

# 統計学 1 中間課題 答案（再提出）

文理学部 情報科学科

5419045 高林 秀

2022 年 12 月 15 日

## 概要

本稿は、後期総合教育科目である統計学 1 の中間課題として与えられた、健康診断データの分析に関するレポートである。

平均や、標準偏差といった各種指標数値の計算には Python のライブラリである Pandas を用いて計算を行った。また、グラフの描画には matplotlib を使用した。

はじめに、各数値データ（年齢、身長、体重、最大血圧、最小血圧）に関して、それぞれ平均や標準偏差、度数分布表、ヒストグラムを作成し、元データ全体がどのような分布になっているかについて考察した。その後、各数値（量的）データと質的データ（以降は「血圧判定」「心電図判定」を指す）の関連性を調べるため、相関比を算出し、結果から導かれることを考察した。

## 目次

1	各データの概要	2
1.1	各数値データの概要	2
1.2	質的データ（血圧判定、心電図判定）の概要	8
2	各数値データの相関比	11
2.1	体重と血圧判定の関係	11
2.2	体重と心電図判定の関係	13

## 1 各データの概要

この章では、健康診断データの全体像を把握するため、平均や標準偏差、ヒストグラム、クロス集計表を用いて、各データの分布を可視化した。

また、可視化したデータを分析し、その結果から導かれることを考察した。

### 1.1 各数値データの概要

まず、以下の表に性別を区別しない場合、男性のみのデータ、女性のみのデータでそれぞれの平均や標準偏差、最大値、最小値を算出した。

表 1 性別の区別がない場合の平均・標準偏差・最大, 最小値

	年齢	身長 (cm)	体重 (kg)	最大血圧 (mmHg)	最小血圧 (mmHg)
平均	41.3	162.5	60.3	124.9	89.4
標準偏差	12.1	10.2	9.4	9.2	15.2
最大値	67.0	185.0	75.0	144.0	129.0
最小値	23.0	141.0	43.0	105.0	65.0
総数	50				

表 2 男性のみのデータの平均・標準偏差・最大, 最小値

	年齢	身長 (cm)	体重 (kg)	最大血圧 (mmHg)	最小血圧 (mmHg)
平均	40.0	170.0	66.0	126.6	87.0
標準偏差	11.2	6.2	5.6	9.3	13.9
最大値	67.0	185.0	75.0	75.0	125.0
最小値	23.0	159.0	53.0	115.0	65.0
総数	24				

表 3 女性のみのデータの平均・標準偏差・最大, 最小値

	年齢	身長 (cm)	体重 (kg)	最大血圧 (mmHg)	最小血圧 (mmHg)
平均	41.7	155.7	55.0	123.3	91.6
標準偏差	13.0	8.2	9.2	8.9	16.3
最大値	65.0	172.0	75.0	143.0	129.0
最小値	23.0	141.0	43.0	105.0	65.0
総数	26				

