos.

Tabela 1 – Medidas de tendência central e de variabilidade

	Idade	Pressão Sanguínea em repouso(kPa)	Colesterol (mg/dl)	Pico de frequência cardíaca(bpm)
Média	54.36	131.62	246.26	149.64
Mediana	55	130	240	153
Moda	58	120	204, 234, 197	162
Desvio Padrão	9.08	17.53	51.83	22.90

Fonte: Produzido pelos próprios autores

Analisando os dados, o primeiro detalhe a ser observado é a idade, onde é perceptível que em média as pessoas apresentam idade quase avançada(perto dos 65 anos), além disso, o desvio padrão indica uma baixa dispersão(9%), logo indicando que existe pouca variação entre as idades (dados homogêneos).

Sobre a pressão sanguínea em repouso, os números parecem estar normais, possuindo valores que são considerados saudáveis. Enquanto isso, desta vez o desvio padrão demonstra que neste caso os dados possuem uma dispersão media.

O nível total de colesterol já demonstra um quadro um pouco mais preocupante, todas as medidas de tendência central demonstra dados que estão muito altos em comparação com os dados de referência. Desta vez, os dados se apresentam heterogêneos, devido a um número alto do desvio padrão.

Por fim, temos o pico da frequência cardíaca. Neste caso, temos uma situação interessante, pois a média e mediana se apresentam apenas um pouco acima do que a referência indica, porém o número mais frequente nestes dados (também conhecido como a moda), é de 162 bpm, o que é um número relativamente alto. Assim como o colesterol, o desvio padrão desses dados sugere uma dispersão alta.

5 Intervalo de confiança para proporção de interesse

De uma base amostral contendo 303 indivíduos, para se aprofundar mais nas características deste grupo é necessário realizar testes mais específicos. Com isso serão realizados testes para se descobrir os intervalos de confiança em determinadas situações. Existem duas formas de testes.

5.1 Proporção Populacional de Interesse

Intervalo de proporção de interesse para os homens presentes na amostragem. É importante ressaltar que o número de indivíduos masculinos é equivalente a 207.

$$\begin{aligned} IC(1 - \alpha)\% &= \hat{P} \pm Z_{\frac{\alpha}{2}} * \sqrt{\frac{\hat{P} * (1 - \hat{P})}{n}} \\ IC(95)\% &= 0,6832 \pm 1,96 * \sqrt{\frac{0,6832 * (0,3168)}{303}} \\ IC(95)\% &= 0,6832 \pm 0,0524 \\ IC(95)\% &= [0,6308; 0,7356] \end{aligned} \tag{1}$$

Com os resultados mostrados, é possível observar que a chance de um homem ter um infarto é de 63,08% a 73,56% com intervalo de confiança de 95%.

Intervalo de proporção de interesse para os indivíduos que apresentaram angina após a prática de atividades físicas. É importante ressaltar que o número de indivíduos com essa condição é equivalente a 99.

$$\begin{aligned} IC(1 - \alpha)\% &= \hat{P} \pm Z_{\frac{\alpha}{2}} * \sqrt{\frac{\hat{P} * (1 - \hat{P})}{n}} \\ IC(95)\% &= 0,3267 \pm 1,96 * \sqrt{\frac{0,3267 * (0,6733)}{303}} \\ IC(95)\% &= 0,3267 \pm 0,0269 \\ IC(95)\% &= [0,2998; 0,3536] \end{aligned} \tag{2}$$

Com os resultados mostrados, é possível observar que a chance de uma pessoa apresentar angina após a execução de atividades físicas é de 29,98% a 35,36% com intervalo de confiança de 95%.

