

Parallel Coordinate Plots for Visualisation of Longitudinal Survey Data

Kevin Pulo

National Computational Infrastructure (NCI)
National Facility

ANU Supercomputing Facility

kevin.pulo@anu.edu.au

Longitudinal study

- Follows the same set of individuals over time
 - Aka "Panel data"
- Eg. employment/study/relationship status over the past 10 years for a group of people

person	year	employment	study	relationship
1	2006	none	full-time	single
1	2007	part-time	part-time	cohabitating
1	2008	full-time	none	married
2	2006	full-time	part-time	cohabitating
2	2007	full-time	none	cohabitating
2	2008	full-time	none	single
...				

Longitudinal study

- Follows the same set of individuals over time
 - Aka "Panel data"
- Eg. employment/study/relationship status over the past 10 years for a group of people

	person		
year	1	2	...
2006	n, f, s	f, p, c	
2007	p, p, c	f, n, c	
2008	f, n, m	f, n, s	
.			
.			
.			

Major Australian longitudinal studies

- Negotiating the Life Course (NLC)
- Household, Income and Labour Dynamics in Australia (HILDA)
- Growing Up in Australia
- Australian Longitudinal Study on Women's Health
- Australian Longitudinal Study of Ageing (ALSA)

Negotiating the Life Course (NLC)

- Based at ANU (RSSS, ADSRI, ASSDA)
- Is interested in:
 - "the changing life courses and decision-making processes of Australian men and women as the family and society move from male breadwinner orientation in the direction of higher levels of gender equity"
- 4 "waves" of data
 - Wave 1 (1997): 2231 respondents
 - Wave 2 (2000): 1768 respondents
 - Wave 3 (2003): 1192 respondents
 - Wave 4 (2006): 1138 respondents + 2000 new resps.

Pilot visualisation project

- Collaboration between ANUSF and ASSDA
- Interested in investigating ways of directly visualising longitudinal data
- Eventual goal is full integration with the data archive website

Employment status

- Initial motivation
- Employment status can be:
 - Employed
 - Unemployed (looking for work)
 - Not in the labour force (not looking for work)
- NLC has retrospective data going back 40 years
- How to visualise the employment paths/decisions that people take over their lives?

Why visualise?

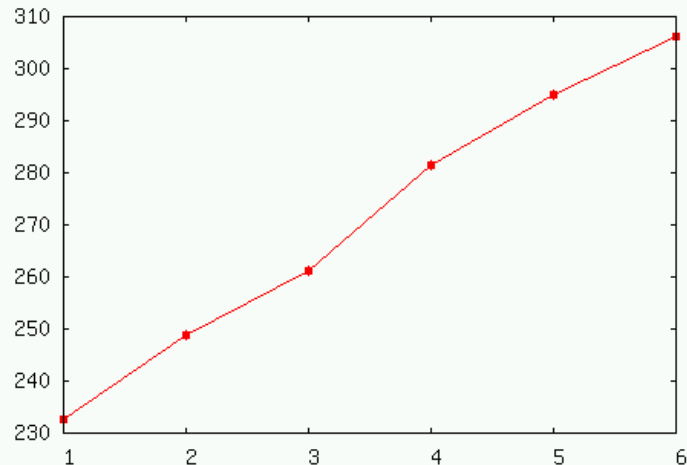
- Make use of the human perceptual system
- Eg. graphs, pie charts, etc

t	value
1	232.56
2	248.89
3	261.22
4	281.55
5	294.88
6	306.21

Why visualise?

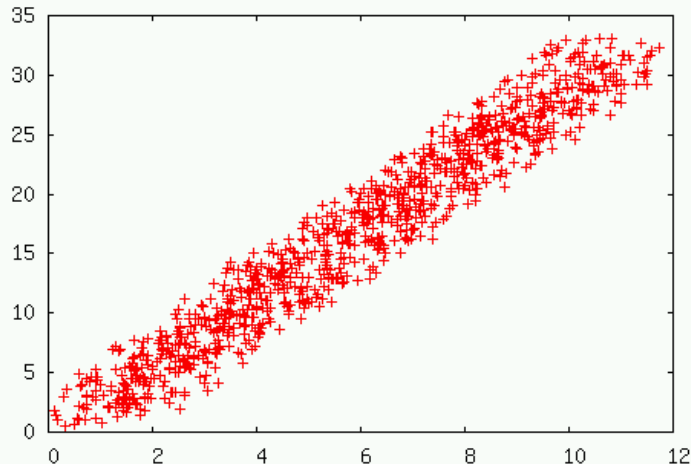
- Make use of the human perceptual system
- Eg. graphs, pie charts, etc

t	value
1	232.56
2	248.89
3	261.22
4	281.55
5	294.88
6	306.21



Why visualise?

