University of Dayton

Parallel Coordinates

CPS 563 – Data Visualization

Dr. Tam Nguyen

tamnguyen@udayton.edu

Multivariate data

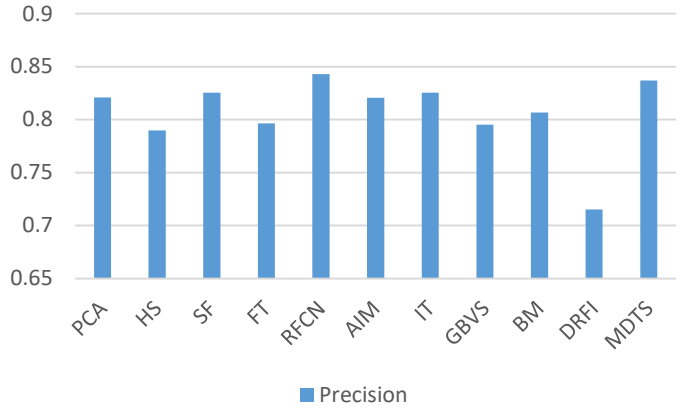
- Multivariate datasets can be expressed as a data table
  - Each entry in table is an observation
  - An observation consists of values of a set of variables, or variates

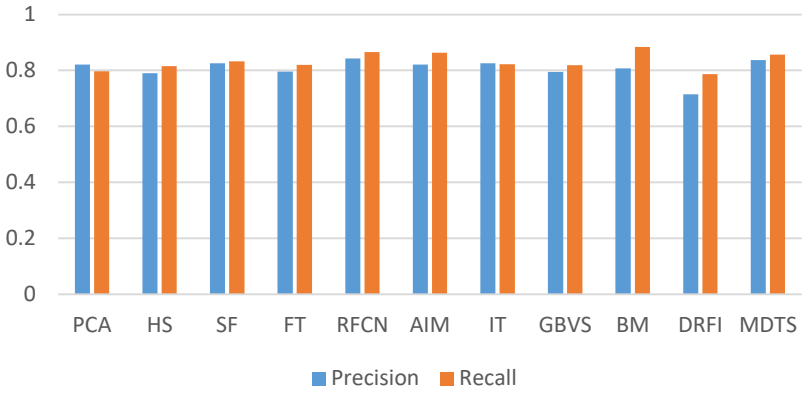
variables

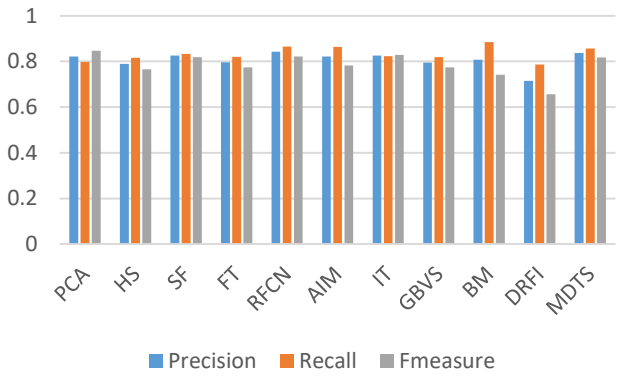
observations

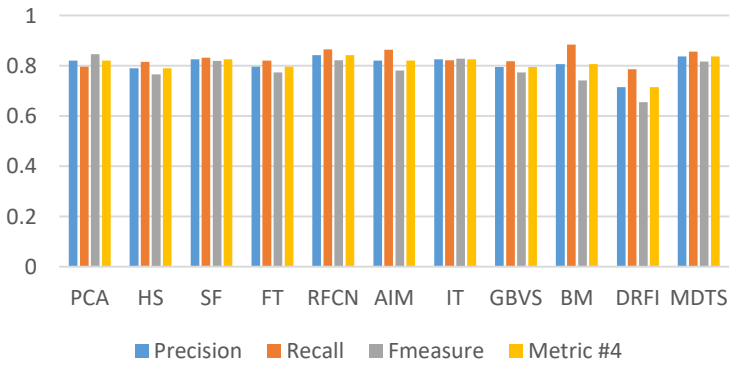
	Precision	Recall	F measure
Method 1	..	..	..
Method 2	..	..	..
Method 3			
Method 4			

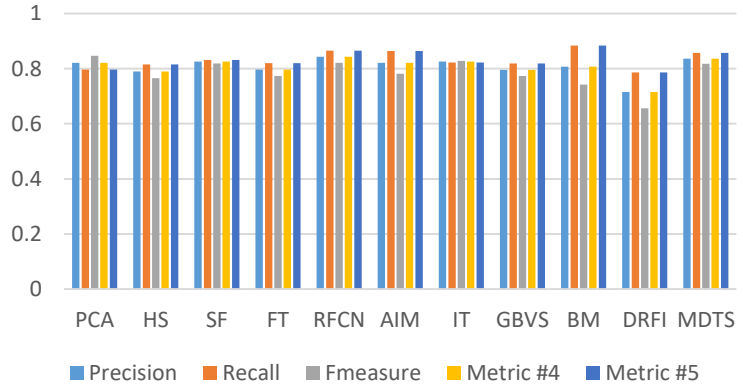
Bar Charts

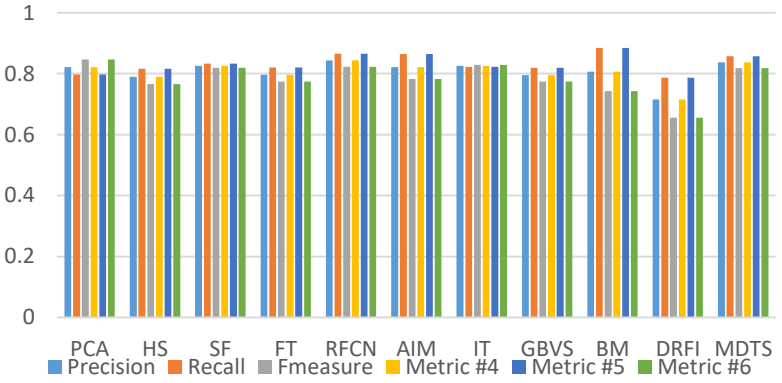

0.9
0.85
0.8
0.75
0.7
0.65
PCA
HS
SF
FT
RFCN
AIM
IT
GBVS
BM
DRFI
MDTS
Precision


1
0.8
0.6
0.4
0.2
0
PCA
HS
SF
FT
RFCN
AIM
IT
GBVS
BM
DRFI
MDTS
Precision
Recall

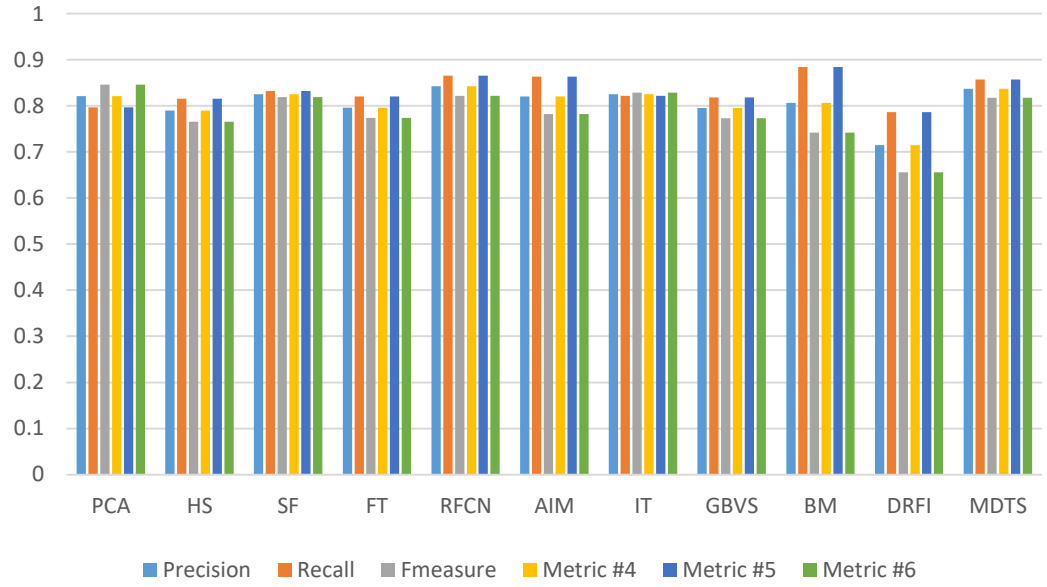

1
0.8
0.6
0.4
0.2
0
PCA
HS
SF
FT
RFCN
AIM
IT
GBVS
BM
DRFI
MDTS
Precision
Recall
Fmeasure


