

## Interstenoランキング分析マクロ 仕様書

### 【基本仕様】

- ・INPUT: Interstenoサイトから取得したランキングデータをExcelシートに貼り付けたファイル
- ・OUTPUT: 以下を含むファイル
  - ・言語別ランキング
    - 各参加者の各言語の偏差値を含む(0点の人を除外した場合と除外しない場合の両方)
  - ・各種統計(以下を含む)
    - 縦軸に参加者、横軸に各言語の得点を記載した一覧表
    - 各参加者の受験言語数、合格言語数、合格率
    - 各言語の受験者数、合格者数、合格率、単純平均点、合格者平均点
    - 日本人に限定した各種統計
  - ・得点の代わりにミスを集計したランキング、各種統計
- ・INPUT、プログラム、OUTPUTはすべて別ファイルとし、いずれかのファイルの肥大化を防ぐ。

### 【元データ】

- ・以下のファイルを、本マクロと同じフォルダ内に設置する。

引数	ファイル名	備考
22	IS2022_Rawdata.xlsx	2022年のIntersteno公開ランキング
21	IS2021_Rawdata.xlsx	2021年のIntersteno公開ランキング
20	IS2020_Rawdata.xlsx	2020年のIntersteno公開ランキング
19	IS2019_Rawdata.xlsx	2019年のIntersteno公開ランキング
18	IS2018_Rawdata.xlsx	2018年のIntersteno公開ランキング
17	IS2017_Rawdata.xlsx	2017年のIntersteno公開ランキング
16	IS2016_Rawdata.xlsx	2016年のIntersteno公開ランキング
15	IS2015_Rawdata.xlsx	2015年のIntersteno公開ランキング

※2006-2014年も同様。

- ・Rawdataシートに、Interstenoサイトからデータを取得する。  
Classification Listsから「言語: 英語」「年齢: 全部」「type: general」を選択し、  
表示されたデータをExcelにコピー&ペーストする。  
このデータは以下のように格納されている。
- ・先頭11行はタイトル行。(注: Win10では先頭10行。よって手動で1行追加する)
- ・1~10行目は集計・分析に使用しないため省略。
- ・11行目は「Place」「Surname」「Town/Nation」「School/Association」  
「Strokes per minute」「Total strokes」「Errors」「error %」「Points」「Language」  
「Software」「Technology」を記述。
- ・12行目以降はデータ。  
但し各参加者につき、以下の順に格納されている。
  - ・1行目(総得点行)には、「Place」「Surname」「Town/Nation」「School/Association」  
「合計points」「Technology」が記載されている。他の項目は「-」表示。
  - ・2行目以降(言語別得点行)には、「Strokes per minute」「Total strokes」「Errors」  
「Error %」「各種目のPoints」「Language」「Software」「Technology」が、  
Pointsの降順に記載されている。  
1~4列目はセルが結合され、何も記載されていない。  
(注: Win10ではnon-breaking spaceが格納されている)
- ・1種目のみに参加した参加者のデータは、1行のみ記載されている。総得点行は記載されない。

## 【ボタン/マクロ】

引数	ボタン名	備考
-	IS_Ranking1	ランキング処理1(ranking1マクロ)
-	IS_Ranking2	ランキング処理2(ranking2マクロ)

## 【ranking1マクロ】

・以下の手順で、言語別ランキングを含むファイルを作成する。

- (1) 元データファイルを開く。ファイルはユーザが選択できる。
- (2) 元データファイルの行数と列数を数え、変数に格納する。
- (3) 元データファイルから必要なデータを配列に格納する。

No.	表記	概要
1	Place	順位
2	Surname	名前
3	Town	都市(元データのTown/Nationから分離)
4	Nation	国(元データのTown/Nationから分離)
5	School/Association	学校/団体
6	SPM	打鍵速度(文字/分)
7	Total strokes	10分間の総打鍵数
8	Error	10分間の総ミス数
9	error %	ミス率
10	Points	得点
11	Language	言語
12	Software	使用ソフトウェア(TAKI/ZAV/JAVA)
13	Technology	使用テクノロジー(KB/CK)

- (a) No.1,2には元データファイルの1,2列目のデータを格納する。  
但し、言語別得点行(元データファイルに何も記載されていない、あるいはnon-breaking spaceが存在する)の場合は、出力ファイルの1行上からコピーする。
  - (b) No.3,4には元データファイルの3列目の「Town/Nation」を「Town」と「Nation」に分離して格納する。  
但し、言語別得点行(元データファイルに何も記載されていない、あるいはnon-breaking spaceが存在する)の場合は、出力ファイルの1行上からコピーする。
  - (c) No.5には元データファイルの4列目のデータを格納する。  
但し、言語別得点行(元データファイルに何も記載されていない、あるいはnon-breaking spaceが存在する)の場合は、出力ファイルの1行上からコピーする。
  - (d) No.6以降には、元データファイルの1つ前の列から格納する。
- (4) 元データファイルを保存せずに閉じる。
  - (5) 出力先ファイルを開く。Excelの設定で、新規ファイルのシート数が1となっていることを仮定する。
  - (6) 出力先ファイルのシート名は、「Ranking1」とする。
  - (7) タイトル行を記述する。
  - (8) データ行を記述する。フォントやフォントサイズ、列の幅の設定はforループの外でまとめて実施する。

