# Everything You've Always Wanted to Know About Card Sorting

William Hudson

Email: william.hudson@syntagm.co.uk

Web Page: www.syntagm.co.uk/design/cardsort.shtml

# Card Sorting

❑ Card sorting allows us to establish how users think about the terminology we plan to use in menus, navigation and information architecture

❑ Simple procedure: participants are given items to organize into related groups

❑ Two types of card sort
- Open – participants make as many groups as they need and give them their own names
- Closed – numbers and names of groups are fixed in advance

❑ Extensions
- Allow participants to query terms, circle alternatives or provide their own
- Get participants to provide a "quality of fit" measure for each item in a second pass

# Card Sorting



| Item name | Fit |
|---|---|
| A line or two to provide missing context (optional) | ☐ Perfect |
| Alternatives: Other words that users might prefer (optional) | ☐ Good |
| | ☐ Fair |
| | 1 |

Sample Item Card with Quality of Fit Measure

# Card Sorting Activity

❑ Closed card sort
- Step 1: Make a note of your participant number
- Step 2: Sort the supplied cards into groups, with a group card on top of each
- Step 3: Return to each group and mark the appropriate "fit" value for each item card

❑ You should end up with a deck of cards as follows:
- Participant card, group card, item cards for group
- (Group and item cards repeated for each group)

❑ If you are not sure where to put an item card:
- Put it in the most likely group
- Or, if it really does not fit, leave it out altogether (keep it, the back is perfect for shopping lists!)

# Hierarchical Cluster Analysis

❑ Technically known as agglomerative monothetic clustering

- Start with small, two-item clusters and merge into larger clusters
- Single characteristic: similarity or proximity – the number of times two items appear in the same group

| | | |
|---|---|---|
| Chardonnay | Beaujolais | Cava |
| Riesling | Cabernet Sauvignon | Champagne |
| Muscat | Claret | |
| Pinot Grigio | Merlot | |
| White Zinfandel | Syrah | |

**Sample Card Sort Items Grouped by a Participant**

---

# Proximity Matrix for Single Participant

| | Beaujolais | Cabernet Sauvignon | Cava | Champagne | Chardonnay | Claret | Merlot | Muscat | Pinot Grigio | Riesling | Syrah | White Zinfandel |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Beaujolais | | 1 | | | | 1 | 1 | | | | 1 | |
| Cabernet Sauvignon | 1 | | | | | 1 | 1 | | | | 1 | |
| Cava | | | | 1 | | | | | | | | |
| Champagne | | | 1 | | | | | | | | | |
| Chardonnay | | | | | | | | 1 | 1 | 1 | | 1 |
| Claret | 1 | 1 | | | | | 1 | | | | 1 | |
| Merlot | 1 | 1 | | | | 1 | | | | | 1 | |
| Muscat | | | | | 1 | | | | 1 | 1 | | 1 |
| Pinot Grigio | | | | | 1 | | | 1 | | 1 | | 1 |
| Riesling | | | | | 1 | | | 1 | 1 | | | 1 |
| Syrah | 1 | 1 | | | | 1 | 1 | | | | | |
| White Zinfandel | | | | | 1 | | | 1 | 1 | 1 | | |

# Proximity Matrix for All Participants

| | Beaujolais | Cabernet Sauvignon | Cava | Champagne | Chardonnay | Claret | Merlot | Muscat | Pinot Grigio | Riesling | Syrah | White Zinfandel |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Beaujolais | | 9 | 3 | 2 | | 11 | 11 | 1 | 4 | 1 | 9 | |
| Cabernet Sauvignon | 9 | | 1 | 1 | 2 | 10 | 9 | 2 | 4 | 1 | 10 | 1 |
| Cava | 3 | 1 | | 9 | 2 | 3 | 1 | 4 | 1 | 1 | 1 | 2 |
| Champagne | 2 | 1 | 9 | | 1 | | 1 | 3 | 2 | 1 | | |
| Chardonnay | | 2 | 2 | 1 | | | 1 | 8 | 6 | 11 | 2 | 11 |
| Claret | 11 | 10 | 3 | | | | 10 | 1 | 5 | | 10 | 1 |
| Merlot | 11 | 9 | 1 | 1 | 1 | 10 | | | 4 | 2 | 10 | 1 |
| Muscat | 1 | 2 | 4 | 3 | 8 | 1 | | | 5 | 8 | 2 | 8 |
| Pinot Grigio | 4 | 4 | 1 | 2 | 6 | 5 | 4 | 5 | | 6 | 3 | 7 |
| Riesling | 1 | 1 | 1 | 1 | 11 | | 2 | 8 | 6 | | 2 | 12 |
| Syrah | 9 | 10 | 1 | | 2 | 10 | 10 | 2 | 3 | 2 | | 2 |
| White Zinfandel | | 1 | 2 | | 11 | 1 | 1 | 8 | 7 | 12 | 2 | |

