

エクセルを活用したアンケート共同集計の実践

土橋 喜

要旨

エクセルをある程度学習した履修者を対象に、授業の中でアンケートを実施し、集計作業を履修者自身が共同で行う例を報告する。授業では履修者がお互いに協力して集計作業を行うことにより、実際の社会調査において活用できるアンケート調査の集計方法を身に付けることを目標にしている。授業で実践した結果から今後の授業展開に役立つ知見が得られた。

キーワード： アンケート、共同集計、エクセル、データ分析、社会調査

1. はじめに

現代社会ではアンケート調査を行う例は多く、愛知大学現代中国学部でも現地研究実習などの授業でアンケート調査を行う機会を設けている。そのためエクセルなどのソフトを使って履修者がある程度自分でアンケートの集計をできるようになっておく必要がある。アンケートの質問項目などは、参加する履修者が毎年自分たちで考えて作成している。アンケートは学内で大学生に対して行う場合、名古屋の中心部で一般人を対象に行う場合、郵送などで企業にお願いして行う場合などがある。回収数の多い場合は400部を超える場合もある。

回収したアンケートの集計にはエクセルを活用している。ほとんどの履修者はエクセルなど表計算ソフトを使ったアンケート集計の経験がなく、加えて統計学の基礎的な学習もほとんど行っていない場合が多い。そのためエクセルの操作方法だけでなく、統計学の基礎的な知識とエクセルの活用法を学ぶ必要がある。

そこでアンケート集計を授業で取り上げるまえに、エクセルの基本操作、表計算の方法、グラフの作成、移動平均、分散、正規分布、偏差値、条件判断、度数分布、散布図、共分散、相関係数、回帰直線、クロス表、乱数とシミュレーション、データベース機能、フィルタ、などを取り上げている。

これらは「社会データ分析入門」として一つの授業科目になっており、上に取り上げた項目を90分授業で11回ほど行っている。これらの学習を行った後で授業中に模擬アンケートを行い、その集計に3回前後の授業回数を確保している。

2. 調査票の準備と回収

アンケートは一定の質問形式で意見などを調べる社会調査の手法の一つである。そのため調査票を作成して何を調べたいかを明確にして行う必要がある。ここでのアンケート調査は、授業を履修している集団を対象に行うもので、特定集団の特徴を調べる調査になる。いくつかの観

点から質問項目を作成し、履修者全体の特徴を把握することを目的としている。アンケートの基本的な属性に相当するような項目は教員が用意したものもあるが、その他の具体的な質問項目は履修者から募集し、それを取捨選択して編集を行い作成している。調査票の質問項目の一部を示すが、調査票の選択肢は集計練習のため、Q1のように数字で回答するものと、Q2のようにアルファベットで回答するものを含んでおり、最後に複数回答の選択肢の例も用意している。これらは履修

者が実際のアンケート調査のときに、数字で回答を求める場合と、アルファベットで回答を求める場合があり、それに対応するためである。

調査票は授業中に配布して、全員の回答が終わったところを見計らって回収する。また調査票の回収方法や回答内容から個人が特定できないように配慮する。回答は個人の目に触れないように裏返して回収し、回収した調査票には通し番号を付して整理する。

表 1. 調査項目の一部

| | |
|---|---|
| 「社会データ分析入門 履修者アンケート」(授業で実施した一部) | |
| 授業の教材として使いますので、ご自身の状況について、以下の項目にお答えください。 該当する記号または番号を○で囲ってください。 | |
| (以下の質問で「30分～1時間以内」のように範囲があるものは「30分を越えて1時間以内」の意味とし、エクセルの度数分布表のしくみと合わせます。つまり、...より大きく～以下の意味とします。) | |
| Q1. 性別 | 1. 男 2. 女 |
| Q2. 現在住んでいる県 | a. 愛知県 b. 岐阜県 c. 三重県 d. 静岡県 e. その他 |
| Q3. 身長は何 cm ですか (少数第一位まで)。 | |
| Q4. 足の大きさは何 cm ですか (少数第一位まで)。 | |
| Q5. 朝食は週何回食べていますか。 | 1. 毎日 2. 1～3日 3. 4～6日 4. 食べない |
| Q6. 通学時間はどのぐらいですか。 | 1. 30分以内 2. 30分～1時間以内 3. 1時間～1.5時間以内 4. 1.5時間～2時間以内 5. 2時間～ |
| Q7. インターネットの検索エンジンでよく使うサイトはどこですか。(複数回答可) | 1. Google 2. goo 3. Yahoo 4. MSN 5. Biglobe |
| (以下略) | |

