

# A Problem-Solving Course in Statistics for Mathematical Science Student

Danny Parsons, Roger Stern, David Stern

IDEMS International, UK

University of Reading & Statistics for Sustainable Development, UK

An “eye opening” course...

that aims to motivate and instill an interest in statistics...

## The Course

exposing students to the role of statistics in solving real problems with data...

presenting the broader processes within statistics including design, collection, management and analysis of data and interpretation and presentation of results.

# Aims and Principles

Students should:

- understand the role of statistics in solving real problems with data
- be comfortable with producing and interpreting simple descriptive tools
- understand the broader processes within statistics from design and collection to interpretation and presentation of results

Support students to:

- engage students in statistics through experiential learning
- use real world problem and research related projects that have genuine complexity (without single “correct” answers) and can be adapted to the level and experience of the students involved.





# Background

African Institute for  
Mathematics Sciences  
(AIMS) Cameroon

“Skills course” at start of  
very broad Mathematics  
Science MSc with stats  
stream supported by RSS

Aim to prepare students with  
skills to solve problems that  
can contribute to the  
development of Africa

# Student Diversity & Background

- 47 MSc students from 12 African countries
- From no statistics courses to full statistics degrees
- Theoretical statistics teaching – analysis focused
- Little use of data/software
- Limited descriptive stats

An “eye opening” course...

- *38/47 students said their view of statistics had changed because of the course*
- *“I now know that statistics is more practical than theoretical as taught in schools”*
- *“Computers can make calculations, but it is for the statistician to give to those calculations a real signification.”*

that aims to motivate and instill an interest in statistics...

- *“this course gave me the interest in statistics, I didn't know before that statistic is a very good field”*
- *“statistics has never been more fun”*

## The Course

exposing students to the role of statistics in solving real problems with data...

- *40 out of 47 students said the course prepared them “a lot” to use statistics to solve problems*
- *“I now understand what a typical statistician should do”,*
- *“now I find that it is a promising field that will help Africa to solve several problems”*
- *“it is useful in solving the problems of society.”*

presenting the broader processes within statistics including design, collection, management and analysis of data and interpretation and presentation of results.

- *“statistics is not all about Data Analysis but it starts even before the problem, data collection until the final conclusion”*

# Week 1

- Broad definition of statistics that includes design, collection, organisation, interpretation and communication
- What does a statistician do? Seven consultancy questions – mostly concerning design.
- Two (analogue) statistical games simulating an experiment and a survey - design the study, collect data, enter data, analyse and then write a report in the style of a short paper

## Tutorials

- Introduce pivot tables in a spreadsheet & a statistics package (R-Instat)

## Quizzes

- What is standard deviation?

## Question

The photograph below shows a crowd of soccer supporters at a match.

The mean weight of the supporters is 80 kg; guess their approximate standard deviation.



Approx standard deviation is  kg

Check

Tell me

Another question

## Message

Guess the standard deviation of the weights of the supporters then type it into the text-edit box above.