5.2 Média Populacional de Interesse

Intervalo de média de interesse para a idade na amostragem. É importante ressaltar que a média equivale 54,36 ; desvio padrão equivale 9,08.

$$IC(1 - \alpha)\% = \bar{x} \pm Z * \frac{\alpha}{2} * \frac{\sigma}{\sqrt{N}} \quad (3)$$

$$IC(95)\% = 54,36 \pm 1,96 * \frac{9.08}{\sqrt{303}}$$

$$IC(95)\% = 54,36 \pm 1,02$$

$$IC(95)\% = [53,34; 55,38]$$

A média da amostragem pode variar de um intervalo entre 53,34 e 55,38 para confinça de 95%

Intervalo de média de interesse para a quantidade de açúcar no sangue ser menor que 120 mg/dl) na amostragem. É importante ressaltar que a média equivale 53,91 ; desvio padrão equivale 9,32 e que a quantidade da amostra é 258 pessoas.

$$IC(1 - \alpha)\% = \bar{x} \pm Z * \frac{\alpha}{2} * \frac{\sigma}{\sqrt{N}} \quad (4)$$

$$IC(95)\% = 53,91 \pm 1,96 * \frac{9.32}{\sqrt{258}}$$

$$IC(95)\% = 53,91 \pm 1,14$$

$$IC(95)\% = [52,77; 55,05]$$

A média da amostragem pode variar de um intervalo entre 52,77 e 55,05 para confinça de 95%

Conclusão

Após trabalhar nessa linguagem por quase um mês, ficou claro para toda a equipe que Haskell é uma linguagem diferenciada. Possui uma história que foi essencial para sua formação, sem esquecer, é claro, que foi uma língua desenvolvida de forma comunitária. Há grandes obstáculos em ingressar nessa língua, principalmente por documentação escassa e a documentação oficial em maioria ser paga ou extremamente complexa para iniciantes.

Por muito tempo, Haskell não foi uma língua unificada, cada pessoa que entrava no projeto criava uma versão diferente sem que um líder principal coordenasse como ela estava evoluindo. Falta de uma unidade atrasou um pouco a língua ser adotada pela comunidade. Por conta da falta de uma documentação única, a dificuldade de aprendizado foi outro grande ponto negativo que impactou diretamente na falta de profissionais que a utilização, ficando apenas fechada em um ambiente científico como universidades.

Entretanto depois de tudo que vimos, a linguagem se apresenta de forma bem mais positiva do que todas essas ideias citadas acima. Ela apresenta vários recursivos interessantes, como cálculos lambda, recursões simples e amarrações de funções em variáveis de forma simples. Há vários tutoriais distribuídos pela internet. Mesmo que poucas pessoas programem nessa língua, possui uma comunidade bem forte que a mantém.

Por fim, o grupo entendeu que Haskell desempenhou seu papel na história da computação. Além disso, várias empresas ainda o adotam por entenderem a importância de seu uso, feito para cálculos científicos e precisos. Ainda dentro do contexto de programação funcional e uso de cálculos lambda, Haskell ainda é uma, se não a melhor, língua para ser utilizado.

Referências

- DESCONHECIDO. *Histograma*. Disponível em: <<https://www.fm2s.com.br/histograma/>>. Acesso em: 28 de maio de 2021. Nenhuma citação no texto.
- DESCONHECIDO. *Qual é a frequência cardíaca ideal durante o exercício?* Disponível em: <<https://veja.abril.com.br/saude/qual-e-a-frequencia-cardiaca-ideal-durante-o-exercicio/>>. Acesso em: 28 de maio de 2021. Nenhuma citação no texto.
- FARIA, B. *Boxplot: Como interpretar?* Disponível em: <<https://operdata.com.br/blog/como-interpretar-um-boxplot/>>. Acesso em: 28 de maio de 2021. Nenhuma citação no texto.
- LIMA, A. L. *Exame de colesterol: como entender e valores de referência*. Disponível em: <<https://www.tuasaude.com/valores-de-referencia-para-o-colesterol/>>. Acesso em: 28 de maio de 2021. Nenhuma citação no texto.
- MAYER, F. de P. *Análise exploratória de dados*. Disponível em: <<https://www.inf.ufsc.br/~andre.zibetti/probabilidade/aed.html>>. Acesso em: 28 de maio de 2021. Nenhuma citação no texto.
- OMS. *Doenças Cardiovasculares*. 2017. Disponível em: <<https://www.paho.org/pt/topicos/doencas-cardiovasculares>>. Acesso em: 28 de maio de 2021. Citado na página 5.
- RAHMAN, R. *Heart Attack Analysis & Prediction Dataset - A dataset for heart attack classification*. 2021. Disponível em: <<https://www.kaggle.com/rashikrahmanpritom/heart-attack-analysis-prediction-dataset/metadata>>. Acesso em: 28 de maio de 2021. Citado na página 6.
- REZENDE, D. *Estudo apresenta dados e impactos das doenças cardiovasculares no Brasil*. 2020. Disponível em: <<https://pressreleases.scielo.org/blog/2020/11/06/estudo-apresenta-dados-e-impactos-das-doencas-cardiovasculares-no-brasil/>>. Acesso em: 28 de maio de 2021. Citado na página 5.
- SIGNOR, D. *Medidas de Tendência Central e Dispersão*. Disponível em: <<https://proeducacional.com/ead/curso-cga-modulo-i/capitulos/capitulo-4/aulas/medidas-de-tendencia-central-e-dispersao/>>. Acesso em: 28 de maio de 2021. Nenhuma citação no texto.

