

Diversity Calculator

Author: Klaus D. Goepel

Overview

The diversity excel template allows to calculate α -, β - and γ -diversity for a set samples (input data), and analyze similarities between the samples based on β -diversity.

The template works under Windows OS and Excel 2010 (xlsx extension). No macros or links to external workbooks are necessary. The workbook consists of an input worksheet for data samples, a calculation worksheet, where all necessary calculations are done, and a worksheet "beta" displaying the results.

Limitations

- Maximal number of data samples: 24
- Maximal number of categories/classes per data sample: 20

Results

Following data will be displayed:

For all samples:

- a) **Shannon Entropy H** (natural logarithm) α -, β - and γ , and corresponding Hill numbers 1D

| Diversity (Shannon - ln) | | 1D |
|--------------------------|-------|---------|
| α -Diversity | 1,557 | 4,745 |
| γ -Diversity | 1,775 | 5,903 |
| β -Diversity | 0,218 | 1,244 |

- b) **Homogeneity Measures**

| Homogeneity Measures | |
|--------------------------|--------|
| MacArthur M | 0,8038 |
| rel. Homogeneity 1S | 78,9% |
| AHP consensus S^* | 61,1% |

Mac Arthur $M = 1/{}^1D_{\beta}$ homogeneity indicator

relative Homogeneity ${}^1S = (M - 1/N) / (1 - 1/N)$; N number of samples (between 0 and 100%)

AHP group consensus S^* : relative homogeneity corrected by AHP scale limitations

- c) **AHP max γ and min β true diversity based on the max score of the AHP scale**

| AHP | |
|-------------------|------|
| scale max | 9 |
| D_{γ} -max | 6,88 |
| D_{α} -min | 3,41 |

Table 1: Shannon α -Entropy, Equitability, Simpson Dominance, Gini-Simpson index and Hill numbers

| Sample | No | Shannon α -Entropy | Equi- tability | Simpson Dominance | Gini- Simpson | Hill Numbers | |
|----------|----|------------------------------|-------------------|----------------------|------------------|--------------|---------|
| | | | | | | 1D | 2D |
| sample-1 | 1 | 1,6216 | 91% | 0,2270 | 77,3% | 5,06 | 4,41 |
| sample-2 | 2 | 1,4544 | 81% | 0,2873 | 71,3% | 4,28 | 3,48 |
| sample-3 | 3 | 1,6684 | 93% | 0,2022 | 79,8% | 5,30 | 4,95 |
| sample-4 | 4 | 1,5728 | 88% | 0,2309 | 76,9% | 4,82 | 4,33 |

Table 2: will show the top 20 pairs of most similar samples

| No | S | Sample | | α | β |
|----|------|----------|-----------|----------|---------|
| | | (1) | (2) | | |
| 1 | 100% | sample-4 | sample-10 | 1,63 | 1,63 |
| 2 | 97% | sample-7 | sample-14 | 1,47 | 1,47 |
| 3 | 96% | sample-9 | sample-13 | 1,45 | 1,45 |
| 4 | 96% | sample-3 | sample-14 | 1,64 | 1,64 |

On page 2 the matrix of pairs of data samples is displayed. It shows either α -, β - or γ -diversity (Shannon) of each pair of data samples, Horn Index or AHP consensus indicator.

