



( My opening Keynote for  
the 2017 U.S. Conference on Teaching Statistics )

Chris Wild  
University of Auckland

*Introducing co-Presenter  
Marcel Marceau*

This talk carries a **Government  
Mental Health Warning**



Chris Wild  
University of Auckland

Crazy Diamond Productions  
Pastiche Productions  
Rank Amateur Films

Powered by 200 Mb of PowerPoint  
& graphics from 

Darkness and  
ignorance  
reign ...

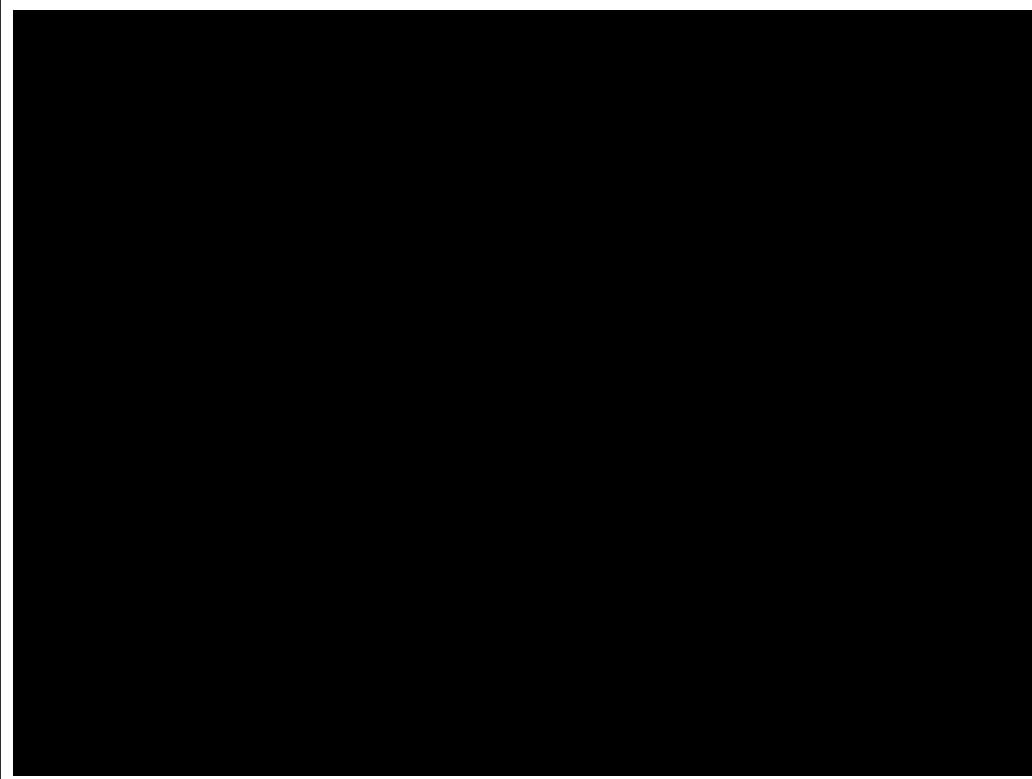
RecoShan

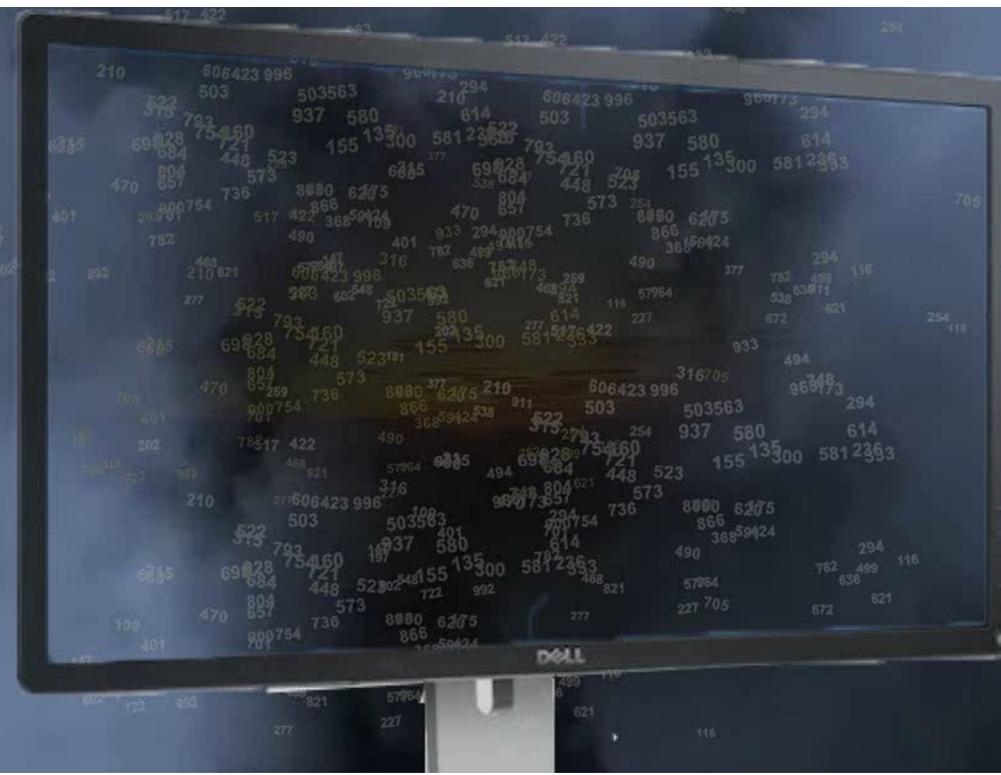


RecoShan



RecoShan





*“Oh Say, Can you see ...?”*

An Edu-tainment

Chris Wild

University of Auckland



*“Oh Say, Can you see ...?”*

Speaking into these elements: Early ...

- multivariate thinking
- visualisation ...
- use of technology
- “Big” data

How Gilding Familiar Lilies ...

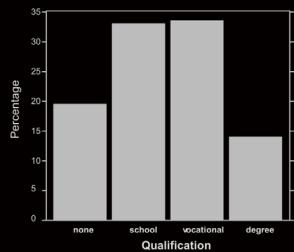


can illuminate our World

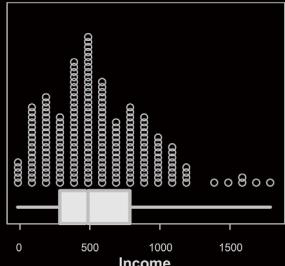




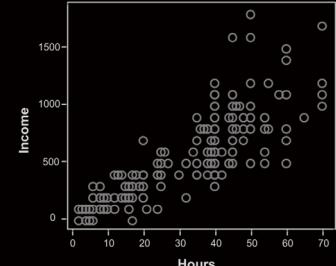
What are these familiar lilies ? ...



Bar chart



Dot plot

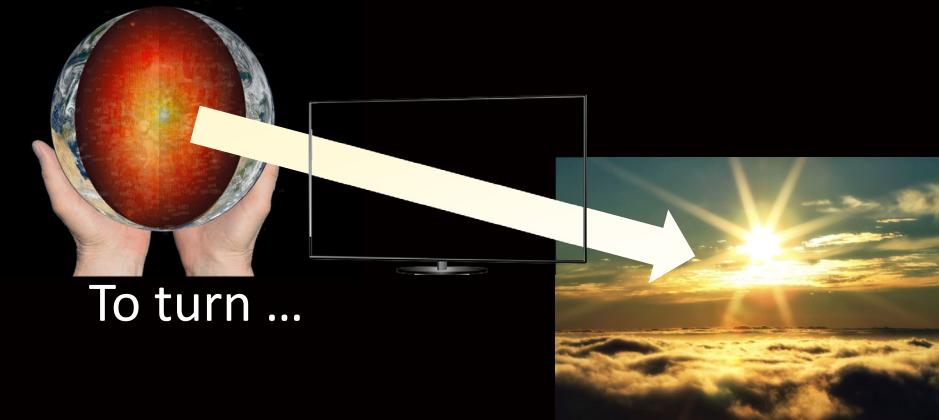


Scatter plot



What are these familiar lilies ? ...

But how can we *gild them* to gain deeper, *multivariate insights*?



But here's an even better image ...

Adding *Zing* to home-cooked meals



with *herbs and spices*



Well Chris, These are some of my favourites ...

So Rachael, What do we *need* for a well-stocked spice rack?

# *Stocking up the spice rack*

- Colour my world
- Size matters
- Subsetting / Faceting
- Stepping through lists
- Motion step on moving
- Identification & brushing
- Transparency
- Bait & Switch



Cooking with spices ...

Less ...



This talk ...

More ...



Why ???

*Holi: Indian festival of Colour*

Chris Wild, 2017 Statistics Teachers' Day

THE UNIVERSITY OF AUCKLAND  
DEPARTMENT OF STATISTICS

What is *the single best thing*  
*we can do for our students?*

*Populate*

*their imaginations*

*with possibilities*

# Along the way, we'll find ...



*You can't always get what you want*

# Along the way, we will find ...



*"Careful with that axe ..."*

*And we have to be a bit careful ...*

## But ours is a statistics ...

*driven by passion!*



Lisa Dierker

PASSION DRIVEN STATISTICS

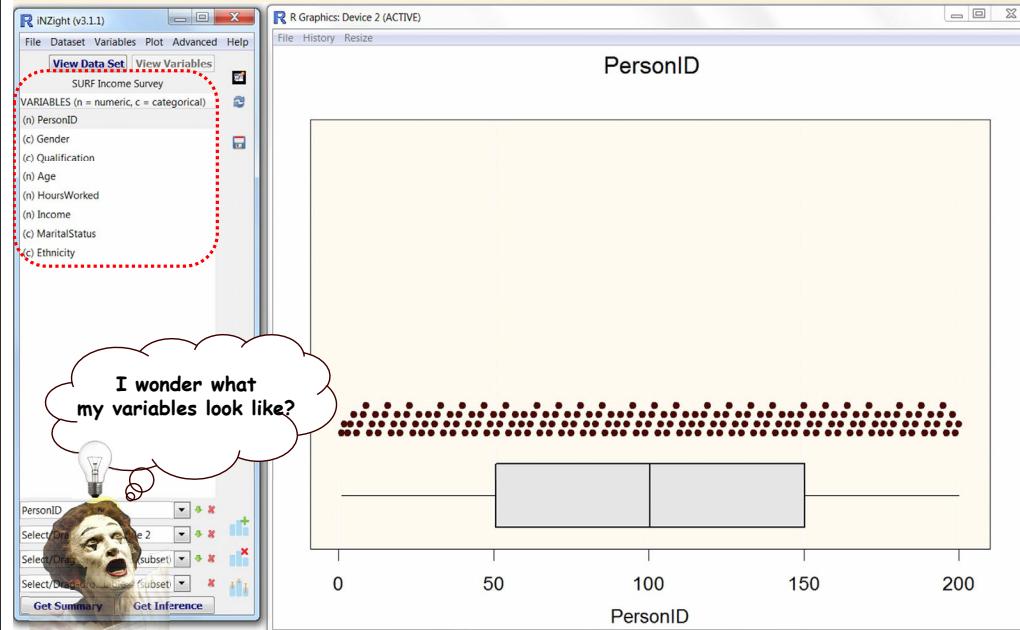
## So let's get cooking !!





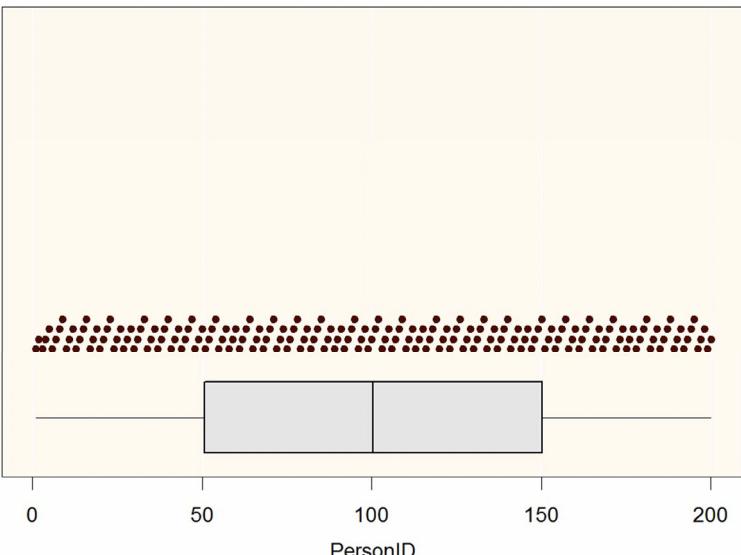
## ... Stepping through Variables

# Stepping through Variables

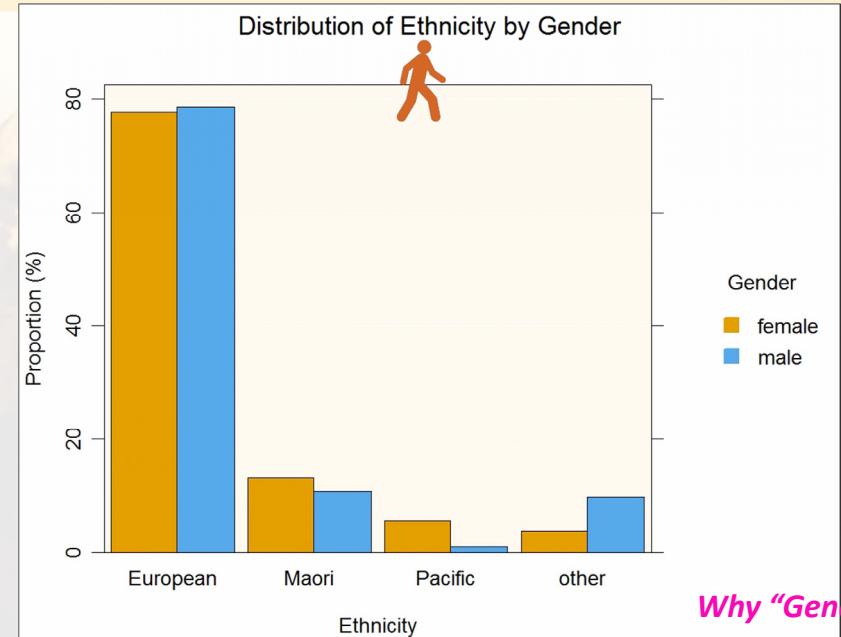


## Stepping through Variables

PersonID



## Relationships to Gender



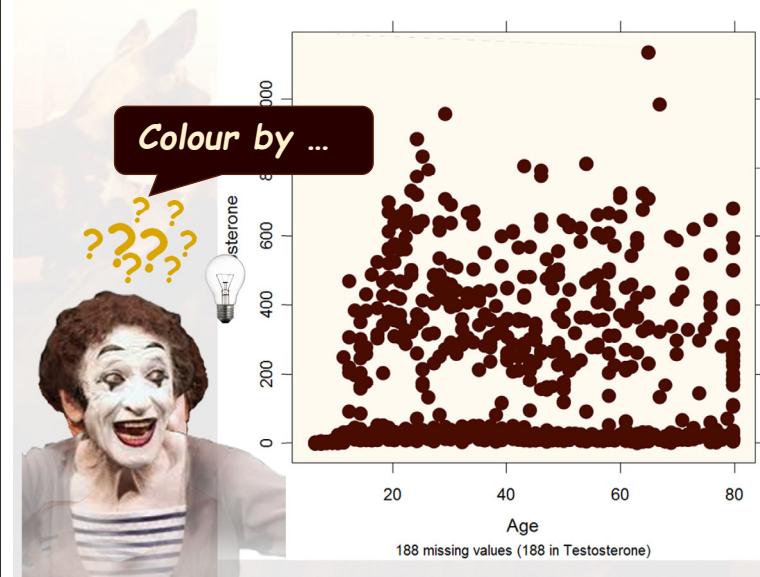
Why "Gender"?