Apêndices

APÊNDICE A – Banco de Dados

age	sex	cp	trtbps	chol	fbs	restecg	thalachh	exng	oldpeak
63	1	3	145	233	1	0	150	0	2.3
37	1	2	130	250	0	1	187	0	3.5
41	0	1	130	204	0	0	172	0	1.4
56	1	1	120	236	0	1	178	0	0.8
57	0	0	120	354	0	1	163	1	0.6
57	1	0	140	192	0	1	148	0	0.4
56	0	1	140	294	0	0	153	0	1.3
44	1	1	120	263	0	1	173	0	0
52	1	2	172	199	1	1	162	0	0.5
57	1	2	150	168	0	1	174	0	1.6
54	1	0	140	239	0	1	160	0	1.2
48	0	2	130	275	0	1	139	0	0.2
49	1	1	130	266	0	1	171	0	0.6
64	1	3	110	211	0	0	144	1	1.8
58	0	3	150	283	1	0	162	0	1
50	0	2	120	219	0	1	158	0	1.6
58	0	2	120	340	0	1	172	0	0
66	0	3	150	226	0	1	114	0	2.6
43	1	0	150	247	0	1	171	0	1.5
69	0	3	140	239	0	1	151	0	1.8
59	1	0	135	234	0	1	161	0	0.5
44	1	2	130	233	0	1	179	1	0.4
42	1	0	140	226	0	1	178	0	0
61	1	2	150	243	1	1	137	1	1
40	1	3	140	199	0	1	178	1	1.4
71	0	1	160	302	0	1	162	0	0.4
59	1	2	150	212	1	1	157	0	1.6
51	1	2	110	175	0	1	123	0	0.6
65	0	2	140	417	1	0	157	0	0.8
53	1	2	130	197	1	0	152	0	1.2

age	sex	cp	trtbps	chol	fbs	restecg	thalachh	exng	oldpeak
41	0	1	105	198	0	1	168	0	0
65	1	0	120	177	0	1	140	0	0.4
44	1	1	130	219	0	0	188	0	0
54	1	2	125	273	0	0	152	0	0.5
51	1	3	125	213	0	0	125	1	1.4
46	0	2	142	177	0	0	160	1	1.4
54	0	2	135	304	1	1	170	0	0
54	1	2	150	232	0	0	165	0	1.6
65	0	2	155	269	0	1	148	0	0.8
65	0	2	160	360	0	0	151	0	0.8
51	0	2	140	308	0	0	142	0	1.5
48	1	1	130	245	0	0	180	0	0.2
45	1	0	104	208	0	0	148	1	3
53	0	0	130	264	0	0	143	0	0.4
39	1	2	140	321	0	0	182	0	0
52	1	1	120	325	0	1	172	0	0.2
44	1	2	140	235	0	0	180	0	0
47	1	2	138	257	0	0	156	0	0
53	0	2	128	216	0	0	115	0	0
53	0	0	138	234	0	0	160	0	0
51	0	2	130	256	0	0	149	0	0.5
66	1	0	120	302	0	0	151	0	0.4
62	1	2	130	231	0	1	146	0	1.8
44	0	2	108	141	0	1	175	0	0.6
63	0	2	135	252	0	0	172	0	0
52	1	1	134	201	0	1	158	0	0.8
48	1	0	122	222	0	0	186	0	0
45	1	0	115	260	0	0	185	0	0
34	1	3	118	182	0	0	174	0	0
57	0	0	128	303	0	0	159	0	0
71	0	2	110	265	1	0	130	0	0