3. 調査票の整理と入力

回答の入力は履修者で分担して行う。いったん回収して整理番号を付した調査票は、ランダムに履修者に戻して回答内容の確認を行いながら、エクセルのワークシートにデータの入力を行う(図1)。

入力データの形式は、ワークシート上部の列に質問項目の番号を入力し、左端の行に回答用紙の整理番号を入力する。さらに実際の回答記号を該当する質問項目ごとに、1つのセルに1つずつ記号で入力する。この段階の入力データは、選択肢の回答記号を入力し、身長などのデータは回答の数値をそのまま入力する。入力したデータは必要に応じて、検索と置換機能を使って、集計段階で置き換え

を行う。

エクセルのファイルは回答用紙ごとに作成し、整理番号をファイル名に使う。例えば整理番号が001番の場合は、エクセルのファイル名は001.xlsのように作成し、回答用紙を必要に応じて参照できるようにしておく。

回答のデータを入力したファイルは、教員側で用意したディレクトリに貼り付けて提出する(図2)。そのディレクトリは履修者全員が書き込みと閲覧およびコピーが可能な状態に設定されている。

データの入力が終了したあとで、入力データの確認を行う必要がある。入力したままのデータでは間違いがある可能性が高い。調査票を履修者から回収し、再

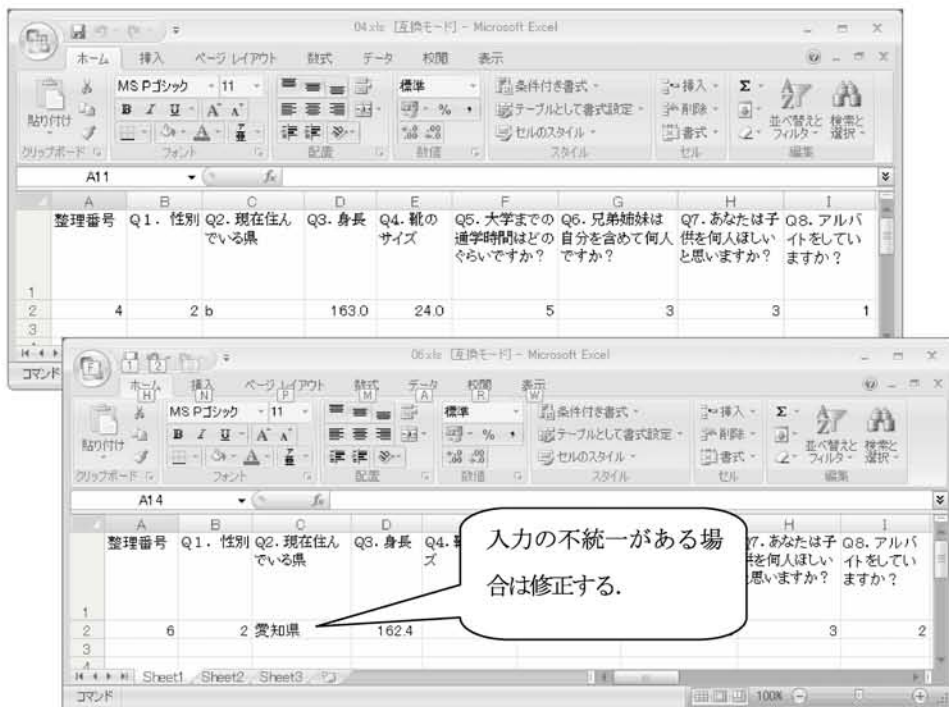


図1 分担して入力したファイルの例

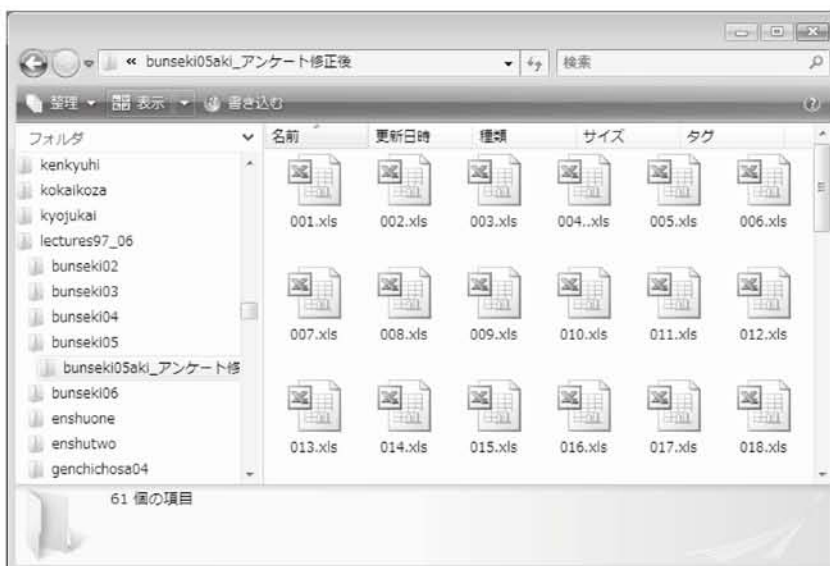


図 2 回答を入力して提出したファイル

度ランダムに配布して、最初にデータを入力した履修者とは別な履修者に回答用紙が渡るようにし、ファイルに入力されたデータに間違いや、入力の不統一などが無いかを確認する。これらの間違いや不統一があればこの段階でワークシートのデータを修正しておく。

4. 一覧表の作成

調査票の回答の入力が終わり、ファイルが全部揃ったことを確認して、履修者各自で一覧表を作成する。ここでは新たにエクセルのファイルを開いて一覧表を作成する（図 3）。

調査票の回答を 1 枚ずつ見ても、回答にどのような傾向があるかはなかなか把握できない。全体的な分析を行うためには統計表やグラフを作成する必要があり、その前段階の作業として集めた調査票を一覧表にまとめ、集計の基礎にな

るデータとして作成しておけば、その後の分析やグラフの作成を容易にすることができる。

一覧表の作成は次の手順で行う。調査票の提出用フォルダに集められたファイルを開き、整理番号順に行単位でデータのコピーを行い、各自のワークシートに貼り付けて一覧表を完成する。このときに教室内では多数の利用者が同じファイルを参照すると時間がかかることがあるので、回答を入れたフォルダを履修者各自のフォルダに貼り付けを行い、一覧表の作成を行った。

この調査票の例では一覧表を作成する場合にはいくつかの手順が考えられる。まず調査票を入力したデータには整理番号が付されているので、データの数が少ない場合は、その番号を頼りに手作業で順番に貼り付けることができる。

| 1 | 履修者アンケート | | | | | | | | | | |
|----|----------|-----|-------|------|----|----|----|----|----|----|-----|
