

データファイルの作成

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	id	x1	x2	x3	x4												
2	1	23	28	23	17												
3	2	18	22	23	14												
4	3	15	12	15	13												
5	4	12	16	22	25												
6	5	5	24	13	28												
7	6	14	16	15	16												
8	7	17	29	9	22												
9	8	17	11	25	6												
10	9	24	11	28	8												
11	10	14	14	18	10												
12	11	21	14	22	13												
13	12	21	18	13	9												
14	13	16	11	19	13												
15	14	19	26	18	21												
16	15	10	15	10	15												
17	16	12	23	16	32												
18	17	18	14	17	12												
19	18	10	25	21	29												
20	19	17	25	14	20												
21	20	20	16	16	15												
22	21	20	18	22	19												
23	22	21	10	15	23												
24	23	13	24	20	21												
25	24	17	20	18	16												
26	25	10	31	15	13												
27	26	23	14	29	6												
28	27	18	23	17	15												
29	28	9	24	21	11												
30	29	12	24	27	9												
31	30	20	8	21	8												
32	31	12	26	23	18												
33	32	10	23	18	20												
34	33	16	21	15	18												
35	34	22	32	5	21												
36	35	17	16	18	19												
37	36	15	14	25	5												
38	37	16	24	8	18												
39	38	21	4	32	6												
40	39	20	22	19	23												
41	40	22	15	15	6												
42	41	21	24	20	23												
43	42	14	17	24	19												

エクセルなどを使ってデータを入力し、CSV形式で保存しておくのが汎用性が高くてよい。

`setwd` で指定する作業ディレクトリにデータファイルを置いておくのがわかりやすくてよい。

1行目に変数名を入れておく。x1, x2のようなものでもよいし, stress, support, のように内容がわかるような変数名でもいい。

1列目にはidを入れておくと、入力ミスの検索や質問紙との照合などに役立つ。id変数の変数名を大文字で「ID」としないほうが無難である。

和文フォントも使えないわけではないが、ときどきRで問題がおきるので、英文フォントにしておいたほうが無難。(Rは世界中のボランティアが開発しているソフトなので、言語のローカル性を考慮していない追加オプションが存在する可能性がある)

欠測の場合は空白セルにしておくといよい(Rで読み込むと自動的にNAとなるが、最初からNAと入れるよりは、空白にしておいたほうがよい)。

CSV形式で保存するには、ファイルを保存するときに、ファイルの種類を「CSV (カンマ区切り) (*.csv)」にする。エクセルとしての装飾が無くなるけどいいのか、など、いろいろ言ってくる面倒だが、とにかくCSV形式で保存する。

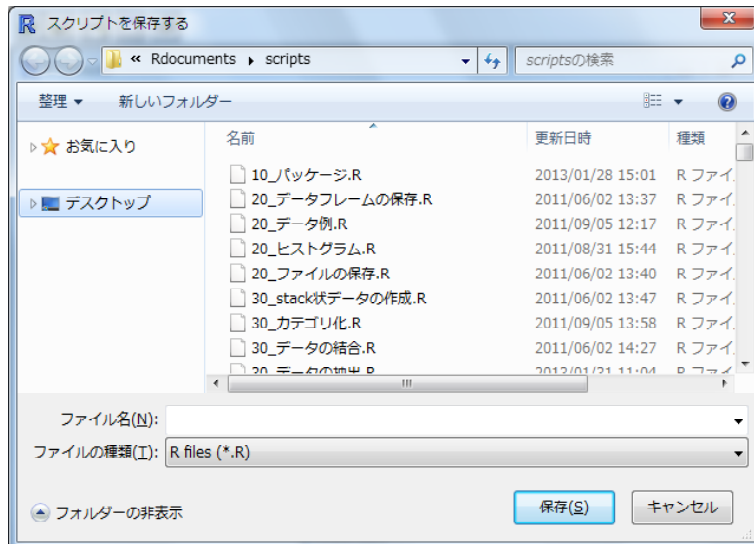
作業ディレクトリの指定

`setwd("ドライブ名¥¥フォルダ名¥¥…¥¥")`

「f:¥¥data¥¥」のように「¥」マーク（画面上では「\」になることも多い）を2つずつ入れて区切る。指定したディレクトリを「"」で囲む。

Windowsでは、¥¥の代わりに /（スラッシュ）を1つ書いて「f:/data/」とすることもできる。

- > #データファイルがあるフォルダの指定
- > `setwd("i:¥¥Rdocuments¥¥scripts¥¥")`
- >



`setwd` を実行しておくと、「ファイル」→「(別名で) 保存」などとしたとき、`setwd`で指定されたディレクトリが最初から表示される。

データフレームの保存（後述）も`setwd`で作業ディレクトリを指定してあれば、その場所に保存される。

`setwd` を実行していないと、「ファイル」→「(別名で) 保存」などとしたとき、個人のドキュメントファイルディレクトリが表示される。ファイルを保存すべきドライブ¥¥フォルダ¥¥まで、いちいち移動しなければならない。

データの読み込み

データフレーム名 <- read.table("ファイル名.csv", header=TRUE, sep=",")

header=TRUE: 1行目に変数名があるという指定

sep=",": データはカンマ区切りのテキストファイル (つまりCSV形式) という指定

ある列の値を行名に設定したいときは, row.names="変数名" または 列番号 とする.

