

高校生っぽいアンケート調査を - 質問票の作成から集計まで -

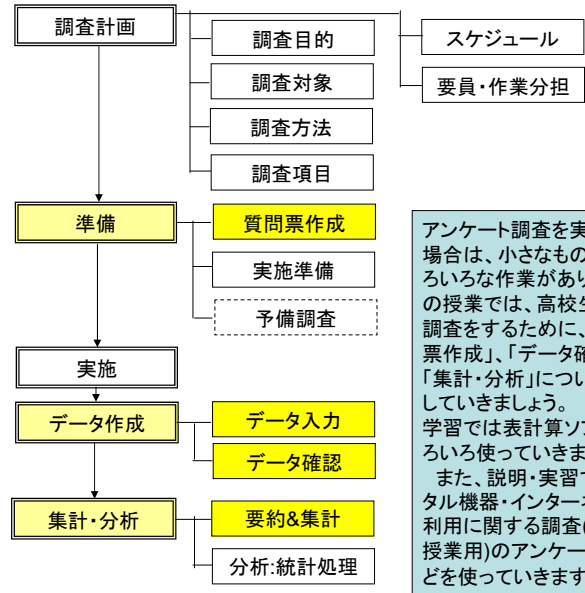
アンケート調査なんて小学校からやっているから平気、平気!



この授業では、少しは高校生らしいアンケート調査をやってみましょう。



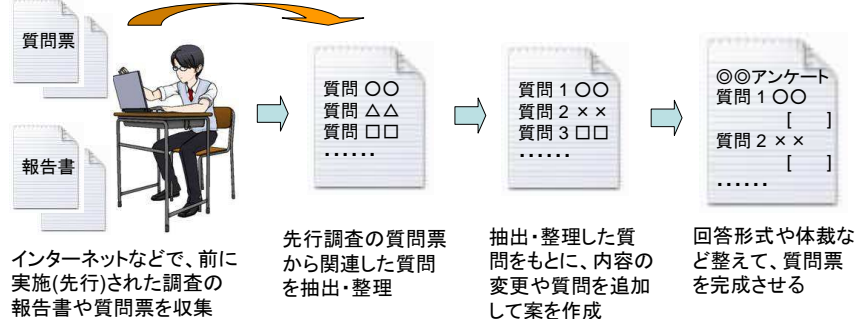
調査の流れ



アンケート調査を実施する場合は、小さなものでもいろいろな作業があります。この授業では、高校生らしい調査をするために、「質問票作成」、「データ確認」、「集計・分析」について学習していきましょう。学習では表計算ソフトをいろいろ使っていきます。また、説明・実習ではデジタル機器・インターネットの利用に関する調査(情報科授業用)のアンケート用紙などを使っていきます。



質問票作成は先行調査の情報収集から



先行調査の質問票を参考につくることはマネや盗作では、ありません。むしろ多くの本格的な研究や調査では先行事例を調べることから始まっています。そして、先行調査を調べることは次のような利点があります。

- ・一から作るよりは、楽に作れる。
- ・無駄な質問を入れたり(例えば回答者が全員同じ回答するような質問)や質問の抜けがなくなる。
- ・報告書の結果などから目的にあった質問が何かわかる。
- ・質問自体がすでに、確立されたもので信頼性が高い場合が多い。

なお、盗作ではありませんが、何をもとにして作成したか明記した方がいい場合があります。収集した質問票などのリストをちゃんと作っておきましょう。



回答形式1: 集計の手間と分析からは選択式

記述式 質問: あなたが使用しているSNSアプリを全部書いてください。
回答欄[]

- ・回答者のすべての答えがでるチャンスがある。・集計分析に非常に手間がかかる
- ・記入が大変で回答が十分得られないことがある。・質問を作るのが簡単

選択式 質問: あなたは使用しているSNSアプリをチェックしてください?

- LINE facebook twitter mixi Google+ カカオトーク その他

- ・選択肢以外の回答がある可能性あり。・集計分析がしやすい。
- ・記入が楽。・選択肢を考えるのが大変

回答形式の大きな区分として、記述式と選択式があります。上記に両方の特徴を示しましたが、集計の手間や分析の観点からできるだけ選択式になるように質問を作ります。また、選択肢で抜けがないか確認したような場合は、予備調査で、例えば「その他」を選択した時にその内容を記述してもらって、本番の調査の質問票に反映させるようにしましょう。



回答形式2: 微妙にちがう態度と印象

態度を聞く質問 (リッカード尺度)

質問1-1: あなたはLINEアプリは高校生に使用制限をつけた方がいいと思いますか?

