ECO 602 Analysis of Environmental Data

FALL 2019 – UNIVERSITY OF MASSACHUSETTS

DR. MICHAEL NELSON

Deterministic Functions: Part 1

Assignment 1

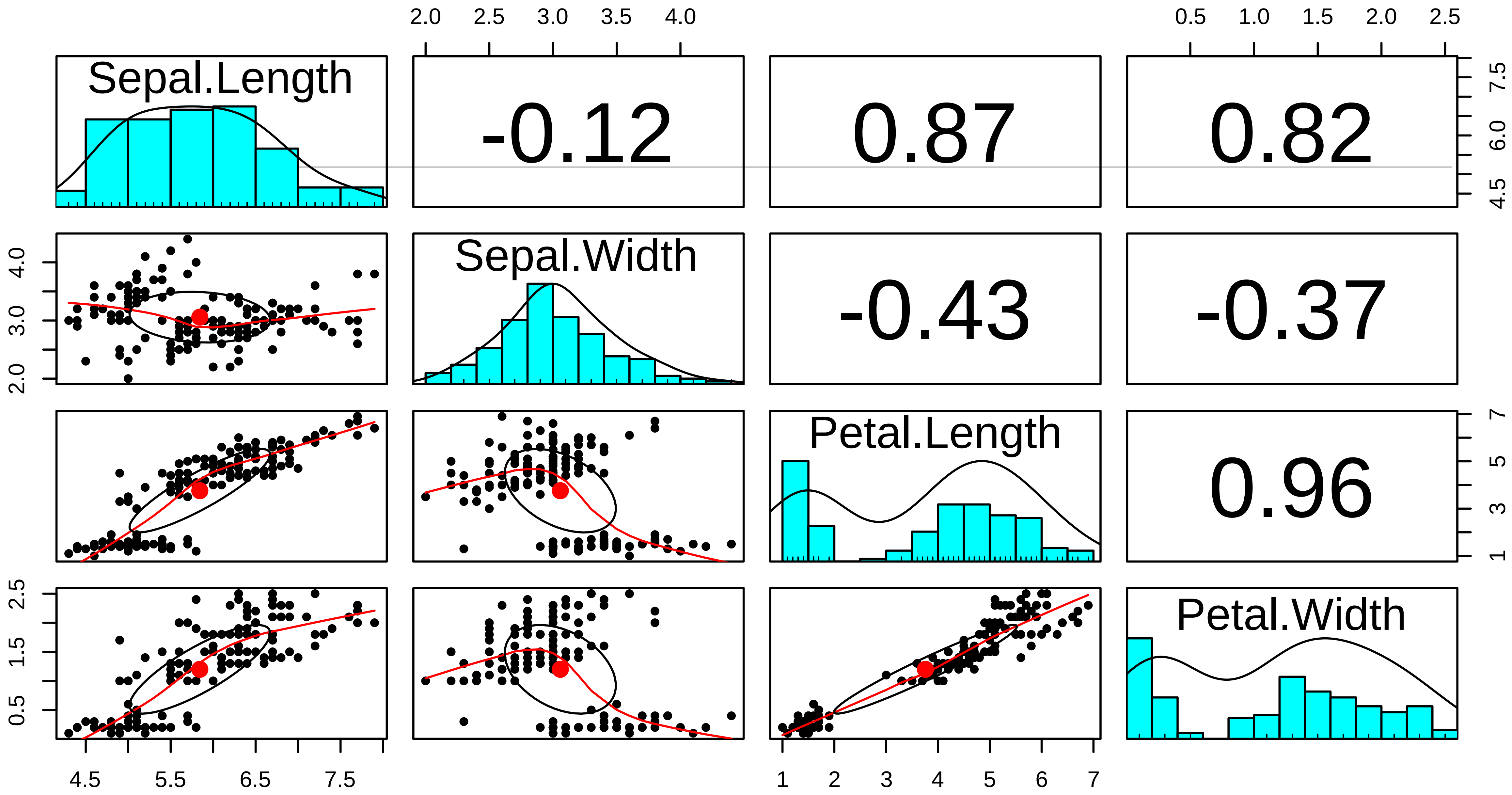
1. Due Sunday night, the 29th, by midnight.
2. Please email me your submissions and be sure to include everyone's names in the document.
3. Thank you to all those who sent drafts!