# SynCaps Dendogram
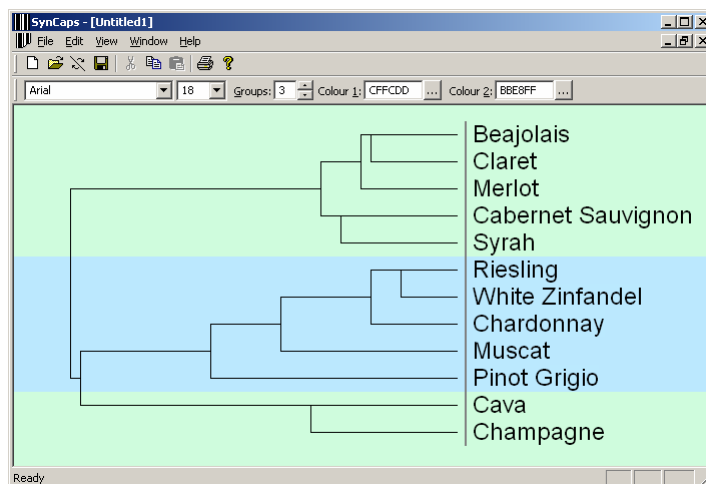
# Manual Cluster Analysis

1. List all pairs in descending score order
2. Work from the top of the list downwards (strictly in descending-score order)
3. Clusters are formed by merging pairs with items in common

| Riesling | 12 | White Zinfandel |
|---|---|---|
| Riesling | 11 | Chardonnay |
| White Zinfandel | 11 | Chardonnay |
| Beaujolais | 11 | Claret |
| Beaujolais | 11 | Merlot |
| Syrah | 10 | Claret |
| ... | | |

**Sample of the "Agglomeration Schedule"**

---

# Manual Cluster Analysis

4. Calculate scores:
   - ❑ For a orphan pairs the score is as shown
   - ❑ For a merging an item with a cluster (or merging two clusters), the score depends on the linkage method chosen: average, single (highest score), complete (lowest score)
   - ❑ Average works best in most cases

| Riesling | 12 | White Zinfandel |
|---|---|---|
| Riesling | 11 | Chardonnay |
| White Zinfandel | 11 | Chardonnay |
| Beaujolais | 11 | Claret |
| Beaujolais | 11 | Merlot |
| Syrah | 10 | Claret |
| ... | | |

**Agglomeration Schedule**

Chardonnay has a score of 11 with both other cluster members

| Riesling | |
|---|---|
| White Zinfandel | 12 |
| Chardonnay | 11 |

**Sample Cluster List**

# Manual Cluster Analysis

5. A score is likely to change when an item is merged into an existing cluster. Move the entry to the right place in the list. (This may mean changing earlier clusters.)

| Riesling | 12 | White Zinfandel |
|---|---|---|
| Riesling | 11 | Chardonnay |
| White Zinfandel | 11 | Chardonnay |
| Beaujolais | 11 | Claret |
| Beaujolais | 11 | Merlot |
| Syrah | 10 | Claret |
| ... | | |

**Agglomeration Schedule**

Syrah has a score of 11 with Merlot and Claret, but only 9 with Beaujolais

| Beaujolais | |
|---|---|
| Claret | 11 |
| Merlot | 10.5 |
| Syrah | 10.3 |

**Sample Cluster List**

# Dendograms (AKA Dendrograms)

❑ The scores correspond to the vertical lines in the dendogram
❑ Higher scores are closer to the labels (to the right in this case)
❑ Happily this doesn't normally have to be done by hand!

