Model Basics

READING:	0.1, 0.2
EXERCISES:	ALL CHAPTER 0
ASSIGNED:	HW 1
PRODUCER:	DR. MARIO

IMG CREDIT: ALEX RIEGE

Course Website / Syllabus

- Access Course Website Through Canvas
- Cover Syllabus
  - Office Hours
  - Grading and Curving
  - Attendance: UNC Check-In App
  - Homework
  - Quizzes
  - Exams
  - PDFs and Gradescope
  - Grade Disputes
  - Honor Code
- Usage of Course Website and Canvas

Preview of Dataset

LEGO Dataset (n=1304)

Set_Name	Theme	Pieces	Price	Amazon_Price	Year	Ages	Minifigures	Unique_Pieces
<chr>	<chr>	<dbl>	<dbl>	<dbl>	<dbl>	<chr>	<dbl>	<dbl>
Imperial Star Destroyer	Star Wars	4784	700.	700.	2019	Ages_16+	2	445
Betrayal at Cloud City	Star Wars	2812	350.	668.	2018	Ages_14+	19	676
Liebherr R 9800	Powered UP	4108	450.	443.	2019	Ages_12+	NA	221
NINJAGO City Docks	NINJAGO	3553	230.	440	2018	Ages_12+	14	690
Hogwarts Castle	Harry Potter	6020	400.	400.	2018	Ages_16+	28	624
Voltron	Ideas	2321	180.	389.	2018	Ages_16+	NA	471
Roller Coaster	Creator Expert	4124	380.	380.	2018	Ages_16+	11	556
Bugatti Chiron	Technic	3599	350.	340.	2018	Ages_16+	NA	306
Mack Anthem	Technic	2595	180.	330	2018	Ages_11-16	NA	253
Farm Adventures	DUPLO	104	60.0	300.	2018	Ages_2-5	3	42

Any interesting questions about LEGO we may want to answer?

Models Help Us…

- Answer Questions
- Make Predictions or Classifications
- Evaluate Treatments or Test Theories
- Understand Relationships

Architecture of a Model

$$Y = f(X) + \varepsilon$$

Response Variable

Explanatory Variable(s)

Error or Deviation from the Model

Generates an Expectation about Y given X

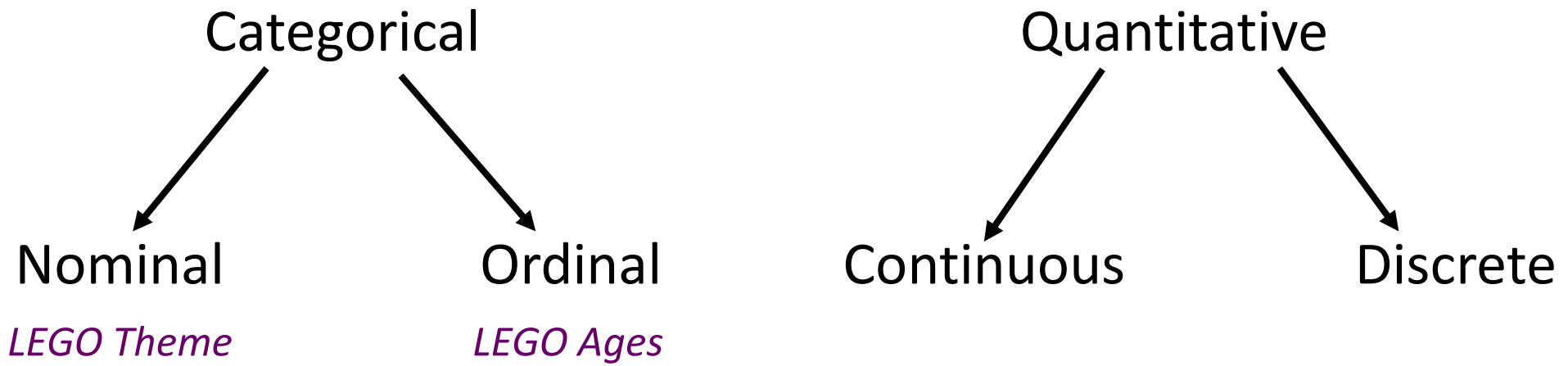
Statistical Modeling

Statistical Modeling is the Process of …

Defining the Function *f(X)* and then **Fitting that Function *f(X)***
to a sample dataset by **Minimizing the Error** the best we possibly can

