



# Exploring data with SAS Insight

A multidimensional  
perspective



**François Mainville**  
**Statistics Canada**

# Today's challenge...

- **Tremendous amounts of data are collected from which...**
  - discoveries will be made
  - decisions will be taken
  - policies will be derived
  - funding will be allocated
  - ...



# Today's challenge...

- **An analyst's perspective**
  - Time available
  - Resources available
  - Tools available vs. required
  
- **The conventional approach...**
  - task -> define needs ->
  - build system -> get results, find answers ...



# Today's challenge...

Can this conventional approach  
answer today's challenges ?

## Dilemma

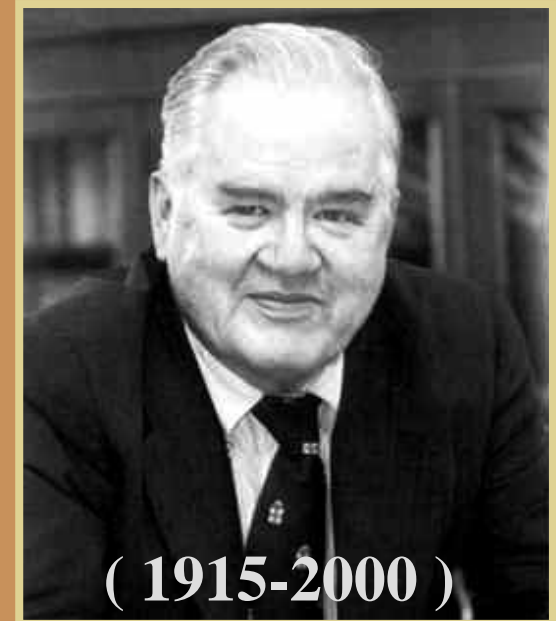
Provide specifications ???

But how can you  
since you don't know  
what your are looking for ...



# A bit of history...

- **John Wilder Tukey**
  - In the 70's,  
laid the foundation of EDA...
  - 25 years later, EDA  
& Informatics crossed roads  
to bring to life Tukey 's work.
  - (SAS Insight and SAS JMP “software”  
are vibrant examples )



( 1915-2000 )

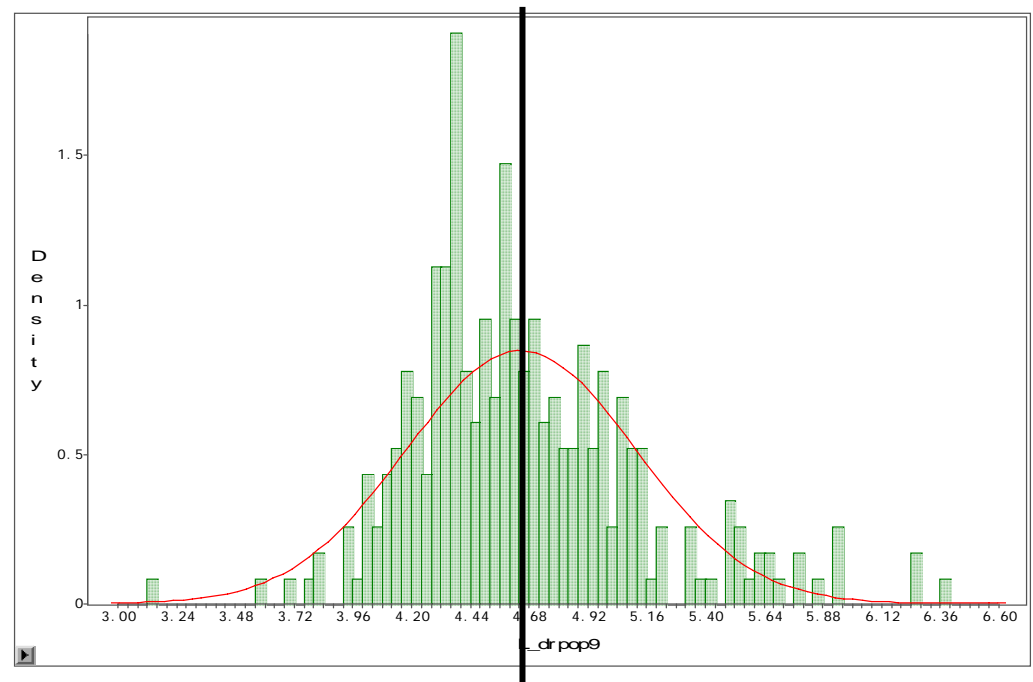


# A bit of philosophy...

“Numerical quantities focus on expected values, graphical summaries on unexpected values”