- Make use of the human perceptual system
- Eg. graphs, pie charts, etc



Mapping data to space

- Mapping data dimensions to space/time runs out at 4 dimensions
 - Eg. animated 3 dimensional scatter plot

Mapping data to space

- Mapping data dimensions to space/time runs out at 4 dimensions
 - Eg. animated 3 dimensional scatter plot
- Map dimensions to some more visual attributes
 - Eg. colour, glyph size, glyph shape

Visualising multi-dimensional data

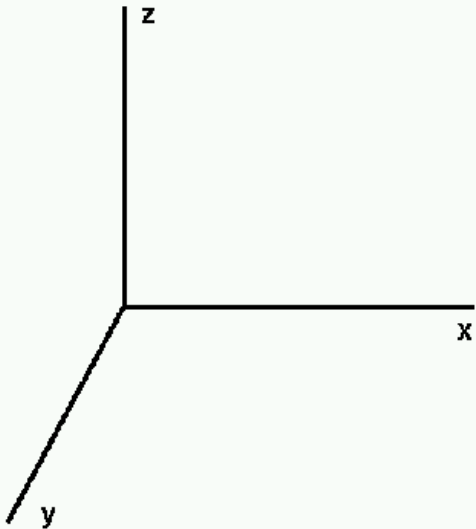
- What to do when the number of dimensions is over ~ 10 (eg. hundreds)?
 - Eg. Relationship and labour force status over the past 40 years (ie. 80 dimensional data)
- Somehow map high dimensional data to low-dimensional space
- "Multidimensional scaling"
 - Sophisticated statistical technique
 - Tries to find (dis)similarities in the data, and then plot similar things near each other in the target space
 - Heavy-weight, difficult to explain/interpret, requires a priori selection of similarity metric

Parallel Coordinate Plots

- Dimensions (axes) are parallel (not orthogonal)

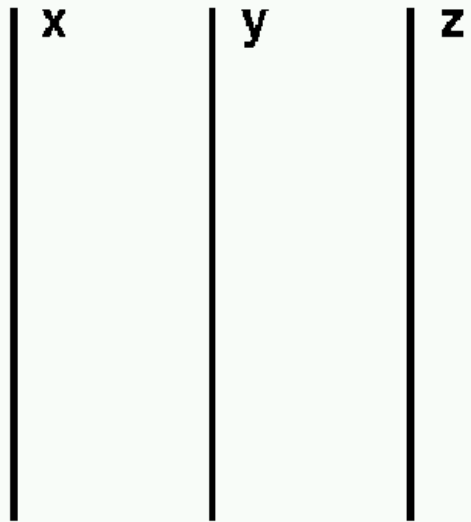
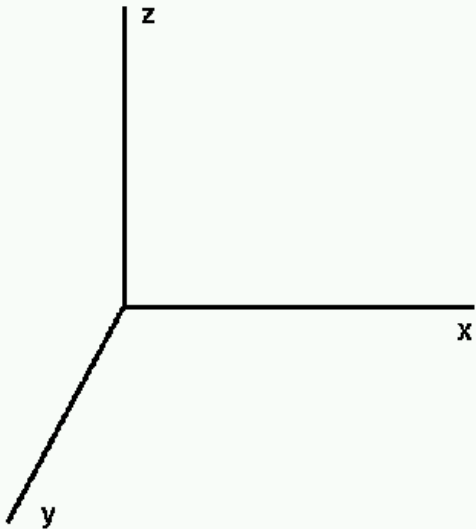
Parallel Coordinate Plots

- Dimensions (axes) are parallel (not orthogonal)