1
0.8
0.6
0.4
0.2
0
PCA
HS
SF
FT
RFCN
AIM
IT
GBVS
BM
DRFI
MDTS
Precision
Recall
Fmeasure
Metric #4


1
0.8
0.6
0.4
0.2
0
PCA
HS
SF
FT
RFCN
AIM
IT
GBVS
BM
DRFI
MDTS
Precision
Recall
Fmeasure
Metric #4
Metric #5


1
0.8
0.6
0.4
0.2
0
PCA
HS
SF
FT
RFCN
AIM
IT
GBVS
BM
DRFI
MDTS
Precision
Recall
Fmeasure
Metric #4
Metric #5
Metric #6

Bar Charts



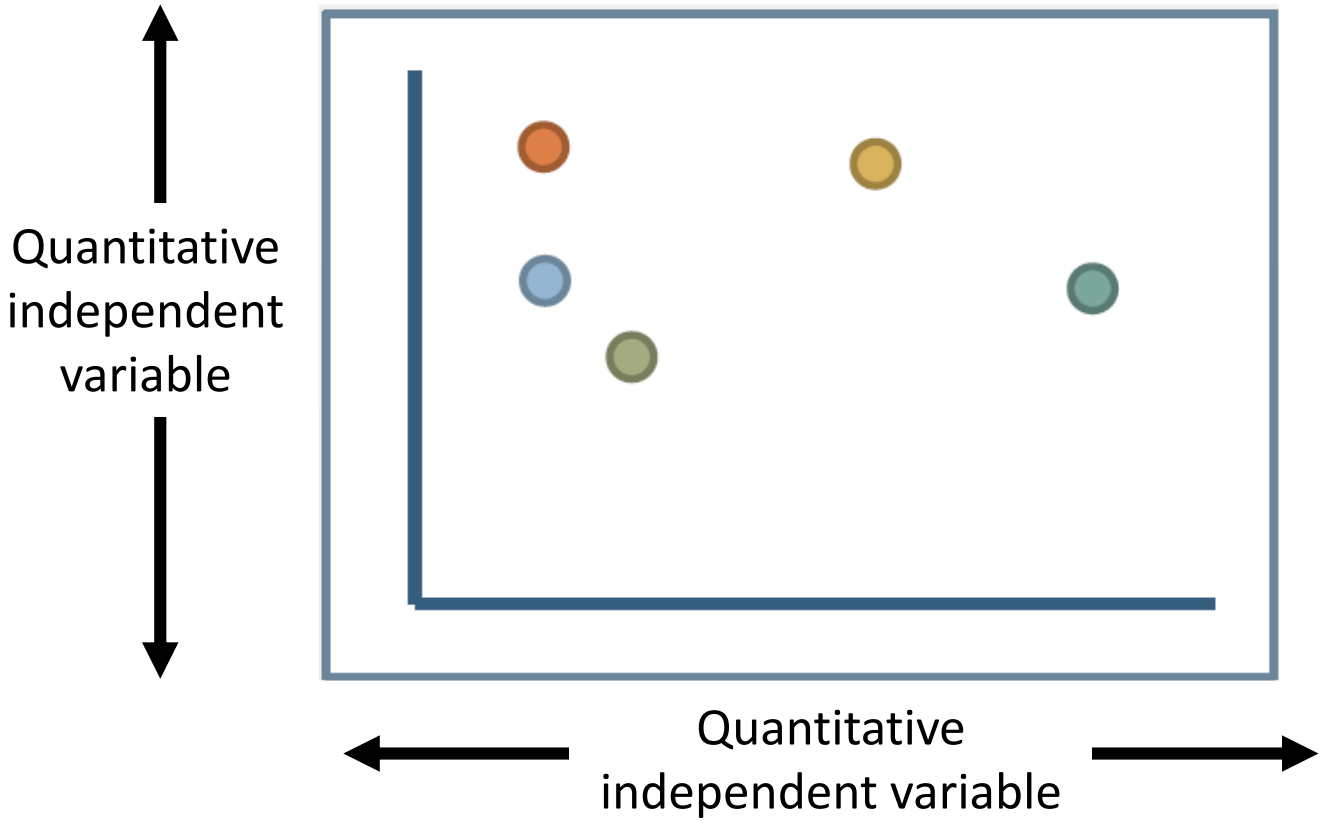
Advantages

+ Good for small comparisons

Disadvantages

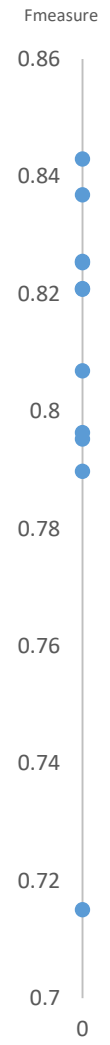
- Hard to observe the correlation between variables