| Beaujolais | |
|---|---|
| Claret | 11 |
| Merlot | 10.5 |
| Syrah | 10.3 |
| Riesling | |
| White Zinfandel | 12 |
| Chardonnay | 11 |
| ... | |



Beajolais
Claret
Merlot
Cabernet Sauvignon
Syrah
Riesling
White Zinfandel
Chardonnay
Muscat
Pinot Grigio
Cava
Champagne
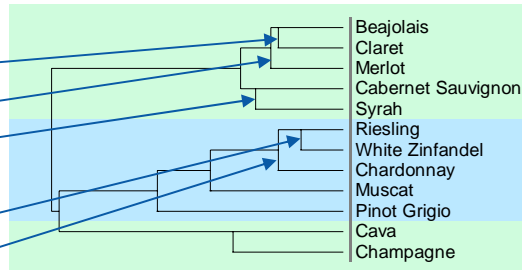
# Contour Map

❑ A simple technique that does not use cluster analysis or special software (just Excel or similar)

❑ Requires a proximity matrix – either created manually or output from a package like SynCaps

❑ Two-step process
   1. Create a contour map
   2. Move related rows and columns in the matrix closer together
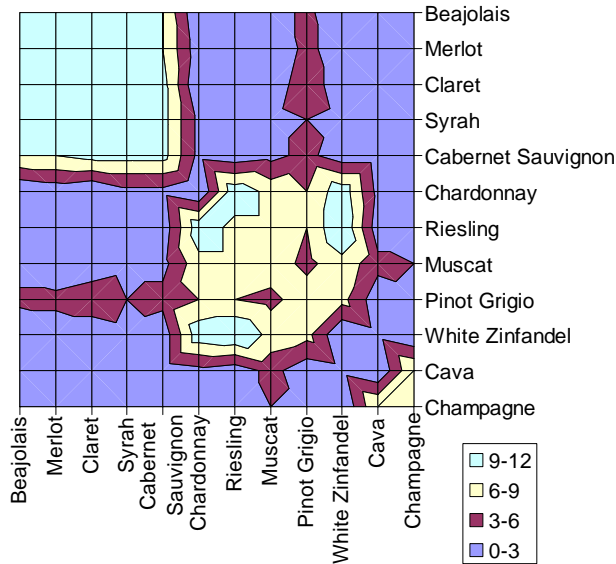
# Initial Contour Map (Excel)
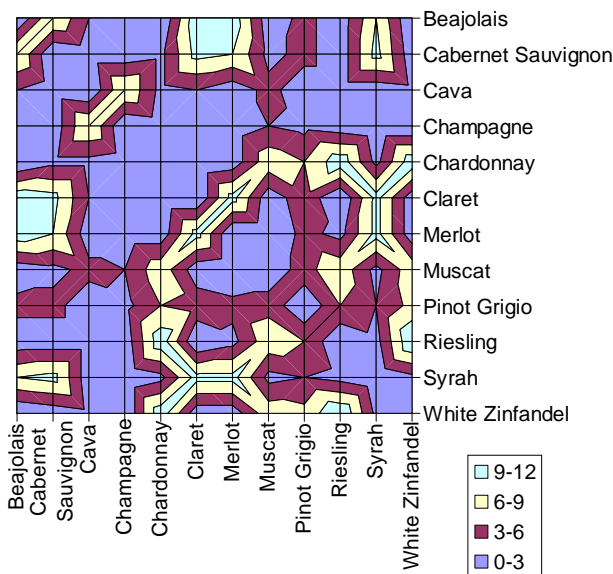
# Map with Reordered Rows & Columns

# Extended Analysis

❑ Basic cluster analysis counts how many times cards appear in the same group

❑ Adding a quality of fit measure lets participants say how well an item fits into its group

# Quality of Fit

❑ Instead of just counting each pair we use the average quality-of-fit between two items

❑ For example, if a participant gave Claret a 3 and Beaujolais 1, the score entered for that pair would be 2