続いて、上記の 3 パターンで、各数値データのヒストグラムを示す。度数分布表は画像のサイズが大きいため、省略する。

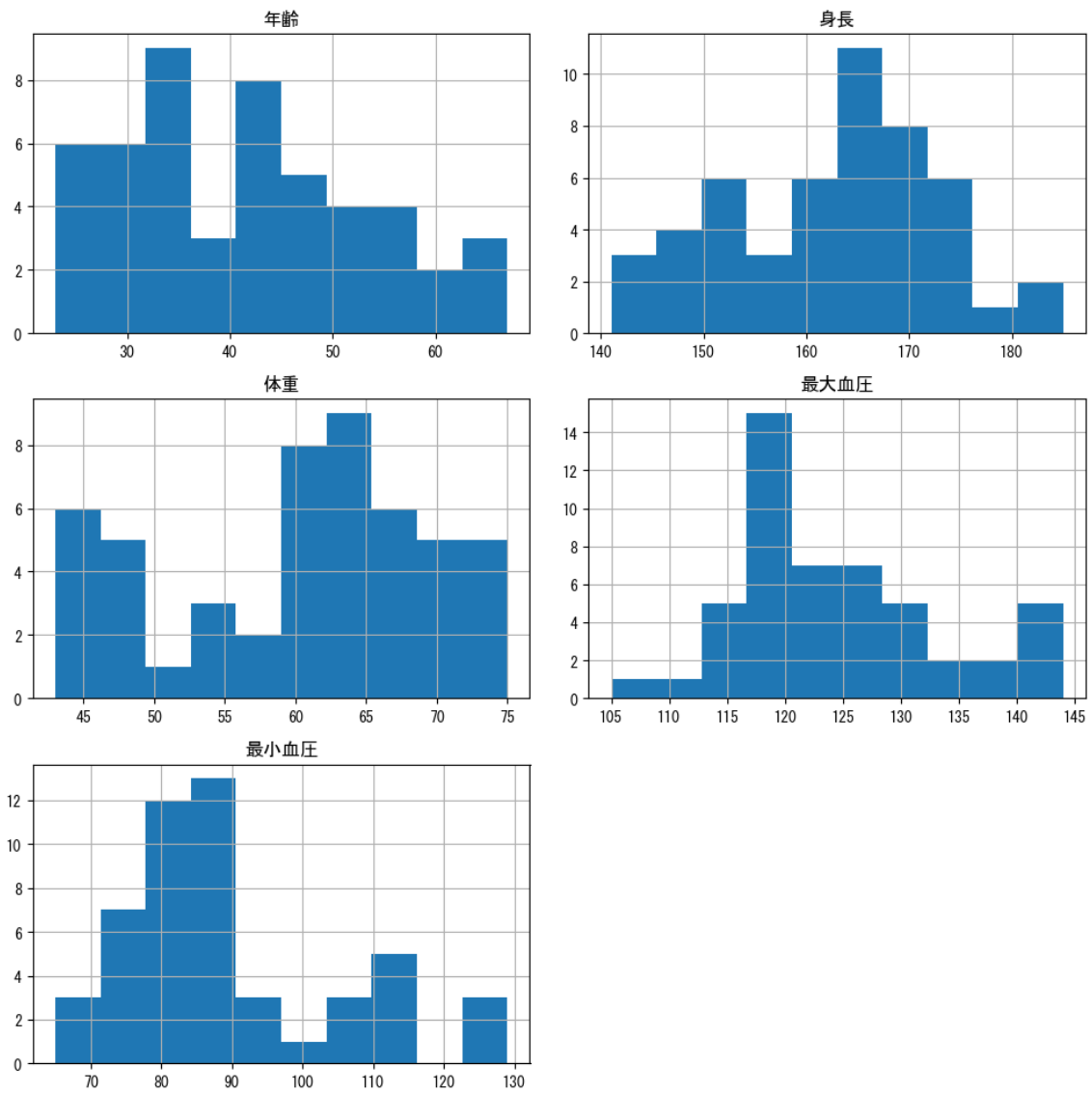


図1 全性別でのヒストグラム

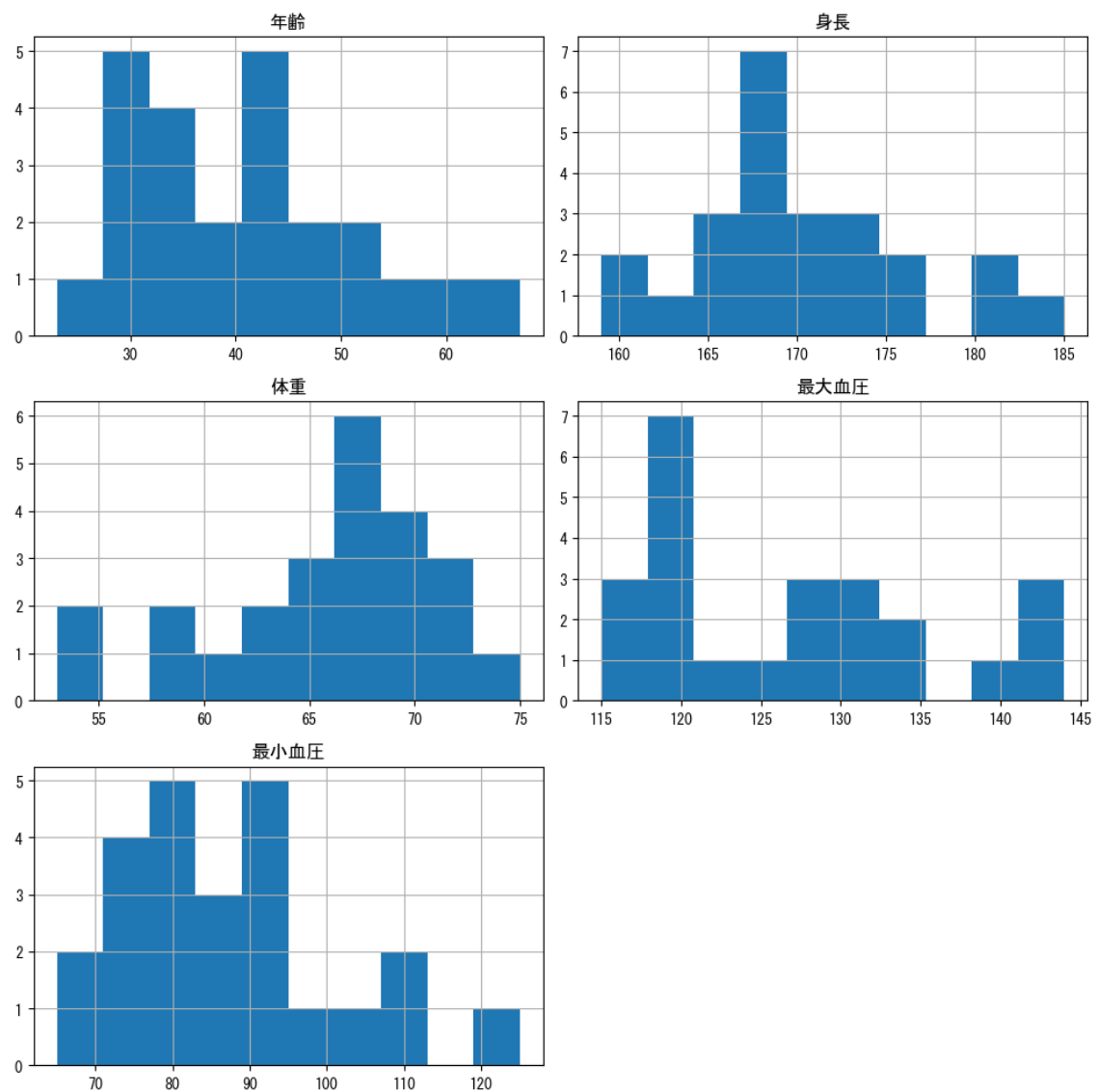


図2 男性データのヒストグラム

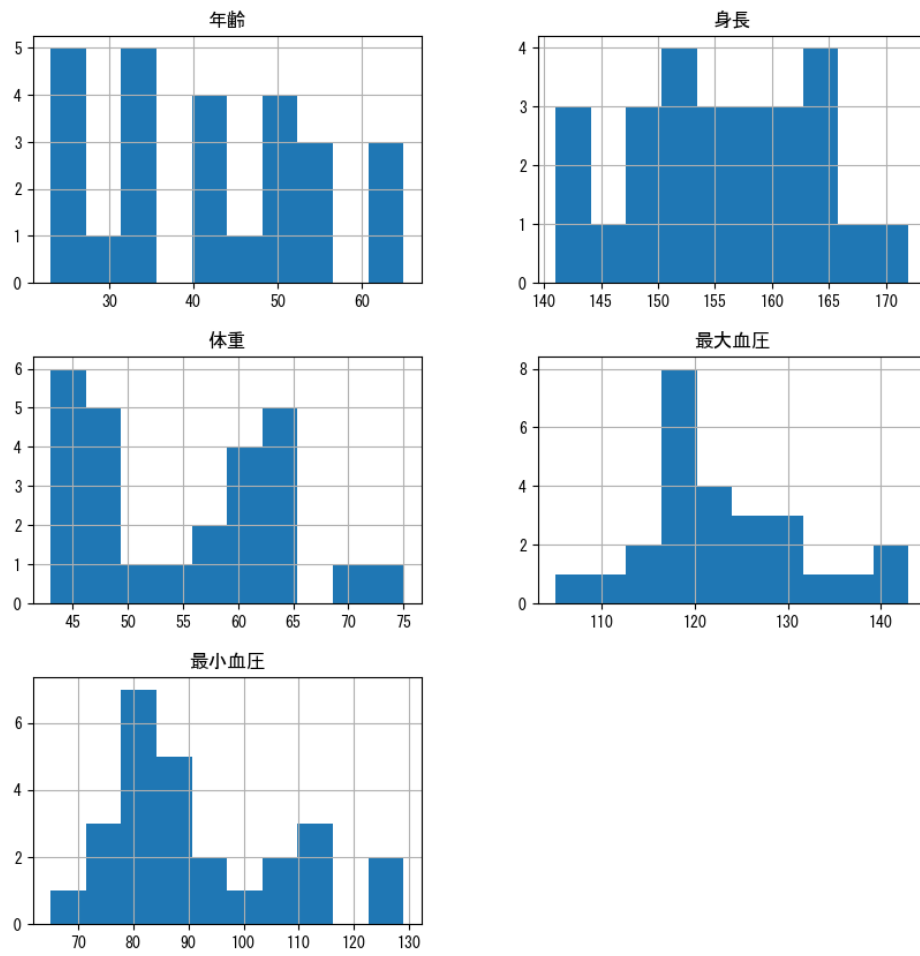


図3 女性データのヒストグラム