age	sex	cp	trtbps	chol	fbs	restecg	thalachh	exng	oldpeak
54	1	1	108	309	0	1	156	0	0
52	1	3	118	186	0	0	190	0	0
41	1	1	135	203	0	1	132	0	0
58	1	2	140	211	1	0	165	0	0
35	0	0	138	183	0	1	182	0	1.4
51	1	2	100	222	0	1	143	1	1.2
45	0	1	130	234	0	0	175	0	0.6
44	1	1	120	220	0	1	170	0	0
62	0	0	124	209	0	1	163	0	0
54	1	2	120	258	0	0	147	0	0.4
51	1	2	94	227	0	1	154	1	0
29	1	1	130	204	0	0	202	0	0
51	1	0	140	261	0	0	186	1	0
43	0	2	122	213	0	1	165	0	0.2
55	0	1	135	250	0	0	161	0	1.4
51	1	2	125	245	1	0	166	0	2.4
59	1	1	140	221	0	1	164	1	0
52	1	1	128	205	1	1	184	0	0
58	1	2	105	240	0	0	154	1	0.6
41	1	2	112	250	0	1	179	0	0
45	1	1	128	308	0	0	170	0	0
60	0	2	102	318	0	1	160	0	0
52	1	3	152	298	1	1	178	0	1.2
42	0	0	102	265	0	0	122	0	0.6
67	0	2	115	564	0	0	160	0	1.6
68	1	2	118	277	0	1	151	0	1
46	1	1	101	197	1	1	156	0	0
54	0	2	110	214	0	1	158	0	1.6
58	0	0	100	248	0	0	122	0	1
48	1	2	124	255	1	1	175	0	0
57	1	0	132	207	0	1	168	1	0

age	sex	cp	trtbps	chol	fbs	restecg	thalachh	exng	oldpeak
52	1	2	138	223	0	1	169	0	0
54	0	1	132	288	1	0	159	1	0
45	0	1	112	160	0	1	138	0	0
53	1	0	142	226	0	0	111	1	0
62	0	0	140	394	0	0	157	0	1.2
52	1	0	108	233	1	1	147	0	0.1
43	1	2	130	315	0	1	162	0	1.9
53	1	2	130	246	1	0	173	0	0
42	1	3	148	244	0	0	178	0	0.8
59	1	3	178	270	0	0	145	0	4.2
63	0	1	140	195	0	1	179	0	0
42	1	2	120	240	1	1	194	0	0.8
50	1	2	129	196	0	1	163	0	0
68	0	2	120	211	0	0	115	0	1.5
69	1	3	160	234	1	0	131	0	0.1
45	0	0	138	236	0	0	152	1	0.2
50	0	1	120	244	0	1	162	0	1.1
50	0	0	110	254	0	0	159	0	0
64	0	0	180	325	0	1	154	1	0
57	1	2	150	126	1	1	173	0	0.2
64	0	2	140	313	0	1	133	0	0.2
43	1	0	110	211	0	1	161	0	0
55	1	1	130	262	0	1	155	0	0
37	0	2	120	215	0	1	170	0	0
41	1	2	130	214	0	0	168	0	2
56	1	3	120	193	0	0	162	0	1.9
46	0	1	105	204	0	1	172	0	0
46	0	0	138	243	0	0	152	1	0
64	0	0	130	303	0	1	122	0	2
59	1	0	138	271	0	0	182	0	0
41	0	2	112	268	0	0	172	1	0

