

# 第 1 章

## 1-1 総務省統計局「社会生活基本調査」から、都道府県別のスポーツの種類別行動者数・行動者率の表を入手しましょう。

### 【解説】

以下の 2 つのアプローチが考えられます。まず①は総務省統計局 HP からのアプローチ、次に②「e-Stat 政府統計の総合窓口」からのアプローチです。

### ①総務省統計局 HP

Google などの検索サイトで「社会生活基本調査」を検索するか、総務省統計局の HP 中の「統計データ」から「社会生活基本調査」を探します。以下の調査の最新ページに到達できるはずですが、本文中の使用データは平成 28 年のものです（過年度のデータを入手したい場合には最下部に「過去の結果」のリンクがあります）。

総務省統計局「令和 3 年社会生活基本調査の結果」

<https://www.stat.go.jp/data/shakai/2021/kekka.html>

「調査の結果」中の「統計表一覧」を見ると、「調査票 A に基づく結果」のうち「生活行動に関する結果」の中に「スポーツ」があります。都道府県別を入手するため、「地域」をクリックすると、「e-Stat 政府統計の総合窓口」の該当ページにジャンプします。表番号 84-1 が「スポーツの種類別行動者数」、表番号 84-2 が「スポーツの種類別行動者

率」で【EXCEL】ボタンをクリックすると該当ファイルを XLSX 形式でダウンロードできます。

The screenshot shows the e-Stat website interface. At the top, there is a navigation bar with the e-Stat logo and the text '統計で見る日本' (View Japan with Statistics). Below this, there are search filters for 'ファイル' (File), '社会生活基本調査' (Social Life Basic Survey), '令和3年社会生活基本調査' (2021 Social Life Basic Survey), '2021年' (2021), '調査票Aに基づく結果' (Results based on Survey Form A), '生活行動に関する結果' (Results related to life activities), '生活行動編 (地域)' (Life Activities Section (Region)), and 'スポーツ' (Sports). The search results are displayed in a table with columns for '表番号' (Table Number), '統計表' (Statistical Table), '調査年月' (Survey Year), '公開 (更新) 日' (Release (Update) Date), and '表示・ダウンロード' (Display/Download). The table lists three data sets related to sports participation, each with an 'EXCEL' download button.

表番号	統計表	調査年月	公開 (更新) 日	表示・ダウンロード
84-1	男女、スポーツの種類別行動者数(10歳以上)～全国、都道府県	2021年	2022-08-31	EXCEL
84-2	男女、スポーツの種類別行動者率(10歳以上)～全国、都道府県	2021年	2022-08-31	EXCEL
84-3	男女、スポーツの種類別平均行動日数(10歳以上)～全国、都道府県	2021年	2022-08-31	EXCEL

## ②e-Stat 政府統計の総合窓口

まず、「e-Stat 政府統計の総合窓口」にアクセスします。

<https://www.e-stat.go.jp/>

「社会生活基本調査」をキーワードとして検索するか、「分野別」で探します。「分野別」の場合、「教育・文化・スポーツ・生活」の中にあります。「社会生活基本調査」のページに入ると、下図のように調査年別に「データベース」と「ファイル」へのリンクがあります。「データベース」では【DB】と【API】が提供されており、「ファイル」では【EXCEL】が提供されています。

24,391 件のデータ

データセット キーワードを入力

検索オプション

提供分類、表盟を検索  データベース、ファイル内を検索

検索のしかた

データ種別

データベース 9,857

ファイル 14,534

政府統計名で絞込み

社会生活基本調査 [24,391]

統計分野 (大分類) で絞込み

組織で絞込み

政府統計名で絞込み

提供統計名で絞込み

提供周期で絞込み

調査年で絞込み

調査月で絞込み

### 社会生活基本調査

一覧形式で表示

社会生活基本調査は、統計法に基づく基礎統計調査として、生活時間の配分や余暇時間における主な活動（学習・自己啓発・訓練、ボランティア活動、スポーツ、趣味・娯楽及び旅行・行楽）を調査し、国民の社会生活の実態を明らかにするための基礎資料を得ることを目的として5年ごとに実施しています。

調査の結果は、仕事と生活の調和（ワーク・ライフ・バランス）の推進、男女共同参画社会の形成など、国民の豊かな社会生活に関する各種行政施策に欠かすことのできない重要な資料となります。

社会生活基本調査	データベース	件数 更新日	ファイル	件数 更新日	概要
令和3年社会生活基本調査		2,901件 2022-08-31		1,607件 2022-08-31	
平成28年社会生活基本調査		1,691件 2017-12-22		1,864件 2018-02-14	
平成23年社会生活基本調査		751件 2012-12-21		4,594件 2018-02-14	
平成18年社会生活基本調査		1,931件 2019-03-08		4,998件 2019-03-08	
平成13年社会生活基本調査		1,255件 2007-10-05		1,455件 2017-01-31	
平成8年社会生活基本調査		639件 2007-10-05		16件 2017-01-31	
平成3年社会生活基本調査		458件 2007-10-05			
昭和61年社会生活基本調査		231件 2007-10-05			