## Tasting Colour ...

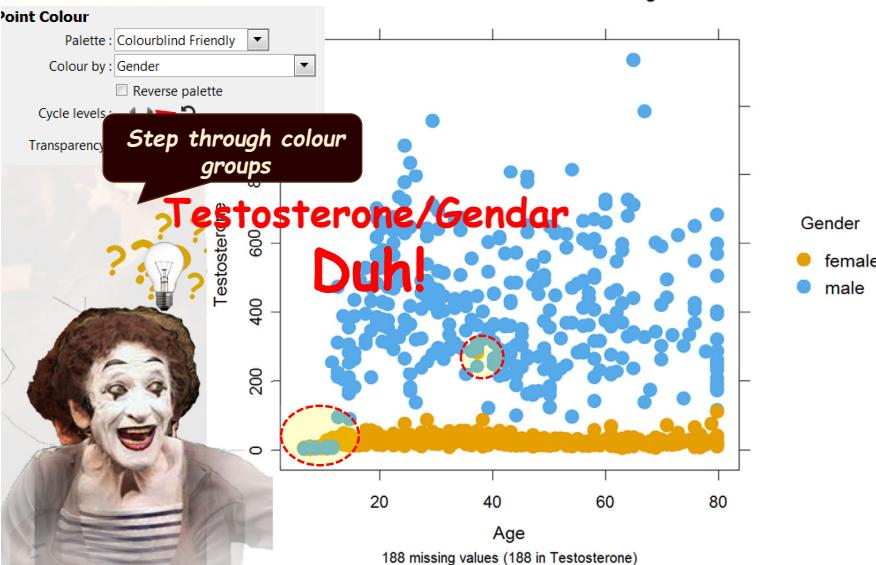
## Testosterone levels by Age

Testosterone versus Age



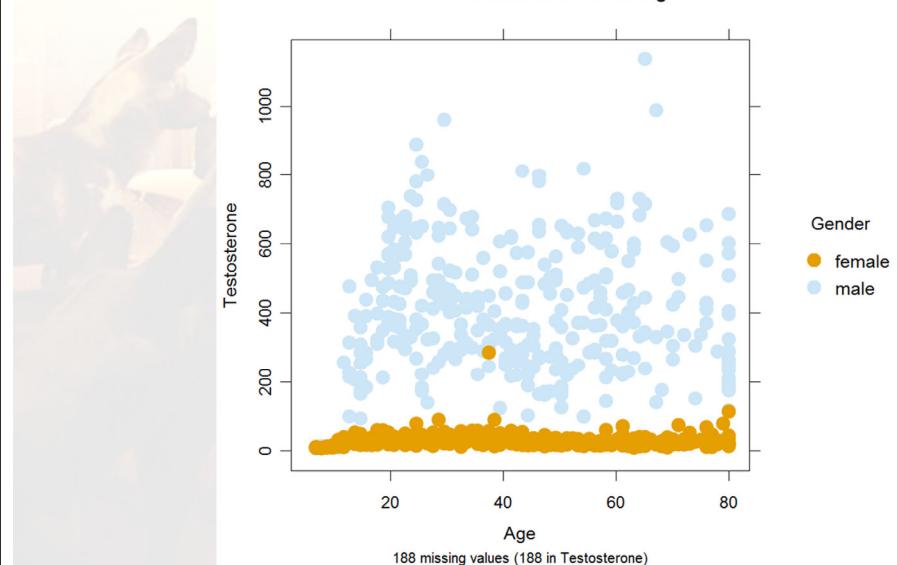
## Colour by Gender

Testosterone versus Age



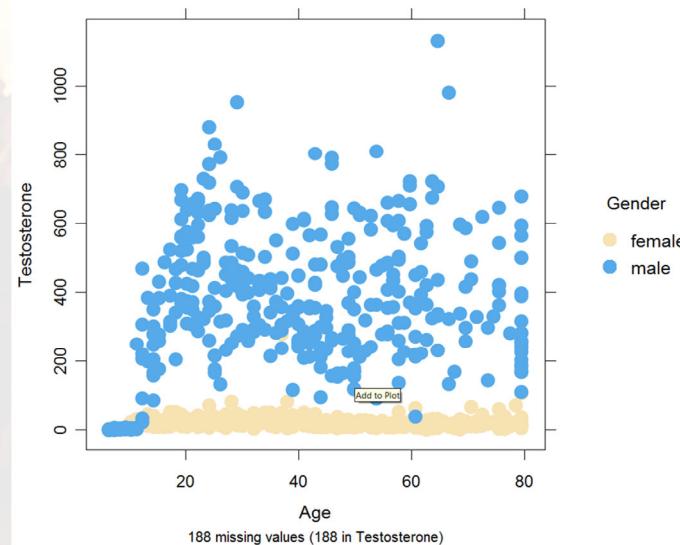
## Highlighting the females

Testosterone versus Age



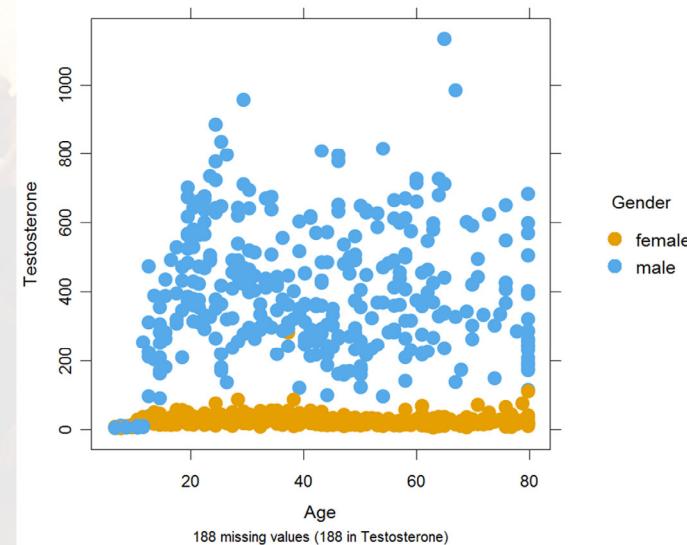
## Highlighting the males

Testosterone versus Age



## Colour by Gender

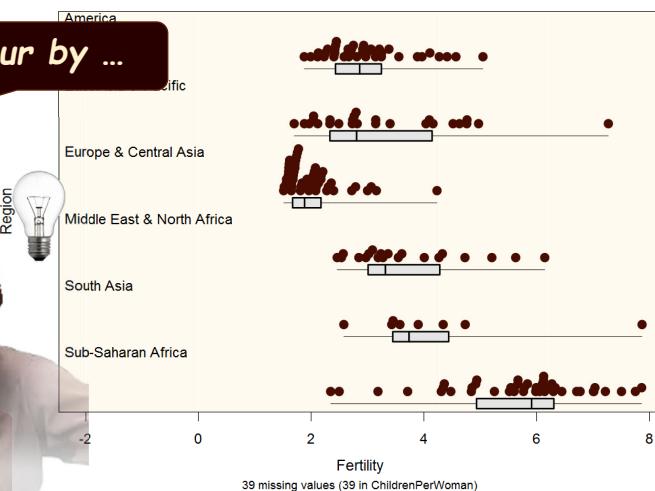
Testosterone versus Age



## Fertility by Region

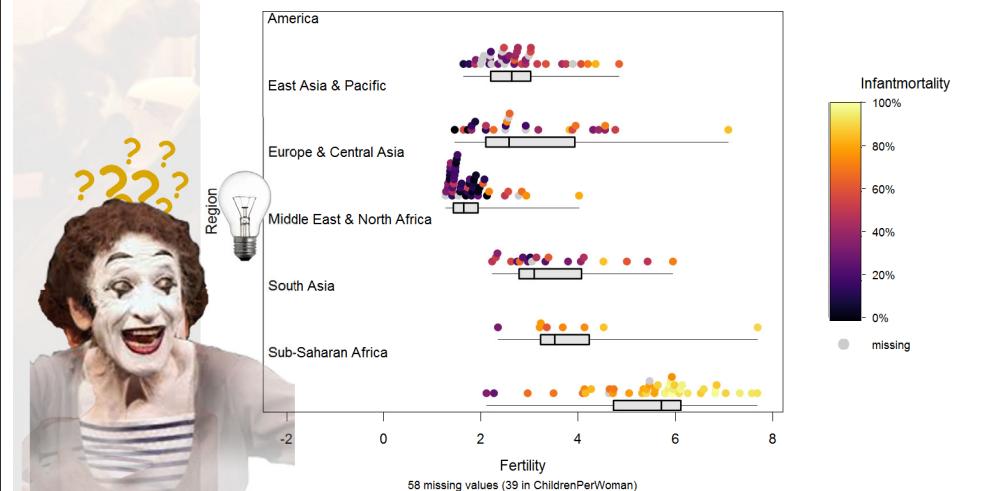
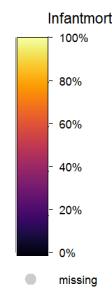
Fertility by Region

Colour by ...



## Colour by Infant Mortality

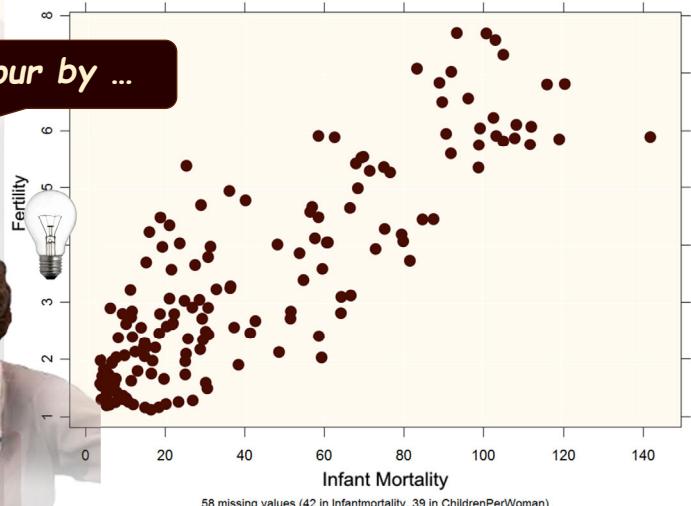
Infant mortality



# Fertility vs Infant Mortality

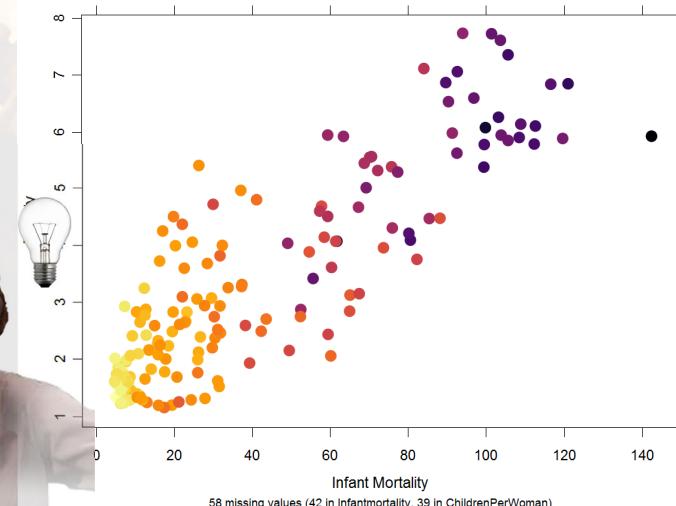
Fertility vs Infant Mortality

Colour by ...



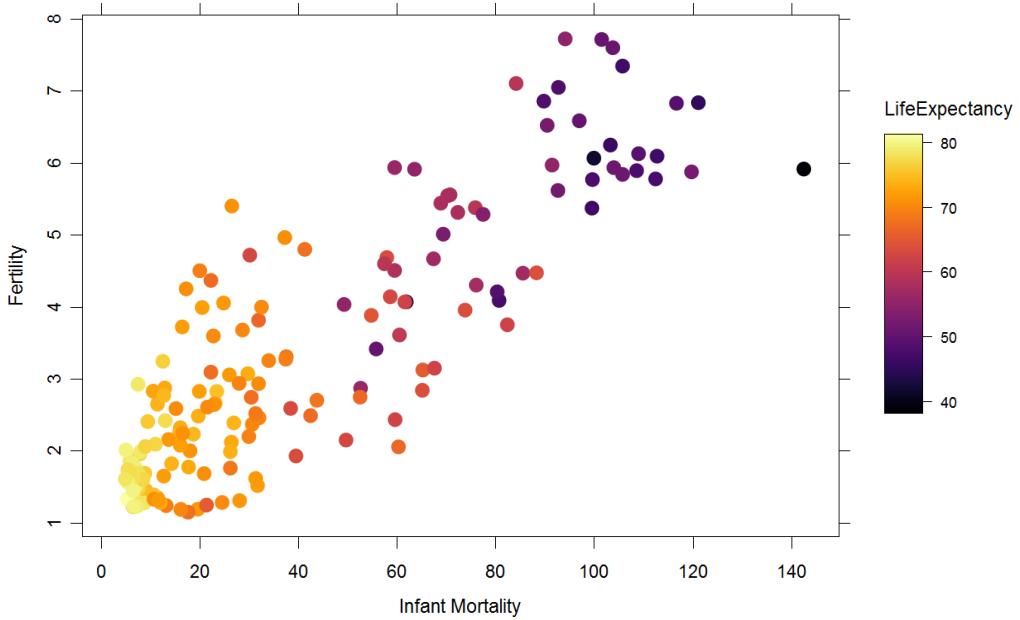
Chris Wild, 2017 Statistics Teachers' Day

# Colour by Life Expectancy



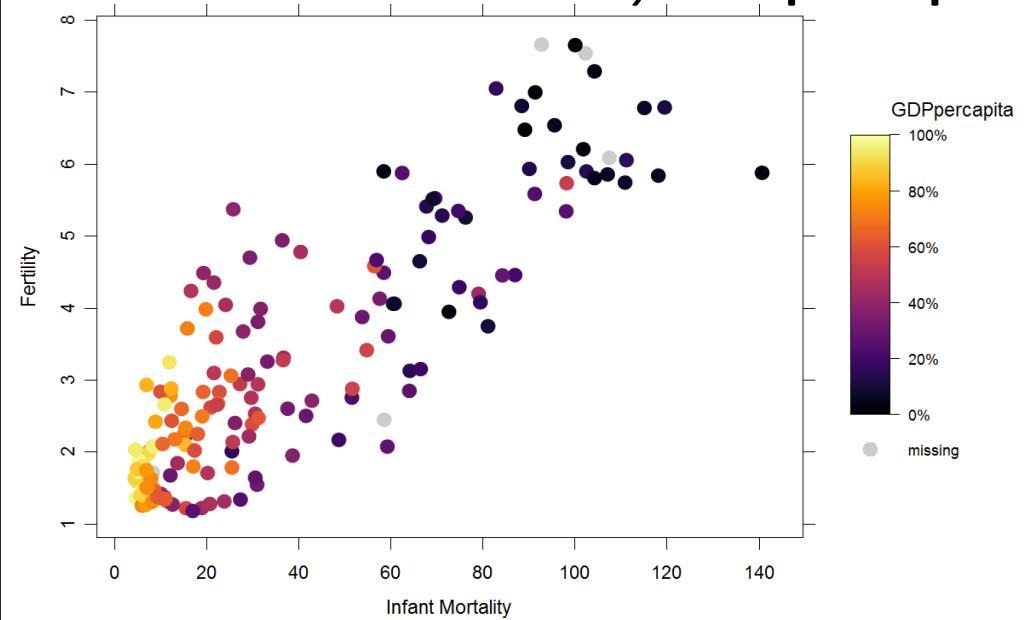
Chris Wild, 2017 Statistics Teachers' Day

# Colour by Life Expectancy



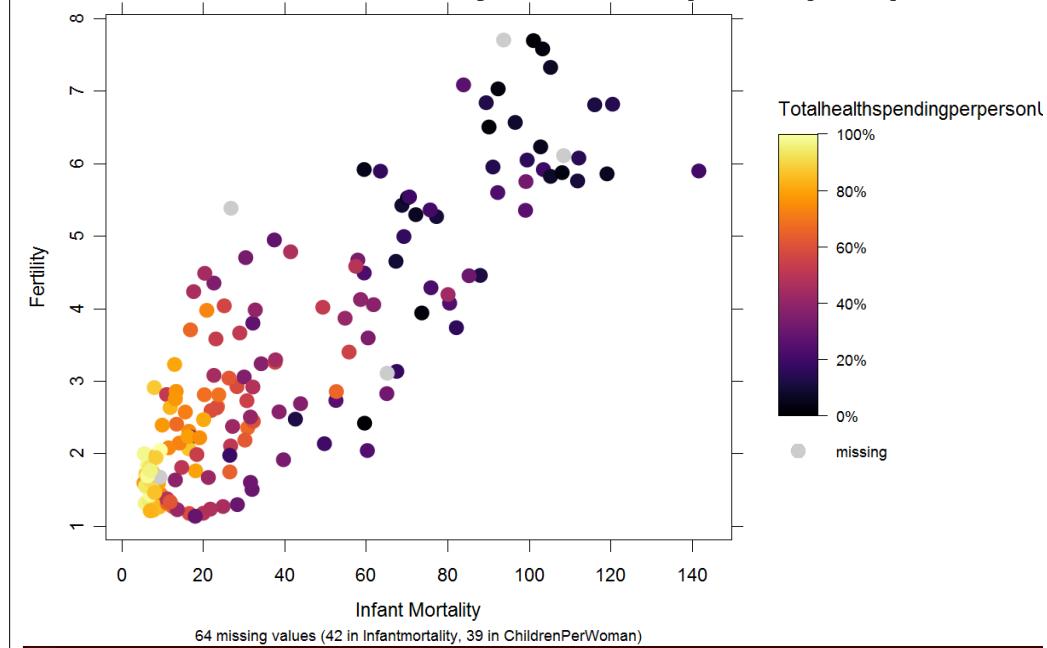
Chris Wild, 2017 Statistics Teachers' Day

