



# From crop fields to baker's shop, data science at every step

Pr Yves Brostaux, Uliège

Cereals & Grains 18, London

AnalysisOfVariance

AlphaRisk Fdistribution

Correlation

LatinSquare

MetaAnalysis

Variance

ExperimentalDesign

Pvalue

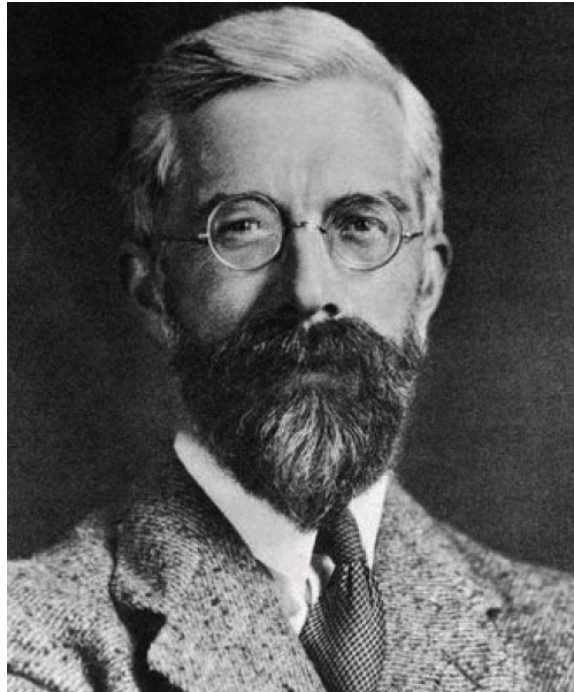
Anova

NullHypothesis DiscriminantAnalysis

MaximumLikelihood



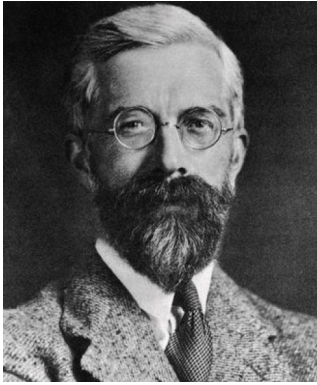
LIÈGE université  
Gembloux  
Agro-Bio Tech



**Ronald A. Fisher**  
(1890-1962)



# The beginning



Ronald A. Fisher  
(1890-1962)



## Rothamsted Experimental Station (1919-1933)



ROTHAMSTED  
RESEARCH



LIÈGE université  
Gembloux  
Agro-Bio Tech

# In the fields



Lattice designs (1936) Split-plot designs

Partially Balanced Incomplete Block (PBIB) designs

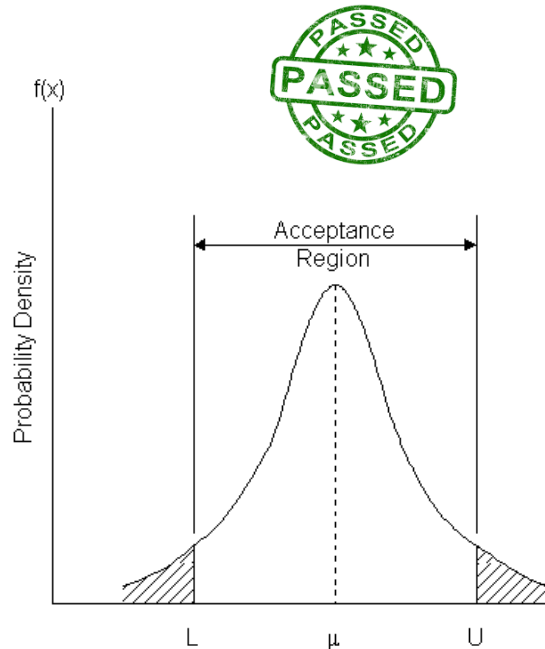
Split-block designs Alpha designs (1978)