【DB】を用いると、ブラウザの画面上で表示項目選択やレイアウト設定を変更することができ、CSV や XLSX 形式でダウンロードすることができます。【API】を利用すると、R や Python など統計データを取得できます。「生活行動編 (地域)」中の「スポーツ」をクリックすると、表番号 84-1-2 が「スポーツの種類別行動者数」、表番号 84-2-2 が「スポーツの種類別行動者率」となっています (84-1-1 と 84-2-1 は推定人口とサンプルサイズ)。下図は「データベース」のページです。

e-Stat 統計で見る日本  
 政府統計の総合窓口 e-Statは、日本の統計が閲覧できる政府統計ポータルサイトです

お問い合わせ | ヘルプ | English ログイン 新規登録

統計データを探す 統計データの活用 統計データの高度利用 統計関連情報 リンク集

トップページ / 統計データを探す / データベース

選択条件: データベース / 社会生活基本調査 / 令和3年社会生活基本調査 / / 調査票Aに基づく結果 / 生活行動に関する結果 /  
 生活行動編(地域) / スポーツ 政府統計一覧に戻る(すべて解除)

データセット キーワードを入力

検索オプション  
提供分類、表題を検索 データベース、ファイル内を検索 検索のしかた

データセット一覧

一覧形式で表示

政府統計名	社会生活基本調査	詳細
提供統計名	令和3年社会生活基本調査	
提供分類1	調査票Aに基づく結果	
提供分類2	生活行動に関する結果	
提供分類3	生活行動編(地域)	
提供分類4	スポーツ	

表番号	統計表	調査年月	公開(更新)日	表示・ダウンロード
84-1-1	男女,スポーツの種類別行動者数(10歳以上)-全国,都道府県	2021年	2022-08-31	DB API
84-1-2	男女,スポーツの種類別行動者数(10歳以上)-全国,都道府県	2021年	2022-08-31	DB API
84-2-1	男女,スポーツの種類別行動者率(10歳以上)-全国,都道府県	2021年	2022-08-31	DB API
84-2-2	男女,スポーツの種類別行動者率(10歳以上)-全国,都道府県	2021年	2022-08-31	DB API

【EXCEL】のある「ファイル」ページは①の方法で最終的に到達したページです。表番号 84-1 が「スポーツの種類別行動者数」、表番号 84-2 が「スポーツの種類別行動者率」で【EXCEL】ボタンをクリックすると該当ファイルを XLSX 形式でダウンロードできます。

## 1-2 上記 1-1 で入手した表を加工し、野球とサッカーの行動者数・行動者率をピックアップして 47 都道府県別の表を作成してみましよう

### 【解説】

ここでは【EXCEL】で XLSX 形式のファイルを入手したケースを考えます。野球(キャッチボールを含む)は 6 列目、ソフトボールは 7 列目、サッカー(フットサルを含む)は 10 列目にあります(表番号 84-1「スポーツの種類別行動者数」と表番号 84-2「スポーツの種類別行動者率」に共通)。各都道府県は「0\_総数」、「1\_人口集中地区」、「2\_人口集中地区以外」の 3 つに分かれています 3 列 9 行目のフィルターのパルダウンをクリックして、「総数」のみにチェックを入れましょう。全国の値が不要であれば、2 列 9 行目のパルダウンで「00\_全国」のチェックを外すこともできます。