# Colour by GDP per capita



Chris Wild, 2017 Statistics Teachers' Day

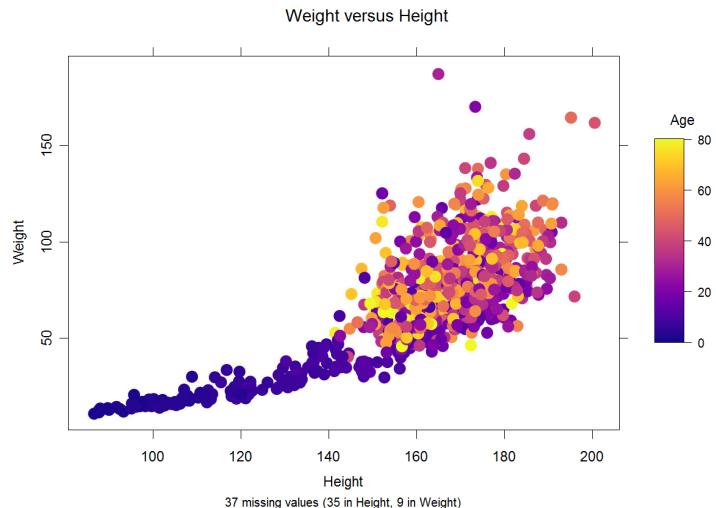
## Colour by Health spend per person



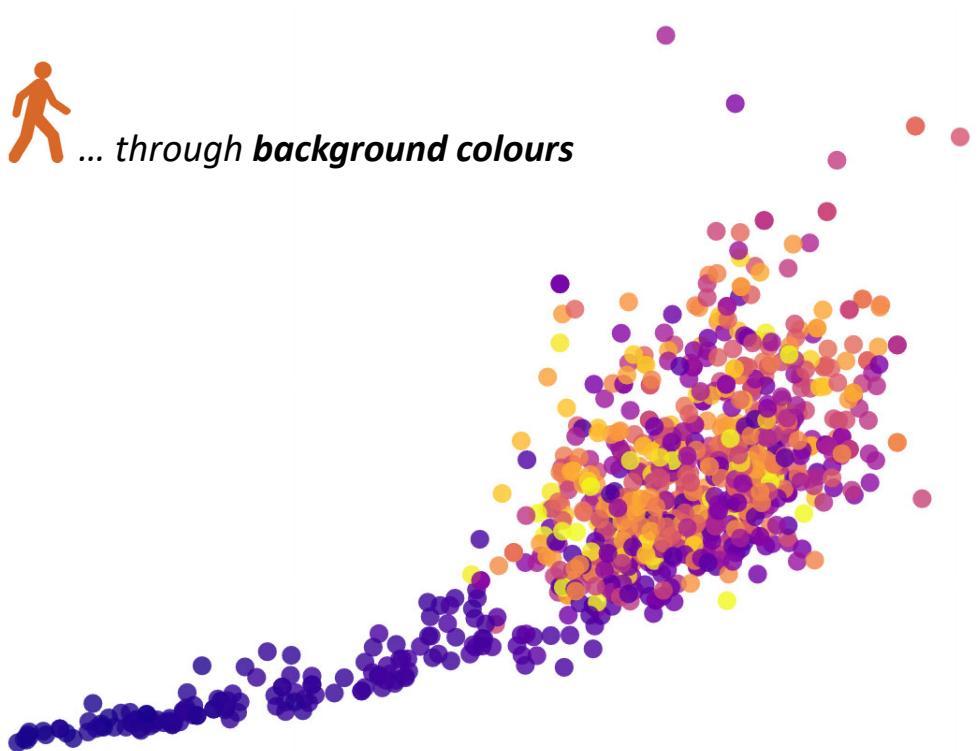
... Stepping in Colour space



... through **background colours**

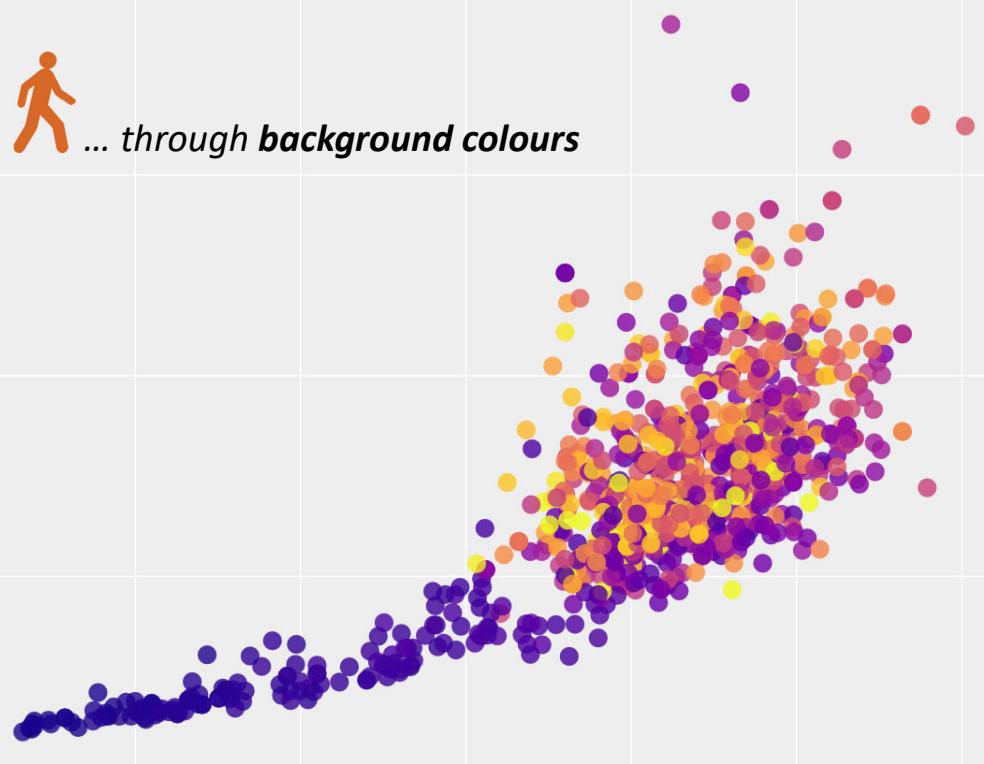


... through **background colours**

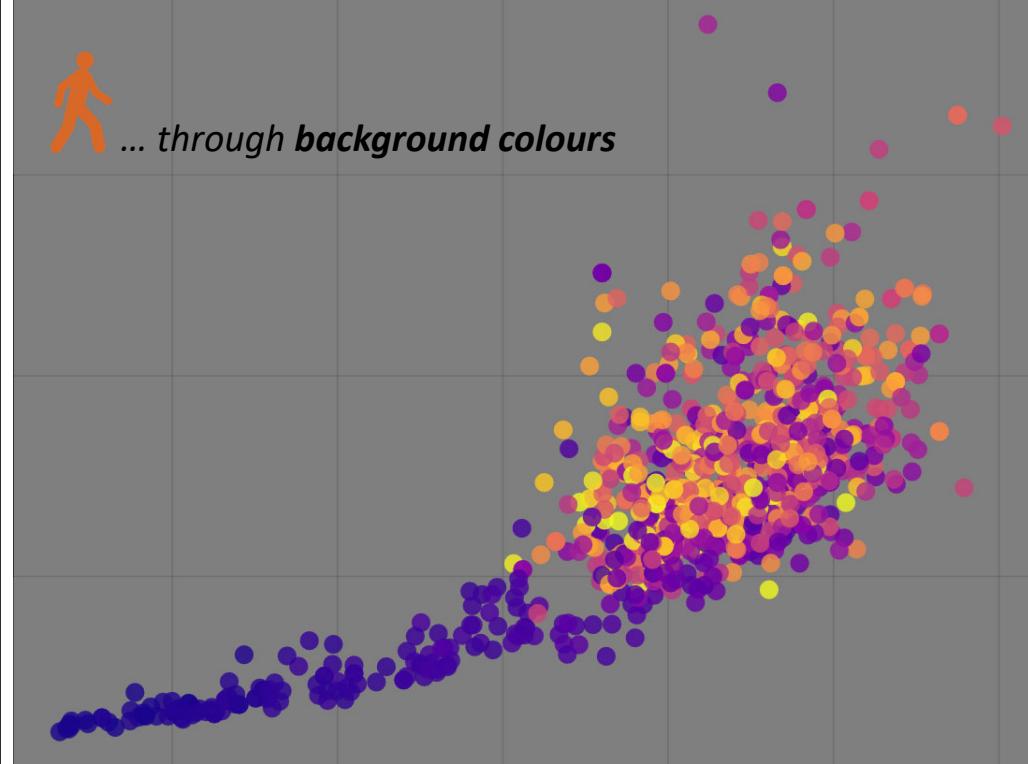




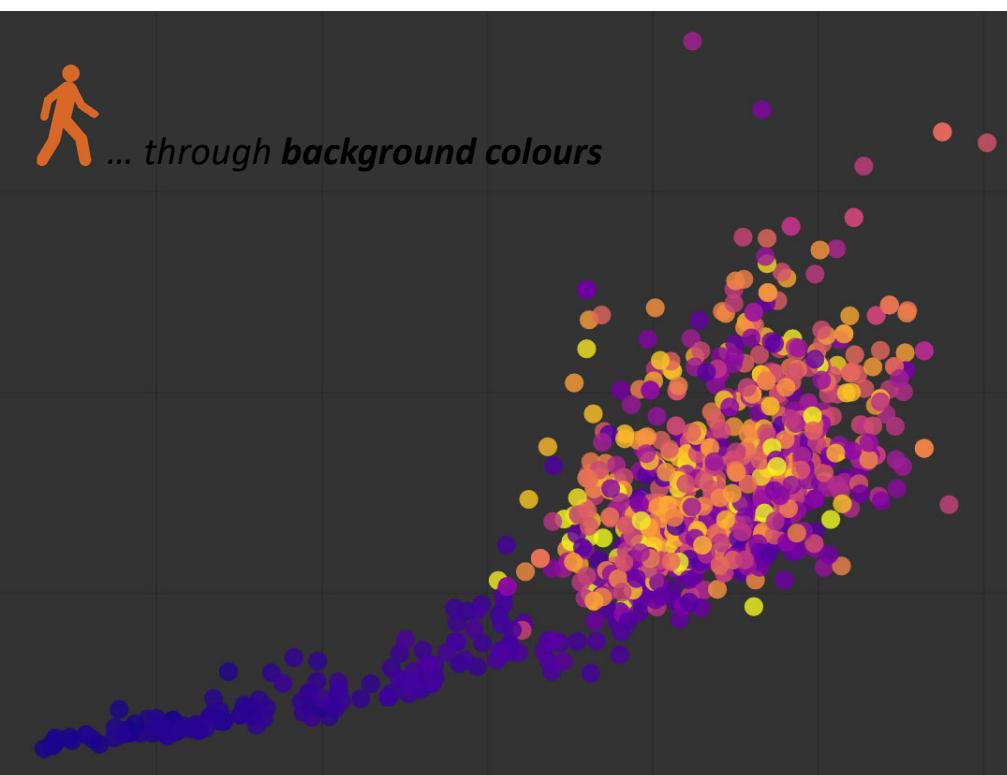
*... through **background colours***



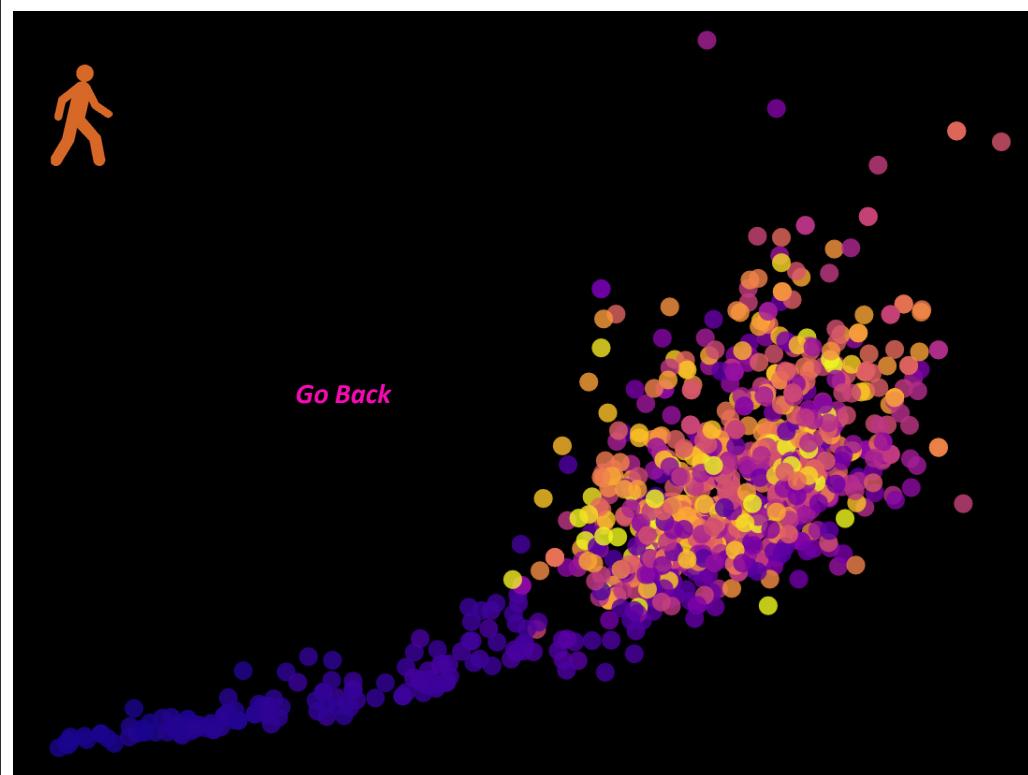
*... through **background colours***