(J.W. Tukey)

Moment s			
N	288. 0000	Sum of Squares	288. 0000
Mean	4. 6524	Sum	1339. 8975
St d Dev	0. 4709	Variance	0. 2218
Skewness	0. 8128	Kurtosis	1. 4095
USS	6297. 4207	CSS	63. 6522
CV	10. 1225	St d Mean	0. 0278

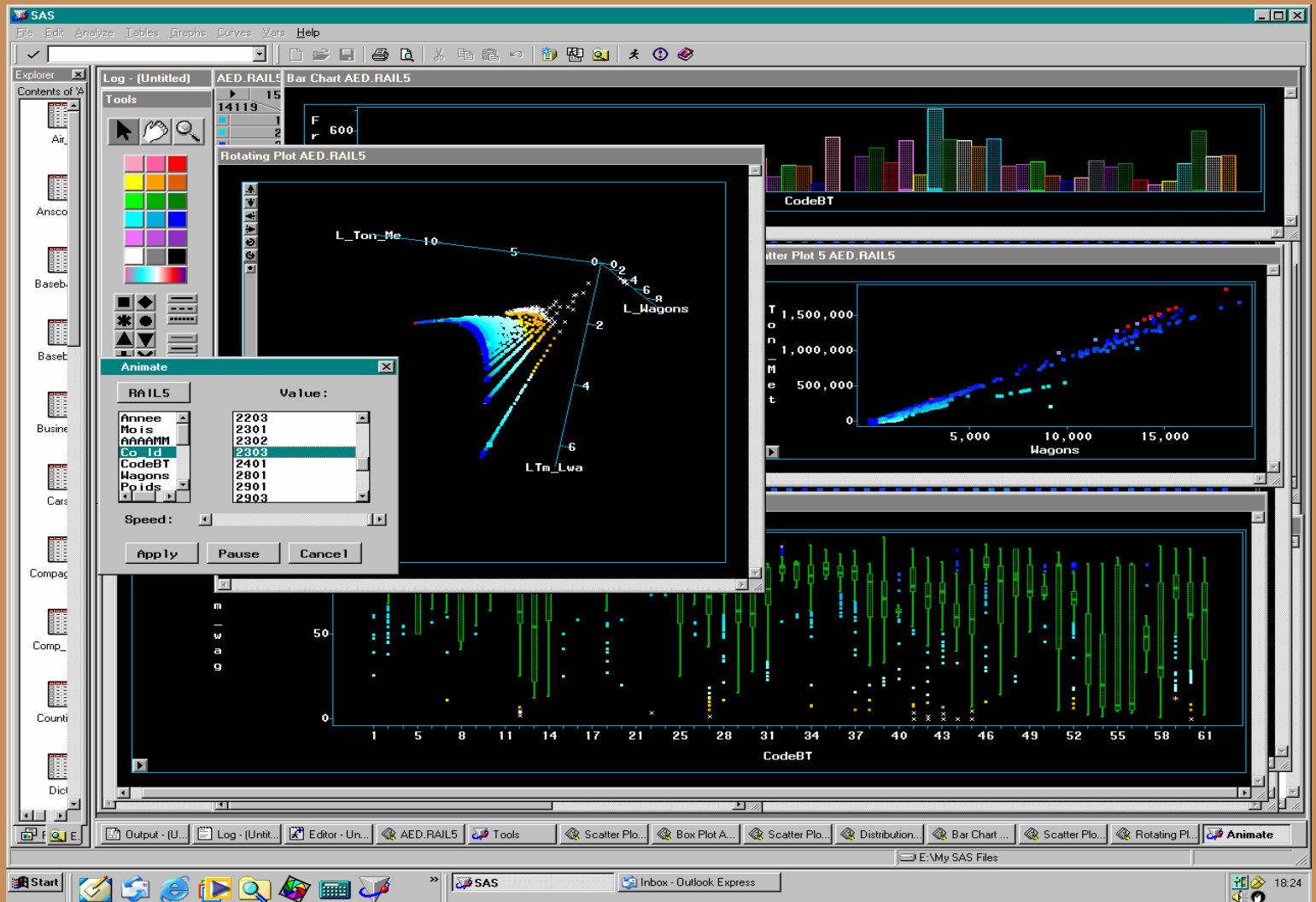


# A conceptual framework

- Dynamic graphs VS Static graphs
- Multidimensional Data Exploration
- Juxtaposition ( simultaneously displayed windows )
- Brushing (selected obs. show in all visualization)
- Data Transformation ( create new variables on the fly)
- Highlighting (use special symbols for selected groups )
- Marking ( tag specific obs. )
- Coloring ( its use provides additional analytical dimensions )
- Extracting ( subset data )
- Animation provides yet another pseudo-dimension

