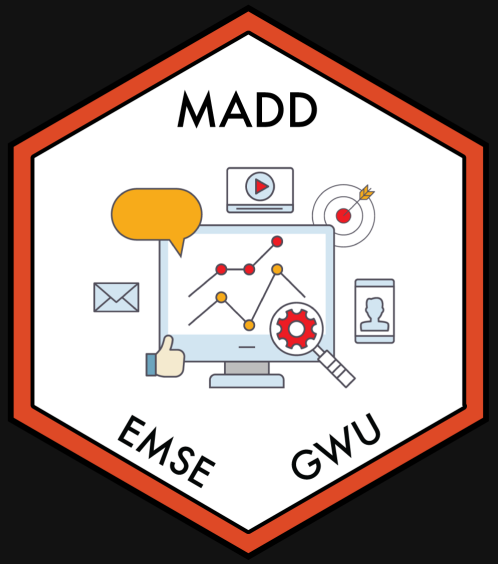


Week 1: *Getting Started*

EMSE 6035: Marketing Analytics for Design Decisions

John Paul Helveston

September 01, 2021

Week 1: *Getting Started*

1. Course orientation

2. Intro to conjoint analysis

3. Introductions

BREAK: Teaming

4. Getting started with R & RStudio

Week 1: *Getting Started*

1. Course orientation

2. Intro to conjoint analysis

3. Introductions

BREAK: Teaming

4. Getting started with R & RStudio

Meet your instructor!

John Paul Helveston, Ph.D.

Assistant Professor, Engineering Management & Systems Engineering

- 2016-2018 Postdoc at Institute for Sustainable Energy, Boston University
- 2016 PhD in Engineering & Public Policy at Carnegie Mellon University
- 2015 MS in Engineering & Public Policy at Carnegie Mellon University
- 2010 BS in Engineering Science & Mechanics at Virginia Tech
- Website: www.jhelvy.com

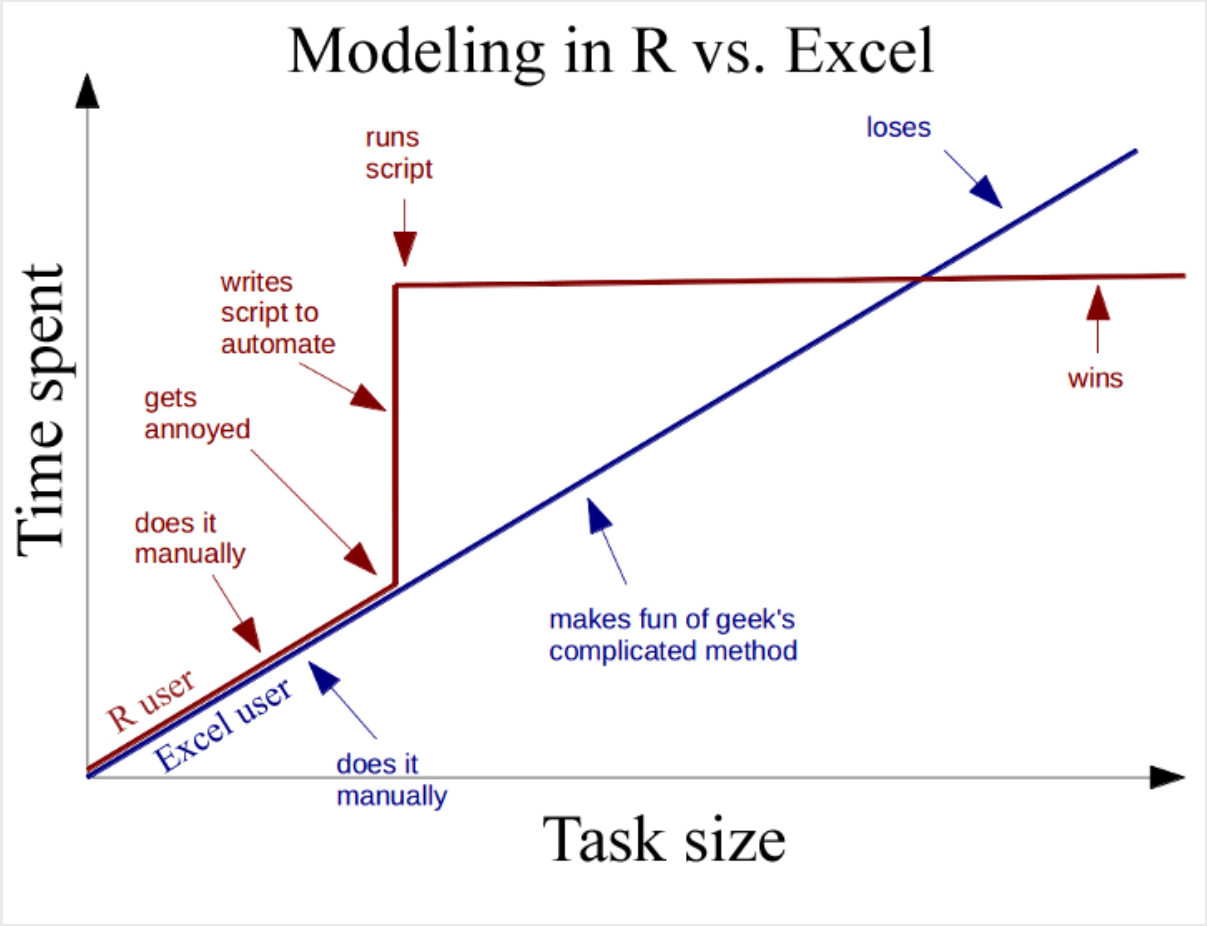
Tools

Course website: https://madd.seas.gwu.edu/2021-Fall/

Course slack: https://emse-madd-f21.slack.com

R & RStudio: installation instructions

Why R?