age	sex	cp	trtbps	chol	fbs	restecg	thalachh	exng	oldpeak
54	0	2	108	267	0	0	167	0	0
39	0	2	94	199	0	1	179	0	0
34	0	1	118	210	0	1	192	0	0.7
47	1	0	112	204	0	1	143	0	0.1
67	0	2	152	277	0	1	172	0	0
52	0	2	136	196	0	0	169	0	0.1
74	0	1	120	269	0	0	121	1	0.2
54	0	2	160	201	0	1	163	0	0
49	0	1	134	271	0	1	162	0	0
42	1	1	120	295	0	1	162	0	0
41	1	1	110	235	0	1	153	0	0
41	0	1	126	306	0	1	163	0	0
49	0	0	130	269	0	1	163	0	0
60	0	2	120	178	1	1	96	0	0
62	1	1	128	208	1	0	140	0	0
57	1	0	110	201	0	1	126	1	1.5
64	1	0	128	263	0	1	105	1	0.2
51	0	2	120	295	0	0	157	0	0.6
43	1	0	115	303	0	1	181	0	1.2
42	0	2	120	209	0	1	173	0	0
67	0	0	106	223	0	1	142	0	0.3
76	0	2	140	197	0	2	116	0	1.1
70	1	1	156	245	0	0	143	0	0
44	0	2	118	242	0	1	149	0	0.3
60	0	3	150	240	0	1	171	0	0.9
44	1	2	120	226	0	1	169	0	0
42	1	2	130	180	0	1	150	0	0
66	1	0	160	228	0	0	138	0	2.3
71	0	0	112	149	0	1	125	0	1.6
64	1	3	170	227	0	0	155	0	0.6
66	0	2	146	278	0	0	152	0	0

age	sex	cp	trtbps	chol	fbs	restecg	thalachh	exng	oldpeak
39	0	2	138	220	0	1	152	0	0
58	0	0	130	197	0	1	131	0	0.6
47	1	2	130	253	0	1	179	0	0
35	1	1	122	192	0	1	174	0	0
58	1	1	125	220	0	1	144	0	0.4
56	1	1	130	221	0	0	163	0	0
56	1	1	120	240	0	1	169	0	0
55	0	1	132	342	0	1	166	0	1.2
41	1	1	120	157	0	1	182	0	0
38	1	2	138	175	0	1	173	0	0
38	1	2	138	175	0	1	173	0	0
67	1	0	160	286	0	0	108	1	1.5
67	1	0	120	229	0	0	129	1	2.6
62	0	0	140	268	0	0	160	0	3.6
63	1	0	130	254	0	0	147	0	1.4
53	1	0	140	203	1	0	155	1	3.1
56	1	2	130	256	1	0	142	1	0.6
48	1	1	110	229	0	1	168	0	1
58	1	1	120	284	0	0	160	0	1.8
58	1	2	132	224	0	0	173	0	3.2
60	1	0	130	206	0	0	132	1	2.4
40	1	0	110	167	0	0	114	1	2
60	1	0	117	230	1	1	160	1	1.4
64	1	2	140	335	0	1	158	0	0
43	1	0	120	177	0	0	120	1	2.5
57	1	0	150	276	0	0	112	1	0.6
55	1	0	132	353	0	1	132	1	1.2
65	0	0	150	225	0	0	114	0	1
61	0	0	130	330	0	0	169	0	0
58	1	2	112	230	0	0	165	0	2.5

age	sex	cp	trtbps	chol	fbs	restecg	thalachh	exng	oldpeak
39	0	2	138	220	0	1	152	0	0
58	0	0	130	197	0	1	131	0	0.6
47	1	2	130	253	0	1	179	0	0
35	1	1	122	192	0	1	174	0	0
58	1	1	125	220	0	1	144	0	0.4
56	1	1	130	221	0	0	163	0	0
56	1	1	120	240	0	1	169	0	0
55	0	1	132	342	0	1	166	0	1.2
41	1	1	120	157	0	1	182	0	0
38	1	2	138	175	0	1	173	0	0
38	1	2	138	175	0	1	173	0	0
67	1	0	160	286	0	0	108	1	1.5
67	1	0	120	229	0	0	129	1	2.6
62	0	0	140	268	0	0	160	0	3.6
63	1	0	130	254	0	0	147	0	1.4
53	1	0	140	203	1	0	155	1	3.1
56	1	2	130	256	1	0	142	1	0.6
48	1	1	110	229	0	1	168	0	1
58	1	1	120	284	0	0	160	0	1.8
58	1	2	132	224	0	0	173	0	3.2
60	1	0	130	206	0	0	132	1	2.4
40	1	0	110	167	0	0	114	1	2
60	1	0	117	230	1	1	160	1	1.4
64	1	2	140	335	0	1	158	0	0
43	1	0	120	177	0	0	120	1	2.5
57	1	0	150	276	0	0	112	1	0.6
55	1	0	132	353	0	1	132	1	1.2
65	0	0	150	225	0	0	114	0	1
61	0	0	130	330	0	0	169	0	0
58	1	2	112	230	0	0	165	0	2.5

