MeasuringU

Practical Statistics

for UX & Customer Research

Jeff Sauro, PhD & Jim Lewis, PhD

Measuring University™

Instructors

Jeff Sauro, PhD

Founding Principal

Jeff Sauro PhD, is the founding principal of MeasuringU. For over twenty years he's been conducting UX research, including benchmarking studies for clients.

Jeff has published over twenty-five peer-reviewed research articles and five other books, including *Benchmarking the User Experience*, *Customer Analytics for Dummies* and *Quantifying the User Experience*.

Jim Lewis, PhD, CHFP

Distinguished User Experience Researcher

Jim is a Certified Human Factors Professional with a Ph.D. in Experimental Psychology (and M.A. in Engineering Psychology, minor in applied statistics).

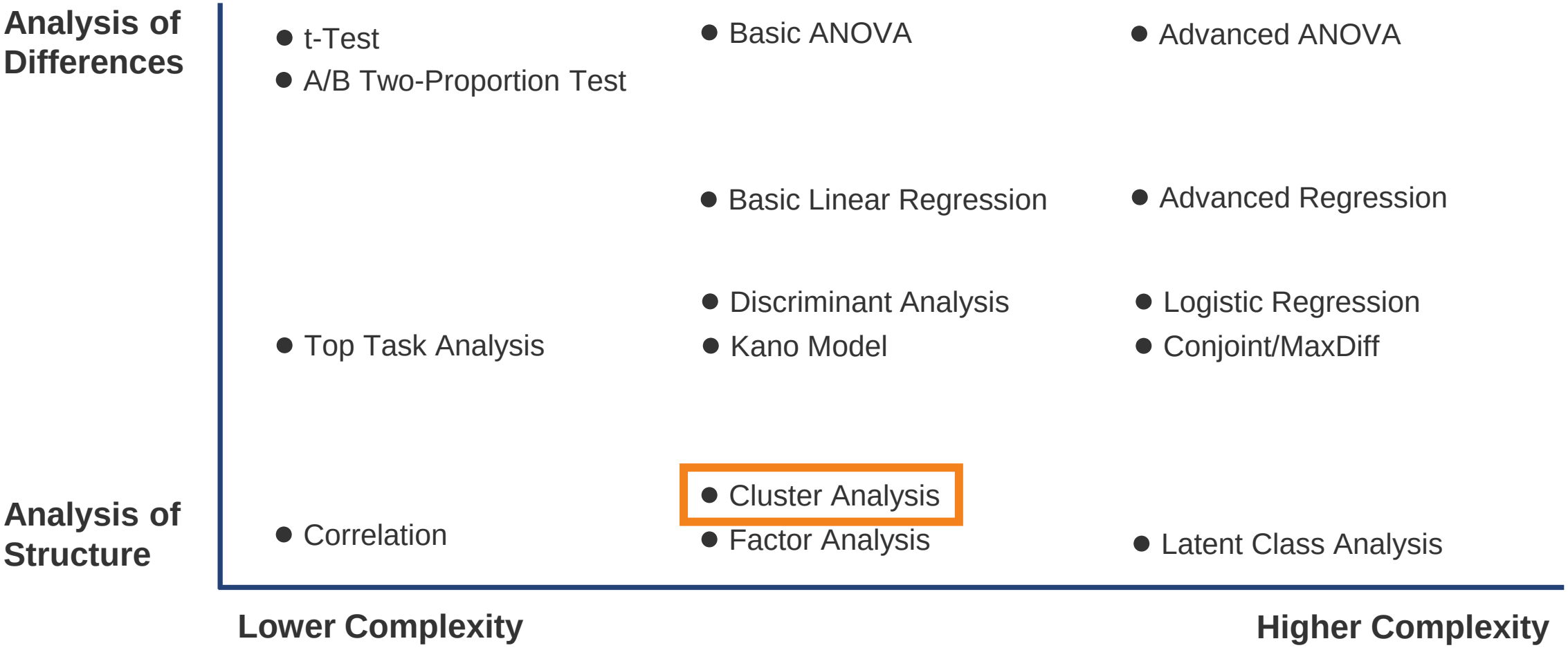
Before joining MeasuringU Jim worked at IBM for nearly 40 years. He is an IBM Master Inventor (> 90 US patents) and has published over 100 articles and papers.