Modeling in R vs. Excel
Time spent
Task size
runs script
loses
writes script to automate
wins
gets annoyed
does it manually
makes fun of geek's complicated method
R user
Excel user
does it manually

Learning Objectives

After this class, you will know how to...

- ...work with data in R
- ...design effective surveys to get rich data
- ...analyze consumer choice data to model consumer preferences
- ...design effective charts to communicate insights

Course prerequisites

This course requires prior exposure to:

- Probability theory
- Multivariable calculus
- Linear algebra
- Regression

Not sure?

Take this self assessment

Reflections (27% of grade)

Do some readings, recorded lectures, practice problems

Write a short reflection

📅 ~Every week (9 total)

🕒 Due 11:59pm Tues. before class

✔ Graded for completion (looking for engagement)

Quizzes (12% of grade)

In class every other week-ish (5 total, lowest dropped)

5 minutes (3-5 questions)

Example quiz

Why quiz at all? There's a phenomenon called the "retrieval effect" - basically, you have to *practice* remembering things, otherwise your brain won't remember them (details in the book "Make It Stick: The Science of Successful Learning").

Exam (10% of grade)

Take home exam, 2nd to last week of class

We'll go over exam solutions on last day of class

Semester Project (51% of grade)

Teams of 3-4 students

Goals:

- Assess market viability of a new technology or design
- Recommend best design choices for target market or application

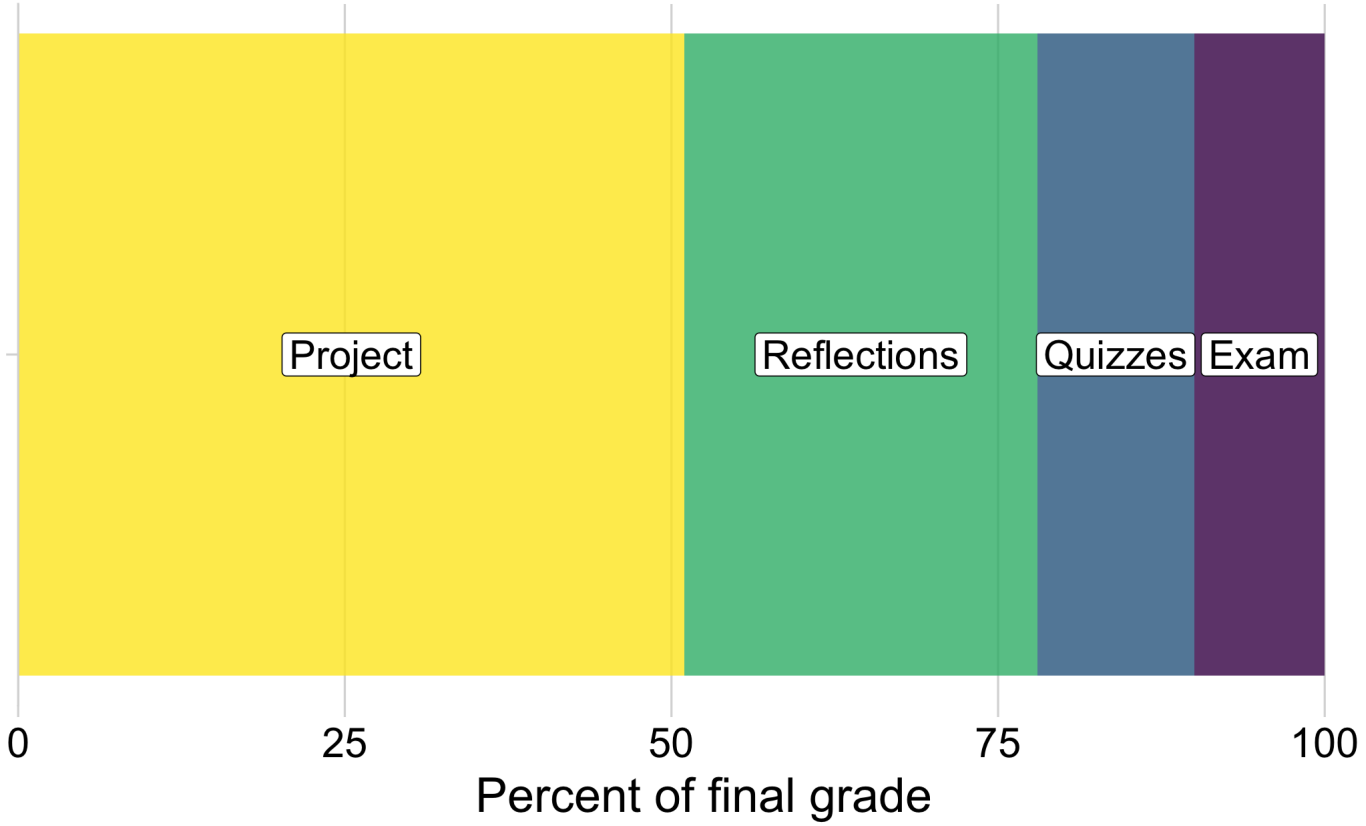
Key deliverables:

Item	Weight	Due
Proposal	7 %	9/26
Survey Plan	4 %	10/05
Pilot Survey	4 %	10/15
Pilot Analysis	9 %	11/07
Final Survey	5 %	11/21
Final Analysis Report	14 %	12/13
Final Presentation	8 %	12/15

Grades

Item	Weight	Notes
Reflections	27 %	Weekly assignment (9 x 3%)
Quizzes	12 %	5 quizzes, lowest dropped
Project Proposal	7 %	Teams of 3-4 students
Survey Plan	4 %	
Pilot Survey	4 %	
Pilot Analysis	9 %	
Final Survey	5 %	
Final Analysis Report	14 %	
Final Presentation	8 %	
Final Exam	10 %	Take home exam

Grades


Project
Reflections
Quizzes
Exam
0
25
50
75
100
Percent of final grade

Course policies

- BE NICE
- BE HONEST
- DON'T CHEAT

Copying is good, stealing is bad

> "Plagiarism is trying to pass someone else's work off as your own. Copying is about reverse-engineering."
>
> -- Austin Kleon, from Steal Like An Artist

Late submissions

- **5** late days - use them anytime, no questions asked
- No more than **2** late days on any one assignment
- Contact me for special cases

How to succeed in this class

- Participate during class!
- Start assignments early and **read carefully**!
- Get sleep and take breaks often!
- Ask for help!

Getting Help

Use Slack to ask questions.

Schedule a meeting w/Prof. Helveston:

- Mondays from 8:00-5:00pm
- Tuesday from 1:00-5:00pm
- Thursdays from 12:00-5:00pm

GW Coders

Week 1: *Getting Started*

1. Course orientation

2. Intro to conjoint analysis

3. Introductions

BREAK: Teaming

4. Getting started with R & RStudio

We want to answers to questions like...

- Higher prices decrease demand, but by how much?

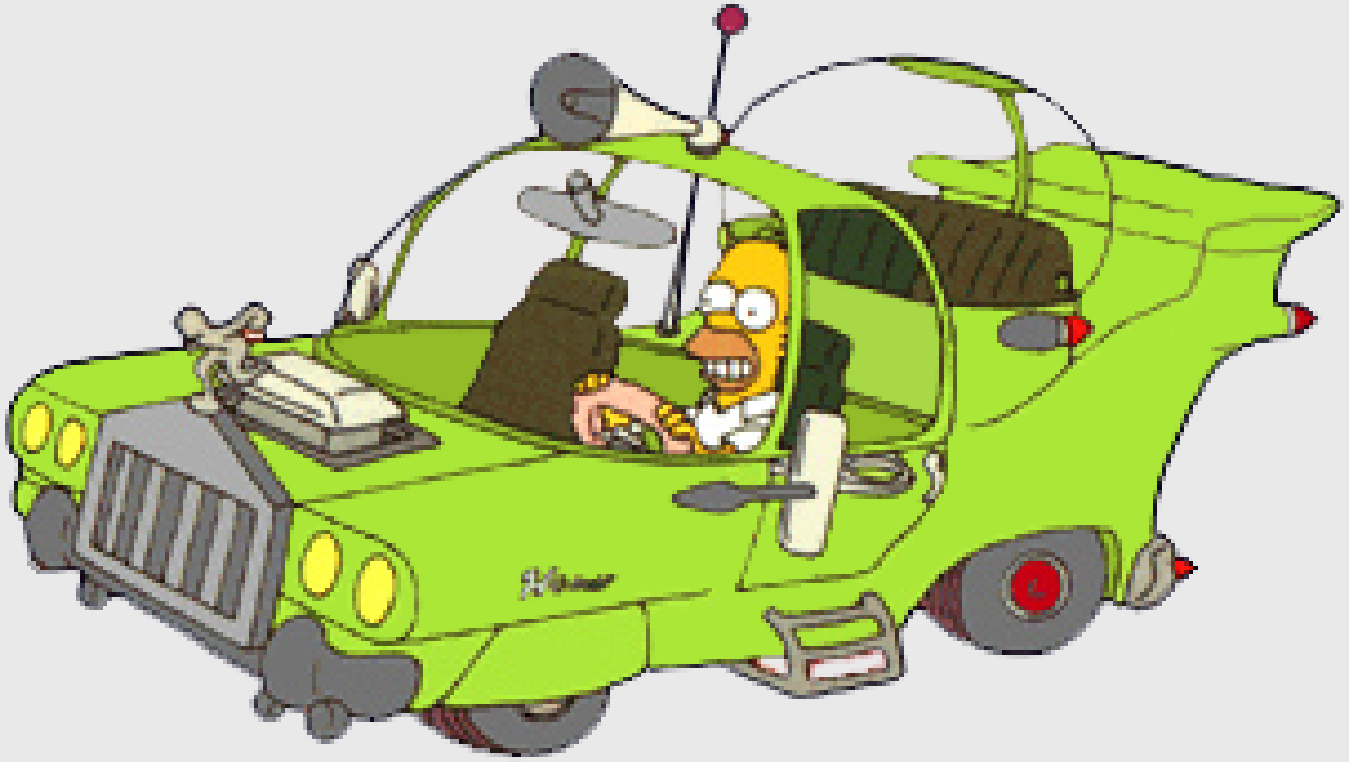
- How much more is a consumer willing to pay for increased performance in X?