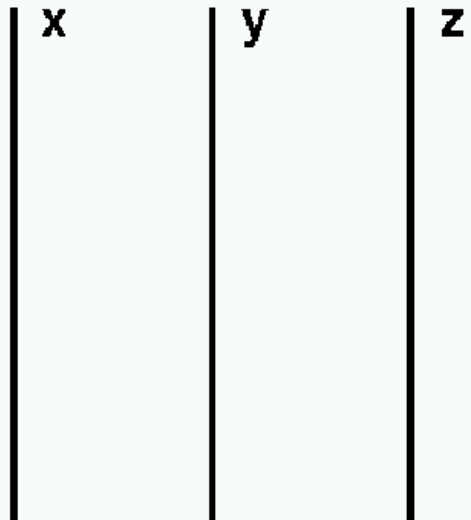
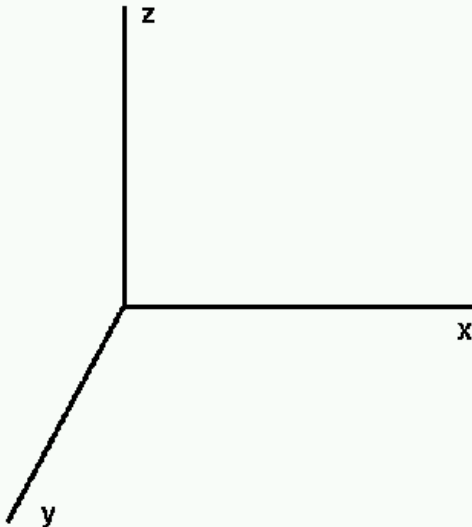
Parallel Coordinate Plots

- Dimensions (axes) are parallel (not orthogonal)



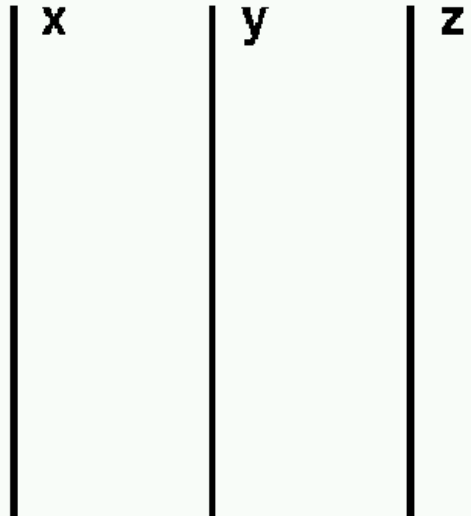
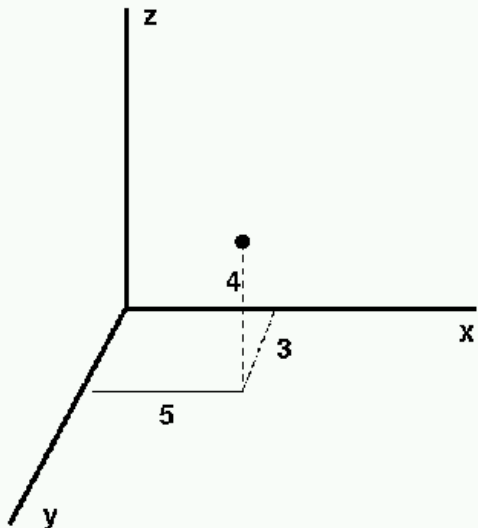
Parallel Coordinate Plots

- Dimensions (axes) are parallel (not orthogonal)
- "Points" in space become lines joining coordinates on the axes