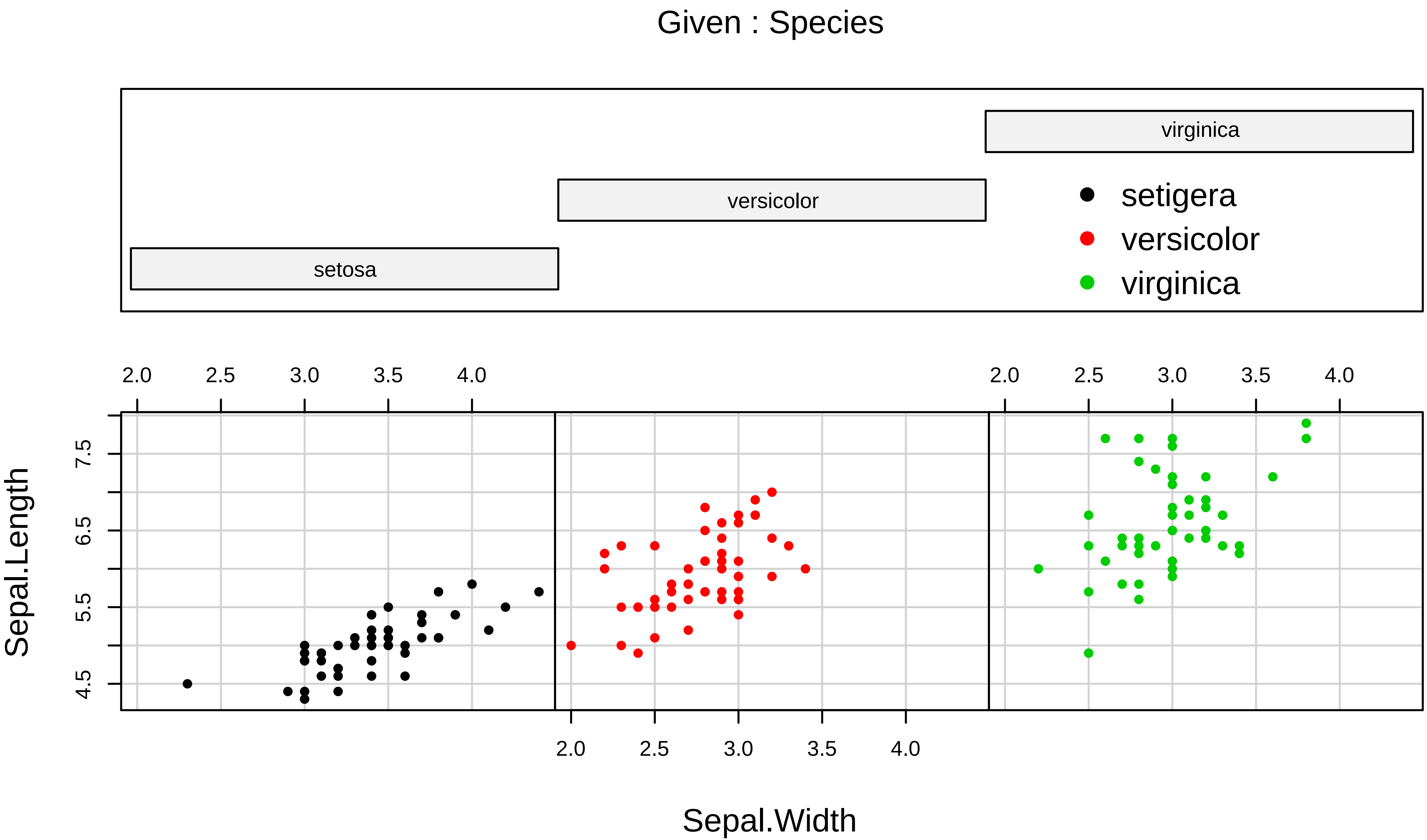
Today’s Agenda

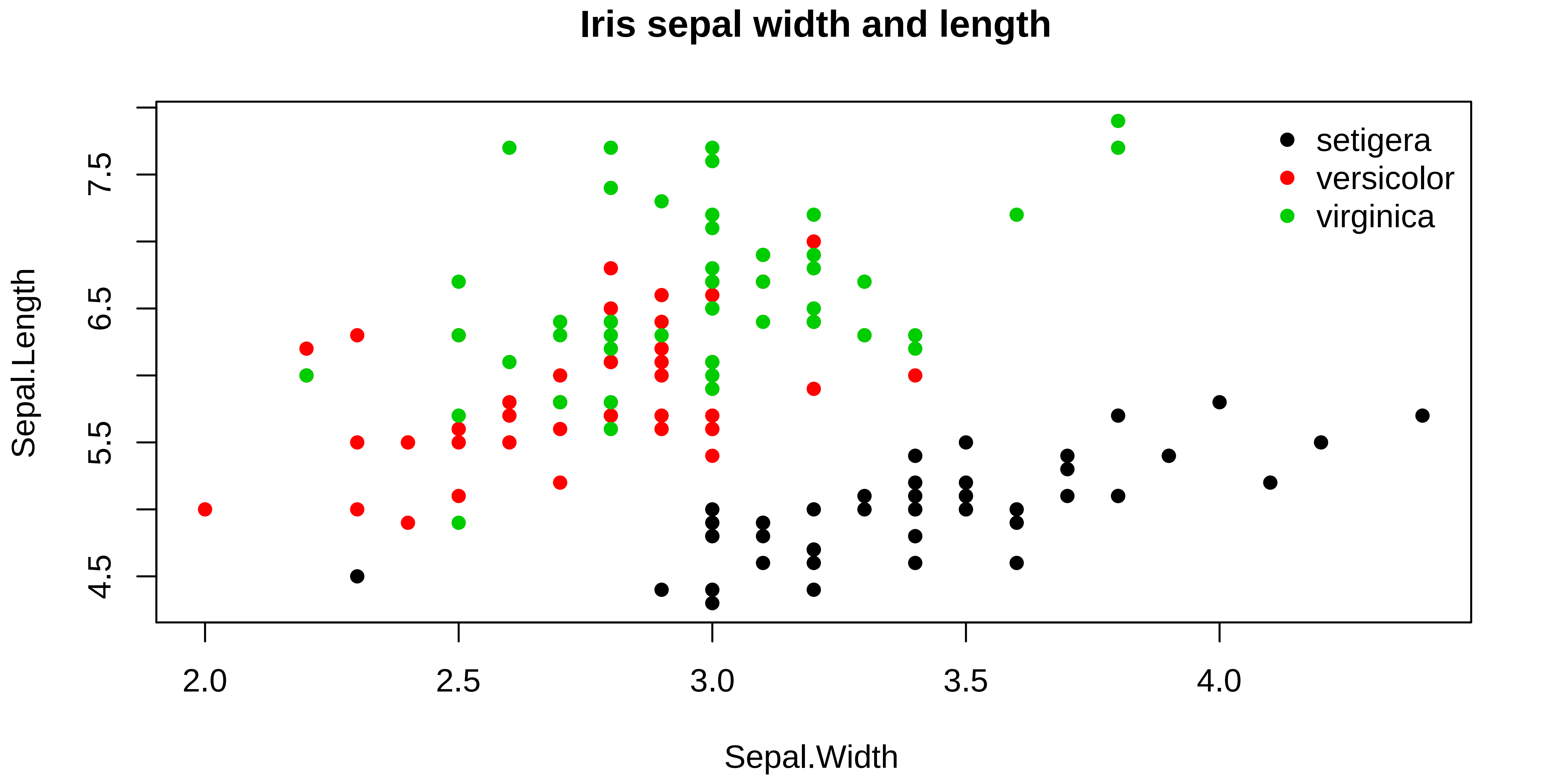
1. Chapter 3 recap
2. Data transformations briefly
3. Corrected plot examples
4. The two-model system
5. Deterministic functions within model thinking context

