

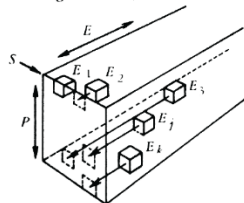


Imagining an interface for Statistical modeling

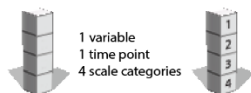
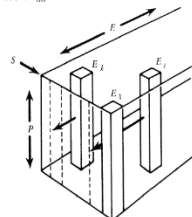
(2) Mixed Set Facet

(a) Simple Mixed Facet

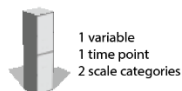
Projection of single cell values onto a facet to give a *mixed facet*.



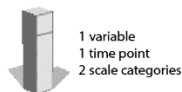
(b) Staggered Facet



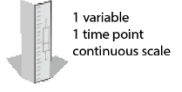
1 variable
1 time point
4 scale categories



1 variable
1 time point
2 scale categories



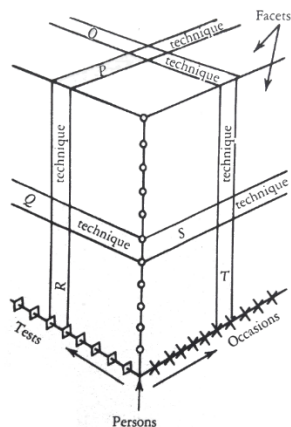
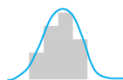
1 variable
1 time point
2 scale categories



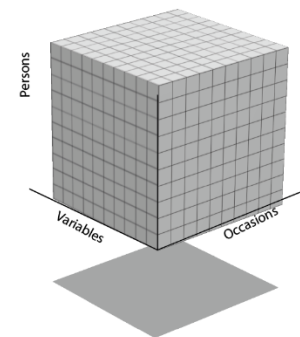
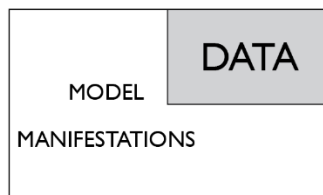
1 variable
1 time point
continuous scale

$$Y_i = f(x)$$

$$Y_i = f(x) + e_i$$



DATA = MODEL + ERROR
DATA = FIT + RESIDUAL
DATA = SMOOTH + ROUGH



1 variable
1 time point
1 person

1 variable
1 time point
3 persons

Scale
Dimension

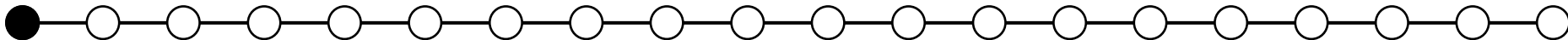
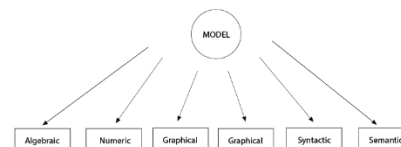
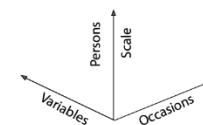