age	sex	cp	trtbps	chol	fbs	restecg	thalachh	exng	oldpeak
39	0	2	138	220	0	1	152	0	0
58	0	0	130	197	0	1	131	0	0.6
47	1	2	130	253	0	1	179	0	0
35	1	1	122	192	0	1	174	0	0
58	1	1	125	220	0	1	144	0	0.4
56	1	1	130	221	0	0	163	0	0
56	1	1	120	240	0	1	169	0	0
55	0	1	132	342	0	1	166	0	1.2
41	1	1	120	157	0	1	182	0	0
38	1	2	138	175	0	1	173	0	0
38	1	2	138	175	0	1	173	0	0
67	1	0	160	286	0	0	108	1	1.5
67	1	0	120	229	0	0	129	1	2.6
62	0	0	140	268	0	0	160	0	3.6
63	1	0	130	254	0	0	147	0	1.4
53	1	0	140	203	1	0	155	1	3.1
56	1	2	130	256	1	0	142	1	0.6
48	1	1	110	229	0	1	168	0	1
58	1	1	120	284	0	0	160	0	1.8
58	1	2	132	224	0	0	173	0	3.2
60	1	0	130	206	0	0	132	1	2.4
40	1	0	110	167	0	0	114	1	2
60	1	0	117	230	1	1	160	1	1.4
64	1	2	140	335	0	1	158	0	0
43	1	0	120	177	0	0	120	1	2.5
57	1	0	150	276	0	0	112	1	0.6
55	1	0	132	353	0	1	132	1	1.2
65	0	0	150	225	0	0	114	0	1
61	0	0	130	330	0	0	169	0	0
58	1	2	112	230	0	0	165	0	2.5

age	sex	cp	trtbps	chol	fbs	restecg	thalachh	exng	oldpeak
50	1	0	150	243	0	0	128	0	2.6
44	1	0	112	290	0	0	153	0	0
60	1	0	130	253	0	1	144	1	1.4
54	1	0	124	266	0	0	109	1	2.2
50	1	2	140	233	0	1	163	0	0.6
41	1	0	110	172	0	0	158	0	0
51	0	0	130	305	0	1	142	1	1.2
58	1	0	128	216	0	0	131	1	2.2
54	1	0	120	188	0	1	113	0	1.4
60	1	0	145	282	0	0	142	1	2.8
60	1	2	140	185	0	0	155	0	3
59	1	0	170	326	0	0	140	1	3.4
46	1	2	150	231	0	1	147	0	3.6
67	1	0	125	254	1	1	163	0	0.2
62	1	0	120	267	0	1	99	1	1.8
65	1	0	110	248	0	0	158	0	0.6
44	1	0	110	197	0	0	177	0	0
60	1	0	125	258	0	0	141	1	2.8
58	1	0	150	270	0	0	111	1	0.8
68	1	2	180	274	1	0	150	1	1.6
62	0	0	160	164	0	0	145	0	6.2
52	1	0	128	255	0	1	161	1	0
59	1	0	110	239	0	0	142	1	1.2
60	0	0	150	258	0	0	157	0	2.6
49	1	2	120	188	0	1	139	0	2
59	1	0	140	177	0	1	162	1	0
57	1	2	128	229	0	0	150	0	0.4
61	1	0	120	260	0	1	140	1	3.6
39	1	0	118	219	0	1	140	0	1.2
61	0	0	145	307	0	0	146	1	1