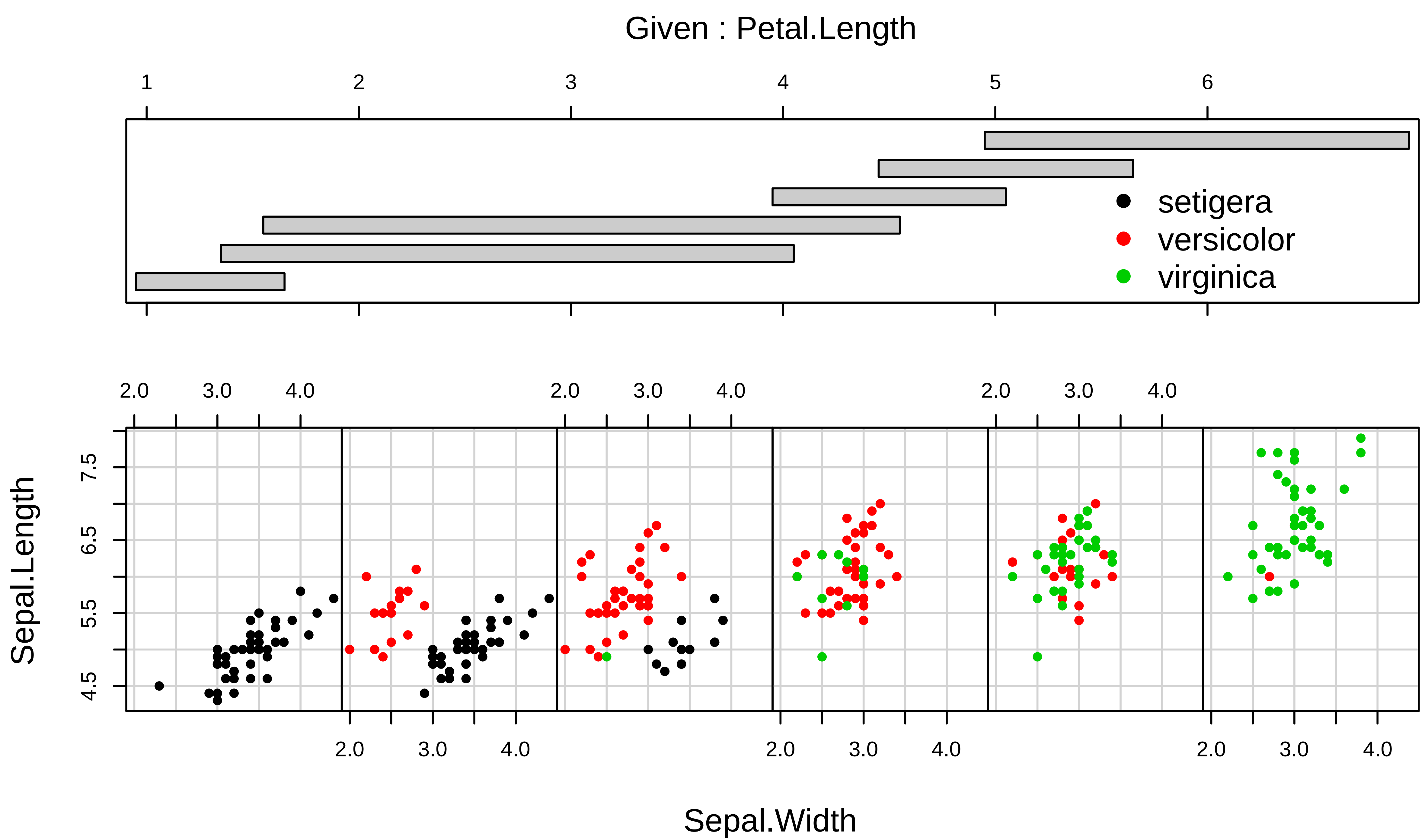
Iris example data in R

	Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species	col
1	5.1	3.5	1.4	0.2	setosa	1
2	4.9	3.0	1.4	0.2	setosa	1
3	4.7	3.2	1.3	0.2	setosa	1
4	4.6	3.1	1.5	0.2	setosa	1
5	5.0	3.6	1.4	0.2	setosa	1
6	5.4	3.9	1.7	0.4	setosa	1


Sepal.Length
-0.12
0.87
0.82
Sepal.Width
-0.43
-0.37
Petal.Length
0.96
Petal.Width
2.0 2.5 3.0 3.5 4.0
0.5 1.0 1.5 2.0 2.5
4.5 6.0 7.5
2.0 3.0 4.0
1 3 5 7
0.5 1.5 2.5
4.5 5.5 6.5 7.5
1 2 3 4 5 6 7


Given : Species
virginica
versicolor
setosa
setigera
versicolor
virginica
Sepal.Length
Sepal.Width
2.0
2.5
3.0
3.5
4.0
4.5
5.5
6.5
7.5


Iris sepal width and length
setigera
versicolor
virginica
Sepal.Length
4.5
5.5
6.5
7.5
2.0
2.5
3.0
3.5
4.0
Sepal.Width



Given : Petal.Length
1
2
3
4
5
6
setigera
versicolor
virginica
Sepal.Length
4.5
5.5
6.5
7.5
2.0
3.0
4.0
Sepal.Width

