

Excel × 亥年現象

「選挙疲れ」って、本当にあるのだろうか

2019.6.8

このスライドの最新版とデータは
github.com/nishioWU/JNPC1906
にあります

目標 = データ整形に慣れる

- ① 「**F4**」で絶対参照と相対参照を使い分ける
- ② VLOOKUPやIFが怖くなる
- ③ 「**名前をつける**」を活用する
- ④ 便利な「**テーブル**」を使う
- ⑤ グループ分けした散布図（層別散布図）を描いてみる

やるからには、参院選がらみのお題で挑戦してみましよう

目標 = データ整形に慣れる

- ① データを入手する
- ② 整形・加工する
- ③ 作図して分析する
- ④ 検定・区間推定する
- ⑤ その結果を持って、また現場取材する

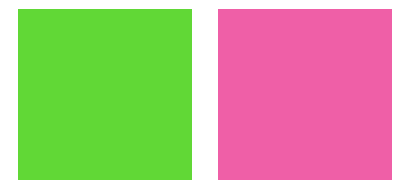
ぐるぐる回る探索プロセスのうち、きょうは②と③です

亥年現象、選挙疲れは本当か

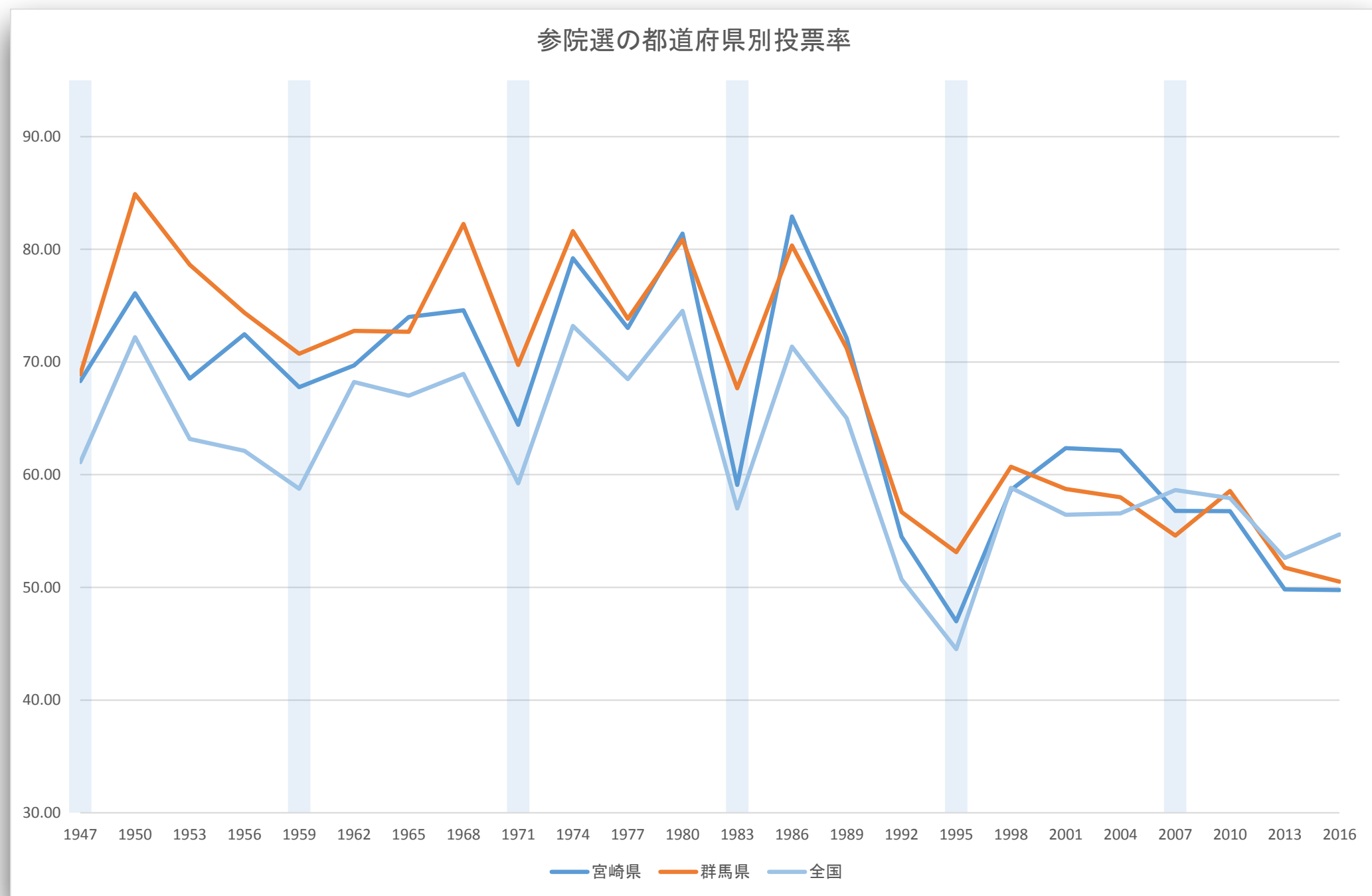
- 亥年は統一地方選の「選挙疲れ」が原因で参院選の投票率が下がる、
と言われている
- では、前回の亥年2007年と、その1回前の2004年の投票率を比べてみる。07年のほうが高いorz
- とはいえ、亥年でなかったら、もっと上がったのかもしれない