Methodology we Use Depends on the Types of Variables

Variable Types

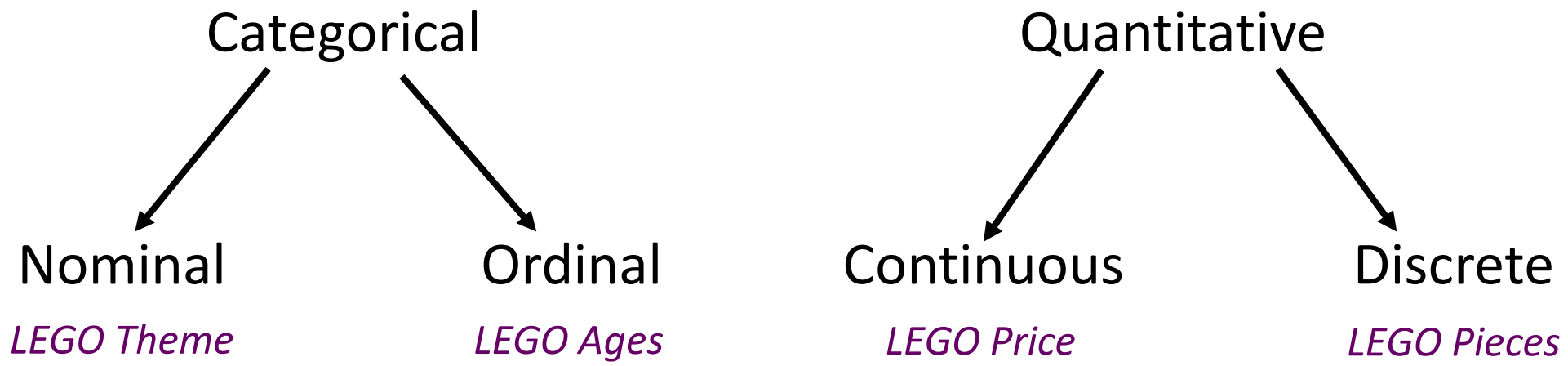


Preview of Dataset

Can we find all the variable types in this dataset?

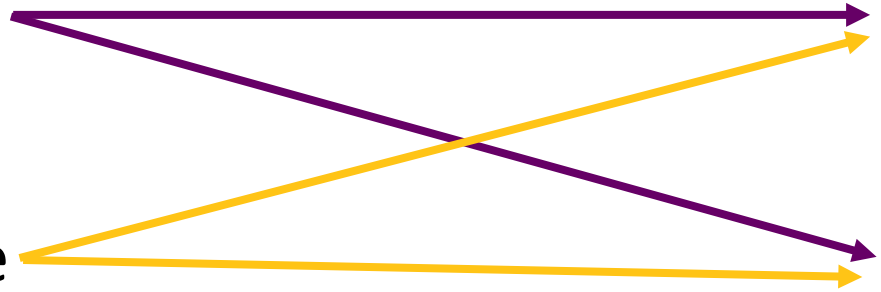
Set_Name	Theme	Pieces	Price	Amazon_Price	Year	Ages	Minifigures	Unique_Pieces
<chr>	<chr>	<dbl>	<dbl>	<dbl>	<dbl>	<chr>	<dbl>	<dbl>
Imperial Star Destroyer	Star Wars	4784	700.	700.	2019	Ages_16+	2	445
Betrayal at Cloud City	Star Wars	2812	350.	668.	2018	Ages_14+	19	676
Liebherr R 9800	Powered UP	4108	450.	443.	2019	Ages_12+	NA	221
NINJAGO City Docks	NINJAGO	3553	230.	440	2018	Ages_12+	14	690
Hogwarts Castle	Harry Potter	6020	400.	400.	2018	Ages_16+	28	624
Voltron	Ideas	2321	180.	389.	2018	Ages_16+	NA	471
Roller Coaster	Creator Expert	4124	380.	380.	2018	Ages_16+	11	556
Bugatti Chiron	Technic	3599	350.	340.	2018	Ages_16+	NA	306
Mack Anthem	Technic	2595	180.	330	2018	Ages_11-16	NA	253
Farm Adventures	DUPLO	104	60.0	300.	2018	Ages_2-5	3	42

Example of Variable Types Using LEGO



Families of Models

Response Variable	Predictor Variable
Categorical	Categorical
Quantitative	Quantitative