# What is standard deviation?

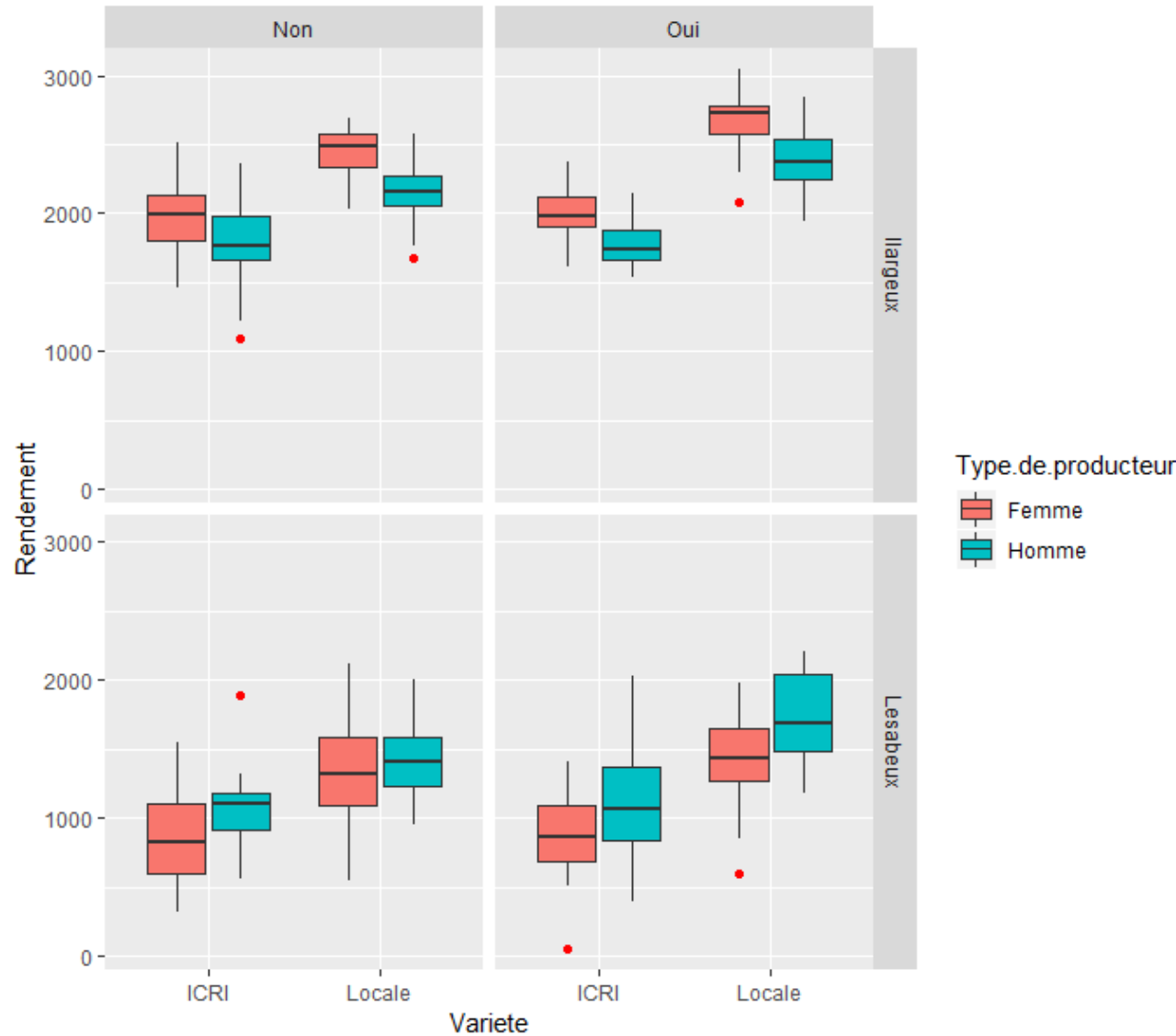
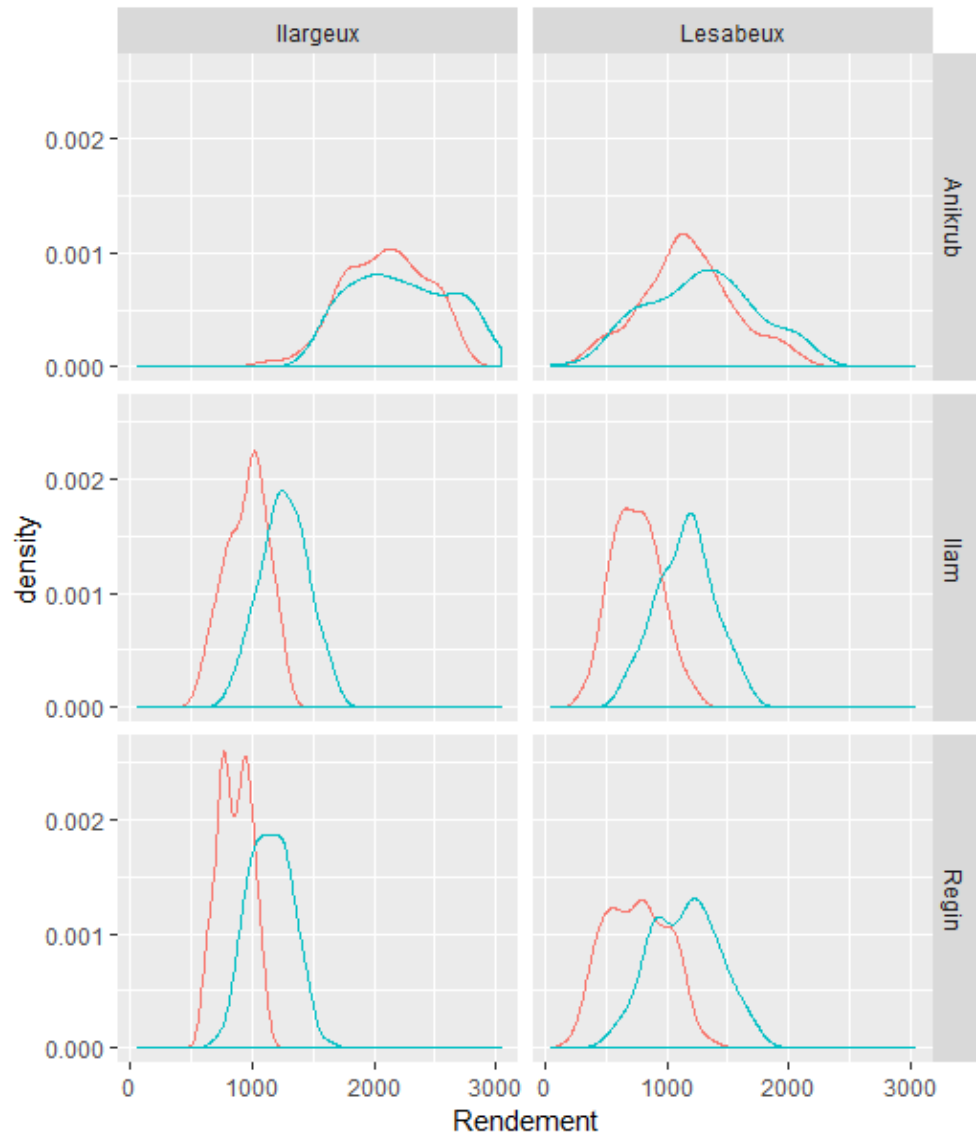
Source: CAST Computer-Assisted Statistics Textbooks, Doug Stirling  
<http://cast.massey.ac.nz/>