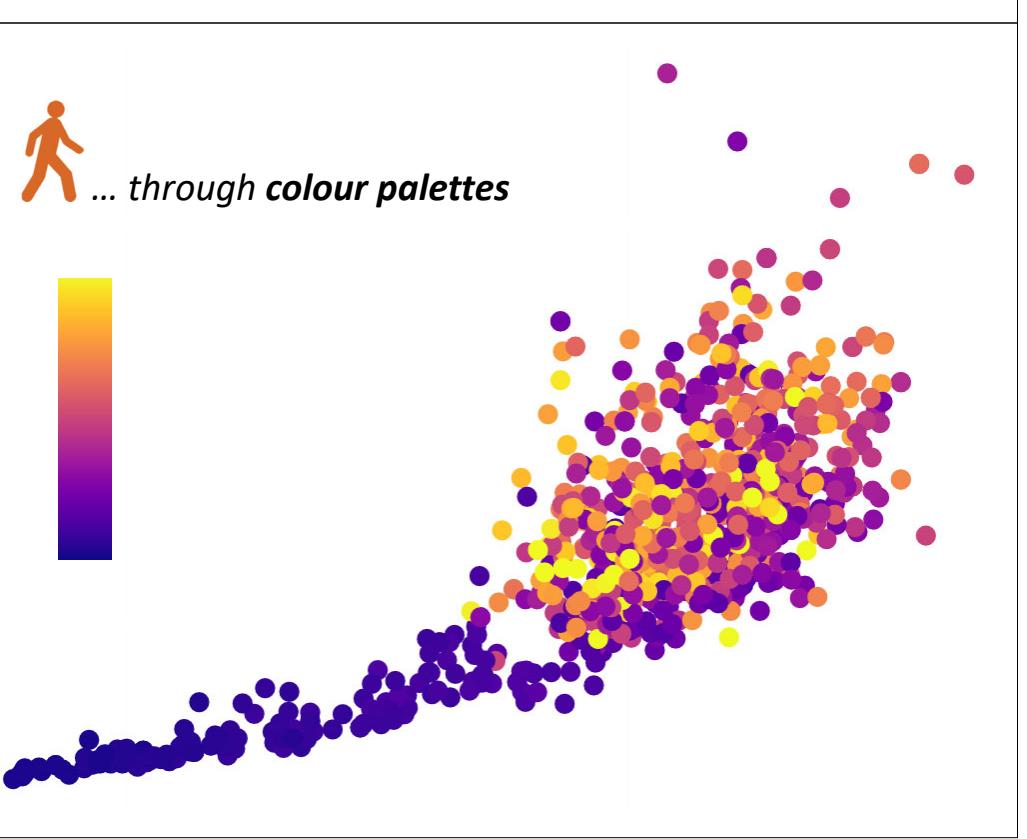
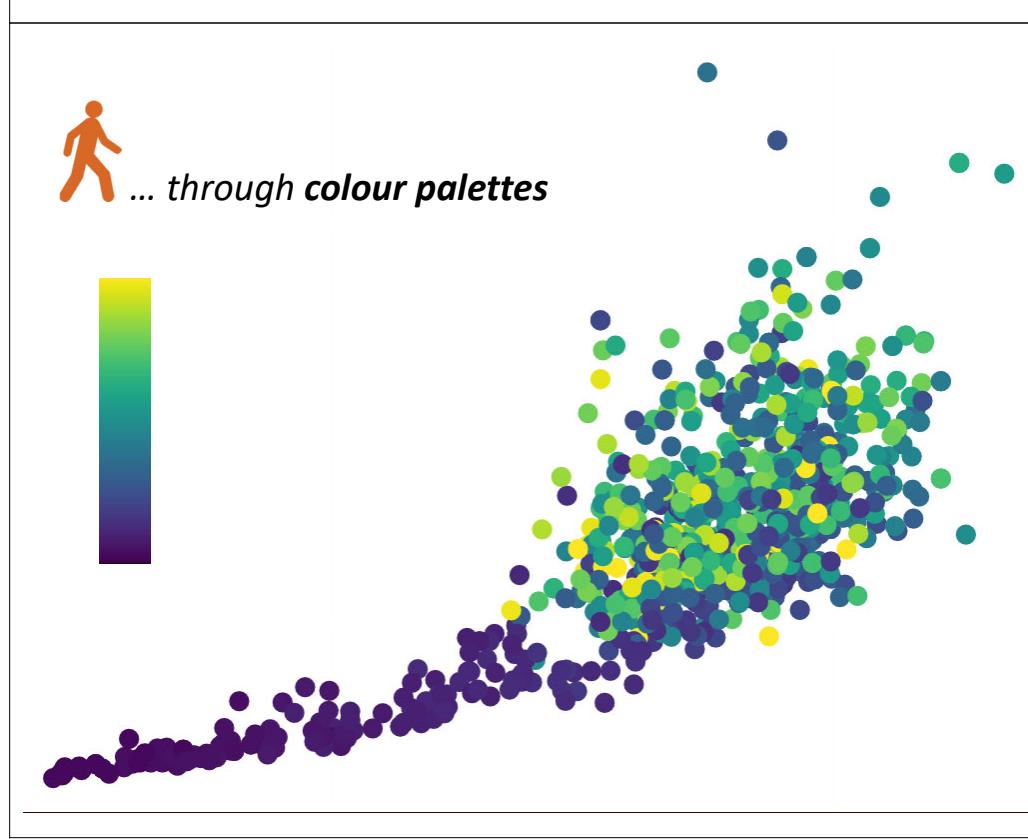
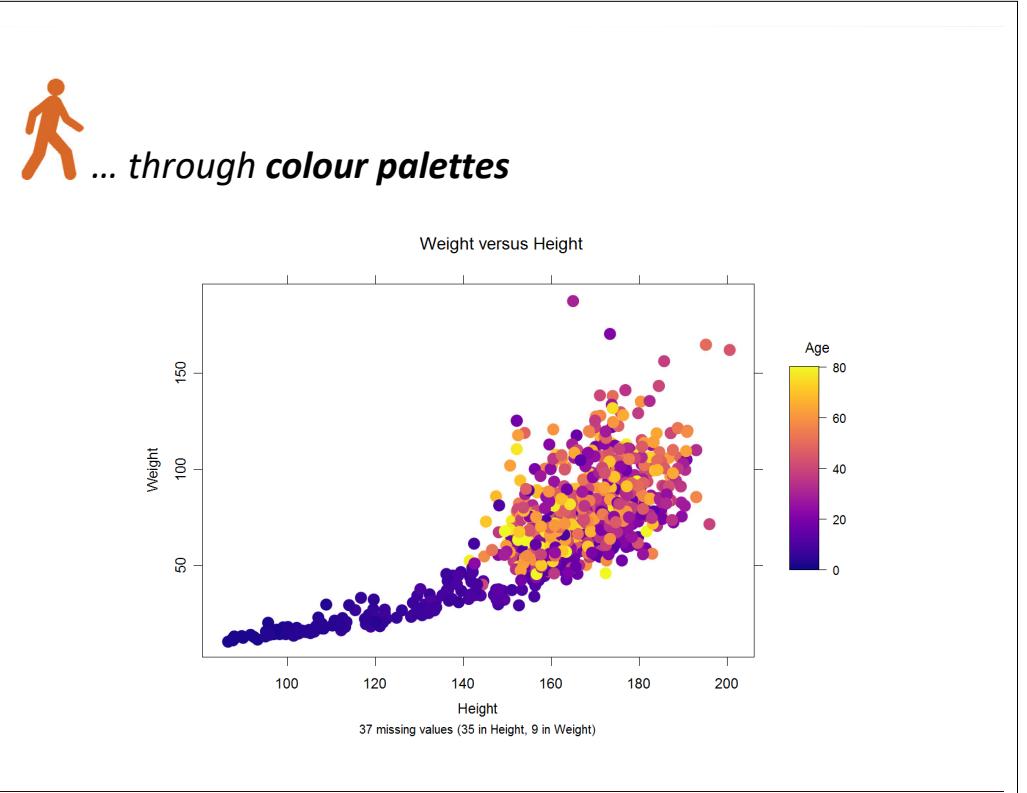


*... through **background colours***



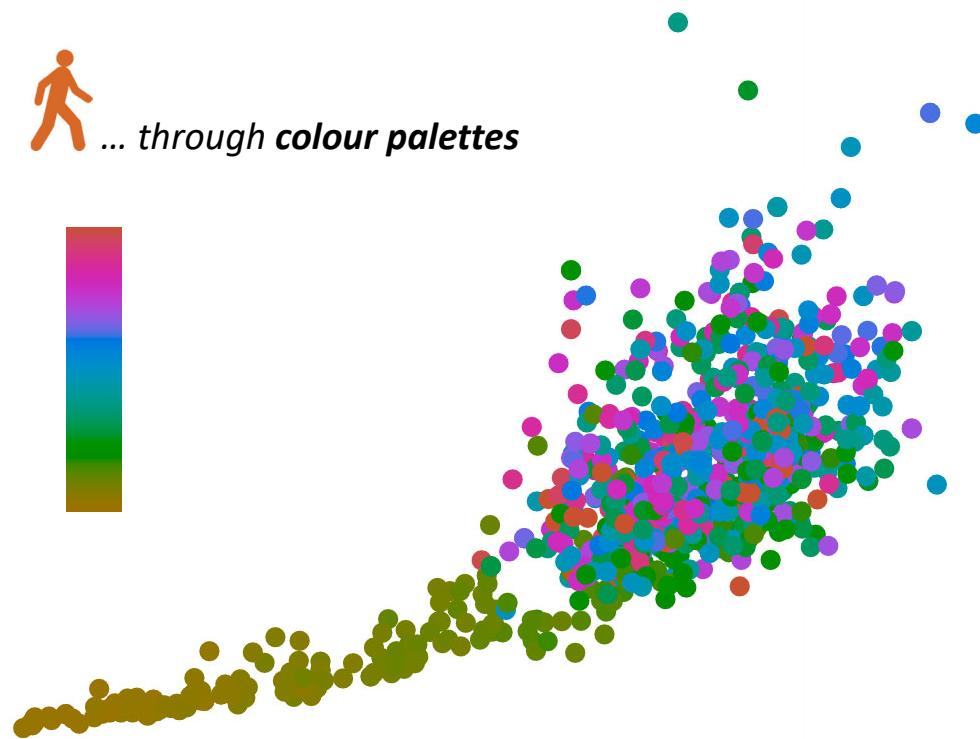
*Go Back*



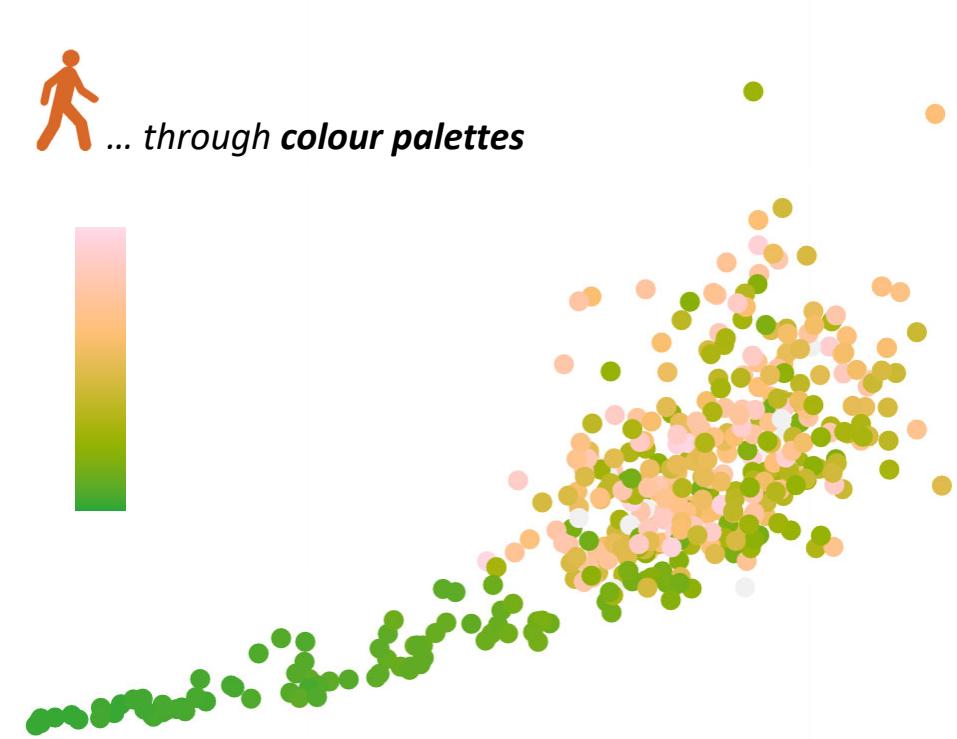




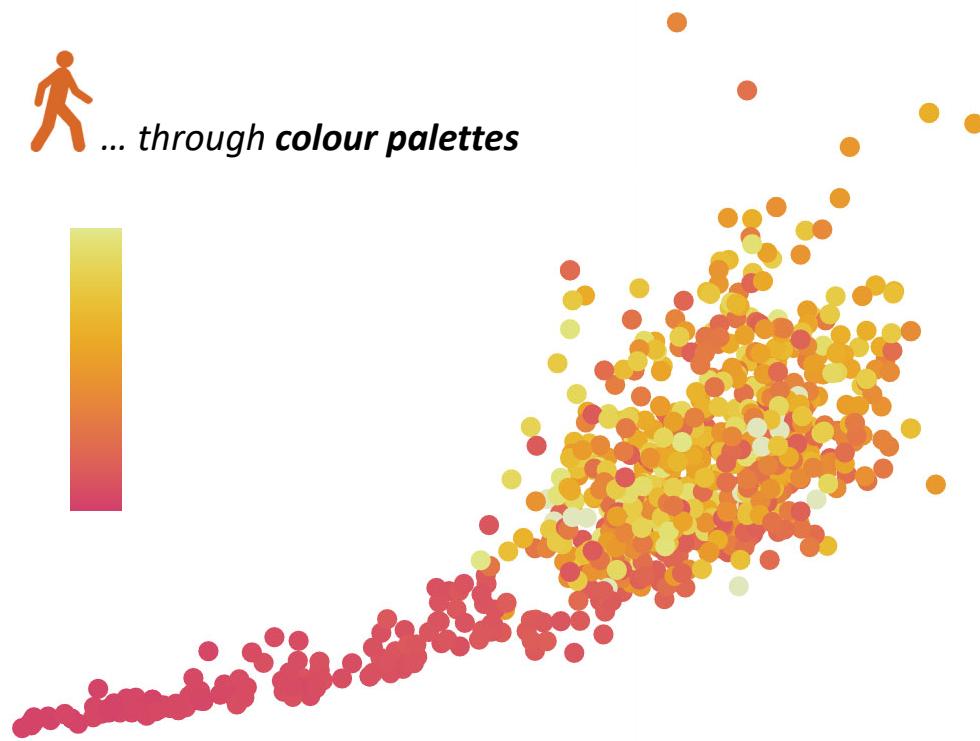
*... through colour palettes*



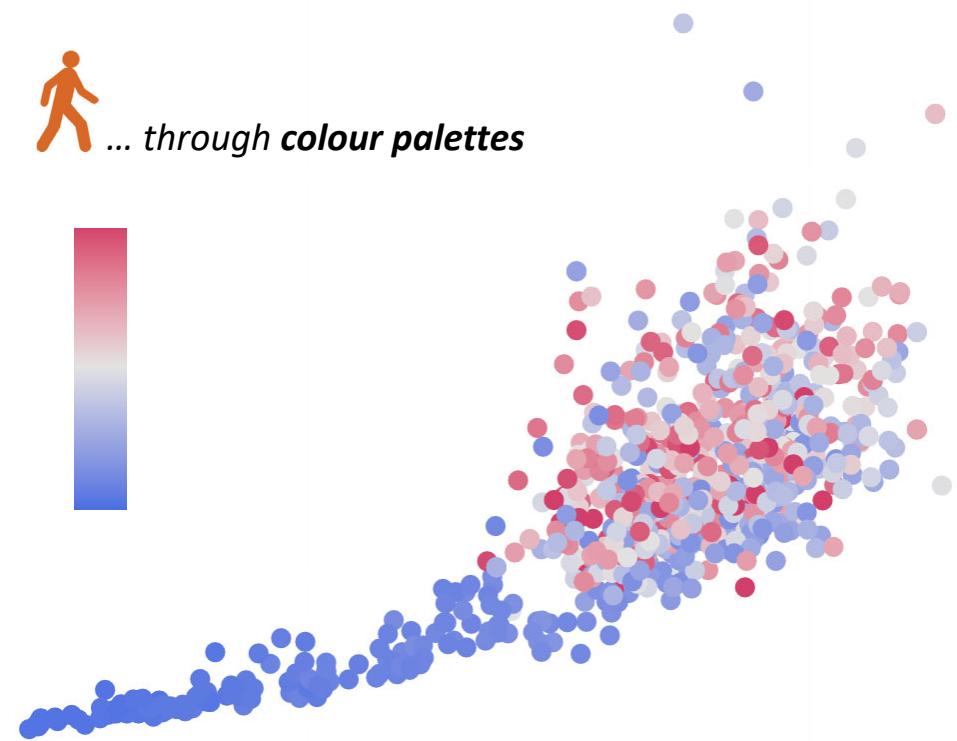
*... through colour palettes*



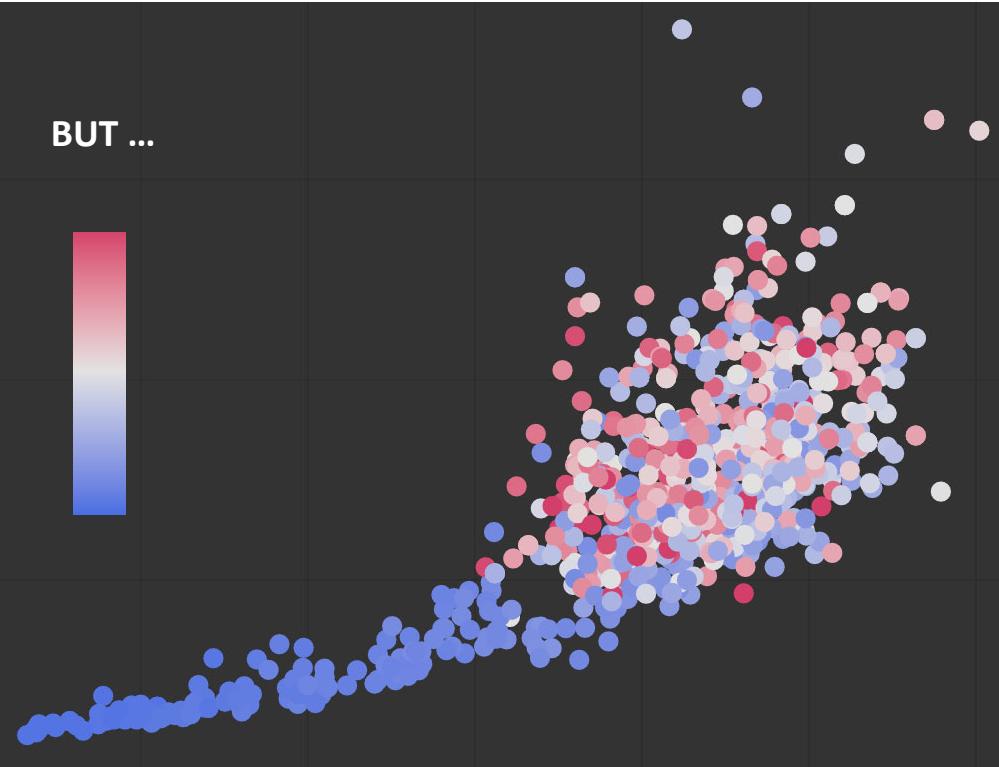
*... through colour palettes*



*... through colour palettes*

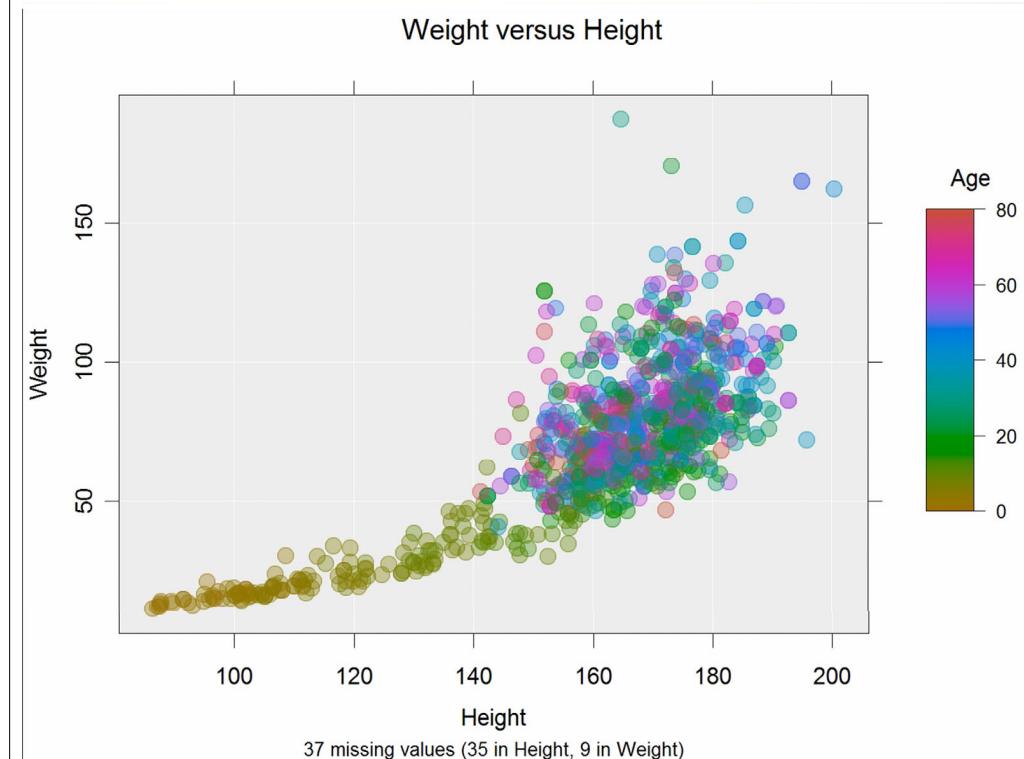


BUT ...