Cluster Analysis

Topics Covered

- History of Cluster Analysis
- Different Types of Cluster Analysis
- Preview of the Steps for Conducting a Cluster Analysis

Overview of advanced UX analytical methods



More info: [5 Advanced Stats Techniques & When to Use Them]

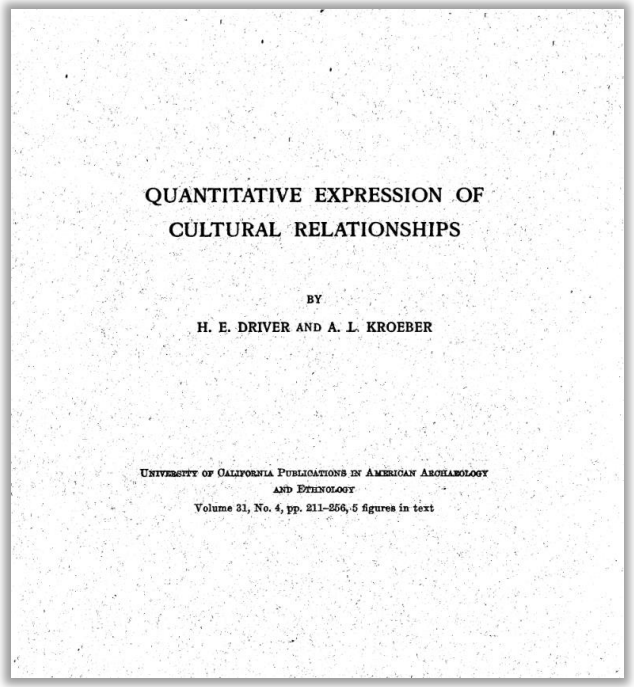
Applications of advanced analytical methods

Question	Methods	Sample Application
Are there significant differences?	Two Proportion Test	A/B testing
	t-Test	Test two designs
	ANOVA	Test multiple designs and interactions
Are there significant similarities?	Correlation	Assess relationships (e.g., CSAT and age)
Are there significant predictors?	Linear Regression	Key driver analysis
Is there latent (hidden) structure?	Cluster Analysis	Persona/segmentation and card sorting
	Factor Analysis	Develop standardized UX questionnaire
	Latent Class Analysis	Advanced persona development
Can we determine membership in classes?	Discriminant Analysis	Customer segment classification tool
	Logistic Regression	Statistical basis for feature prioritization
What are the most important features/tasks?	Conjoint Analysis	Exhaustive feature prioritization
	MaxDiff Analysis	Streamlined feature prioritization
	Kano Model	Alternative feature prioritization method
	Top Task Analysis	Identify most important tasks

More info: 5 Advanced Stats Techniques & When to Use Them

Earliest Use Of Cluster Analysis

Alfred Kroeber & Unknown Peruvian



QUANTITATIVE EXPRESSION OF
CULTURAL RELATIONSHIPS

BY

H. E. DRIVER AND A. L. KROEBER

UNIVERSITY OF CALIFORNIA PUBLICATIONS IN AMERICAN ARCHAEOLOGY
AND ETHNOLOGY
Volume 31, No. 4, pp. 211-256, 5 figures in text

Driver and Kroeber 1932

"We inquire whether the cultures carried or possessed by such ethnic groups are more or less similar to one another, on the basis of their containing or not containing traits such as matrilineate, avoidance, self-torture vows, the fire drill, sinew-backed bow, twined weaving, ridged houses, etc."