### 1.1.1 年齢ごとの身長・体重・最大, 最小血圧値の対応

続いて、年齢ごとの身長・体重・最大, 最小血圧値の推移を折れ線グラフで可視化した。その結果を以下に示す。

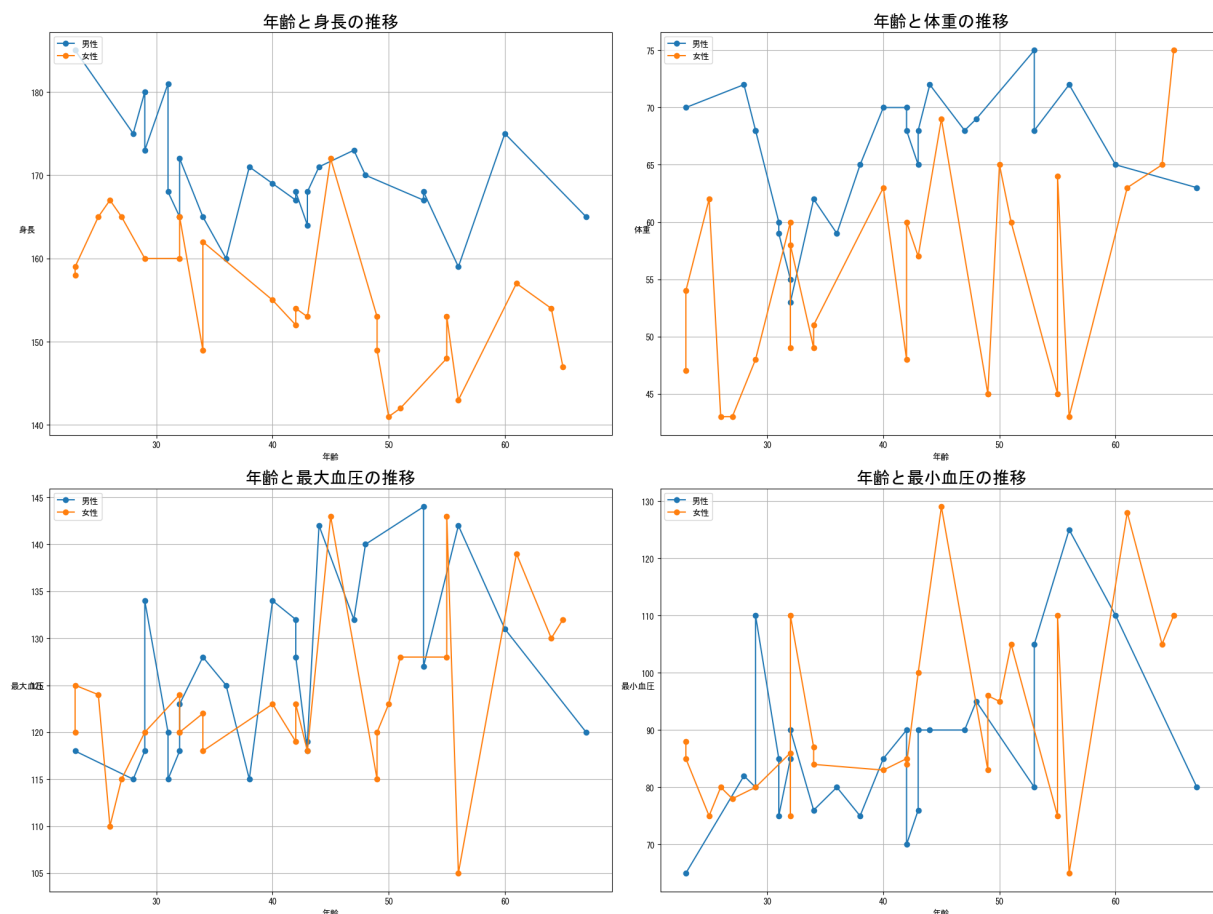


図4 年齢ごとの身長・体重・最大, 最小血圧値の対応

このグラフから、本課題で使用したデータにおいて、年齢とそれ以外の各数値データの関係性に関して以下のことが言える。以下の考察はあくまで本データに対してのみであり、一般的なことは限らない。

■**年齢と身長** 男性の場合は、年齢が上昇するにつれて全体として身長が下がっている傾向があるのが分かる。30代以下の人の身長の値は大体 180cm 170cm の間であるのに対し、50代以上になると 160cm 170cm の間の人が多く、年齢が大きくなるにつれて身長が下がっていく傾向がある。また、一般的に「高身長」といわれる身長 175cm 以上の人は、今回のデータでは 30代以下がほとんどで、それ以上になると 170cm 程度の人が多いことが分かる。