Parallel dimensions
Parallel dimensions

Person
Dimension

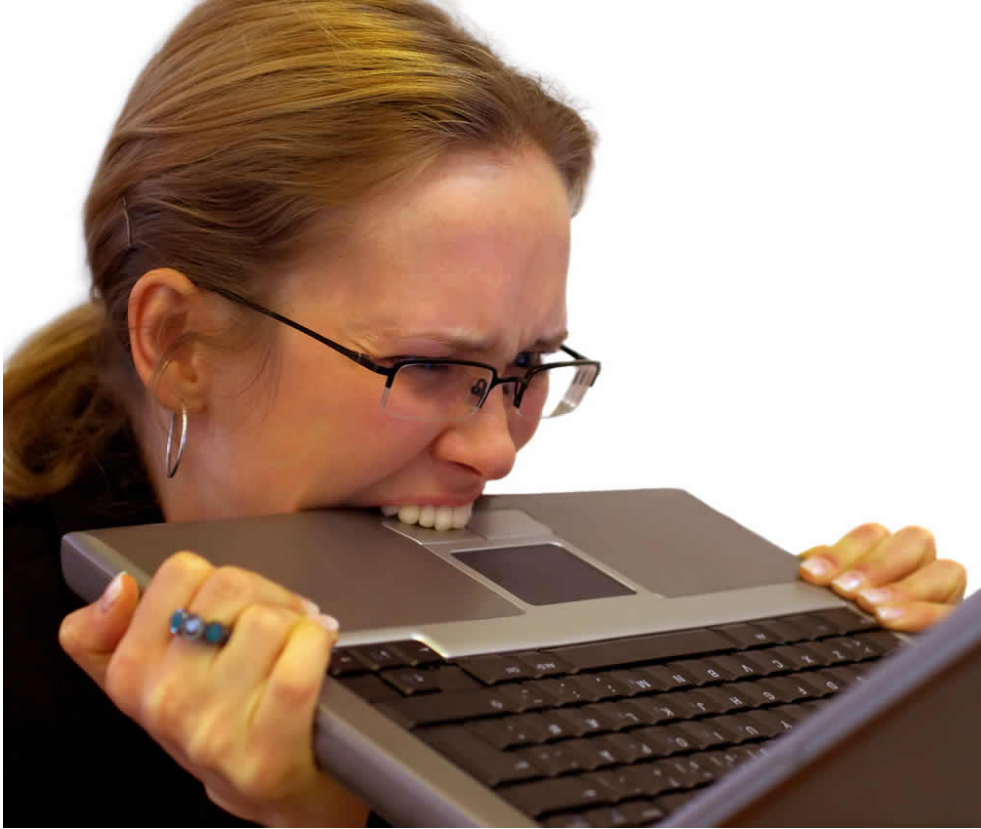


1 variable
1 time point
16 persons

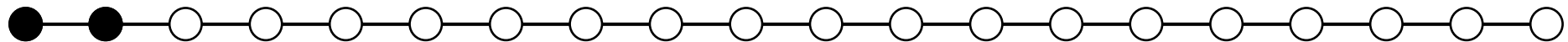




Statistical modeller at work

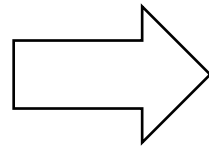


- Attention span
- Levels of abstraction
- Visual processing
- Mental space architecture
- Concept management





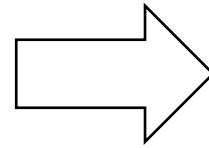
Interface for **flying**



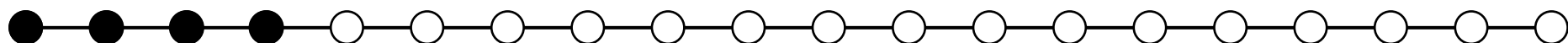
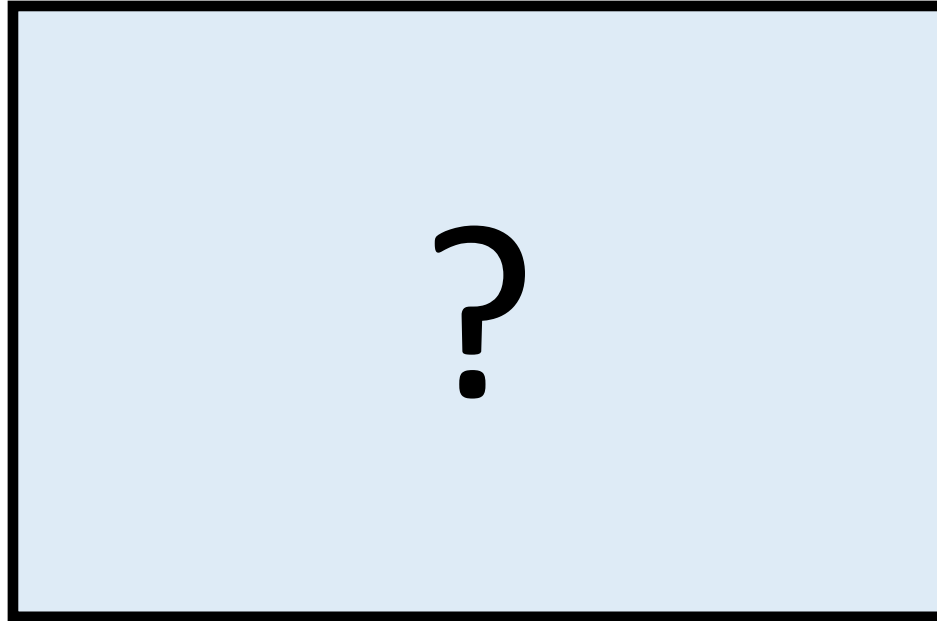
Operations with concepts of
speed, azimuth, yaw, roll, pitch



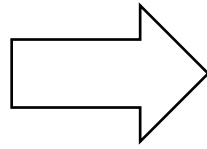
Interface for **modeling**



Operations with concepts of
fit, significance, parsimony



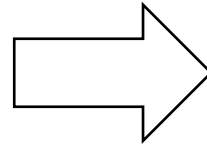
Interface for **flying**



Operations with concepts of speed, azimuth, yaw, roll, pitch

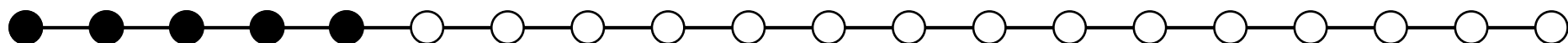


Interface for **modeling**



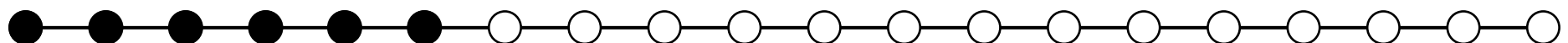
Operations with concepts of fit, significance, parsimony

- Framework?
- Discipline?
- Software?
- Data Types?