Scatter Plot



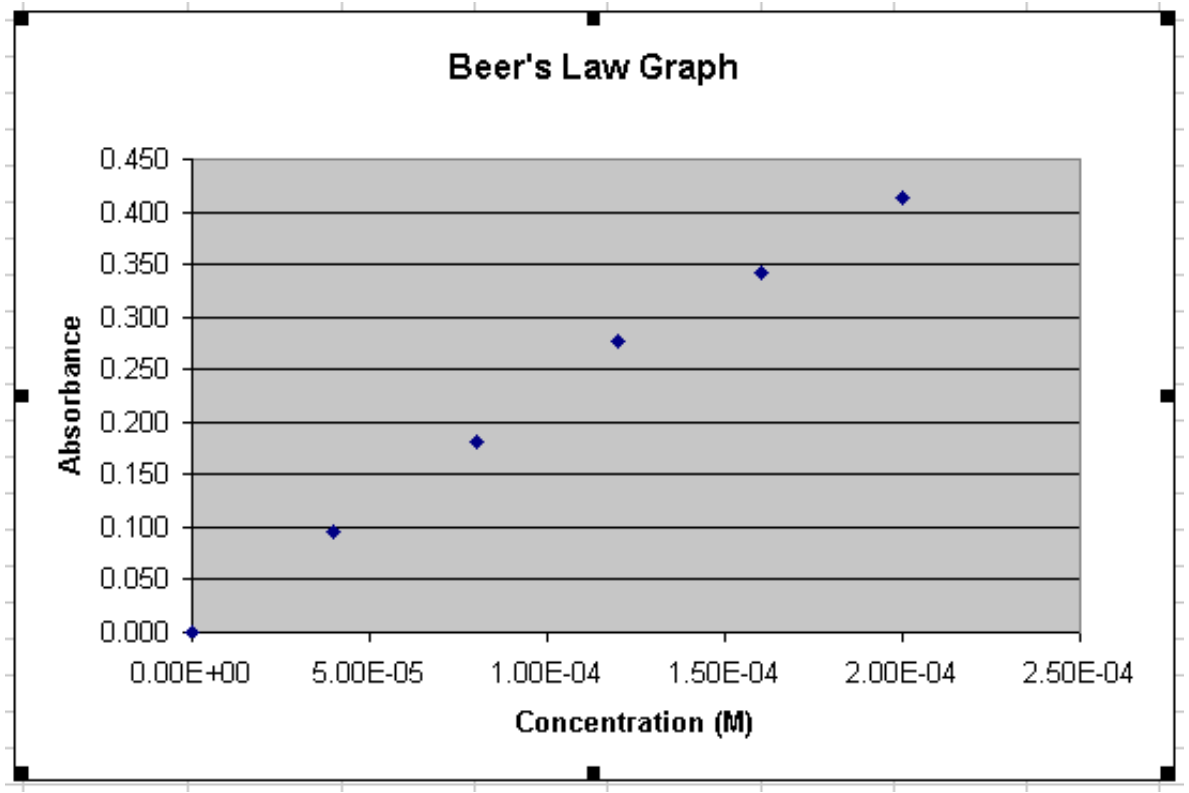
Scatter Plot

- Uni-variate data can be plotted on a line



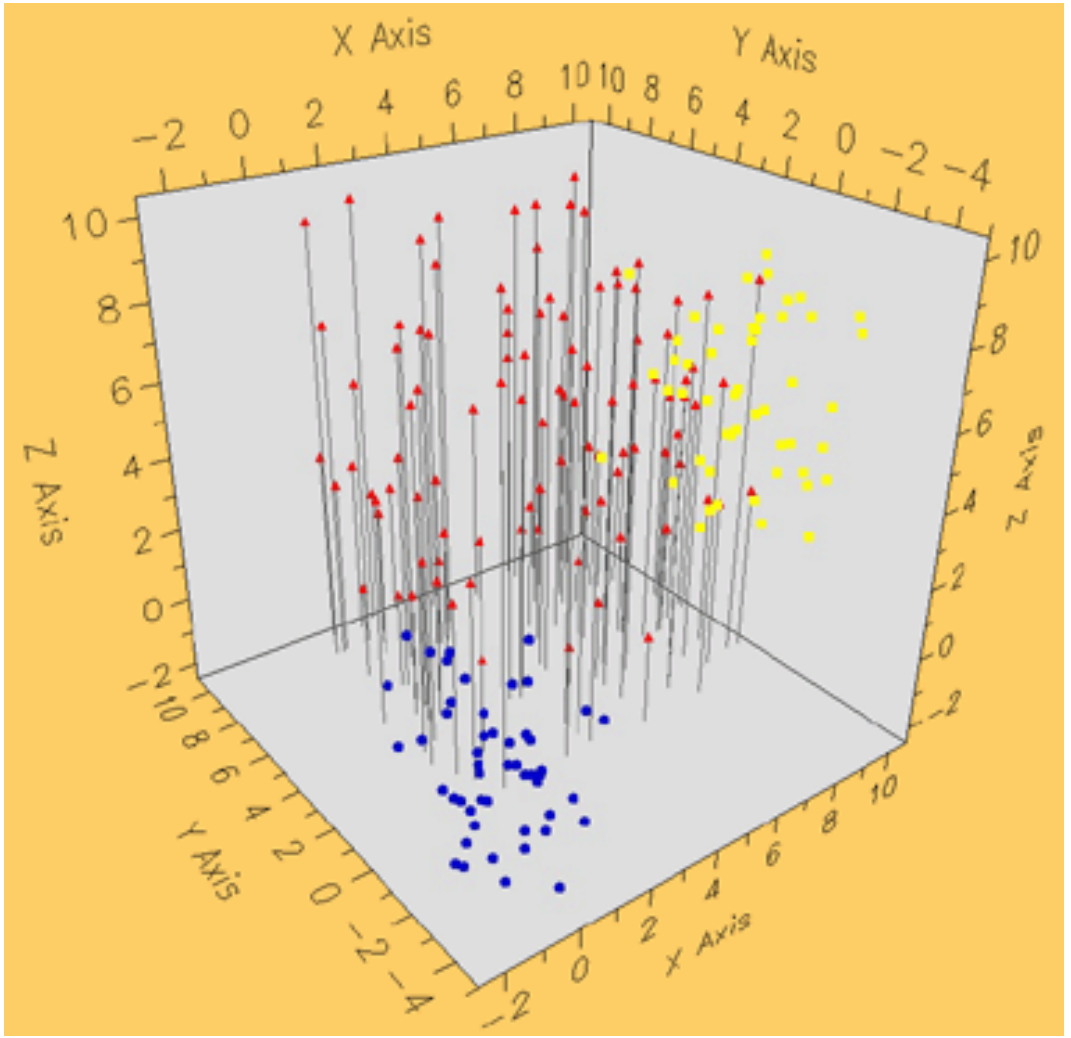
Scatter Plot

- For bivariates, we have already met the scatter plot technique
- It is useful for showing what happens to one variable as another changes…



3D Scatter Plot?

- Third variate expressed as another axis
- How about 4D, 5D, 6D …?



What are these?

Multivariate Visualization

- Example of Iris data set
  - 150 observations of 4 variables (length, width of petal and sepal)
  - We aim to display relationships between variables

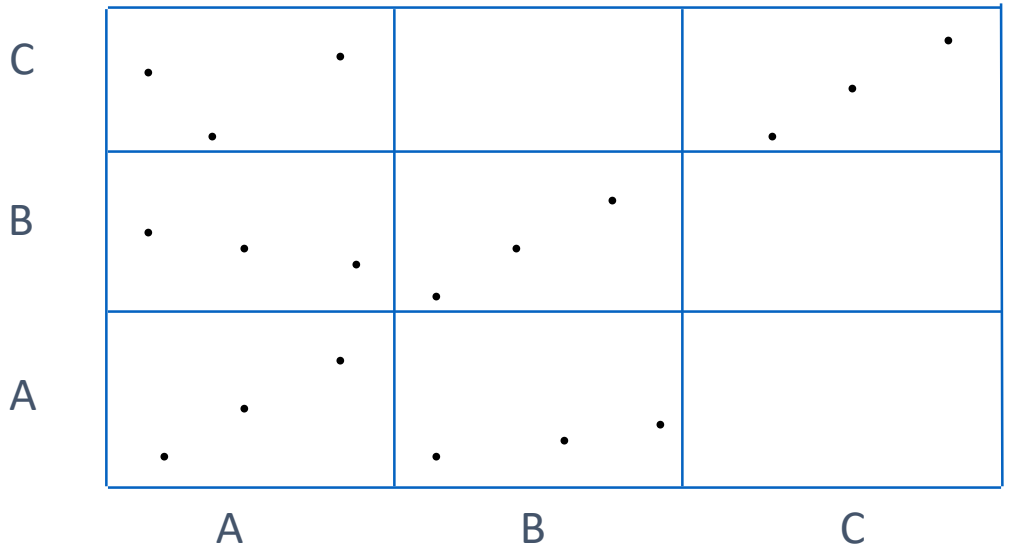
```
4 150
sepal_length
sepal_width
petal_length
petal_width
4.3 7.9 5
2.0 4.4 5
1.0 6.9 5
0.1 2.5 5
5.1   3.5   1.4   0.2
4.9   3     1.4   0.2
4.7   3.2   1.3   0.2
4.6   3.1   1.5   0.2
5     3.6   1.4   0.2
5.4   3.9   1.7   0.4
4.6   3.4   1.4   0.3
5     3.4   1.5   0.2
4.4   2.9   1.4   0.2
4.9   3.1   1.5   0.1
5.4   3.7   1.5   0.2
4.8   3.4   1.6   0.2
4.8   3     1.4   0.1
4.3   3     1.1   0.1
5.8   4     1.2   0.2
```

How to load Iris data in Matlab?