- How will my product compete against competitors in the market?

Answers depend on knowing what people want

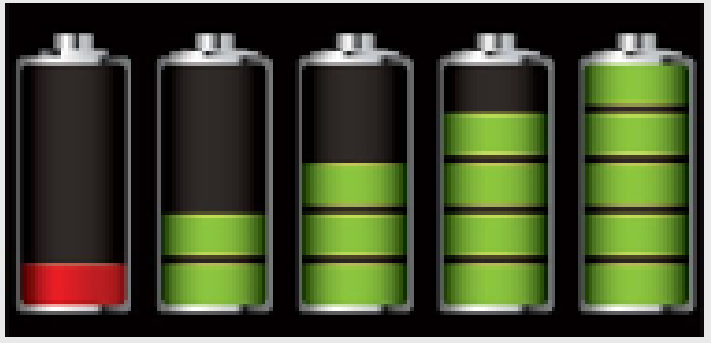
Directly asking people what they want isn't always helpful

(People want everything)



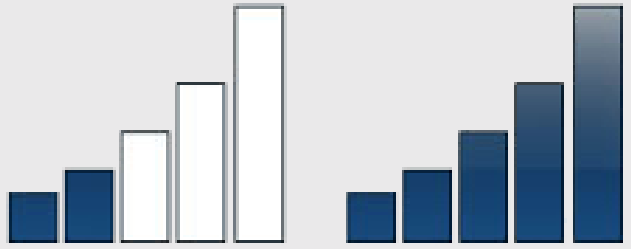
Which feature do you care more about?

Battery Life?

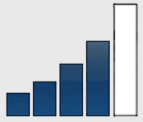
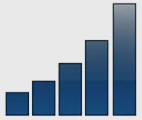
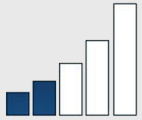


Brand?

Signal quality?



Conjoint approach:
Use consumer choice data to model preferences

Attribute	Phone 1	Phone 2	Phone 3
Price	$400	$450	$350
Brand			
Battery Life			
Signal Quality			
N chosen:	350	250	400

Use random utility framework to predict probability of choosing phone *j*

1. $u_j = \beta_1 \mathrm{price}_j + \beta_2 \mathrm{brand}_j + \beta_3 \mathrm{battery}_j + \beta_4 \mathrm{signal}_j + \varepsilon_j$

2. Assume $\varepsilon_j \sim$ iid extreme value

3. Probability of choosing phone *j*: $P_j = \frac{e^{\beta' x_j}}{\sum_k^J e^{\beta' x_k}}$

4. Estimate $\beta_1, \beta_2, \beta_3, \beta_4$ by minimizing $-L = -\sum_n^N \sum_j^J y_{nj} \ln P_{nj}$

Willingness to Pay

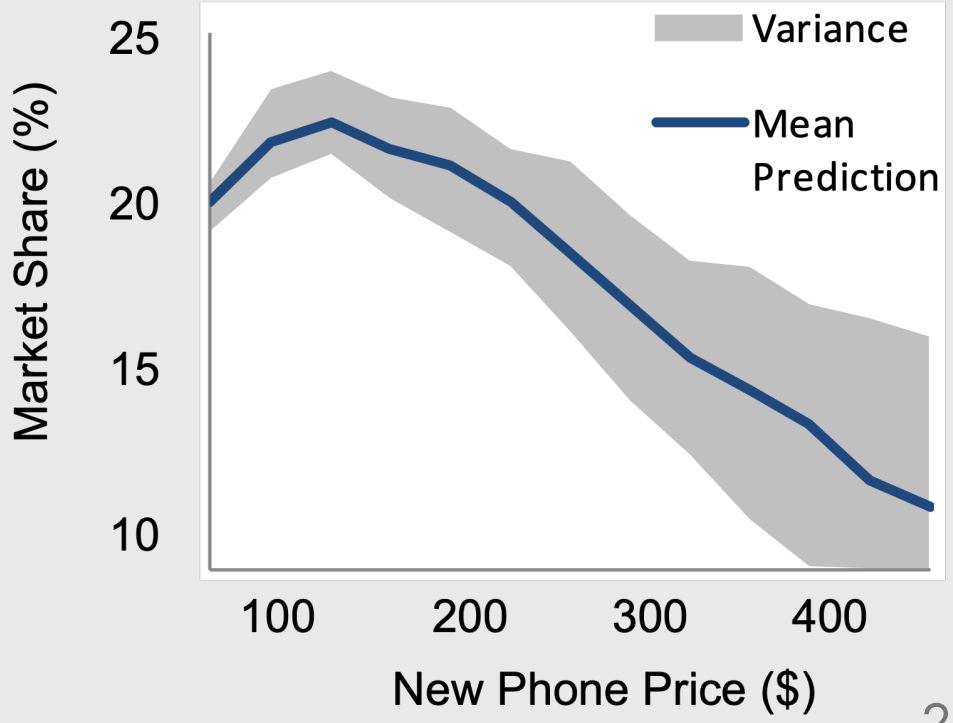
$$u_j = \beta' x_j + \alpha p_j + \varepsilon_j$$