Models are becoming more

- Complex
- Sophisticated
- Numerous



```
Proc MIXED;
class id;
model attend = timec/
              solution;
random INTERCEPT timec/
      SUB=id TYPE=UN G;
```

Ssas

```
MODEL:
int by   timec1@1 timec2@1 timec3@1 timec4@1 timec5@1
        timec6@1 timec7@1 timec8@1 timec9@1 timec10@1 timec11@1;
linear by timec1@0 timec2@1 timec3@2 timec4@3 timec5@4
        timec6@5 timec7@6 timec8@7 timec9@8 timec10@9 timec11@10;
        [timec1-timec11@0]; timec1-timec11;
        [int linear]; int linear;
        int with linear;
```

Mplus

$$y_{ij} = \beta_{0j} + \beta_{1j} \text{timec} + \varepsilon_{ij} \quad \varepsilon_{ij} \sim N\left([0], [\sigma^2]\right)$$

$$\beta_{0j} = \gamma_{00} + u_{0j} \quad \begin{bmatrix} u_{0j} \\ u_{1j} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \tau_{00} & \\ & \tau_{10} \quad \tau_{11} \end{bmatrix}\right)$$

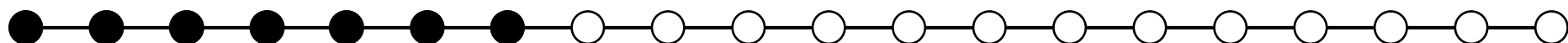
$$\beta_{1j} = \gamma_{10} + u_{1j}$$

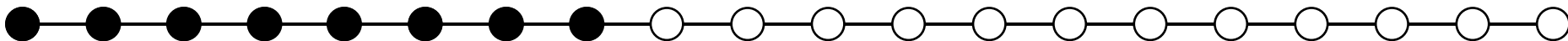
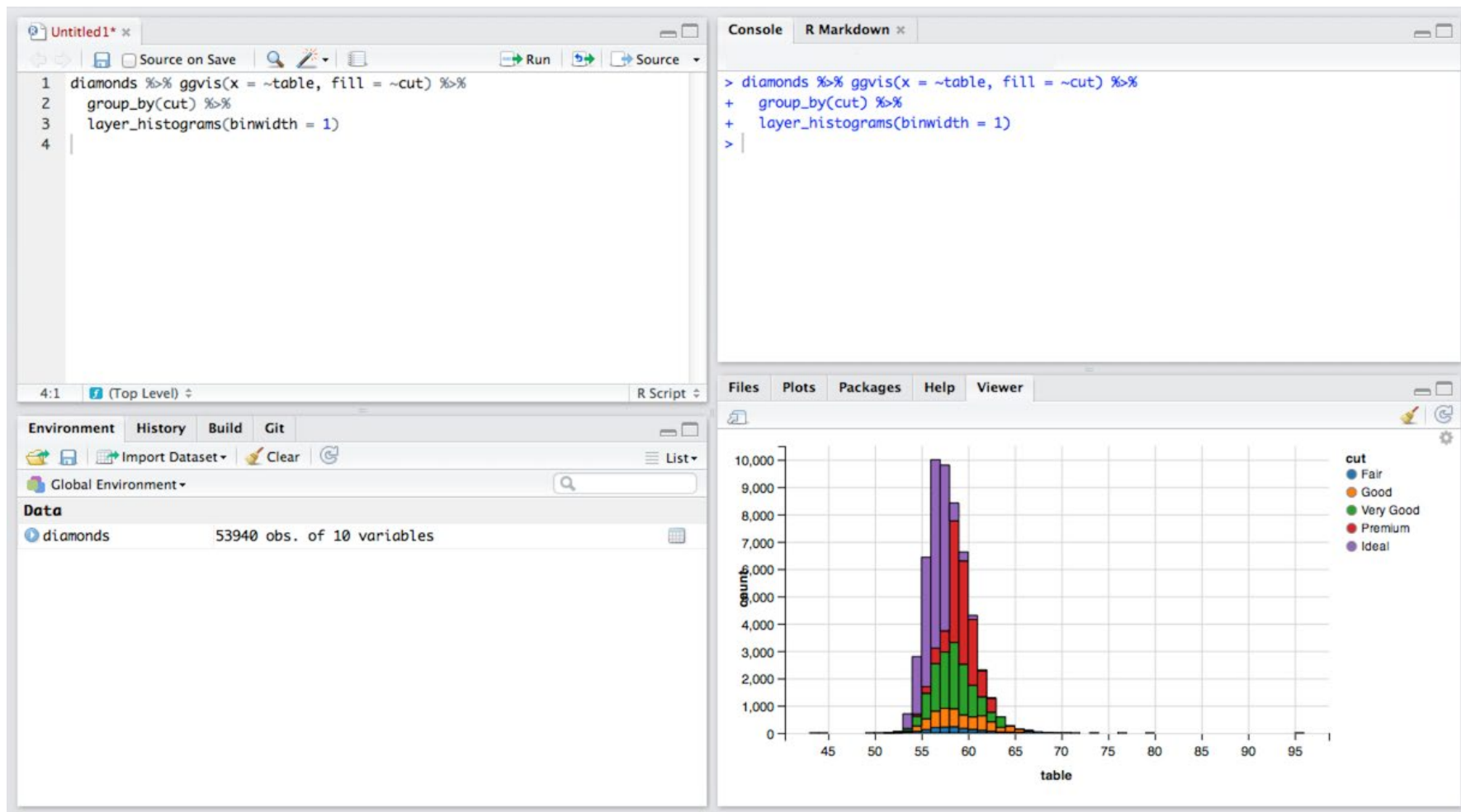