❑ We can also keep QOF averages per item and per participant

**Example Quality of Fit Ratings for a Card Sorting Participant**

| Group: Full-Bodied Reds | QOF |
|---|---|
| Claret | 3 |
| Merlot | 2 |
| Beaujolais | 1 |

Quality of Fit: 1 = Fair, 2 = Good, 3 = Perfect

# Matrix for Single Participant with QOF

| | Beaujolais | Cabernet Sauvignon | Cava | Champagne | Chardonnay | Claret | Merlot | Muscat | Pinot Grigio | Riesling | Syrah | White Zinfandel |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Beaujolais | | 2.0 | | | | 2.0 | 1.5 | | | | 1.5 | |
| Cabernet Sauvignon | 2.0 | | | | | 3.0 | 2.5 | | | | 2.5 | |
| Cava | | | | 3.0 | | | | | | | | |
| Champagne | | | 3.0 | | | | | | | | | |
| Chardonnay | | | | | | | | 3.0 | 2.5 | 2.5 | | 2.0 |
| Claret | 2.0 | 3.0 | | | | | 2.5 | | | | 2.5 | |
| Merlot | 1.5 | 2.5 | | | | 2.5 | | | | | 2.0 | |
| Muscat | | | | | 3.0 | | | | 2.5 | 2.5 | | 2.0 |
| Pinot Grigio | | | | | 2.5 | | | 2.5 | | 2.0 | | 1.5 |
| Riesling | | | | | 2.5 | | | 2.5 | 2.0 | | | 1.5 |
| Syrah | 1.5 | 2.5 | | | | 2.5 | 2.0 | | | | | |
| White Zinfandel | | | | | 2.0 | | | 2.0 | 1.5 | 1.5 | | |

# Effect of Quality of Fit on Dendogram
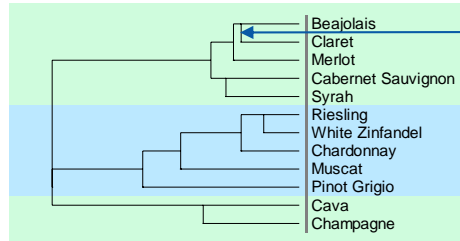
| | |
|---|---|
| Beajolais<br>Claret<br>Merlot<br>Cabernet Sauvignon<br>Syrah<br>Riesling<br>White Zinfandel<br>Chardonnay<br>Muscat<br>Pinot Grigio<br>Cava<br>Champagne | **Without QOF**, Beaujolais has quite a strong relationship with the other red wines |
| +++ Cabernet Sauvignon<br>+++ Claret<br>++ Merlot<br>++ Syrah<br>++ Beajolais<br>++ Cava<br>+++ Champagne<br>++ Chardonnay<br>++ Riesling<br>++ White Zinfandel<br>++ Muscat<br>++ Pinot Grigio | **With QOF**, Cabernet Sauvignon and Claret are now strongest as the group was named "full-bodied reds". The Beaujolais relationship is much weaker |
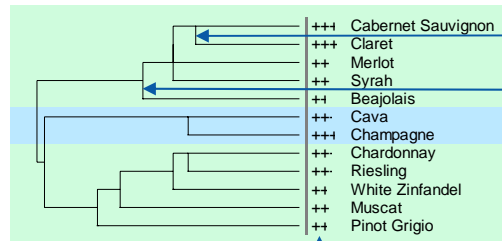
Average QOF for each item (from SynCaps)

19

# Extending Card Sort Methods

❑ Paper card sorting
- Advantages: Flexible, relatively inexpensive
- Disadvantages: Manual data capture

❑ Computer card sorting (EZSort, for example)
- Advantages: Quick, no separate data capture
- Disadvantages: One computer per participant, potentially confusing to participants, lacks flexibility

❑ Traditional card sort analyses
- Examination of the proximity matrix
- Hierarchical cluster analysis and dendogram

❑ Some shortcomings
- Dendograms are only part of the picture
- Dendogram output usually not easy to include in a report