何をとっかかりにしたらよい？

07年に投票率が下がった県もある。なぜ？



トレンドを眺めてみる



表示する県を選べる折れ線グラフ：[zenkoku trend only.xlsx](#)

県議選と知事選の有無に注目

- 統一選と言っても、すべての都道府県議選があるわけではない。県議選があったかどうかで、グループに分けて比較してみたら？
- とはいえ、県議選の有無で分けると44対3。知事選を入れるにしても、統一選だけ見ると13対34
- でも、統一選からは外れていても、同じ年で、かつ参院選よりは前なら、「亥年現象」に寄与するのでは？ 調べてみる価値あり

県議の無投票当選に注目

- 道府県議選の選挙区単位で見れば、無投票当選が結構ある
- 無投票だった市区町村と、選挙戦になった自治体とで、その後の夏の参院選の投票率 or 3年前と比べた高低に差があるか？
- 2019年は岐阜などで無投票当選が多かった。今夏の国政選の結果が出たときに、同じ手法を適用したら、何か見えてくるかも
- ただし、統計学的には問題あり。県議選がある選挙区で無投票になるかどうかは、その地区の国政選の投票率と無関係に決まっている（ランダムにコントロールされている）とは考えにくい

ところで、始める前に…

www.menti.com

をクリックして下さい。
もしコードを訊かれたら

47 40 68

または、このQRから▶▶▶



サイトでデータを見つけたら

- ダウンロードできるものは、する
- サイトに載っているだけで、ダウンロードできないときは、コピー & ペーストで、Excelのシートに貼り付ける
- Chromeを使っているなら、「Copytables」を入れておくと便利。Altキーを押しながら、マウスで範囲指定すると、さほど困らない。簡易集計もできる
- 貼り付けるときは、Ctrl + Vではなく、**Alt + Ctrl + V**で。「テキスト」を選ぶと、リンクも書式もない状態に変換して貼り付けることができる

CSV形式のデータとは

- 拡張子「.csv」はコンマ区切りのテキストのこと
- テキストエディタで中身を確認しておくとい
- 文字コード違いで化ける。**UTF-8**か、**S-JIS**（CP932）か
- 先頭の0を削るなど、Excelが気を利かせすぎるのが困りもの。日付も要注意。年がない場合は今年にされてしまう

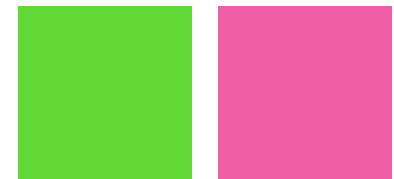
なので、CSVの**ダブルクリックはNG！** 文字化けしない場合でも

Excelを先に起動しておき、そこにデータを取り込む

整形データの保存形式としてはオススメ。受け渡ししやすい

CSVの読み込み手順 1



- ① Excel左上の「データ」タブで
- ② 「テキストファイル」を選択
- ③ CSVファイルを選ぶ。「data」の下「sangiin_csv_UTF8」フォルダにある「zenkoku_hinagata.csv」。「開く」が「インポート」に変わるので、そのボタンをクリック
- ④ 「Unicode (UTF-8)」を選ぶ
- ⑤ 取り込み開始行を指定
- ⑥ 「先頭行をデータの見出しとして使用」に✓
- ⑦ 「次へ」