4-diemensional slice plots

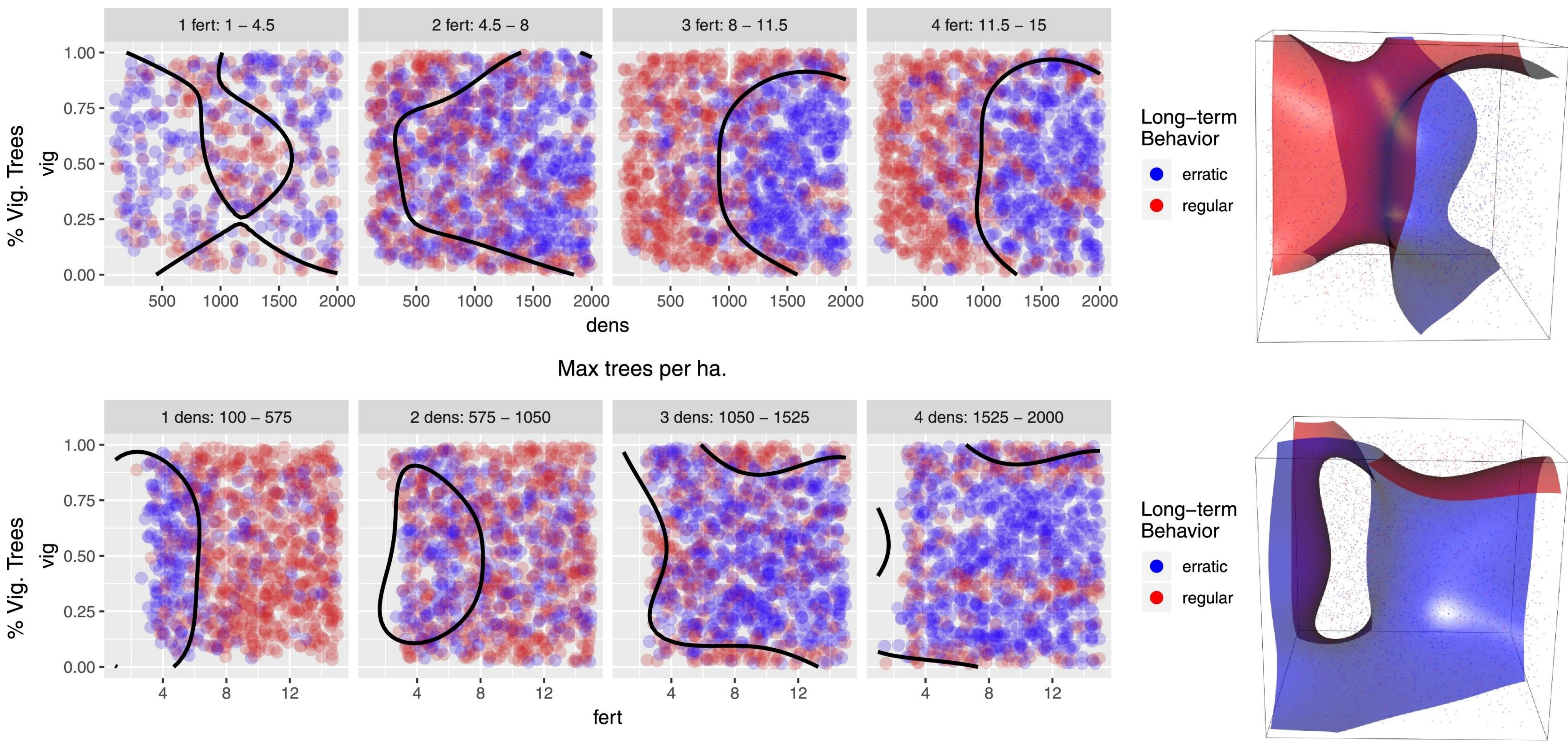
3 model continuous parameters: predictors on x, y, z axes

- Tree vigor
- Beetle fertility
- Tree density

1 binary response variable: Epidemic return interval

- Red = regular intervals, ca. 80 years
- Blue = erratic return intervals

Support vector machine decision boundary


1 fert: 1 – 4.5
2 fert: 4.5 – 8
3 fert: 8 – 11.5
4 fert: 11.5 – 15
% Vig. Trees
vig
1.00
0.75
0.50
0.25
0.00
500
1000
1500
2000
dens
Max trees per ha.
Long–term Behavior
erratic
regular
1 dens: 100 – 575
2 dens: 575 – 1050
3 dens: 1050 – 1525
4 dens: 1525 – 2000
4
8
12
fert