$$\omega = \frac{\beta}{-\alpha}$$

"Respondents on average are willing to pay $XX to improve battery life by XX%"

Make predictions

$$P_j = \frac{e^{\hat{\beta}' x_j}}{\sum_k^J e^{\hat{\beta}' x_k}}$$



Example: *Pocket Charge*

A Flexible, Portable Solar Charger

Product Diagram

Durability

Degradation Protections

Attribute Units

Price– USD
Weight– Kg
Power Output – Watts
Durability – Months
Portability – LxWxH

Power Output

Power Density

Packing Design

Decision Variable Units

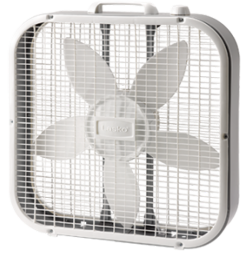
Power Density – W/Kg
Degradation Rate – Hours
Packing Design – Cm^3

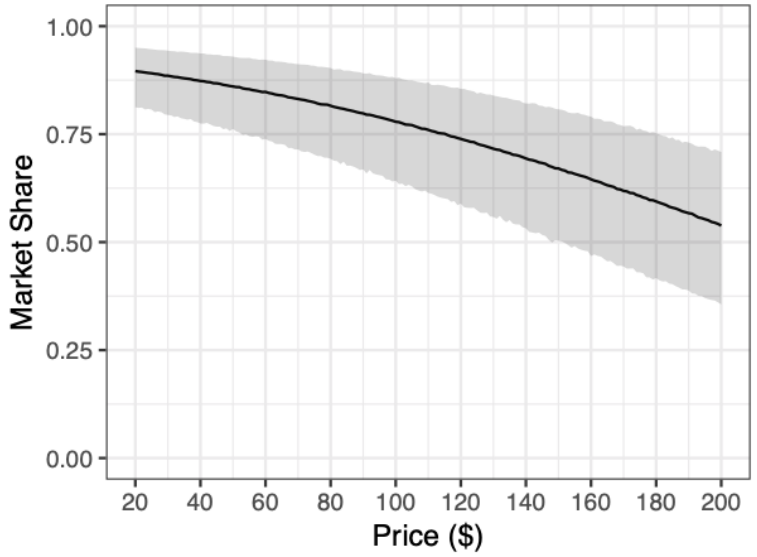
Weight

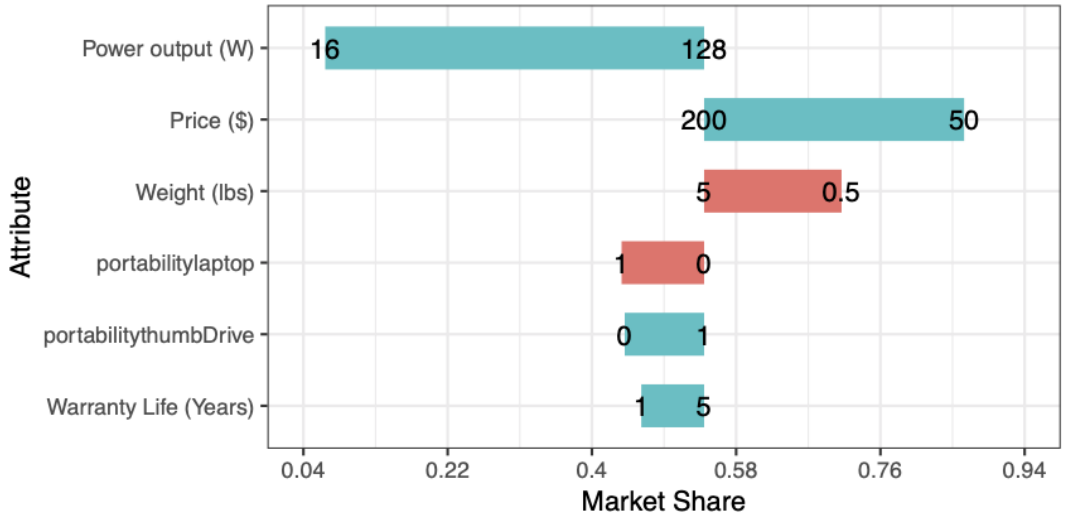
Image Sources:
1.https://www.deviantart.com/kota3214/art/Rainy-Sun-244479413
2.https://www.pennscale.com/products/ntep-certified-scales/label-printing-
3. https://www.thiswhyimbro
foldable-solar-panel-charger/