| | | 2005 | 2006 | Q3 2007 | Q1 2008 | Q4 2007 | Q2 2008 | Q2 2009 | Q3 2008 | Q1 2009 | Q4 2008 | Q3 2009 | Q4 2009 | Q1 2010 | Q2 2010 | Q3 2010 | Q4 2010 | Q1 2011 | Q2 2011 | Q3 2011 | Q4 2011 | Q4 2012 | Q1 2012 | Q3 2012 | Q2 2012 | |
|--------------------------------|------|------|------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|------|
| Shannon α -Diversity | 1 | 1,55 | 1,34 | 1,36 | 1,49 | 1,56 | 1,55 | 1,53 | 1,56 | 1,57 | 1,57 | 1,6 | 1,59 | 1,59 | 1,58 | 1,57 | 1,59 | 1,58 | 1,57 | 1,56 | 1,6 | 1,6 | 1,59 | 1,6 | 1,57 | 1,55 |
| | 4,72 | 3,8 | 3,9 | 4,4 | 4,8 | 4,7 | 4,6 | 4,8 | 4,8 | 4,8 | 4,8 | 4,9 | 4,9 | 4,9 | 4,8 | 4,8 | 4,9 | 4,8 | 4,8 | 4,7 | 5,0 | 5,0 | 4,9 | 4,9 | 4,8 | 4,7 |
| 2005 | 1 | 0,00 | 0,00 | 0,04 | 0,05 | 0,06 | 0,06 | 0,08 | 0,08 | 0,07 | 0,07 | 0,09 | 0,09 | 0,11 | 0,12 | 0,11 | 0,12 | 0,13 | 0,14 | 0,10 | 0,10 | 0,11 | 0,10 | 0,12 | 0,14 | |
| 2006 | 2 | 0,00 | 0,00 | 0,03 | 0,05 | 0,05 | 0,06 | 0,08 | 0,07 | 0,07 | 0,07 | 0,09 | 0,08 | 0,10 | 0,11 | 0,11 | 0,12 | 0,12 | 0,13 | 0,10 | 0,10 | 0,10 | 0,10 | 0,12 | 0,14 | |
| Q3 2007 | 3 | 0,04 | 0,03 | 0,00 | 0,01 | 0,01 | 0,02 | 0,03 | 0,03 | 0,03 | 0,03 | 0,03 | 0,03 | 0,04 | 0,04 | 0,05 | 0,05 | 0,06 | 0,06 | 0,05 | 0,05 | 0,06 | 0,05 | 0,07 | 0,08 | |
| Q1 2008 | 5 | 0,05 | 0,05 | 0,01 | 0,00 | 0,00 | 0,00 | 0,01 | 0,01 | 0,01 | 0,00 | 0,01 | 0,01 | 0,01 | 0,02 | 0,02 | 0,02 | 0,03 | 0,02 | 0,02 | 0,03 | 0,03 | 0,03 | 0,04 | 0,04 | |
| Q4 2007 | 4 | 0,06 | 0,05 | 0,01 | 0,00 | 0,00 | 0,00 | 0,01 | 0,01 | 0,01 | 0,01 | 0,01 | 0,01 | 0,01 | 0,01 | 0,02 | 0,02 | 0,02 | 0,03 | 0,02 | 0,02 | 0,03 | 0,03 | 0,05 | 0,04 | |
| Q2 2008 | 6 | 0,06 | 0,06 | 0,02 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,01 | 0,01 | 0,01 | 0,02 | 0,02 | 0,02 | 0,02 | 0,02 | 0,03 | 0,02 | 0,03 | 0,04 | 0,04 | 0,05 | 0,04 | |
| Q2 2009 | 10 | 0,08 | 0,08 | 0,03 | 0,01 | 0,01 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,01 | 0,01 | 0,01 | 0,01 | 0,01 | 0,01 | 0,01 | 0,02 | 0,01 | 0,02 | 0,02 | 0,03 | 0,04 | 0,02 | |
| Q3 2008 | 7 | 0,08 | 0,07 | 0,03 | 0,01 | 0,01 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,01 | 0,01 | 0,01 | 0,01 | 0,01 | 0,01 | 0,02 | 0,01 | 0,01 | 0,02 | 0,02 | 0,03 | 0,02 | |
| Q1 2009 | 9 | 0,07 | 0,07 | 0,03 | 0,01 | 0,01 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,01 | 0,01 | 0,01 | 0,01 | 0,01 | 0,01 | 0,01 | 0,02 | 0,01 | 0,02 | 0,02 | 0,02 | 0,03 | 0,02 | |
| Q4 2008 | 8 | 0,07 | 0,07 | 0,03 | 0,00 | 0,01 | 0,01 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,01 | 0,01 | 0,01 | 0,01 | 0,01 | 0,02 | 0,01 | 0,01 | 0,01 | 0,01 | 0,02 | 0,01 | |