4-diemensional slice plots

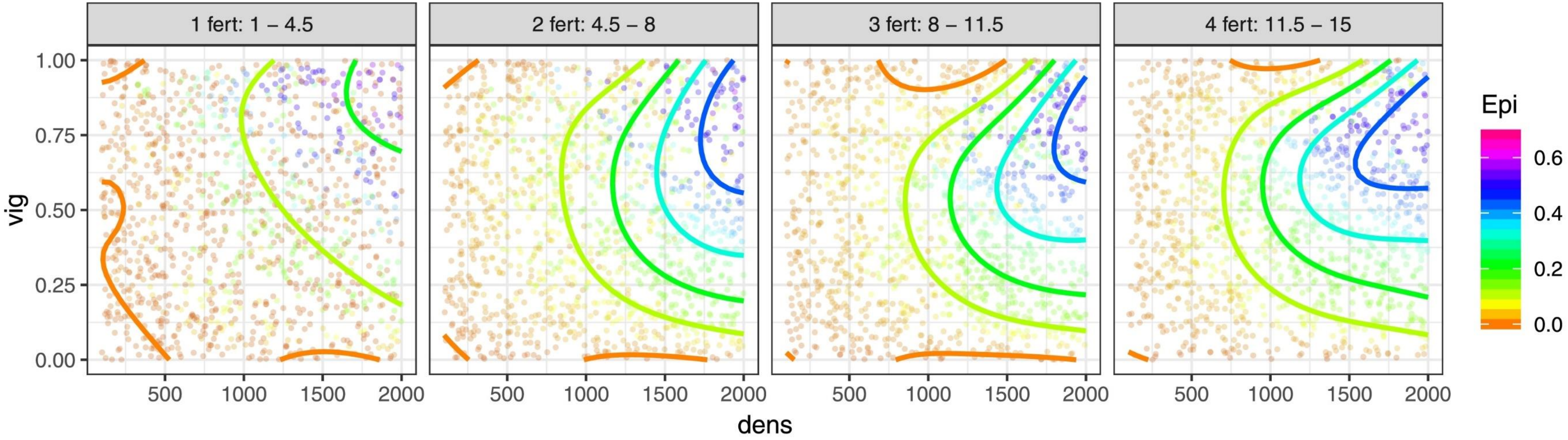
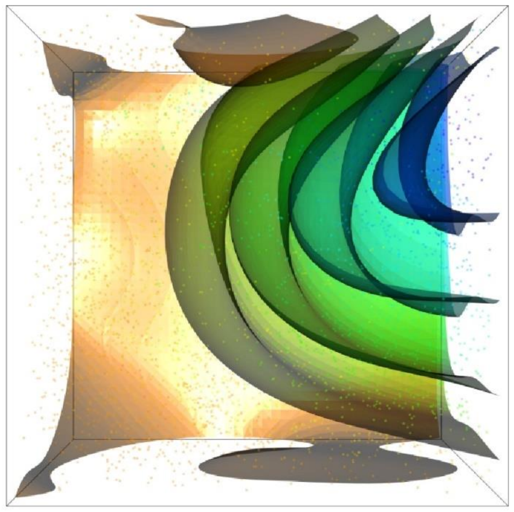
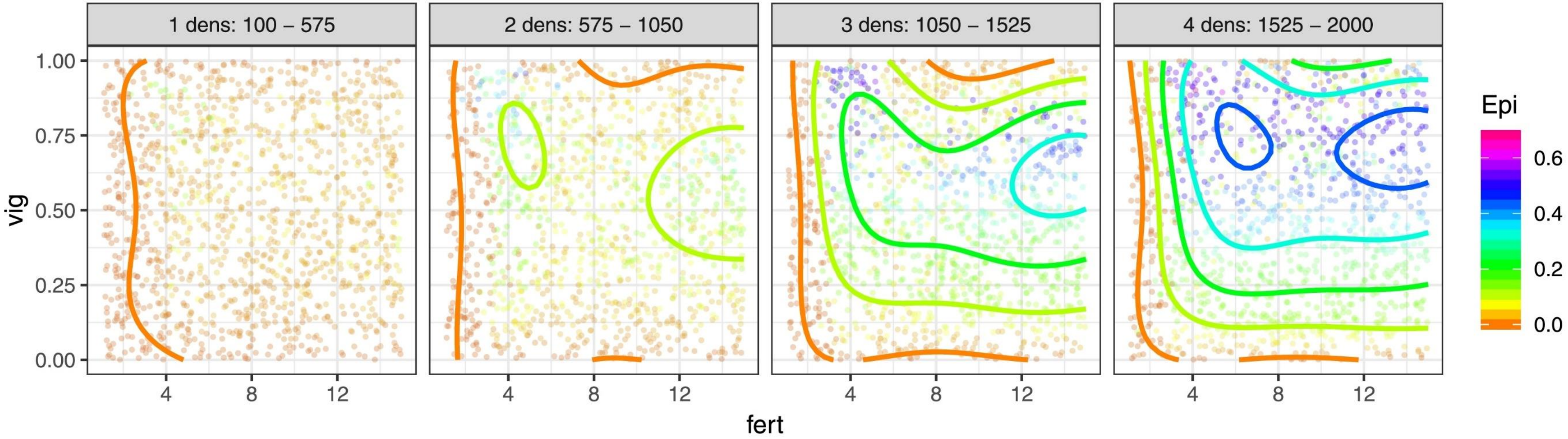
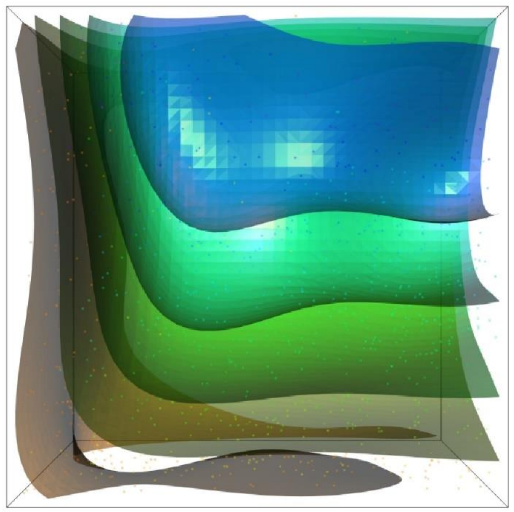
3 model continuous parameters: predictors on x, y, z axes

- Tree vigor
- Beetle fertility
- Tree density

1 continuous response variable: Epidemic index

- Unitless, 0 – 1 range

Support vector machine regression


1 fert: 1 – 4.5
2 fert: 4.5 – 8
3 fert: 8 – 11.5
4 fert: 11.5 – 15
vig
1.00
0.75
0.50
0.25
0.00
500
1000
1500
2000
dens
Epi
0.6
0.4
0.2
0.0


1 dens: 100 – 575
2 dens: 575 – 1050
3 dens: 1050 – 1525
4 dens: 1525 – 2000
vig
1.00
0.75
0.50
0.25
0.00
4
8
12
fert
Epi
0.6
0.4
0.2
0.0


Variable sufficiency

For categorical variables: if very few sampling units fall into a category, that category may not be sufficiently represented in the sample.