1	2	3	4	5	6	7	8	9	10
令和3年社会生活基本調査	生活行動-地域(調査票A)								
第94-1表	男女、スポーツの種類別行動者数(10歳以上)-全国、都道府県								
			推定人口	行動者数	行動者数	行動者数	行動者数	行動者数	行動者数
				スポーツの種類	スポーツの種類	スポーツの種類	スポーツの種類	スポーツの種類	スポーツの種類
					01_野球(キャッチボールを含む)	02_ソフトボール	03_バレーボール	04_バスケットボール	05_サッカー(フットサルを含む)
			(千人)	00_総数	(千人)	(千人)	(千人)	(千人)	(千人)
男女	地域区分	人口集中地							
0_総数	00_全国	0_総数	112,482	74,785	7,051	1,667	3,906	4,051	5,339
0_総数	00_全国	1_人口集中地	81,801	56,282	5,257	1,143	2,760	2,971	4,123
0_総数	00_全国	2_人口集中地	30,681	18,504	1,793	524	1,146	1,080	1,216
0_総数	01_北海道	0_総数	4,614	2,869	336	50	188	177	204
0_総数	01_北海道	1_人口集中地	3,764	2,367	265	37	136	146	167
0_総数	01_北海道	2_人口集中地	849	502	71	12	52	31	36
0_総数	02_青森県	0_総数	1,090	567	63	6	32	43	39
0_総数	02_青森県	1_人口集中地	552	313	36	3	14	25	23
0_総数	02_青森県	2_人口集中地	538	254	27	3	18	19	15
0_総数	03_岩手県	0_総数	1,075	635	63	18	38	45	46
0_総数	03_岩手県	1_人口集中地	392	265	22	6	15	20	20
0_総数	03_岩手県	2_人口集中地	683	371	41	12	23	25	26
0_総数	04_宮城県	0_総数	2,065	1,390	145	25	84	96	102
0_総数	04_宮城県	1_人口集中地	1,390	954	100	18	62	70	80
0_総数	04_宮城県	2_人口集中地	675	376	45	8	22	26	22
0_総数	05_秋田県	0_総数	849	485	59	8	33	45	31
0_総数	05_秋田県	1_人口集中地	344	219	28	4	14	23	18
0_総数	05_秋田県	2_人口集中地	506	266	32	4	19	22	13
0_総数	06_山形県	0_総数	944	551	50	15	38	34	42
0_総数	06_山形県	1_人口集中地	477	287	21	5	16	17	21
0_総数	06_山形県	2_人口集中地	467	264	29	10	22	17	21
0_総数	07_福島県	0_総数	1,632	971	91	41	59	57	53
0_総数	07_福島県	1_人口集中地	793	508	50	23	29	31	32
0_総数	07_福島県	2_人口集中地	839	463	41	17	31	27	21
0_総数	08_茨城県	0_総数	2,572	1,692	155	37	94	94	147
0_総数	08_茨城県	1_人口集中地	1,177	831	40	16	56	50	85

サッカーと野球のみで整理した表は下のようになります。ここでは、A列に都道府県名、B・C列に行動者数、D・E列に行動者率を配置しました。下表の「サッカー」は元データ10列目の「サッカー(フットサルを含む)」であり、「野球」は元データ6列目の野球(キャッチボールを含む)と7列目のソフトボールを足したものです。サッカーを左、野球を右に配置したのは、次に図を作成する際に便利だからです。

	A	B	C	D	E
1		行動者数		行動者率	
2		サッカー	野球	サッカー	野球
3	北海道	204	386	4.4	8.4
4	青森	39	69	3.5	6.4
5	岩手	46	81	4.3	7.5
6	宮城	102	170	5	8.2
7	秋田	31	67	3.7	7.9
8	山形	42	65	4.5	6.9
9	福島	53	132	3.3	8.1
10	茨城	147	192	5.7	7.4
11	栃木	76	107	4.4	6.1
12	群馬	77	123	4.5	7.1
13	埼玉	372	549	5.6	8.2
14	千葉	261	366	4.6	6.5
15	東京	622	975	4.9	7.7
16	神奈川	489	701	5.9	8.4