20

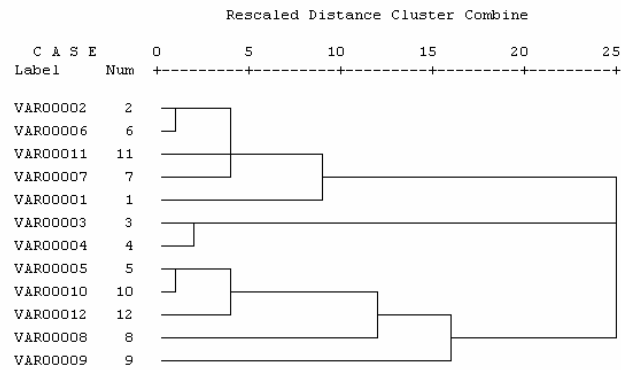# Computer Card Sort (USort)
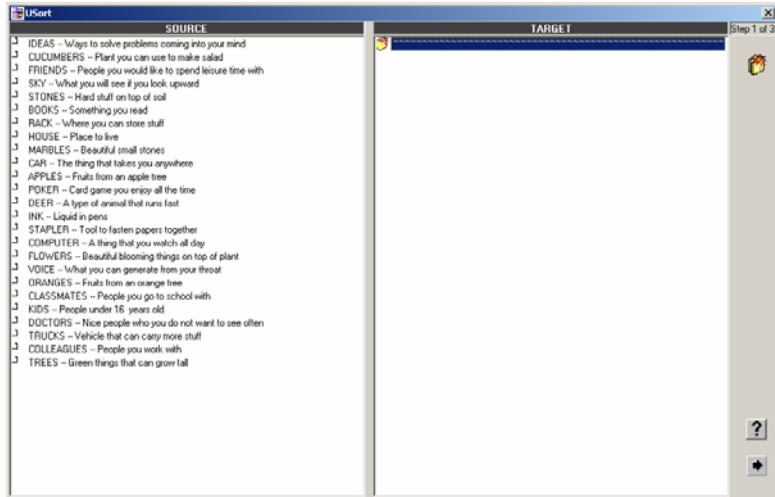
# EZCalc Dendogram
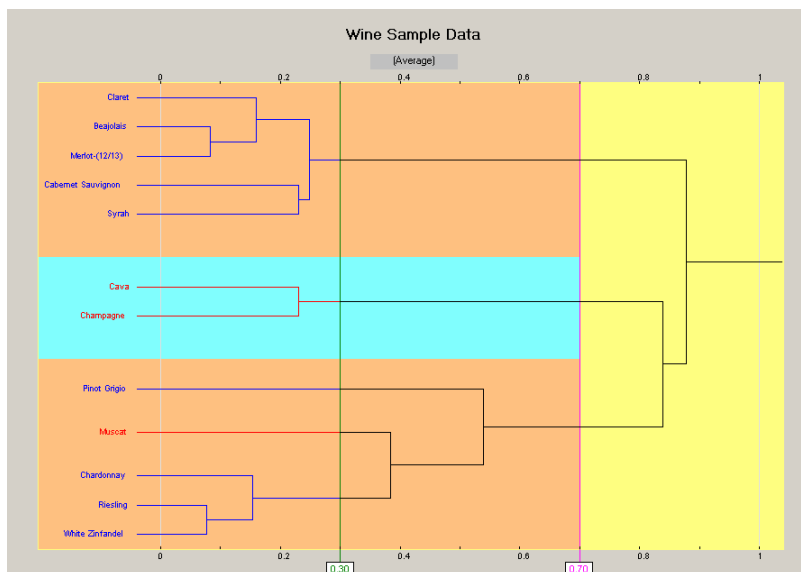
# SPSS Dendogram

```
* * * * * H I E R A R C H I C A L   C L U S T E R   A N A L Y S I S * * * * *

Dendrogram using Average Linkage (Between Groups)

                       Rescaled Distance Cluster Combine

    C A S E      0         5        10        15        20        25
 Label    Num   +---------+---------+---------+---------+---------+

 VAR00002    2
 VAR00006    6
 VAR00011   11
 VAR00007    7
 VAR00001    1
 VAR00003    3
 VAR00004    4
 VAR00005    5
 VAR00010   10
 VAR00012   12
 VAR00008    8
 VAR00009    9
```

# Extending Card Sort Methods

❑ Computer Aided Paper Sorting (CAPS)
- ▪ Paper cards with bar codes
- ▪ Advantages: Flexible, relatively inexpensive, fast and accurate data capture (over 100 cards per minute)
- ▪ Disadvantages: Requires bar code reader