- (9) 全データを「言語の昇順」「得点の降順」の優先順位でソートする。
- (10) 言語別シートを追加する。  
全データの11列目(言語)でフィルタして、重複を排除して言語名を取得する。  
独仏蘭**伊日克**に関しては、統合ランキングのシートを追加する。
- (11) 言語別シートのタイトル行を記述する。“Ranking1”シートと同様にする。
- (12) 言語別シートのデータ行を記述する。  
(a) “Ranking1”シートの言語欄でフィルタし、シート名と一致するデータをまとめてコピーする。  
(b) 言語列は不要なので削除する。但し独仏蘭**伊日克**に関しては方言列として残す。

- ・以下の手順で、各言語・各人の偏差値を計算する。(CalcDev関数)
- (13) 独仏蘭**伊日克**の統合ランキングのみ、得点の降順でソートする。
- (14) 平均と標準偏差を計算するため、以下の処理を実施する。  
(a) Sum関数を用いて、変数meanPointに、各データの合計値を格納する。  
(b) 変数stdevPointに、各データの得点の2乗を順次加算する。  
(c) CountIf関数を用いて、変数nLanguageRow2に、0点を除くデータ数を格納する。
- (15) 各言語の平均と標準偏差を計算する。  
(a) 平均=各言語の総得点÷各言語のデータ数  
(b) 標準偏差=(各言語の得点の2乗の平均－各言語の得点の平均の2乗)の平方根  
(c) 平均2=各言語の総得点÷各言語のデータ数(但し得点が0のデータは除外)  
(b) 標準偏差2=(各言語の得点の2乗の平均－各言語の得点の平均の2乗)の平方根  
(但し得点が0のデータは除外)
- (16) 各言語につき、各参加者の偏差値を計算し、表示する。  
(a) 偏差値=50+10×(得点－平均)÷標準偏差  
(b) 偏差値2=50+10×(得点－平均2)÷標準偏差2 (但し得点が0の時は偏差値も0)  
標準偏差が0の場合、偏差値=50とする。標準偏差2=0の場合、偏差値2=50とする。  
⇒得点が1以上で偏差値がマイナスになるケースも考えられる。その場合も0とすべきか。

- (17) “Ranking1”シートのフィルタを解除する。
- (18) 各シートのウィンドウ枠を固定する。(A2セルをアクティブにした状態で行う)
- (19) 出力先ファイルをカレントフォルダに保存し、閉じる。

No.	ファイル名
22	IS2022_ranking1.xls
21	IS2021_ranking1.xls
20	IS2020_ranking1.xls
19	IS2019_ranking1.xls
18	IS2018_ranking1.xls
17	IS2017_ranking1.xls
16	IS2016_ranking1.xls
15	IS2015_ranking1.xls

※2006-2014年も同様。

- (20) 終了時に画面上にメッセージを表示する。

## 【ranking2マクロ】

・前処理を実施する。

- (1) 元データファイルとして、ISYYYY\_Ranking1.xlsを開く。ファイルはユーザが選択可能である。
- (2) 元データファイルのRanking1シートのデータの行数と列数を数え、変数に格納する。  
但し、総得点行は除外する。
- (3) 全データを「順位」の昇順でソートする。
- (4) 元データファイルから必要なデータ(順位、名前、国、ミス数、得点、言語)を配列に格納する。
- (5) 元データファイルを保存せずに閉じる。
- (6) 出力先ファイルを開く。Excelの設定で、新規ファイルのシート数が1となっていることを仮定する。
- (7) 出力先ファイルのシート名は、“Points”とする。

・以下の手順で、縦軸に参加者、横軸に言語を配置した一覧表を作成する。(WriteData2関数)

- (1) タイトル行を記述する。

列番号	表記	概要
1	Place	順位
2	Surname	名前
3	Nation	国
4	Points	総合得点
5~21	CZ~TR	言語別得点(方言は一つにまとめる)
22	受験数	参加した言語数(失格した言語も含む)
23	合格数	合格した言語数
24	合格言語平均	合格した言語の平均得点