そう思う 1 2 3 4 5 思わない

質問1-2 あなたはタブレットが高齢者に役立つと思いますか?

役に立つ 1 2 3 4 5 役に立たない

印象を聞く質問 (SD法)

質問2 あなたはLINEアプリにどのような印象を持っていますか?

明るい 1 2 3 4 5 暗い
 やわらかい 1 2 3 4 5 固い
 暖かい 1 2 3 4 5 冷たい

上の質問のように1~5を選択するような質問がありますが、実は全くことなる二つの使い方があります。ひとつは態度など聞く質問で、質問で聞いたことについて賛成か反対かなど回答者の意見を聞くものです。もう一つは、質問で提示した商品、芸術、人物、ものなどの印象を聞く場合に、対で提示した形容詞のどのような感じを持つか答えるものです。
 回答欄自体は似たようなものですが、使い分けしましょう。なお、前者については、程度を選択肢として質問することができます。(1 非常に思う 2. 思う、3. どちらでもない)



5

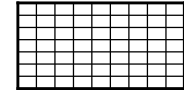
データ入力は結構大変で、誤りも起きる



回答を記入したアンケート用紙



データ入力



データ(表計算、データベース、統計ソフト)

アンケート用紙の工夫

入力が誤りやすい

質問: あなたが一番使う通信方法について一つ〇をつけてください。

1. 電話 2. メール 3. SNS 4. その他

入力が誤りにくい

質問: あなたが一番使う通信方法の番号を一つ記入してください。

1. 電話 2. メール 3. SNS 4. その他
 回答欄[]

データ入力は思ったより大変な作業です。アンケートをとる人数や質問数など十分に考えておきましょう。また、入力時にミスも起きやすいです。そのためアンケート用紙自体も入力のしやすさを考慮しましょう。

Webでの入力フォームで回答してもらおうとデータ入力の手間がはげけます。



6

データ確認: 異常値だけは確認しよう

異常値例	回答	誤りの判談	対応
質問: あなたの性別は? 1. 男子 2. 女子	3	・記入又は入力ミス	・アンケート用紙の確認 ・回答者/記入者に問い合わせ ・異常値として印をつける又は未回答に変更
質問: あなたは平日に何時間ゲームをしますか(分で回答してください)?	2000分 400分	・記入又は入力ミス 又は ・実は正しい値	・上記と同様の対応 ・正しい値の場合でも、あまりにも他のデータと違った場合は集計・統計の処理からは外す



アンケート調査は質問票の作成から実施、データ入力までいろいろ手間がかかっています。それでもデータの中におかしな値が入っていると、それまでの苦労が無駄になるかもしれません。
 そんな理由で最低でもデータの中に異常値が無い確認しましょう。異常値には、記入や入力ミスによる、明らかな間違ったデータがあります。また、ミスなのか、例えば正しい値だけで他と比べて非常に大きな値などがあります。この時、データの平均などとりつくと、正しい区判断できないことなどあるので、正しいデータでも集計や統計処理するとき外すことがあります。似たような考えかたは、オリンピックなどのフィギアや体操等の採点で、一番良い点と悪い点を除くというのも、極端な点数をはずすという意味があると考えられます。

実習: 異常値を見つけよう

	A	B	C	D	E	F
1	No	q01	q02	q03a	q03b	q03c
2	1	1	2	7	0	0
3	2	1	4	2	0	0
38	37	2	4	16	0	43
39	38	2	1	10	0	0
40	39	2	4	0	10	87
41	40	2	3	10	0	9
42	Count	40	40	40	40	40
43	CountA	40	40	40	40	40
44	Ave	1.425	2.975	15.425	2	10.25
45	Min	1	1	0	0	0
46	Max	2	4	46	20	87

- =COUNT(範囲)
- =COUNTA(範囲)
- =AVERAGE(範囲)
- =MIN(範囲)
- =MAX(範囲)