More info: Quantitative Expression of Cultural Relationships [PDF]

Cultural groups were clustered by their shared attributes

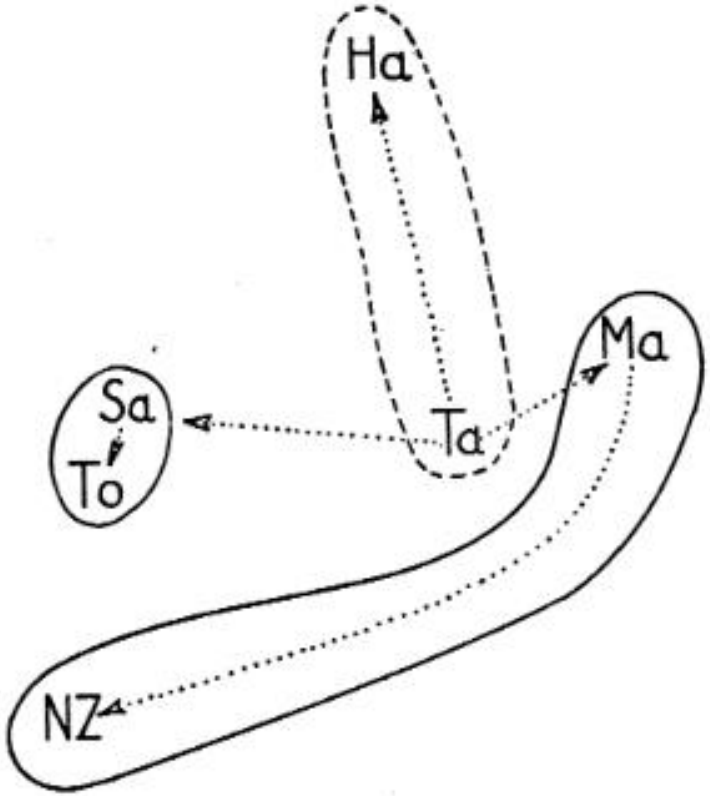


Fig. 1. Polynesian relationships. The distances are those on the map. Cultural similarities are indicated by enclosures of different strength; probable courses of cultural relationship, by arrows.

Different algorithms used for grouping

Over 100 Different Algorithms and the “best” one depends on the context and results

Common “Hierarchical” Clustering Algorithms

Names	Formula
Euclidean distance	$\|a-b\|_2 = \sqrt{\sum_i (a_i - b_i)^2}$
Squared Euclidean distance	$\|a-b\|_2^2 = \sum_i (a_i - b_i)^2$
Manhattan (or city block) distance	$\|a-b\|_1 = \sum_i \|a_i - b_i\|$
Maximum distance (or Chebyshev distance)	$\|a-b\|_\infty = \max_i \|a_i - b_i\|$
Mahalanobis distance	$\sqrt{(a-b)^\top S^{-1}(a-b)}$ where S is the Covariance matrix

