

# Archaeological Geometric Morphometrics and R / 考古学における幾何形体測定学と R

As part of the #StayHomeButStudy Workshop Series

Dr Christian Steven Hoggard (University of Southampton, United Kingdom)

## Introduction

This guide provides a “hands-on” step-by-step introduction into the application of geometric morphometric (GMM) methodologies in archaeological science, as conducted the R Environment. Using a published dataset this workflow will guide the reader through four key GMM procedures: 1) data importing (and creation), 2) data transformation, 3) data analysis and 4) data visualisation. A Japanese translation of this documentation is also provided.

これは、R 環境下における考古学のための幾何形態測定学（GMM）の方法を、ステップ・バイ・ステップで学ぶハンズオン・ワークショップの実習ガイドです。このガイドは、公表されているデータセットを利用して、読者に GMM の 4 つの重要な手順：1) データ読み込み（作成）、2) データ変換、3) 分析、4) 可視化を紹介します。

I will first demonstrate the actions or functions on Zoom (using this markdown document) and then allow time for all participants to run the function (3-5 minutes per function). To run a ‘chunk’, a shaded area of function we can press the “Run selected chunk” button, represented by a play button or use the shortcut `ctrl + enter`. Should there be any queries then please let us know in the Slack workspace. Conversely, when you complete a function please could you use a “thumbs up” emoji on Slack. We are allowing time between functions to ensure that all participants keep up, if you finish a particular process early explore the functions in the packages through the ‘Help’ tab in the ‘Packages’ window.

ワークショップでは、この R マークダウン資料を使用して、最初に Zoom でコードの動作・関数を実演し、その後で 3~5 分ほどみなさんにも実行してもらいます。マークダウンの中で灰色の背景の範囲が実行可能なコードの「チャンク」です。チャンクの右上の再生ボタンをクリックするか、またはショートカット `ctrl + enter` で実行できます。疑問・質問があると思うので、その際は Slack のこのワークショップのチャンネルに書き込んでください。問題なく実行できた人は、Zoom で「いいね」のリアクションをしてください。参加者全員が、コードの実行を確認する時間があります。慌てないでください。また、早く確認できた人は、RStudio の右下の Package ペインのパッケージ・ライブラリ名をクリックすると関数のヘルプを見ることができるので、それらをご覧ください。

This practical constitutes the second workshop of the #StayHomeButStudy event, organised by Dr. Atsushi Noguchi, and is tailored for Japanese archaeologists, researchers and enthusiasts.

この実習は、野口 淳, クリスチャン・ホガード, ベン・マーウィックが主催する #StayHomebutStudy オンライン・ワークショップの第2回を構成するものであり、日本の考古学者、研究者、愛好家のために準備されたものである。

### ### About the Code, Packages and Data / コード・パッケージ・データについて

The data used throughout this guide originates from Ivanovaité et al (2020): “*All these Fantastic Cultures? Research History and Regionalization in the Late Palaeolithic Tanged Point Cultures of Eastern Europe*”, published in open-access in the European Journal of Archaeology (<https://doi.org/10.1017/ea.2019.59>). The data can be found on a GitHub repository (<https://github.com/CSHoggard/-Eastern-Europe-Tanged-Points>), in addition to the Open Science Framework (<https://osf.io/agrwb/>).

このワークショップで使用されるデータは、Ivanovaité et al (2020) 「これらの考古学的文化はなにか? 東ヨーロッパ後期旧石器時代の剥片尖頭器文化をめぐる研究史と地域性」『ヨーロッパ考古学』(オープンアクセス) で公表されたものです (<https://doi.org/10.1017/ea.2019.59>). 同論文のオリジナルデータは GitHub リポジトリ, (<https://github.com/CSHoggard/-Eastern-Europe-Tanged-Points>) OSF (<https://osf.io/agrwb/>) /

All code, and data, including the markdown document (in HTML and PDF format) for this practical can be found on GitHub ([https://github.com/CSHoggard/-japanworkshop2020tree/master/workshop\\_2](https://github.com/CSHoggard/-japanworkshop2020tree/master/workshop_2)).

この実習で用いるすべてのコード, データ, マークダウン文書 (HTML と PDF フォーマットを含む) は, GitHub 上のワークショップのリポジトリで公開されています。

The GMM procedure detailed below is grounded on two-dimensional outline analysis. In conducting outline analysis for this practical the following two packages are necessary:

ここで実行する GMM の方法は, 2D のアウトライン分析です. 下記の2つのパッケージ (および付随してインポートされるパッケージ群) を必要とします.

- **Momocs** (Version 1.3.0) <https://cran.r-project.org/web/packages/Momocs/index.html>
- **tidyverse** (Version 1.3.0) <https://cran.r-project.org/web/packages/tidyverse/index.html> ##  
**Software Installation / ソフトウェアのインストール**

Following the installation of R and RStudio, we can now install the required packages:

R および RStudio のインストールに続き、必要なパッケージをインストールします：

```
if(!require("Momocs")) install.packages('Momocs', repos='http://cran.us.r-project.org') # Momocs のラ
if(!require("tidyverse")) install.packages('tidyverse', repos='http://cran.us.r-project.org') #tidy
if(!require("rio")) install.packages('rio', repos='http://cran.us.r-project.org') #rio のチェックとイ
```

As the tidyverse and Momocs packages may take time to install given the size of the files *please ensure that these are downloaded prior the workshop.*

tidyverse と Momocs パッケージの読み込みは、ファイルサイズの大きさから多少時間がかかるかも知れません。できる限りワークショップの開始前にダウンロードしてください。

To bring the data (from Github) into R/Rstudio we can use the `import` function from the `rio` package, and extract the data from the repository:

R/RStudio にデータをダウンロードする際には `rio` パッケージの `import` ファンクション（関数）を使用して、GitHub リポジトリからデータを抽出することができます：

```
database <- rio::import("https://github.com/CSHoggard/-workshopjapan2020/raw/master/workshop_2/databa
tpsdata <- rio::import("https://github.com/CSHoggard/-workshopjapan2020/raw/master/workshop_2/tpsline")
```

Once installed, the packages can be activated through the `library()` function:

一度インストールされたパッケージは、`library()` 関数によってアクティベートされます：

```
library(Momocs) #Momocs のアクティベート
library(tidyverse) #tidyverse のアクティベート
```

## About the Data / データについて

This data was composed to assess the robustness of cultural taxonomies in the Final Palaeolithic period of Eastern Europe, as portrayed through tanged point variants. It consists of 250 tanged point outlines and produced in the TPS Suite (<https://life.bio.sunysb.edu/morph/soft-dataacq.html>), using the **outline object** function. As this dataset was produced in the TPS Suite the file format is **.tps**. In their composition these outlines are semilandmarks, an algorithm-produced series of equidistant points are each shape. A database of all examples and their respective cultural assignment is also provided.

このデータは、東ヨーロッパの晩期旧石器時代における考古学的文化の分類のロバストネス（頑健性）を、剥片尖頭器のバリエーションから評価するために作成されました。250 点の剥片尖頭器のアウトラインデータで、TPS Suite(<https://life.bio.sunysb.edu/morph/soft-dataacq.html>) \*outline object\* 関数を使用して作成しました。ファイルフォーマットは **.tps** です。アウトラインは、アルゴリズムによって生成された等間隔の点群セミランドマークとして構成されています。すべての資料の詳細、その考古学的文化の帰属もデータベース化されています。

### ## Importing GMM Data: Alternative Approaches

There are a number of ways which landmark and outline morphometric data can be imported into the R Environment. Here, for ease and replicability, the outline data (in **.tps** format) was stored on a GitHub repository and directly fed into the R environment, utilising the `Momocs::import_tps()`

function in a .rds file. Other ways to import .tps data (if saved locally) include the above function, `geomorph::readland.tps()` and rewriting tools in Momocs e.g. `Momocs::rw_rule()`. Data from stereomorph can also be imported through the `Momocs::import_StereoMorph_ldk()` and `Momocs::import_StereoMorph_curve()` functions.

Note: for the purpose of this workshop I will detail in-text the function and its constituent package e.g. `geomorph::readland.tps()`, however only the function is what will be ‘used’ so-to-speak e.g. `readland.tps()`. This helps you to understand what packages the functions originate from.

Within Momocs, outlines can be extracted from silhouette data through the `Momocs::import_jpg()` and `Momocs::import_jpg1()` functions. See their respective helpfiles for more details on these functions. These will be demonstrated at the end of the workshop.

## Examining the Data / データの検証

With our data now in the R Environment we can now call our tpsdata object through the `base::View` functions. The `base::View()` function will highlight the three constituent parts of the tps file: the 1) *Coo* (coordinate data), 2) *cur* (the curve data if necessary), and 3) *scale* (the scale data if present). It is the *Coo* and scale data which we will take forward, with the database, to examine shape variation among our tanged points. It is best not to call the tps data as R will stream all coordinate data for each example. We can also inspect our database using the `head()` function, and examine the different components of our dataset.

R に読み込んだ tps データは、`base::View` 関数で呼び出し、表示できます。`base::View()` 関数は、tps の 3 種類の要素：1) *Coo* (座標データ), 2) *cur* (曲線：必要な場合), 3) *scale* (縮尺：記録されている場合) を表示します。今回は、*Coo* と *scale* データを、データベースとともに、剥片尖頭器の「かたち」の多様性を検証するために使用します。ただし tps の、たとえば座標データをすべて呼び出す必要はありません。また、`head()` 関数を使ってデータベースを確認し、データセットの異なる部分を検証することもできます。

```
head(database) #database の先頭数行を呼び出す
```

```
## # A tibble: 6 x 10
##       ID Site Context Longitude Latitude Country Archaeological_~ File_Name
##   <dbl> <chr> <chr>      <dbl>    <dbl> <chr>    <chr>             <chr>
## 1     1 Balt~ Baltaš~    24.0     54.0 Lithua~ Baltic Magdalen~ Bal.1
## 2     2 Baro~ Barouka    30.3     53.5 Belarus Grensk        Bor.1
## 3     3 Baro~ Barouka    30.3     53.5 Belarus Grensk        Bor.3
## 4     4 Baro~ Barouka    30.3     53.5 Belarus Grensk        Bor.4
## 5     5 Baro~ Barouka    30.3     53.5 Belarus Grensk        Bor.5
## 6     6 Baro~ Barouka    30.3     53.5 Belarus Grensk        Bor.6
## # ... with 2 more variables: Reference <chr>, Notes <chr>
```

We can observe that the group data we want to examine (Archaeological\_Unit) is <chr>, that is to say of type ‘character’, and not <fctr> (‘factor’), as required for our analysis. This can be corrected through the `base::as_factor()` function and confirmed through the `base::is_factor()` function:

表示されたデータベースでは、検証したい‘Archaeological\_Unit(考古学的単位(文化))’のデータ型がと表示されています。はデータが文字列型であることを示し、分析対象となるつまり因子型ではありません。そこで `base::as_factor()` 関数で因子型に修正し、`base::is_factor()` 関数で確認することができます(結果が [1]TRUE と返されれば OK)。

```
database$Archaeological_Unit <- as.factor(database$Archaeological_Unit) #Archaeological_Unitを因子型
is.factor(database$Archaeological_Unit) # check to see the data is now of type 'character' / データが文
## [1] TRUE
```

We can also inspect the number of different archaeological units within our dataset through the `base::summary()` function. This highlights the number of tanged points in each group. With certain taxonomic units rarely used this is reflected in the low sample sizes for certain groups e.g. Vyshegorian.

`base::summary()` を使い、考古学的文化ごとに含まれる資料数を確認できます。これは、考古学的に文化ごとの、データベースに含まれる剥片尖頭器の数です。たとえば Vyshegorian 文化のサンプル数が少ないことから分かる通り、剥片尖頭器は稀な型式です。

```
summary(database$Archaeological_Unit) #Archaeological_Unitを集計
```

```
##      Baltic Magdalenian Bromme (Eastern Europe) Bromme (Western Europe)
##                36                      9                      49
##      Grensk      Krasnosillya      Perstunian
##                55                29                4
##      Pitted Ware (Type A)      Podolian      Vyshegorian
##                24                14                8
##      Wolkushian
##                22
```

## GMM Procedure 1: Outline File Creation / GMM の実行 1：アウトライン・ファイルの作成

Central to Momocs are a specific suite of shape classes for: 1) *outlines* (OutCoo), *open outlines* (OpnCoo) and *landmarks* (LdkCoo), with often one class specific to your own dataset. While some operations in Momocs are generic and do not depend on one of these classes, many functions require your data to be one of these specific ‘S3 objects’. In this instance our tps data is comprised of outlines, and so we wish for our data to be OutCoo, as to enable efourier (elliptic Fourier) analyses. Other analyses including rfourier

(radii Fourier) or tfourier (tangent angle Fourier) analyses can be conducted through this process but for this workshop we're only going consider elliptic Fourier analysis (EFA).

Momocs パッケージの中心となるのは、特定の「かたち」のクラスです。それらは、(閉じた) アウトライン (OutCoo), 閉じていない (開いた) アウトライン (OpenCoo) と ランドマーク (LdkCoo) で、多くの場合、データセットはひとつのクラスで構成されています。Momocs の機能のいくつかは、(ジェネリックな関数なので) 「かたち」データのクラスに依拠せず実行できますが、多くの関数は「S3 オブジェクト」として特定のクラスに対して実行することになります。ここでは、tps データはアウトラインなので、楕円フーリエ (efourier) 分析を行なうための OutCoo として扱います。そのほかの半径フーリエ (rfourier) 分析や、タンジェント・フーリエ (tfourier) 分析も同じ手順で実施できますが、ここでは楕円フーリエ分析 (EFA) のみを取り扱います。Through this lens, the coordinate data (coo) must therefore be turned into outline data through the Momocs::Out() function for the workflow to work. Once performed, we can then enter the object (here titled 'shape') and examine its properties.

この「レンズ」を通して、今回のワークフローに即して Momocs::Out() 関数により座標データ (coo) はアウトラインデータに変換されます。これを実行すると、'shape' と命名したオブジェクトにアウトラインデータが格納され、その属性 (プロパティ) を分析できるようになります。

```
shape <- Out(tpsdata$coo, fac = database) # incorporating our database as our factors / tps 座標データ
shape # call the object / オブジェクトの呼び出し
```

```
## Out (outlines)
##   - 250 outlines, 1543 +/- 1370 coords (in $coo)
##   - 10 classifiers (in $fac):
## # A tibble: 250 x 10
##       ID Site Context Longitude Latitude Country Archaeological_~ File_Name
##   <dbl> <chr> <chr>         <dbl>    <dbl> <chr>    <fct>             <chr>
## 1     1 Balt~ Baltaš~      24.0     54.0 Lithua~ Baltic Magdalen~ Bal.1
## 2     2 Baro~ Barouka    30.3     53.5 Belarus Grensk           Bor.1
## 3     3 Baro~ Barouka    30.3     53.5 Belarus Grensk           Bor.3
## 4     4 Baro~ Barouka    30.3     53.5 Belarus Grensk           Bor.4
## 5     5 Baro~ Barouka    30.3     53.5 Belarus Grensk           Bor.5
## 6     6 Baro~ Barouka    30.3     53.5 Belarus Grensk           Bor.6
## # ... with 244 more rows, and 2 more variables: Reference <chr>, Notes <chr>
##   - also: $ldk
```

This tells us that in our Out file there are a total of 250 outlines, with a mean number of 1543 landmarks and 10 different factors (longitude, Latitude, Archaeological\_Unit, etc.). We are only going consider Archaeological\_Unit within these factors.