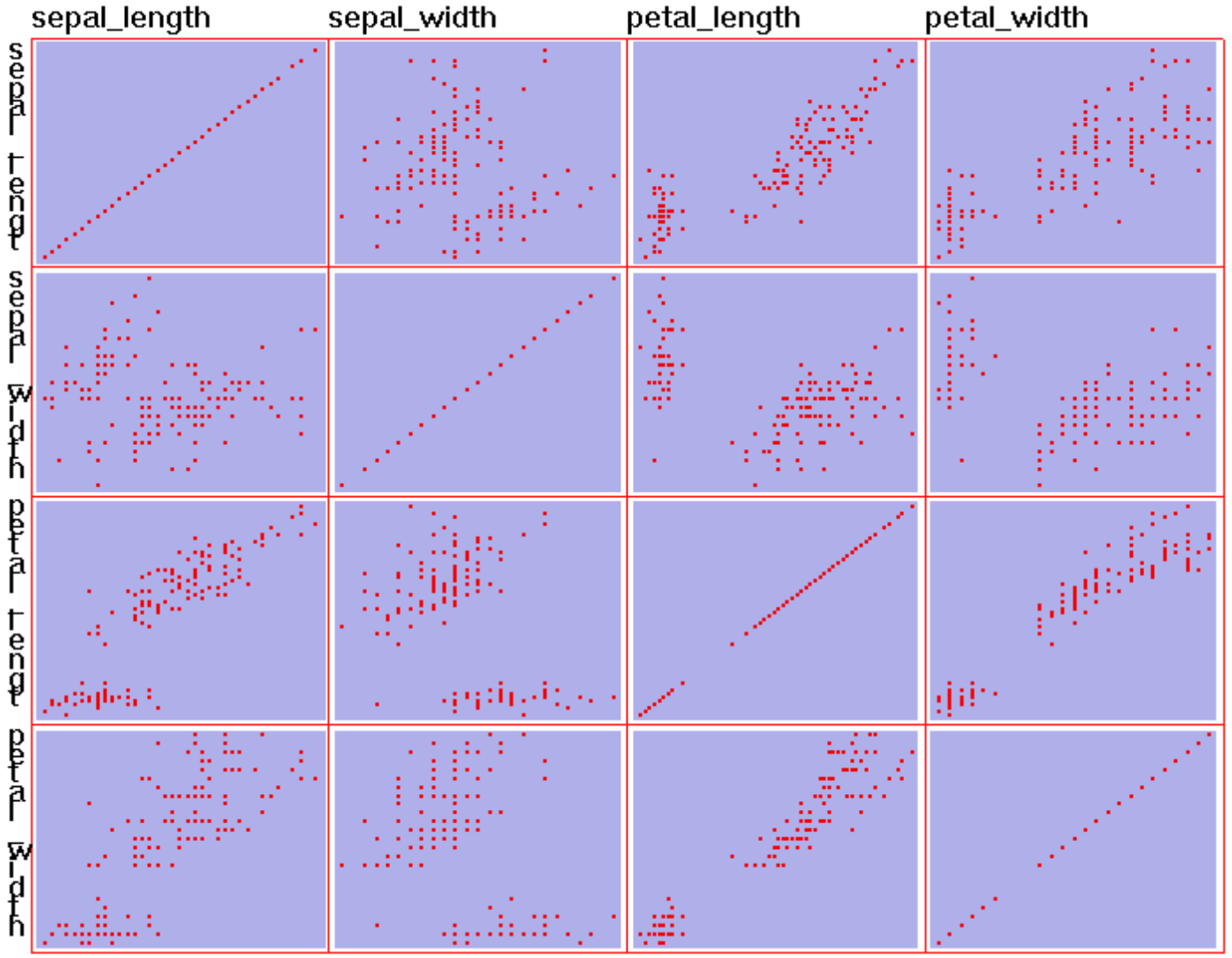
load fisheriris;