SPSS

```
MIXED attend WITH timec
/FIXED= INTERCEPT timec
/RANDOM= INTERCEPT timec | SUBJECT(id)
COVTYPE(UN)
```

```
model <- lmer (attend ~ 1 +
timec + (1 + timec | id))
```





Tabular

id	time	attend	model
1	0	1	2.788
1	1	6	2.732
1	2	2	2.675
1	3	1	2.618
1	4	1	2.562
1	5	1	2.505
1	6	1	2.449
1	7	1	2.392
1	8	1	2.335
1	9	1	2.279
1	10	1	2.222
1	11	1	2.166
4	0	2	2.788
4	1	1	2.732

Algebraic

$$y_{it} = \beta_0 + \beta_1 \text{time}_t + \varepsilon_{it}$$

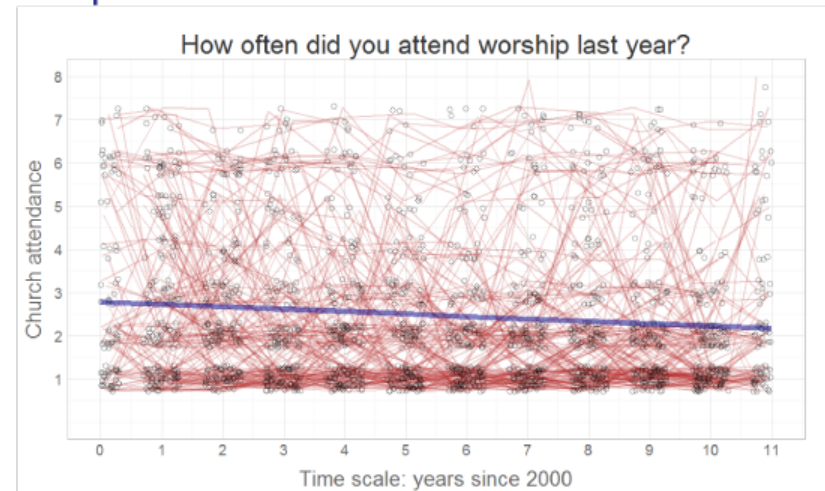
$$\beta_0 = \gamma_{00}$$

$$\beta_1 = \gamma_{10}$$

Semantic

In 2000 respondents attended church less than once a month (2.79) and gradually declined in their attendance since (.06 per year).

Graphical



Syntactic

```
nlme::gls(attend ~ 1 + time, data=dsM)
```

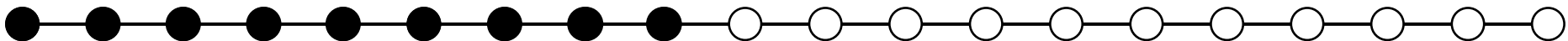
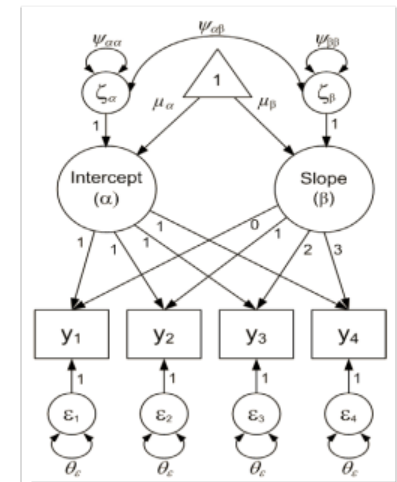
Numeric

Coefficients:

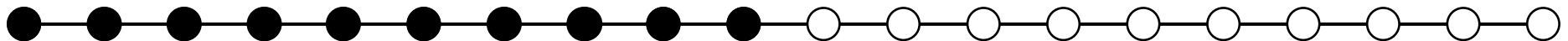
	Value	Std.Error	t-value	p-value
(Intercept)	2.7882	0.07774	35.86	0
time	-0.0566	0.01197	-4.73	0

	modelB
logLik	-3719
deviance	7438
AIC	7444
BIC	7461
df.resid	1858
N	1860
p	2
ids	155

Schematic



Scientists make digital art
by combining model manifestations
to tell stories about data





Meaning

cells → tissues

How do **models** comprise a **meaning**?

cells



tissue

Model

How do **variables** comprise a **model**?

molecules



cell

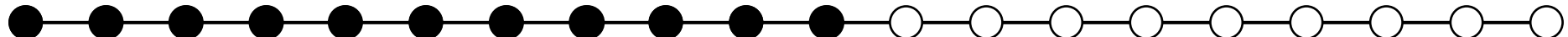
Data

How do **data** comprise the **variable**?