Resolvable Incomplete Block designs

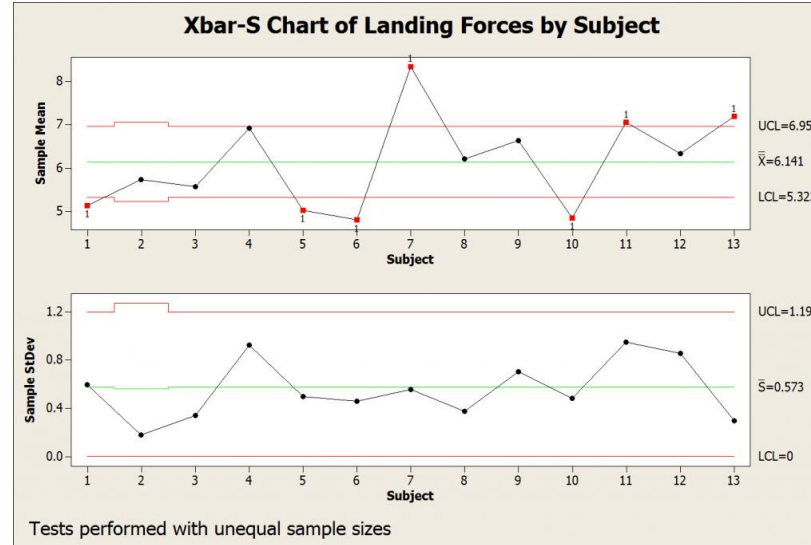


# Meanwhile, in the industry...

## Process Capability Index Control Charts



Shewhart, W. A. (1931)

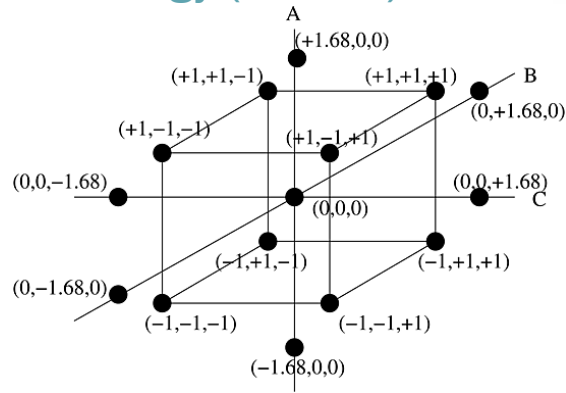


## Quality Control upon Reception Statistical Quality Control Six Sigma

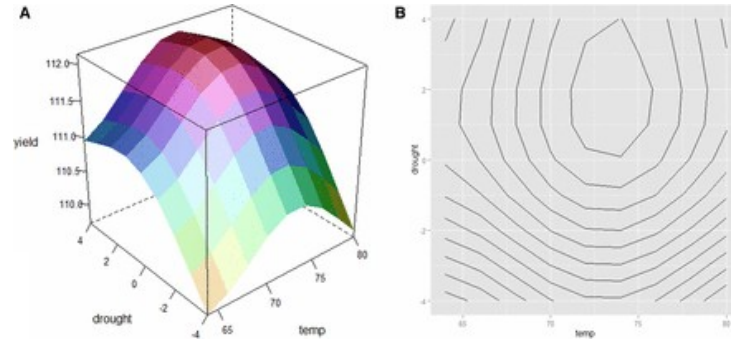
# Meanwhile, in the industry...

Packett-Burman  
designs (1946)

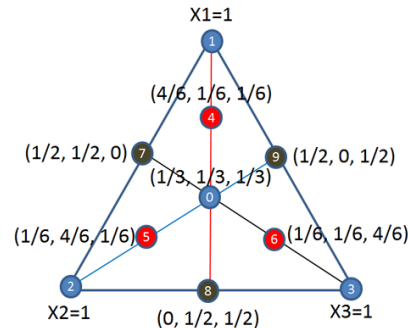
Taguchi Robust Design  
Strategy (1950's)



Mixture designs  
(Snee & Marquardt, 1976)



Response Surface Methods  
(Box & Wilson, 1951)