| 2 | 整理番号 | Q1 | Q2 | Q3 | Q4 | Q5 | Q6 | Q7 | Q8 | Q9 | Q10 |
| 3 | 1 | 2 a | 157.0 | 24.0 | 1 | 4 | 4 | 3 | 3 | | |
| 4 | 2 | 2 a | 165.2 | 24.0 | 3 | 2 | 4 | 2 | 2 | | |
| 5 | 3 | 2 a | 154.0 | 23.0 | 3 | 3 | 2 | 2 | 6 | | |
| 6 | 4 | 1 a | 176.5 | 25.0 | 2 | 1 | 6 | 2 | 2 | | |
| 7 | 5 | 1 a | 170.5 | 26.0 | 2 | 1 | 5 | 2 | 3 | | |
| 8 | 6 | 1 a | 171.2 | 25.5 | 4 | 2 | 1 | 2 | 2 | | |
| 9 | 7 | 1 a | 167.0 | 26.0 | 4 | 1 | 3 | 3 | 6 | | |
| 10 | 8 | 2 a | 168.4 | 24.5 | 2 | 3 | 4 | 1 | 7 | | |
| 11 | 9 | 1 b | | | 1 | 4 | 4 | 3 | 3 | | |
| 12 | 10 | 1 c | 172.3 | 26.5 | 2 | 5 | 1 | 1 | 2 | | |
| 13 | 11 | 1 a | 166.0 | 26.0 | 1 | 3 | 5 | 2 | 3 | | |
| 14 | 12 | 2 a | 150.0 | 23.5 | 1 | 4 | 4 | 1 | 2 | | |
| 15 | 13 | 2 a | 160.0 | 23.0 | 1 | 4 | 3 | 3 | 2 | | |
| 16 | 14 | 2 a | 165.0 | 24.5 | 4 | 1 | 3 | 1 | 3 | | |
| 17 | 15 | 2 b | 150.1 | 22.5 | 1 | 5 | 4 | 3 | 3 | | |
| 18 | 16 | 2 c | 156.0 | 23.5 | 2 | 5 | 4 | 1 | 4 | | |
| 19 | 17 | 2 a | 162.0 | 23.0 | 1 | 4 | 3 | 2 | 2 | | |
| 20 | 18 | 2 a | 154.1 | 23.0 | 1 | 2 | 6 | 2 | 2 | | |
| 21 | 19 | 2 a | 151.5 | 24.0 | 1 | 2 | 5 | 2 | 3 | | |
| 22 | 20 | 1 a | 176.7 | 27.5 | 1 | 4 | 5 | 3 | 5 | | |
| 23 | 21 | 2 a | 161.0 | 24.0 | 1 | 4 | 4 | 3 | 3 | | |
| 24 | 22 | 1 a | 167.7 | 25.0 | 4 | 1 | 3 | 3 | 4 | | |
| 25 | 23 | 1 a | 174.8 | 26.0 | 1 | 3 | 4 | 3 | 3 | | |
| 26 | 24 | 1 a | 170.0 | 27.0 | 1 | 3 | 4 | 2 | 2 | | |