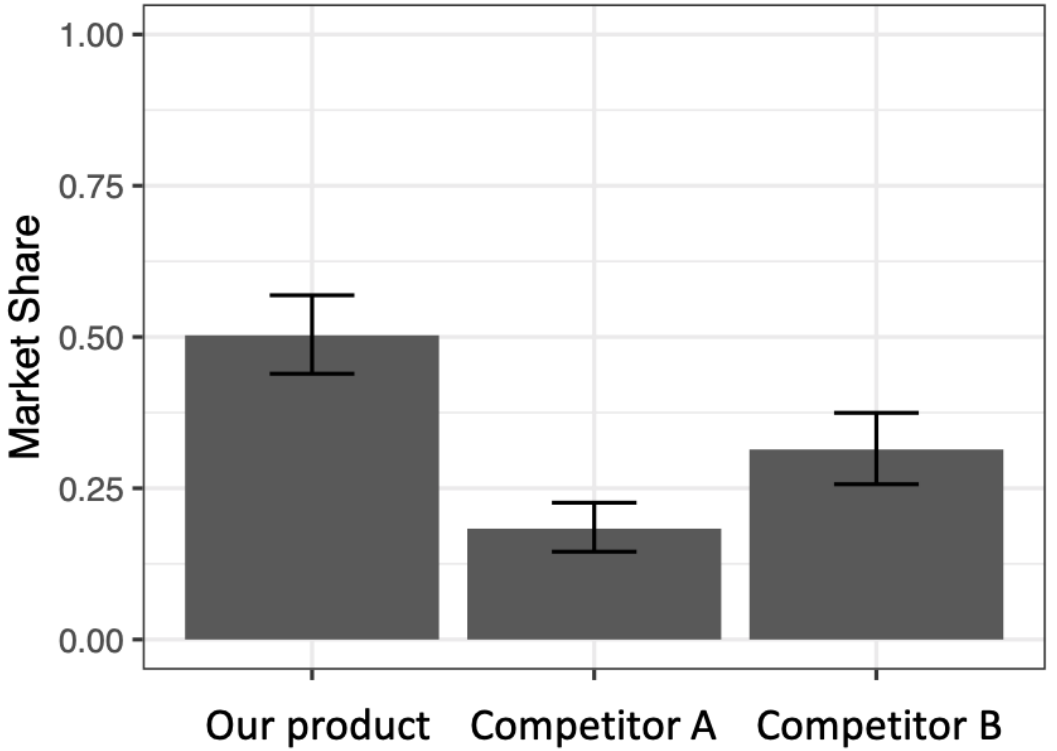
Example survey choice question

	Choice 1	Choice 2	Choice 3
Price (USD)	**200**	**50**	**100**
Weight (lbs)	**3**	**0.5**	**0.5**
Power Output (Watts)	**16 (One cellphone in 2 hours)**	**16 (One cellphone in 2 hours)**	**16 (One cellphone in 2 hours)**
Years of Manufacturers Warranty	**1**	**1**	**3**
Portability	**Compacted size of a thumb drive**	**Compacted size of a deck of cards**	**Compacted size of a box fan**
	○	○	○




1.00
0.75
0.50
0.25
0.00
Market Share
20
40
60
80
100
120
140
160
180
200
Price ($)


Power output (W)
Price ($)
Weight (lbs)
portabilitylaptop
portabilitythumbDrive
Warranty Life (Years)
Attribute
16
128
200
50
5
0.5
1
0
0
1
1
5
0.04
0.22
0.4
0.58
0.76
0.94
Market Share


1.00
0.75
0.50
0.25
0.00
Market Share
Our product
Competitor A
Competitor B

Your project starts now!

View project Ideas

Week 1: *Getting Started*

1. Course orientation

2. Intro to conjoint analysis

3. Introductions

BREAK: Teaming

4. Getting started with R & RStudio

Introduce yourself

- Preferred name
- Degree program
- Prior experience
- What do you hope to gain from this class?
- Project interests?