Scatter Plot Matrices

- For table data of M variables, we can look at pairs in 2D scatter plots



Scatter Plot Matrix – Iris Data Set



Scatter Plot

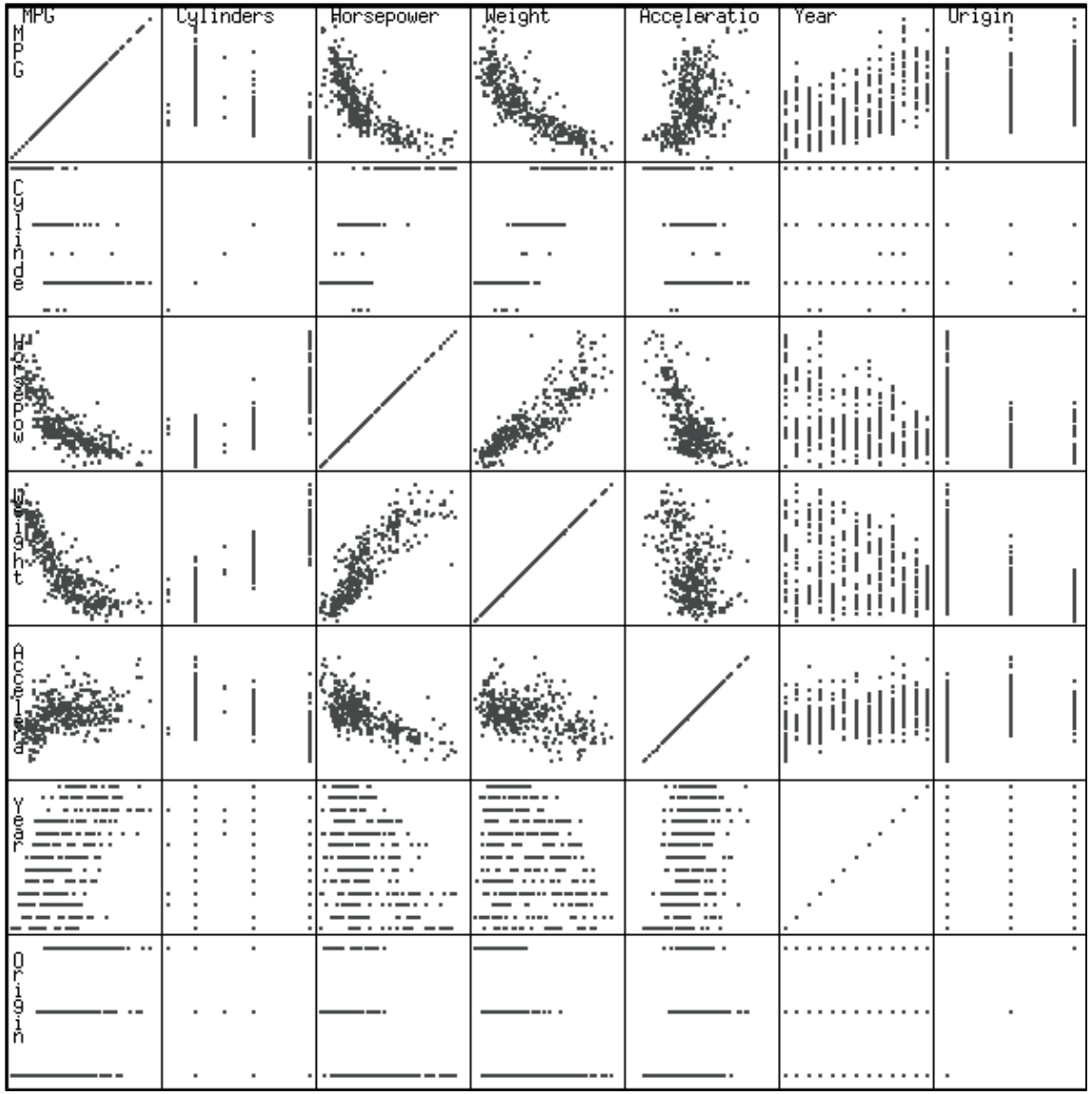
Advantages

+ Useful for looking at all possible two-way interactions between dimensions

Disadvantages

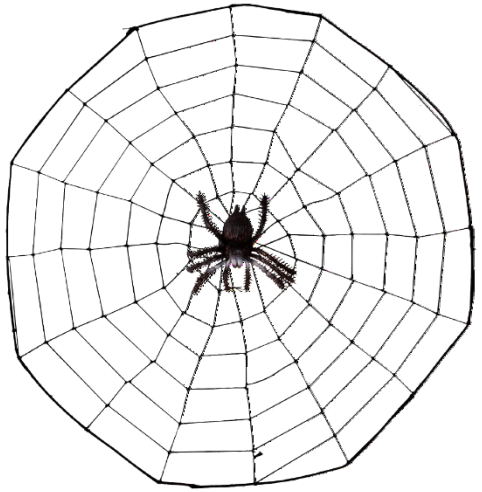
- Becomes inadequate for medium to high dimensionality

Scatter Plot Matrix – Car Data Set



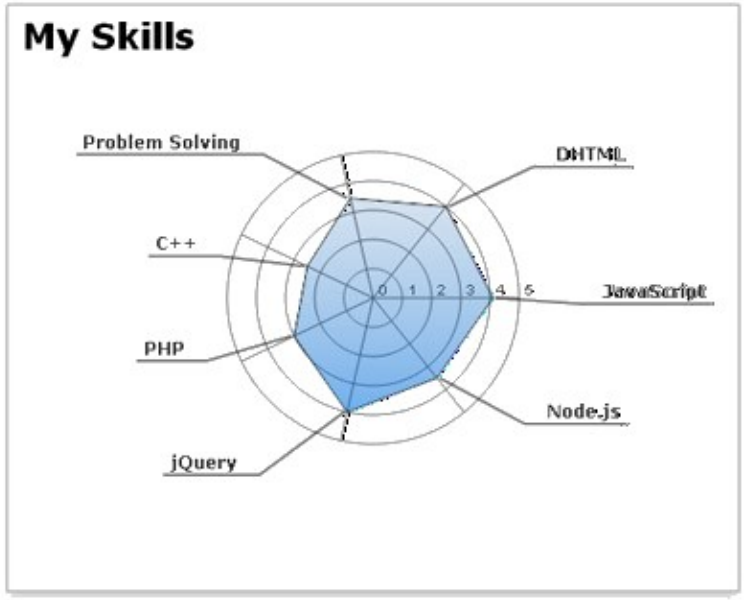
Data represents
7 aspects of cars:
what relationships
can we notice?

Radar Chart



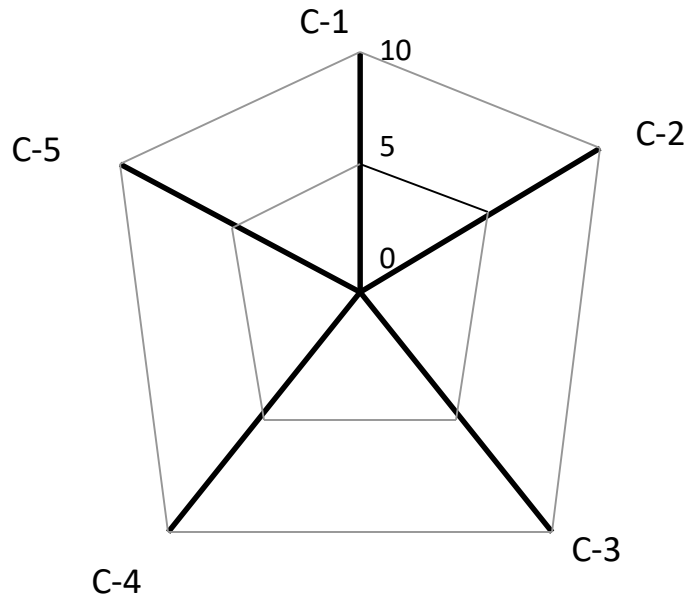
- A graphical way to compare data
- Displayed in a “web-like” form (spider chart)
- Used to evaluate multiple alternatives based on multiple criteria

Radar Chart (in real applications)



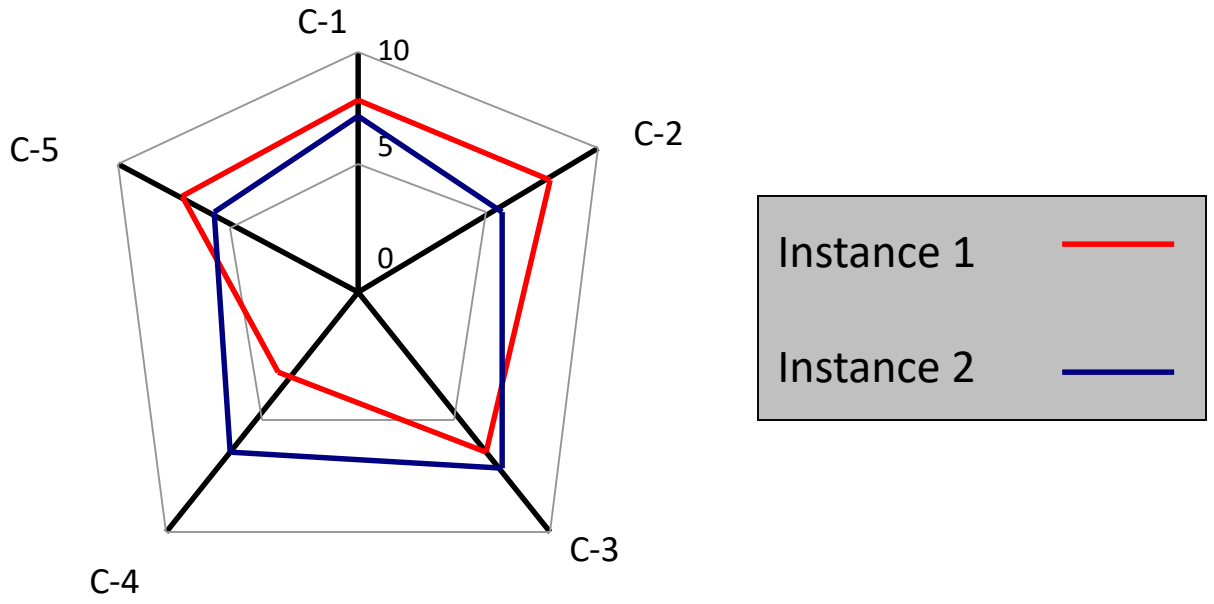
Radar Chart

- Draw and label the axis arms of the chart (one arm for each criterion)
    - If there are five criteria (C) on a scale of 0-10:



Radar Chart

- Draw and label each alternative's ratings on the chart, connecting between arms
    - Using different colors for each instance is best



Radar Chart

Advantages

+ Combines properties of glyphs and parallel coordinates making pattern recognition easier
+ Compact