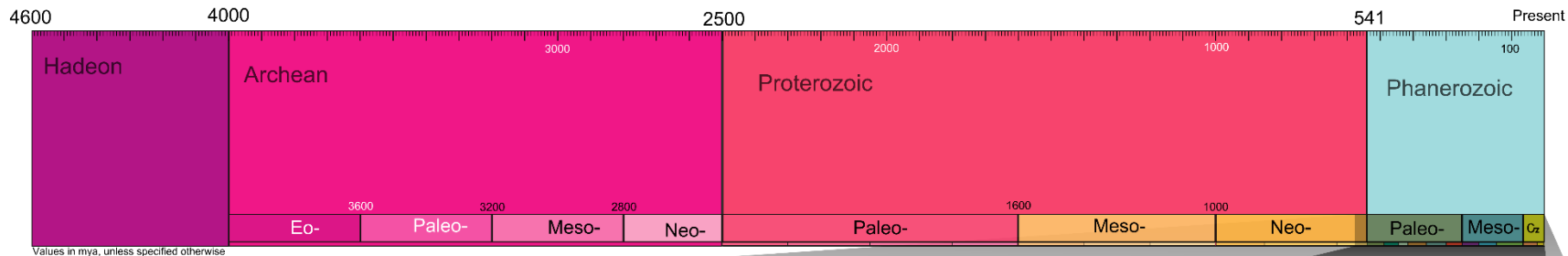
atoms



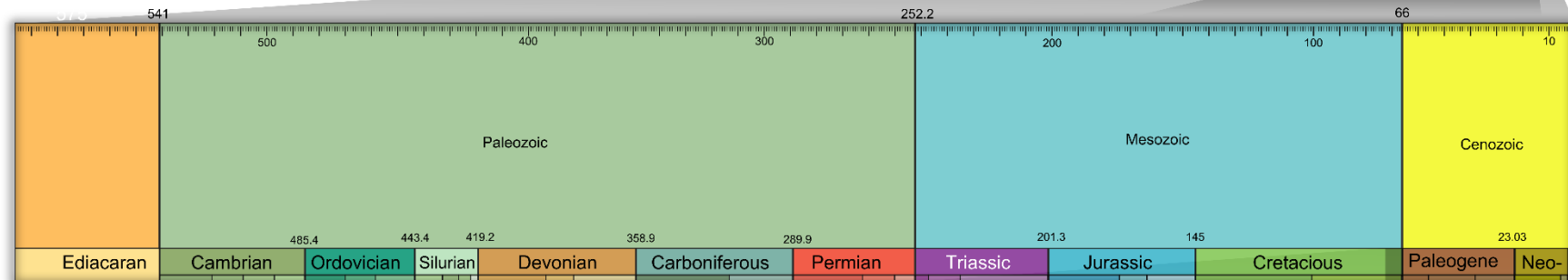
molecule



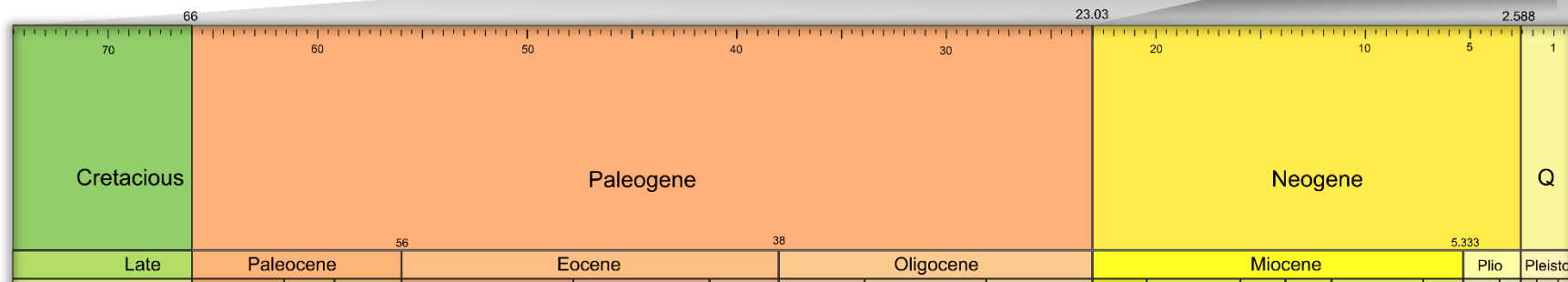
Meaning