Break: Teaming

Week 1: *Getting Started*

1. Course orientation

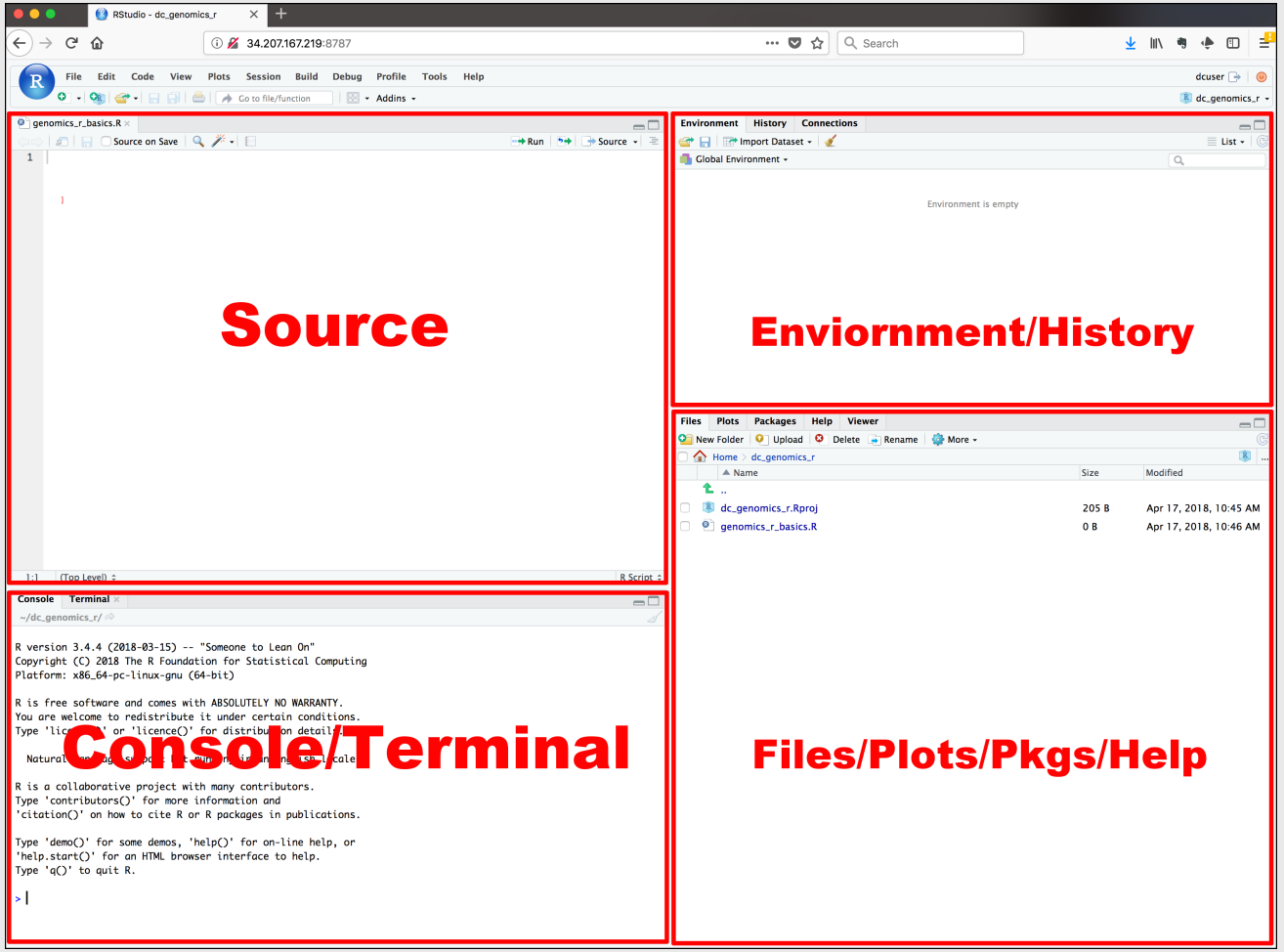
2. Intro to conjoint analysis

3. Introductions

BREAK: Teaming

4. Getting started with R & RStudio

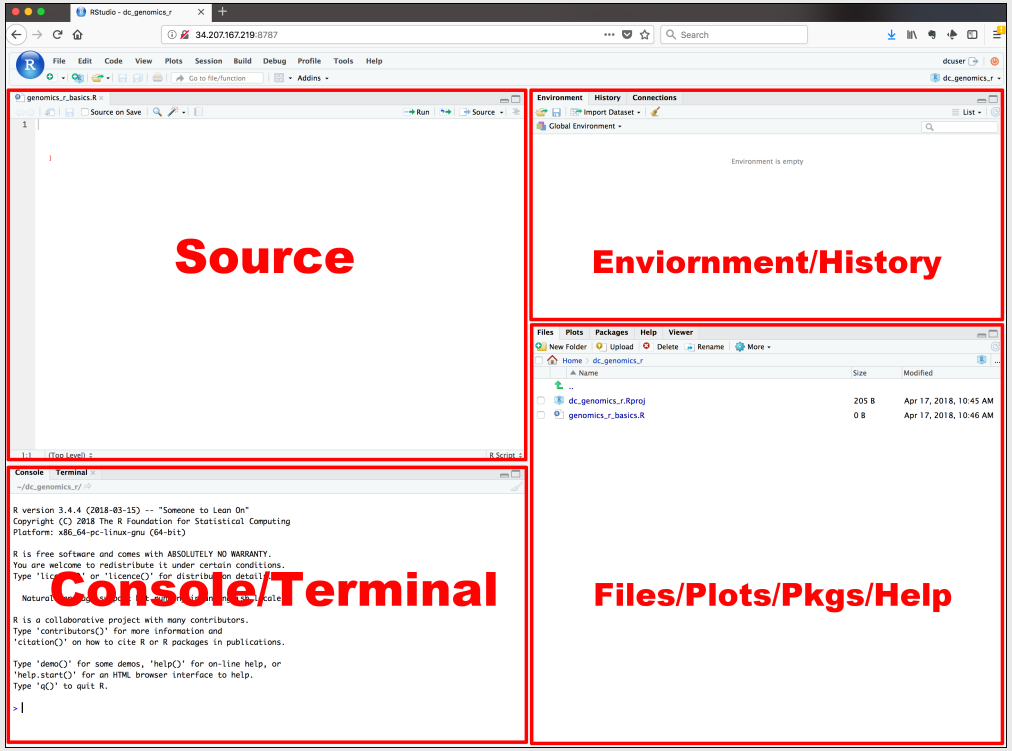
RStudio Orientation



- Know the boxes
- Customize the layout
- Customize the look
- Extra themes

Open `intro_to_R.R` file and follow along

View prior code in history pane



Use "up" arrow see previous code

Staying organized

1) Save your code in .R files

File > New File > R Script

2) Keep work in R Project files

File > New Project...