変数名 (header) が 1セル左にずれたデータを読み込むと, 1列目のデータが自動的に行名になる.

read.csv という関数や, エクセルファイルを読み込む関数などもある.

```
> setwd("i:¥¥Rdocuments¥¥scripts¥¥")
```

```
>
```

```
> d1 <- read.table("データ例.csv", header=TRUE, sep=",")
```

```
> #最初の数行を表示
```

```
> head(d1)
```

```
id x1 x2 x3 x4
```

```
1 1 23 28 23 17
```

```
2 2 18 22 23 14
```

```
3 3 15 12 15 13
```

```
4 4 12 16 22 25
```

```
5 5 5 24 13 28
```

```
6 6 14 16 15 16
```

```
>
```

	A	B	C	D	E
1	id	x1	x2	x3	x4
2	1	23	28	23	17
3	2	18	22	23	14
4	3	15	12	15	13
5	4	12	16	22	25
6	5	5	24	13	28
7	6	14	16	15	16
8	7	17	29	8	22
9	8	17	11	25	6
10	9	24	11	28	8
11	10	14	14	18	10
12	11	21	14	22	13
13	12	21	18	13	9
14	13	16	11	19	13
15	14	19	26	18	21
16	15	10	15	10	15
17	16	12	23	16	32
18	17	19	14	17	12
19	18	10	25	21	29
20	19	17	25	14	20
21	20	20	16	16	15

```
> # 項目名を行名にした読み込み
```

```
> p.all <- read.table("irt_項目パラメタデータ.csv", row.names="item", header=TRUE, sep=",")
```

```
> p.all
```

```
aj bj cj
```

```
s03 0.7519081 -1.3479054 0
```

```
s04 0.6336668 -2.6447011 0
```

```
s05 1.2141256 -1.0870115 0
```

```
s06 1.4940753 -0.9263137 0
```

```
s07 1.7233574 -1.7290162 0
```

```
s08 0.9815908 3.1497698 0
```

```
s09 1.0866067 -0.3036178 0
```

```
s12 1.4498429 -0.1380151 0
```

```
s14 1.0159834 -0.2558882 0
```

```
s15 0.9046749 -0.4216555 0
```

```
s16 1.4752955 0.2216158 0
```

```
>
```

item 列の値が行名になっている

	A	B	C	D
1	item	aj	bj	cj
2	s03	0.752	-1.348	0
3	s04	0.634	-2.645	0
4	s05	1.214	-1.087	0
5	s06	1.494	-0.926	0
6	s07	1.723	-1.729	0
7	s08	0.982	3.150	0
8	s09	1.087	-0.304	0
9	s12	1.450	-0.138	0
10	s14	1.016	-0.256	0
11	s15	0.905	-0.422	0
12	s16	1.475	0.222	0

```
> p.all[1,1]
```

```
[1] 0.7519081
```

p.all というデータの (1,1) 要素は 0.75... であり, s03 ではない. s03 は 1 行目に付けられた行名

```
> # 変数名が1セル左にずれたデータの読み込み
> # write.table で, row.names=TRUE にした場合, 1行目に変数名, 1列目に行名が保存されるが,
> # 行名に対する変数名は無いため, 変数名が1セル左にずれて保存される.
> # このデータを読み込むと, 2行目以下の1列目のデータが, 自動的に行名になる
```

```
> # 1行目に変数名があるCSV形式ファイルの読み込み
> d3 <- read.table("データ例_rowあり.csv", header=TRUE, sep=",")
> head(d3)
```

```
  x1 x2 x3 x4
1 23 28 23 17
2 18 22 23 14
3 15 12 15 13
4 12 16 22 25
5  5 24 13 28
6 14 16 15 16
```

1列目が行番号のデータ
変数名が1セル左にずれている

	x1	x2	x3	x4	
1					
2	1	23	28	23	17
3	2	18	22	23	14
4	3	15	12	15	13
5	4	12	16	22	25
6	5	5	24	13	28
7	6	14	16	15	16
8	7	17	29	8	22
9	8	17	11	25	6
10	9	24	11	28	8
11	10	14	14	18	10
12	11	21	14	22	13
13	12	21	18	13	9

読み込んだデータの確認

head(データフレーム名)