❑ Extended analyses
- ▪ Quality of fit adjustments to dendogram
- ▪ Average quality of fit by item and participant
- ▪ Deviations by participant against "expert" and group
- ▪ Group sizes and number of items by participant

❑ Improved reporting
- ▪ Appearance of dendogram is user adjustable
- ▪ Image is a metafile rather than a bitmap, meaning that it can easily be resized or printed without loss of quality

# Caps Cards



Card Showing Bar Codes and Quality of Fit

# Caps Cards

❑ Caps cards
- Produced with Microsoft Word mail merge templates (free download from Syntagm web site)
- Laser or inkjet printed
- Pre-perforated A4 or US Letter sheets
- Can be bought from any supplier (we can supply the A4 version more cheaply than most other UK sources)

❑ Bar code reader
- Relatively inexpensive (about GBP £100/EUR €100/USD $100)
- We supply a suitable scanner in the UK (with free technical support for Caps) but many USB/Code 3 of 9 scanners will work
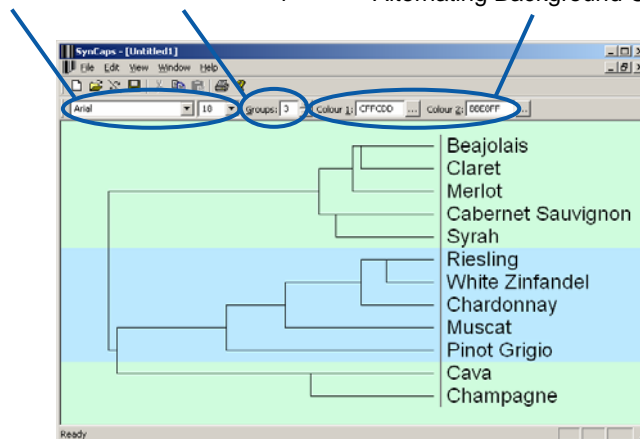
# SynCaps

❑ Data preparation and analysis

- ▪ Processes bar code (or typed) data
- ▪ Produces dendogram that can dragged and dropped (or copy any pasted) into Word documents or similar
- ▪ Allows what-if analyses by allowing easy control over the number of groups
- ▪ Generates output files for further analysis and charting by Excel (or other packages that read .csv files), SPSS and EZCalc
- ▪ Creates proximity matrix files with and without QOF weighting
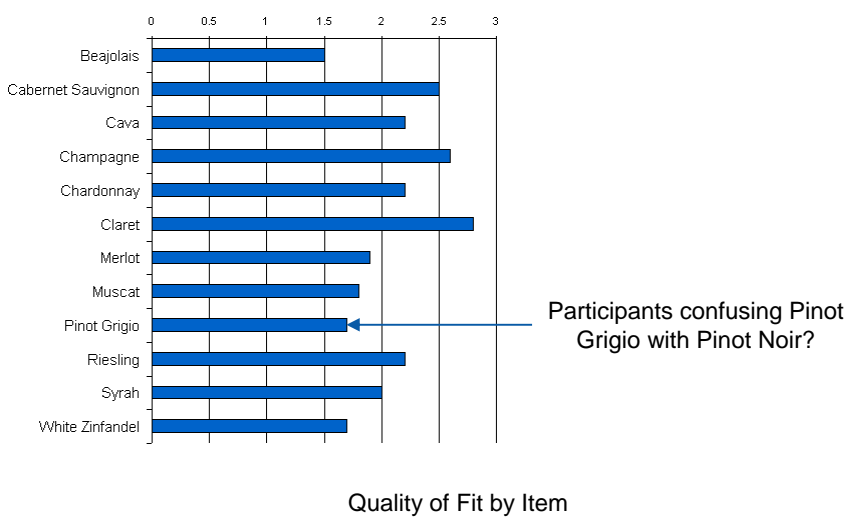- ▪ Free to try for 30 days, licensed per machine (GBP £40 – about EUR €60/USD $70)