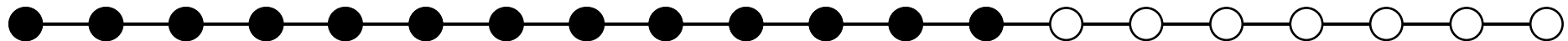
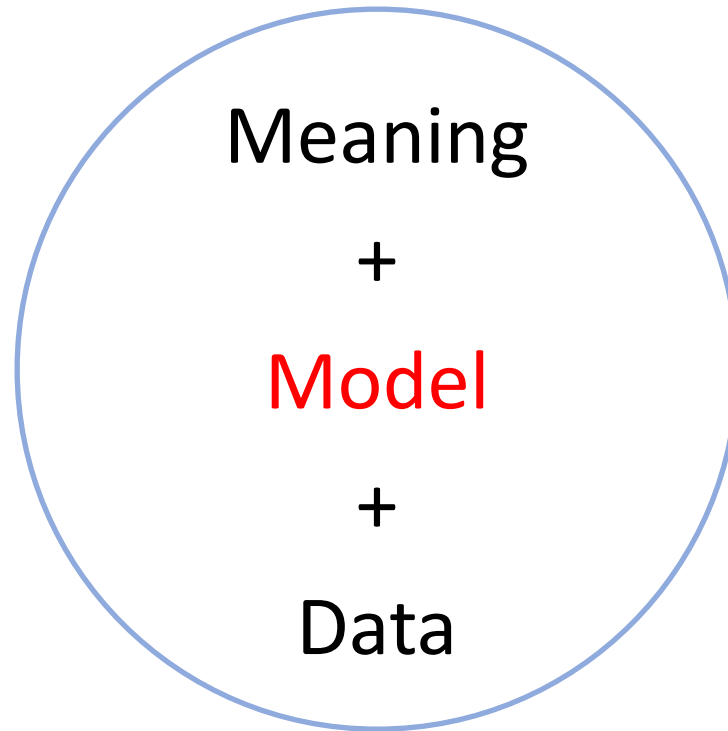
Model

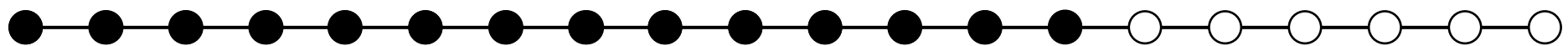
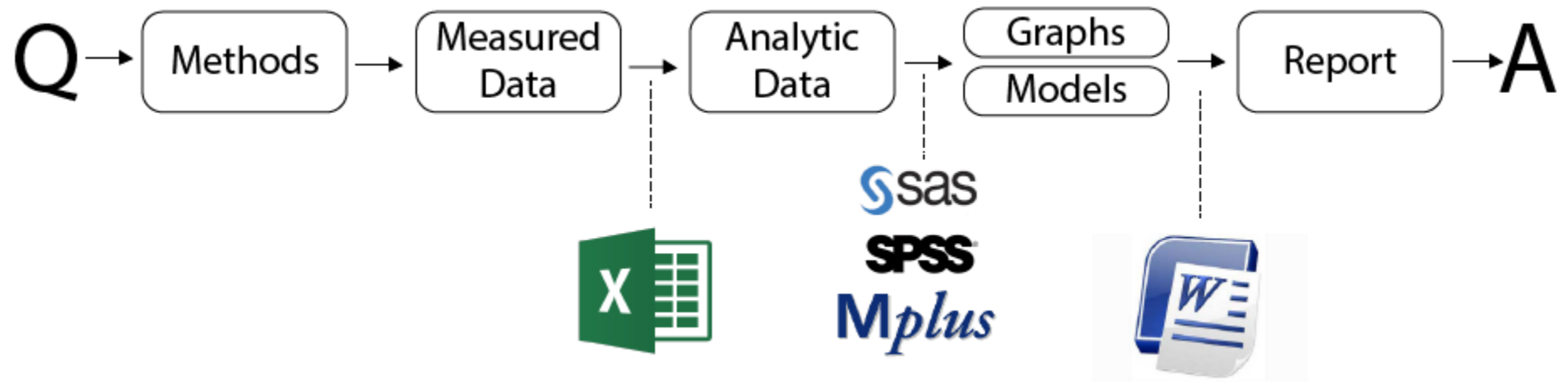


