Functional Data Analysis

Jiguo Cao, PhD

Canada Research Chair in Data Science

Director of the Pacific Blue Cross Health Informatics Laboratory

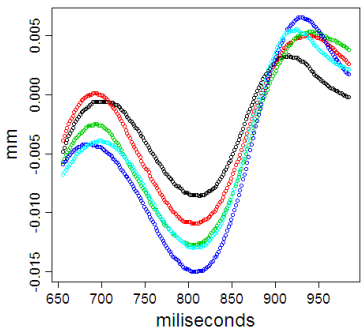
Associate Professor, Department of Statistics and Actuarial Science

Associate Faculty Member, School of Computing Science

Simon Fraser University, Vancouver, Canada

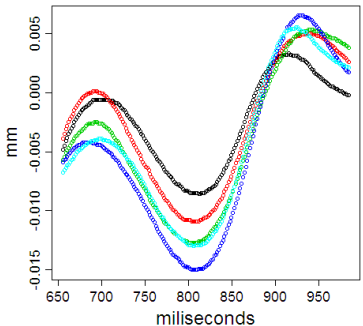
What is Functional Data?

What are the most obvious features of these data?



What is Functional Data?

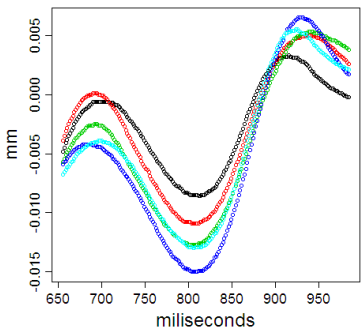
What are the most obvious features of these data?



- quantity

What is Functional Data?

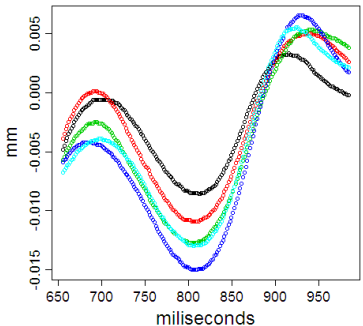
What are the most obvious features of these data?



- quantity
- frequency (resolution)

What is Functional Data?

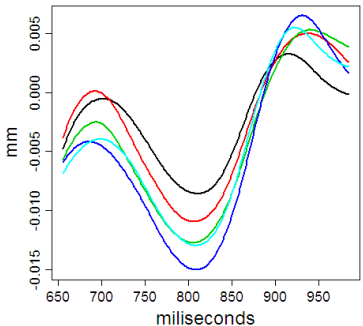
What are the most obvious features of these data?



- quantity
- frequency (resolution)
- similar trends

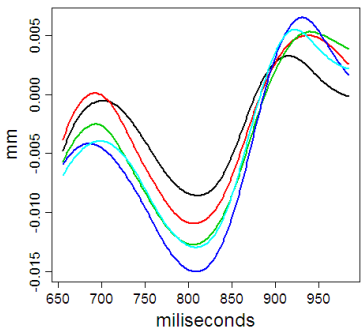
What is Functional Data?

Most important: **smoothness**



What is Functional Data?

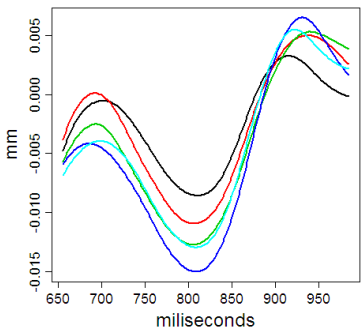
Most important: **smoothness**



These data describe (nearly) a process that changes smoothing, and continuously over time.

What is Functional Data?

Most important: **smoothness**

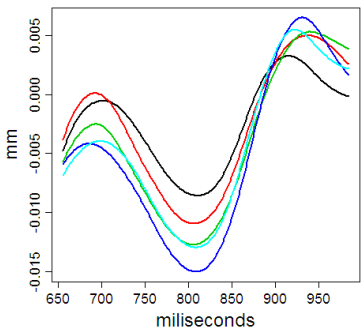


These data describe (nearly) a process that changes smoothing, and continuously over time.

Functional Data Analysis = Analysis of data that are functions.

What is Functional Data?

Most important: **smoothness**



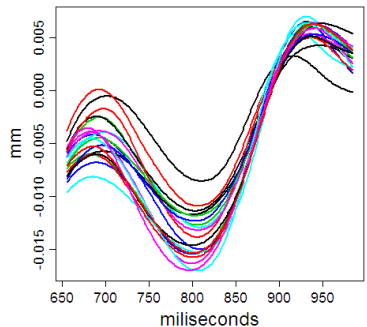
These data describe (nearly) a process that changes smoothing, and continuously over time.