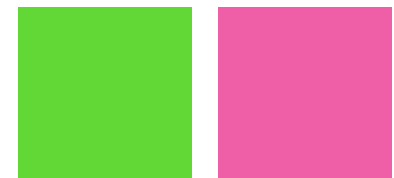


CSVの読み込み手順 2

- ⑧ 「区切り文字」を変更。「カンマ」に✓を
- ⑨ プレビュー画面を確認して「次へ」
- ⑩ 列のデータ形式を指定。もし、Excel任せでは危ない列があったら、その列を選択してから、上のラジオボタン●で「文字列」にする。0で始まる数字や、年の入っていない日付など。「標準」のままでは、勝手に今年にされてしまう
- ⑪ 今回はデフォルトのままで「完了」
- ⑫ データの貼り付け先を適宜指定し、「OK」

CSVの読み込み手順 3 完

- ⑭ OKを押せば、シートの左上隅を起点に貼りつく  
- ⑮ 左上隅以外や、新しくシートを増やして貼る場合には、その旨指定を
- ⑯ 「data」 の下の 「sangiin_csv_UTF8」 フォルダにある
「chihousen_nittei_UTF8.csv」 も、同様に読み込む
- ⑰ 分かりやすいシート名に適宜変更



ここまで済んだ状態：[zenkoku-00.xlsx](#)

Excelの勘どころ 1

▶ 行と列

横を行、縦を列とかカラムと呼んで区別している。

▶ 式は小文字で

Excelは、大文字でも小文字でも命令を聞いてくれる。なので、関数は小文字で入力するとよい。正しく認識されれば大文字に変換される。小文字のまま残ったら、打ち間違いだと分かる。

▶ 絶対参照

式をコピーすると、気を利かして、計算対象の行や列をずらしてくれる。それが便利だからだが、困る場合もある。そのときは、ずらされては困るものに「\$」マークをつけると、コピー先でもずれない。これが絶対参照。式の入力中に「F4」キーを押すと、行と列の両方またはどちらかに、\$がついたり消えたりして切り替えられる

Excelの勘どころ 2

▶ 名前をつける

「数式」タブ → 「名前の定義」で。シート上のある範囲に、名前をつけ、その名前を数式の中で使うことができる。VLOOKUPの第2引数（探しに行く範囲）を、このセルからこのセルまで、といちいち書くのは煩わしいが、この方法なら分かりやすい。式で使う場合には、自動補完が働くので入力が楽

▶ ウィンドウ枠の固定

スクロールしても、選挙が行われた年の行や、都道府県名の列がいつも見えているほうが作業しやすい、という方にオススメ。

「表示」タブ → 「ウィンドウ枠の固定」で。選択したセルの上の行・左の列が固定される

絶対参照と\$と「F4」

セルの式をコピー＆ペーストやフラッシュフィルで増殖させると、気を利かして、計算対象の行や列をずらしてくれる。それが便利だからで、このデフォルトの動作を「相対参照」という。

困る場合もある。そのときは、ずらされては困るものに「\$」マークをつけると、コピー先でもずれなくなる。これが「絶対参照」。式の入力中に「F4」キーを押すと、行と列の両方またはどちらかに、\$がついたり消えたりして切り替えられる。

VLOOKUPの2番目の引き数である「探す範囲」は、絶対参照にする。でないと、だんだんずれていってしまい、困るはず。

「範囲に名前を付ける」という方法もおすすめ。名前をつけた場合は、ずれずに固定される

能率が上がるショートカット

- ▶ シート名のタブをつかんで、Ctrl + ドラッグ

そのシートのコピーを作成

- ▶ ブック間で、シートをドラッグ

シートごと移動。別シート参照の数式は、値だけにしておくこと

- ▶ Shift + ドラッグ

行や列を選択し、その境目をつかみながらだと、並び替え

- ▶ Ctrl + 1

セルの書式設定。エルではなくて数字の一（テンキーの1はダメ）

- ▶ Ctrl + A

シート全体/表全体を選択（選択中のセルが表の外/内の場合）

- ▶ Ctrl + C

コピー

ショートカット 2

▶ Ctrl + V

通常の貼り付け。セル幅以外すべて引き継ぐ。もう一度押すと、貼り付けの形式を選ぶ

▶ Alt + Ctrl + V

形式を選択して貼り付け。関数を使って整形をした後、貼り直して「値だけ」に固定するのに便利（Vかテキストを選ぶ）

▶ Alt + ;