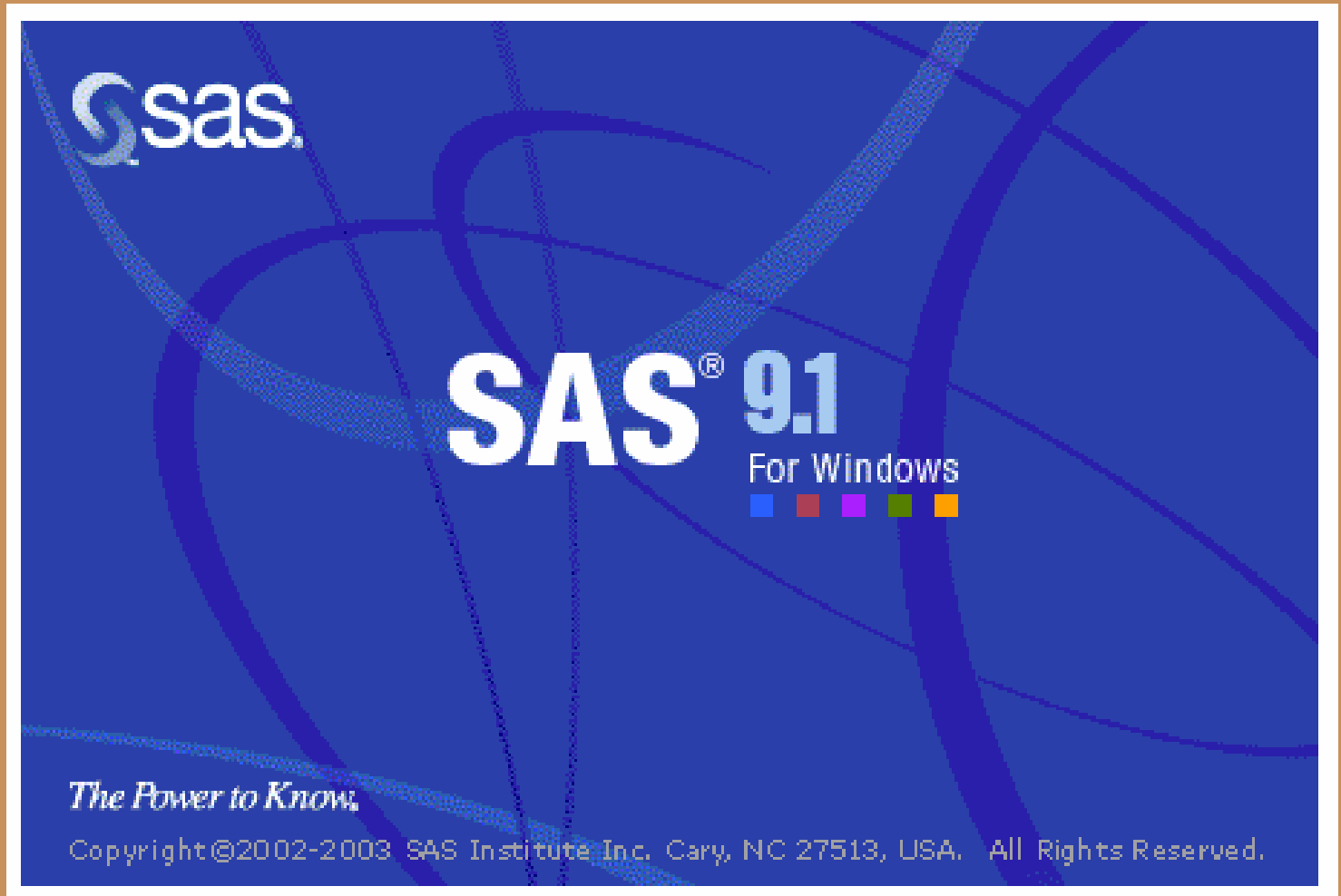


# Providing interactive multi dimensional views





# A live demo ...



# Features of Insight

- Enhanced data grid
  - indicates # of rows and columns
  - allows definition of data and modeling types
    - character, numeric, continuous, categorical
    - default role: group, label, frequency or weight.
  - manipulation of rows and columns
    - moving, sorting, searching and editing.
  - examination of data by single or group

