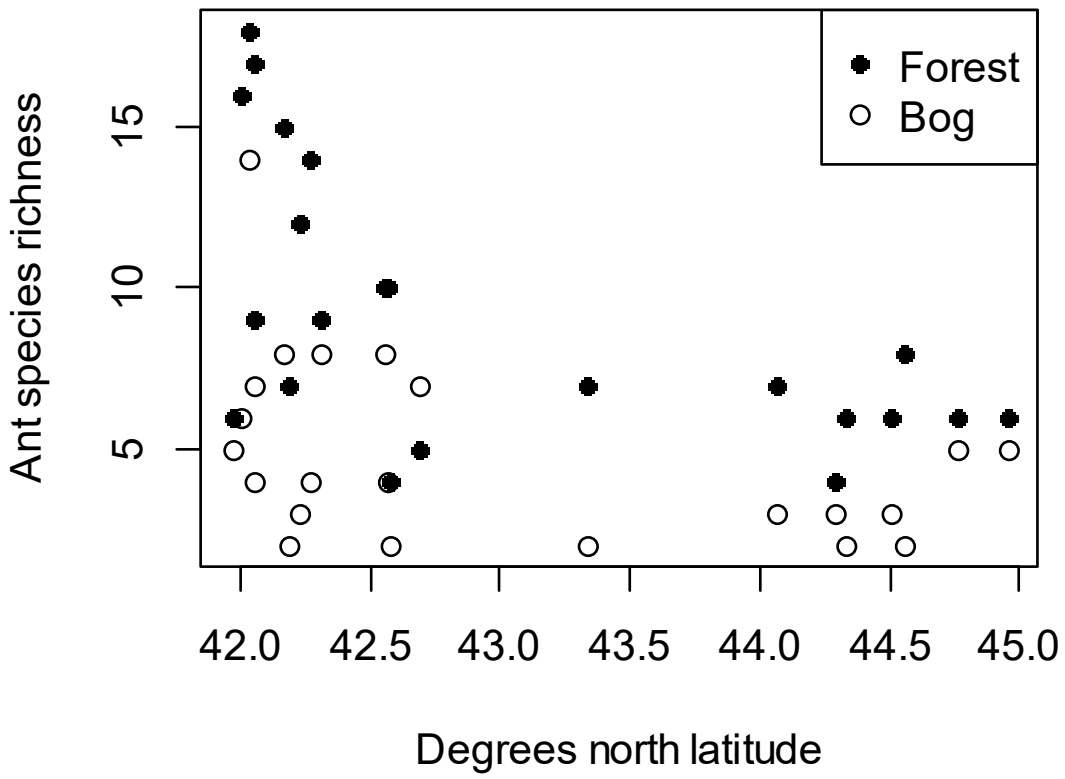
# Dataset to analyze




What will the data-generating model be?
Ignore pairs for now

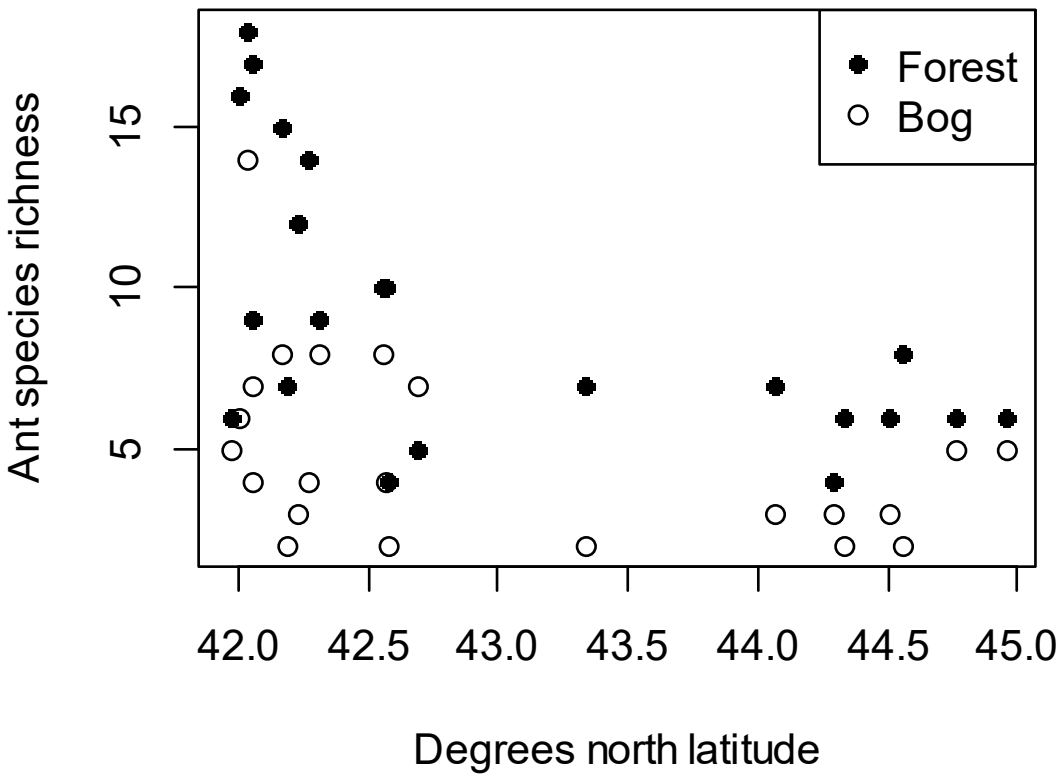
Scientific questions:

How different is species richness between habitats?