1-3 上記 1-2 で作成した表を用いて、Excel で図 1-2・1-3 と同様の図を作成しましょう。

【解説】

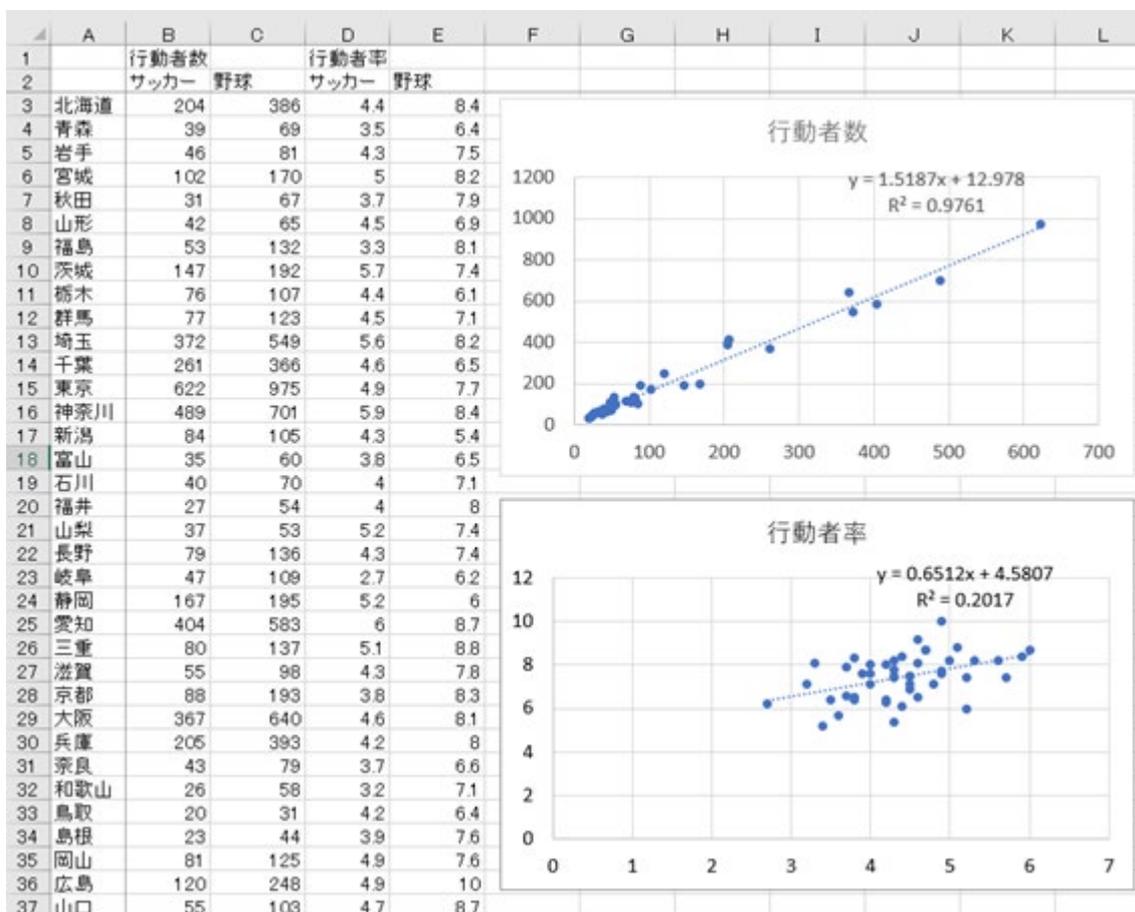
上の表から散布図を作成します。作成イメージは次の問題の図を参照してください。ただし、ここで作成した図は令和 3 年（2021 年）のデータを使用しているため、平成 28 年（2016 年）のデータを使用した本文中の図と異なります。

なお、EXCEL で散布図を作成する場合、サッカー（フットサルを含む）を左側（番号の若い列）、野球（キャッチボール、ソフトボールを含む）を右側に配置しておく、横軸と縦軸に自動的に配置されます。横軸と縦軸を逆に作成してしまった場合も、グラフの参照列を変更することで横軸と縦軸を容易に変更することが可能です。

1-4 上記 1-3 で作成した図に回帰直線（近似曲線の線形近似）を加えてみましょう。数式や R-2 乗値も表示してみてください。

【解説】

散布図上の点のいずれかを右クリックすると、「近似曲線の追加」メニューが現れます。チェックボックスにチェックを入れると、数式や R-2 乗値を表示することも可能です。作成イメージは下図を参照してください。



## 【補足】

散布図の各点に都道府県名（A 列）をラベル付けすることも可能です。「行動者率」の散布図を用いて確認してみましょう。

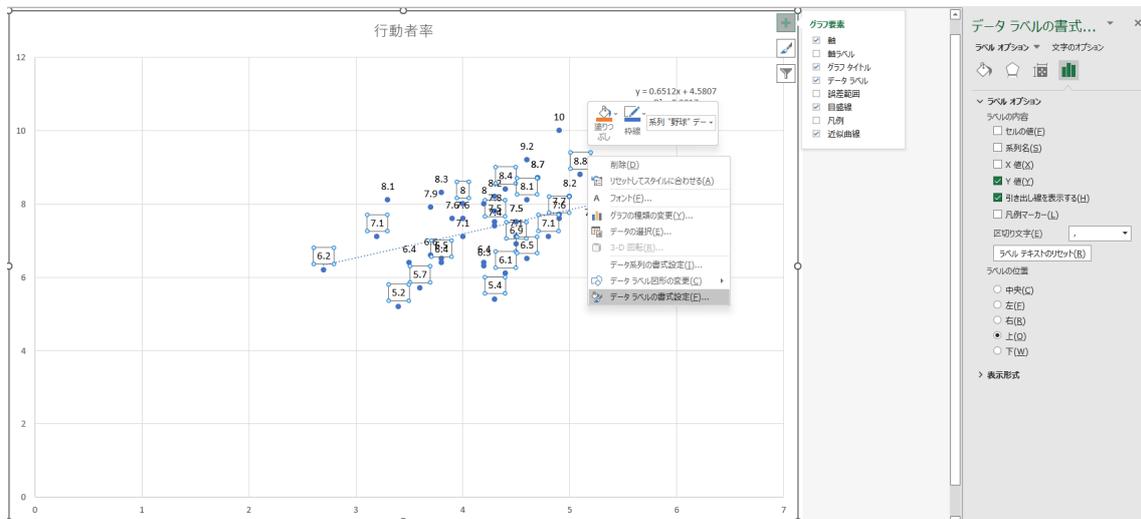
散布図の点をクリックすると、右上に「+」が現れます。この「+」をクリックするとグラフ要素のチェックボックスが現れますので、「データラベル」にチェックを入れます。各点のすぐ上に縦軸の値が表示されるはずです。