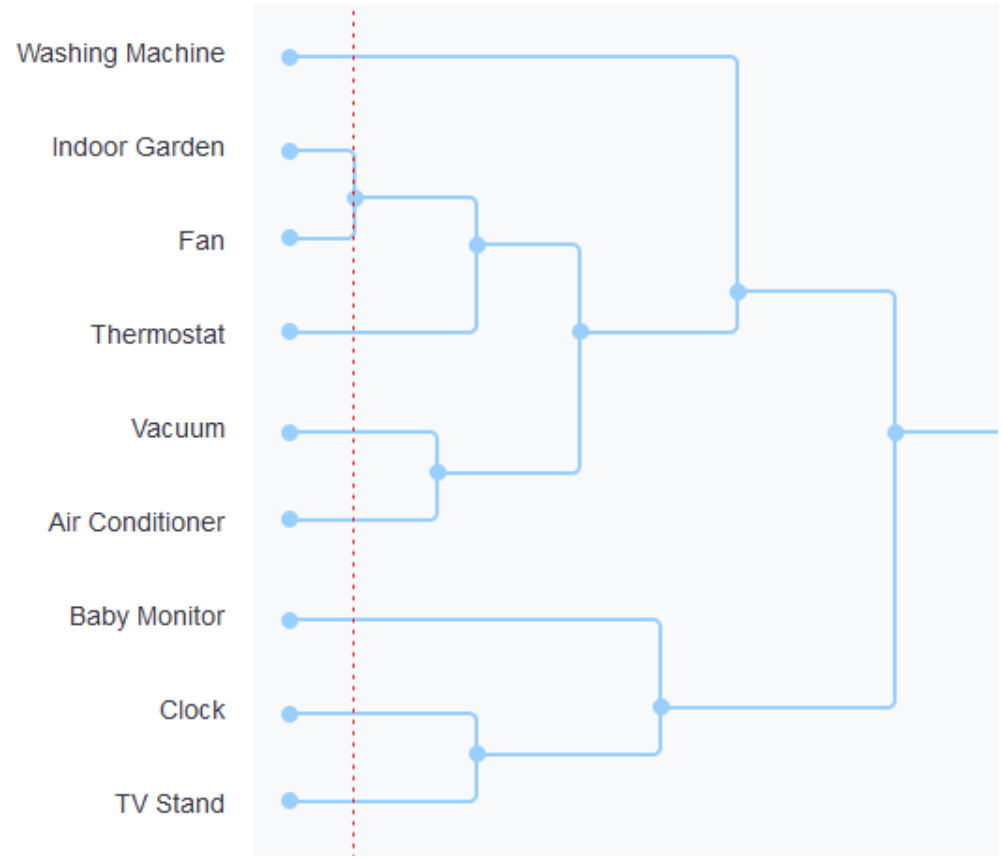
Ward’s Method is a popular approach used in SPSS based on the Squared Euclidian distance