age	sex	cp	trtbps	chol	fbs	restecg	thalachh	exng	oldpeak
56	1	0	125	249	1	0	144	1	1.2
43	0	0	132	341	1	0	136	1	3
62	0	2	130	263	0	1	97	0	1.2
63	1	0	130	330	1	0	132	1	1.8
65	1	0	135	254	0	0	127	0	2.8
48	1	0	130	256	1	0	150	1	0
63	0	0	150	407	0	0	154	0	4
55	1	0	140	217	0	1	111	1	5.6
65	1	3	138	282	1	0	174	0	1.4
56	0	0	200	288	1	0	133	1	4
54	1	0	110	239	0	1	126	1	2.8
70	1	0	145	174	0	1	125	1	2.6
62	1	1	120	281	0	0	103	0	1.4
35	1	0	120	198	0	1	130	1	1.6
59	1	3	170	288	0	0	159	0	0.2
64	1	2	125	309	0	1	131	1	1.8
47	1	2	108	243	0	1	152	0	0
57	1	0	165	289	1	0	124	0	1
55	1	0	160	289	0	0	145	1	0.8
64	1	0	120	246	0	0	96	1	2.2
70	1	0	130	322	0	0	109	0	2.4
51	1	0	140	299	0	1	173	1	1.6
58	1	0	125	300	0	0	171	0	0
60	1	0	140	293	0	0	170	0	1.2
77	1	0	125	304	0	0	162	1	0
35	1	0	126	282	0	0	156	1	0
70	1	2	160	269	0	1	112	1	2.9
59	0	0	174	249	0	1	143	1	0
64	1	0	145	212	0	0	132	0	2
57	1	0	152	274	0	1	88	1	1.2

age	sex	cp	trtbps	chol	fbs	restecg	thalachh	exng	oldpeak
56	1	0	132	184	0	0	105	1	2.1
48	1	0	124	274	0	0	166	0	0.5
56	0	0	134	409	0	0	150	1	1.9
66	1	1	160	246	0	1	120	1	0
54	1	1	192	283	0	0	195	0	0
69	1	2	140	254	0	0	146	0	2
51	1	0	140	298	0	1	122	1	4.2
43	1	0	132	247	1	0	143	1	0.1
62	0	0	138	294	1	1	106	0	1.9
67	1	0	100	299	0	0	125	1	0.9
59	1	3	160	273	0	0	125	0	0
45	1	0	142	309	0	0	147	1	0
58	1	0	128	259	0	0	130	1	3
50	1	0	144	200	0	0	126	1	0.9
62	0	0	150	244	0	1	154	1	1.4
38	1	3	120	231	0	1	182	1	3.8
66	0	0	178	228	1	1	165	1	1
52	1	0	112	230	0	1	160	0	0
53	1	0	123	282	0	1	95	1	2
63	0	0	108	269	0	1	169	1	1.8
54	1	0	110	206	0	0	108	1	0
66	1	0	112	212	0	0	132	1	0.1
55	0	0	180	327	0	2	117	1	3.4
49	1	2	118	149	0	0	126	0	0.8
54	1	0	122	286	0	0	116	1	3.2
56	1	0	130	283	1	0	103	1	1.6
46	1	0	120	249	0	0	144	0	0.8
61	1	3	134	234	0	1	145	0	2.6
67	1	0	120	237	0	1	71	0	1
58	1	0	100	234	0	1	156	0	0.1

age	sex	cp	trtbps	chol	fbs	restecg	thalachh	exng	oldpeak
47	1	0	110	275	0	0	118	1	1
52	1	0	125	212	0	1	168	0	1
58	1	0	146	218	0	1	105	0	2
57	1	1	124	261	0	1	141	0	0.3
58	0	1	136	319	1	0	152	0	0
61	1	0	138	166	0	0	125	1	3.6
42	1	0	136	315	0	1	125	1	1.8
52	1	0	128	204	1	1	156	1	1
59	1	2	126	218	1	1	134	0	2.2
40	1	0	152	223	0	1	181	0	0
61	1	0	140	207	0	0	138	1	1.9
46	1	0	140	311	0	1	120	1	1.8
59	1	3	134	204	0	1	162	0	0.8
57	1	1	154	232	0	0	164	0	0
57	1	0	110	335	0	1	143	1	3
55	0	0	128	205	0	2	130	1	2
61	1	0	148	203	0	1	161	0	0
58	1	0	114	318	0	2	140	0	4.4
58	0	0	170	225	1	0	146	1	2.8
67	1	2	152	212	0	0	150	0	0.8
44	1	0	120	169	0	1	144	1	2.8
63	1	0	140	187	0	0	144	1	4
63	0	0	124	197	0	1	136	1	0
59	1	0	164	176	1	0	90	0	1
57	0	0	140	241	0	1	123	1	0.2
45	1	3	110	264	0	1	132	0	1.2
68	1	0	144	193	1	1	141	0	3.4
57	1	0	130	131	0	1	115	1	1.2
57	0	1	130	236	0	0	174	0	0