非表示の列は無視して、見えているセルだけをコピー元にする

▶ Ctrl + S

ファイルを上書き保存

▶ 「F12」

ファイルを別名で保存

ショートカット 3

- ▶ Ctrl + Z

直前の変更を元に戻す

- ▶ 「F2」

セルの編集

- ▶ Ctrl + F

検索

- ▶ Ctrl + H

置換

- ▶ Ctrl + カーソルキー

空白セルは飛ばし、その次にデータの入っているセルにジャンプ

- ▶ Ctrl + Home

A1セル（左上）にジャンプ

ショートカット 4完

▶ Ctrl + End

データの入っている最終セルにジャンプ

▶ Ctrl + ;

きょうの日付を入力。便利

▶ Ctrl + :

現在の時刻を入力。便利

▶ Ctrl + *

データが入っている範囲を選択。離れ小島は選択されない

▶ Ctrl + T

テーブルにする

▶ Ctrl + Enter

複数のセルに同じデータを入れる。一括して修正するときに便利

▶ Alt + 下矢印

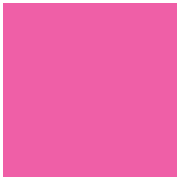



そのカラムに入力済みのデータのリストから選ぶ

別のブックからシートを移動

- ① データがまだ入っていない、21～24回参院選の元ファイルは「data」の下に「sangiin_original」フォルダにある、「san_senkyoku_hikaku21-20.xls」から「san_senkyoku_hikaku24-23.xls」までの4つ
- ② 1枚目のシートをCtrl + ドラッグで複写。シート名は西暦（21回なら2007年、24回なら2016年）の4桁数字にする
- ③ そのシート全体をコピーして、値だけを貼り付け直す
- ④ 20回目までのデータが入っているブックに、シートをそのままドラッグ。複写元のブックは閉じてよい。これを繰り返す



INDIRECT 何をしている？

- ① 2列目に、第何回目の参院選かの数字を入れる。なるべく楽な方法で
- ② 3列目に、年を入れる。B3のセルなら「=year(b\$1)」
とする。これも、横に増殖させる
- ③ V列以降に、VLOOKUPで別シートの値を取ってくる。V4のセルに、「=vlookup(\$a4, indirect(v\$3&"!\$a\$8:\$a\$55"), 1, FALSE)」と打つ。打ち間違いがなければ、式は大文字に変換され、セルに何か出るはず
- ④ V列の下まで増殖させる
- ⑤ INDIRECTごちゃごちゃは、何をしている？ 「！」がヒント

ここまで済んだ状態：[zenkoku-01.xlsx](#)

VLOOKUP(ア, イ, ウ, エ)

ア：ここに入っているのと同じものを

イ：この範囲の左端の列で探して、一致するものがあつたら

ウ：その右にこれだけ行ったところのデータを取ってくる

エ：FALSE

※アを探しに行く、イのキー列は範囲の左端でないとダメ

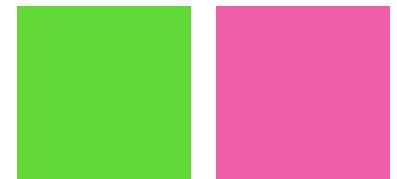
※イのキーは一意でないといけない。同名のものが複数回出てくるのはNG。一意かどうかは、条件付き書式で確認できる

※イは範囲に「名前をつける」のも一案

※エは「FALSE」で決め打ち。省略時は完全一致検索にならない

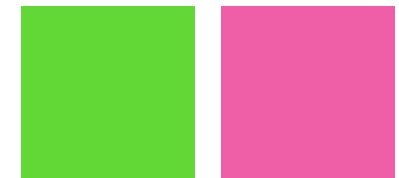
VLOOKUPに探させる

- ⑥ 21～24回のデータを、2007～2016のシートから取ってくる。
実は都道府県の並び順が揃っているので、今回はコピー＆ペーストでもできないことはない。が、練習のため、ぐっところえて、先ほどのVLOOKUPで
- ⑦ データのあるシートは4枚とも同じ形。なので……
- ⑧ 第3の引数（何列右に行ったところか）は、全部同じ
- ⑨ 第2の引数（探しに行く範囲）も工夫できないか、と考えた。
「シート名!セルの範囲」を指定するのにINDIRECTを使った。そのために、シート名を西暦年にしておいた
- ⑩ 引数を正しく直して、Y列まで増殖させる