137 行目のコードで shape オブジェクトを呼び出すと、Out 関数で、平均して 1543 のランドマークと 10 の

異なる分析可能な因子型データ (経度, 緯度, 考古学的文化など) を含む, 合計 250 のアウトラインが取り出され格納されたことが分かります. 今回は, このうち考古学的文化 (Archaeological\_Unit) だけを検討の対象とします.

## ## GMM Procedure 2: Outline Visualisation / GMM の実行 2: アウトラインの可視化

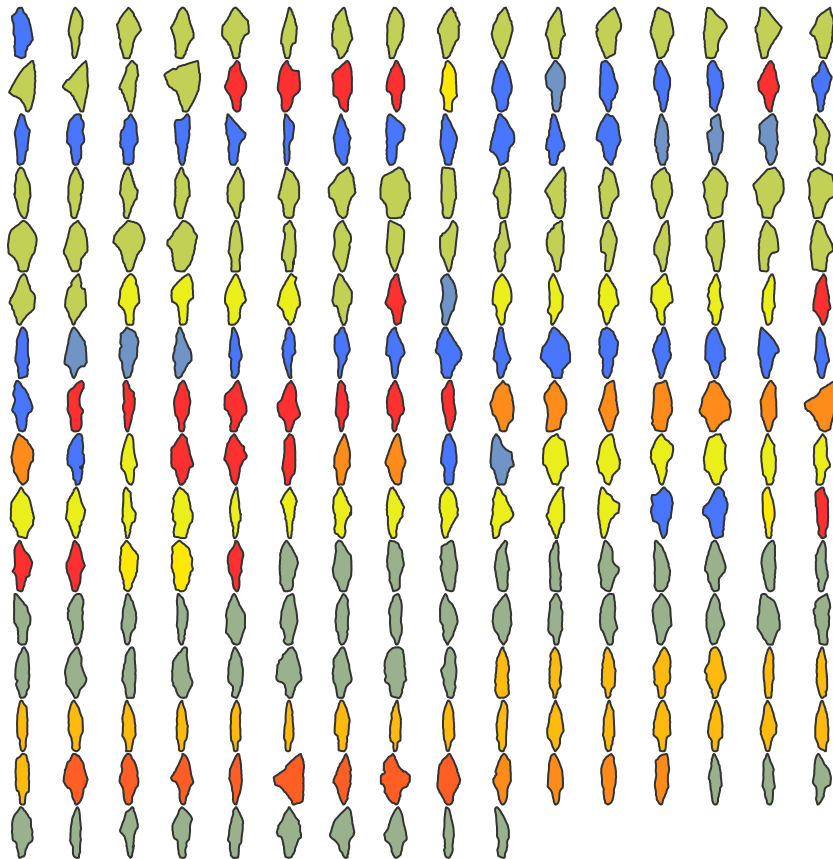
Now our data is in the R environment and in the appropriate class required for Momocs, we can examine the outline shapes. We can first look at all outlines through the `Momocs::panel()` function. Factors can also be coloured in using the `fac` argument.

ここまでで, データが R に取り込まれ, Momocs が要求するクラスが与えられたので, いよいよ「かたち」のアウトラインを検証できます. まずはじめに `Momocs::panel()` 関数で, すべてのアウトラインを見てみましょう. `fac` 引数により因子型データ (ここでは考古学的文化) の種別に色分けすることもできます.

An example using the `Momocs::panel()` function is seen below.

`Momocs::panel()` 関数の使い方の一例を挙げます.

```
panel(shape, main = "", fac = 'Archaeological_Unit') # 「かたち」のデータを, 考古学的文化ごとに色分けして可
```



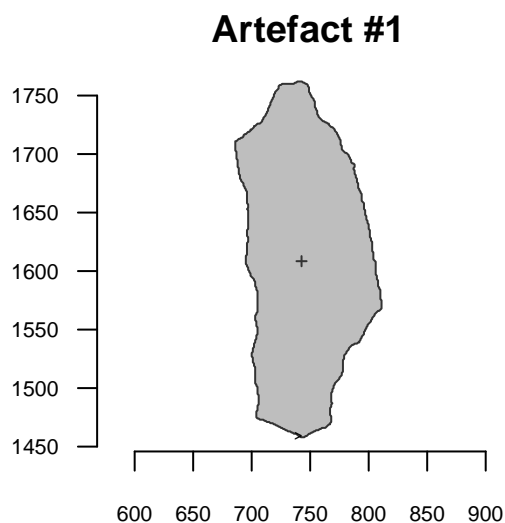
An alternative to the `Momocs::panel()` function is `Momocs::mosaic()`, an updated display function (which will soon replace panel). This does include a legend, unlike the panel function, however the legend drawing options are limited, and are currently being improved for further package versions.

`Momocs::panel()` 関数の代わりに `Momocs::mosaic()` 関数を使用することもできます。これはアップデートされた関数で、`panel()` 関数に取って代わるでしょう。

We can also draw individual shapes of interest using the `Momocs::coo_plot()` function. A number of aesthetic or stylistic changes (including line colour and fill) are possible.

`Momocs::coo_plot()` 関数で、個別資料の「かたち」を描画することもできます。線の色や塗りつぶしなどの表現の変更も可能です。

```
coo_plot(shape[1], col = "grey", main = "Artefact #1") # 資料 1(shape[1]) の「かたち」を灰色に塗りつぶし
```



### GMM Procedure 3: Outline Normalisation / GMM の実行 3: アウトラインの正規化

Normalisation, as stressed by Claude (2008), has long been an issue in the elliptic Fourier process. Normalisation can be performed through the actual elliptic Fourier transformation (using what is known as the “first ellipse”). As we noted in the first workshop, this process (normalisation and elliptic fitting to coefficients) is equivalent to the Procrustes Superimposition for landmark data./

Claude (2008) が強調する通り、正規化は長い間 EFA の課題でした。正規化は、「第 1 楕円」を使用する楕円フーリエ変換により行うことができます。第 1 回ワークショップで指摘したとおり、このプロセス（正規化と係数の楕円フィッティング）は、ランドマークのプロクラステス重ね合わせと同義です。

It is recommended to normalise (standardise) and align your shapes before the `Momocs::efourier()` process. Rotation was considered before outline digitisation, however rotation could also be explored in



Momocs through the `Momocs::coo_aligncalliper()` function. Here we will explore three transformation processes: 1) `Momocs::coo_center()`, 2) `Momocs::coo_scale()` and 3) `Momocs::coo_close()`.

「かたち」の正規化 (標準化) と整列は, Momocs での EFA の実施 (`Momocs::efourier()`) の前に行なうべきです. 向き (回転) は, アウトラインのデジタル化の前に考慮すべきですが, `Momocs::coo_aligncalliper()` 関数で実施することも可能です. ここでは, 以下の 3 つの変換の過程を行ないます. それらは, 1) `Momocs::coo_center()`, 2) `Momocs::coo_scale()`, 3) `Momocs::coo_close()` です.

These three functions perform the following actions:

これらは, 以下のように動作します:

\* `Momocs::coo_center()`: This action centres coordinates on a common origin (common centroid).

共通の基準点 (セントロイド=重心) に揃える \* `Momocs::coo_scale()`: This action scales the coordinates by their 'scale' if provided, or centroid size if 'scale' is not provided.

スケールが与えられている場合はそれに従って, そうでない場合はセントロイド・サイズによって揃えます.

\* `Momocs::coo_close()`: Closes unclosed shapes (precautionary).

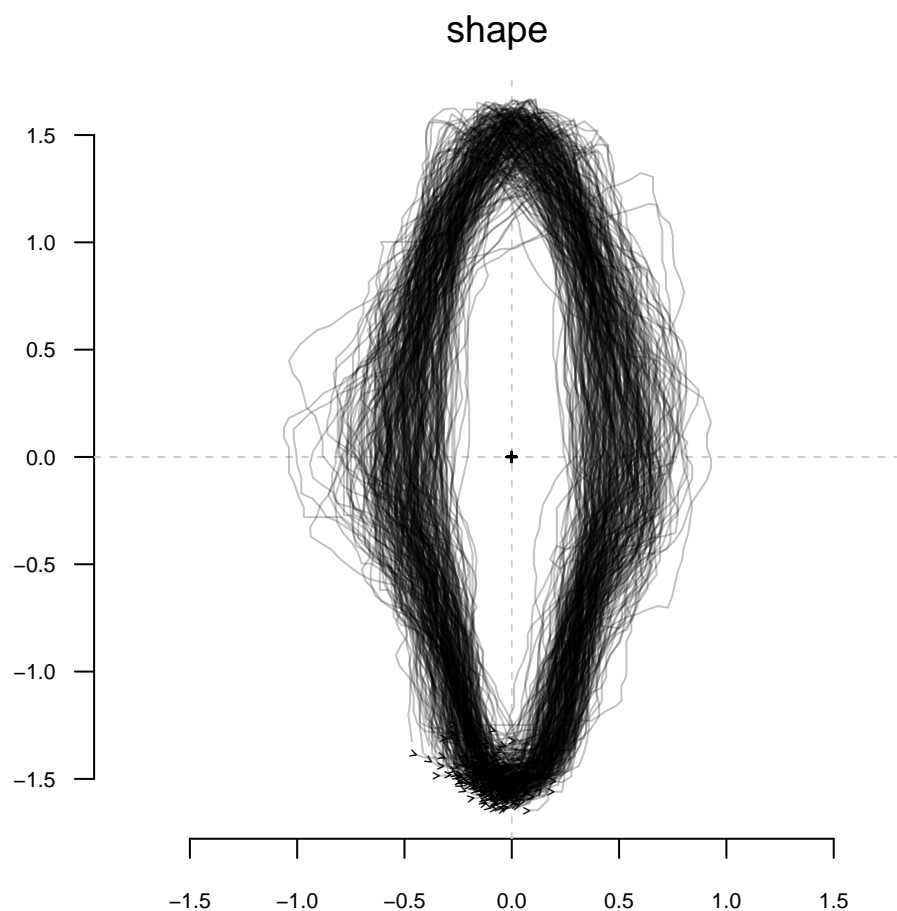
輪郭が閉じていないアウトラインを閉曲線にします (EFA は閉曲線にしか適用できないので, エラーを出さない用心のため)

We can then use the `Momocs::stack()` function to inspect all outlines, now according to a common centroid and of a common scale:

セントロイドとスケールを揃えた結果は, `Momocs::stack()` 関数を使用して確認することができます.

```
shape <- coo_center(shape) # 基準点を揃える
shape <- coo_scale(shape)  # スケールを揃える
shape <- coo_close(shape)  # 開曲線を閉じる

stack(shape, main = "")    # shape の重ね合わせ表示
```



#### GMM Procedure 4: Elliptic Fourier Transformation / GMM の実行 4：楕円フーリエ変換

Elliptic Fourier Analysis (EFA) is one of a number of Fourier based methods of curve composition derived from the first series by Jean Baptiste Joseph Fourier (1768-1830), and developed by Giardina and Kuhl (1977) and Kuhl and Giardina (1982). In practice, a set of four parametric equations (grounded on sine and cosine transformations) are used to define the x and y Cartesian landmarks into curves (Fourier harmonic amplitudes). The coefficients (termed A,B,C and D), when summed together, represent the approximation of artefact form, and are the framework for further analyses. This level of detail depends on the number of harmonics you use. The first harmonic (first ellipse) is responsible for rotation and defines an ellipse in the plane, with which all other harmonics fit onto. The greater the number of harmonics, the greater the level of detail, and the closer the curves resemble the shape. However, a considerable level of statistical noise is produced if there is too much detail (and thus too many harmonics), and so an appropriate level of harmonics are necessary.

EFA は, J.B. ジョセフ. フーリエ (仏: 1766~1830) による「フーリエ変換」の方法のひとつで, Giar-

dina&Kuhl(1977) および Kuhl&Giardina(1982) により発展されました。実際には、サイン・コサイン変換にもとづく4つのパラメトリック方程式の集合により、xy 直交座標系のランドマークを、フーリエ調和振動によって曲線に当てはめます。係数 A,B,C,D を合計が対象資料の形態の近似を表し、さらなる分析の基礎となります。近似の度合いは、採用するハーモニクス (調和) 数に依ります。第1調和 (第1楕円) は、対象資料の向き (回転) に対応した楕円で、より高次の調和がそこに当てはめられます。調和数が増加し、詳細さが増大すると、曲線は「かたち」に近づきます。しかし、調査数が大き過ぎ、詳細さが過剰になると統計的に考慮すべきレベルのノイズが発するので、適当なレベルの調和数が必要となります。

When a level of harmonic power is determined by the researcher (95%, 99%, 99.9%, 99.99%), a series of procedures can be implemented to test how many harmonics are necessary:

95%, 99%, 99.9%, 99.99% というように調和度のレベルを決定したら、以下の手順で必要な調和数をテストできます。

\* `Momocs::calibrate_harmonicpower_efourier()`: This function estimates the number of harmonics required for the elliptic Fourier process (and all other Fourier processes).

楕円フーリエ変換 (およびその他のすべてのフーリエ変換) に必要な調和数を推定します。

\* `Momocs::calibrate_reconstructions_efourier()`: This procedure calculates reconstructed shapes for a series of harmonic numbers. This process best demonstrates the harmonic process.

各調和数により再現される「かたち」を計算します。調和のプロセスをよく表します。

\* `Momocs::calibrate_deviations_efourier()`: This procedure calculates deviations from the original and reconstructed shapes for a series of harmonic numbers. 調和数ごとに、資料のオリジナルの「かたち」と再現された「かたち」の偏差を計算します。

```
calibrate_harmonicpower_efourier(shape, id = 4, nb.h = 20, plot = FALSE) # 資料番号 4 (id=4) の調和数を
```

```
## $gg
## [1] NA
##
## $q
##
##          h1          h2          h3          h4          h5          h6          h7          h8
## 4) Barouka 6.27197 77.4212 80.2774 97.35366 97.45275 98.0372 98.43572 99.14469
##
##          h9          h10          h11          h12          h13          h14          h15
## 4) Barouka 99.25864 99.40648 99.55404 99.68398 99.72722 99.81056 99.89088
##
##          h16          h17          h18 h19
## 4) Barouka 99.92353 99.93879 99.96744 100
##
## $minh
##   90%   95%   99% 99.9%
##    5    5    9   17
```

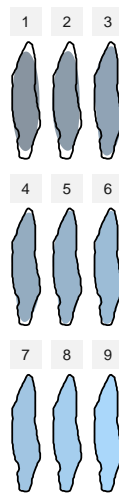
This first procedures highlights how much shape (harmonic power) is represented by the individual harmonics. Here we assessed it on one example and only considered the first twenty harmonics. Typically,

the process is performed on all shapes, however one is used here to detail the components obtained from the function. These three calibrate functions can also take some time to process, please be patient while they load.