### データラベルの表示



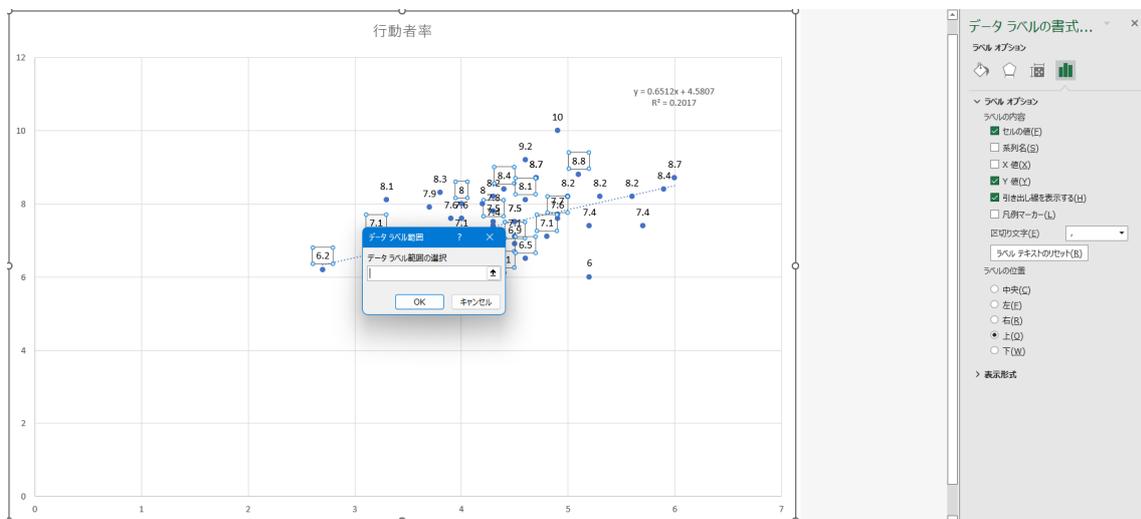
表示されたラベルを右クリックします。現れたメニューから「データラベルの書式設定」を選択すると、右端に「データラベルの書式設定」が表示されます。

## データラベルの書式設定



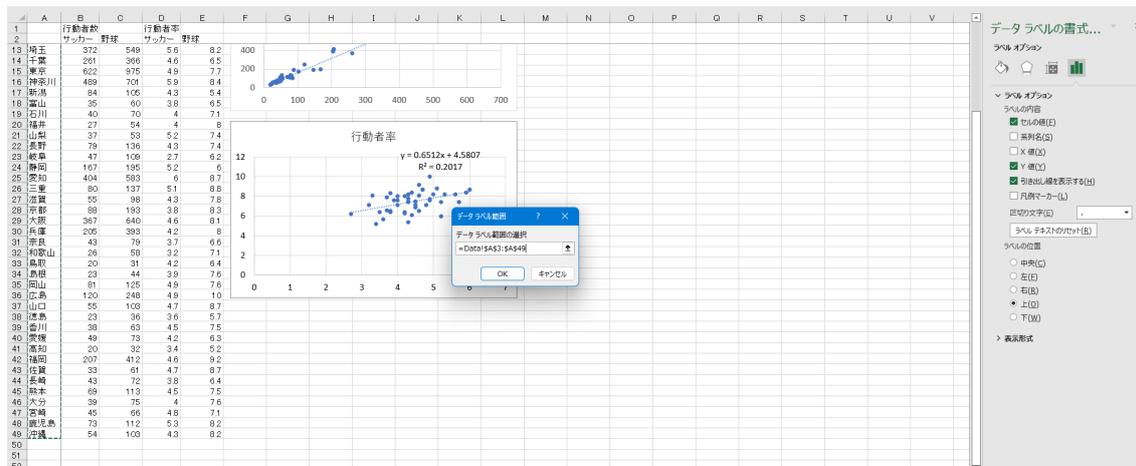
表示された「データラベルの書式設定」中の「ラベルの内容」で「セルの値(F)」をチェックすると、グラフ中央に「データラベル範囲の選択」ボックスが現れます。