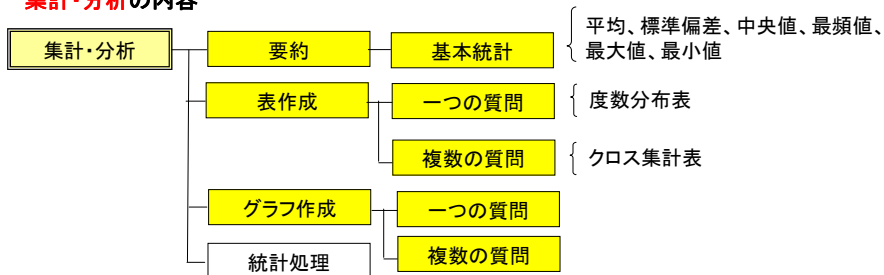
異常値を見つけるためには、そんなに難しいことはないで、入力されているデータの数、最大値、最小値を調べていきます。通常より異常に大きな数があれば異常値として扱うことになります。あと、あまり未入力が多い質問は分析するかどうか検討することもあります。同様に未入力の多い回答も集計の対象からははずすこともあります。



この段階で異常値を見つけたら、セルの色を変えておきましょう。色を変えた場合は、検索、ソート、フィルターの対象になって探すことができます。

8

集計・分析の内容



集計と分析の区別はなかなか難しく、はっきりしないこともあります。この授業では上図のような作業に分けています。要約は各質問ごとに、そのデータ全体を代表して表す値を算出します。表作成は各質問のデータをいくつかの区分に分割して、度数(各区分にいくつデータがあるか)を算出します。次に実際に作業してみましょう。



要約のできる質問、できない質問

問題のタイプ (回答の数字の意味)	問題例	要約			表作成
		平均・標準偏差	中央値	最頻値	
分類や状況を示す便宜的(名義的)な番号	質問: あなたの性別は? 1. 男子 2. 女子	×	×	○	◎ 必須
番号は大小の順番を表すが、意味する間隔が違う	質問: あなたはデジタル機器を一日どの位の時間利用していますか? 0: 使わない, 1: 1-30分, 2: 30分-1時間, 3: 1-2時間, 4: 2-3時間, 5: 3時間以上	×	○	○	◎ 必須
回答の数字が計算できるもの	質問: あなたはメールを一日どのぐらいの時間利用していますか? []分	◎ 必須	○	○	○ 階級幅

質問の回答は処理しやすいように、通常は数字で行いますが、問題によってその意味が違ってきます。一番上の例だと、例えば1.3という平均が出たとしても、意味がないですね。また、次の例だと1と2は30分間隔、3と4は60分間隔なので、例えば、1.8という平均は、では何分か考えた場合ははっきりしません。このように、質問に要約ができるものと、できないものがあることを意識しましょう。また表作成の場合は、上の二つは回答数字で直接集計できますが、最後のものは階級幅といって、数字をどう区切るか考える必要があります。



実習: 要約した値を求めよう。

No	q01	q02	q03a	q03b	q03c
1	1	2	7	0	0
2	1	4	2	0	0
37	2	4	16	0	43
38	2	1	10	0	0
39	2	4	0	10	87
40	2	3	10	0	9
データ数	40	40	40	40	40
平均	1.43	2.98	15.43	2.00	10.25
標準偏差	0.50	0.92	13.78	5.04	19.48
中央値	1	3	10	0	0
最頻値	1	3	0	0	0
最小値	1	1	0	0	0
最大値	2	4	46	20	87

異常値の2種類の対応
1. 異常値のセルをブランク又は異常値であることを記す文字を入力にする
 AVERAGEなどの関数はブランクや文字を除いて計算してくれるので、左下の指定で問題なく計算してくれます。
2. 異常値を残す又は、異常値であることを示す数値(例えばマイナスの数を入れておく)
 AVERAGEやSUMIに対してAVERAGEIFやSUMIFなどの特定の条件にはまる数値のみを対象に計算する関数があります。これで異常値を計算しないようにします。

- =COUNT(範囲)
- =AVERAGE(範囲)
- =STDEV(範囲)
- =MEDIAN(範囲)
- =MODE(範囲)
- =MIN(範囲)
- =MAX(範囲)

異常値の対応では、簡単にするにはブランクにしたり文字を入力したりすることがお勧めです。ただし、異常値であったことを記録しておくためにセルの色は変えておいた方がいいですね。



実習: 一つの質問の表を作成しよう: 度数分布表 (回答が便宜的(名義的)又は順番を表す数字。)