女性の場合も同様で、年齢が上昇するにつれて全体として身長が下がっている傾向があるのが分かる。30代以下の人の身長の値は大体 160cm 170cm の間であるのに対し、50代以上になると 140cm 150cm の間の人が多く、年齢が大きくなるにつれて身長が下がっていく傾向がある。

また、性別によって、同じ年代でも身長大きさにはかなり開きがあることがこのグラフから明らかになった。特に、30代以下の人と、50代以上の人のデータにおいて、男性の身長は女性の身長よりも大きい傾向があることが分かる。

**■年齢と体重** 男性の場合は、女性に比べ比較的年代ごとに体重は同じような値をとっていることがグラフから分かる。例えば、40代50代の人は大体重65kg 70kgの間にある人が多い。反対に、女性の場合は同じ年代の人同士でも、体重の値に明かな開きがあることが分かる。例えば40代50代の人、体重の値が、下は45kg程度で上は70kg弱といったように、体重の値にかなりの開きがある。

また、男性の場合は、50代後半の人から、体重の値が下がっていく傾向にあるのに対し、女性の場合は反対に増えているのが分かる。この原因として以下の文章の現象が関係していそうなので引用する。

閉経を迎えると女性ホルモンよりも男性ホルモン優位となることが関係し、中性脂肪をため込みやすい体質に変化。また、基礎代謝量の減少により60代を超えるとBMIは23台へと数値が上がる。

※引用元:<https://womanslabo.com/c-marketing-tips-190614-8>

一般的に個人差はあるが、女性の閉経時期は50代中ごろと言われている。引用文にある通り、閉経によるホルモンバランスの変化により中性脂肪を貯えやすい体質に変化し、加齢による基礎代謝量の減少により体重が上昇していると考えられる。実際にグラフから、50代後半の人の体重データは、それ以下の年代の人に比べて上昇傾向にあるのが分かる。

**■年齢と最大、最小血圧** 最大、最小血圧の値に関しては、身長、体重のグラフと異なり、若干重なっているように見え、あまり開きがない。これは、男性と女性の最大、最小血圧の値にはあまり乖離がないことを意味する。しかし、体重の部分でも述べたように、一般的に女性の閉経が始まるとされる50代後半以降のグラフをみると、極端に男性と女性の血圧値に開きがある。加えて、その年代以降の女性は、それ以前の年代の人よりも血圧値が高めになっているのが分かる。その原因として、以下の文章の現象が関係していそうなので引用する。

閉経によりエストロゲンに変換されなくなり、相対的に男性ホルモンが多くなります。すると自律神経の1つである交感神経の働きが活発化したり、血圧上昇を促すレニンというホルモンの産生が上昇したりします。つまり閉経にともなう相対的男性ホルモンの働きの増加も高血圧を引き起こす原因になります。

※引用元:[https://www.senshiniryo.net/column\\_a/29/index.html#:~:text=%E9%96%89%E7%B5%8C%E3%81%AB%E3%82%88%E3%82%8A%E3%82%A8%E3%82%B9%E3%83%88%E3%83%AD%E3%82%B2%E3%83%B3%E3%81%AB%E5%A4%89%E6%8F%9B,%E5%BC%95%E3%81%8D%E8%B5%B7%E3%81%93%E3%81%99%E5%8E%9F%E5%9B%A0%E3%81%AB%E3%81%AA%E3%82%8A%E3%81%BE%E3%81%99%E3%80%82](https://www.senshiniryo.net/column_a/29/index.html#:~:text=%E9%96%89%E7%B5%8C%E3%81%AB%E3%82%88%E3%82%8A%E3%82%A8%E3%82%B9%E3%83%88%E3%83%AD%E3%82%B2%E3%83%B3%E3%81%AB%E5%A4%89%E6%8F%9B,%E5%BC%95%E3%81%8D%E8%B5%B7%E3%81%93%E3%81%99%E5%8E%9F%E5%9B%A0%E3%81%AB%E3%81%AA%E3%82%8A%E3%81%BE%E3%81%99%E3%80%82)

以上の文にあるように、一般的に女性の場合、閉経にともなう相対的男性ホルモンの働きの増加により血圧が上昇傾向にあるようだ。このことから、グラフの「50代後半以降のグラフをみると、極端に男性と女性の血圧値に開きがある。加えて、その年代以降の女性は、それ以前の年代の人よりも血圧値が高めになっている」部分に関して納得が出来る。

## 1.2 質的データ（血圧判定、心電図判定）の概要

まずは、それぞれの判定結果がどれほどデータ内に存在するか把握するため、血圧判定、心電図判定のクロス集計表を作成した。以下はその表である。縦軸が血圧判定、横軸が心電図判定の結果で、それぞれの度数をセル内に示した。判定内容は、1,2,3,4 の順にそれぞれ、正常、要観察、要指導、要医師相談である。