図3 一覧表の例(選択肢の記号で入力した例)

次にデータの数が多く場合にソートを使う別な方法として、整理番号は気にしないで貼り付けを行い、後からソートを使って整理番号順に並べ替えを行うこともできる。この場合は表の先頭部に記入する一覧表のタイトルや質問の番号などは、データのソートが終わった後に、必要な行を挿入して入力するとやりやすい。

5. 記号データの置換

上の一覧表は記号で入力してあるため、データの並びから何らかの意味を読み取るようなことは難しいかも知れない。しかしパソコンの場合は記号でデータ入力してあれば、編集メニューにある置換機

能を使って、文章で表現した具体的な質問項目に置き換えることが容易である。例えばB列は回答者の男女別を尋ねたものであり、1は男、2は女のデータを示している。この部分を数字から漢字に置き換えるときは、検索と置換機能を使えば簡単にできる。なお置換するときは「オプション」から検索方向を列に指定し、「セル内容が完全に同一であるものを検索する」もチェックする。また置換したい列を選択しておいてから行うとやりやすい(図4と図5)。

操作に失敗すると他の列まで無用な置換を行うこともあるので、そのような場合は一旦元に戻るよう指示した。



図4 エクセルの検索と置換の画面



図5 一覧表の例(回答の内容に置換した例)

数字の記号だけで入力したものはエクセルではピボットテーブルでは集計しにくいですが、文字列を含んだ回答に置き換えると、エクセルではピボットテーブルで簡単に集計することができる。またアルファベットの文字の場合は、ヒストグラムでの集計がしにくいので、COUNTIF関数を使うなどの工夫が必要になる。エクセルでアンケート調査表の集計を行う場合は、回答を数字やアルファベットで入力した一覧表と、内容を文字列で入力した一覧表の2つを用意しておくとう便利である（図3および図5）。これらの一覧表から必要な部分を抜き出し、分析目的に合わせて集計やグラフの作成を行うとやりやすい。

6. 集計とグラフ

調査票の内容をもとに集計したものを

統計表と呼ぶ。統計表には単純集計表とクロス集計表がある。単純集計は1つの質問項目について度数を求めたものである。クロス集計表は2つ以上の質問項目について度数を求め、それらをまとめて表にしたもので、質問項目間の関係などが分かりやすい。どのような統計表を作成するかは、アンケート調査の目的や質問の内容などに合わせた工夫を行う必要がある。

一覧表の上で集計やグラフの作成を行うと、入力してあるデータが見にくくなることが多いので、必要な部分を別なワークシートに貼り付けて作業を行わせた。以下に集計とグラフ作成の例を簡単に示すが、実際の授業では簡単な復習を交えて、最初にいくつか集計とグラフ作成をやって見せることにしている。