	A	B	C	D	E	F	G	AC	AD	AE
1	No		q01		q02		q07a			
2	1		1		2		0			
3	2		1		4		0			
41	40		2		3		1			
42	データ数		40		40		40			
43	平均		1.43		2.98		0.175			
44	標準偏差		0.50		0.92		0.308			
45	中央値		1		3		0			
46	最頻値		1		3		0			
47	最小値		1		1		0			
48	最大値		2		4		1			
49		度数	%		度数	%	度数	%		
50	1. 男	23	57.5%	0:使わない	0	0.0%	1:はい	7	17.5%	
51	2. 女	17	42.5%	1:1-30分	4	10.0%	0:はい又	33	82.5%	
52				2:30分-1時間	5	12.5%				
53				3:1-2時間	19	47.5%				
54				4:2-3時間	13	30.0%				
55				5:3時間以上	0	0.0%				

便宜的又は順番を表す数字の場合は、個々の値の数をカウントしてテーブルを作ります。具体的にはCOUNTIF関数を使えばカウントする条件を指定できます。選択枝の数だけCOUNTIF関数を使うことになりますね。

- =COUNTIF(範囲, "1")
- =COUNTIF(範囲, "2")

- =COUNTIF(範囲, "0")
- =COUNTIF(範囲, "1")
- =COUNTIF(範囲, "2")
- =COUNTIF(範囲, "3")
- =COUNTIF(範囲, "4")
- =COUNTIF(範囲, "5")

- =COUNTIF(範囲, "1")
- =COUNTIF(範囲, "0")

%は各度数をデータ数で割った値です。[書式]表示形式]はパーセンテージを指定しています)



実習: 一つの質問の表を作成しよう: 度数分布表 (回答が数値)

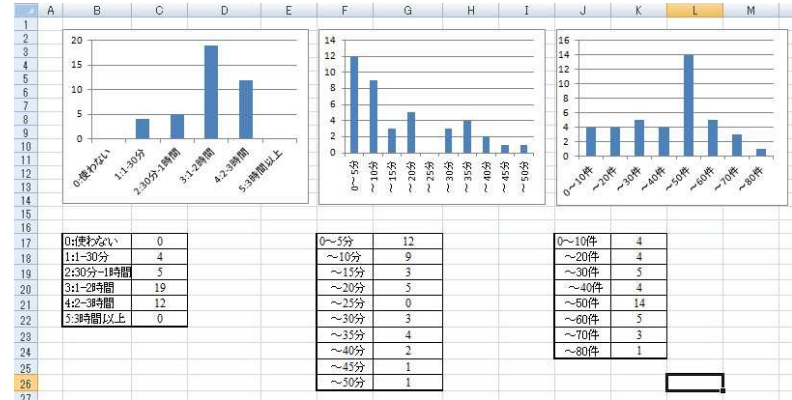
	A	H	I	J	V	W	X
1	No		q03a			q04	
2	1		7			43	
3	2		2			53	
40	39		0			24	
41	40		10			35	
42	データ数		40			40	
43	平均		15.43			37.93	
44	標準偏差		13.78			19.28	
45	中央値		10			44	
46	最頻値		0			0	
47	最小値		0			0	
48	最大値		46			79	
49							
50		度数	%		度数	%	
51		0~5分	12	30.0%	0~10件	4	10.0%
52		~10分	9	22.5%	~20件	4	10.0%
53		~15分	3	7.5%	~30件	5	12.5%
54		~20分	5	12.5%	~40件	4	10.0%
55		~25分	0	0.0%	~50件	14	35.0%
56		~30分	3	7.5%	~60件	5	12.5%
57		~35分	4	10.0%	~70件	3	7.5%
58		~40分	2	5.0%	~80件	1	2.5%
59		~45分	1	2.5%			
		~50分	1	2.5%			

度数のカウントは同様にCOUNTIF関数を使ってできます。ただし単純にカウントできないので、個々の階級の範囲に入るデータをカウントしなければなりません。そのため、階級までの個数 - 1つ前までの階級の個数で計算します。

- =COUNTIF(範囲, "<=10")
- =COUNTIF(範囲, "<=20") - COUNTIF(範囲, "<=10")
- =COUNTIF(範囲, "<=30") - COUNTIF(範囲, "<=20")
- =COUNTIF(範囲, "<=40") - COUNTIF(範囲, "<=30")
- =COUNTIF(範囲, "<=50") - COUNTIF(範囲, "<=40")
- =COUNTIF(範囲, "<=60") - COUNTIF(範囲, "<=50")
- =COUNTIF(範囲, "<=70") - COUNTIF(範囲, "<=60")
- =COUNTIF(範囲, "<=80") - COUNTIF(範囲, "<=70")