ここでは、個別の調和によってどの程度「かたち」および調和度が代表されるかが示されます。1 点の資料について 1~20 の調和数を検討しています。典型的には、この過程は全資料に対して実施しますが、ここでは関数の実行で得られる結果を詳しく見るために、1 点だけ実施しました。これら 3 つの過程は、処理時間が多少かかりますので、完了まで少し待ってください。

```
calibrate_reconstructions_efourier(shape) # shape の復元を計算
```

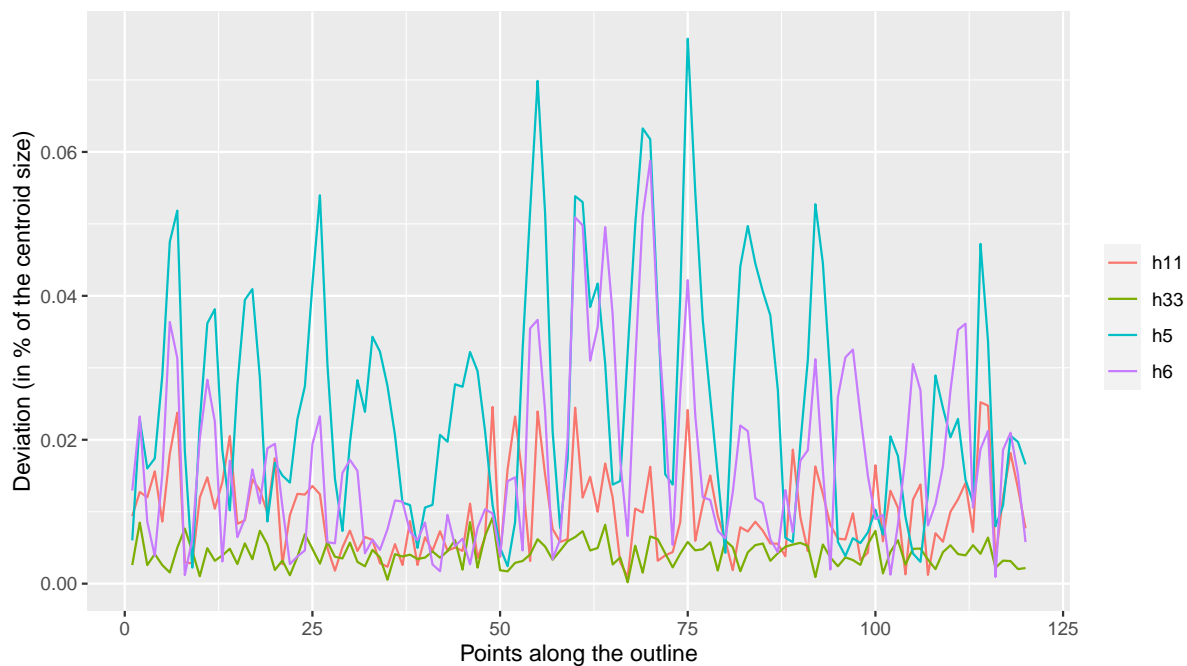
176) Hjarup Mose



This second function best exemplifies the harmonic concept: as the number of harmonics increase, so the approximation of shape is closer to the digitised artefact. This function also highlights the elliptic fitting in the first harmonic.

これは調和 (ハーモニクス) の概念をよく示している関数です。調和数が増加すると、(計算結果による)「かたち」は実際にデジタル化された資料の「かたち」に近似していきます。

```
calibrate_deviations_efourier(shape) # オリジナルと計算結果の「かたち」の偏差を計算
```



The third and final function provides another means of examining the role of harmonic power on deviation in shape.

3つ目の関数は、「かたち」の偏差における調和度の役割を検討する別の手段です。

Once we know how many harmonics are required we can use the `Momocs::efourier()` function to generated out `OutCoe` (outline coefficients) object.

どれだけの調和を利用すべきか分かったら, `Momocs::efourier()` 関数が対象資料の `OutCoe`(アウトライン係数) を生成します。

```
efashape <- efourier(shape, nb.h = 11, smooth.it = 0, norm = TRUE) # shapeのアウトライン係数を第 11 調和まで計算
```

## GMM Procedure 5: Principal Component Analysis (PCA) / GMM の実行 5 : 主成分分析 (PCA)

With our elliptic fourier coefficients we can now begin the exploratory and analytical procedure. We will start by exploring the main theoretical differences in shape through a Principal Component Analysis (PCA). Please refer to the first workshop for a detailed explanation of PCA. We first need to convert out `OutCoe` class object to a PCA class object through the `Momocs::PCA()` function. We can then explore the main sources of shape variation through the `Momocs::PCcontrib()` function.

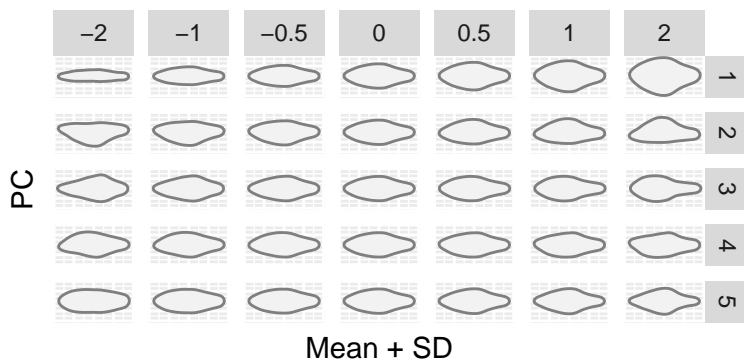
楕円フーリエ係数によって, 探索と分析を開始することができるようになりました. ここでは主成分分析 (PCA) によって「かたち」の主要な理論的差異を検討したいと思います. PCA の詳細な説明については, 第 1 回ワーク

ショップを思い出してください。まず OutCoe(アウトライン係数) クラス・オブジェクトを、Momocs::PCA() 関数によって、PCA クラス・オブジェクトに変換する必要があります。これにより、Momocs::PCcontrib() 関数を使用して、「かたち」の変異の主要要素を探索することができるようになります。

The proportion can also be retrieved through calling the Momocs::Scree() function.

Momocs::Scree() 関数を使用して、(主成分の) 割合を取得することも可能です。

```
pcashape <- PCA(efashape) #EFA の結果を PCA クラスに変換し pcashape に格納
PCcontrib(pcashape, nax = 1:5) #EFA の主成分分析結果を第 1～第 5 主成分軸まで計算表示
```



We can see through this function that Principal Component 1 (PC1), i.e. the main source of shape variation among the tanged points, range from thin tanged points to wider-tanged examples, and that Principal Component 2 (PC2), i.e. the second main source of shape variation, extends from left-exaggerated tangs to right-exaggerated tangs. This function can be set to display as many sources of shape variation as required by the researcher.

この関数の実施により、第 1 主成分軸 (PC1)、つまり剥片尖頭器の「かたち」の変異の主要要素を、細いもの (SD マイナス側) から幅広いもの (同プラス側) へという変化として、また「かたち」の変異の 2 番目に主要な要素としての PC2 を、左に凸から右に凸な基部 (基部) への変化として理解することができます。この関数は、必要な限りいくつでも「かたち」の変異の要素を示すことができます (nax=1:x の x を大きくする)。

While we can observe the main changes in artefact shape, at present we are unsure how much variation these components account for. Using the Momocs::Scree() function we can find out that PC1 accounts for 57.7% of all shape variation, and that the first two axes account for 74.1% (almost three quarters of all shape variation within our dataset). 95% of all shape variation can be accounted for in the first ten principal components (an observation we will come back to afterwards).

このように資料の「かたち」の主要な変化を観察することができますが、この時点では、これらの変異がどれ

だけ資料の組み合わせ全体を説明しているか定かではありません。そこで `Momocs::Scree()` 関数を使用することで、PC1 が全ての「かたち」の変異の 57.7% を説明し、また PC1+PC2 が 74.1%，すなわちデータセット全体の 3/4 を説明することを確認できます。そして PC1~PC10 により、全体の 95% が説明されます。この点については、のちにまた触れます。

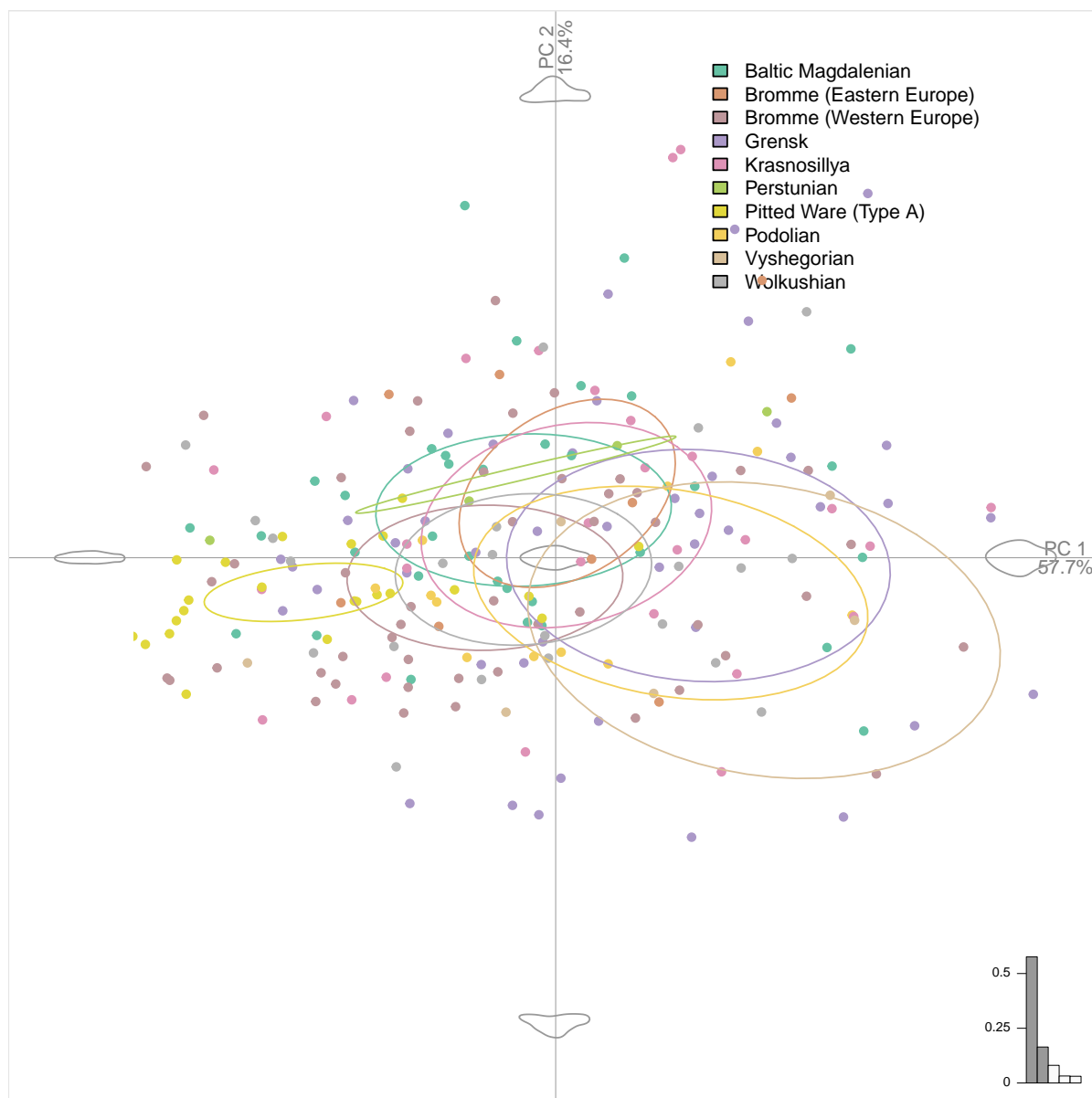
```
scree(pcashape) #pcashape の各主成分軸の寄与度を計算
```

```
## # A tibble: 44 x 3
##   axis proportion cumsum
##   <int>      <dbl> <dbl>
## 1     1      0.577  0.577
## 2     2      0.164  0.741
## 3     3      0.0806 0.821
## 4     4      0.0319 0.853
## 5     5      0.0307 0.884
## 6     6      0.0251 0.909
## 7     7      0.0195 0.929
## 8     8      0.0106 0.939
## 9     9      0.00918 0.948
## 10    10      0.00719 0.956
## # ... with 34 more rows
```

Now we know the main sources of shape variation, and the importance of each axis, we can now observe how each tanged point is reflected in the theoretical shape space through the `Momocs::plot_PCA()` function.

ここまでで、「かたち」の変異のおもな要素と各主成分軸の重要度を確認し、`Momocs::plot_PCA()` 関数によってそれぞれの剥片尖頭器が理論的な形状空間にどのように反映されるかを見ることができます。

```
plot_PCA(pcashape, axes = c(1,2), ~Archaeological_Unit, morphospace_position = "full_axes", zoom = 2
```



In this diagram we can observe the different distributions of each archaeological unit within the morphospace, and the relative clustering of each unit within this graph. It's important to remember that this graph only represents the first two principal components, we may wish to examine other sources of shape variation (some which may be of importance to archaeologists).