| Q3 2009 | 11 | 0,09 | 0,09 | 0,03 | 0,01 | 0,01 | 0,01 | 0,01 | 0,00 | 0,01 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,01 | 0,01 | 0,01 | 0,01 | 0,01 | 0,01 | 0,02 | 0,01 | |
| Q4 2009 | 12 | 0,09 | 0,08 | 0,03 | 0,01 | 0,01 | 0,01 | 0,01 | 0,01 | 0,01 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,01 | 0,01 | 0,00 | 0,00 | 0,01 | 0,01 | 0,02 | 0,02 | |
| Q1 2010 | 13 | 0,11 | 0,10 | 0,04 | 0,01 | 0,01 | 0,02 | 0,01 | 0,01 | 0,01 | 0,01 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,01 | 0,01 | 0,02 | 0,02 | 0,03 | 0,01 | |
| Q2 2010 | 14 | 0,12 | 0,11 | 0,04 | 0,02 | 0,01 | 0,02 | 0,01 | 0,01 | 0,01 | 0,01 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,01 | 0,01 | 0,02 | 0,02 | 0,03 | 0,01 | |
| Q3 2010 | 15 | 0,11 | 0,11 | 0,05 | 0,02 | 0,02 | 0,02 | 0,01 | 0,01 | 0,01 | 0,01 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,01 | 0,01 | 0,02 | 0,01 | |
| Q4 2010 | 16 | 0,12 | 0,12 | 0,05 | 0,02 | 0,02 | 0,02 | 0,01 | 0,01 | 0,01 | 0,01 | 0,00 | 0,01 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,01 | 0,01 | 0,02 | 0,02 | 0,01 | |
| Q1 2011 | 17 | 0,13 | 0,12 | 0,06 | 0,02 | 0,02 | 0,02 | 0,01 | 0,01 | 0,01 | 0,01 | 0,01 | 0,01 | 0,01 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,01 | 0,01 | 0,02 | 0,02 | 0,01 | |
| Q2 2011 | 18 | 0,14 | 0,13 | 0,06 | 0,03 | 0,03 | 0,03 | 0,02 | 0,02 | 0,02 | 0,02 | 0,01 | 0,01 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,01 | 0,01 | 0,02 | 0,02 | 0,03 | 0,01 | |
| Q3 2011 | 19 | 0,10 | 0,10 | 0,05 | 0,02 | 0,02 | 0,02 | 0,01 | 0,01 | 0,01 | 0,01 | 0,00 | 0,00 | 0,01 | 0,01 | 0,01 | 0,01 | 0,01 | 0,00 | 0,00 | 0,00 | 0,00 | 0,01 | 0,01 | 0,01 | |
| Q4 2011 | 20 | 0,10 | 0,10 | 0,05 | 0,02 | 0,02 | 0,03 | 0,02 | 0,01 | 0,02 | 0,01 | 0,01 | 0,00 | 0,01 | 0,01 | 0,00 | 0,01 | 0,01 | 0,01 | 0,01 | 0,00 | 0,00 | 0,00 | 0,01 | 0,01 | |
| Q4 2012 | 24 | 0,11 | 0,10 | 0,06 | 0,03 | 0,03 | 0,04 | 0,02 | 0,02 | 0,02 | 0,01 | 0,01 | 0,01 | 0,02 | 0,02 | 0,01 | 0,01 | 0,01 | 0,02 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,01 | |
| Q1 2012 | 21 | 0,10 | 0,10 | 0,05 | 0,03 | 0,03 | 0,04 | 0,03 | 0,02 | 0,02 | 0,01 | 0,01 | 0,01 | 0,02 | 0,02 | 0,01 | 0,02 | 0,02 | 0,02 | 0,01 | 0,00 | 0,00 | 0,00 | 0,00 | 0,02 | |
| Q3 2012 | 23 | 0,12 | 0,12 | 0,07 | 0,04 | 0,05 | 0,05 | 0,04 | 0,03 | 0,03 | 0,02 | 0,02 | 0,02 | 0,03 | 0,03 | 0,02 | 0,02 | 0,03 | 0,03 | 0,01 | 0,01 | 0,00 | 0,00 | 0,00 | 0,02 | |
| Q2 2012 | 22 | 0,14 | 0,14 | 0,08 | 0,04 | 0,04 | 0,04 | 0,02 | 0,02 | 0,02 | 0,02 | 0,02 | 0,01 | 0,02 | 0,01 | 0,01 | 0,01 | 0,01 | 0,01 | 0,01 | 0,01 | 0,01 | 0,02 | 0,02 | 0,00 | |
| Column Average | | 1,1 | 1,1 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | |
| exp | | 1,1 | 1,1 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | |

On the last page you will find the three diagrams.

Diagram 1 showing Gini-Simpson index and Shannon Equitability

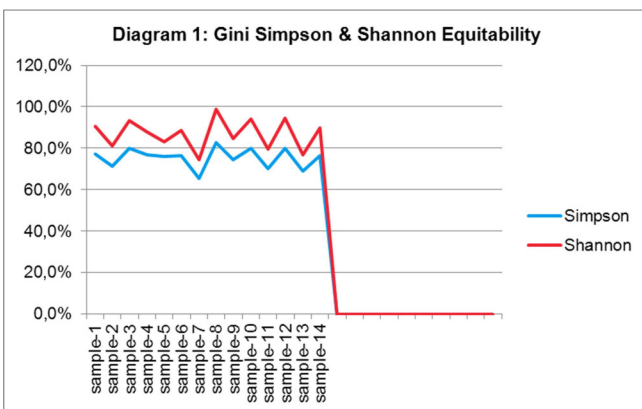
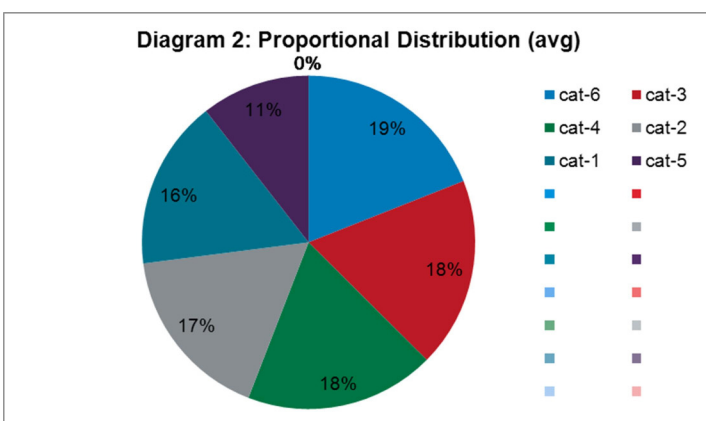


Diagram 2 showing the average proportional distribution for all classes/categories:



2. Select data sheet (input name of worksheet containing input data table)

Worksheet

Parameter Selections:

AHP

AHP
scale max

If your input data are priority distributions based on AHP (Analytic Hierarchy process), you can specify the maximum score of the AHP scale used. Standard is the linear 1-9 scale with scale max =9

Partitioning (Matrix)

Weight: Sort: Parameter: **Horn**

Weight (future extension)

- 0: all samples same weight
- 1: weighted according weights given in input sheet

Sort

- 0: unsorted Sequence of samples as given in the input sheet
- 1: sorted by α -diversity From max. to min alpha diversity (left to right)
- 2: sorted by β -diversity From max. to min beta diversity (left to right)
- 3: manual Manual as given by sequence in column s-sort of input sheet
- 4: cluster Clustered: Arranged by consecutive similarity
The second parameter determines the number of the first sample to be displayed in the matrix

Parameter (selects parameter displayed in matrix for all pairs of classes/categories):

- 25 – alpha Shannon α -diversity
- 26 – gamma Shannon γ -diversity
- 27 – beta Shannon β -diversity
- 28 – Horn index
- 29 – AHP consensus indicator

Clustering

To find clusters of similar data samples use Sort parameter “4” and vary the first sample to be displayed in the matrix, until you see clusters of low beta or high similarity along the matrix diagonal.

Example of clusters with high similarity

| | | | | | | |
|------|------|------|------|------|------|------|
| | 1,00 | 0,99 | 0,99 | 0,99 | 0,98 | 0,98 |
| 1,00 | | 1,00 | 1,00 | 1,00 | 0,97 | 0,97 |
| 0,99 | 1,00 | | 1,00 | 1,00 | 0,96 | 0,96 |
| 0,99 | 1,00 | 1,00 | | 1,00 | 0,96 | 0,96 |
| 0,99 | 1,00 | 1,00 | 1,00 | | 0,96 | 0,96 |
| 0,98 | 0,97 | 0,96 | 0,96 | 0,96 | | 1,00 |
| 0,98 | 0,97 | 0,96 | 0,96 | 0,96 | 1,00 | |