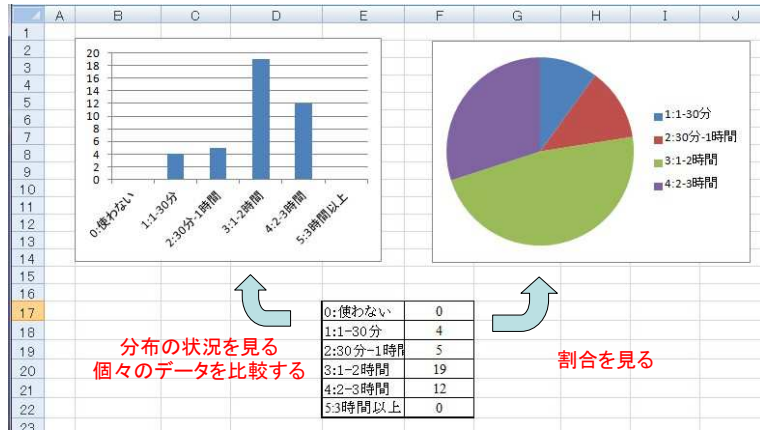
まず、最大値などに注目し、階級幅(左では5分単位、右では10件単位)を決めます

実習: グラフの作成: 度数分布表から度数分布図(ヒストグラム)



度数分布表を作ったあとに、グラフを作るとデータの状況がより判りやすくなります。Excelでは表ができていれば、比較的簡単に度数分布図(ヒストグラム)を作成することができます。
[グラフ]-[縦棒]-[2D縦棒]-[集合縦棒]

実習: グラフの作成: 比率を見る場合

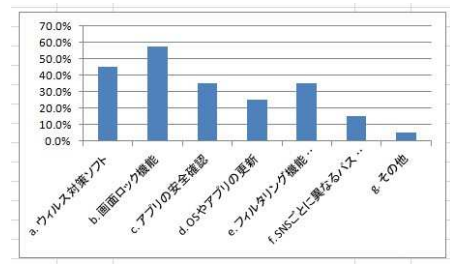


分布の状況を見る
個々のデータを比較する

割合を見る

同じ表を使っても、何をみたいかで、作成するグラフを変えましょう。比率をみたいときには円グラフが適しています。

実習: グラフの作成: 複数選択の場合

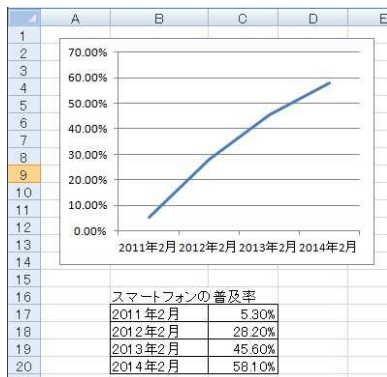


1つの質問で、複数選択可能な場合は、それぞれの選択肢がどの位の割合で選ばれたかを比較して見ることが多く、棒グラフで比較すると理解しやすいことが多いです。
また、並び替えをして大きな数のものから表示するのも一つの手です。

問題	セキュリティ対策	%
q06a	a. ウィルス対策ソフト	45.0%
q06b	b. 画面ロック機能	57.5%
q06c	c. アプリの安全確認	35.0%
q06d	d. OSやアプリの更新	25.0%
q06e	e. ファイル共有機能の利用	35.0%
q06f	f. SNSごとに異なるパスワードの使用	15.0%
q06g	g. その他	5.0%



実習: グラフ作成: 時の流れの変化を見る。



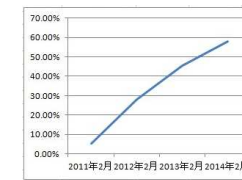
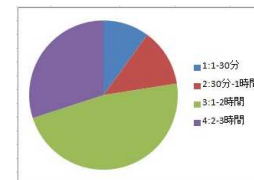
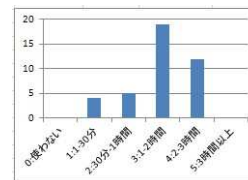
年による変化や、年齢による違いなど見る場合は、折れ線グラフがぴったりですね。



ソース:博報堂DYグループ・スマートデバイス・ビジネスセンター、「全国スマートフォンユーザー1000人定期調査」第9回分析結果報告
<http://www.hakuhodo.co.jp/archives/newsrelease/16784>