表 4 全性別での血圧判定、心電図判定のクロス集計表

血圧/心電図	1	2	3	4
1	28	5	2	0
2	1	1	5	0
3	0	0	1	5
4	0	0	2	1

表 5 男性での血圧判定、心電図判定のクロス集計表

血圧/心電図	1	2	3	4
1	13	0	1	0
2	0	1	3	0
3	0	0	0	4
4	0	0	1	1

表 6 女性での血圧判定、心電図判定のクロス集計表

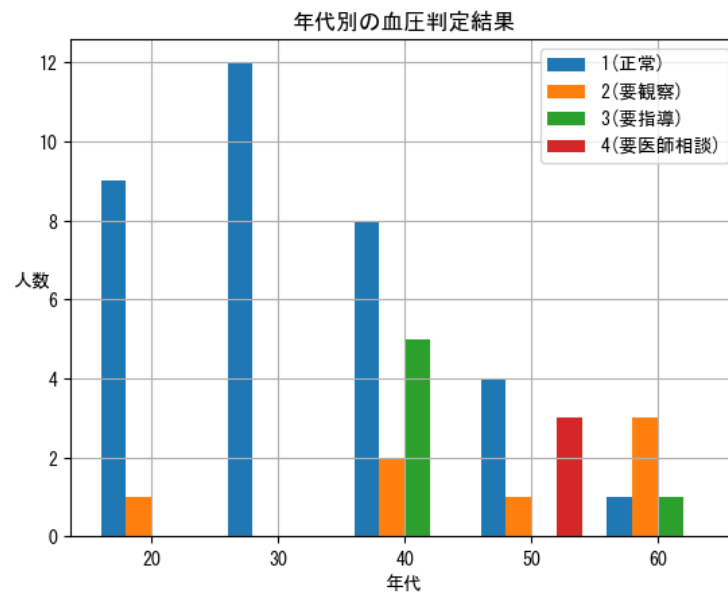
血圧/心電図	1	2	3	4
1	15	5	0	0
2	1	0	2	0
3	0	0	1	1
4	0	0	1	0

全体的には、血圧判定が 1, 心電図判定が 1 の血圧、心電図ともに正常判定の人が多数を占めているのが表から分かる。

次に全性別にわたって年代別に、血圧判定、心電図判定の結果を集計し、それぞれの判定結果が年代別でどの程度いるのかを棒グラフで集計した。



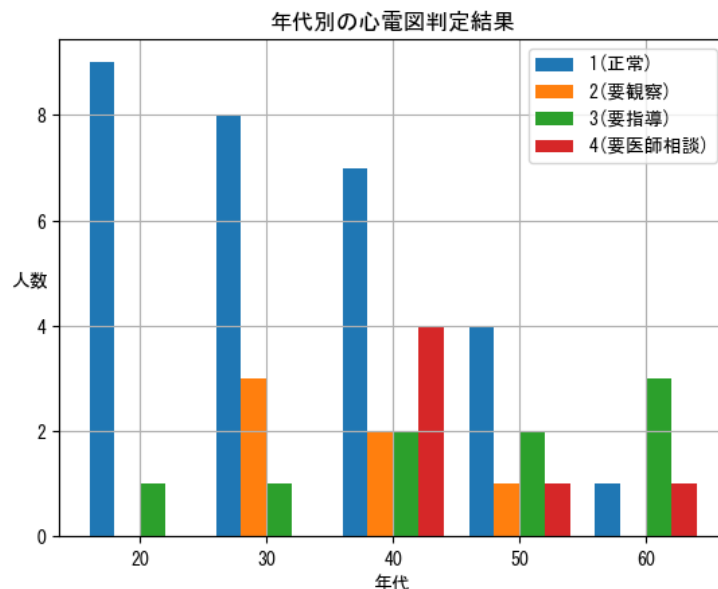
### 1.2.1 年代別の血圧判定



以下のラフから分かるように、20代から30代にかけては基本的に血圧に問題の無い人しかいない。しかし、40代以上の年代になると、要観察、要指導、要医師相談のひとの割合が増え、正常な人の割合が減少する傾向にあるのが分かる。特に、50代のデータで、要医師相談の人の割合が正常の人の割合と近いくらいにまで増えている。加えて、60代では要観察の人の割合が正常な人の割合を上回っているのが分かる。

このことから、年齢が上がるにつれて、血圧に問題があると判定される人の割合が増えていることが分かる。よって、加齢による老化が原因で、血圧に異常が発生すると考えられる。