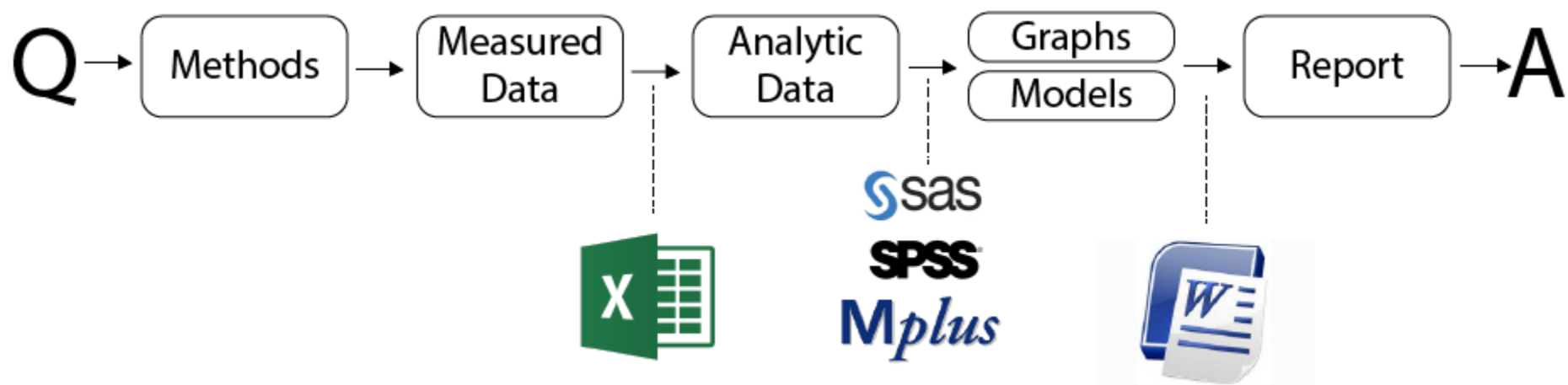
Data



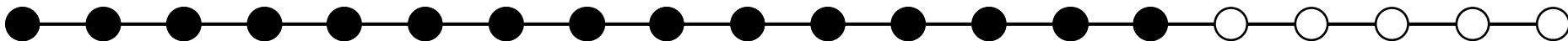
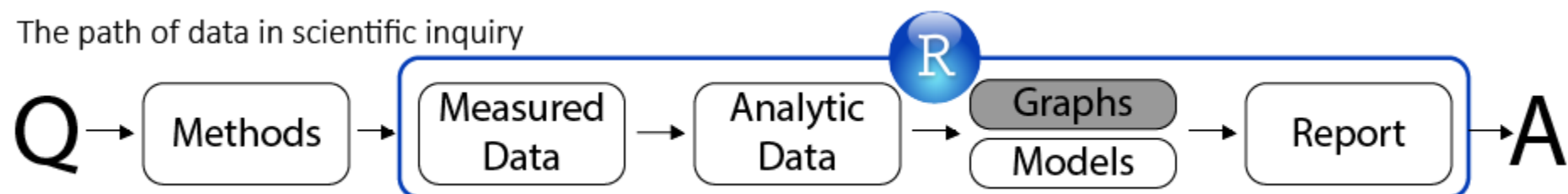
Dynamic Reporting







The path of data in scientific inquiry



Data Manipulation

R

Statistical Modeling

RStudio

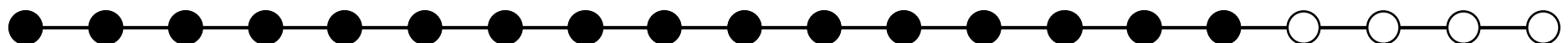
Graph Production

Git

Dynamic Reporting

GitHub

The path of data in scientific inquiry



Prototype of a model sequencer

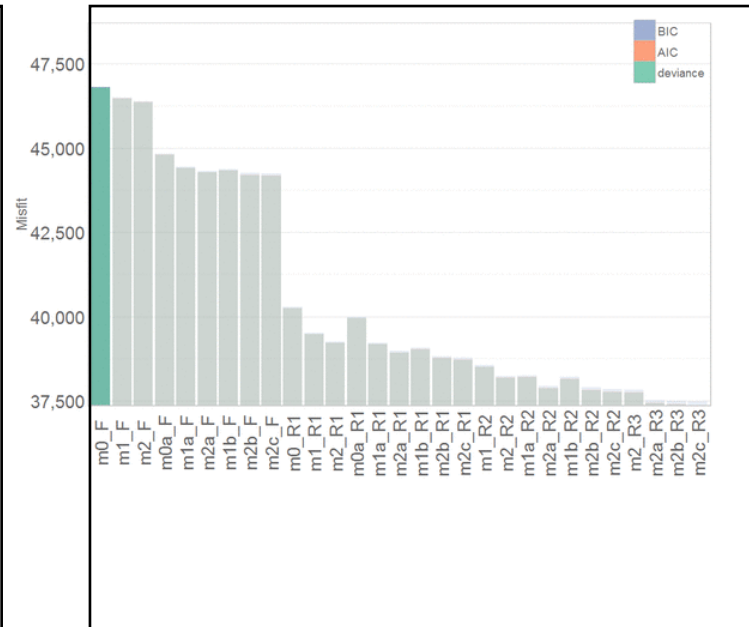
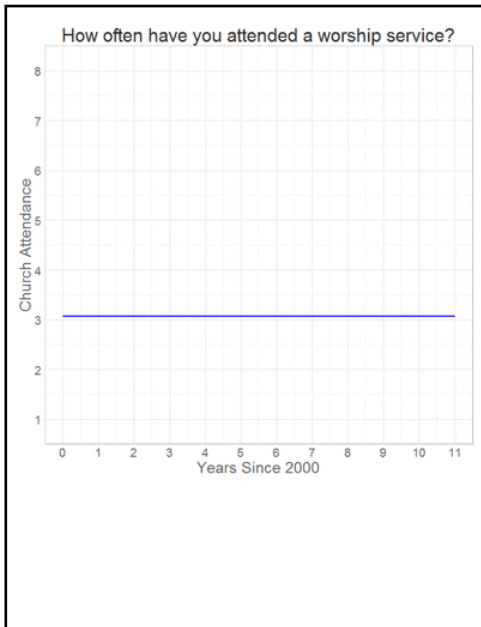
$$y_{it} = \beta_{0i} + \varepsilon_{it}$$