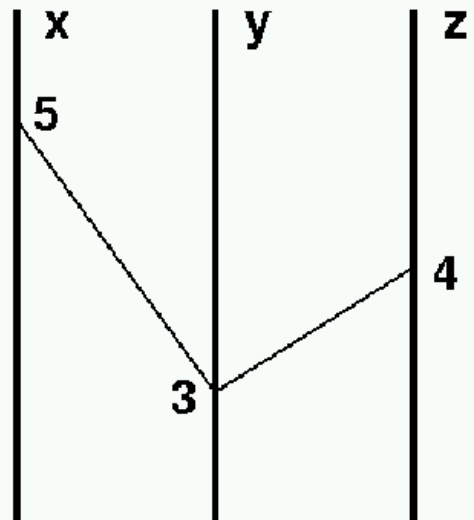
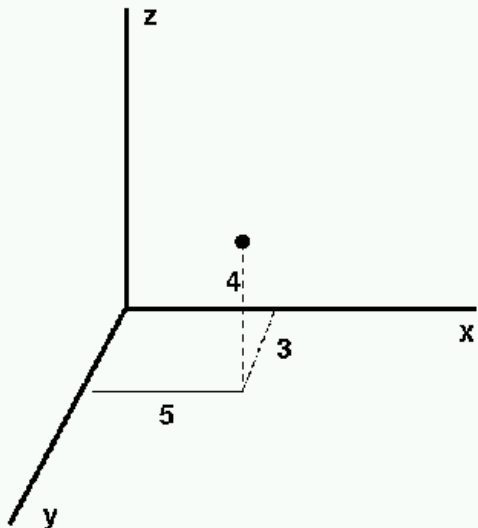
Parallel Coordinate Plots

- Dimensions (axes) are parallel (not orthogonal)
- "Points" in space become lines joining coordinates on the axes



Parallel Coordinate Plots

- Dimensions (axes) are parallel (not orthogonal)
- "Points" in space become lines joining coordinates on the axes

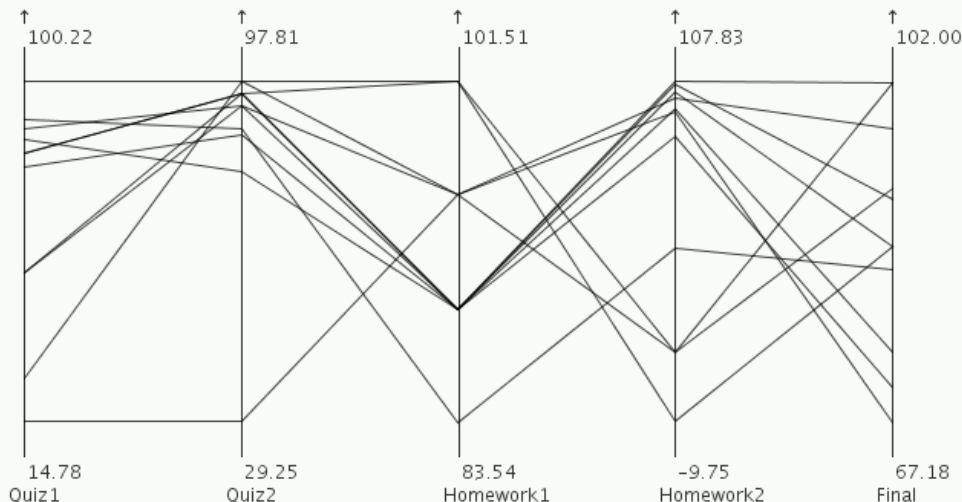


Advantages

- Straightforward mapping of dimensions to 2D
- Allows many dimensions
- Human perception is good at seeing
 - adjacency (ie. lines showing connectivity)
 - patterns (eg. groups of lines)

Advantages

- Straightforward mapping of dimensions to 2D
- Allows many dimensions
- Human perception is good at seeing
 - adjacency (ie. lines showing connectivity)
 - patterns (eg. groups of lines)



Disadvantages

- Not great for categorical data
 - Solution: Spreading, curves
- Axis ordering/scaling/sign can matter (a lot)
 - Solution: Interactive axis manipulation
- Can get very messy
 - Solution: Interactive brushing/filtering/labouring
- The best axes/view is not always clear
 - Solution: Interactive builtin recoding

Application to NLC

- Looking at employment/labour force status (and studying)
- W1-W3 (unbalanced)
- Labour force status at each wave
 - Employed, unemployed, NILF
- Retrospective work history
 - Full-time, part-time, none
- Retrospective study history
 - Full-time, part-time, none

Employment at each wave

Retrospective work/study

- Separately

Retrospective work/study

- Separately
- Combined
 - by year

Retrospective work/study

- Separately
- Combined
 - by year
 - by age

Retrospective work/study

- Separately
- Combined
 - by year
 - by age
 - transitions