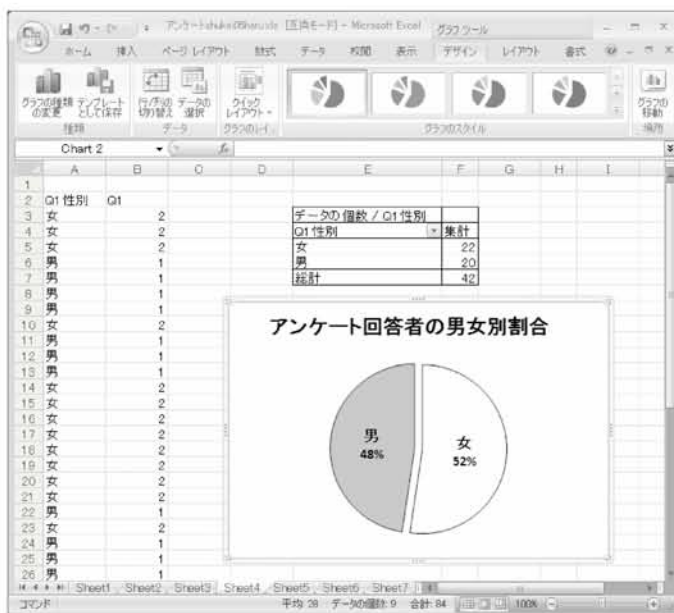


図6 アンケート回答者の男女別割合

(1) 男女別集計の例

ここでは性別のデータからピボットテーブルを使って男女別に集計する例を示す。男女別に回答者を単純集計したいときは、男女別のデータを入力した列を別なワークシートに貼り付けて作業を行うが、履修者は事前に授業のなかでピボットテーブルの操作を学習しているので、ここでは操作方法の説明は略する(図6)。図6は男女別の漢字の文字列データを使ってピボットテーブルで集計したものである。回答のデータと集計結果およびグラフを同一ワークシート上に表示するため、ピボットテーブルのデータを使って円グラフを作成した。集計結果と円グラフからは履修者には女が若干多いことが分かる。

(2) 身長と足の大きさの関係

図7は身長と足の大きさを回答してもらい、散布図を作成して身長と足の大きさに相関関係があるかどうかを調べたものである。散布図を見るとデータの分布は明らかに右上がりになっているので、履修者の身長と足の大きさには相関関係があるといえる。全体の相関係数を求めると0.80654となり、強い正の相関関係が認められる。

なお散布図には左下と右上に2つの部分的な集団があることが分かり、散布図の元のデータには男女が混在していることから、男女の差が現れていると想定される。これらことは男女別に集計しなおすと確認することができるし、男女間の相関係数に違いがあるかどうかは男女別