## データラベル範囲の選択



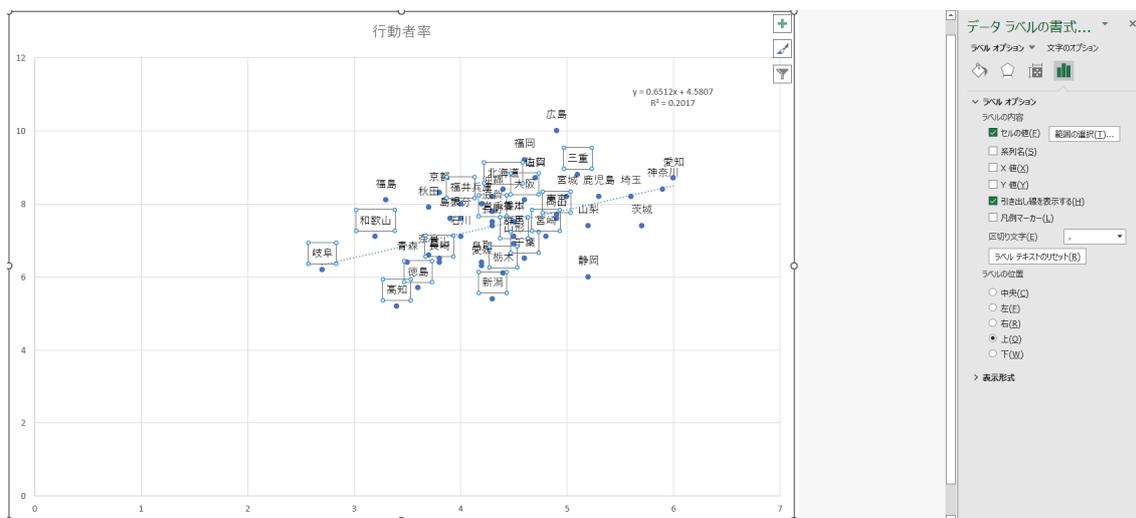
データのあるシートを選択し、都道府県名の入力されている範囲（ここでは A3～A49）を選びます。

## 都道府県名の選択



最後に、「ラベルの内容」で「Y 値(Y)」のチェックを外します。これで、都道府県名のみがラベルとして残ったはずです。

完了



1-5 上記 1-2 で作成した表を利用して、R (RStudio、R コマンドー) で図 1-2・1-3 と同様の図を作成しましょう。

### 【解説】

ここでは、RStudio を使って作図してみましよう。最初に、R で読み込みやすいように、1-2 で作成したデータを加工します。コツは、下図のように 1 行目を変数名、2 行目以降をデータとすることです。ここでは、行動者数と行動者率を区別するため、行動者数に N、行動者率に P を付しました。回帰分析などの統計分析を行うことを考えると、変数名はアルファベット表記がベターかもしれません。その場合には、「都道府県」

を「Pref」、行動者数「サッカーN」と「野球 N」をそれぞれ「soccer\_n」「baseball\_n」、行動者率「サッカーP」と「野球 P」をそれぞれ「soccer\_p」「baseball\_p」などにしてください。

## R 読み込みデータの作成

	A	B	C	D	E	F
1	都道府県	サッカーN	野球N	サッカーP	野球P	
2	北海道	204	386	4.4	8.4	
3	青森	39	69	3.5	6.4	
4	岩手	46	81	4.3	7.5	
5	宮城	102	170	5	8.2	
6	秋田	31	67	3.7	7.9	
7	山形	42	65	4.5	6.9	
8	福島	53	132	3.3	8.1	
9	茨城	147	192	5.7	7.4	
10	栃木	76	107	4.4	6.1	
11	群馬	77	123	4.5	7.1	
12	埼玉	372	549	5.6	8.2	
13	千葉	261	366	4.6	6.5	
14	東京	622	975	4.9	7.7	
15	神奈川	489	701	5.9	8.4	

次に、上記のように作成したファイルを RStudio で読み込みます。RStudio の File メニュー「Import Dataset」から「From Excel」を選び、作成したファイルを読み込んでください。データの読み込みに成功したら、グラフの作成のために ggplot2 パッケージをインストールします。Console で以下を実行するか、

```
install.packages("ggplot2")
```

または「Tools」メニューから「Install Packages」を選んで「ggplot2」と入力してインストールを実行します。ggplot2 パッケージは tidyverse パッケージ群に含まれるので、以下を実行しても構いません。

```
install.packages("tidyverse")
```

「Tools」メニュー「Install Packages」から「tidyverse」のインストールを実行しても同じことです。インストールした ggplot2 パッケージを実際に使用可能な状態にするには、以下の読み込みを行う必要があります。

```
library(ggplot2)
```