このグラフでは、それぞれの考古学的文化 (に帰属する資料が)、形状空間の中で異なった分布を示していることが確認できます。なお、このグラフでは、PC1 と PC2 しか反映されていないことに注意しておく必要があります。「かたち」の変異のほかの要素が、考古学者にとって重要なものかもしれないので、それらも検討した方が良いでしょう。

Note: pipes (%>%) are used here to processes multiple arguments at the same time. Momocs supports piping with the whole process able to be 'piped'. For teaching purposes we are doing the 'long way' of



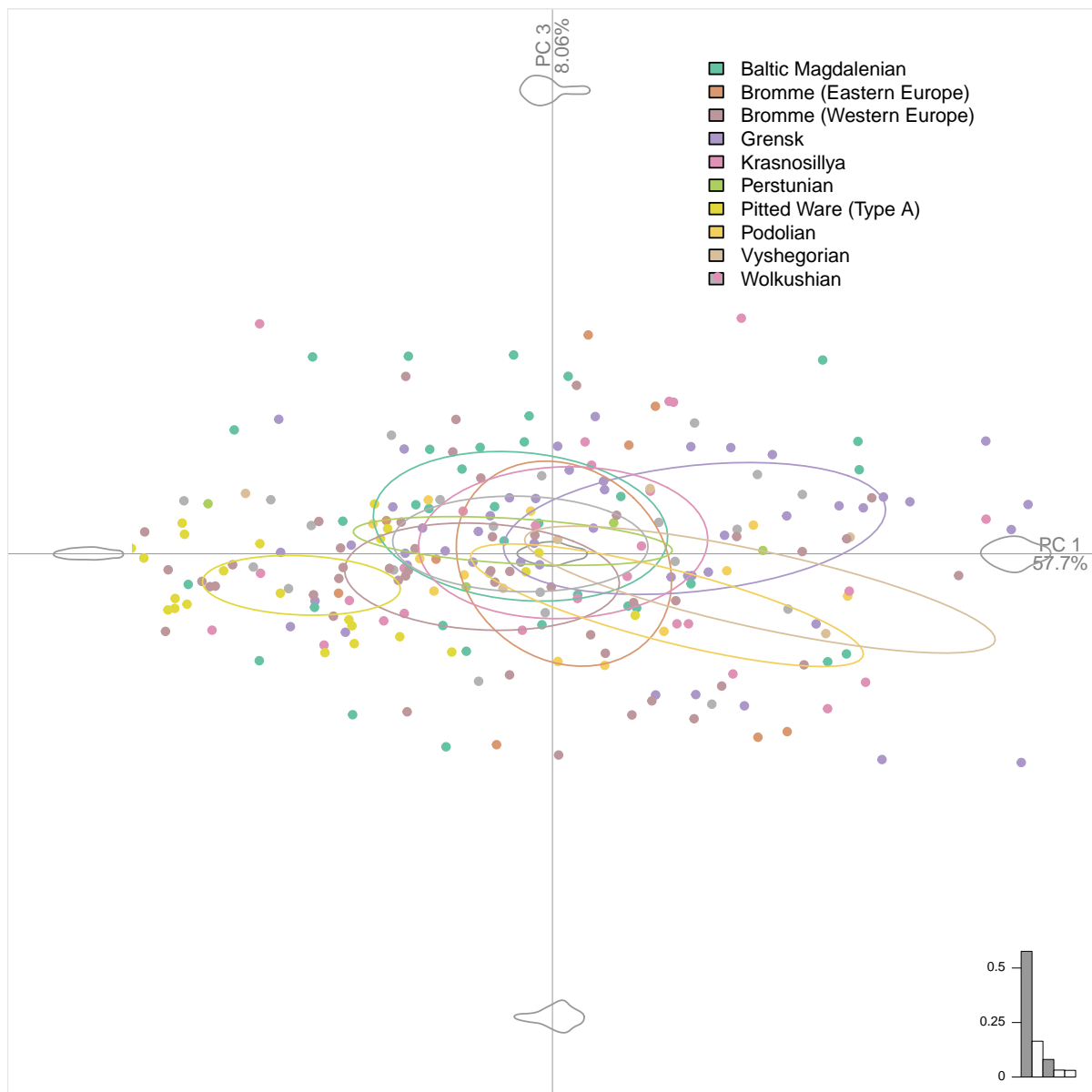
GMM.

注：ここではパイプ演算子 (`%>%`) は、同時に異なる引数を実行するために使用しています。Momocs は、パイプ演算子によりすべてのプロセスを「つなぐ」ことが可能です。

If we wish to examine the relationship between different principal components we can use the `axes` argument to change our graph configuration. For example, if we wish to examine differences in shape between PC1 and PC3 we can specify the `axes` argument in the following way:

異なる主成分軸間の関係を検討したい場合には、`plot_PCA()` 関数中の `axes` 引数を決定します。たとえば、PC1 と PC3 の関係を検討したい時には、次のようにします：

```
plot_PCA(pcashape, axes = c(1,3), ~Archaeological_Unit, morphospace_position = "full_axes", zoom = 2
```



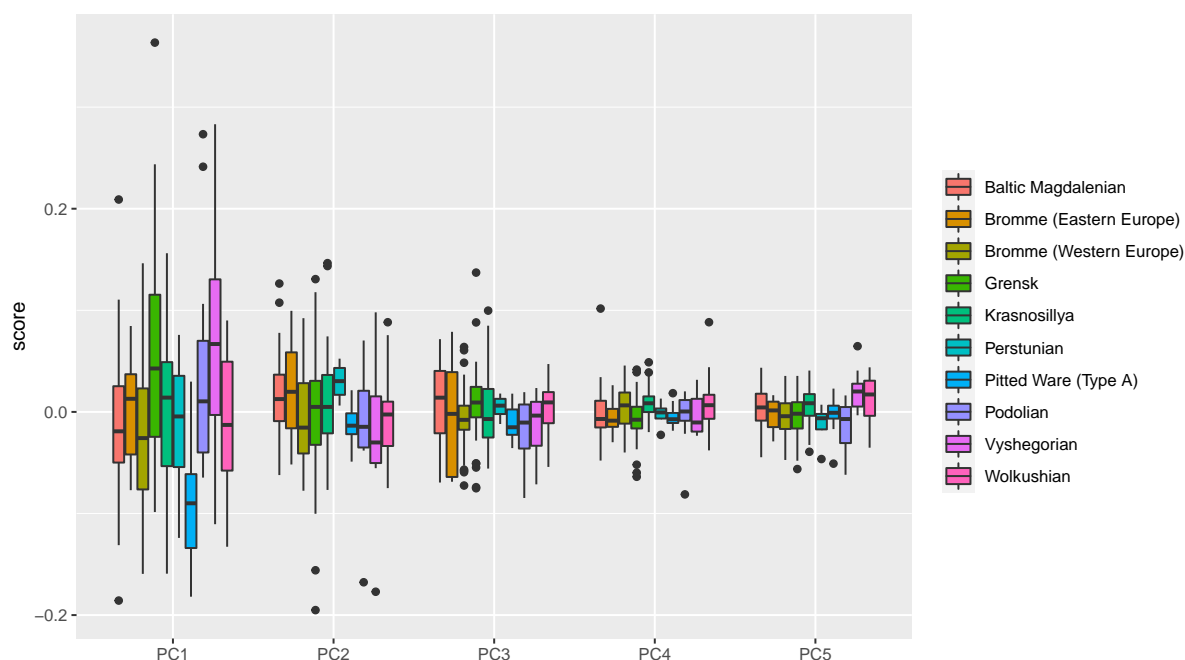
There are also a number of other visualisation options including the addition of confidence axes, convex hulls, and morphospace layouts (not explored here).

ほかにも、信頼軸や凸包、形状空間のレイアウトの追加など、可視化・グラフ描画のさまざまなオプションがありますがここでは触れません。

We can also visualise these principal components, and the variance within different archaeological units, in an alternative way, through the `Momocs::boxplot()` function:

また、`Momocs::boxplot()` 関数により、各主成分における異なる考古学的文化ごとの分散を可視化することもできます。

```
boxplot(pcshape, ~Archaeological_Unit, nax = 1:5) # pcshape の考古学的文化ごとの箱ひげ図を描画
```



## GMM Procedure 6: Discriminant Analysis (LDA/CVA) / GMM の実行 6：判別分析 (LDA/CVA)

As we highlighted in the first workshop, PCA explores differences in shape variation irrespective of group composition (i.e. *a priori* groupings). Through a discriminant analysis we can examine differences in shape as based on their maximum group separation (between-group variation in contrast to within-group variation). In Momocs, we use the `Momocs::LDA()` function on either the elliptic Fourier coefficients or the PCA scores to produce our class accuracy, plots and correction scores. There is no correct answer as to which to use, it depends on the data you wish to examine. In using the PCA scores it is possible to retain a number of components that are deemed important, this can be either: 1) the first *n*th components, 2) the number of components representing a certain level of shape variance (e.g. 95%, 99%, 99.9%), or 3) all principal components. The coefficients, in contrast would encapsulate all shape data.

第1回ワークショップで取り上げたように、PCA はあらかじめ設定された分類・グループに関係なく、「かたち」の変異を検討します。判別分析を使用すると、「かたち」の差異を、グループ間の変異に対するグループ内の変異にもとづく分離の最大化にもとづいて検討することができます。Momocs では、`Momocs::LDA()` 関数により、楕円フーリエ係数または PCA スコアを利用して、私たちの考古学的分類の精確さを検討することができます。どの方法を使えばよいかの正しい答えはありません。検討したいデータによります。PCA スコアを使用する際には、重要と考えられるいくつかの成分を利用できますが、それらは、1) 最初の *n* 番目の成分、2) 95%, 99%, 99.9% といった確実なレベルで「かたち」の変異を代表する成分、3) すべての主成分、のどれでも可能です。対照的に、係数はすべての形状データをカプセル化します。

With greater levels of data you may include a degree of statistical importance, with smaller unimportant variables taking precedence, and so an optimal level of data is necessary.

データのサイズが大きくなると統計的に重要度の高い変数を含めることができますが、小さいときには重要でない変数が優先されるので、最適なデータのサイズが必要です。

When we produced a scree table (the table with documented the percentage variance for each component) we observed that the first ten components defined 95% cumulative shape variance. We can produce a discriminant analysis on just these ten components if we wish.

特徴量の表 (各成分の分散パーセンテージを示した表) では、第 1～第 10 成分が 95% の累積的な「かたち」の分散を定めていることが確認できました。必要であれば、これらの 10 成分で判別分析を行うことができます。

First we create the object:

まずオブジェクトを作成します:

```
dashape <- LDA(pcashape, ~Archaeological_Unit, retain = 0.95) # pcashape と「考古学的文化」に対し線形判
```

We can now examine different aspects of our discriminant analysis data, including the cross-validation table (actual vs. predicted categories for artefacts) and the proportion of correctly classified individuals.

これにより、実際の資料と想定された分類間の交差検証 (クロスバリデーション) 表など、判別分析データの異なる様相を検討できます。

```
dashape$CV.correct #dashape を交差検証
```

```
## [1] 0.324
```

```
dashape$CV.ce # 交差検証表を作成
```

```
##      Baltic Magdalenian Bromme (Eastern Europe) Bromme (Western Europe)
##      0.2222222      0.0000000      0.5510204
##      Grensk      Krasnosillya      Perstunian
##      0.4727273      0.1379310      0.0000000
##      Pitted Ware (Type A)      Podolian      Vyshegorian
##      0.5000000      0.0000000      0.0000000
##      Wolkushian
##      0.1818182
```

When we use the `CV.correct` argument we see that 32.4% of tanged points can be correctly classified. We can examine this in further detail through the `CV.ce` argument.

`CV.correct` 引数によって、全資料のうち 32.4% の剥片尖頭器が「正しく」分類されていた (判別分析により主成分分析によるグルーピングと「考古学的文化」が一致していた) ことを確認できます。その詳細を、`'CV.ce'` 引数によって検証できます (全体ではなく、考古学的文化ごとの交差検証)。

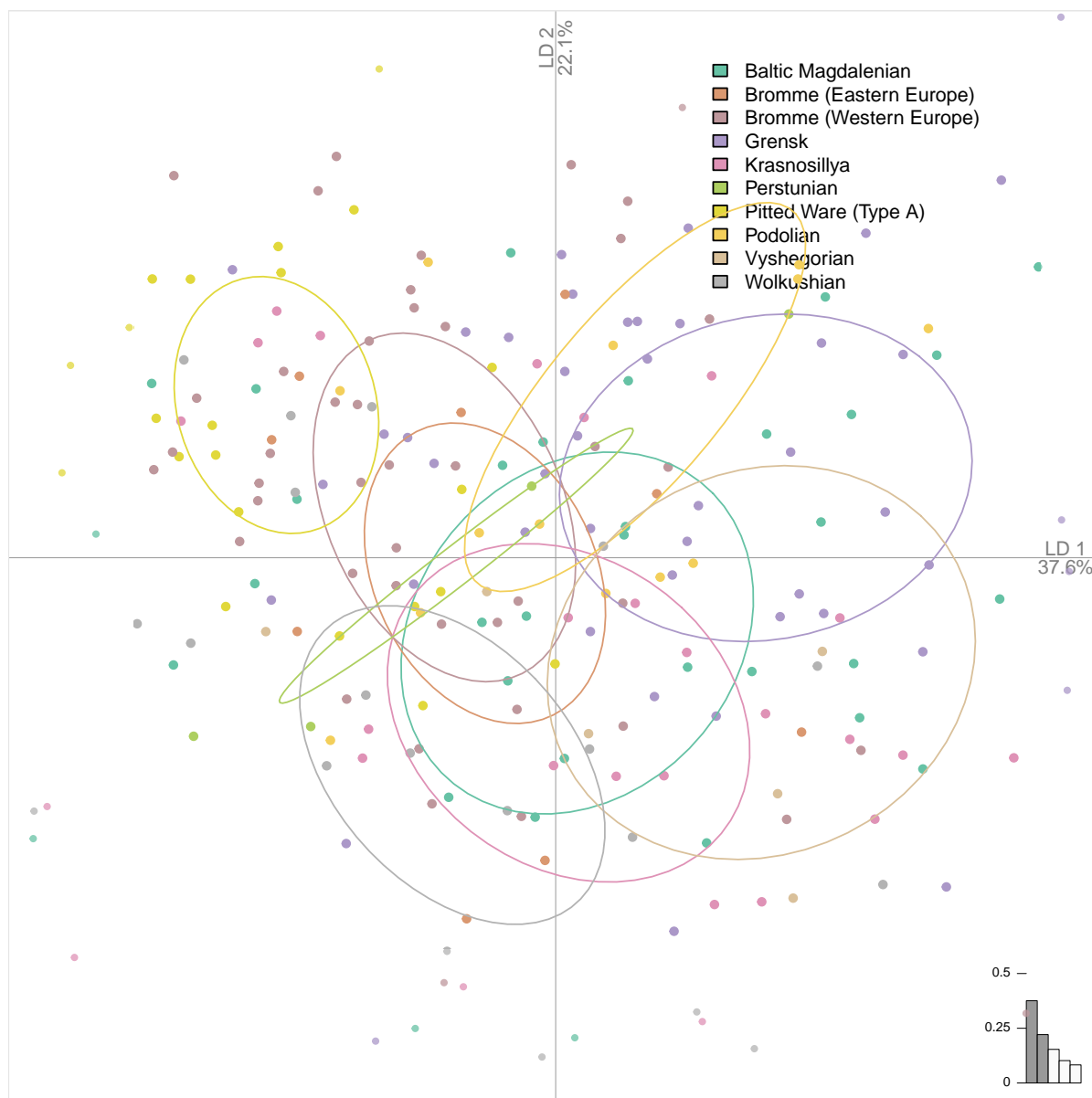
We can see through our classification error table that certain archaeological units are better defined in two-dimensional shape than others e.g. Pitted Ware (Type A), Grensk and Bromme (Western Europe). More detailed metrics are included in the `Momocs::classification_metrics()` function (not covered here).

この分類の誤差の表にもとづいて、いくつかの「考古学的文化」—たとえば Pitted Ware (Type A) や Grensk, 西欧 Bromme 文化などは他の考古学的文化に比べて、剥片尖頭器の 2D の「かたち」によってよく定義されていることを確認できます。より詳細な測定は、`Momocs::classification_metrics()` 関数で実施できますが、ここでは取り上げません。

If we wish to visualise our plot, as is common in exploratory procedures we can use the `Momocs::plot_LDA()` function, using similar arguments to `Momocs::plot_PCA()`:

この結果を散布図に描画するには、`Momocs::plot_PCA()` と同じ引数を使う、`Momocs::plot_LDA()` 関数を共通する手順として使用することができます。

```
plot_LDA(dashape, axes = c(1,2), zoom = 2, chull = FALSE) %>% layer_points(cex = 1) %>% layer_ellips
```



Here we can see how Pitted Ware (Type A), a control from a different period (but of similar technique) can be differentiated from all other archaeological units.

この散布図から、時期が異なるが同じ技法を用いている Pitted Ware (Type A) が、以下にほかの考古学的文化と異なっているかを確認できます。

Note: for the elliptic Fourier coefficient discriminant analysis it is relatively straight forward to impose the shapes onto the graph using the `layer_morphospace_LDA()` argument.