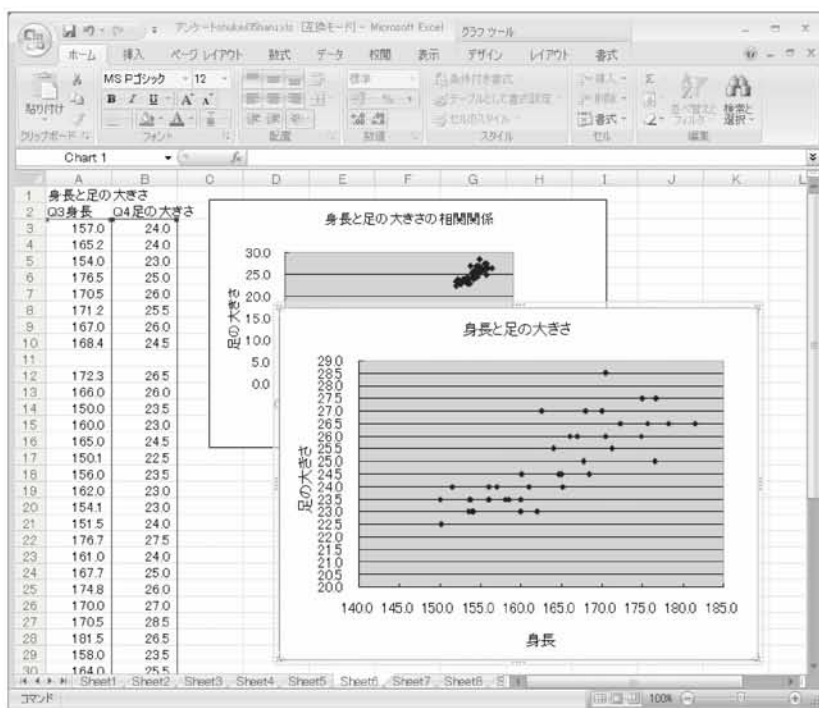


図7 身長と足の大きさの関係

に相関係数を求めると明らかになる。

(3) 通学時間と男女の差

履修者の大学までの通学時間がどのくらいかかるかを 30 分ごとに調べたものが図 8 であり、さらに男女別に調べたものが図 9 である。

図 8 は通学時間を単純集計したものであるため、それぞれの時間帯に回答者が分布していることは分かる。これに男女のデータを加えてクロス集計して図 9 を作成してみると、男女の通学時間に差があることがグラフから読み取れる。男は通学時間が長くなると次第に人数が少なくなっている。これに対して女は 1.5 時間～2 時間以内にもっとも回答が多く、多

少時間がかかっても通学していることが読み取れる。これらの現象については、女子学生は多少通学時間がかかっても親元から通学していることが想像されるが、これを確認するためには自宅から通学しているか、アパートなどから通学しているかなどを調査する必要がある。また年度を変えて調査した場合にどうなるかも検討すべき課題になる。