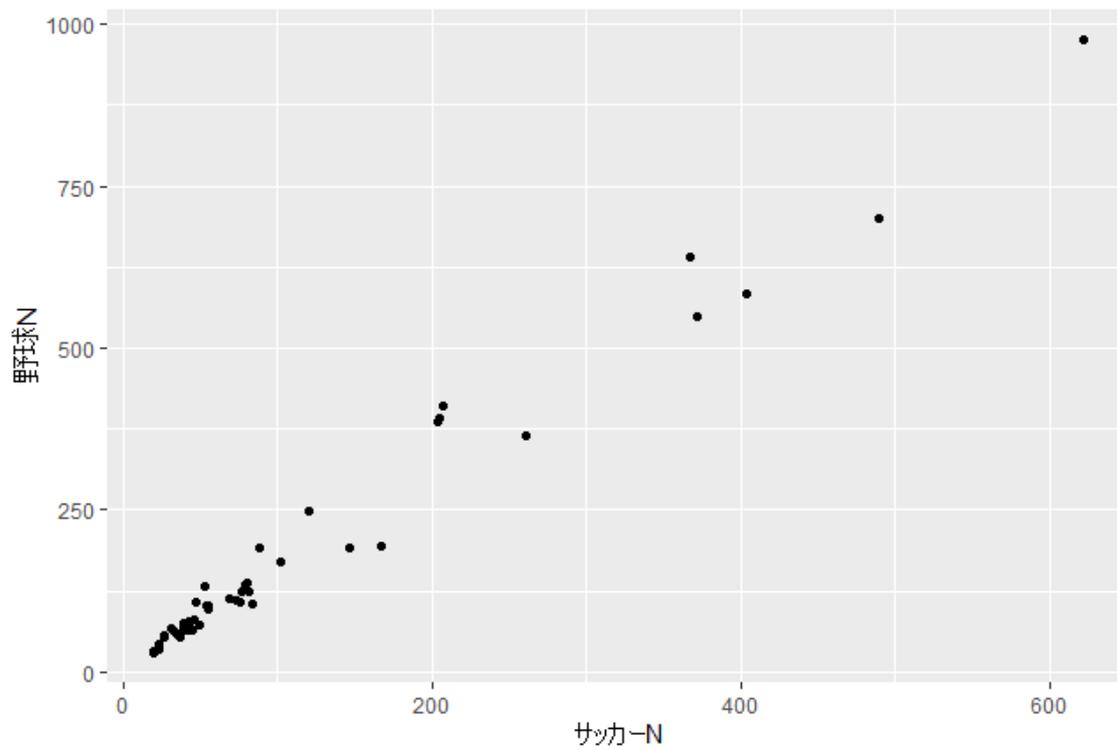
以下のように、tidyverse パッケージ群を読み込んでも構いません。

```
library(tidyverse)
```

以上で準備が整いました。グラフ作成のため以下を実行しましょう。ただし、「データ名」は皆さんのデータファイルの名前（RStudio のデータタブに表示）です。

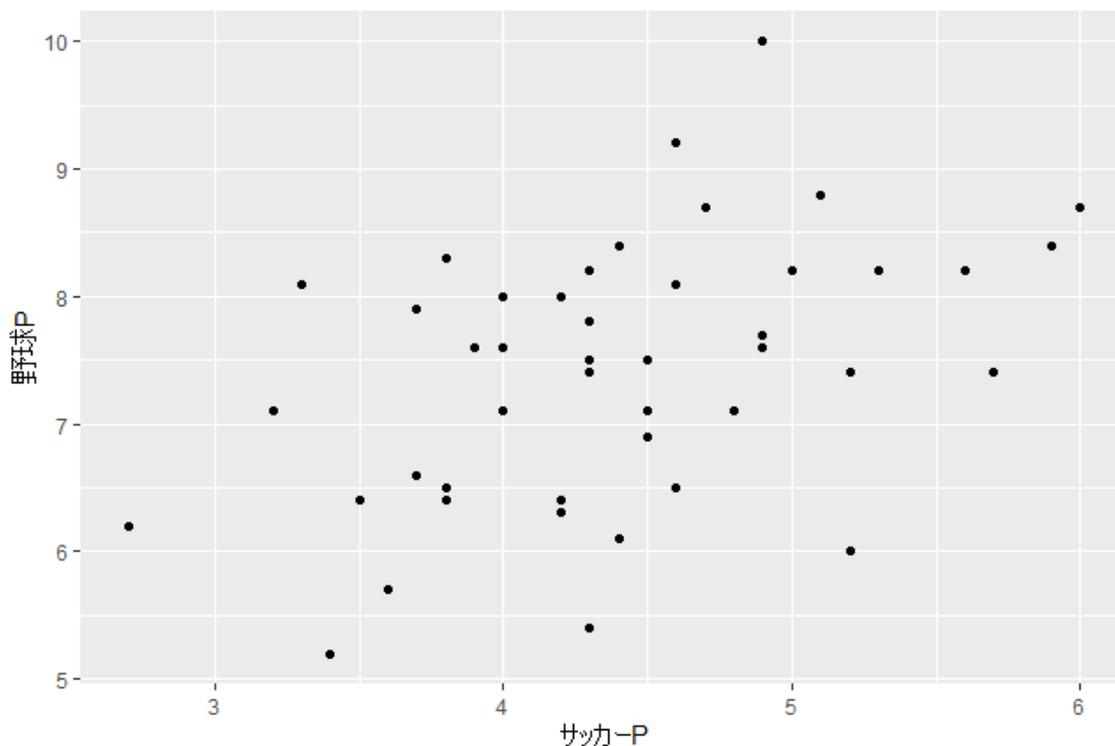
```
ggplot(データ名, aes(x=サッカーN, y=野球 N))+geom_point()
```