$$\beta_{0i} = \gamma_{00}$$

*F

m0_*	m1_*	m2_*
m0a_*	m1a_*	m2a_*
	m1b_*	m2b_*
		m2c_*

	Estimate	Std.Error	t.value	SD	tau0	tau1	tau2	tau3	sigma
(Intercept)	3.07	0.02	158.09		0.00				2.03
timec					0.00				
timec2									
-									
attendPR									
timec:attendPR									
timec2:attendPR									

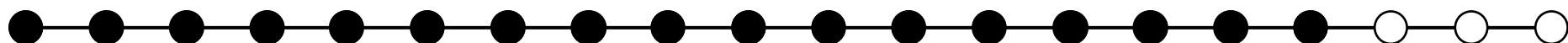


Model Specification

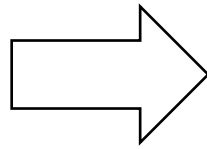
Estimates

Predicted
Values

Model
Fit



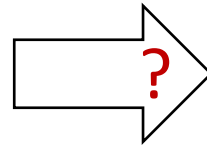
Interface for **flying**



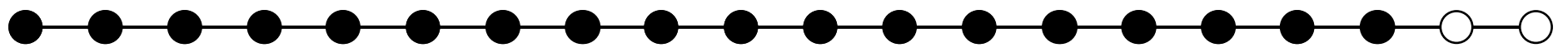
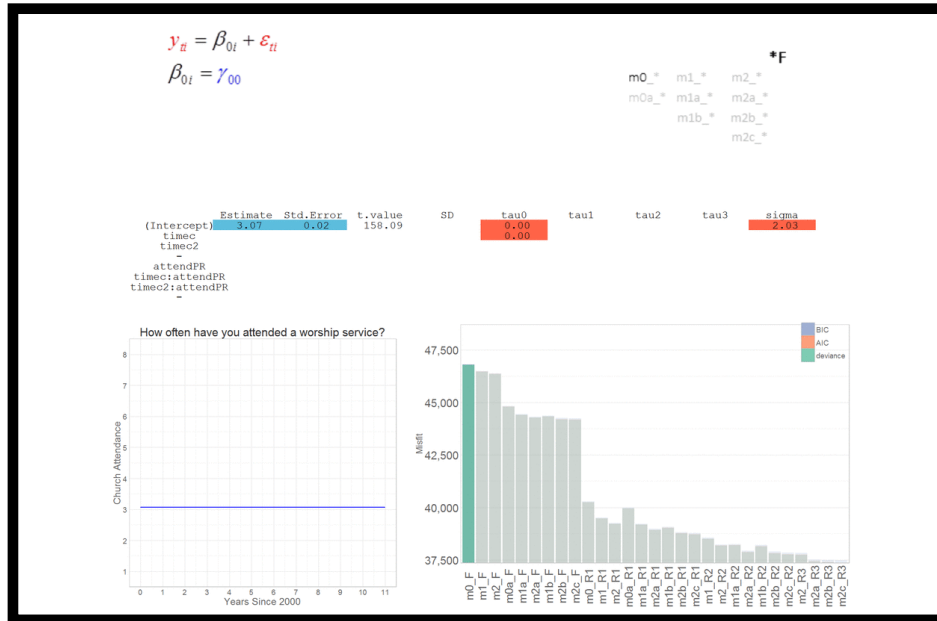
Operations with concepts of
speed, azimuth, yaw, roll, pitch



Interface for **modeling**



Operations with concepts of
fit, significance, parsimony



Specification

Evaluation

Communication

Meaning

How do **models** comprise a **meaning**?

Model

How do **variables** comprise a **model**?

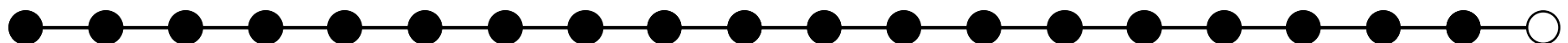
Data

How do **data** comprise the **variable**?

Algebra
Code

Within
Among

Print
Web



$$y_{ti} = \beta_{0i} + \varepsilon_{ti}$$

$$\beta_{0i} = \gamma_{00}$$

*F

m0_* m1_* m2_*
m0a_* m1a_* m2a_*
m1b_* m2b_*
m2c_*

	Estimate	Std.Error	t.value
(Intercept)	3.07	0.02	158.09
timec			
timec2			
-			
attendPR			
timec:attendPR			
timec2:attendPR			
-			

SD	tau0	tau1	tau2	tau3	sigma
	0.00				2.03
	0.00				