More info: 5 Advanced Stats Techniques & When to Use Them

Cluster Analysis Application 1: **Card Sorting**

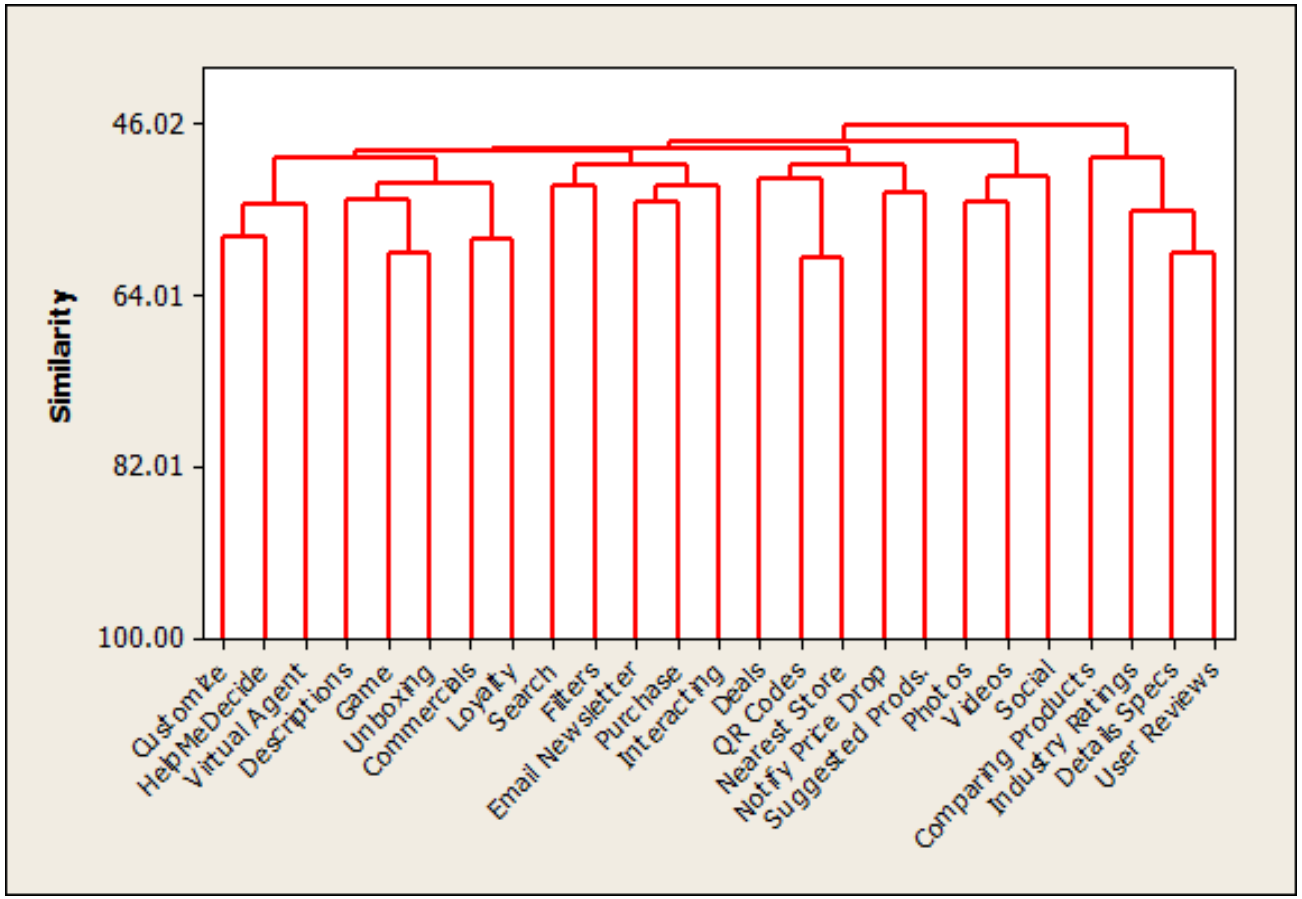
Air Conditioner											
23%	Blender										
23%	98%	Food Processor									
23%	98%	97%	Hand Mixer								
7%	5%	4%	5%	Headphones							
69%	35%	35%	34%	5%	Washing Machine						
29%	13%	12%	12%	32%	22%	Baby Monitor					
25%	96%	96%	95%	3%	36%	11%	Coffee Maker				
13%	4%	3%	4%	64%	9%	40%	2%	Controller			
24%	99%	99%	98%	4%	36%	12%	97%	3%	Deep Fryer		
40%	80%	80%	79%	2%	54%	10%	81%	3%	81%	Dishwasher	
12%	5%	4%	5%	66%	8%	50%	5%	63%	4%	3%	Drone
67%	22%	22%	21%	10%	48%	41%	22%	17%	21%	22%	14% Fan

Similarity Matrix



Dendrogram

Cluster Analysis Application 2: **Segmentation**



Dendrogram from a similarity matrix on how smartphone users rated features

Steps to Running a Cluster Analysis

1. Examine Correlation/Similarity Matrix
2. Select Cluster Algorithm in Software
3. Inspect Dendrogram for Clusters
4. Determine Clusters Based on Visual Inspection

} Covered in Extended Examples

Summary

There are many approaches to cluster analysis

- Cluster analysis includes a variety of methods for identifying groups of items
- Two-step, k-means, hierarchical
- Hierarchical cluster analysis using Ward's method works well for UX research

Cluster analysis useful in UX research to infer groups from data

- Grouping cards after card-sorting study
- Grouping people in persona research
- Some subjectivity involved in determining the number of clusters

Interpret cluster analysis with visual inspection

- Graphs useful in interpretation include dendogram and profile of means

MeasuringU

Moderated UX Studies
(in our Denver labs or remotely)

MUIQ
Unmoderated Studies
(using our MUIQ platform)

Participant Recruiting
(US & International)

Eye Tracking &
Facial Expression Analysis

Navigation Testing
(Card-Sorting/Tree-Testing)

Survey Design & Analysis
(including MaxDiff & Kano)

SUPR-Q
Statistical Analysis &
Measurement Advising

MeasuringU UX BOOTCAMP
Training &
Workshops

MeasuringU is a research firm based in Denver, Colorado
focusing on quantifying the user experience.

Weekly Newsletter: MeasuringU.com
LinkedIn | Twitter : @MeasuringU