... through *colour groups*

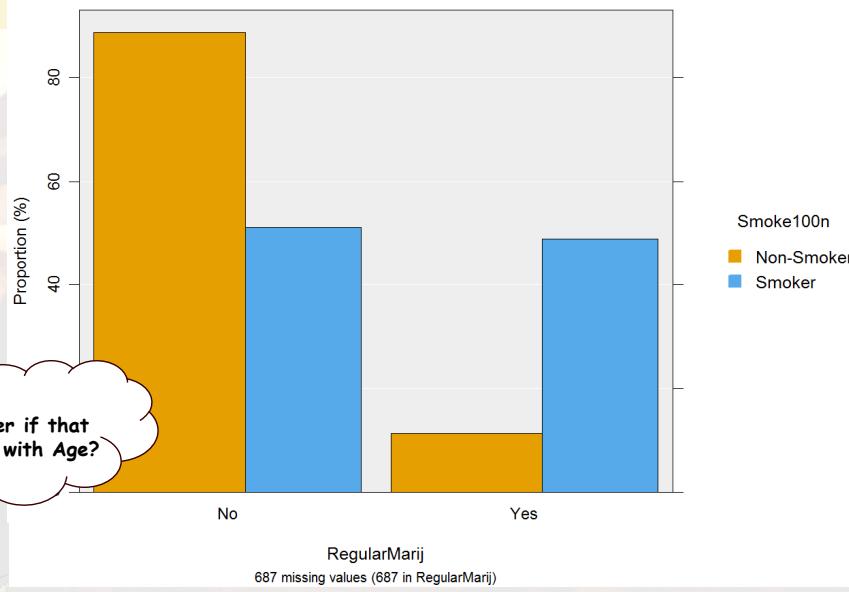
Weight versus Height



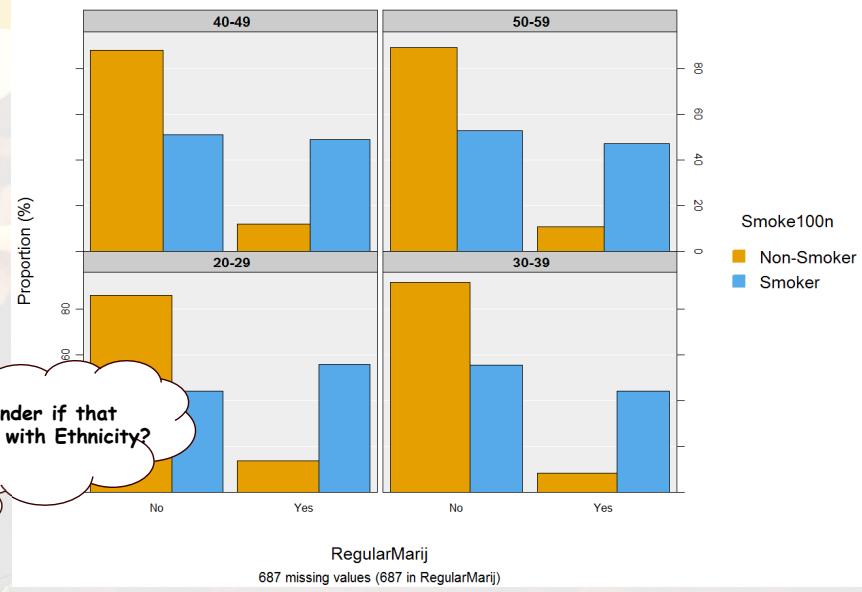
Subsetting ...



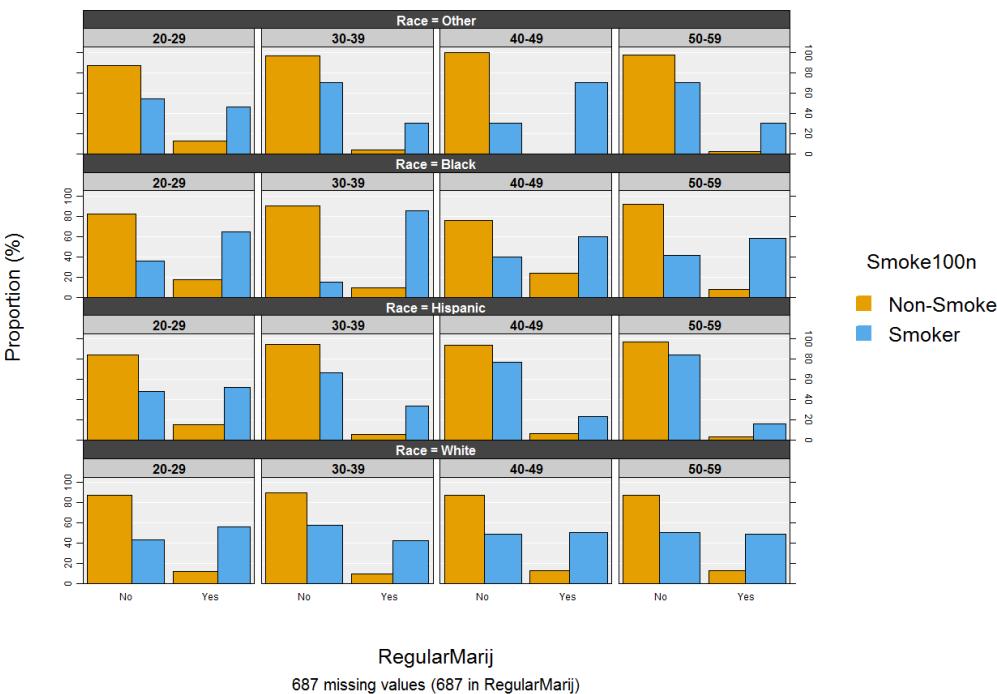
Distribution of RegularMarij by Smoke100n



Distribution of RegularMarij by Smoke100n subset by AgeDecade



Distribution of RegularMarij by Smoke100n subset by AgeDecade and Race

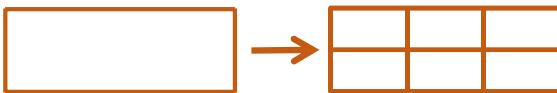


# Motion ...



## Mathematically speaking ...

Subsetting



+



... Stepping through subsets

=

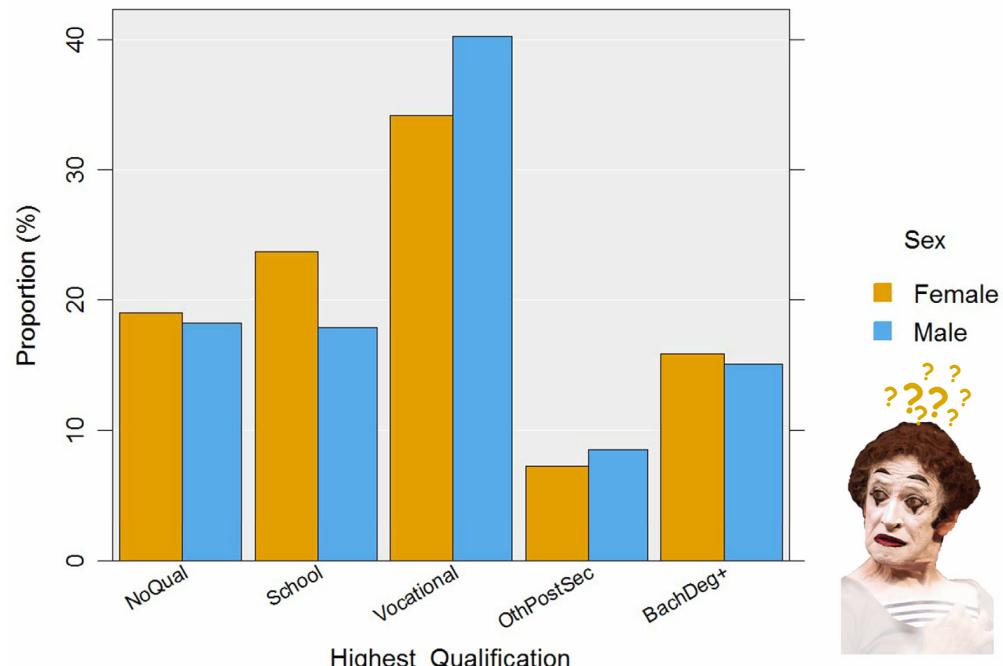


Motion ...



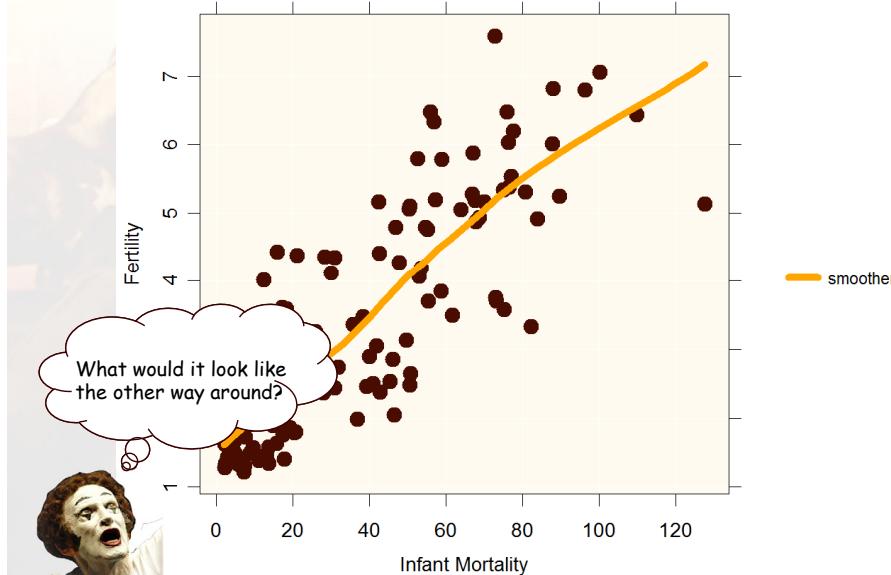
Bait & switch ...

## Distribution of Highest Qualification by Sex

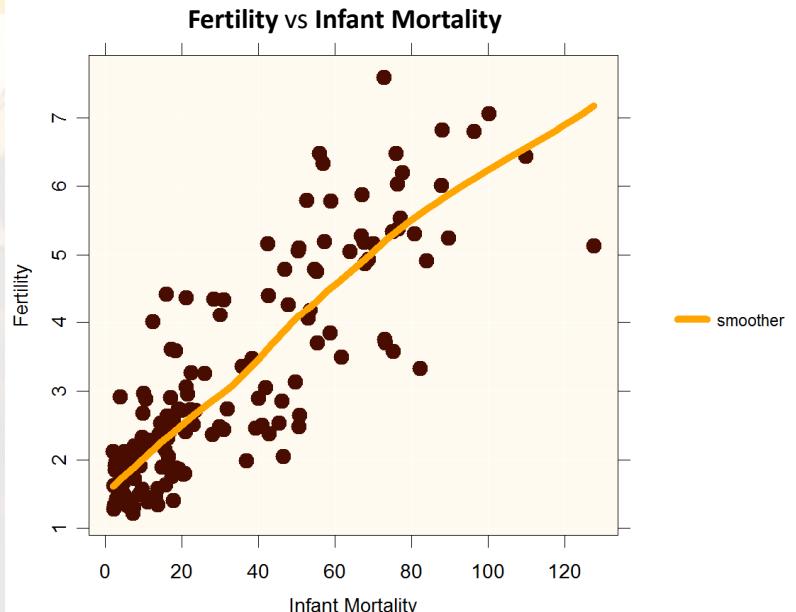


## Switching the roles of variables: Switching “X and Y”

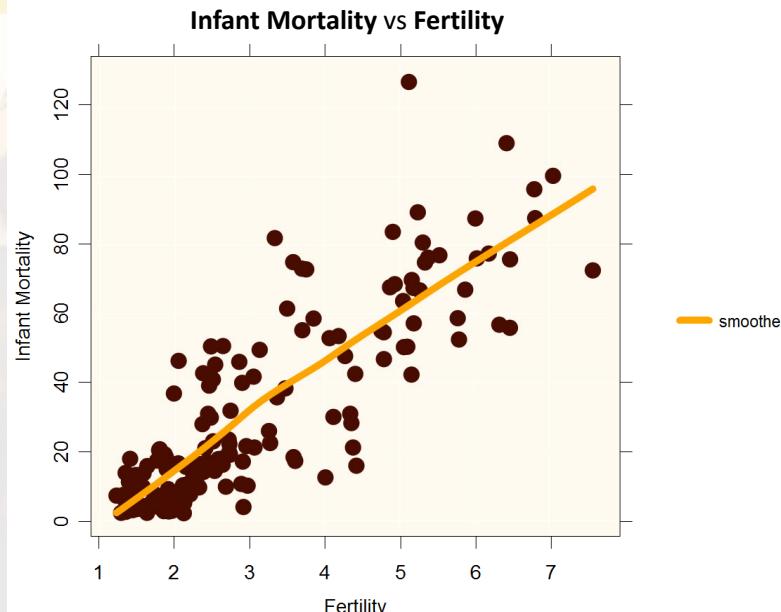
Fertility vs Infant Mortality



## Switching the roles of variables: Switching “X and Y”



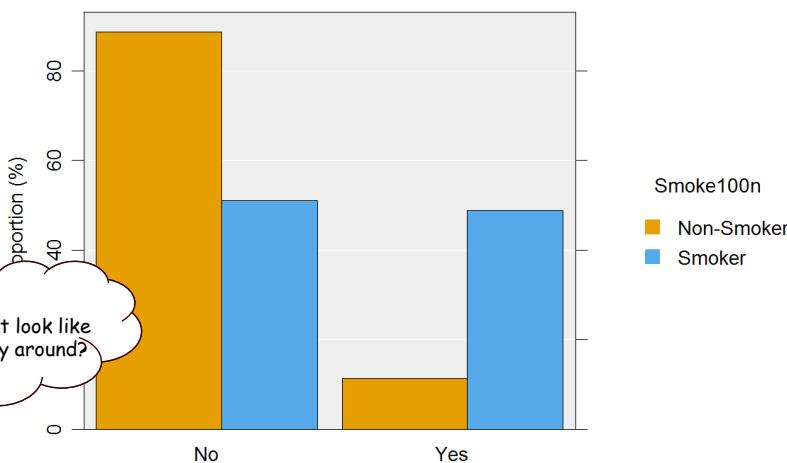
## Switching the roles of variables: Switching “X and Y”