楕円フーリエ係数の判別分析のために、`layer_morphospace_LDA()` 引数を使用したグラフに「かたち」を重ね合わせることは比較的容易です。

## GMM Procedure 7: Multiple Analysis of Variance (MANOVA) / GMM の実行 7: 多変

## 量分散分析 (MANOVA)

So far we have explored the differences in shape within the whole group of artefacts and explored how well they can be separated through their group variance. Now we need to test, within an statistical framework, whether there is a difference in the PC scores (representative of shape) within and between the different archaeological units. A MANOVA will be our required test given we have multiple groups and multiple column data (PC scores).

ここまで、対象資料全体の「かたち」の差異と、それがグループの分散によりどのように区分されるのかを検討してきました。続いて、統計学的な枠組みにおいて、考古学的文化内あるいは文化間において、「かたち」を代表する主成分スコアに差異があるのかどうかを検証する必要があります。MANOVA は、複数の分類群と PC スコアのような複数列のデータを検証するのに適しています。

Once we have chosen a desired alpha level as of marker of difference (that is to say the boundary with which we are able to reject the null hypothesis of same populations) e.g. 0.05 we can use the `Momocs::MANOVA()` function, noting “Archaeological\_Unit” to be our factor which we want to consider:

差異の示標として期待されるアルファ・レベル、つまり同一母集団に対する帰無仮説を棄却することができる境界として、たとえば 0.05 を選ぶと、「考古学的文化」を検討する因子として `Momocs::MANOVA()` 関数を使用することができます。

```
MANOVA(pcashape, ~Archaeological_Unit, retain = 0.95) # pcashape と「考古学的文化」の MANOVA を実施
```

```
##           Df Hotelling-Lawley approx F num Df den Df      Pr(>F)
## fac           9           1.0137    2.5819     90   2063 1.593e-13 ***
## Residuals 240
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Note how we are still using 95% cumulative shape variance as represented by our principal component scores. Once we perform the MANOVA we can see that the null hypothesis is rejected as the p value is below the 0.05 level (or any of the other significance levels). We can examine this in finer detail through pair-wise MANOVA analyses, using the `Momocs::MANOVA_PW()` function:

ここでは、主成分スコアで表される 95% 累積形状分散をまだ使用していることに注目してください。MANOVA を実行したところ、p 値が 0.05 レベル（または他の有意水準のいずれか）を下回っていたため、帰無仮説が棄却されていることがわかります。これをより詳細に調べるには、`Momocs::MANOVA_PW()` 関数を用いて、ペアワイズ法 MANOVA 分析を行うとよいでしょう。

```
MANOVA_PW(pcashape, ~Archaeological_Unit, retain = 0.95) # ペアワイズ法 MANOVA を実施
```

```
## $stars.tab
##           Baltic Magdalenian Bromme (Eastern Europe)
```

```

## Baltic Magdalenian -
## Bromme (Eastern Europe)
## Bromme (Western Europe)
## Grensk
## Krasnosillya
## Perstunian
## Pitted Ware (Type A)
## Podolian
## Vyshegorian

## Bromme (Western Europe) Grensk Krasnosillya Perstunian
## Baltic Magdalenian ** * - -
## Bromme (Eastern Europe) - - -
## Bromme (Western Europe) *** * -
## Grensk ** -
## Krasnosillya -
## Perstunian
## Pitted Ware (Type A)
## Podolian
## Vyshegorian

## Pitted Ware (Type A) Podolian Vyshegorian Wolkushian
## Baltic Magdalenian *** * * -
## Bromme (Eastern Europe) *** - - -
## Bromme (Western Europe) ** - * -
## Grensk *** - - **
## Krasnosillya *** - . -
## Perstunian *** - - -
## Pitted Ware (Type A) *** *** ***
## Podolian - -
## Vyshegorian .
##

## $summary (see also $manovas)

## Df Pillai approx F num Df
## Baltic Magdalenian - Bromme (Eastern Europe) 1 0.02214 0.22642 4
## Baltic Magdalenian - Bromme (Western Europe) 1 0.17590 4.26888 4
## Baltic Magdalenian - Grensk 1 0.12880 3.17862 4
## Baltic Magdalenian - Krasnosillya 1 0.07292 1.17975 4
## Baltic Magdalenian - Perstunian 1 0.01123 0.09936 4
## Baltic Magdalenian - Pitted Ware (Type A) 1 0.43615 10.63610 4
## Baltic Magdalenian - Podolian 1 0.19030 2.64405 4
## Baltic Magdalenian - Vyshegorian 1 0.22929 2.90074 4

```



|  |     |         |           |   |
|--|-----|---------|-----------|---|
| ## Baltic Magdalenian - Wolkushian                   | 1   | 0.08186 | 1.18136   | 4 |
| ## Bromme (Eastern Europe) - Bromme (Western Europe) | 1   | 0.10733 | 1.59313   | 4 |
| ## Bromme (Eastern Europe) - Grensk                  | 1   | 0.05268 | 0.82029   | 4 |
| ## Bromme (Eastern Europe) - Krasnosillya            | 1   | 0.12193 | 1.14560   | 4 |
| ## Bromme (Eastern Europe) - Perstunian              | 1   | 0.10234 | 0.22801   | 4 |
| ## Bromme (Eastern Europe) - Pitted Ware (Type A)    | 1   | 0.51283 | 7.36881   | 4 |
| ## Bromme (Eastern Europe) - Podolian                | 1   | 0.20740 | 1.17751   | 4 |
| ## Bromme (Eastern Europe) - Vyshegorian             | 1   | 0.25072 | 1.00384   | 4 |
| ## Bromme (Eastern Europe) - Wolkushian              | 1   | 0.11384 | 0.83503   | 4 |
| ## Bromme (Western Europe) - Grensk                  | 1   | 0.23821 | 7.73923   | 4 |
| ## Bromme (Western Europe) - Krasnosillya            | 1   | 0.14484 | 3.09097   | 4 |
| ## Bromme (Western Europe) - Perstunian              | 1   | 0.09731 | 1.29366   | 4 |
| ## Bromme (Western Europe) - Pitted Ware (Type A)    | 1   | 0.21505 | 4.65731   | 4 |
| ## Bromme (Western Europe) - Podolian                | 1   | 0.12311 | 2.03571   | 4 |
| ## Bromme (Western Europe) - Vyshegorian             | 1   | 0.19687 | 3.18671   | 4 |
| ## Bromme (Western Europe) - Wolkushian              | 1   | 0.08053 | 1.44514   | 4 |
| ## Grensk - Krasnosillya                             | 1   | 0.17393 | 4.15845   | 4 |
| ## Grensk - Perstunian                               | 1   | 0.04248 | 0.59899   | 4 |
| ## Grensk - Pitted Ware (Type A)                     | 1   | 0.36709 | 10.73009  | 4 |
| ## Grensk - Podolian                                 | 1   | 0.10425 | 1.86218   | 4 |
| ## Grensk - Vyshegorian                              | 1   | 0.07279 | 1.13823   | 4 |
| ## Grensk - Wolkushian                               | 1   | 0.18226 | 4.01179   | 4 |
| ## Krasnosillya - Perstunian                         | 1   | 0.07646 | 0.57950   | 4 |
| ## Krasnosillya - Pitted Ware (Type A)               | 1   | 0.49924 | 11.96361  | 4 |
| ## Krasnosillya - Podolian                           | 1   | 0.17440 | 2.00682   | 4 |
| ## Krasnosillya - Vyshegorian                        | 1   | 0.23935 | 2.51739   | 4 |
| ## Krasnosillya - Wolkushian                         | 1   | 0.03428 | 0.40820   | 4 |
| ## Perstunian - Pitted Ware (Type A)                 | 1   | 0.59901 | 8.58949   | 4 |
| ## Perstunian - Podolian                             | 1   | 0.31055 | 1.46393   | 4 |
| ## Perstunian - Vyshegorian                          | 1   | 0.28762 | 0.70655   | 4 |
| ## Perstunian - Wolkushian                           | 1   | 0.11196 | 0.66190   | 4 |
| ## Pitted Ware (Type A) - Podolian                   | 1   | 0.54690 | 9.95782   | 4 |
| ## Pitted Ware (Type A) - Vyshegorian                | 1   | 0.57670 | 9.19628   | 4 |
| ## Pitted Ware (Type A) - Wolkushian                 | 1   | 0.44407 | 8.18756   | 4 |
| ## Podolian - Vyshegorian                            | 1   | 0.19298 | 1.01626   | 4 |
| ## Podolian - Wolkushian                             | 1   | 0.18524 | 1.76197   | 4 |
| ## Vyshegorian - Wolkushian                          | 1   | 0.27536 | 2.37503   | 4 |
| ##   | den | Df      | Pr(>F)    |   |
| ## Baltic Magdalenian - Bromme (Eastern Europe)      | 40  |         | 9.220e-01 |   |
| ## Baltic Magdalenian - Bromme (Western Europe)      | 80  |         | 3.501e-03 |   |

|  |              |
|--|--------------|
| ## Baltic Magdalenian - Grensk                       | 86 1.740e-02 |
| ## Baltic Magdalenian - Krasnosillya                 | 60 3.289e-01 |
| ## Baltic Magdalenian - Perstunian                   | 35 9.820e-01 |
| ## Baltic Magdalenian - Pitted Ware (Type A)         | 55 1.865e-06 |
| ## Baltic Magdalenian - Podolian                     | 45 4.572e-02 |
| ## Baltic Magdalenian - Vyshegorian                  | 39 3.407e-02 |
| ## Baltic Magdalenian - Wolkushian                   | 53 3.296e-01 |
| ## Bromme (Eastern Europe) - Bromme (Western Europe) | 53 1.897e-01 |
| ## Bromme (Eastern Europe) - Grensk                  | 59 5.174e-01 |
| ## Bromme (Eastern Europe) - Krasnosillya            | 33 3.524e-01 |
| ## Bromme (Eastern Europe) - Perstunian              | 8 9.151e-01  |
| ## Bromme (Eastern Europe) - Pitted Ware (Type A)    | 28 3.469e-04 |
| ## Bromme (Eastern Europe) - Podolian                | 18 3.539e-01 |
| ## Bromme (Eastern Europe) - Vyshegorian             | 12 4.432e-01 |
| ## Bromme (Eastern Europe) - Wolkushian              | 26 5.153e-01 |
| ## Bromme (Western Europe) - Grensk                  | 99 1.810e-05 |
| ## Bromme (Western Europe) - Krasnosillya            | 73 2.081e-02 |
| ## Bromme (Western Europe) - Perstunian              | 48 2.858e-01 |
| ## Bromme (Western Europe) - Pitted Ware (Type A)    | 68 2.210e-03 |
| ## Bromme (Western Europe) - Podolian                | 58 1.012e-01 |
| ## Bromme (Western Europe) - Vyshegorian             | 52 2.047e-02 |
| ## Bromme (Western Europe) - Wolkushian              | 66 2.290e-01 |
| ## Grensk - Krasnosillya                             | 79 4.151e-03 |
| ## Grensk - Perstunian                               | 54 6.649e-01 |
| ## Grensk - Pitted Ware (Type A)                     | 74 6.508e-07 |
| ## Grensk - Podolian                                 | 64 1.279e-01 |
| ## Grensk - Vyshegorian                              | 58 3.476e-01 |
| ## Grensk - Wolkushian                               | 72 5.405e-03 |
| ## Krasnosillya - Perstunian                         | 28 6.799e-01 |
| ## Krasnosillya - Pitted Ware (Type A)               | 48 8.025e-07 |
| ## Krasnosillya - Podolian                           | 38 1.131e-01 |
| ## Krasnosillya - Vyshegorian                        | 32 6.065e-02 |
| ## Krasnosillya - Wolkushian                         | 46 8.018e-01 |
| ## Perstunian - Pitted Ware (Type A)                 | 23 2.153e-04 |
| ## Perstunian - Podolian                             | 13 2.692e-01 |
| ## Perstunian - Vyshegorian                          | 7 6.123e-01  |
| ## Perstunian - Wolkushian                           | 21 6.253e-01 |
| ## Pitted Ware (Type A) - Podolian                   | 33 2.129e-05 |
| ## Pitted Ware (Type A) - Vyshegorian                | 27 8.007e-05 |
| ## Pitted Ware (Type A) - Wolkushian                 | 41 5.989e-05 |