Disadvantages

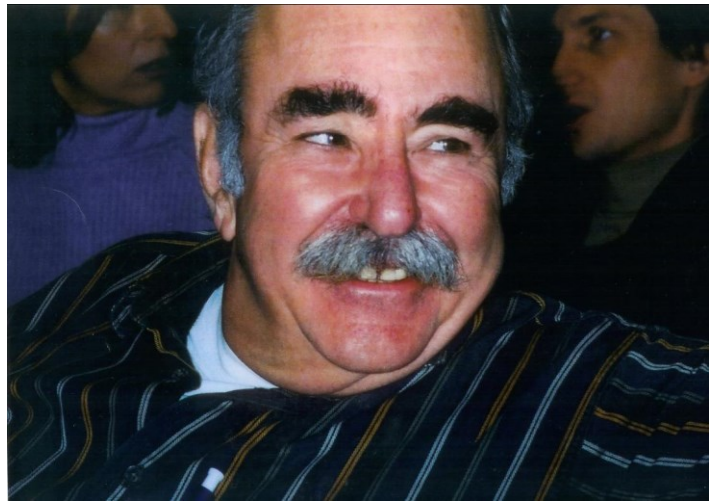
- Cluttering near center
- Harder to interpret relations between each pair of dimensions than parallel coordinates

Parallel Coordinates

- Concept due to Alfred Inselberg
- Conceived the idea as a research student in 1959…
- … idea gradually refined over next 40 years

Al - 1959

Paper "Parallel Coordinates for Visualizing Multi-Dimensional Geometry"

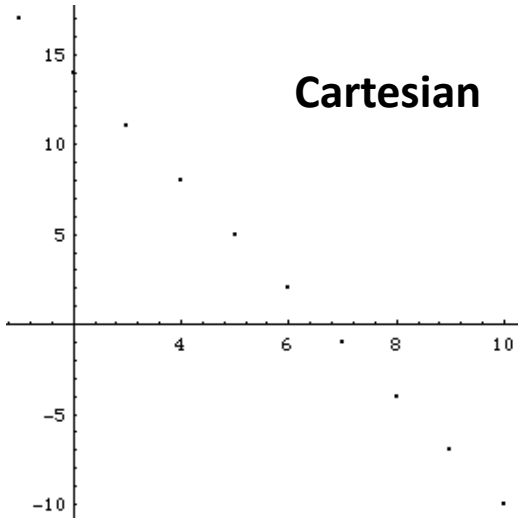


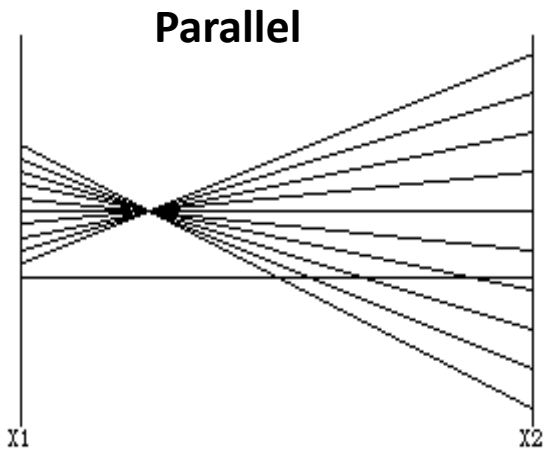
Al - now

http://www.math.tau.ac.il/~aiisreal/

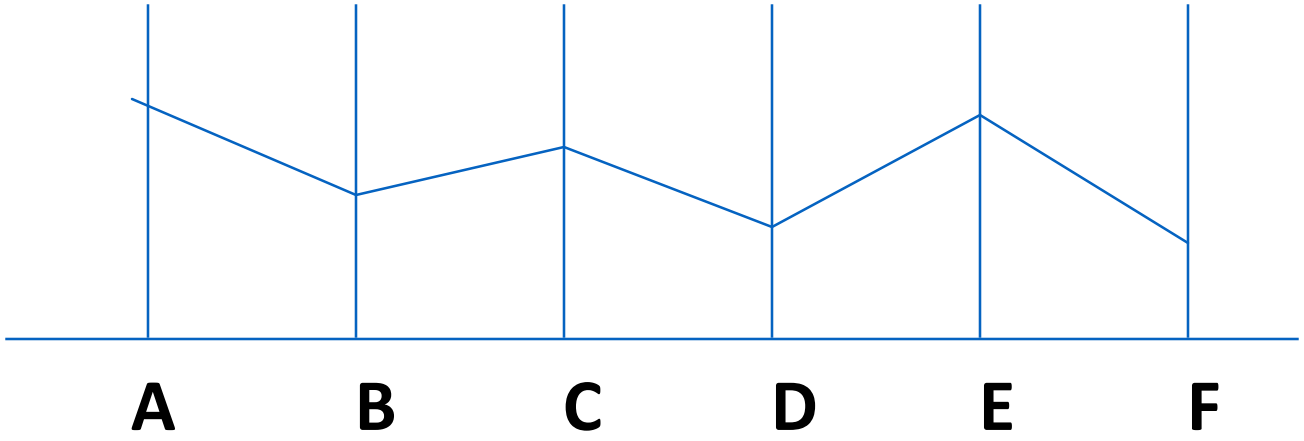
Cartesian vs. Parallel Coordinates

- Cartesian Coordinates:
  - All axes are mutually perpendicular
- Parallel Coordinates:
  - All axes are parallel to one another
  - Equally spaced