Please make a reference to the author, when using the template in your work:

Further reading (AHP consensus index):

Goepel, Klaus D. (2013). Implementing the Analytic Hierarchy Process as a Standard Method for Multi-Criteria Decision Making In Corporate Enterprises – A New AHP Excel Template with Multiple Inputs, *Proceedings of the International Symposium on the Analytic Hierarchy Process 2013*

For questions, feedback, suggestions please contact the author

Annex Mathematical relations and formulas used

A. General

1. *Alpha Entropy* for N categories/classes ($N = 2$ to 20) and K samples ($K = 2$ to 24)

$$H_{\alpha} = -w_1 \sum_{i=1}^N p_{i1} \ln p_{i1} + -w_2 \sum_{i=1}^N p_{i2} \ln p_{i2} + \dots + w_K \sum_{i=1}^N p_{iK} \ln p_{iK}$$

with p_{ij} relative abundance (frequency, priority, share) of class i and sample j

and w_i statistical weights of samples; $\sum_{i=1}^K w_i = 1$

2. *Gamma Entropy* for K samples

$$H_{\gamma} = \sum_{i=1}^K - (w_1 p_{i1} + w_2 p_{i2} + \dots + w_K p_{iK}) \ln (w_1 p_{i1} + w_2 p_{i2} + \dots + w_K p_{iK})$$

3. *Beta Entropy*

$$H_{\beta} = H_{\gamma} - H_{\alpha}$$

4. *True Diversity Order 1*

$${}^1D_{\alpha,\beta,\gamma} = \exp H_{\alpha,\beta,\gamma}$$

B. Homogeneity measures

5. *MacArthur M*

$$M = 1/{}^1D_{\beta}$$

6. *Relative Homogeneity 1S for N classes*

$${}^1S = [M - 1/N] / [(1 - 1/N)]$$

7. *AHP consensus $S^* S^*$* = $[M - \exp(H_{\alpha \min}) / \exp(H_{\gamma \max})] / [1 - \exp(H_{\alpha \min}) / \exp(H_{\gamma \max})]$

with $H_{\alpha \min} = -\frac{M}{N+M-1} \ln(\frac{M}{N+M-1}) - (N-1) \frac{1}{N+M-1} \ln \frac{1}{N+M-1}$

and $H_{\gamma \max} = (N-K)(-\frac{1}{N+8} \ln \frac{1}{N+8}) - (\frac{M+K-1}{M+N-1}) \ln(\frac{1}{K} \frac{M+K-1}{M+N-1})$

N number of classes, K number of samples, M max. Score of AHP scale (lin. = 9)

C. In Table 1

1a *Shannon Equitability*

$$E_H = H_{\alpha} / \ln N$$

1b *Simpson Dominance*

$$\lambda_{k \text{ Simpson}} = \sum_{i=1}^N p_i^2$$

1c *Gini-Simpson Index*

$$\lambda^{-1} = 1/\lambda_{k \text{ Simpson}} = 1/\sum_{i=1}^N p_i^2$$

1d *Hill number 1st and 2nd order*

$${}^1D = \exp H_{\alpha}$$

$${}^2D = 1/\lambda_{k \text{ Simpson}}$$

D. Matrix

1. Alpha Entropy for N categories/classes ($N = 2$ to 20) and 2 samples k and l

$$H_{\alpha kl} = -w_k \sum_{i=1}^N p_{ik} \ln p_{ik} - w_l \sum_{i=1}^N p_{il} \ln p_{il}$$

$$H_{\alpha} = \frac{w_k H_{\alpha k} + w_l H_{\alpha l}}{w_k + w_l}$$

For diagonal $k = l$: $H_{\alpha} = H_{\alpha k}$

2. Gamma Entropy for k categories/classes 2 samples k and l

$$H_{\gamma kl} = -\sum_{i=1}^N \frac{w_k p_{ik} + w_l p_{il}}{w_k + w_l} \ln \frac{w_k p_{ik} + w_l p_{il}}{w_k + w_l}$$

For diagonal $k = l$

$$H_{\gamma kk} = -\sum_{i=1}^N p_{ik} \ln p_{ik} = H_{\alpha k}$$