Problems with detection limits, sampling error, etc.

Key take-home: Inference on low-frequency events is difficult

Data transformations

Usually we want to make a nonlinear relationship linear.

Variance stabilization, we'll see why this is desirable later.

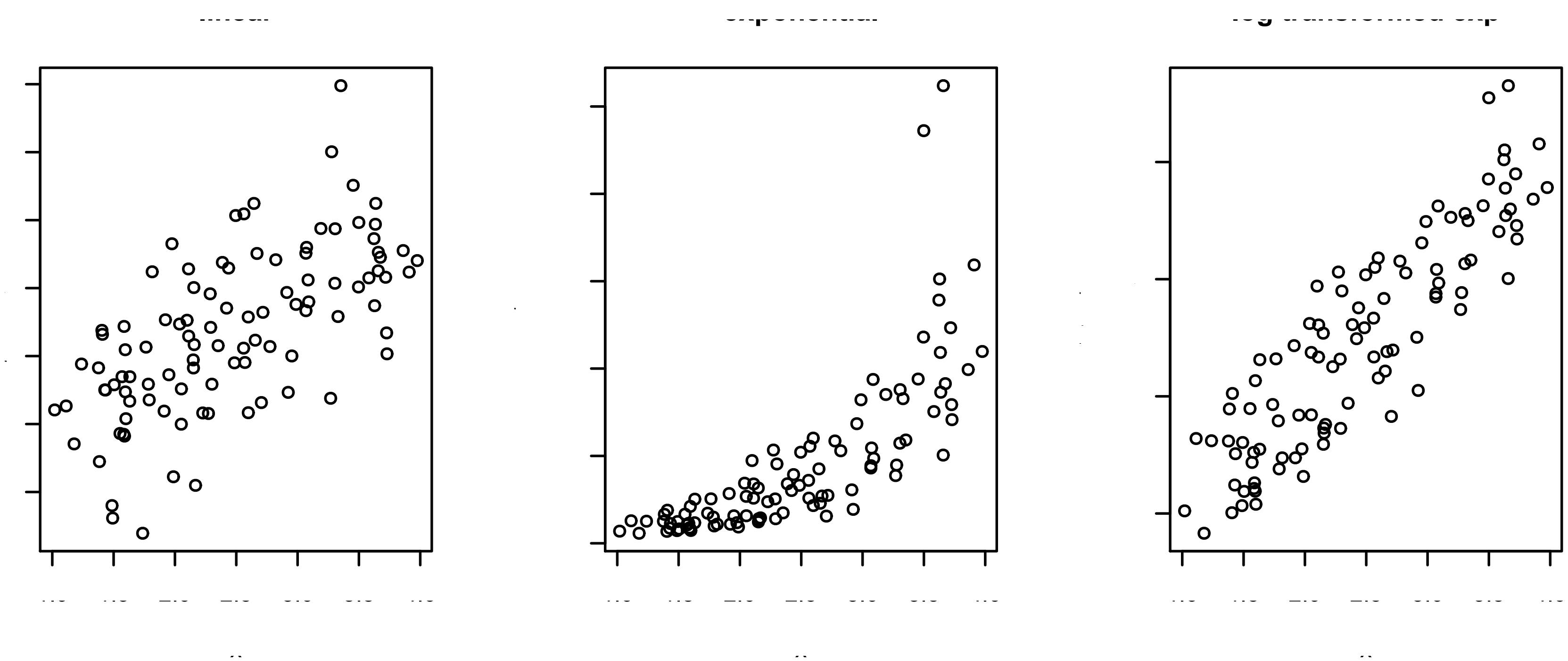
Linear relationships are analytically simple.

Linear relationships are easy to interpret.

Interpretation of transformed variables can be difficult.

We'll discuss standardizations later.

Log transformations



Extreme values

Extreme values may be due to error:

- Measurement
- Data entry
- Transcription

Extreme values may be real

Extreme values can violate inference assumptions.

We'll revisit extreme values many times.

Assignment 2

In groups, choose a paper from the abstract examples from last Thursday's in-class activity.

Assignment instructions and papers are on Moodle.

Model Thinking

How do we describe a linear association verbally?

What are the essential components of a conceptual model of a linear association?

