

自動コーディング用の入力ファイルの形式について

(1) 入力ファイルの全体の概要

自動コーディングシステムは、以下の形式に則った、1行が1つの職業（あるいは産業）分のデータとなるように入力されたCSV形式のファイルが必要です。

ただし、そのファイルでは「項目名」は付けず、1行目からデータを入力して下さい。

A列：通し番号*	(数字)	半角
B列：学歴	(選択回答)	半角
C列：地位・役職	(選択回答)	半角
D列：産業（従業先事業の種類）**	(自由回答)	全角
E列：職業（仕事の内容）*	(自由回答)	全角
F列：従業先の規模	(選択回答)	半角
G列：SSM職業コード***	(数字)	半角

・*印の項目は必須。**印の項目は産業コードには必須。***印の項目は、利用者側ですでにSSM職業コードを与えたデータに対し、新たにISCO-88コードを付与したい場合に必須。

- ・*印以外の項目は必須ではありませんが、自動コードの回答精度に影響します。
- ・「.」（全角ピリオド）、「.」（半角ピリオド）、「全角空白」、「?」、「〒」、「☆」などの特殊文字は、処理上問題になるため、削除してください。
- ・回答の最後に「。」を付けないで下さい（途中に入っても問題ありません）。
悪い例：保健会社の事務。

ファイルの入力例

11	10	9	保険会社の支店	保険会社の事務	11
12	12	8	会社の売店	販売員	8
62	9	7	夫の社会保険事務所	社会保険事務所の総務、経理	3
121	10	2	コンピューターメーカー	コンピューター会社の技術系	8
289	9	10	農業	野菜を作っている	2
465	9	1	訪問介護事業	訪問介護の経営、介護福祉士	4
702	13	14	無回答	小学校の教員	13
1093	13	14	無回答	営業(外回り)	13
1527	13	14	無回答	重機のオペレーター	13

(2) 各項目の説明

①A列：通し番号

- ・一つの職業に一つの通し番号が必要となります。
- ・回答者1人に対して複数の職業についての回答がある場合、それぞれの職業について別のファイルとして作成して下さい。（例：回答者一人に現職と初職の回答がある場合、現職用ファイルと初職用ファイルの二つをご用意下さい）

②B列：学歴の選択肢

学歴についてはシステム上、日本版総合社会調査（JGSS）で用いられた以下のコードを用いて

処理しています。そのため、他の選択肢で行った調査データの場合は、以下のコードに合わせてリコードしてください。

1	旧制尋常小学校（国民学校を含む）	8	新制中学校
2	旧制高等小学校	9	新制高校
3	旧制中学校・高等女学校	10	新制短大・高専
4	旧制実業学校	11	新制大学
5	旧制師範学校	12	新制大学院
6	旧制高校・旧制専門学校・高等師範学校	13	わからない
7	旧制大学・旧制大学院		

③C列：従業上の地位・役職の選択肢

従業上の地位や役職についてはシステム上、以下のコードを用いて処理しています。他の選択肢を用いて行った調査の場合は、以下のコードに合わせてリコードしてください。

1	経営者・役員	8	臨時雇用・パート・アルバイト
2	常時雇用の一般従業者 役職なし	9	派遣社員
3	〃 職長、班長、組長	10	自営業主・自由業者
4	〃 係長、係長相当職	11	家族従業者
5	〃 課長、課長相当職	12	内職
6	〃 部長、部長相当職		
7	〃 役職はわからない	14	わからない

④D列：産業（従業先事業の種類）、E列：職業（仕事の内容）

回答の自由記述の内容を「全角（のみ）」で入力して下さい。

またシステム処理の関係上、以下の諸点にご注意下さい。

- ・半角文字を含めないこと
- ・英字は大文字のこと。
- ・空白（含「全角空白」）やピリオド（.）、特殊な記号（?、〒、→、@、☆など）を含めないこと。

以上の問題があるファイルは適切に処理できないため、特にご注意ください。

⑤F列：従業先の規模（企業規模）の選択肢

従業先の規模（企業規模）についてはシステム上、日本版総合社会調査（JGSS）で用いられた以下のコードを用いて処理しています。そのため、他の選択肢で行った調査データの場合は、以下のコードに合わせてリコードしてください（※）。

1	1人	8	500～999人
2	2～4人	9	1,000～1,999人
3	5～9人	10	2,000～9,999人
4	10～29人	11	1万人以上
5	30～99人	12	官公庁
6	100～299人	13	わからない
7	300～499人		

※) 実施された調査の選択肢が、JGSS よりも粗い企業規模カテゴリとなっているデータの場合 (たとえば「10～99 人」という選択肢がある場合など)、上記のコードをそのまま当てはまることはできません。

その場合の対処としては、

①自動コーディングシステムが職業コードを付与する際に、「30 人以上の規模の企業か、以下の企業か」で管理職の処理が異なり、「1 人、5 人未満、30 人未満、100 人未満、官公庁」という区分が、仮コード修正に一定程度用いられるので、この情報を考慮の上、どのようにコードを与えるべきかを判断する

②機械的に中央値で置き換える (10 人～99 人の場合は、 $(10+99) \div 2 \approx$ 「55 人」→ 「30～99」に含める)

等の方法が考えられます。どのような方法が適切かは、研究目的や内容によるかと思しますので、利用者さまの方で適宜ご判断ください。

⑥G 列：SSM 職業コード

利用者側で SSM 職業コードをすでに付与しているデータに対して、ISCO コードを新たにコーディングしたい場合に記入ください。不明や無回答の場合は「999」と記入し、無回答のないようにしてください。