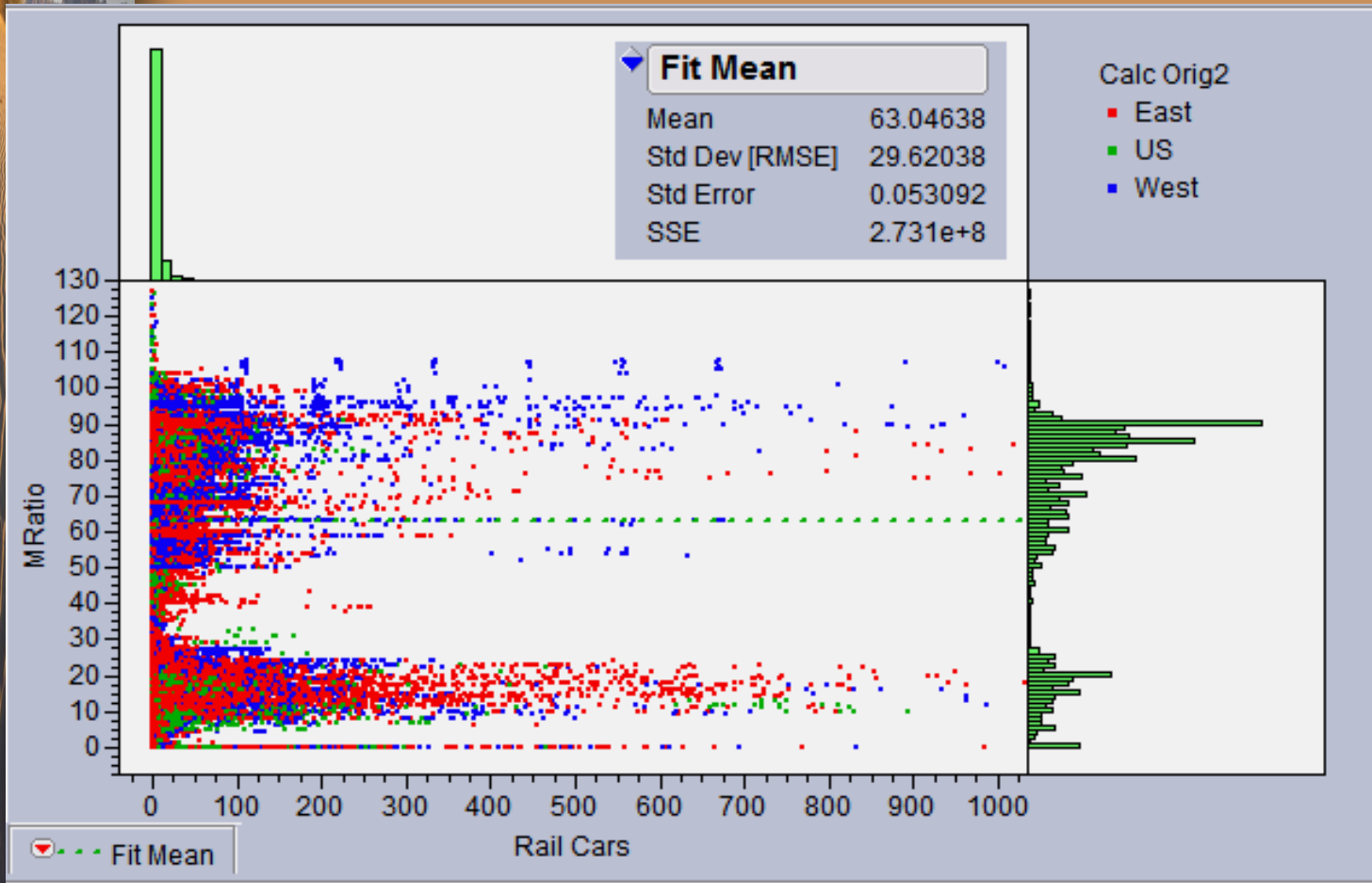


# Some features of Insight

- Easy and powerful visualisations
- Dynamic exploration of data
- Access to summary statistics
- Creation of derived variables
- Animation



# More this afternoon...




Discovery.™ From SAS.

**JMP 6.0.3**

on: STATISTICS CANADA

[www.jmp.com](http://www.jmp.com)

Copyright © 2006 SAS Institute Inc. All Rights Reserved.  
 Do not copy or distribute except under the license terms.



# Thanks You !



# Questions / Comments



Statistics      Statistique  
Canada        Canada

## **François Mainville**

Analyst / Analyste  
Transportation Division  
Division des transports

francois.mainville@statcan.ca  
(613) 951-3504

Canada 