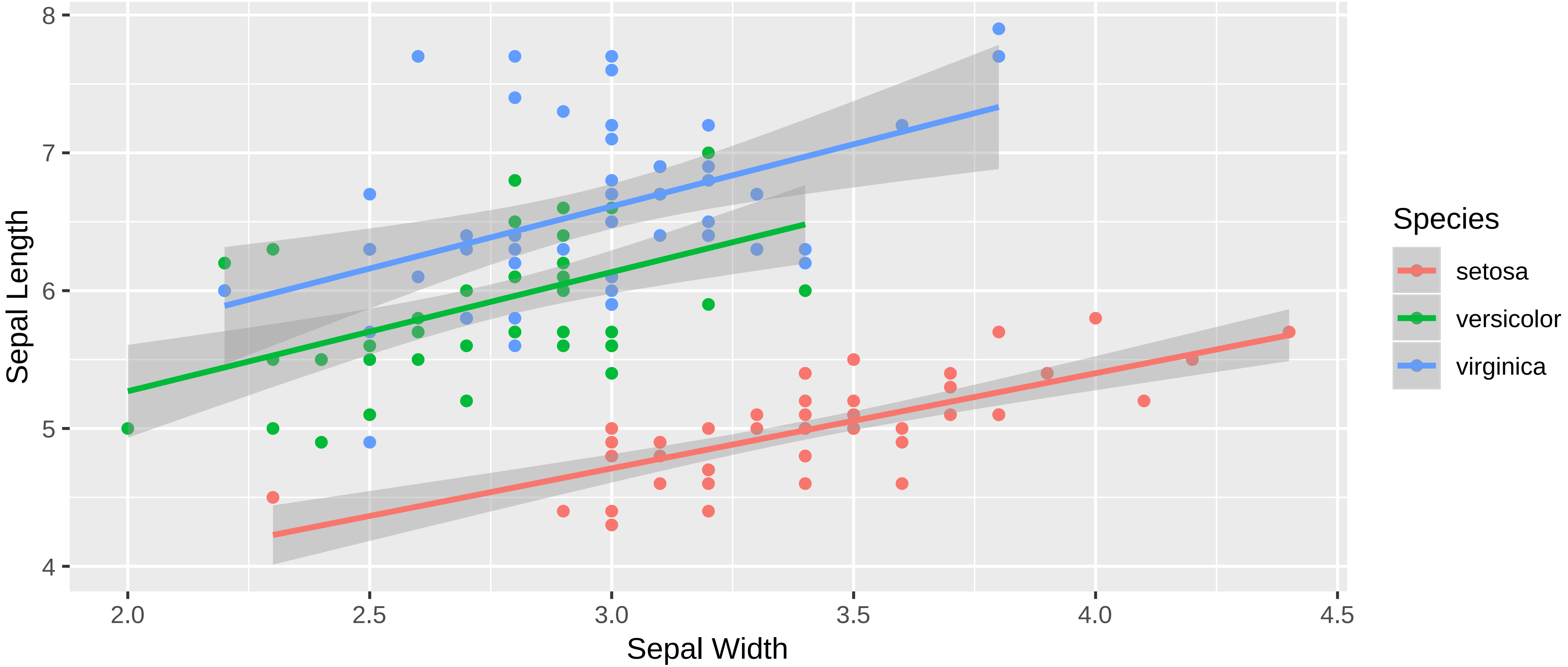
Model Thinking: Dual Data Model

We will fit 2 models to our data:

Deterministic model

Stochastic model

Model Thinking: 2 models



Model Thinking: 2 models

Deterministic **functions**

Probability **distributions**

What is a deterministic function?

Possible inputs? Domain

Possible outputs? Range

Deterministic Models

Phenomenological

Mechanistic

Mechanistic model building

1. Describe system context and research questions
   1. Identify possible components to include in the model
   2. Make list of candidate predictors and response
2. Hypothesize some candidate deterministic models
   1. Always start with a linear model!
   2. Describe what model terms to include, but don’t worry about exact mathematical form at first.

Deterministic model exercise

Scenario: Invasive birds introduced to an archipelago

Populations reproduce rapidly, can migrate to nearby islands

Eradication is possible if the population less than 50% of island's capacity

Question: Can we make a model of an island's bird population to help managers decide which islands are worth allocating limited resources to?

Mechanistic model building

1. Describe system context and research questions
   - Identify possible components to include in the model
   - Make list of candidate predictors and response

Key terms for deterministic functions

1. Monotonic
2. Continuous
3. Smooth
4. Asymptotic
5. Linear/nonlinear
6. Parameters/variables

Mechanistic model building

2. Propose candidate deterministic models
   - Always start with a linear model!
   - Describe what model terms to include, but don’t worry about exact mathematical form at first.

Deterministic model groups

1. Scenario: hare and lynx population cycles
   1. Shamelessly stolen from one of your groups.
   2. Don’t let previous knowledge of Lotka-Volterra models bias your model thinking.
2. Objectives
   1. Conceptual model of population dynamics with list of predictors, responses
   2. Candidate deterministic functions

For next time:

McGarigal Chapter 4: Keep slogging through but don’t get stuck on details.

We’ll discuss graphical intuition about the important families of deterministic functions.

We’ll discuss how to look at function equations, and hopefully develop a little bit of intuition.