To prepare for test: in addition to your own notes and the Cleveland text ...

1. Todays lecture
2. See also two previous tests:
    They are given with answers in the notes for 2003 Oct 10 and Oct 20.

Test is mainly on Ch 4 but may include questions relating to the Bimbo exercise case study.

1. What are the new questions that arise in the data analysis of trivariate data as compared to bivariate data?

2. Why does a matrix plot contain redundant plots?

3. What is the purpose of Fig 4.2 on p 183 as an introduction to chapter 4?

4. Why is the loess fit added to the panels of Fig. 4,3 p 185.?

5. What considerations determine the alpha value for the loess fits, for example in Fig 4.3 and 4.4/

6. The obvious plot for an n x p data matrix is a matrix plot. What is it that Fig 4.5 does not reveal?

7. With reference to Fid 4.6, what is an interaction?

8.   In Fig 4.6, what determines the widths of the ER intervals?

9. ER and CR interact in determining NOX.   Using Fig 4.7, what is this interactions?

10. What is the role of "brushing" in data analysis?
  Comment with respect to the data set presented on pp 192-193.

11. Why is the coplot of NOX vs CR|ER different from the coplot of the estimated surface NOX = lin combo of CR an ER?

12. Cropping as in Fig 4.15: Why is it advisable?

13. What role does loess have in examining a residual plot?

14. Why is the aspect ratio a consideration in examining residual plots?

15. Why do you think the data of Fig 4.43 was collected in straight lines?

16. In the context of the data, what does Fig 4.,44 say?

17. What is the novelty in Fig 4.46 compared with Fig 4.,45?

18. Describe the relationship of Fig 4.50 And FFig. 4.55.

19. What is the discussion on pp 242-244 about?

20. For practical purposes, when would wireframe plots be most useful and when would contour plot of surfaces be more useful?"

21. Why do Fig.s 4.60, 4.61, 4.62 seem to be so different?

22. What is the role of the streriogram in Fig 4.63?

23,. The colourplot on p 266 is appealing to the eye. What is the role of colour in this graph? Would more colours improve the display? Explain.

24. Is Fig 4.69 (p 269) better or worse than the coplot p 265? Discuss briefly.

25. What is the most important idea in Ch 4?

Data from R:

```
LifeCycleSavings
Format

A data frame with 50 observations on 5 variables.
```
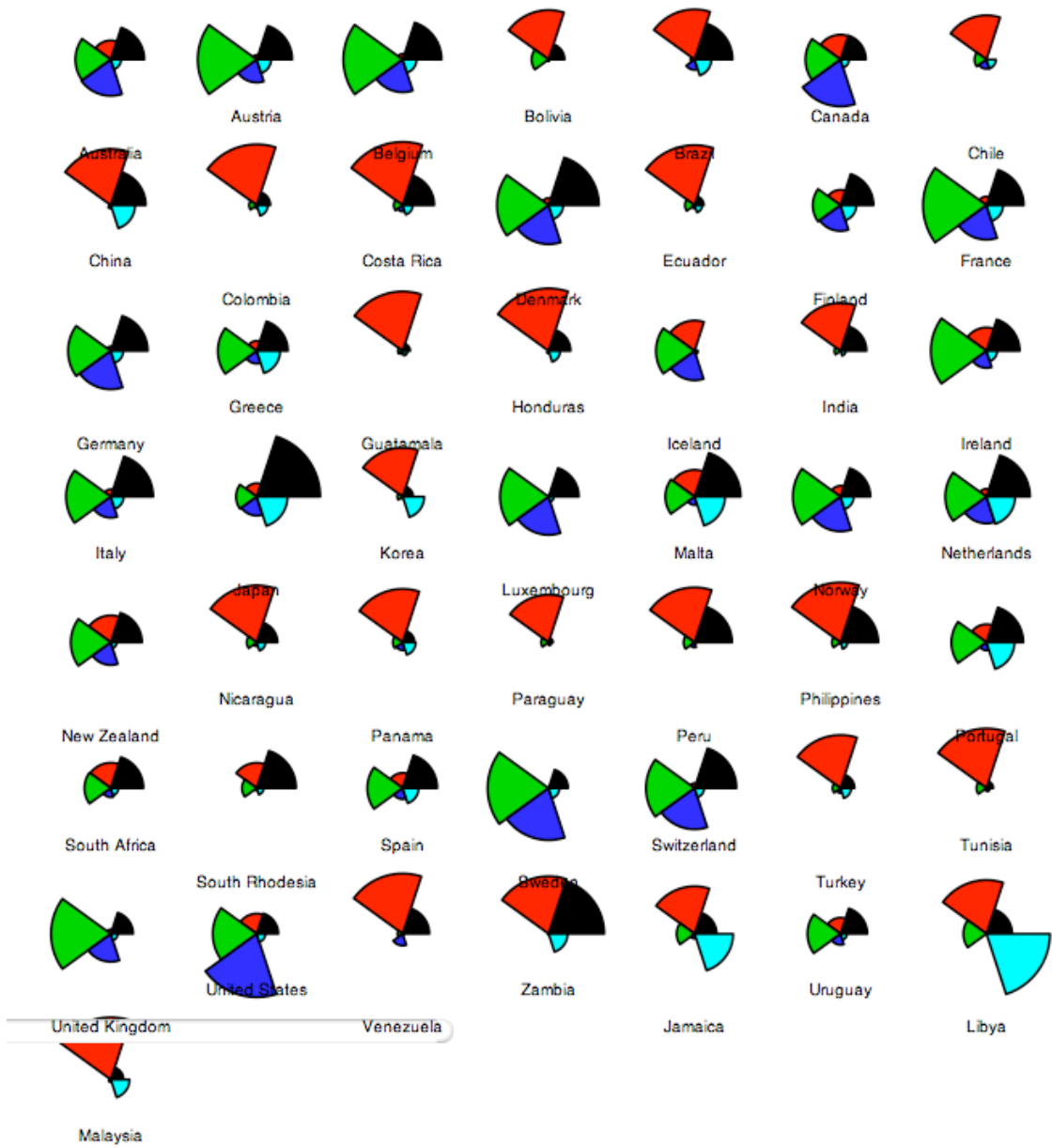
[,1]	sr	numeric	aggregate personal savings
[,2]	pop15	numeric	% of population under 15
[,3]	pop75	numeric	% of population over 75
[,4]	dpi	numeric	real per-capita disposable income
[,5]	ddpi	numeric	% growth rate of dpi

Details

```
Under the life-cycle savings hypothesis as developed by Franco
Modigliani, the savings ratio (aggregate personal saving divided by
disposable income) is explained by per-capita disposable income, the
percentage rate of change in per-capita disposable income, and two
demographic variables: the percentage of population less than 15 years
old and the percentage of the population over 75 years old. The data
are averaged over the decade 1960–1970 to remove the business cycle or
other short-term fluctuations.
```

Austria
Bolivia
Canada
Australia
Belgium
Brazil
Chile
China
Costa Rica
Ecuador
France
Colombia
Denmark
Finland
Greece
Honduras
India
Germany
Guatamala
Iceland
Ireland
Italy
Korea
Malta
Netherlands
Japan
Luxembourg
Norway
Nicaragua
Paraguay
Philippines
New Zealand
Panama
Peru
Portugal
South Africa
Spain
Switzerland
Tunisia
South Rhodesia
Turkey
United States
Zambia
Uruguay
Libya
United Kingdom
Venezuela
Jamaica
Sweden
Malaysia

Q: How would you organize this data to better show the ordering induced by the development status of a nation?