(4) 複数回答の集計

よく使う検索エンジンの質問は複数回答の集計例である。アンケートの調査票では複数回答の質問も頻繁に行われる(図 10)。

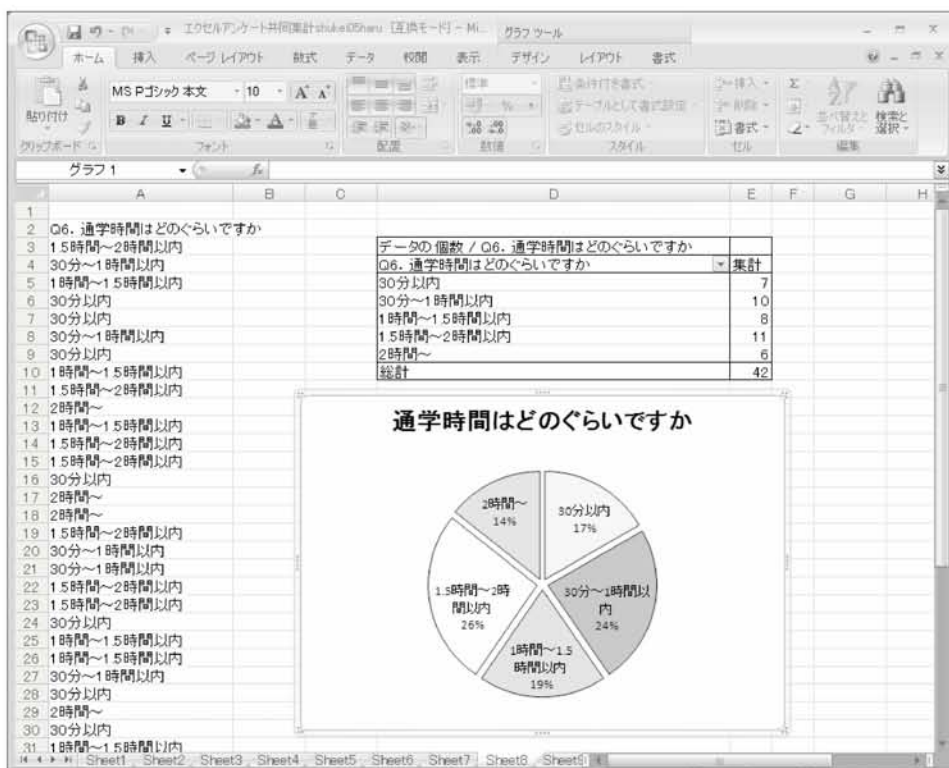


図 8 通学時間の集計

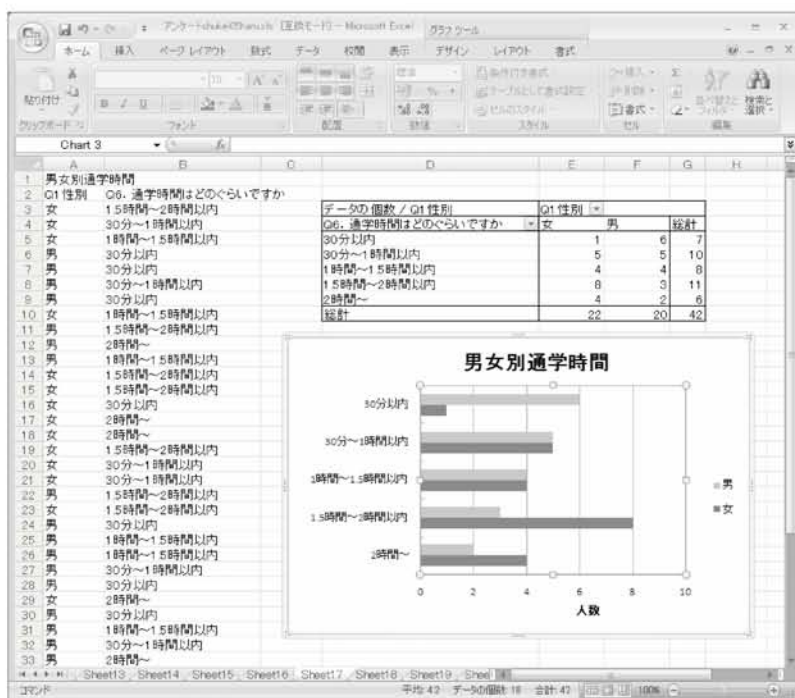


図9 男女別通学時間

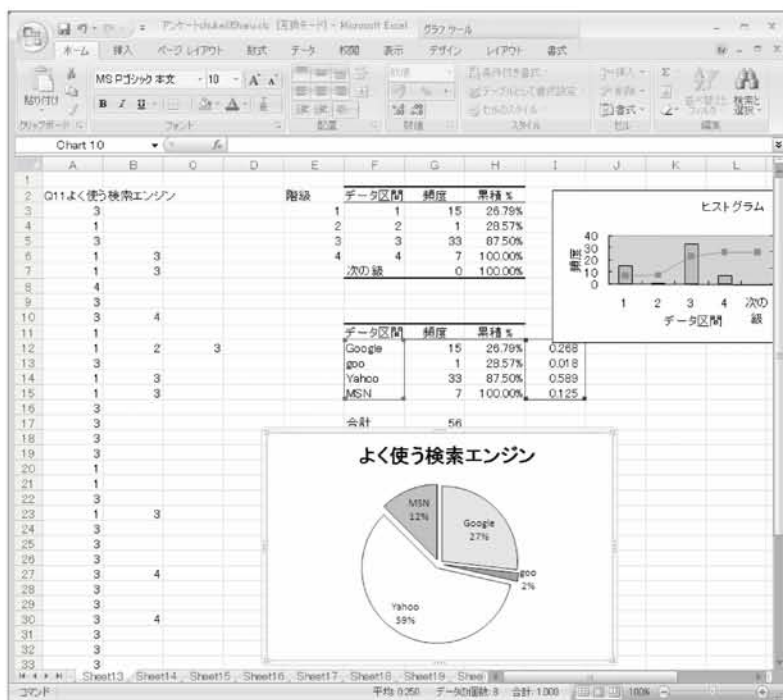


図10 複数回答の集計とグラフ(よく使う検索エンジンの例)

エクセルでは1つのセルに1つのデータ入力の基本である。そのため複数回答があるとき、例えば2つ回答があるときは、1番の回答はA列のセルへ、2番の回答はB列のセルへというように、回答の数に合わせて列とセルを確保しておくとうりやすい。あるいは複数回答が可能な5つの選択肢を設定しているようなときは、一覧表で5列を確保しておき、回答のあったセルへ入力することも可能である。

図10の良く使う検索エンジンの例では、複数回答を含めると56の回答があり、回答者全体は42人であるから14の複数回答があった。そのうち最多の複数回答が3つであったので、A列からC列まで3列を確保し、回答者ごとにA列から回答を入力した。

またエクセルのピボットテーブルで複数列の集計を行うとき、文字列のデータを入力すると、先頭のラベルの処理が面倒で集計しにくい。ここでは回答番号を数字で入力し、ヒストグラムを使って出現頻度を集計した。階級として使っているのは選択肢に付した回答番号の数字であり、下限や上限は不要である。