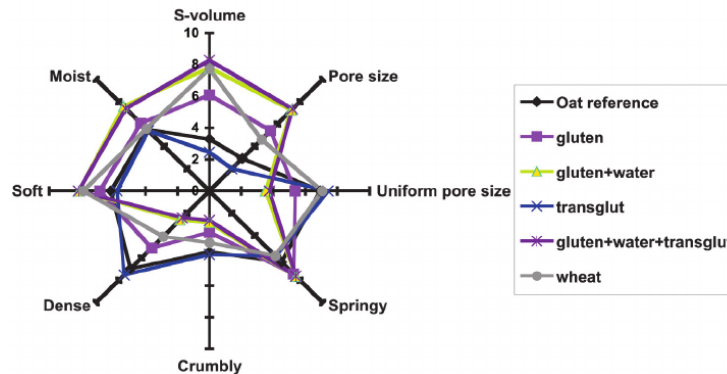
# ... and on the consumer side



Profiling tests  
Hedonic analysis  
Ranking tests



Differentiation tests  
Napping Triangle tests



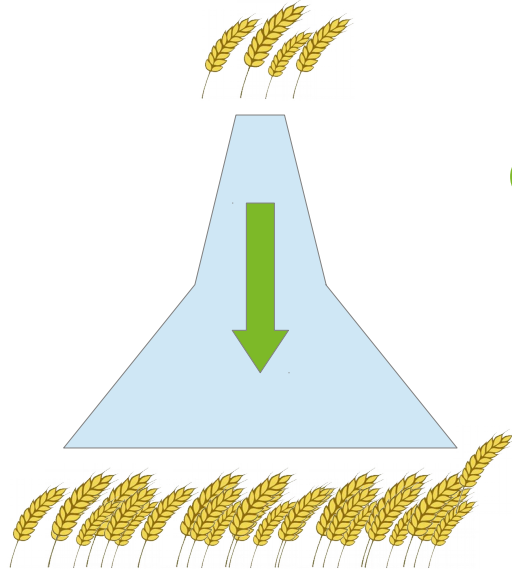
## Sensory analysis

Cairncross, S. E., &  
Sjostrom, L. B. (1950)



# The classical inference paradigm

- Infer properties of a population, based on the observation of a sample of individuals
- Useful when data are *s c a r c e*



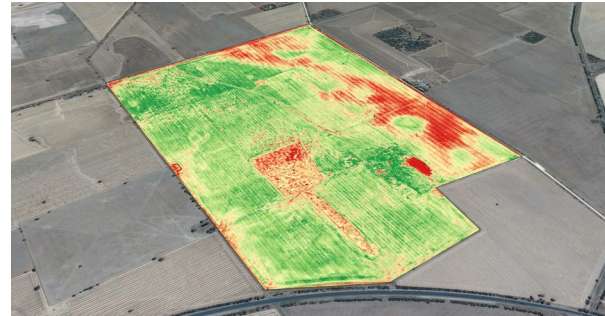
“From the few, thou shall conclude about the many”



# 2000's, the rise of the machines



GPS devices  
UAV imagery

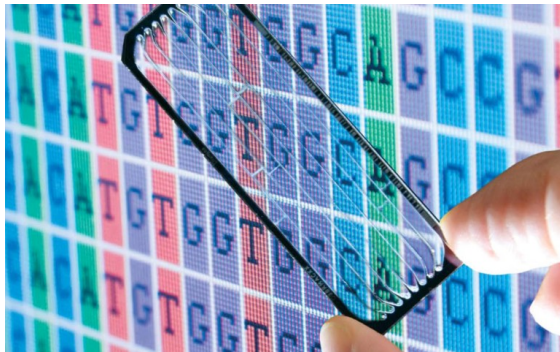


Remote sensing  
**Smart Farming**



Satellite imagery  
**Spatial data analysis**

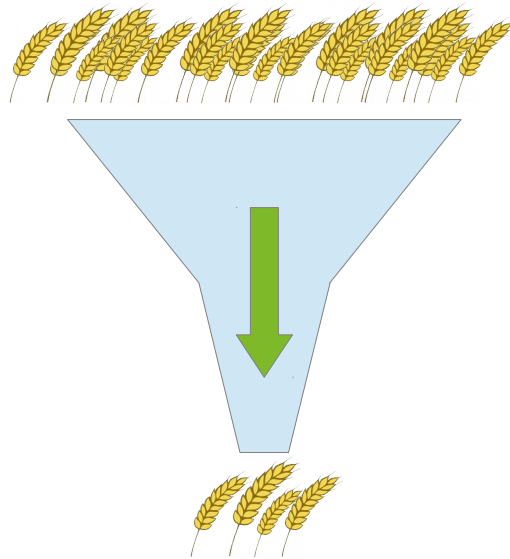




A word cloud visualization of genomics-related terms. The most prominent words are 'Genotyping', 'Genomics', 'Sequencing', 'Microarrays', 'Bioinformatics', 'NGS', 'Methylation', 'Variant', 'Gene', 'DNA-seq', 'RNA-seq', 'miRNA', 'Exome', 'NanoString', 'nCounter', 'SNP', 'ChIP-seq', 'Epigenetics', 'Transcriptomics', 'RNA', 'Covariates', 'fluorescence-luminescence', 'arrays', 'Tiling', 'Real-time', 'discrimination', 'reader', 'DNA', 'TapeStation', 'Bioanalyzer', 'FFPE', 'Green', 'cBot', 'ChIP-string', 'TagMan', 'Q-PCR', 'Affymetrix', 'Syr', 'BeadStation', 'TaqMan', 'PCR', 'illumina', 'Caliper', 'allelic', 'CNV', 'expression analysis', 'Qubit', 'samples', 'profiling', 'chip-seq'.

# Here comes the Big Data

- Paradigm inversion on classical inference
- Refine useful information from huge amount of data



“From the many, thou shall find the chosen ones”







# What's next ?

*From fields to  
baker's shop,  
data science  
at every step*