Your turn

A. Practice getting organized

1. Open RStudio and create a new R project called `week1`.
2. Create a new R script and save it as `practice.R`.
3. Open the `practice.R` file and write your answers to these questions in it.

10:00

B. Creating & working with objects

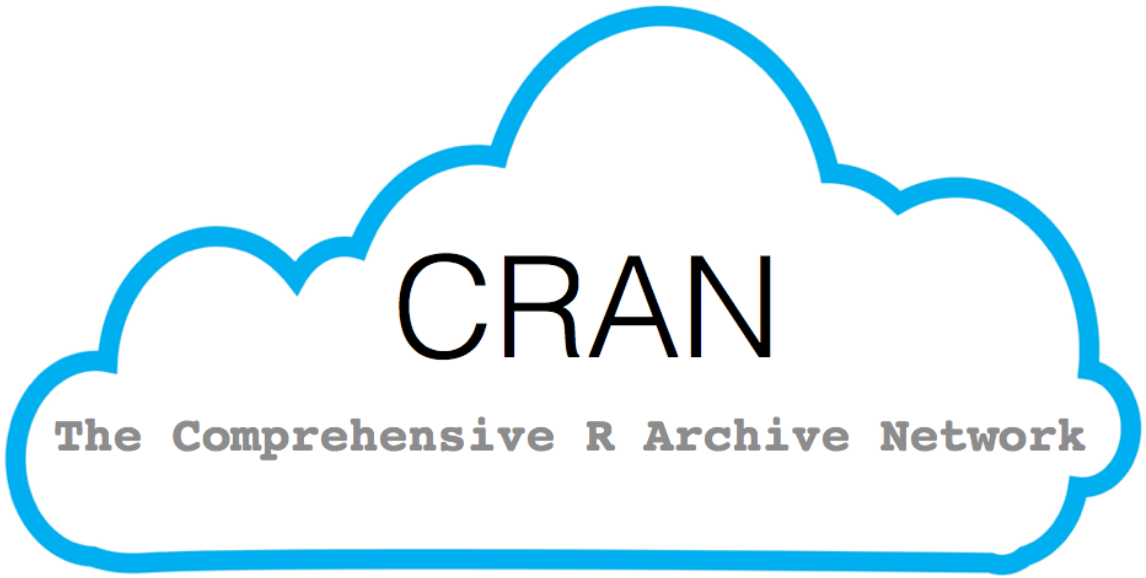
1). Create objects to store the values in this table:

City	Area (sq. mi.)	Population (thousands)
San Francisco, CA	47	884
Chicago, IL	228	2,716
Washington, DC	61	694

2) Using the objects you created, answer the following questions:

- Which city has the highest density?
- How many *more* people would need to live in DC for it to have the same population density as San Francisco?

>15,000 packages on the CRAN



Installing packages

```
install.packages("packagename")
```

(The package name **must** be in quotes)

```
install.packages("packagename") # This works
install.packages(packagename)   # This doesn't work
```

You only need to install a package once!

Loading packages

`library(packagename)`: Loads all the functions in a package

(The package name *doesn't* need to be in quotes)

```
library("packagename") # This works
library(packagename)   # This also works
```

You need to *load* the package every time you use it!

Installing vs. Loading

INSTALL ONCE:

```
install.packages("light")
```

USE MANY TIMES:

```
library("light")
```

Example: **wikifacts**

Install the Wikifacts package, by Keith McNulty:

```
install.packages("wikifacts")
```

Load the package:

```
library(wikifacts) # Load the library
```

Use one of the package functions

```
wiki_randomfact()
```

```
#> [1] "Here's some news from 28 March 2019. The Islamic State of Iraq and the Levant
loses its last territory in Syria following a defeat by the Syrian Democratic Forces and
the U.S.-led coalition. (Courtesy of Wikipedia)"
```

Example: **wikifacts**

Now, restart your RStudio session:

> Session -> Restart R

Try using the package function again:

```
wiki_randomfact()
```

```
#> Error in wiki_randomfact(): could not find function "wiki_randomfact"
```

Using only *some* package functions

You don't always have to load the whole library.

Functions can be accessed with this pattern:

```
packagename::functionname()
```

```
wikifacts::wiki_randomfact()
```

```
#> [1] "Here's some news from 01 February 2020. Floods and landslides in southeastern
Brazil kill at least sixty people and leave thousands others homeless. (Courtesy of
Wikipedia)"
```

If you haven't yet, install these packages

Back `intro_to_R.R` for the rest of class!