以上のようにして回答を単純集計してグラフを作成してみると、履修者のなかではYahoo!の検索エンジンが58%と最も使われていることが分かる。なぜYahoo!が最も多く使われるかについては更なる調査が必要になる。

7. 考察とまとめ

このアンケートの集計は2004年度の

春学期から行っており、これまでに5回実施した。授業では、講義の目的や学習する統計学の基礎概念に加えて、エクセルによるデータ処理をまとめた教科書を用意して行った。上で紹介した部分は教科書のアンケート集計で使っている部分を手直ししたものである。これまでの経験から以下のようなことが明らかになったので、いくつかの観点から考察を加えておきたい。

履修者は学期によって異なるが、秋学期の授業では1年生が大半を占め、次に2年生が多く、3年生は少ない。春学期の授業では2年生と3年生がほぼ半々となる。4年生はいずれの学期も若干名である。履修者数はほぼ55人前後（最大60人）である。

(1) エクセルの操作と慣れ

まずエクセルの操作であるが、表作成あるいは簡単な計算や集計など多少の経験のある履修者が増えてきているが、ほとんどはアンケート集計ができるほどの経験はない。しかし毎回エクセルを使いながら冒頭で紹介した内容で14回ほどの授業を行うと、履修者全員が自分自身で授業中の指示に従って集計表を作成し、簡単なグラフを作成することができるようになる。

(2) データの入力ミス

ワークシートへの入力ミスは、半角文字と全角文字を間違える、小数点とカンマを間違える、選択肢の数字で入力するべきところを言葉で入力する、数値のみ入力すべきところに単位を付ける、雛形

の列を削除してしまう、複数回答の入力指示を守らず 1 つのセルに複数のデータを入力する、などの間違いが多い。これらの間違いはワークシートで処理を行う場合に、いろいろな不都合が発生するので履修者どうして修正するようにしている。

データの入力ミスは 1 割程度の履修者に発生し、一回の修正で全部なおることは少ない。入力ミスを防ぐために、集計のための雛形を作成している。雛形を作成しないで説明だけでアンケートの集計を進めようとする、履修者ごとにばらばらな集計表ができるので、その後の作業が混乱し支障が出やすい。

(3) 集計とグラフについて

アンケートの質問数は毎回 30 弱の項目を用意している。質問の項目間には関係が深いものとそうでないものが混在しているが、男女別などのクロス集計を可能にしている。しかしほとんどの履修者は単純集計で済ませている場合が多く、クロス集計を行っている例は少ない。また靴のサイズなどヒストグラムを作成すべき回答もあるが、平均だけ求めている例が多い。クロス集計に取り組む履修者が少ないので、グラフも単純な形式が多く、クロス集計のグラフ作成例は少ない。

(4) 実際の社会調査への応用

現代中国学部では毎年現地研究実習と称して、国際社会調査を中国で行っているが、本科目の履修者の大半はこの国際社会調査に参加しており、授業で学んだことを活用することができる。2007 年は

事前調査として学内アンケートを実施し、350 部以上のアンケートを収集し、各自で集計を行うことができた。また最近の 3 年間に、中国現地研究実習報告書として出版されている「学生が見た中国社会」シリーズにも、多くのアンケート集計結果が記載されており、全員がここで紹介した授業を履修していることから、ある程度の成果をあげることができたと考えている。

注・文献

- (1) 東京大学教養学部統計学教室編：統計学入門，東京大学出版会，（基礎統計学 I），pp. 307(1991)．
- (2) 縄田和満：Excel による統計入門，朝倉，pp. 196(2000)．
- (3) 上田太一郎：Excel でできるデータマイニング入門，同友館，pp. 340(2001)．
- (4) 上田太一郎：Excel(エクセル)でできるデータマイニング演習，同友館，pp. 213(2001)．
- (5) 白石修二：例題で学ぶ Excel 統計入門，森北出版，pp. 145(2001)．
- (6) 佐藤博樹ほか：社会調査の公開データ -2 次分析への招待-，東京大学出版会，pp. 260(2000)．
- (7) 福田剛志：データマイニング，共立出版，pp. 169(2001)．
- (8) 木下栄蔵：社会現象の統計分析—手法と実例—，朝倉書店，pp. 197(1998)．