### 1.2.2 年代別の心電図判定



心電図判定のグラフも、血圧判定のグラフと同様に正常な人の割合は、年齢を上げるにつれて減少し、反対に要観察や要指導、要医師相談の人の割合が増加していることが分かる。

ここで、40代のグラフに注目してほしい。血圧判定のグラフでは、40代の場合要医師相談の人の割合はなかったが、心電図判定のグラフでは、要医師相談の人の割合が急増しており、正常な人の割合の半分程度に昇っているのが分かる。

このことから、血圧異常の前に、心電図に医療機関の受診を勧めるレベルの異常が発生すると考えられる。加えて、40代より上の年代では、心電図が正常な人の割合が激減している。このことから、40代後半より心臓になんらかの異常をきたし、その結果血圧にも徐々に異常をきたすようになってしまおうと考えられる。これを裏付ける論文に以下の記載があるので引用する。

以後40歳ごろまではあまり年齢との関係性を認めない(とくに男子)が、45歳を過ぎるころからは加齢とともに直線的に漸次減少の一途をたどり、男女の別なく50歳前後では約半数になんらかの異常所見がみられるようである

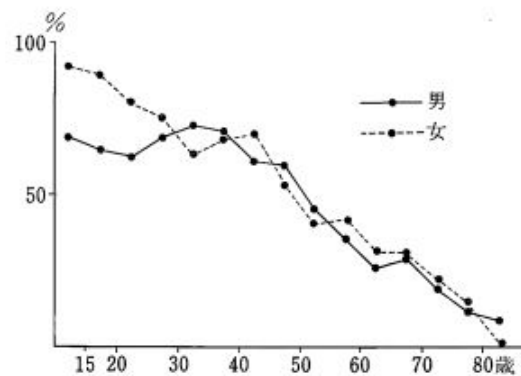


図 1 正常心電図

図 5 正常な心電図の年代別割合の推移

※引用元: 「加齢と心電図」(著) 久保 田昌良, 武信満喜夫, 島本多喜雄 [https://www.jstage.jst.go.jp/article/shinzo1969/2/3/2\\_278/\\_pdf](https://www.jstage.jst.go.jp/article/shinzo1969/2/3/2_278/_pdf)。

よって、上の論文と本データのグラフから、年齢を重ねるにつれて心電図の正常な人の割合は減少し、特に 40 歳後半から心臓になんらかの異常が発生する人の割合が増えるため、その後 50 代、60 代で血圧の異常に結びついていると考えられる。

## 2 各数値データの相関比

全性別において、各数値（量的）データと質的データの相関比を算出したのでその結果と考察を以下に示す。

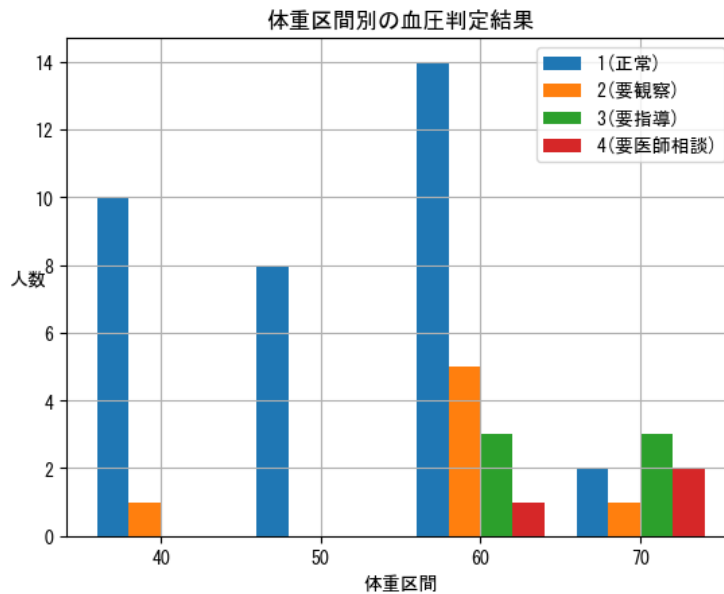
### 2.1 体重と血圧判定の関係

	血圧判定との相関比
年齢	0.54
身長	0.19
体重	0.53
最大血圧	0.85
最小血圧	0.55

表 7 全性別における各数値データと血圧判定の相関比

この結果から、血圧以外の各数値データのうち、血圧判定と関連性のある属性は「年齢」と「体重」であることが分かる。年齢と血圧判定の関係性に関しては、先に示した年齢と血圧判定の集計のグラフから読み取ることが出来る。