グラフの種類と使い方

見方	扱う数値	棒グラフ	円グラフ 帯グラフ	折れ線 グラフ
データの比較	数量	○		△
	割合		○	
データの変化	数量/割合	△		○



**実習: 二つの質問の表を作成しよう: クロス集計表
 (回答が便宜的(名義的)又は順番を表す数字。)**

No	q01	q02
1		
2	1	2
3	2	4
40	2	4
41	2	3

	度数	%	度数	%
1.男子	23	57.5%	0:使わない	0
2.女子	17	42.5%	1:1-30分	4
			2:30分-1時間	5
			3:1-2時間	19
			4:2-3時間	12
			5:3時間以上	0

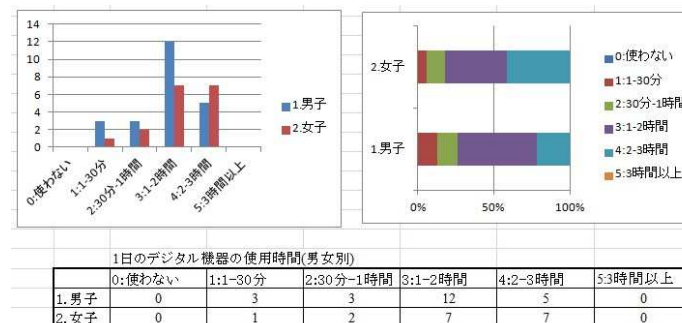
クロス集計表で、回答が便宜的又は順番を表す数字の場合は、単純に条件に合うデータ数をカウントします。COUNTIFS関数を使うと複数の条件を指定できます。

=COUNTIFS(C2:C41,"=1",F2:F41,"=0")	=COUNTIFS(C2:C41,"=1",F2:F41,"=1")
=COUNTIFS(C2:C41,"=2",F2:F41,"=0")	=COUNTIFS(C2:C41,"=2",F2:F41,"=1")



Excelには度数分布表やクロス集計表をつくる関数や機能もあります。これらは使い方はすこし難しいですが、慣れると手早く表を作成することができます。興味があったらWebなどで調べてみてください。

実習: グラフ作成: クロス集計表のグラフ化



単純な集計では全体の様子が判りますが、クロス集計表にすると、細かい要因ごとの特徴が明確になります。またグラフを作成するとより理解しやすくなります、1つの質問をグラフ化した時と同様に、棒グラフは数の比較、円グラフ又は帯グラフは割合を示します。

実習: 二つの質問の表を作成しよう: クロス集計表
(回答が数字の場合)

	A	B	C	D	E	F	G	H	I	J	
1	No		g01			g04					
2	1		1			43					
3	2		1			53					
40	39		2			24					
41	40		2			35					
42	データ数		40			40					
49											
50			度数	%		度数	%				
51	1.男子	23	57.5%	0~10件	4	10.0%					
52	2.女子	17	42.5%	~20件	4	10.0%					
53				~30件	5	12.5%					
54				~40件	4	10.0%					
55				~50件	14	35.0%					
56				~60件	5	12.5%					
57				~70件	3	7.5%					
58				~80件	1	2.5%					
59											
60			1日の出すメールの件数(男女別)								
61			0~10件	~20件	~30件	~40件	~50件	~60件	~70件	~80件	
62	1.男子(累積)	3	5	7	10	17	19	22	23		
63	(度数)	3	2	2	3	7	2	3	1		
64	2.女子(累積)	1	3	6	7	14	17	17	17		
65	(度数)	1	2	3	1	7	3	0	0		

男子(累計)	=COUNTIFS(C2:C41,"=1",F2:F41,"<=10")	=COUNTIFS(C2:C41,"=1",F2:F41,"<=20")	...
(度数)	=C62	=D62-C62	...

=COUNTIFS(C2:C41,"=1",F2:F41,"<=20") - COUNTIFS(C2:C41,"=1",F2:F41,"<=10")でも可 21

ここまで、アンケート調査を行うための基本的な知識や技能を学習してきました。アンケート調査は思ったより集計に手間がかかるものです、質問票を作るときは、先行事例や調査の目的などを十分に考慮して、適当な質問数を考えるとともに、最終的にどのようなクロス集計を行うか考えておきましょう。
では、がんばってアンケート調査をしてみましょう。