```
## Podolian - Vyshegorian          17 4.268e-01
## Podolian - Wolkushian            31 1.617e-01
## Vyshegorian - Wolkushian         25 7.926e-02
```

This rather large amount of information provides the p values for each combination of archaeological units and depicts level of significance in star form. In terms of analysis this data highlights, as previously the degree to which specific archaeological units can be distinguished from others in terms of their two-dimensional outline shape.

この大量の情報は、各考古学文化間の p 値を示し、有意水準を \*(アスタリスク) で示しています。分析の観点から、このデータが、先に述べたように特定の考古学的文化が、そこに含まれる剥片尖頭器の 2D のアウトラインの「かたち」によって他の文化とどの程度区別できるかを示しています。

### ## GMM Procedure 8: Hierarchical Cluster Analysis and K-Means / GMM の実行 8 : 階層的クラスタ分析と k-means 法

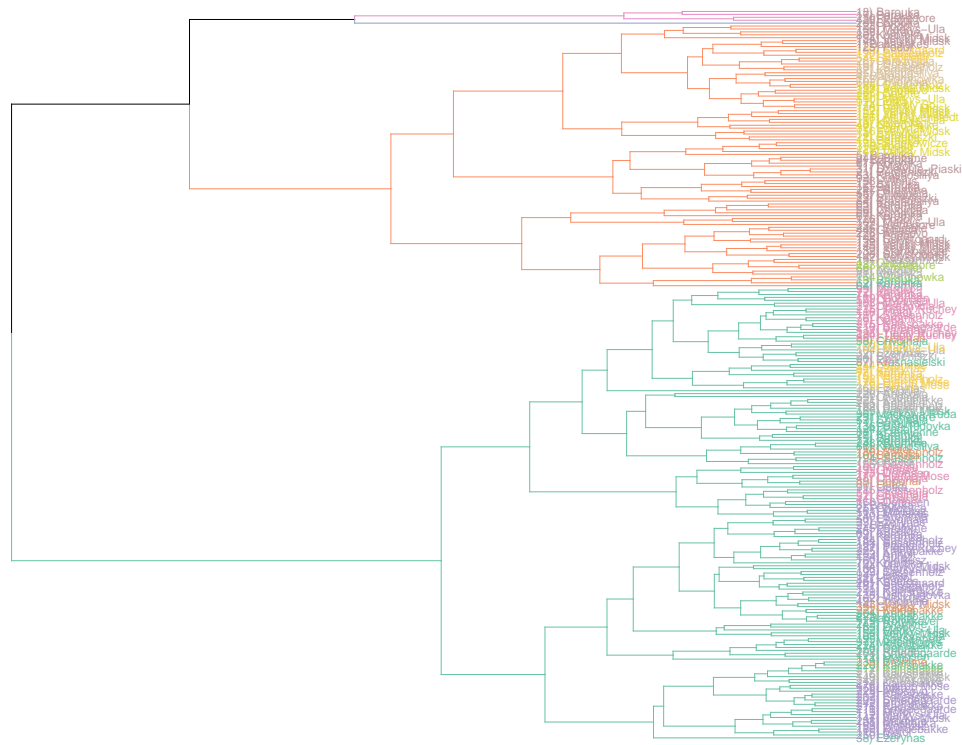
We can now use the elliptic Fourier coefficients and PCA data to examine, irrespective of previous groupings, how similar objects relate to one another within the overall set of examples. The endpoint is a set of clusters, where each cluster is distinct from each other cluster, and the objects within each cluster are broadly similar i.e. of similar shape. This can be done through two different methods in Momocs: Hierarchical Cluster Analysis (through its various subcategories), where the structure is provided, or through a K-Means analysis where partitions the shapes into k groups.

楕円フーリエ係数と PCA データを、従来の分類に関係なく、類似した資料が、対象資料全体の中でどのように相互に関連しているかを検証するために利用できます。その結果は、各クラスターが他のクラスターから区別され、また各クラスター内の資料が大まかに類似している、つまり類似した「かたち」となるクラスターの集合として示されます。これは、Momocs では 2 つの異なる方法で行うことができます。多様なサブカテゴリを用いた階層的クラスタ分析と、「かたち」を k 個のグループに分割する K-Means 法です。

To perform a Hierarchical Cluster Analysis we can use the `Momocs::CLUST()` function, a wrapper of `stats::dist()` and `stats::hclust()`. We can specify what type of shape we wish for our tree to be using the `type` argument (horizontal as default), and the specific `hclust` (complete as default) and `dist_method` (euclidean as default). Again, we can retain the number of PCA scores we find suitable or use the elliptic Fourier coefficients.

階層型クラスタ分析を行うには、`stats::dist()` と `stats::hclust()` を含む `Momocs::CLUST()` 関数を用います。引数 `type` (デフォルトは水平) と、個別の `hclust` (デフォルトは完全) と `dist_method` (デフォルトはユークリッド) を用いて、クラスタ樹形図をどのような形状にするか指定できます。ここでも適切と思われる PCA スコアの数を保持するか、楕円フーリエ係数を使用することができます。

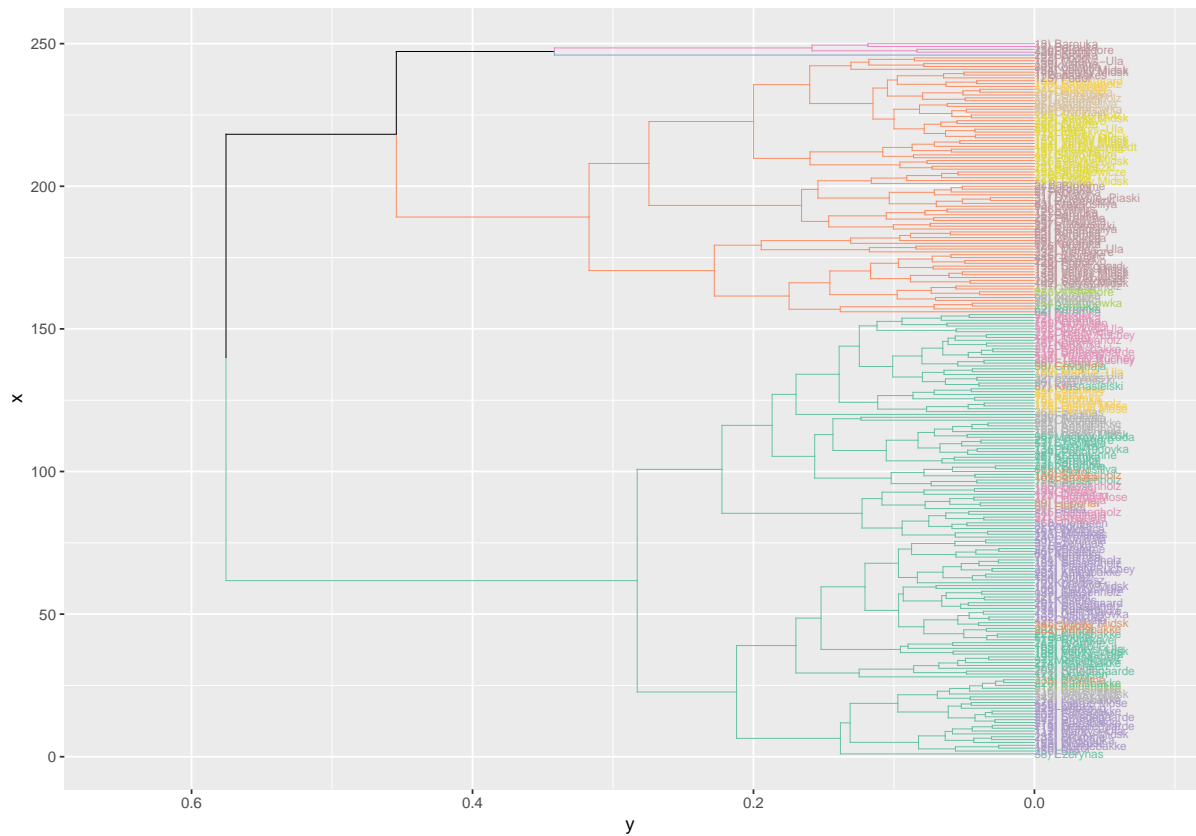
```
CLUST(pcashape, ~Archaeological_Unit, dist_method = "euclidean", hclust_method = "complete", k = 4,
```



Using the `k` argument, I've also specified what the best four groupings would be. We can also modify the aesthetic further through arguments in the tidyverse.

`k` 引数を使って、最もよく分類-結合される 4 つのグループを特定しました。tidyverse の引数を利用して、見た目を変更することもできます。

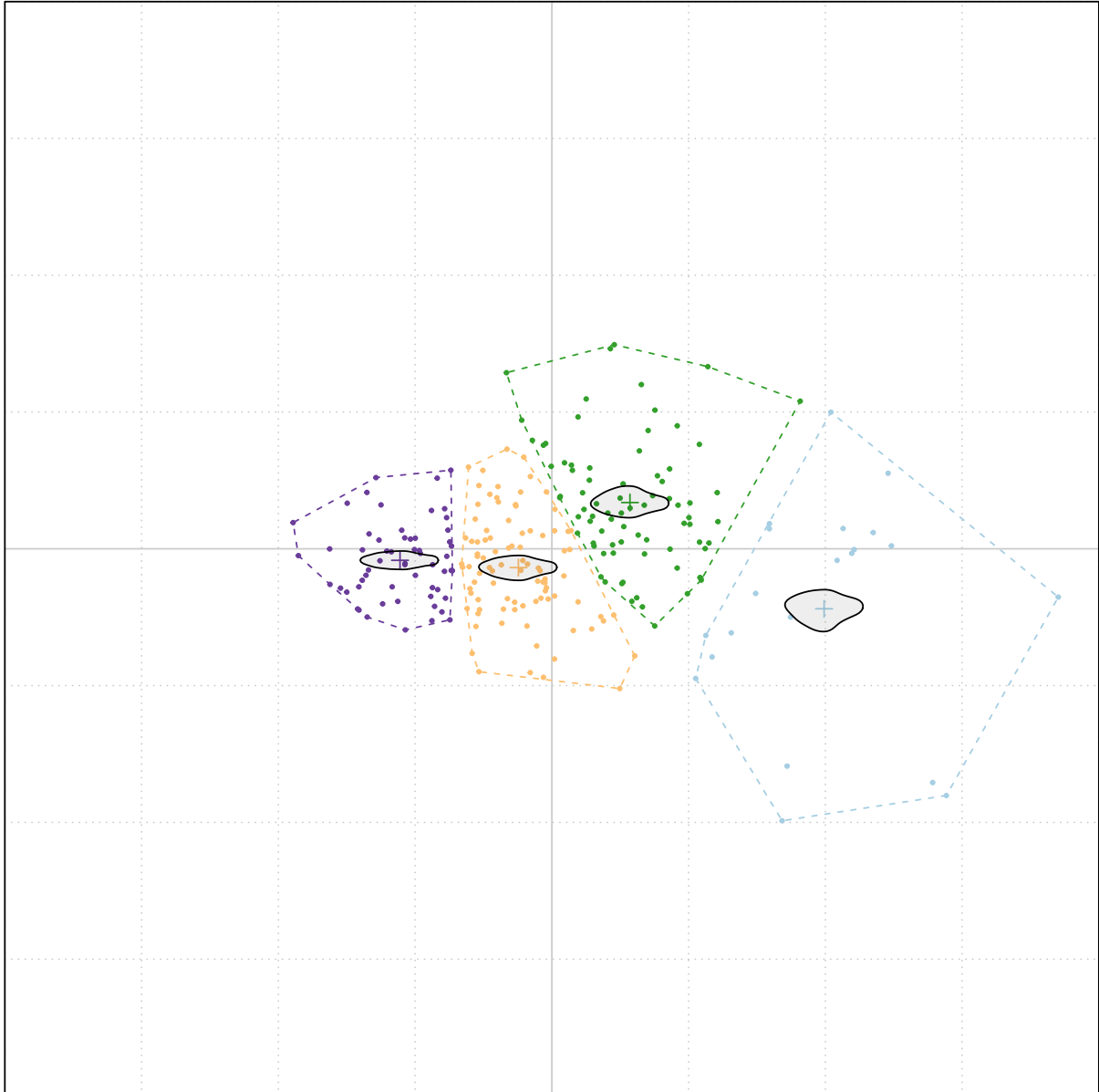
```
CLUST(pcashape, ~Archaeological_Unit, dist_method = "euclidean", hclust_method = "complete", k = 4,
```



Alternatively we can use the `Momocs::KMEANS()` function to derive four groups from the data.

また、'Momocs::KMEANS()' 関数を使い、データを群に分類することもできます。