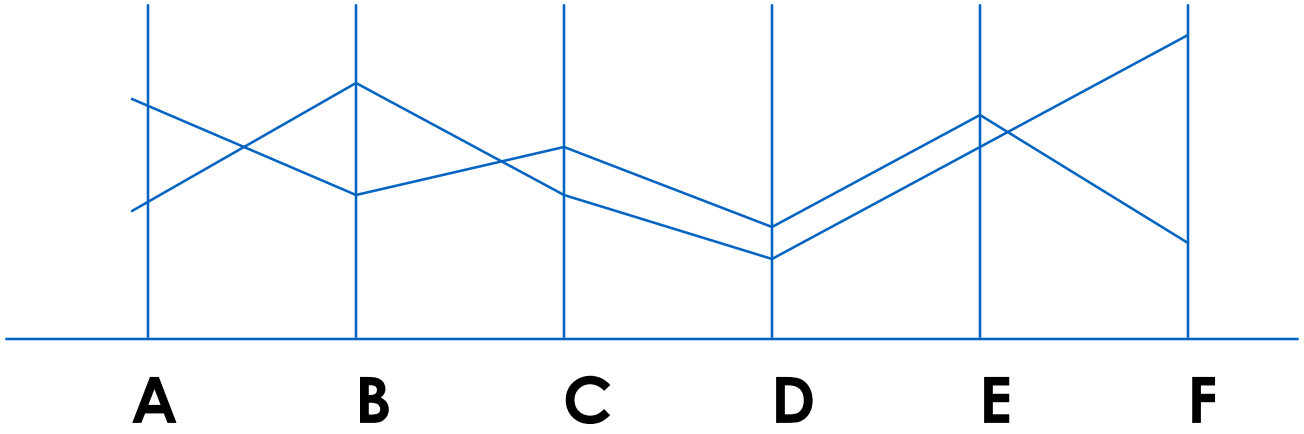


Parallel Coordinates



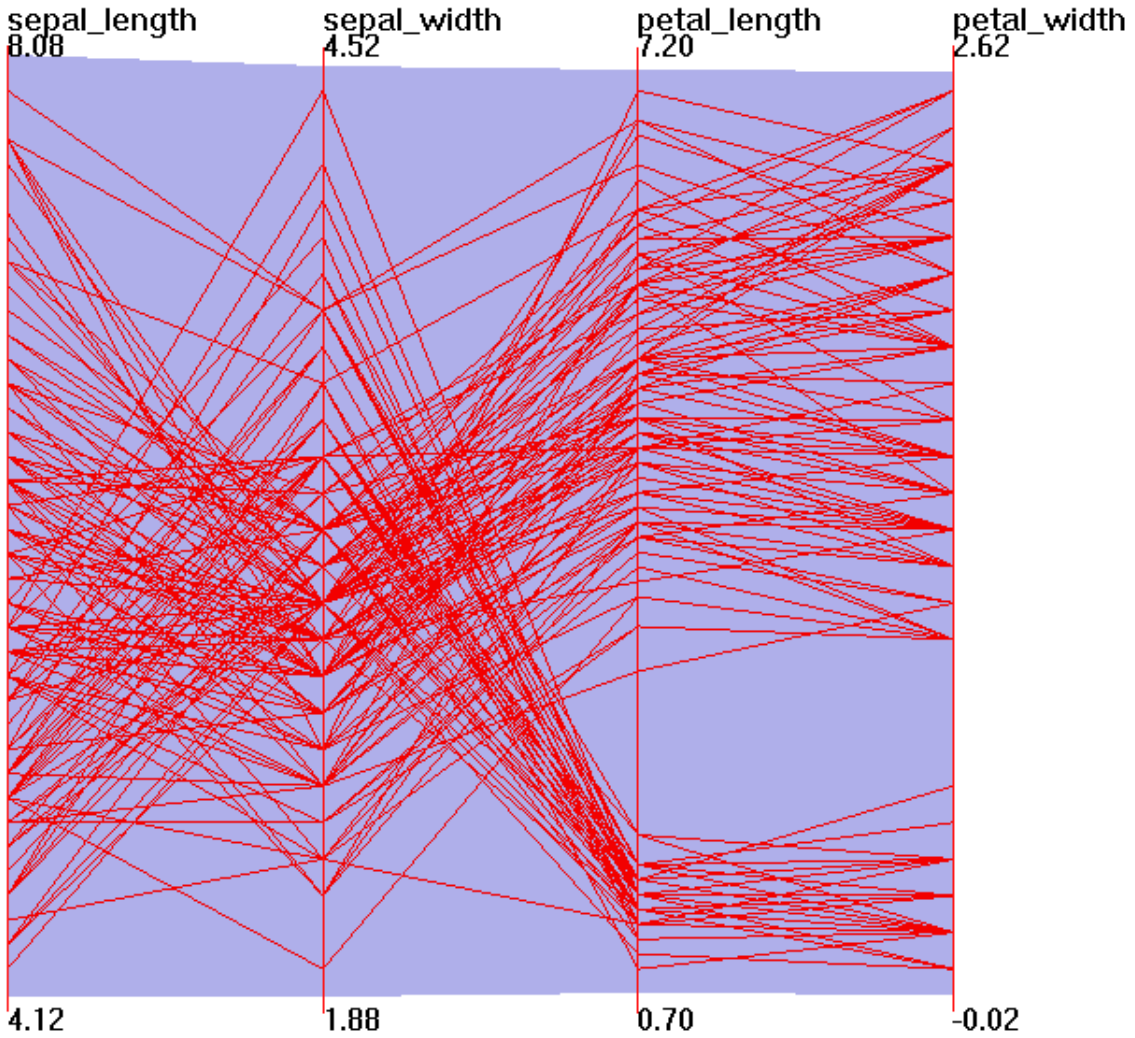
- create M equidistant vertical axes, each corresponding to a variable
- each axis scaled to **[min, max]** range of the variable
- each observation corresponds to a line drawn through point on each axis corresponding to value of the variable

Parallel Coordinates



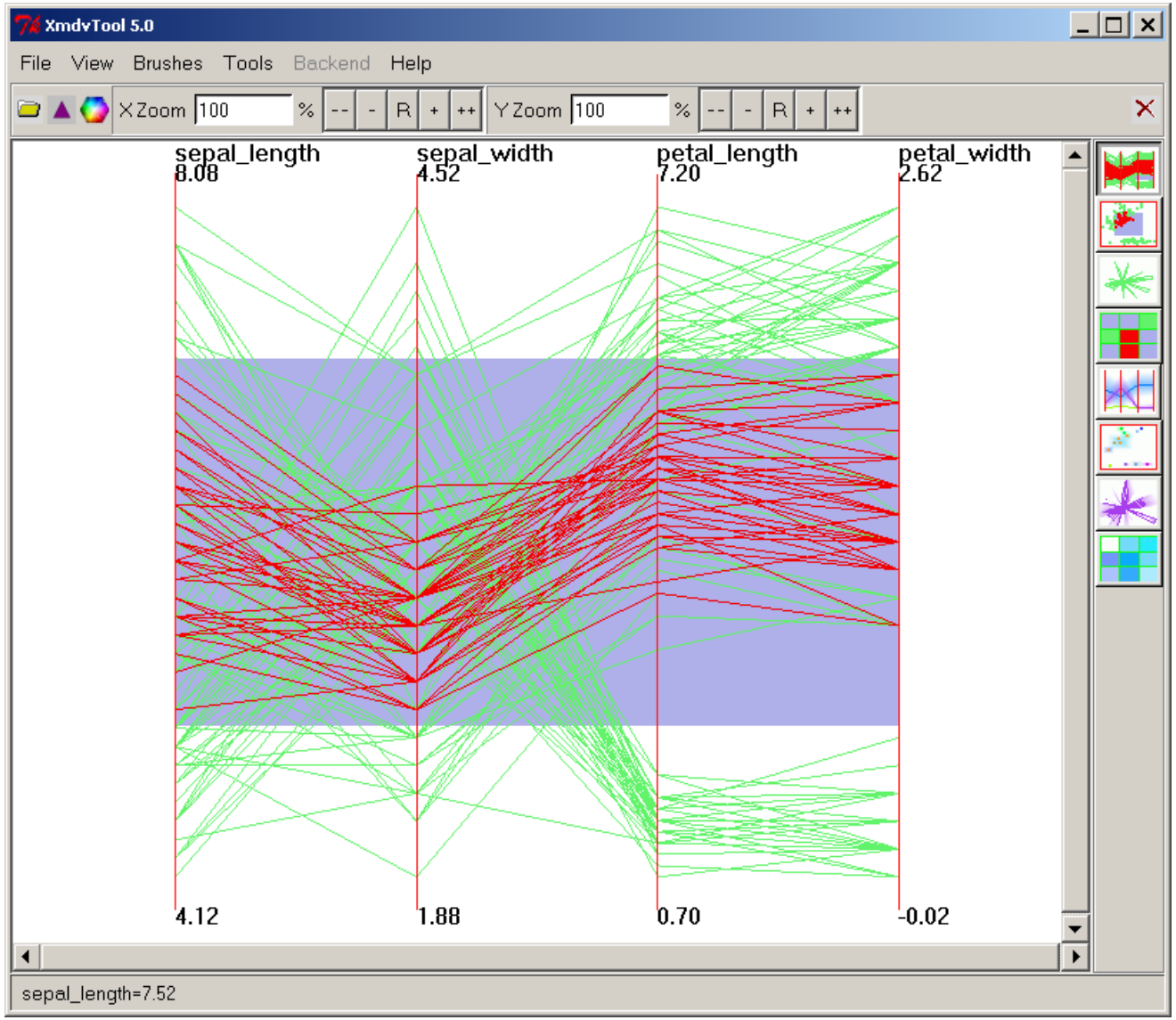
- correlations may start to appear as the observations are plotted on the chart
- for example: there appears to be negative correlation between values of A and B