## 2-way table switching “X and Y”



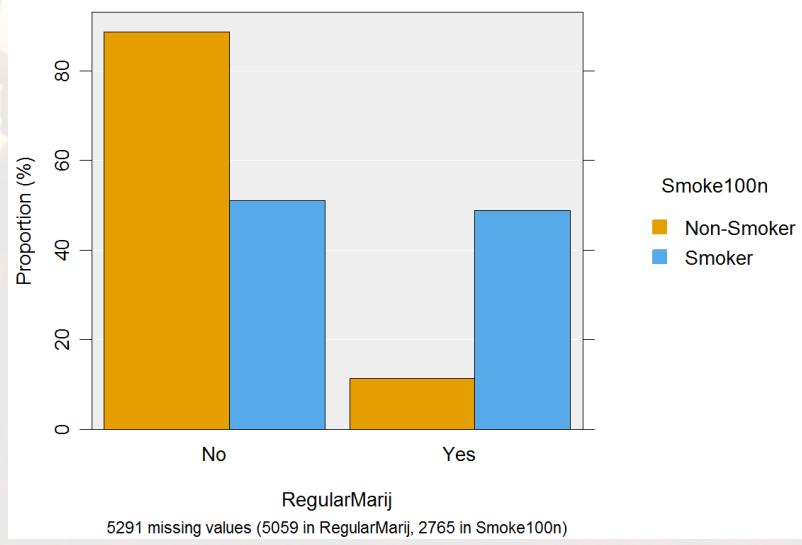
Distribution of RegularMarij by Smoke100n



5291 missing values (5059 in RegularMarij, 2765 in Smoke100n)

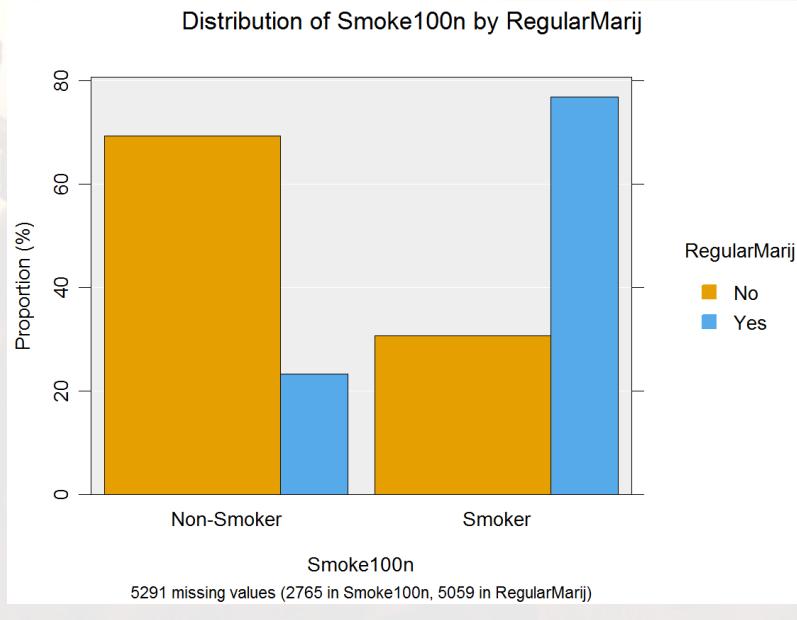
## 2-way table switching “X and Y”

Distribution of RegularMarij by Smoke100n

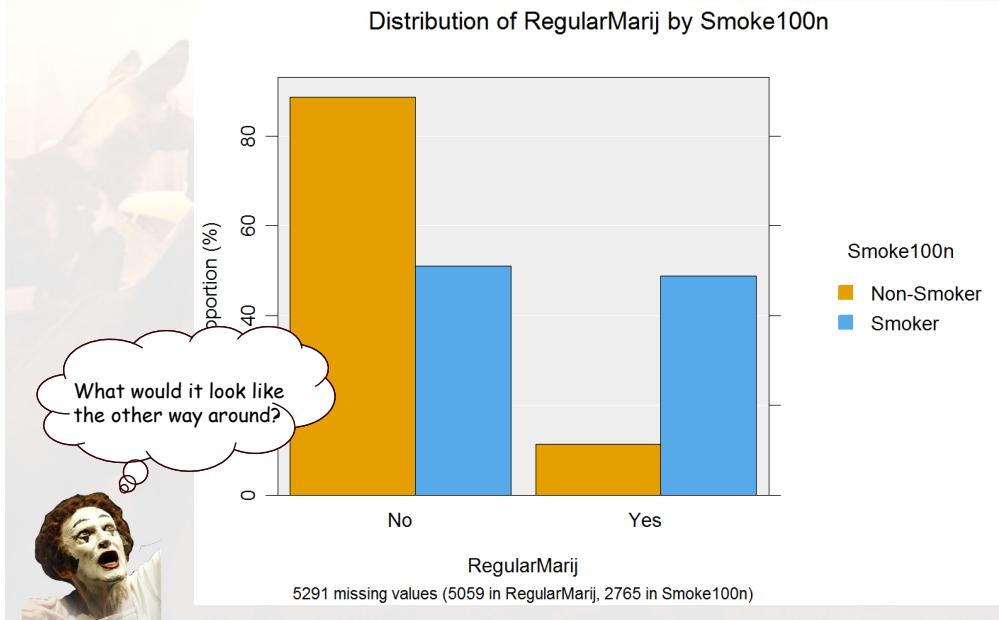


5291 missing values (5059 in RegularMarij, 2765 in Smoke100n)

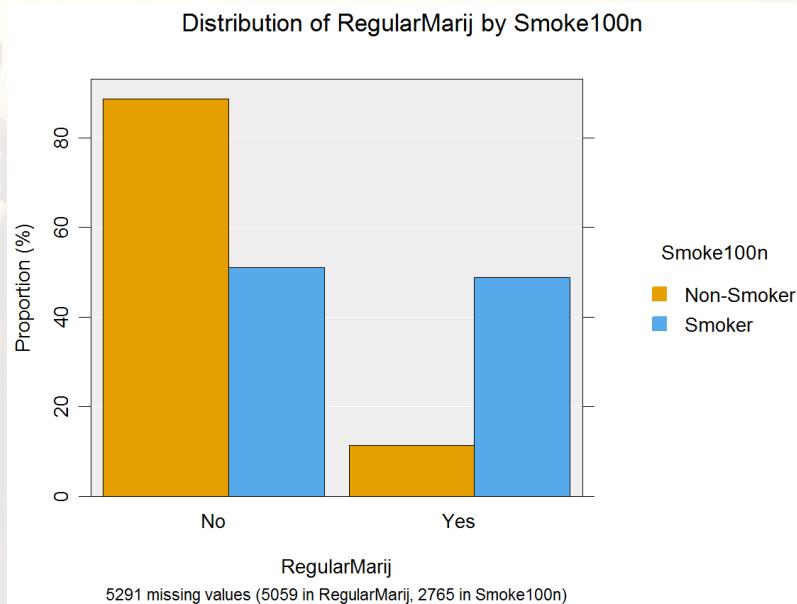
## 2-way table switching “X and Y”



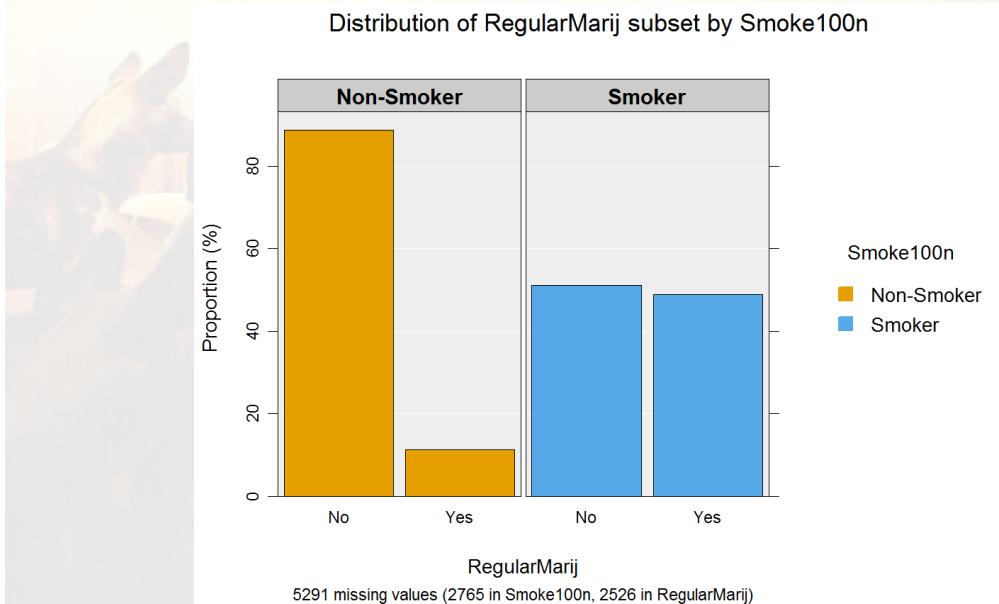
## 2-way table switching representations



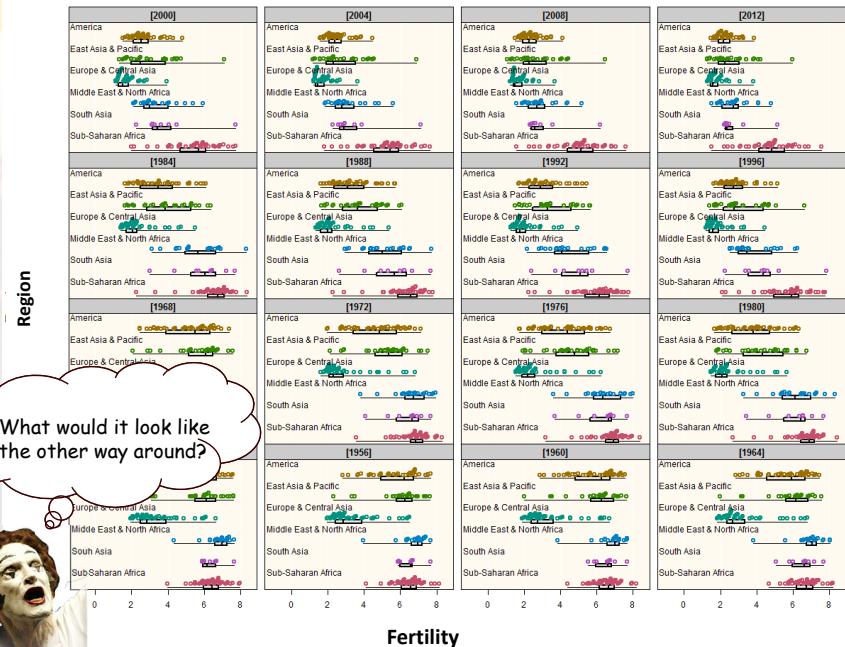
## 2-way table switching representations



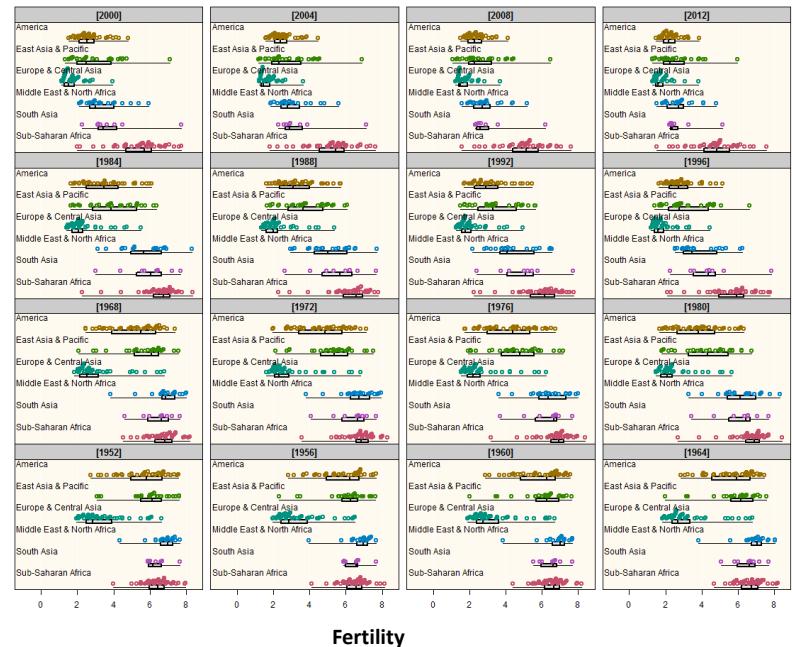
## 2-way table switching representations



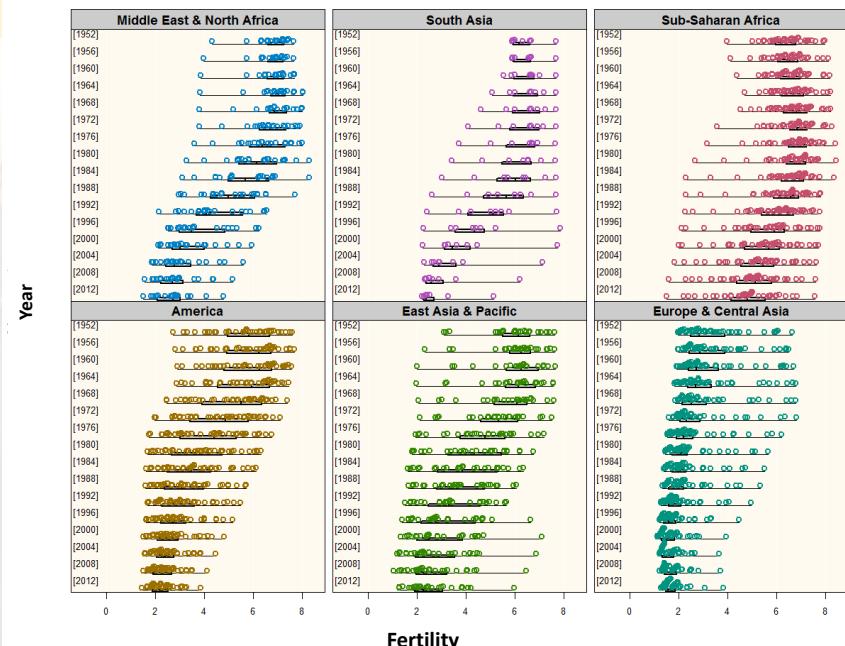
## Fertility by Region subset by Year



## Fertility by Region subset by Year



## Fertility by Year subset by Region



Identification & brushing ...