テーブルにする

- ⑪ マッチしない箇所の手直しを
- ⑫ シートを複写。全体を選択して、値だけを張り直す
- ⑬ 先頭（都道府県名より前）に1列追加し、通し番号を振る。00～09のように、0埋めした2桁数字にする（都道府県コードになる）のでもよい
- ⑭ テーブルにする。Ctrl + A や Ctrl + * で範囲指定してCtrl + T。もし、表を解除したいときは、テーブルのエリア内で右クリックして、「テーブル」「範囲に変換」を選ぶ。セルの背景色はそのままでは消えない



ここまで済んだ状態：[zenkoku-02.xlsx](#)

折れ線グラフを描く

- ① 都道府県名から2016年までの列を選択
- ② 「挿入」から、折れ線グラフを選ぶ。横軸が年で、時系列の折れ線が、都道府県の数だけ引かれるものを。選択肢に出て来ないときは、「おすすすめグラフ」から選ぶか、グラフエリアを右クリックして、「グラフの種類の変更」を試す
- ③ トレンドはどの県も似ているが、投票率は県によって高低の差があることが分かる
- ④ 表のフィルター機能やソート機能を使い、注目する県だけに数を絞るとグラフが見やすくなる

ここまで済んだ状態：[zenkoku-03.xlsx](#)

IF(ア, イ, ウ)

ア：この条件が成立しているか、真偽を調べて

イ：成立している（つまり「TRUE」のとき）なら、セルの中身をこれに

ウ：成立していない（つまり「FALSE」のとき）なら、これに

※アの部分には、AND関数やOR関数を使った、複合条件を入れることもできる

※ウは省略可能

AND(ア, イ)とOR、NOT

アとイがどちらもTRUEのときだけ、ANDはTRUE

アかイの少なくとも一方がTRUEであれば、ORもTRUE

NOT(ア)は、アの真偽値を反転する。TRUEはFALSEに、FはTに

地方選の有無を取り込む 1

- ① 2007年に、その都道府県で議員選・知事選が行われたかを取り込みたい。県議はそのままVLOOKUPで。今回は「名前をつける」方法で試してみても
- ② 知事選は？ 単に07年というだけでなく、参院選より前に行われたものだけを判別したいので、条件式を使う。知事選と県議選の日程のシートに1列増やし、「=or(d2, and(e2 <> "", e2 < ○○)」というような式を入れる。○○が参院選投票日。式が複雑なのは、空白が「小なり」判定されてしまうため。セルに何か入っていること、を課している
- ③ もちろん、この件数なら、手作業でもよい

ここまで済んだ状態：[zenkoku-04.xlsx](#)

地方選の有無を取り込む 2

- ① データの突き合わせはVLOOKUPで。07年かつ参院選より前に「県議選があった」「知事選があった」をそのまま引っ張ってくる作業列を作るはず
- ② AND(「県議選があった」, 「知事選があった」)は、つまり両方ともあった、ということ
- ③ AND(「県議選があった」, NOT(「知事選があった」))は、県議選だけがあった、ということ
- ④ AND(NOT(「県議選があった」), NOT(「知事選があった」))は、どちらもなかった、ということ
- ⑤ 等々、作業列を元に、グループ分けに使っていく

ここまで済んだ状態：[zenkoku-04.xlsx](#)

グループ分けした散布図

- ① 都道府県議選だけがあった、知事選だけがあった、両方があった、どちらもなかった、の4通りにグループ分けして、色分けした散布図を描きたい。「層別散布図」という
- ② 横軸を04年の投票率、縦軸を07年の投票率とする
- ③ 07年のほうを、グループごとに新しく作る計4列に分ける
- ④ IF関数を使って、その列の条件に当てはまるものは07年の数値を入れ、当てはまらなければ空欄にする。IFの中で、ANDやOR、NOTを使う
- ⑤ 04年と、4列に分けた07年の計5列を選択し、次ページの方法で散布図にする

途中経過：[zenkoku-05.xlsx](#)

ここまで済んだ状態：[zenkoku-06.xlsx](#)