How does species richness vary with latitude?

Is this relationship different between habitats?

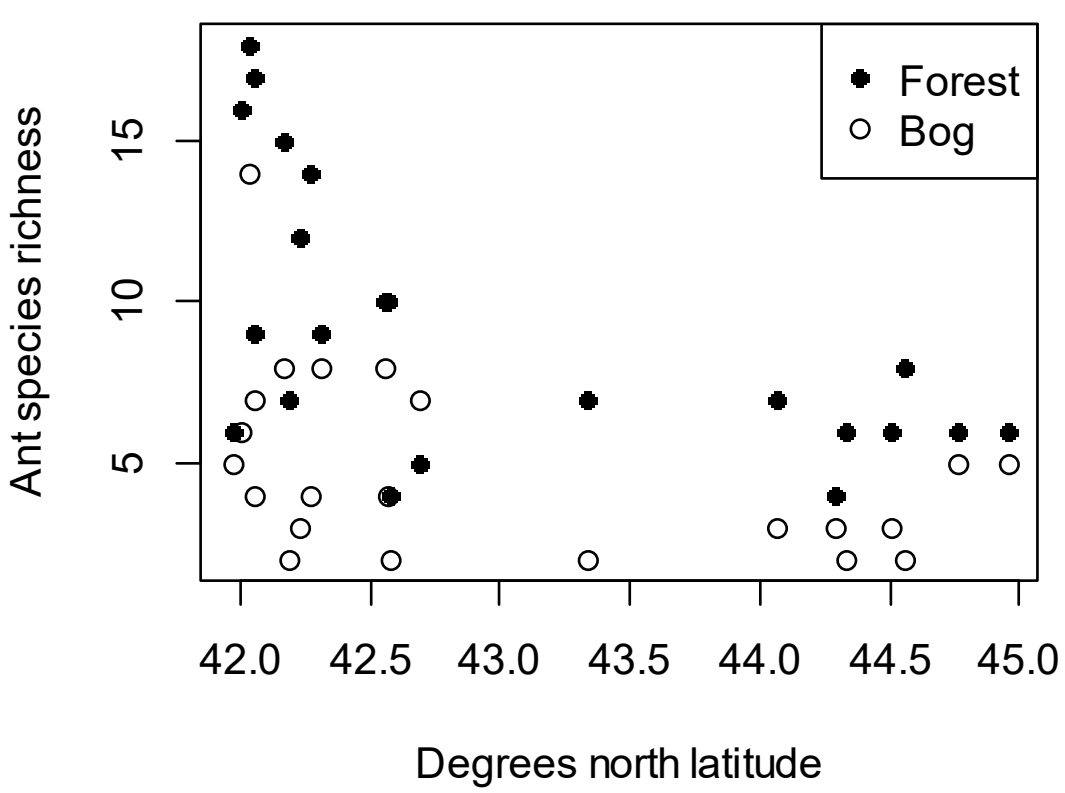
# Dataset to analyze




Model:

GLM
Poisson
Log link

# Dataset to analyze




Model:

Distribution

$$y_i \sim \text{Poisson}(\mu_i)$$

Link function

$$\log(\mu_i) = \beta_0 + \beta_1 forest_i + \beta_2 latitude_i + \beta_3 forest_i \times latitude_i$$

Inverse link function

$$\mu_i = e^{\eta_i}$$

# Model matrix

$$\eta_i = \beta_0 + \beta_1 forest_i + \beta_2 latitude_i + \beta_3 forest_i \times latitude_i$$

```
fit <- glm(richness ~ habitat + latitude + habitat:latitude,
           family=poisson(link="log"), data=ant)
```

## Data

| habitat | latitude | richness |
|---|---|---|
| forest | 42 | 16 |
| forest | 42.56 | 10 |
| forest | 43.33 | 7 |
| forest | 44.76 | 6 |
| bog | 42.17 | 8 |
| bog | 42.57 | 4 |
| bog | 44.06 | 3 |
| bog | 44.95 | 5 |

## Design matrix `model.matrix(fit)`

| intercept | forest | latitude | forest:latitude |
|---|---|---|---|
| 1 | 1 | 42 | 42 |
| 1 | 1 | 42.56 | 42.56 |
| 1 | 1 | 43.33 | 43.33 |
| 1 | 1 | 44.76 | 44.76 |
| 1 | 0 | 42.17 | 0 |
| 1 | 0 | 42.57 | 0 |
| 1 | 0 | 44.06 | 0 |
| 1 | 0 | 44.95 | 0 |

$$\eta_i = \beta_0 intercept_i + \beta_1 forest_i + \beta_2 latitude_i + \beta_3 forest_i \times latitude_i$$