# Week 2

- Working with data using a spreadsheet and a statistics package & interpreting results
- Randomised, simulated data from large (digital) version of experimental “game”. A different dataset of 1000 plots for each student. Based on a current study in Niger, West Africa.
- A good analysis explains variability – ANOVA as a descriptive tool
- A good analysis relates to the objectives of the study

# Visualisations from experimental game simulated data



# Week 3

8 project topics chosen out of 14:

- Corruption “red flags” in public procurement. Using open World Bank data (200,000 records from over 140 countries.)
- Cameroon climatic data analysis using daily data from Cameroon Met Service.
- Analysis of timber trees using multi-level data from farms and plots.
- Designing and administering a survey using ODK.
- IFAD poverty survey data of 1,300 respondents and over 400 variables from 2018 in Lesotho.
- Moving into R. This used a guide on how to write R commands from within R-Instat, and how to transfer to writing scripts in R/RStudio itself.
- Tidy data. Based on a paper by Hadley Wickham and including his data from Mexico (500,000 cases) + messy Cameroon climatic data.
- Data from a 2017 on-farm trial of low cost fertilisers involving 1,700 mainly women farmers in Niger.

# Cameroon climatic data analysis

The data are from Garoua where we have data for 3 elements, namely tmax, tmin and rainfall. The data from Douala are just for rainfall, though we have requested data for other elements.

In each case, the data as provided are not “tidy”. This will be discussed and you may also like to consider the issue of tidy data with the group investigating that topic for their assignment.

- a) In each case the data go “across” with a single row for each month and a column for each day of the month.
- b) The Garoua data are in separate Excel files for each element and each file contains multiple sheets with just a few years on each sheet. The data also contain some curious symbols. It seem likely the data were obtained by OCR, but with no subsequent editing of strange symbols.
- c) The Douala data has more minor complications, including non-numeric values, e.g. TR and – within the data. There are also blank lines for missing years and a second file that has some overlapping data, but that completes some of the missing years. There are also many missing months in the recent data – for an airport this is difficult to understand.

- Examine how to present the results with clear graphs that are useful and also simple enough to be shared with users, including farmers.

- Examine the possible trends in the temperature data.
- Are they consistent through the year or are some months changing faster than others?
- Consider summaries of the rainfall data that could be of interest to farmers. Including the start and end of the rains, rainfall totals and number of rain days. Consider carefully how to cope with missing values.

Prerequisites are not a barrier to students working on interesting, complex, real world data problems

- *Contrast between challenges with basic concepts in Week 1 and complexity of projects tackled in Week 3*

**But..!**

- *“I do not know the purpose of the course and (it) is not theory, but just practical.”*

## Key lessons

The problems can be the same, the approach can be adapted to students' previous experience

- *Games + projects from Week 3 also used with high school students*

Problems with genuine complexity and without single “correct” solutions can engage students in statistics

- *“now I find that it is a promising field that will help Africa to solve several problems”*
- *“it is useful in solving the problems of society.”*

Students can have a fun and enjoyable experience with statistics in a short period without previous experience or interest

- *“This course was wonderful. The fact of having done work ranging from data collection to analysis and then talking to us about data backup has completely opened my mind and made me see statistics from a different angle.”*
- *“They (lecturers) were able to stimulate my interest in just three weeks. I wonder if they had more weeks, then I might have forfeited my love for mathematics, just kidding”*

# What's next?

- Improve & teach the course again in October 2019
- Share “statistical problem solving” concept with others - AIMS network, undergraduate, school students
- How does statistical problem solving fit into statistical and data literacy?
- How could this course and other similar inspiring & motivating courses reach more people?
- Could/should the content be adapted for a MOOC?



## Acknowledgements

Thanks for the Royal Statistical Society for supporting the series of statistics courses delivered at AIMS Cameroon.

Thanks for the students, tutors and all staff at AIMS Cameroon for giving us such an enjoyable and fulfilling teaching experience.