散布図を描く方法

- ① 表にある列を最低2つ選ぶ。2列目が離れているなら、1列目を選んだ後、Ctrl + クリックで追加
- ② 「挿入」タブから「グラフ」の「散布図」を選ぶ
- ③ 選択中の左端の列がX座標、残りは何列あってもY座標になる
- ④ 空欄がY座標の0扱いされるので、X軸のところに無用のドットがたまる。Y軸の目盛りを適正な範囲に変えるか……
- ⑤ または、ソートして空欄をまとめてしまい、列全体ではなく、データがある箇所だけを選択して散布図を描くとよい
- ⑥ 元データにはあるがグラフには不要な行があれば、表のフィルター機能で非表示にしておくと、グラフからも消える

散布図 手動で頑張るなら

- ① 今あるグラフを手直しするか、列を選ばずに空っぽのままの散布図を挿入。グラフエリア右の漏斗のアイコンから、「データの選択」に進み、左側の「凡例項目」の窓で指示する
- ② 「追加」で開く「系列の編集」パネルで、上から順に項目名（じか打ちしても、入っているセルを指定しても可）、X座標のデータ範囲、Y座標のデータ範囲、を指定。XとYは先頭のイコールを残しておく。「OK」で完了
- ③ やり直す場合は「編集」。不要な列があれば「削除」

参考) 散布図をひとひねり

- ① グループごとに大まかな分布を見る「箱ひげ図」の代用めいた図にもできる。横軸は地方選の有無、縦軸は07年と04年の投票率の差にする。
- ② X軸は新たに1列作る。「県議選あり」「知事選あり」「両方あり」「なし」を、1～4とした
- ③ Y軸も新たに1列作る。07年の投票率から04年の投票率を引いたものを入れる
- ④ ②と③の計2列を選択して、散布図にする
- ⑤ 点の重なり（オーバープロット）が発生するので、透明度を上げて薄くしておく

ここまで済んだ状態 + α : [zenkoku-08complete.xlsx](#)

参考になるサイト

- ・「エクセルによる層別散布図の作り方」

<https://hitorimarketing.net/tools/stratified-scatterplot.html>

- ・「エクセル 散布図グラフの作り方」

<https://www.tipsfound.com/excel/05036>

- ・「統計グラフ Excelで簡単に箱ひげ図を作る方法」

<https://did2memo.net/2017/02/12/excel-easy-box-and-whisker-plot/>

2016以降では、箱ひげ図がデフォルトで用意されている

参考) 文字化けをIEで直す

CSVかTXT形式のデータが文字化けしているときは、エンコードの違いが原因。UTF-8かSHIFT-JISを試してみる。実はInternet Exploreでコードを変換して保存し直す手がある。知られていないが、覚えておくと役に立つ

- ① 拡張子が「.csv」の場合は「.txt」に変えて保存。ピリオドまで消さないように注意
- ② IEでファイルを開く。化けていれば、画面を右クリックしてエンコードを直す。たいてい、自動認識してくれる
- ③ 保存したい形式（上記2通りのどちらか）を選び、別名で保存。別名にしないと、元が消えてしまう
- ④ 拡張子を「.csv」に戻す

SUBSTITUTE(ア, "イ", "ウ")

ア：このセルのうち

イ："ここに書いた文字（または文字列）"を

ウ："こちらの文字（または文字列）"に置き換える

※ウで、引用符の中に何も入れない（引用符を続けて2つ打つ）ようにすれば、イを見つけたら削除する、という動作になる

これらの関数も便利

ASC	全角文字を半角に変換。反対はJIS
COUNTA	空白以外のセルの数をカウント。COUNTなら数値の入ったセルのみ
EXACT	2つの文字列が等しいかどうか判定
FIND	ある文字列が他の文字列の中にあるか検索。大文字と小文字を区別。ワイルドカードは使えない
JIS	半角文字を全角に変換。反対はASC
LEFT	文字列の左から、指定された字数を取り出す
LEN	文字列の字数が分かる
MID	文字列の途中の指定の位置から、指定の字数を取り出す
RIGHT	文字列の右から、指定された字数を取り出す
SEARCH	ある文字列が他の文字列の中にあるか検索。FINDと違い、大文字と小文字は区別せず。ワイルドカードが使える
TRIM	前後の余分な空白を取り除く。途中の空白は1つだけ残す

延長戦@大阪府 1

- ① Excelのファイルを新規作成。空っぽのシートを用意しておく
- ② ブラウザで2007年の大阪府議選のデータを開く。今回知りたいのは、投票があったのか無投票当選だったのか、なので「投票速報」をクリック。ダウンロード用のデータは用意されていないので、サイトの表をコピーして貼り付けるしかない
- ③ Excel のシートへのペーストは、Alt + Ctrl + Vで。「テキスト」を選ぶこと
- ④ 表が複数に分かれているときは、これを繰り返す
- ⑤ 表の範囲外でも必要な事項はコピー＆ペーストしておく