- (2) データ行を記述する。

- (a) 1,2,3列目は配列からコピーする。  
但し、1列目を前の行と比較して同じ場合は、コピーしない。
  - (b) 5~21列目は元データファイルの言語別得点行の10列目からコピーする。  
出力先ファイルのタイトル行の表記と、配列内の“Language”の表記を比較し、同じ場合に、配列から得点をコピーする。但し、方言は一つにまとめる。  
参加していない場合は“-”を表示する。
  - (c) 4列目には、(b)の処理中に計算した総得点を記載する。
  - (d) 22列目には、5~21列目のうち“-”以外の値が記載されている列の数を数えて表示する。
  - (e) 23列目には、5~21列目のうち1以上の値が記載されている列の数を数えて表示する。
  - (f) 24列目には、4列目の値を23列目の値を割った値を表示する。  
但し、合格した言語が0の場合は、“-”を表示する。
  - (g) フォント設定、小数点以下の桁設定は、forループの外でまとめて実施する。
- (3) 各言語のデータに、各言語TOP10に入っている場合は着色する。

・以下の手順で、統計結果を作成する。(WriteStats関数)

- (1) 新たにシートを追加する。名前は“PointStats”とする。
- (2) タイトル行を記述する。

列番号	表記	概要
1	項目	統計項目
2	Points	総合得点
3~19	CZ~TR	言語別得点(方言は一つにまとめる)
20	受験数	参加した言語数(失格した言語も含む)
21	合格数	合格した言語数
22	合格言語平均	合格した言語の平均得点

- (3) データ行を記述する。
- (a) 1列目には、上から順に「単純平均」「合格者平均」「受験者数」「合格者数」「合格率」を記述する。  
1行空けて、「TOP100」「単純平均」「合格者平均」を記述する。  
さらに1行空けて、「TOP200」「単純平均」「合格者平均」を記述する。
  - (b) 2列目には、「総合得点の平均値」「全受験者数」を記載する。他の項目には“-”を記載する。
  - (c) 3～19列目には、各言語の「単純平均」「合格者平均」「受験者数」「合格者数」「合格率」を計算し、記載する。但しTOP100に関しては総合順位1～100位の参加者のみを対象とする。  
TOP200に関しては総合順位1～200位の参加者のみを対象とする。
  - (d) 20～22列目には、「受験言語数の単純平均」「合格言語数の単純平均」「合格言語平均の単純平均」のみを記載する。他の項目には“-”を記載する。
  - (e) フォント設定、小数点以下の桁設定は、forループの外でまとめて実施する。

・以下の手順で、縦軸に参加者(日本人限定)、横軸に言語を配置した一覧表を作成する。(WriteData2\_JP関数)

- (1) 新たにシートを追加する。名前は“Points\_JP”とする。
- (2) 元データとして、Pointsシートを用いる。データの列数を数え、変数に格納する。
- (3) タイトル行を記述する。Pointsシートからコピーする。
- (4) データ行を記述する。
  - (a) フィルタ機能を用いて、Nationが“JP”のデータのみ、Pointsシートからコピーする。
  - (b) 言語列は自明なので削除する。
  - (c) フォント設定、小数点以下の桁設定は、forループの外でまとめて実施する。

・以下の手順で、統計結果(日本人限定)を作成する。(WriteStats\_JP関数)

- (1) 新たにシートを追加する。名前は“PointStats\_JP”とする。
- (2) タイトル行を記述する。PointStatsシートからコピーする。
- (3) データ行を記述する。
  - (a) 1列目には、上から順に「単純平均」「合格者平均」「受験者数」「合格者数」「合格率」を記述する。
  - (b) 2列目には、「総合得点の平均値」「全受験者数」を記載する。他の項目には“-”を記載する。
  - (c) 3～19列目には、各言語の「単純平均」「合格者平均」「受験者数」「合格者数」「合格率」を計算し、記載する。
  - (d) 20～22列目には、「受験言語数の単純平均」「合格言語数の単純平均」「合格言語平均の単純平均」のみを記載する。他の項目には“-”を記載する。
  - (e) フォント設定、小数点以下の桁設定は、forループの外でまとめて実施する。

・上記の処理を「得点」の代わりに「ミス数」を用いて繰り返す。注意点は以下の通り。

- (1) シート名には“Point”の代わりに“Error”を用いる。
- (2) 失格となった場合を判定するため、Pointsシートの同じ位置のデータを読み込む必要がある。  
得点が0であれば失格、“-”であれば未受験と判定する。  
このため、事前にPoints、Points\_JPシートの値を配列に格納する。
- (3) ミス数だけで比較すると、例えば「3000文字1ミス」と「6000文字ミス1」が同列に扱われるため、不公平である。  
そこで、ミス率の欄を設ける。

・後処理を実施する。

- (1) “Points”、“Errors”シートのフィルタを解除する。
- (2) 各シートのウィンドウ枠を固定する。(A2セルをアクティブにした状態で行う)
- (3) 出力先ファイルをカレントフォルダのISYYYY\_Ranking2.xlsに保存し、閉じる。
- (4) 終了時に画面上にメッセージを表示する。

### 【今後の仕様変更/バグ修正案】

- (1) 各言語で算出した「偏差値1」「偏差値2」を合計して総得点とした場合に、ランキングがどう変わるか見極めたい
- (2) 21歳以上の参加者の合格基準である「240cpm以上、ミス率0.5%以内」に限定した統計を取りたい