```
KMEANS(pcashape, centers = 4) # クラスタ中心 (分離数) を 4 として k-means 法を実施
```



```
## K-means clustering with 4 clusters of sizes 21, 79, 92, 58
##
## Cluster means:
##          PC1          PC2
## 1  0.19540358 -0.043032957
## 2  0.05604376  0.033196722
## 3 -0.02402585 -0.013482066
## 4 -0.10897507 -0.008249981
##
## Clustering vector:
```

|    |                 |                 |                     |                 |
|----|-----------------|-----------------|---------------------|-----------------|
| ## | 1) Baltasiskes  | 2) Barouka      | 3) Barouka          | 4) Barouka      |
| ## | 2               | 4               | 2                   | 2               |
| ## | 5) Barouka      | 6) Barouka      | 7) Barouka          | 8) Barouka      |
| ## | 1               | 4               | 2                   | 3               |
| ## | 9) Barouka      | 10) Barouka     | 11) Barouka         | 12) Barouka     |
| ## | 2               | 2               | 3                   | 1               |
| ## | 13) Barouka     | 14) Barouka     | 15) Barouka         | 16) Barouka     |
| ## | 2               | 2               | 2                   | 1               |
| ## | 17) Barouka     | 18) Barouka     | 19) Barouka         | 20) Barouka     |
| ## | 1               | 1               | 3                   | 1               |
| ## | 21) Burdeniszki | 22) Burdeniszki | 23) Burdeniszki     | 24) Burdeniszki |
| ## | 2               | 2               | 2                   | 3               |
| ## | 25) Chilczyce   | 26) Dereznycia  | 27) Dereznycia      | 28) Duba        |
| ## | 3               | 2               | 3                   | 3               |
| ## | 29) Duba        | 30) Duba        | 31) Dziewule-Piaski | 32) Ezerynas    |
| ## | 3               | 3               | 2                   | 3               |
| ## | 33) Ezerynas    | 34) Ezerynas    | 35) Ezerynas        | 36) Ezerynas    |
| ## | 4               | 3               | 3                   | 3               |
| ## | 37) Ezerynas    | 38) Ezerynas    | 39) Glukas          | 40) Glyno Pelke |
| ## | 2               | 4               | 3                   | 2               |
| ## | 41) Gribasa     | 42) Gribasa     | 43) Gribasa         | 44) Gribasa     |
| ## | 3               | 2               | 3                   | 2               |
| ## | 45) Kasetos     | 46) Kasetos     | 47) Katra           | 48) Chvojnjaja  |
| ## | 4               | 3               | 3                   | 4               |
| ## | 49) Chvojnjaja  | 50) Chvojnjaja  | 51) Chvojnjaja      | 52) Chvojnjaja  |
| ## | 3               | 3               | 3                   | 3               |
| ## | 53) Chvojnjaja  | 54) Chvojnjaja  | 55) Chvojnjaja      | 56) Chvojnjaja  |
| ## | 3               | 2               | 1                   | 1               |
| ## | 57) Chvojnjaja  | 58) Chvojnjaja  | 59) Chvojnjaja      | 60) Chvojnjaja  |
| ## | 3               | 3               | 3                   | 2               |
| ## | 61) Koromka     | 62) Koromka     | 63) Koromka         | 64) Koromka     |
| ## | 2               | 2               | 1                   | 2               |
| ## | 65) Koromka     | 66) Koromka     | 67) Koromka         | 68) Koromka     |
| ## | 1               | 2               | 1                   | 1               |
| ## | 69) Koromka     | 70) Koromka     | 71) Koromka         | 72) Koromka     |
| ## | 4               | 4               | 3                   | 2               |
| ## | 73) Koromka     | 74) Koromka     | 75) Koromka         | 76) Koromka     |
| ## | 3               | 4               | 3                   | 3               |
| ## | 77) Koromka     | 78) Koromka     | 79) Koromka         | 80) Koromka     |
| ## | 3               | 3               | 2                   | 2               |

|    |                   |                   |                   |                   |
|----|-------------------|-------------------|-------------------|-------------------|
| ## | 81) Koromka       | 82) Koromka       | 83) Krasnosillya  | 84) Krasnosillya  |
| ## | 2                 | 2                 | 2                 | 3                 |
| ## | 85) Krasnosillya  | 86) Krasnosillya  | 87) Krasnasielski | 88) Krzemienne    |
| ## | 2                 | 2                 | 2                 | 3                 |
| ## | 89) Lieporiai     | 90) Lipa          | 91) Liutka        | 92) Liutka        |
| ## | 3                 | 3                 | 4                 | 2                 |
| ## | 93) Liutka        | 94) Liutka        | 95) Liutka        | 96) Mackowa Ruda  |
| ## | 2                 | 4                 | 4                 | 3                 |
| ## | 97) Marcinkonys   | 98) Margiu        | 99) Margiu        | 100) Margiu       |
| ## | 4                 | 2                 | 2                 | 3                 |
| ## | 101) Maskauka     | 102) Markys-Ula   | 103) Markys-Ula   | 104) Markys-Ula   |
| ## | 4                 | 4                 | 4                 | 3                 |
| ## | 105) Markys-Ula   | 106) Markys-Ula   | 107) Markys-Ula   | 108) Markys-Ula   |
| ## | 2                 | 4                 | 1                 | 2                 |
| ## | 109) Markys-Ula   | 110) Markys-Ula   | 111) Markys-Ula   | 112) Markys-Ula   |
| ## | 3                 | 2                 | 2                 | 4                 |
| ## | 113) Mitriskes    | 114) Motol        | 115) Motol        | 116) Motol        |
| ## | 2                 | 3                 | 4                 | 3                 |
| ## | 117) Motol        | 118) Motol        | 119) Motol        | 120) Motol        |
| ## | 2                 | 2                 | 4                 | 3                 |
| ## | 121) Plaska       | 122) Podol        | 123) Podol        | 124) Podol        |
| ## | 4                 | 2                 | 2                 | 3                 |
| ## | 125) Podol        | 126) Podol        | 127) Podol        | 128) Podol        |
| ## | 2                 | 1                 | 3                 | 1                 |
| ## | 129) Podol        | 130) Rudnia       | 131) Rudnya       | 132) Stankowicze  |
| ## | 2                 | 3                 | 3                 | 2                 |
| ## | 133) Stankowicze  | 134) Suraz        | 135) Ust-Tudovka  | 136) Ust-Tudovka  |
| ## | 2                 | 4                 | 3                 | 3                 |
| ## | 137) Varena       | 138) Varene       | 139) Velyky Midsk | 140) Velyky Midsk |
| ## | 3                 | 2                 | 1                 | 2                 |
| ## | 141) Velyky Midsk | 142) Velyky Midsk | 143) Velyky Midsk | 144) Velyky Midsk |
| ## | 2                 | 2                 | 3                 | 3                 |
| ## | 145) Velyky Midsk | 146) Velyky Midsk | 147) Velyky Midsk | 148) Velyky Midsk |
| ## | 2                 | 3                 | 4                 | 2                 |
| ## | 149) Velyky Midsk | 150) Velyky Midsk | 151) Velyky Midsk | 152) Velyky Midsk |
| ## | 4                 | 4                 | 2                 | 3                 |
| ## | 153) Velyky Midsk | 154) Velyky Midsk | 155) Velyky Midsk | 156) Velyky Midsk |
| ## | 2                 | 2                 | 3                 | 2                 |
| ## | 157) Vilnius      | 158) Vilnius      | 159) Wolkusz      | 160) Wolkusz      |
| ## | 2                 | 1                 | 4                 | 4                 |



|    |                    |                     |                    |                    |
|----|--------------------|---------------------|--------------------|--------------------|
| ## | 161) Wolkusz       | 162) Wolkusz        | 163) Woronowka     | 164) Woronowka     |
| ## | 2                  | 3                   | 2                  | 2                  |
| ## | 165) Zusno         | 166) Alt Duvenstedt | 167) Dohnsen       | 168) Dohnsen       |
| ## | 3                  | 2                   | 2                  | 3                  |
| ## | 169) Dohnsen       | 170) Dohnsen        | 171) Dohnsen       | 172) Dohnsen       |
| ## | 3                  | 4                   | 4                  | 3                  |
| ## | 173) Dohnsen       | 174) Elemly So      | 175) Hjarup Mose   | 176) Hjarup Mose   |
| ## | 3                  | 2                   | 3                  | 4                  |
| ## | 177) Hjarup Mose   | 178) Hjarup Mose    | 179) Rolykkevej    | 180) Rundebakke    |
| ## | 3                  | 3                   | 4                  | 4                  |
| ## | 181) Sassenholz    | 182) Sassenholz     | 183) Sassenholz    | 184) Sassenholz    |
| ## | 2                  | 3                   | 4                  | 4                  |
| ## | 185) Sassenholz    | 186) Sassenholz     | 187) Sassenholz    | 188) Sassenholz    |
| ## | 3                  | 3                   | 3                  | 3                  |
| ## | 189) Sassenholz    | 190) Sassenholz     | 191) Sassenholz    | 192) Sassenholz    |
| ## | 2                  | 3                   | 2                  | 3                  |
| ## | 193) Sassenholz    | 194) Sassenholz     | 195) Sassenholz    | 196) Sassenholz    |
| ## | 3                  | 2                   | 4                  | 3                  |
| ## | 197) Sassenholz    | 198) Solystgaard    | 199) Solystgaard   | 200) Solystgaard   |
| ## | 3                  | 1                   | 3                  | 2                  |
| ## | 201) Solystgaard   | 202) Anholt         | 203) Anholt        | 204) Anholt        |
| ## | 3                  | 4                   | 4                  | 4                  |
| ## | 205) Anholt        | 206) Sotofte        | 207) Follenslev    | 208) Smedegaarde   |
| ## | 3                  | 2                   | 4                  | 4                  |
| ## | 209) Smedegaarde   | 210) Smedegaarde    | 211) Smedegaarde   | 212) Kainsbakke    |
| ## | 4                  | 3                   | 4                  | 4                  |
| ## | 213) Kainsbakke    | 214) Kainsbakke     | 215) Kainsbakke    | 216) Kainsbakke    |
| ## | 4                  | 4                   | 3                  | 4                  |
| ## | 217) Kainsbakke    | 218) Kainsbakke     | 219) Kainsbakke    | 220) Kainsbakke    |
| ## | 4                  | 4                   | 3                  | 4                  |
| ## | 221) Kainsbakke    | 222) Kainsbakke     | 223) Kainsbakke    | 224) Kainsbakke    |
| ## | 3                  | 4                   | 3                  | 4                  |
| ## | 225) Kainsbakke    | 226) Anosovo        | 227) Anosovo       | 228) Anosovo       |
| ## | 3                  | 2                   | 3                  | 3                  |
| ## | 229) Anosovo       | 230) Vishegore      | 231) Vishegore     | 232) Vishegore     |
| ## | 4                  | 1                   | 3                  | 1                  |
| ## | 233) Vishegore     | 234) Tieply Ruchey  | 235) Tieply Ruchey | 236) Tieply Ruchey |
| ## | 2                  | 3                   | 3                  | 3                  |
| ## | 237) Tieply Ruchey | 238) Bromme         | 239) Bromme        | 240) Bromme        |
| ## | 3                  | 4                   | 4                  | 2                  |

```
##          241) Bromme          242) Bromme          243) Bromme          244) Bromme
##                2                4                3                3
##          245) Bromme          246) Bromme          247) Bromme          248) Bromme
##                4                2                1                2
##    249) Trollesgave          250) Bro
##                4                4
##
## Within cluster sum of squares by cluster:
## [1] 0.2064188 0.2765445 0.2051620 0.0936057
## (between_SS / total_SS = 71.3 %)
##
## Available components:
##
## [1] "cluster"      "centers"      "totss"        "withinss"     "tot.withinss"
## [6] "betweenss"    "size"         "iter"         "ifault"
```

If more computationally-intensive tree-building exercises were investigated we could explore the principal components through **maximum likelihood** in the RPhylib package (this requires the Phylip software to be installed on a computer already). All these trees can also be imported into the **ggtree** for full customisation, or analysed for their structures (phylogenetic or otherwise) in the **ape** package.

もし、より計算量の多い樹形図を作成したいのであれば、RPhylib パッケージの**最尤法**を使って主成分を探索することができます。そのためには Phylip パッケージをインストールする必要があります。これらの樹形図はすべて、**ggtree** にインポートしてカスタマイズしたり、**ape** パッケージで構造が系統的かどうかを分析したりすることができます。

## ## GMM Procedure 9: Constructing Mean Shapes / GMM の実行 9：平均的な「かたち」の構築

If we wish, we can retrieve mean shapes for a provided factor (e.g. “Archaeological\_Unit”), using the elliptic Fourier coefficients or PCA scores. This is done through the **Momocs::MSHAPES()** function with the object first being made.

もし必要ならば、楕円フーリエ係数または PCA スコアを用いて「考古学的文化」のような因子として平均的「かたち」を取得することもできます。これまでに作成したオブジェクトを用いて、**Momocs::MSHAPES()** 関数を使用できます。

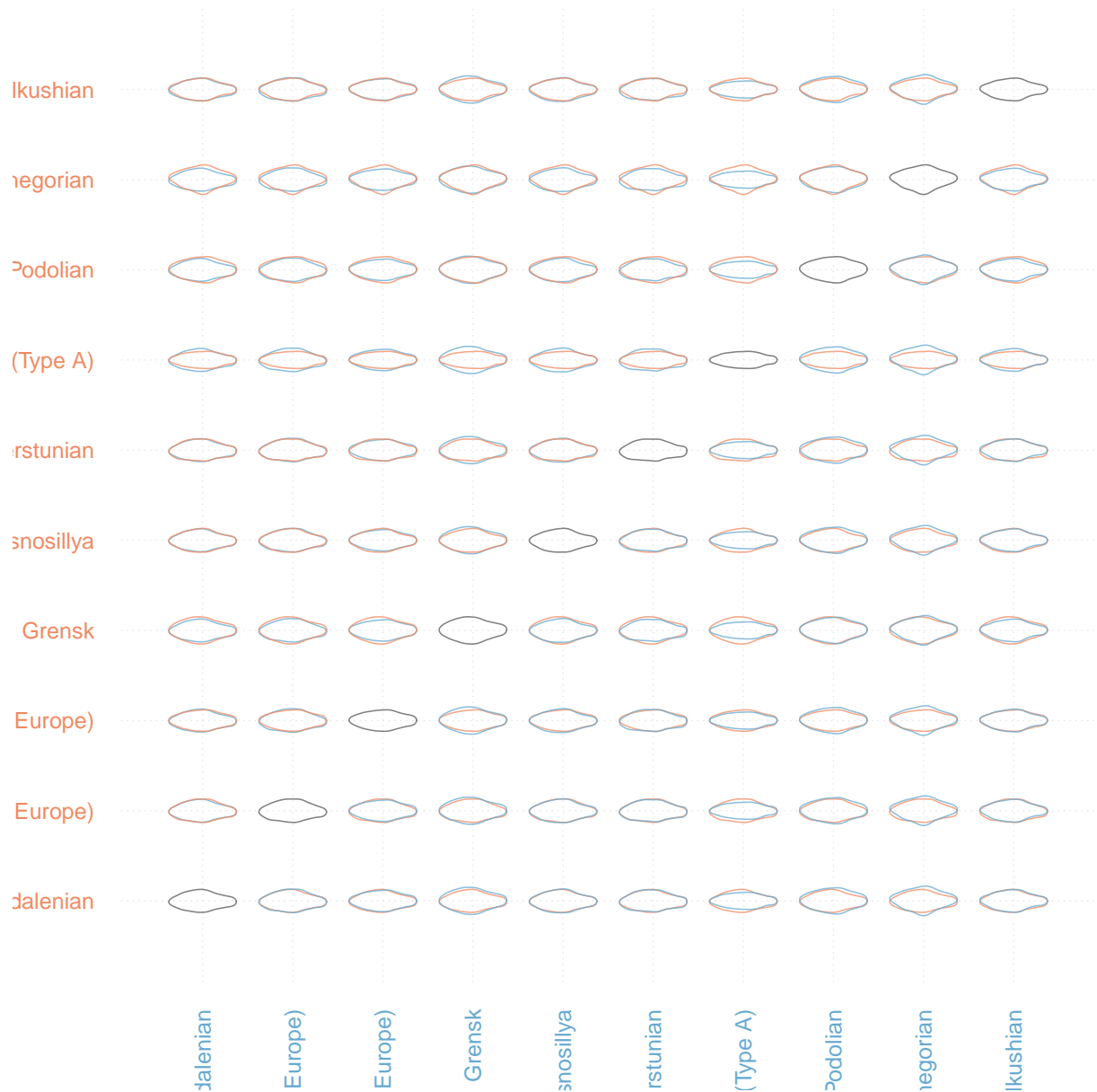
```
meanshapes <- MSHAPES(efashape, ~Archaeological_Unit) # efashape と「考古学的文化」により平均的「かたち」
```

The **Momocs::plot\_MSHAPES()** function is particularly useful for displaying the mean shapes for all the archaeological units and the visualisation of different configurations of mean shapes.

**Momocs::plot\_MSHAPES()** 関数は、すべての考古学的文化の平均的「かたち」の表示と、平均的「かたち」の

異なる構成の視覚化に便利です。

```
plot_MSHAPES(meanshapes, size = 0.75) # meanshapes を描画, 左下～右上中央が個別の「考古学的文化」の剥片尖頭
```



## GMM Procedure 10: Further Work: Incorporating Size...

In this example we have so far only examined shape, however we still have size data (as all images have a scale). From this we can extract various different measures including length or symmetry through the `Momocs::coo_length()` option (noting that the converted value is pixels!).

ここまで「かたち」のみを調べてきましたが、すべての元画像にはスケールがあるので、データセットにはサイズもあります。このデータから、`Momocs::coo_length()` オプションを使って、長さや対称性を含む様々な尺度を抽出することができます（ただし変換された値はピクセル単位であることに注意！）。

Centroid size is perhaps the best measure of size, incorporating the distance from all points of interest in relation to the shape. This can be extracted from the original shape data, using the `Momocs::coo_centsize()` function. We can then take this data and the principal component scores, and merge them into one database. There are a variety of ways this can be done, this is just one example.

セントロイド・サイズは、おそらく最も優れたサイズの尺度であり、形状に関連する注目すべき点からの距離が組み込まれています。`Momocs::coo_centsize()` 関数を用いて、オリジナルの「かたち」データから抽出することが可能です。このサイズのデータと主成分スコアを取り出して、一つのデータベースに統合することができます。これには様々な方法がありますが、ここでは一例を示します。

```
centroidsize <- as_tibble(coo_centsize(shape)) # shape の外形のセントロイド・サイズを行列にして centroidsize
centroidsize <- rename(centroidsize, cs = "value") # データ行列内の cs 列を数値型に変換
pcascores <- as_tibble(pcashape$x) # pcashape の x を行列にして pcascores に格納
databasedata <- cbind(database, centroidsize, pcascores) # database, centroidsize, pcascores を連結
head(databasedata) # databasedata の先頭 6 行を表示
```

| ##   | ID | Site        | Context     | Longitude | Latitude | Country   | Archaeological_Unit |
|------|----|-------------|-------------|-----------|----------|-----------|---------------------|
| ## 1 | 1  | Baltašiškės | Baltašiškės | 23.99     | 54.03    | Lithuania | Baltic Magdalenian  |
| ## 2 | 2  | Barouka     | Barouka     | 30.26     | 53.51    | Belarus   | Grensk              |
| ## 3 | 3  | Barouka     | Barouka     | 30.26     | 53.51    | Belarus   | Grensk              |
| ## 4 | 4  | Barouka     | Barouka     | 30.26     | 53.51    | Belarus   | Grensk              |
| ## 5 | 5  | Barouka     | Barouka     | 30.26     | 53.51    | Belarus   | Grensk              |
| ## 6 | 6  | Barouka     | Barouka     | 30.26     | 53.51    | Belarus   | Grensk              |

| ##   | File_Name | Reference  | Notes |
|------|-----------|--|-------|
| ## 1 | Bal.1     | Rimantiene 1971; 1974                              | <NA>  |
| ## 2 | Bor.1     | Kolasau 2018 (Krasnosillya by Zalyzniak (Borovka)) |       |
| ## 3 | Bor.3     | Kolasau 2018 (Krasnosillya by Zalyzniak (Borovka)) |       |
| ## 4 | Bor.4     | Kolasau 2018 (Krasnosillya by Zalyzniak (Borovka)) |       |
| ## 5 | Bor.5     | Kolasau 2018 (Krasnosillya by Zalyzniak (Borovka)) |       |
| ## 6 | Bor.6     | Kolasau 2018 (Krasnosillya by Zalyzniak (Borovka)) |       |

| ##   | cs       | PC1         | PC2           | PC3          | PC4         | PC5           |
|------|----------|-------------|---------------|--------------|-------------|---------------|
| ## 1 | 1.000807 | 0.02723891  | 0.0580419170  | -0.018821385 | 0.01075311  | -0.0362432286 |
| ## 2 | 1.000206 | -0.09857714 | -0.0005637827 | 0.048457316  | -0.00511936 | 0.0053152139  |
| ## 3 | 1.000258 | 0.07921519  | 0.0482961705  | 0.035729991  | 0.01293937  | -0.0002026685 |
| ## 4 | 1.000276 | 0.05615152  | 0.0292043739  | -0.006578647 | 0.01480096  | 0.0095477862  |