# SynCaps

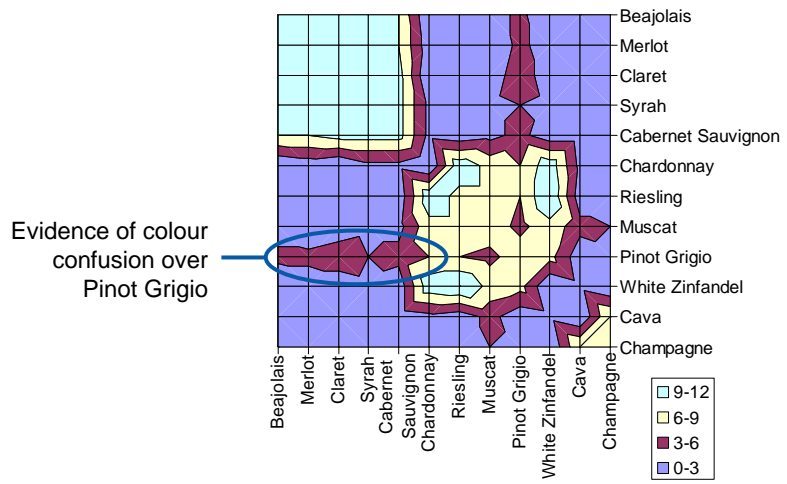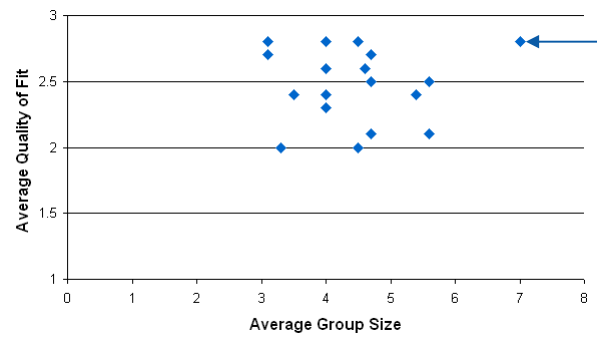Font and size   Number of Groups          Alternating Background Colours

# Sample Excel Analysis



Participants confusing Pinot Grigio with Pinot Noir?

Quality of Fit by Item

# Sample Excel Analysis



Evidence of colour confusion over Pinot Grigio

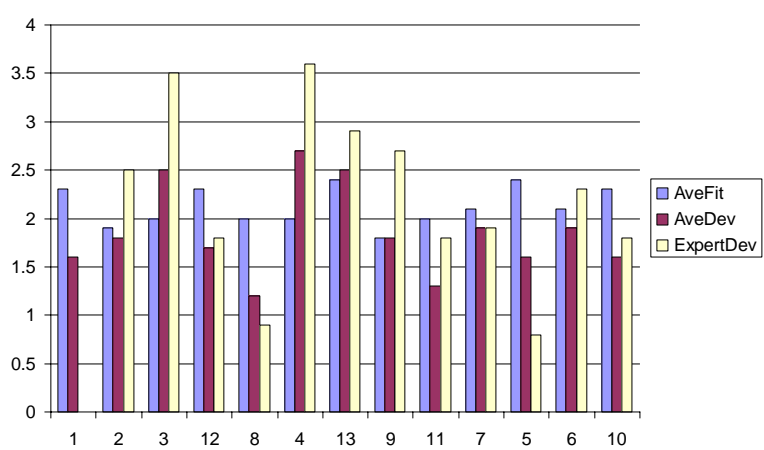Contour Map Using SynCaps Proximity Matrix

# Sample Excel Analysis



Slightly unusual participant?

Quality of Fit by Group Size per Participant

# Sample Excel Analysis



"Expert"

Participants 3 and 4 vary from most other participants and the expert results

# Summary

❑ The basic principles of card sorting are not complex although cluster analysis software and dendograms can be intimidating

❑ There are many other ways of looking at card sorting data: quality of fit and statistics calculated for each item or participant can reveal a wealth of information behind a dendogram

❑ Paper card sorting need not be tedious, although there is a small investment needed

❑ The principles of the extended analysis are straightforward but omitted by most cluster analysis software

❑ See www.syntagm.co.uk/design/cardsort.shtml for more information and free downloads