延長戦@大阪府 2

- ⑥ Excelの別のシートに、07年と04年の参院選の投票率データも持ってくる。空っぽのシートを追加
- ⑦ ブラウザで07年の参院選大阪選挙区のデータを開く。「大阪府選出選挙」の「投票速報」をクリック。前ページの方法で、シートに貼り付けていく
- ⑧ 04年の参院選大阪選挙区のデータも同様に

ここまで済んだ状態：[osaka-01.xlsx](#)

延長戦@大阪府 3

- ⑨ 3種のデータを付き合わせるための、作業用シートを作りたい。
「form」フォルダの「form.xlsx」を開き、シートをコピーする。
シート名のタブをつかみ、Ctrl + ドラッグすると早い
- ⑩ この練習で必要なのは大阪府の分だけ。都道府県の列で「並び替えとフィルター」を使う。漏斗のマークを押し、下の窓で「大阪府」と「都道府県」だけに✓を入れると、絞り込まれる。左端の行番号が青くなるのが、フィルターが効いている目印
- ⑪ Ctrl + A で全範囲指定。Ctrl + C でコピー
- ⑫ 新しいシートに、Ctrl + V で貼り付ける

延長戦@大阪府 4

- ⑬ このシートを、先ほどのExcelのブック（参院選と府議選のデータを貼り付けたもの）に、ドラッグする。シートがそのままコピーされた

今回はこれでOKだが、別シートを参照する数式が入っていた場合は、ほかのブックに移すと、正しく動かなくなる。その場合は「値だけにして貼り付け直す」作業をして数式を消しておくこと

Ctrl + A （全範囲選択）

Alt + Ctrl + V（オプションを指定する貼り付け）

V （数式を消し、値のみのテキストに）

ここまで済んだ状態：[osaka-02.xlsx](#)

延長戦@大阪府 5

- ⑮ 自治体名をキーに、参院選2回分と府議選のデータを突合する。
VLOOKUPを使う。並び順がまちまちだから、コピー&ペーストでは無理
- ⑯ キー列は一意か？ 大阪市と堺市の両方に「西区」がある。一意かどうかは、「条件付き書式」でチェック可能
- ⑰ 投票率の空欄は、府議選が無投票だったから、で間違いないか
- ⑱ 欄外に無投票の選挙区が載っているので、確かめておく
- ⑲ 空白の有無、「ケ」の大小の違い（他県ではよくある）も、マッチしない原因。SUBSTITUTEやTRIMを活用する

延長戦@大阪府 6 完

- ②⑩ データが揃ったら、シートを複写して、値だけを張り直す。数式の入ったシートも残しておくのと、手直しや点検がしやすい
- ②⑪ 層別散布図を描いてみる。府議選の有無で分けるだけなので、2グループでいい。X軸とY軸に何を選ぶか？

一例です：[osaka-04complete.xlsx](#)

で、選挙疲れなのか

- 亥年の参院選投票率が地方選の影響で下振れする傾向は、07年時点では少しある
- ただし、必ず前回より投票率が下がるとは言い切れない。もっと伸びるはずが、伸び切らなかった、という程度の下振れも含む
- 統一地方選とは別日程で行われる知事選の影響もうけている
- それが選挙「疲れ」によるものなのか、疲れるのは誰なのか、は現場を取材しない限り分からない。有権者が地方選で「選挙に満足」してしまうのだったり、地方選で投票できないと参院選でハッスルする「選挙へのあこがれ」があるのだとしても、同じ現象は起きる
- 今年の地方選と今夏の参院選のデータで調べてみたい。県議選も知事選も、統一選からの離脱が3ずつ増えた

お疲れさまでした！

受講ありがとうございます

- ・ 質問はみんなの宝物です。ご遠慮なく質問を。分からないのは、説明が行き届かないから。たぶん、ほかの方も困っています
- ・ 内容改善につなげるために、ぜひフィードバックをお願いします
- ・ そのほか、お問い合わせは
yoshito.nishio+JNPC@gmail.com へ

ご協力いただいたアドバイザーのみなさま、いつも助けていただいている事務局のみなさま、そして受講いただいたみなさまに深謝します。