```

## 5 1.000237 0.15608400 0.0143529019 0.040597347 -0.02173636 0.0343686118
## 6 1.000253 -0.07456266 0.0133997919 -0.028260524 -0.02272342 0.0165848492
##          PC6          PC7          PC8          PC9          PC10          PC11
## 1 0.04039065 0.002221293 -0.003887411 0.0141624637 0.011570382 -0.001929899
## 2 -0.01360586 -0.001471208 0.006836667 -0.0219516587 -0.008226835 -0.001947317
## 3 -0.02995656 0.027029617 0.029296167 0.0112590246 0.004525287 -0.006822262
## 4 0.03045869 0.030115081 0.001359413 -0.0004625243 0.022532871 0.013166032
## 5 -0.05445377 0.006164025 0.007580372 -0.0002689392 -0.012458161 0.001668338
## 6 -0.01298230 -0.008691961 0.009880353 0.0067560334 0.003571735 0.020928347
##          PC12          PC13          PC14          PC15          PC16
## 1 -0.012675160 0.008215481 -0.018238173 -0.004578046 -0.0046666536
## 2 0.006375108 -0.004841490 0.001690275 0.007453436 -0.0025290329
## 3 0.010398928 0.005558982 0.009781070 -0.004088408 0.0001736873
## 4 0.001044667 -0.007329051 0.009168578 -0.009269989 0.0032797327
## 5 0.009341902 0.010488130 0.006770371 -0.004617734 0.0034287629
## 6 -0.005352459 -0.004258540 -0.009235560 -0.005278896 0.0035957133
##          PC17          PC18          PC19          PC20          PC21
## 1 -0.001487122 -8.519330e-04 -0.006141183 0.0025277633 0.0057051817
## 2 -0.004681512 -1.978794e-03 -0.005500646 0.0032431352 -0.0026323367
## 3 -0.003117830 -1.777459e-03 0.001190171 -0.0003838019 -0.0002687512
## 4 0.003038680 -8.907413e-05 0.001670598 -0.0048705880 -0.0007977494
## 5 -0.003071278 -6.090379e-03 -0.002698053 0.0005167888 -0.0038296692
## 6 0.001855152 2.056026e-03 0.003133939 0.0082911548 0.0029166073
##          PC22          PC23          PC24          PC25          PC26
## 1 -0.002199005 0.0031339125 0.0005578041 0.0057015579 -0.0000918038
## 2 0.002709378 0.0006669588 0.0011875057 0.0015199619 -0.0016203800
## 3 -0.003902373 0.0066457039 -0.0007067034 0.0015648038 -0.0023431863
## 4 0.001760109 -0.0005884933 0.0045755061 0.0030173512 -0.0034482293
## 5 -0.011159395 0.0008422671 0.0021378488 0.0046062839 -0.0024967472
## 6 0.002358739 0.0051735558 0.0042347471 0.0008474177 0.0031718262
##          PC27          PC28          PC29          PC30          PC31
## 1 -0.0071442912 0.0001140081 -1.975473e-03 0.0028551797 -0.002676492
## 2 -0.0008167792 0.0008976529 3.809636e-05 0.0010231558 -0.001679553
## 3 -0.0011043431 0.0006605308 -4.871148e-03 0.0038723982 0.006917407
## 4 0.0039970805 0.0021537871 -2.660179e-03 0.0009938206 0.001471461
## 5 -0.0072786490 0.0010303600 -4.185127e-04 -0.0013743188 0.003132933
## 6 0.0006651295 0.0032590344 -1.864137e-04 -0.0008202582 0.001806728
##          PC32          PC33          PC34          PC35          PC36
## 1 -0.0003583197 -0.0013943507 -0.0035560834 0.0011660315 1.413841e-03
## 2 0.0007690033 -0.0017011470 0.0014070157 0.0013656916 6.617025e-05

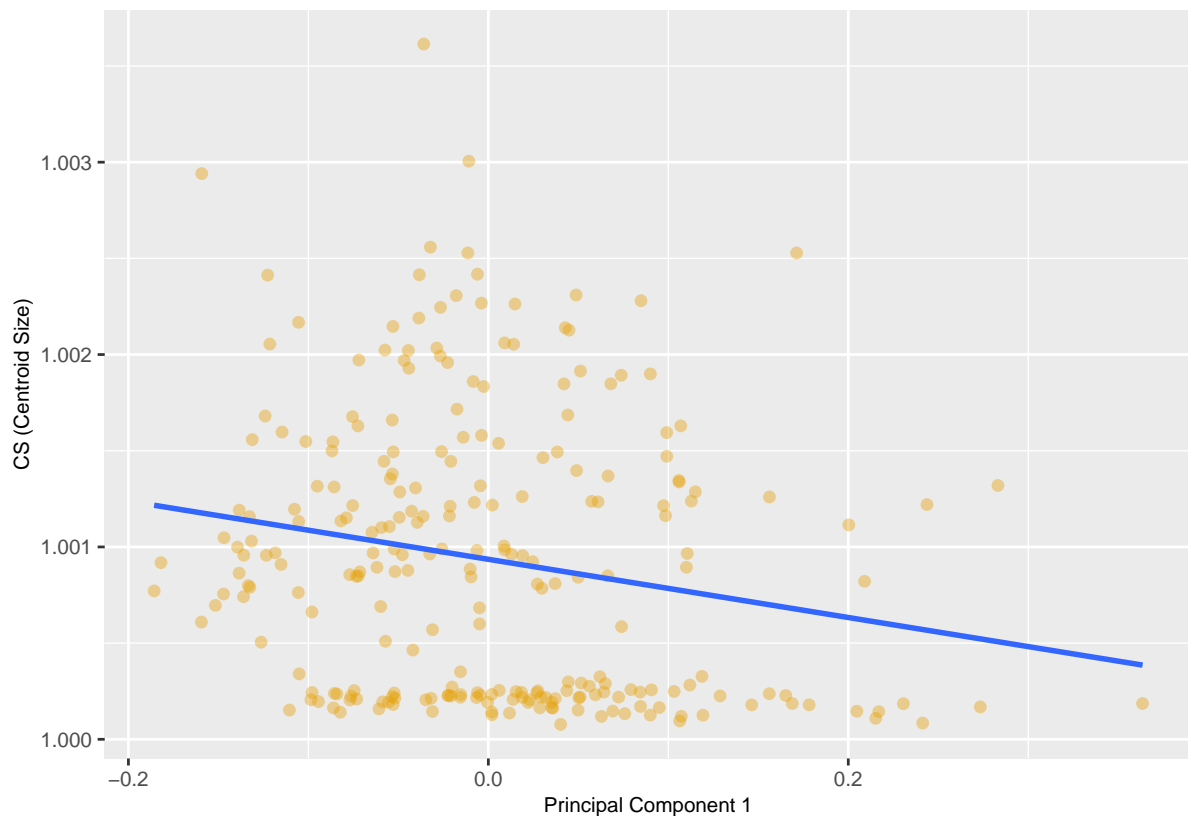
```

```
## 3 -0.0005061710 0.0006146251 0.0011405345 0.0005078773 7.766481e-04
## 4 0.0024797067 -0.0003838746 0.0013412707 -0.0010007059 2.011609e-03
## 5 -0.0049452251 0.0014702001 -0.0032072038 0.0004185099 3.483295e-04
## 6 -0.0014203187 0.0024718622 0.0009630608 0.0013234195 -3.258149e-04
##          PC37          PC38          PC39          PC40          PC41
## 1 7.546544e-04 -0.0031047438 -1.617456e-03 2.178502e-03 1.554697e-03
## 2 1.987449e-04 0.0011389019 1.250215e-03 1.028497e-03 5.193192e-06
## 3 2.977223e-04 0.0004859952 6.423081e-04 -1.728807e-03 -6.067506e-04
## 4 5.702769e-04 0.0006614252 5.309941e-05 -2.063167e-04 -6.627334e-04
## 5 8.492301e-05 0.0013030088 6.746118e-04 -3.132884e-05 -7.200236e-04
## 6 -3.118316e-04 -0.0021313201 -1.226630e-03 1.109502e-03 -3.353373e-04
##          PC42          PC43          PC44
## 1 -7.011628e-17 -5.429479e-18 1.541621e-17
## 2 -1.060826e-16 1.264940e-17 -1.332717e-17
## 3 1.778452e-16 -1.660799e-17 1.308076e-17
## 4 -2.005661e-16 -1.539559e-17 1.441835e-18
## 5 9.469562e-17 -2.072481e-17 1.801843e-17
## 6 -8.611479e-17 3.961242e-18 -1.640800e-17
```

We can now explore these through regression and correlation based analyses. For example, using the ggplot functions we can create a scatter plot and add a regression line (these functions will be detailed by Prof. Ben Marwick in a forthcoming workshop).

さらに相関-回帰分析も実施できます。例えば、ggplot 関数を使って散布図を作成したり、回帰直線を追加したりすることができます。これらの関数については、Ben Marwick 教授が今後のワークショップで詳しく説明します。

```
ggplot(databasedata, aes(PC1, cs)) + geom_point(size = 2, pch = 16, alpha = 0.4, colour = "#E69F00",
```



We can then perform a correlation (and test) using the `cor` and `cor.test` functions:

次に、`cor` `cor.test` 関数を用いて相関の確認と検証を行なえます。

```
cor(databasedata$PC1, databasedata$cs) #PC1 とセントロイドサイズの相関
```

```
## [1] -0.1913964
```

```
cor.test(databasedata$PC1, databasedata$cs) # 同上を検証
```

```
##
## Pearson's product-moment correlation
##
## data:  databasedata$PC1 and databasedata$cs
## t = -3.0709, df = 248, p-value = 0.002372
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
## -0.30814620 -0.06896712
## sample estimates:
##      cor
```

## -0.1913964

## Concluding Remarks / まとめ

This workshop was designed to highlight how geometric morphometrics (outline analysis) can be examined for archaeological material in the R Environment, from data importing to visualisation and analysis. It is worth stressing that Momocs is only one of a number of packages in the R Environment, and the methods showcased here are only one way (and one style) of conducting GMM. Landmark analysis can be incorporated into Momocs but there are a number of functions in **Geomorph** which are particularly impressive and powerful. Similarly, a number of other packages have been referenced throughout this workflow. Only through exploring the packages and their functions will you be able understand what workflow works best for your research question and process.

このワークショップでは、データのインポートから可視化、解析まで、Rを使用した考古学資料の幾何形態測定学(アウトライン分析)をどのように行なうのかを紹介しました。ただし、ここで紹介した方法は、GMMを実施するための一つの方法(そして一つのスタイル)に過ぎないことを強調しておきたいと思います。ランドマーク分析はMomocsに組み込むことができますが、**Geomorph**にも印象的で強力な関数がいくつかあります。また、今回のワークフローを通して、Momocs以外のパッケージも参照しました。これらのパッケージとその機能を探求することで、あなたの研究課題や研究プロセスに最も適しているワークフローを見つけ出すことができるでしょう。

If there are any questions please feel free to contact me: C.S.Hoggard@soton.ac.uk

もし何か質問があれば、C.S.Hoggard@soton.ac.uk

## References / 参考文献 ##

For literature pertaining to outline analysis see / アウトライン分析について:

- \* Claude, J. (2008). *Morphometrics with R*. Springer Publishing.
- \* Bonhomme, V., Picq, S., Gaucherel, C., & Claude, J. (2014). Momocs: Outline analysis using R. *Journal of Statistical Software*, 56: 1-24.
- \* Caple, J., Byrd, J., & Stephan, C. N. (2017). Elliptical Fourier analysis: Fundamentals, applications, and value for forensic anthropology. *International Journal of Legal Medicine*, 131 (6): 1675-1690.
- \* Ferson, S., Rohlf, F. J., & Koehn, R. K. (1985). Measuring shape variation of two-dimensional outlines. *Systematic Zoology*, 34 (1): 59-68.
- \* Kuhl, F. P., & Giardina, C. R. (1982). Elliptic Fourier features of a closed contour. *Computer Graphics and Image Processing*, 18 (3): 236-258.
- \* Yoshioka, Y. (2004). Analysis of petal shape variation of *Primula sieboldii* by elliptic fourier descriptors and principal component analysis. *Annals of Botany*, 94 (5), 657-664.
- \* Zahn, C. T., & Roskies, R. Z. (1972). Fourier descriptors for plane closed curves. *IEEE Transactions on Computers*, C-21 (3): 269-281.

For literature pertaining to GMM by the author (including code and data) see / 講師の業績:



- \* Hoggard, C.S., Lauridsen, L. and Witte, K.B. (2019). The Potential of Geometric Morphometrics for Danish Archaeology: Two Case Studies. *Arkaeologisk Forum*, 40: 30-42. (<http://www.archaeology.dk/16738/Nr.%2040%20-%202019>). OSF: <https://osf.io/en5d2/>.
- \* Hoggard, C.S., McNabb, J. and Cole, J.N. (2019). The application of elliptic Fourier analysis in understanding biface shape and symmetry through the British Acheulean. *Journal of Paleolithic Archaeology*, 2 (2): 115-133. (<https://doi.org/10.1007/s41982-019-00024-6>). OSF: <https://osf.io/td92j/>.
- \* Ivanovaite, L., Swertka, K., Hoggard, C.S., Sauer, F. and Riede, F. (2020). All these fantastic cultures? Research history and regionalisation in the Late Palaeolithic tanged point cultures of Eastern Europe. *European Journal of Archaeology*. (<https://doi.org/10.1017/eea.2019.59>). OSF: <https://osf.io/agrwb/>.
- \* Vestergaard, C. and Hoggard, C.S. (2019). A Novel Geometric Morphometric (GMM) Application to the Study of Bronze Age Tutuli. *Danish Journal of Archaeology*, 8: 5-28. (<https://tidsskrift.dk/dja/article/view/112494/164318>). OSF: <https://osf.io/fcp43/>.
- \* Riede, F., Hoggard, C.S. and Shennan, S. (2019). Reconciling material cultures in archaeology with genetic data requires robust cultural evolutionary taxonomies. *Nature: Palgrave Communications*, 5 (1): 55. (<https://doi.org/10.1057/s41599-019-0260-7>). OSF: <https://osf.io/vtdf2/>.