Functional Data Analysis = Analysis of data that are functions.

Domain is usually time, but can be anything: space, energy ...

What is Functional Data?

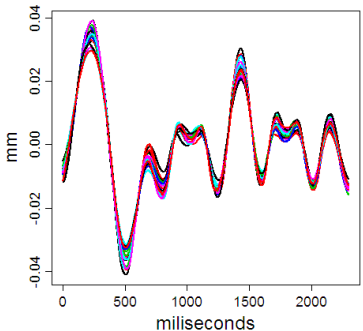
20 replications



Functional data analysis involves repeated measures of the same process.

What is Functional Data?

20 replications, 1401 observations within replications

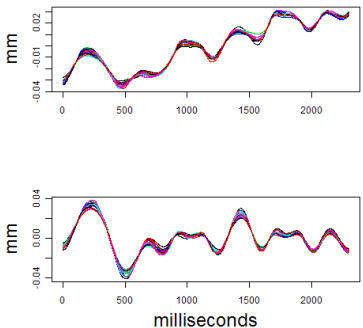


Functional data is often complicated:

- not easily described by mathematical formulae
- variation between replications even harder to describe

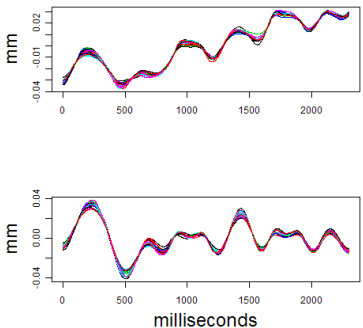
What is Functional Data?

20 replications, 1401 observations within replications, 2 dimensions



What is Functional Data?

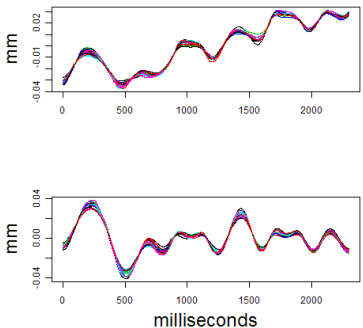
20 replications, 1401 observations within replications, 2 dimensions



Functional data is often complex:

What is Functional Data?

20 replications, 1401 observations within replications, 2 dimensions

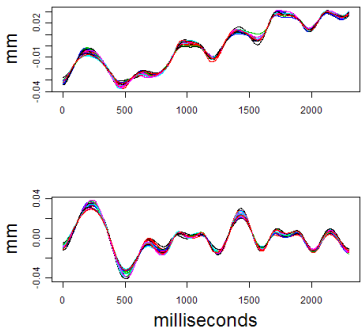


Functional data is often complex:

- often a large number of related quantities

What is Functional Data?

20 replications, 1401 observations within replications, 2 dimensions

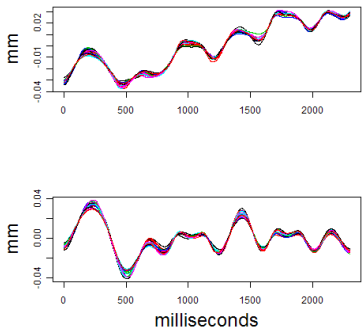


Functional data is often complex:

- often a large number of related quantities
- viewing each replication as a single observation can make the data easier to think about (once we have the right machinery)

What is Functional Data?

20 replications, 1401 observations within replications, 2 dimensions



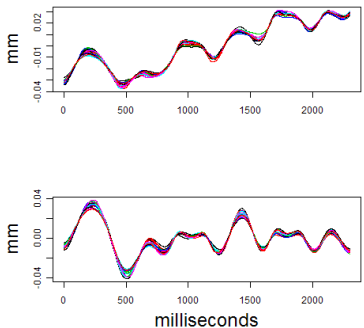
Functional data is often complex:

- often a large number of related quantities
- viewing each replication as a single observation can make the data easier to think about (once we have the right machinery)

- What are these data, anyway?

What is Functional Data?