行動者数の散布図



```
ggplot(データ名, aes(x=サッカーP, y=野球 P))+geom_point()
```

## 行動者率の散布図



各点にラベルとして都道府県名を付したい場合は `ggrepel` パッケージを使用します。  
`ggrepel` パッケージをインストールし読み込んでください。

```
library(ggrepel)
```

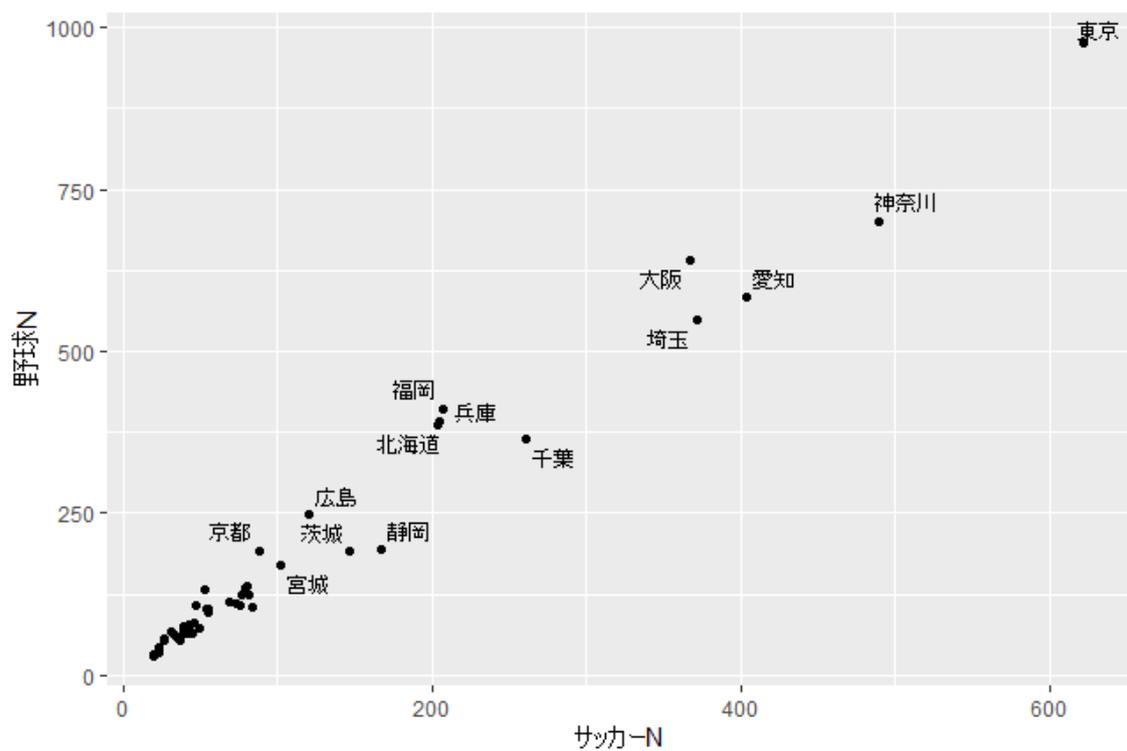
まず `ggplot` の描画を任意の名称のオブジェクトに格納します。ここではオブジェクト名を「a」とします。

```
a <- ggplot(Data1_5, aes(x=サッカーN, y=野球 N))+geom_point()
```

次に `ggrepel` パッケージの機能により、a に格納された散布図にラベルを付け加えます。ここでは、ラベルをデータ内の「都道府県」変数として設定します。

```
a + geom_text_repel(aes(label = 都道府県), size = 3)
```

行動者数の散布図（都道府県ラベル付）



行動者率についても同様の手順を踏めば、都道府県名をラベルとした散布図を描くことができます。

```
b <- ggplot(Data1_5, aes(x=サッカーP, y=野球 P))+geom_point()  
b + geom_text_repel(aes(label = 都道府県), size = 3)
```

行動者率の散布図（都道府県ラベル付）