Parallel Coordinates – Iris Data

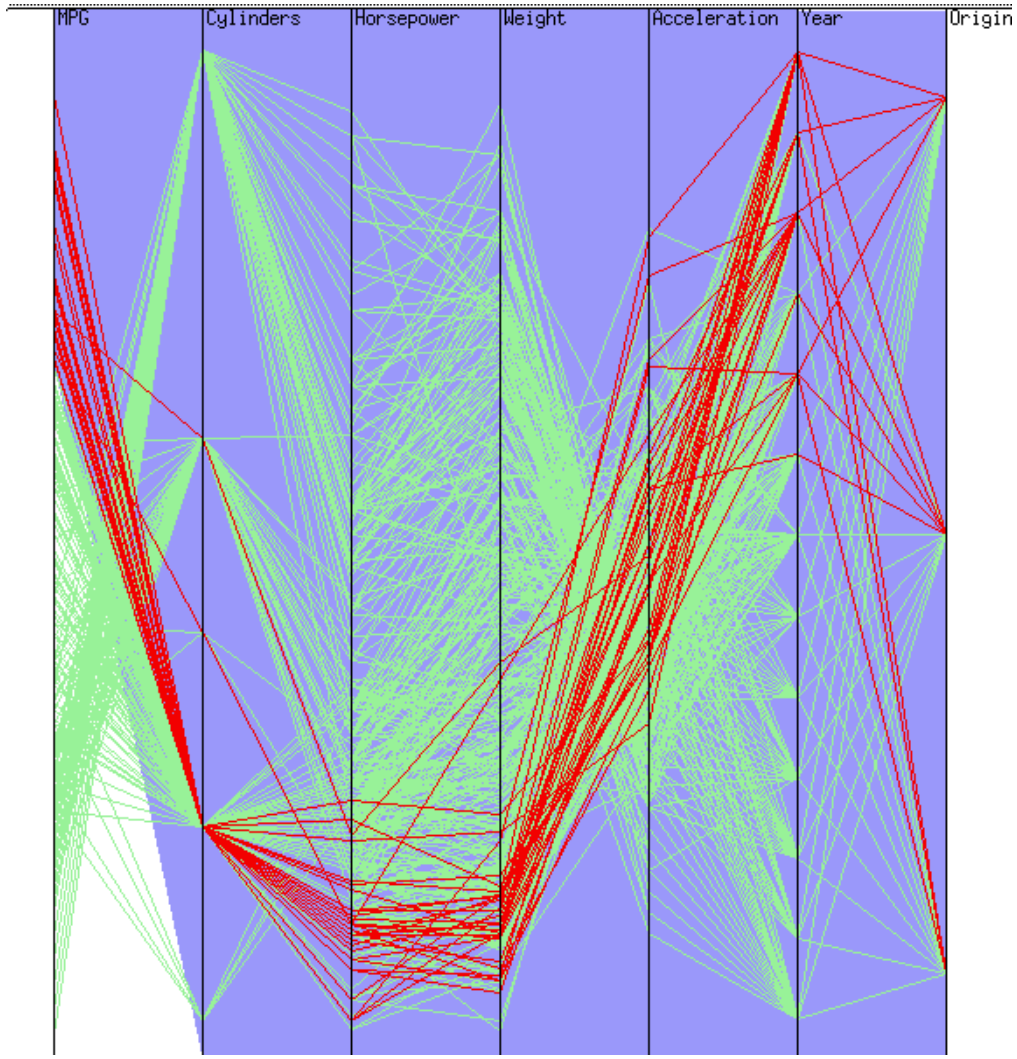


Brushing as a Solution

- **Brushing** selects a restricted range of one or more variables
- Selection then highlighted



Parallel Coordinates



Brushing picks out the high MPG data

Parallel Coordinates

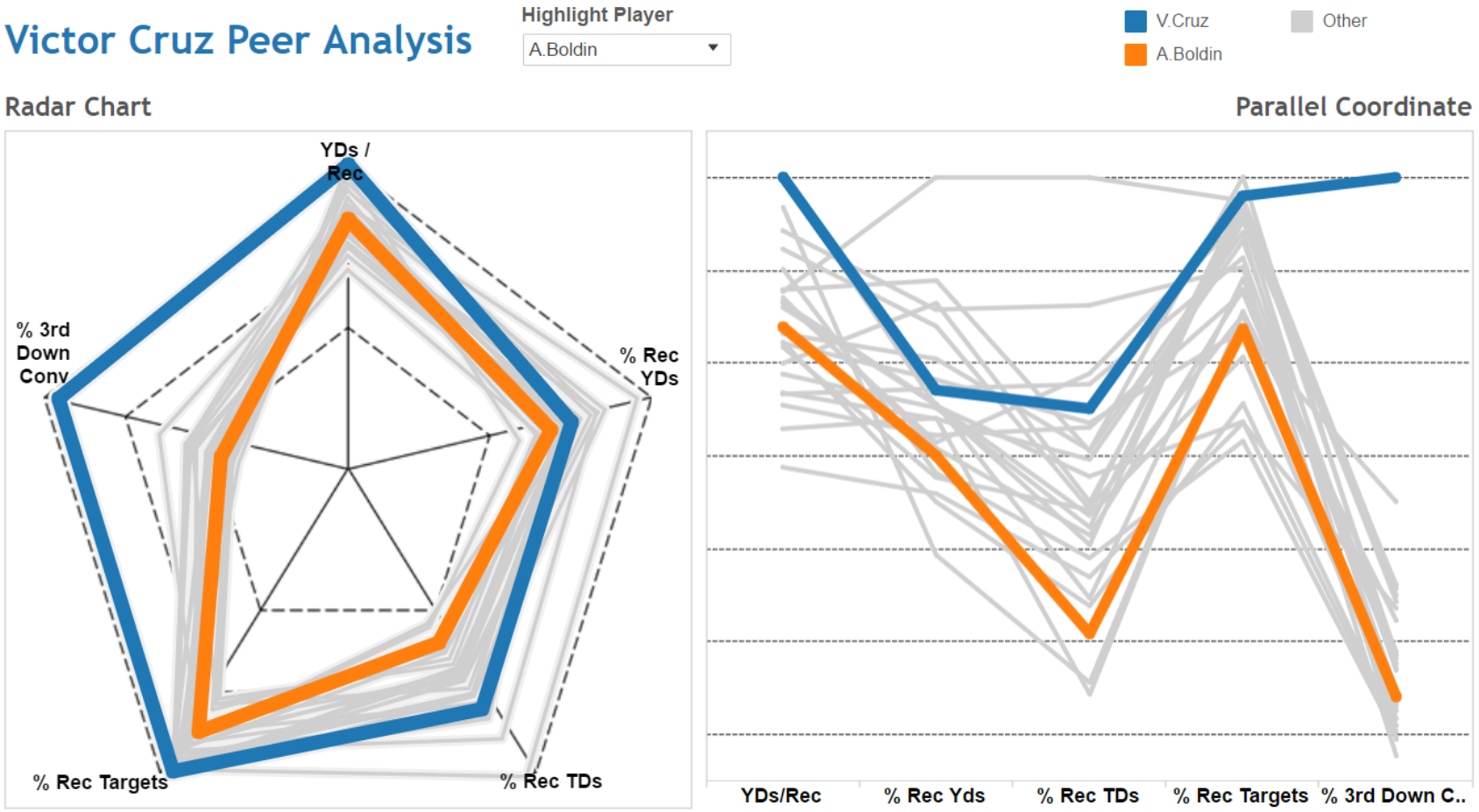
Advantages

+ Many connected dimensions are seen in limited space
+ Can see trends in data

Disadvantages

- Become inadequate for very high dimensionality
- Cluttering

Radar Chart vs. Parallel Coordinates



Q&A