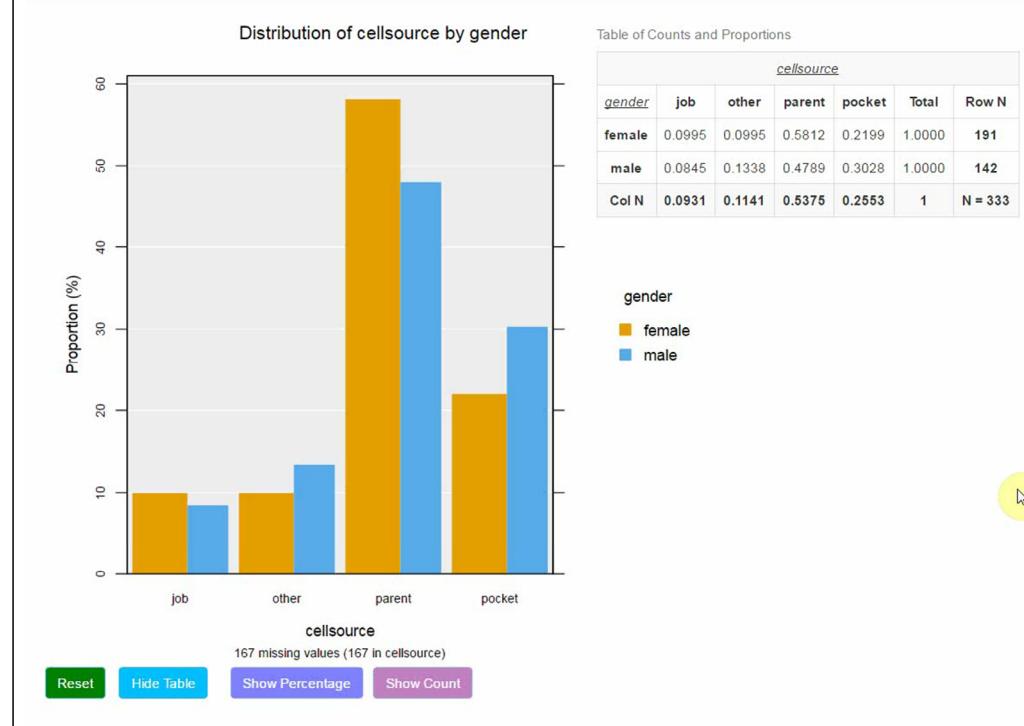
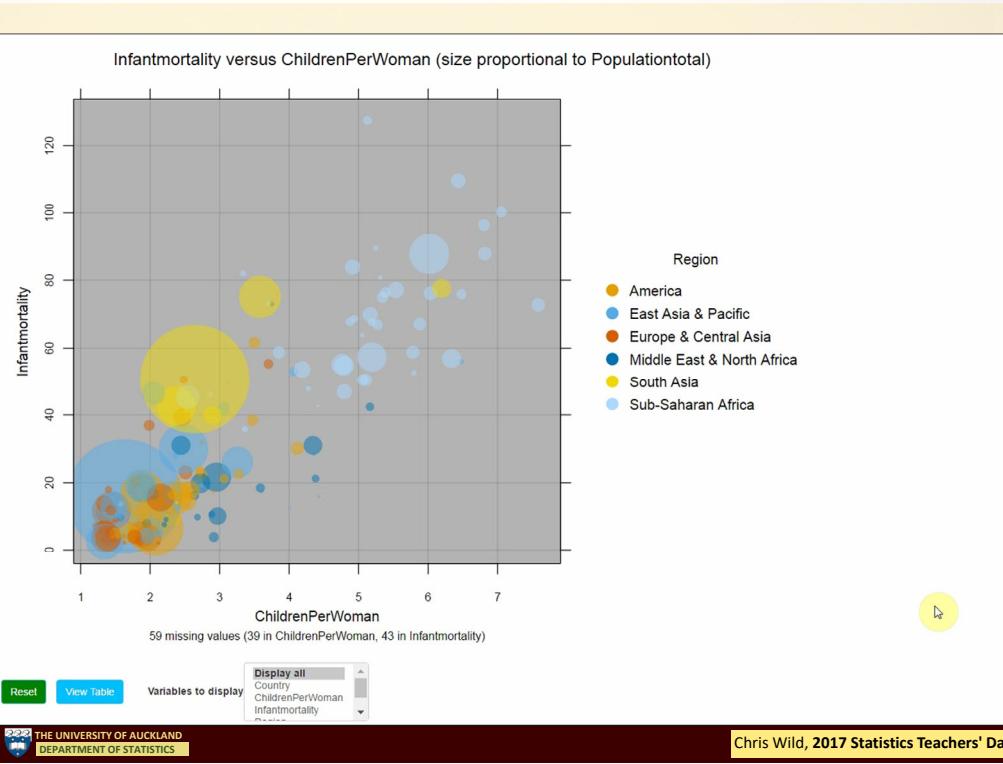
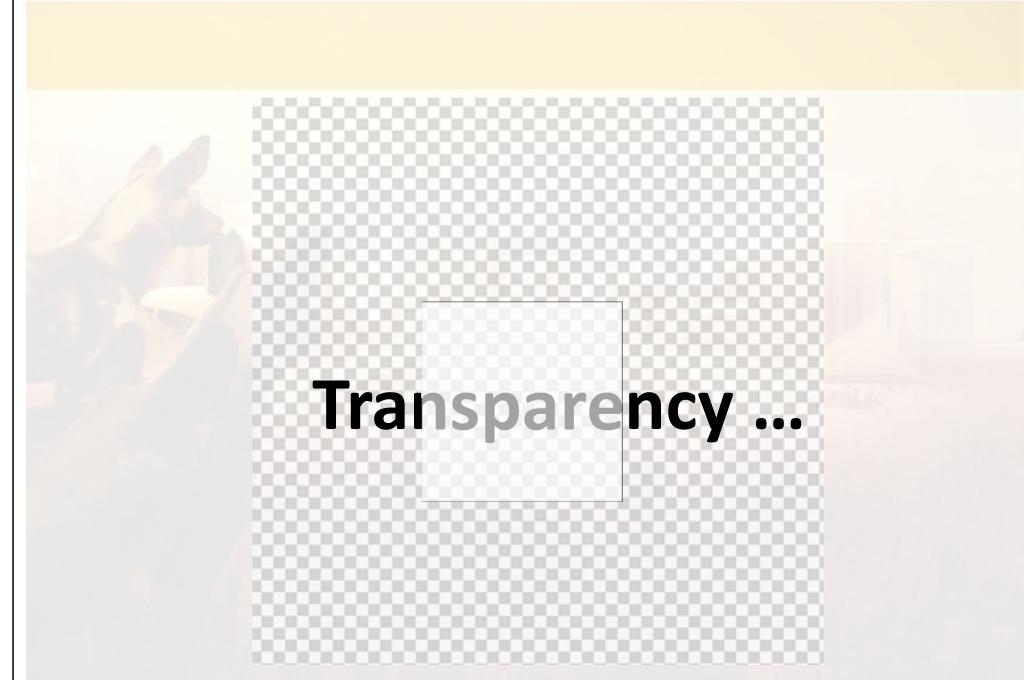
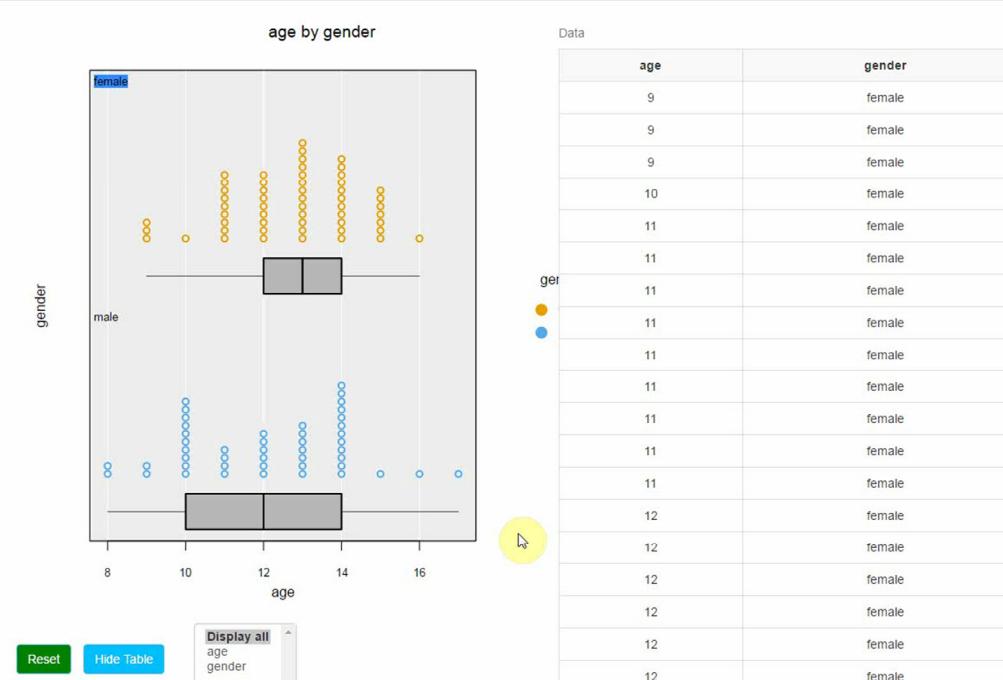


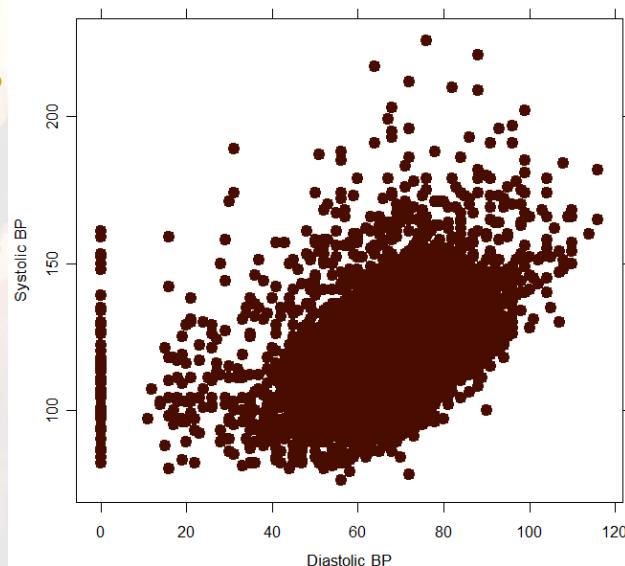
Table of Counts and Proportions

	cellsouce					Total	Row N
gender	job	other	parent	pocket			
female	0.0995	0.0995	0.5812	0.2199	1.0000	191	
male	0.0845	0.1338	0.4789	0.3028	1.0000	142	
Col N	0.0931	0.1141	0.5375	0.2553	1	N = 333	



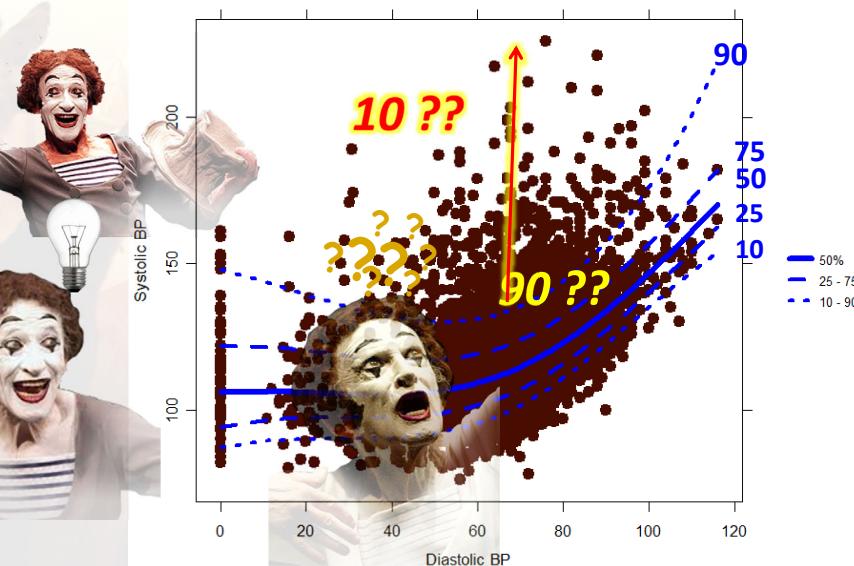
# Blood pressures for 10,000 people

Systolic vs Diastolic Blood Pressures



# Blood Pressures for 10,000 people

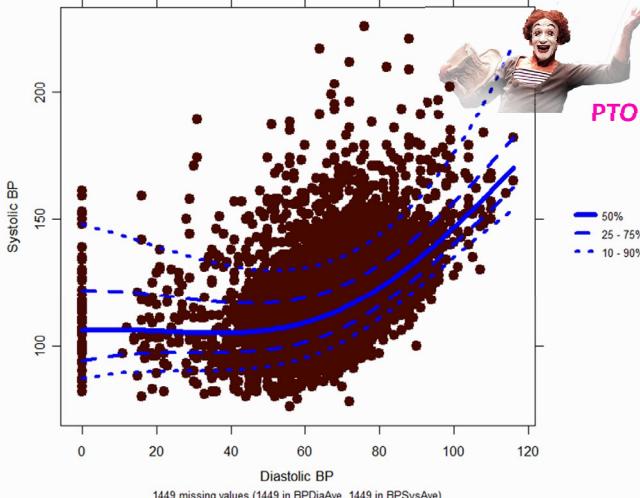
Systolic vs Diastolic Blood Pressures



## Transparency

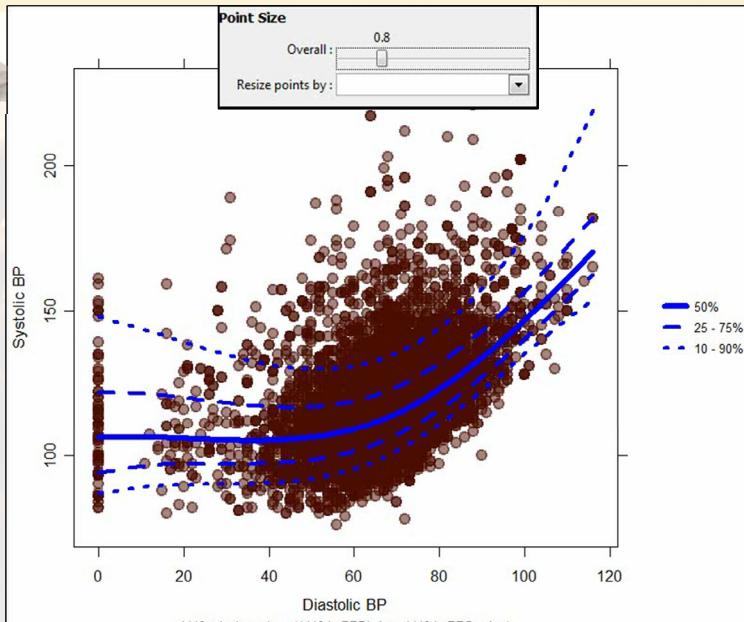
Transparency : 0

Systolic vs Diastolic Blood Pressures



## Reducing Point Size

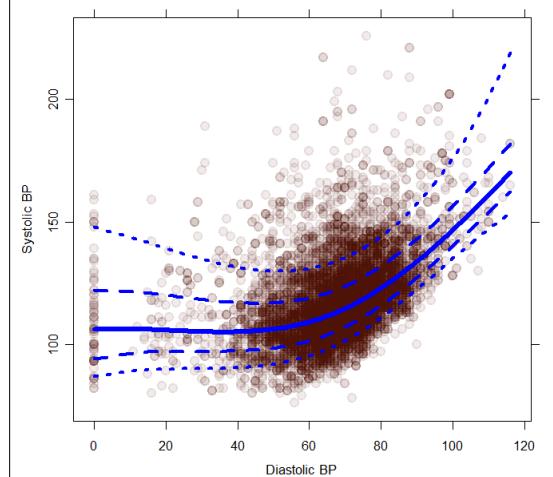
Point Size  
Overall : 0.8  
Resize points by :



# Emphasizing ...

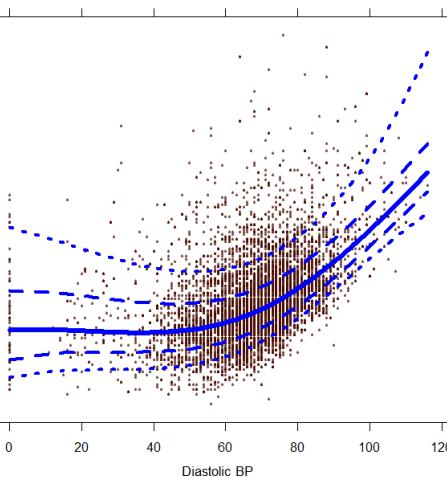
## Density

Systolic vs Diastolic Blood Pressures



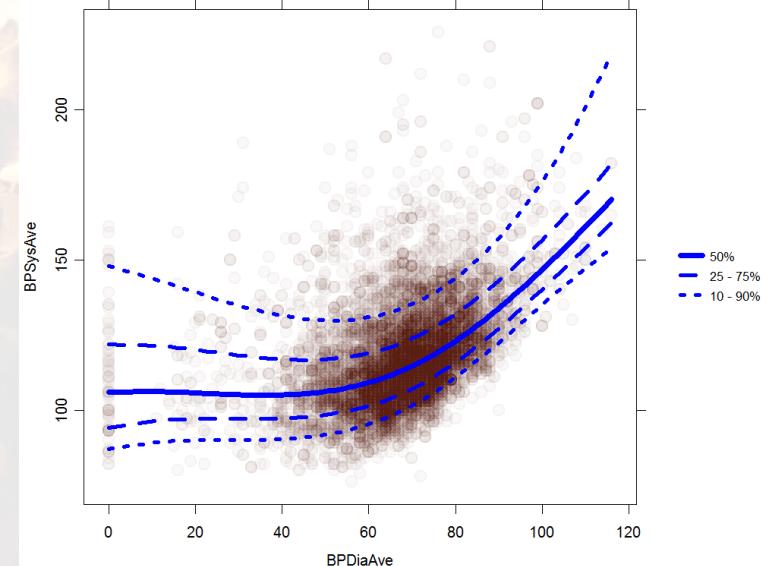
## Discreteness

Systolic vs Diastolic Blood Pressures



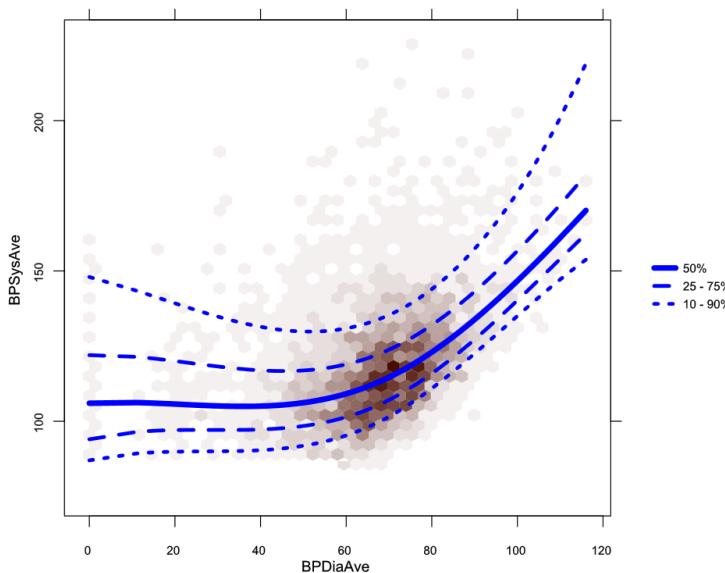
# Scatter plot with transparency

BPSysAve versus BPDiaAve



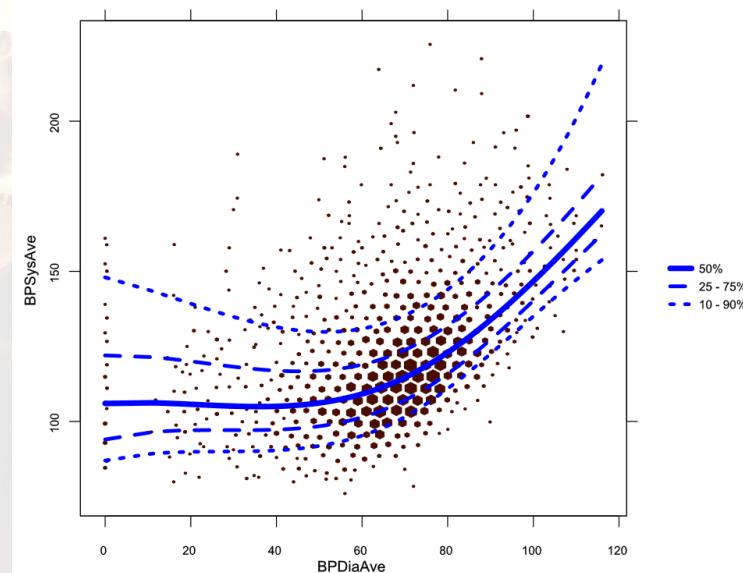
# "Big data" version ... Hexbin (alpha)