データフレームの最初の6行を表示する.

n=10 など行数を指定するオプションを付けることもできる.

```

> setwd("i:¥¥Rdocuments¥¥scripts¥¥")
>
> d1 <- read.table("データ例.csv", header=TRUE, sep=",")
>
> head(d1)
  id x1 x2 x3 x4
1  1 23 28 23 17
2  2 18 22 23 14
3  3 15 12 15 13
4  4 12 16 22 25
5  5  5 24 13 28
6  6 14 16 15 16
>

> #最初の15行を表示する
> head(d1, n=15)
  id x1 x2 x3 x4
1  1 23 28 23 17
2  2 18 22 23 14
3  3 15 12 15 13
4  4 12 16 22 25
5  5  5 24 13 28
6  6 14 16 15 16
7  7 17 29  8 22
8  8 17 11 25  6
9  9 24 11 28  8
10 10 14 14 18 10
11 11 21 14 22 13
12 12 21 18 13  9
13 13 16 11 19 13
14 14 19 26 18 21
15 15 10 15 10 15
>

```

	A	B	C	D	E
1	id	x1	x2	x3	x4
2	1	23	28	23	17
3	2	18	22	23	14
4	3	15	12	15	13
5	4	12	16	22	25
6	5	5	24	13	28
7	6	14	16	15	16
8	7	17	29	8	22
9	8	17	11	25	6
10	9	24	11	28	8
11	10	14	14	18	10
12	11	21	14	22	13
13	12	21	18	13	9
14	13	16	11	19	13
15	14	19	26	18	21
16	15	10	15	10	15
17	16	12	23	16	32
18	17	19	14	17	12
19	18	10	25	21	29
20	19	17	25	14	20
21	20	20	16	16	15

データの保存

```
write.table(データ名, "ファイル名.csv", row.names=FALSE, sep=",")
```

row.names=FALSE にしておくと、行名は保存されない。

sep="," でデータをカンマ区切りのテキストファイル（つまりCSV形式）で出力する指定をしている。

setwd で作業ディレクトリを設定していない場合は、意に反したディレクトリに保存されることもある。

```
> setwd("i:¥¥Rdocuments¥¥scripts¥¥")
>
> d1 <- read.table("データ例.csv", header=TRUE, sep=",")
>
```

```
> head(d1)
  id x1 x2 x3 x4
1  1 23 28 23 17
2  2 18 22 23 14
3  3 15 12 15 13
4  4 12 16 22 25
5  5  5 24 13 28
6  6 14 16 15 16
```

```
> #分析に必要な変数だけを取り出す
> #d2 <- d1[,c(-1)] としても同じ。 c(-1)は「1列目は削除」を意味する
> d2 <- d1[,c("x1", "x2", "x3", "x4")]
>
```

```
> head(d2)
  x1 x2 x3 x4
1 23 28 23 17
2 18 22 23 14
3 15 12 15 13
4 12 16 22 25
5  5 24 13 28
6 14 16 15 16
>
# id変数が削除されている。
```

> # データの保存

```
> # row.names=FALSEの場合
> write.table(d2, "データ例_rowなし.csv", row.names=FALSE, sep=",")

> # row.names=TRUEの場合
> write.table(d2, "データ例_rowあり.csv", row.names=FALSE, sep=",")
```

新たに保存したファイル

行番号なし
変数名とデータの列が一致

	A	B	C	D
	x1	x2	x3	x4
	23	28	23	17
	18	22	23	14
	15	12	15	13
	12	16	22	25
	5	24	13	28
	14	16	15	16
	17	29	8	22
	17	11	25	6
	24	11	28	8
	14	14	18	10
	21	14	22	13
	21	18	13	9
	16	11	19	13
	19	26	10	21

	A	B	C	D	E
1	id	x1	x2	x3	x4
2	1	23	28	23	17
3	2	18	22	23	14
4	3	15	12	15	13
5	4	12	16	22	25
6	5	5	24	13	28
7	6	14	16	15	16
8	7	17	29	8	22
9	8	17	11	25	6
10	9	24	11	28	8
11	10	14	14	18	10
12	11	21	14	22	13
13	12	21	18	13	9
14	13	16	11	19	13
15	14	19	26	18	21
16	15	10	15	10	15
17	16	12	23	16	32
18	17	19	14	17	12
19	18	10	25	21	29
20	19	17	25	14	20
21	20	20	16	16	15

行番号あり

1 列目は行番号。変数名が1セル左にずれる。

	x1	x2	x3	x4
1				
2	1	23	28	23
3	2	18	22	23
4	3	15	12	15
5	4	12	16	22
6	5	5	24	13
7	6	14	16	15
8	7	17	29	8
9	8	17	11	25
10	9	24	11	28
11	10	14	14	18
12	11	21	14	22
13	12	21	18	13