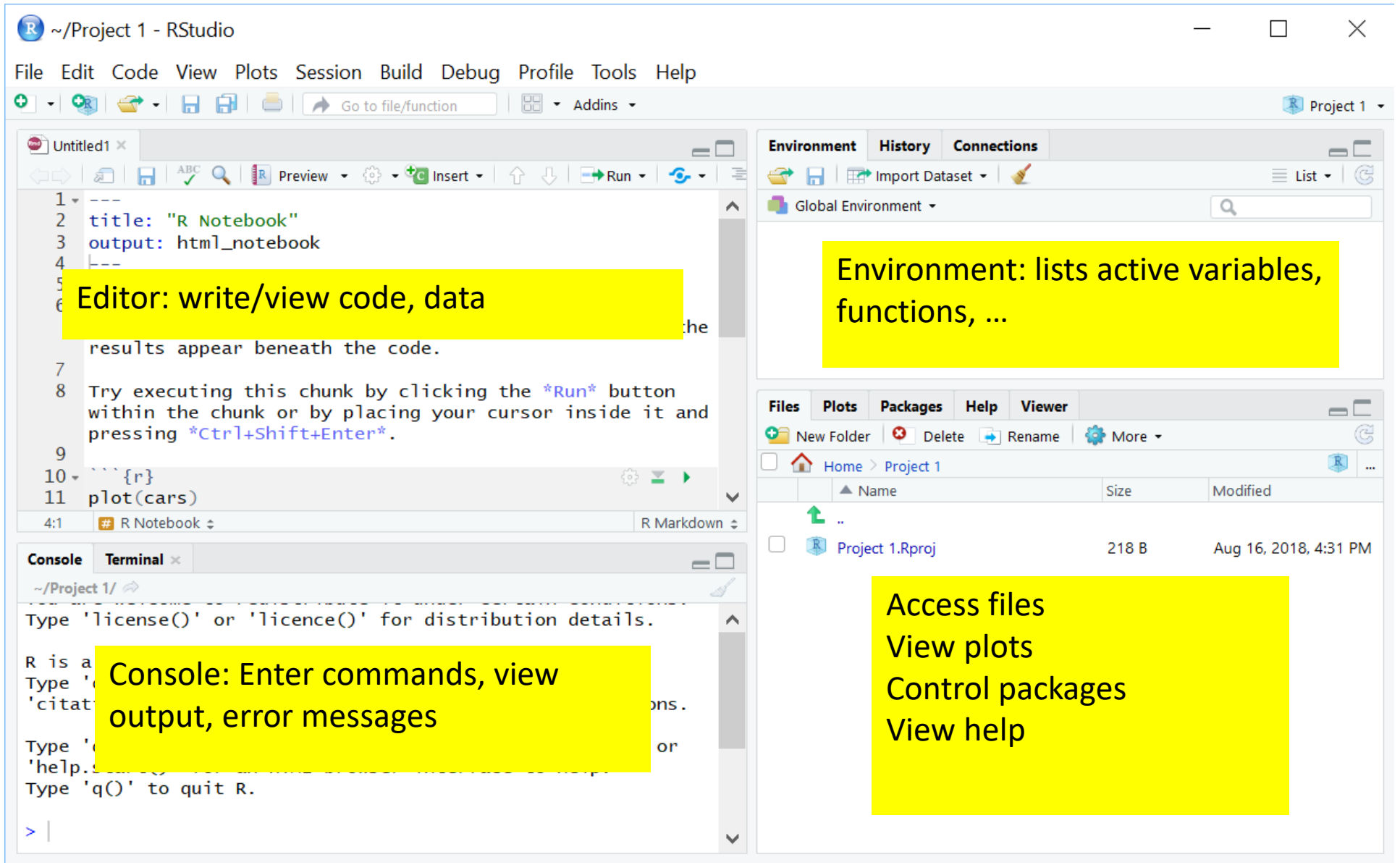


*Complicated by **multiple** predictor variables and/or response variables*

Technology We Will Use

- *R* = a free, widely used, open source, language and environment for statistical computing and graphics
- *RStudio* = an interface for *R* (Integrated Development Environment)
- *RMarkdown* = a tool in R for creating documents that combine R code with text

RStudio


~/Project 1 - RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function
Addins
Project 1
Untitled1
Preview
Insert
Run
---
title: "R Notebook"
output: html_notebook

Editor: write/view code, data
results appear beneath the code.
Try executing this chunk by clicking the *Run* button
within the chunk or by placing your cursor inside it and
pressing *Ctrl+Shift+Enter*.
```{r}
plot(cars)
4:1
R Notebook
R Markdown
Console
Terminal
~/Project 1/
Type 'license()' or 'licence()' for distribution details.
Console: Enter commands, view output, error messages
Type 'q()' to quit R.
Environment
History
Connections
Import Dataset
List
Global Environment
Environment: lists active variables, functions, …
Files
Plots
Packages
Help
Viewer
New Folder
Delete
Rename
More
Home
Project 1
Name
Size
Modified
Project 1.Rproj
218 B
Aug 16, 2018, 4:31 PM
Access files
View plots
Control packages
View help
```