20 replications, 1401 observations within replications, 2 dimensions



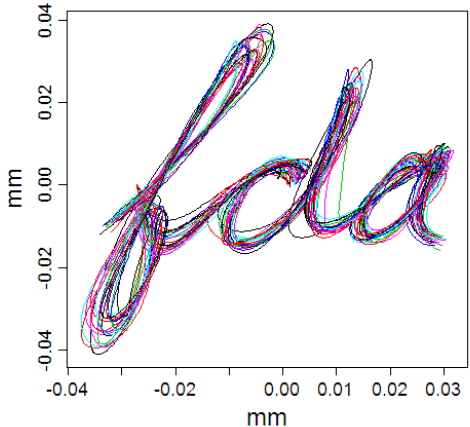
Functional data is often complex:

- often a large number of related quantities
- viewing each replication as a single observation can make the data easier to think about (once we have the right machinery)

- What are these data, anyway?
- What if I plot one component against another?

Classical Functional Data

Measures of position of nib of a pen writing "fda". 20 replications, measurements taken at 200 hertz.



Characteristics

- Data are measurements of smooth processes over time
- We usually do not want to make parametric assumptions about those processes.
- Often have multiple measurements of the same process
- We are interested in describing the variation of processes.
- Frequently, collected data have high resolution and low noise.
- Can be applied to any estimate of a smooth process.

About Functional Data Analysis

1. FDA is New
   - First named in Dalzell & Ramsay, 1991
   - Relatively little penetration into applied fields (= easy publication)
   - Several competing methodologies (we focus on one)
   - Limited public software/resources
   - data analysis rather than inference
2. Functional Data is Complex
   - Requires more thought/judgement than a t-test
   - data needs pre-processing
   - parametric inference is rarely available/appropriate

What will this course offer?

Audience: application areas with functional data

Focus:
- What can Functional Data Analysis do?
- How do I make it happen?

Software: packages in `R`, Matlab

Goals: Enabling you to
- Understand and interpret the result of FDA applied to real data
- Use existing FDA libraries to analyze functional data
- Evaluate its usefulness/correctness
- Extend the methods in existing software if you need to

Not Covered: reproducing-kernel Hilbert spaces, asymptotics, theorems...

Pre-requisites and Recommendations

Pre-requisites: multiple linear regression

Useful: Life will be easier if you do not need to learn some of the following:

- R/Matlab or other programming experience
- Calculus
- Matrix algebra
- Multivariate statistics
- Computational statistics

Any necessary material will be covered in class, but will be out of context.

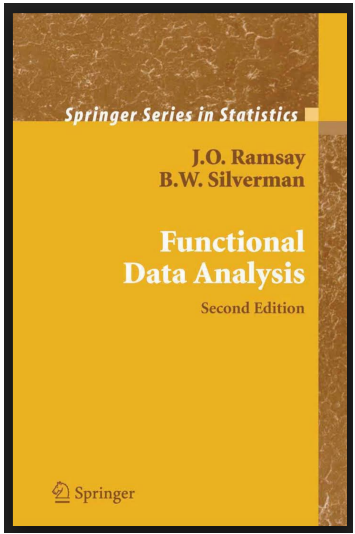
Resources

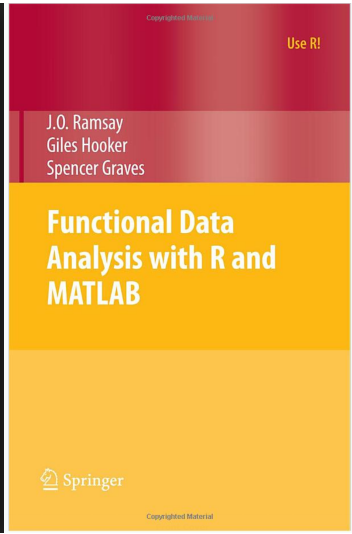
Textbook: Ramsay and Silverman, 2005, Functional Data Analysis, Springer.

Books:
- Ramsay and Silverman, 2002, Applied Functional Data Analysis, Springer.
- Ramsay, Hooker and Graves, 2009, Functional Data Analysis with R and MATLAB, Springer

Online:
- http://www.functionaldata.org for FDA
- http://www.r-project.org a general site for R
- https://github.com/caojiguo/FDAworkshop

  All workshop notes, exercises, computer codes, etc will be posted here.


Springer Series in Statistics
J.O. Ramsay
B.W. Silverman
Functional
Data Analysis
Second Edition
Springer


Use R!
J.O. Ramsay
Giles Hooker
Spencer Graves
Functional Data
Analysis with R and
MATLAB
Springer

Policies

- Questions are always very welcomed. No questions are naive. You are helping other students by asking questions.
- Except asking questions publicly, please do not talk with your neighbours. Please share your questions with all of us. Please respect other students.
- Turn off your cellphone.