では、体重と血圧判定に関してはどうだろうか。相関比の数値上では年齢と同程度の関連性があると読み取れるが、実際にグラフに書き起こすと以下のような結果を得た。



一見では、年齢と同程度の関連性を見出すことはできない。しかし、体重 60kg を超えた人から、徐々にではあるが要観察、要指導、要医師相談の人の割合が増加しているように見える。加えて、70kg 代のデータにおいては、正常と判定された人の割合が激減している。

このことから、体重と血圧測定の関係に関して、体重 60kg までならば血圧に異常はなく、体重が 70kg 以上になると血圧に問題を生じる確率が高くなると考えることが出来る。これを裏付ける資料として以下の文を引用する。

内蔵脂肪型肥満では、脂肪細胞が分泌する物質や交感神経の活動亢進（活発になること）による血管の収縮、過食や塩分の取りすぎによる体液量の増加、およびインスリンの効きが悪くなるために過剰に分泌されたインスリンにより、腎臓での塩分調整障害が起こり体液量が増加するなど、血圧を上げる働きが多くなり結果として血圧が上昇します。

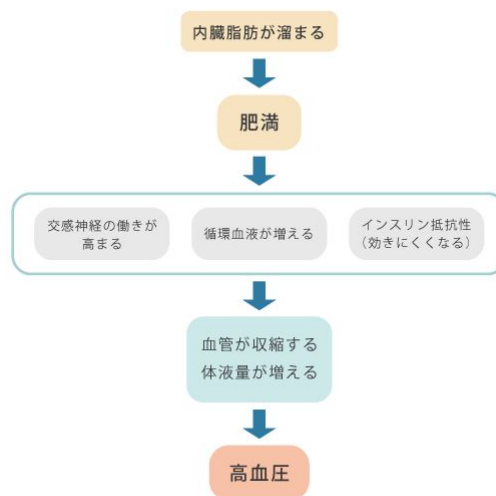


図 6 肥満から高血圧へ至るまでのフロー図

※引用元：<https://fukushima-mimamori.jp/physical-examination/column/06.html>

つまり、体重が 70kg 以上のいわゆる「肥満」に該当する人の場合、交感神経の活動による血管の収縮やインスリンの過剰分泌により、血圧に異常をきたすという事だ。

よって、グラフと引用文から、体重と血圧判定の相関比が年齢と同程度の関連性になった理由として、体重の増加による、体内環境の変化（交感神経の活動、インスリンの過剰分泌）により、高血圧となる人が増えたため、と考えることが出来る。

## 2.2 体重と心電図判定の関係

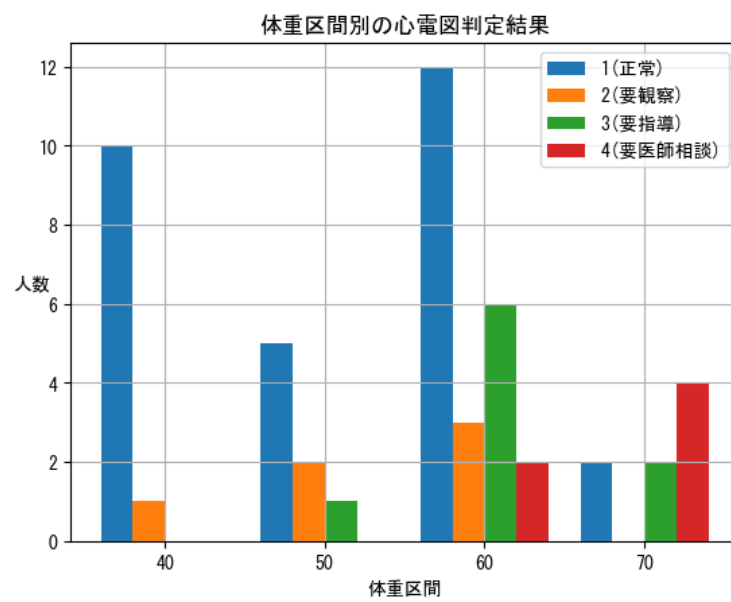
次に、各数値データと心電図判定の相関比を算出した結果を示す。

	心電図判定との相関比
年齢	0.44
身長	0.22
体重	0.56
最大血圧	0.78
最小血圧	0.66

表 8 全性別における各数値データと心電図判定の相関比

この結果から、各数値データのうち、心電図判定と関連性のある属性は「年齢」と「体重」に加え「最大・最小血圧」であることが分かる。年齢と心電図判定の関係性に関しては、先に示した年齢と心電図判定の集計のグラフから読み取ることが出来る。

では、体重、最大最小血圧値に関してはどうだろうか。実際にグラフに書き起こすと以下のような結果を得た。



■体重と心電図判定