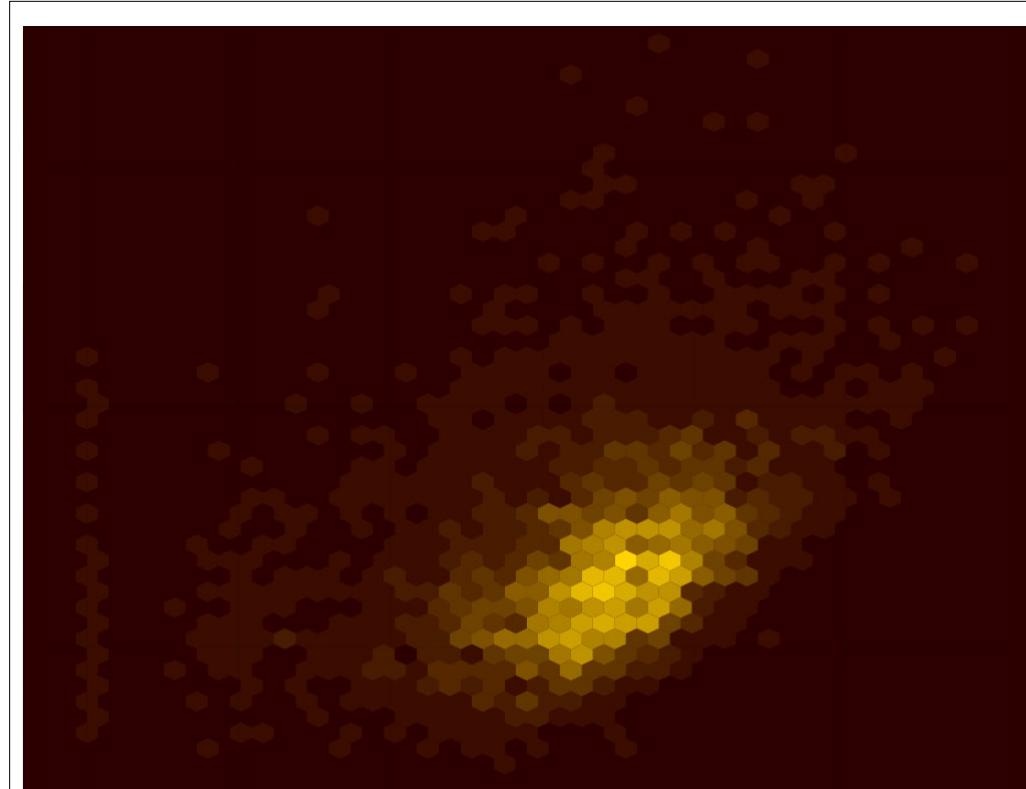
BPSysAve versus BPDiaAve



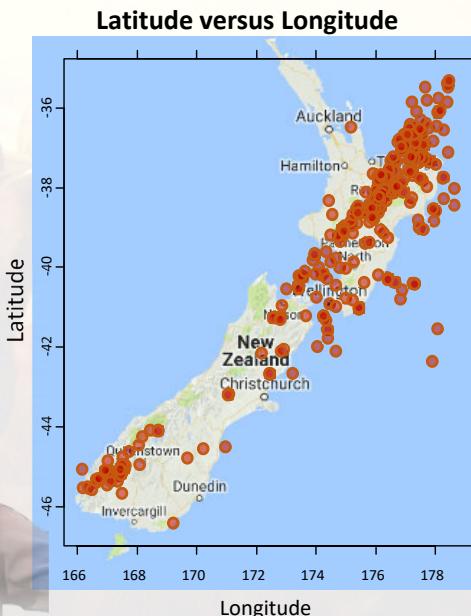
# "Big data" version ... Hexbin (size)

BPSysAve versus BPDiaAve



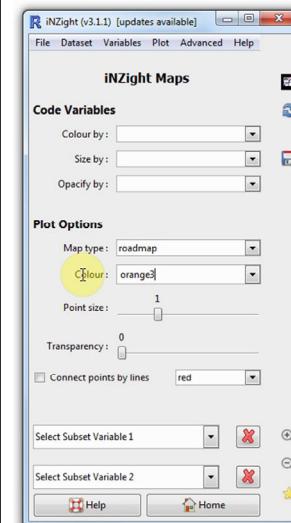


## New Zealand Earthquakes (in 2000)



Earthquakes ...

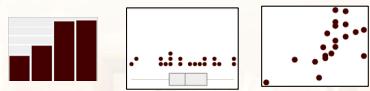
## Locations of Earthquakes



# Today's cooking demonstration

## RAW INGREDIENTS

- Bar charts
- Dot plots
- Scatter plots



## ADDED SPICES

- Colouring
- Sizing
- Transparency
- Subsetting
- Identification
- Bait & switch

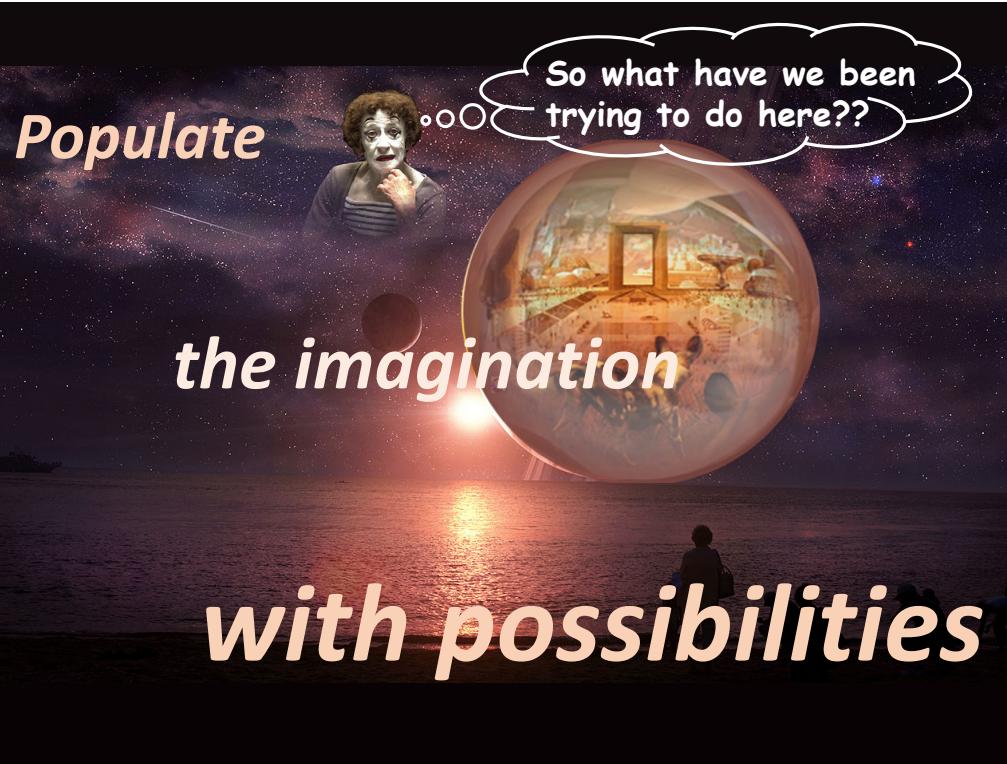
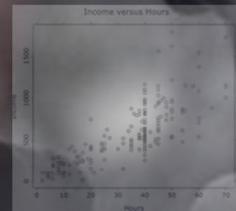
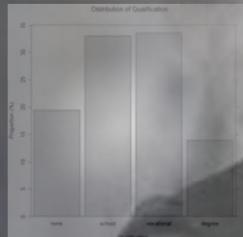


## MAGIC SUPER SPICE

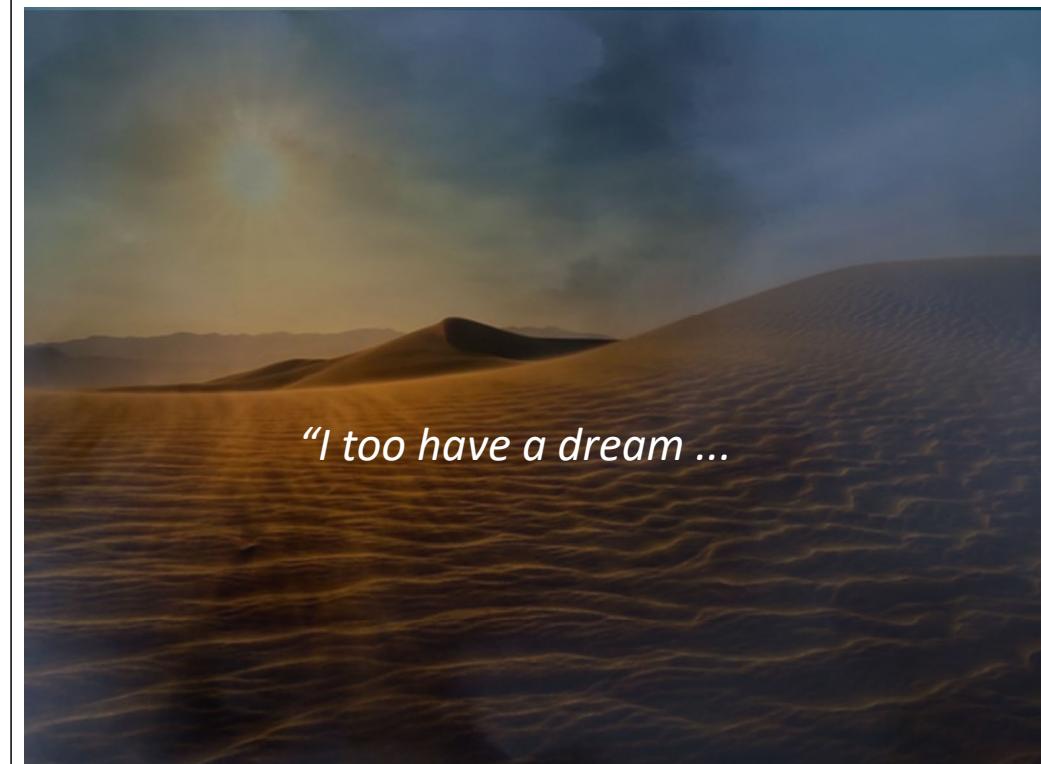
- Stepping



Chris Wild, 2017 Statistics Teachers' Day



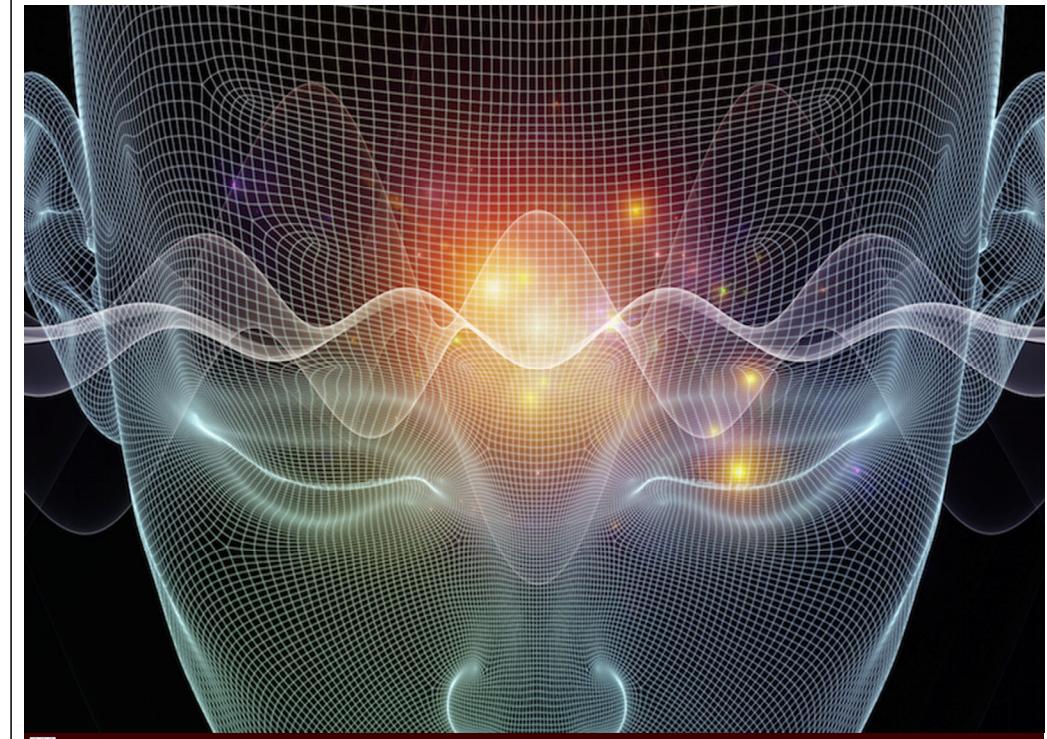
*"I too have a dream ..."*



*"I have a dream of students spellbound by the broad vistas of the data landscape*

*I have a dream of their swooping over it effortlessly on magic carpets exploring its nooks and crannies in search of its hidden treasures*

*I have a dream of students empowered to look at data and keep , like Alice, crying "*Curiouser and curiouser!*", and have the ability and confidence to go where that curiosity leads*

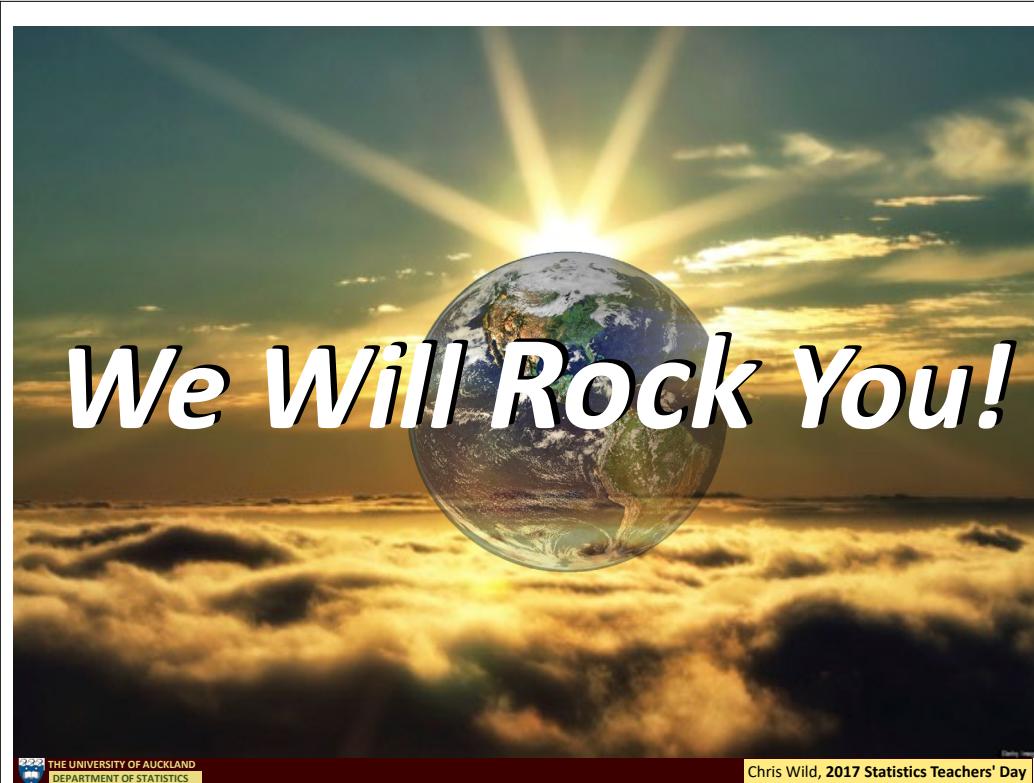




Please Rise

for the official Anthem  
of Statistics Teachers Day

We Slap, Slap, Clap,  
We will Rock you ... / We will Plot you ...



We Will Rock You!