Week 1: *Getting Started*

🏛 EMSE 6035: Marketing Analytics for Design Decisions

👤 John Paul Helveston

📅 August 27, 2025

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 Helveston, Ph.D.

- 2025 - Present: Associate Professor, EMSE
- 2018 - 2025: Assistant Professor, EMSE
- 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

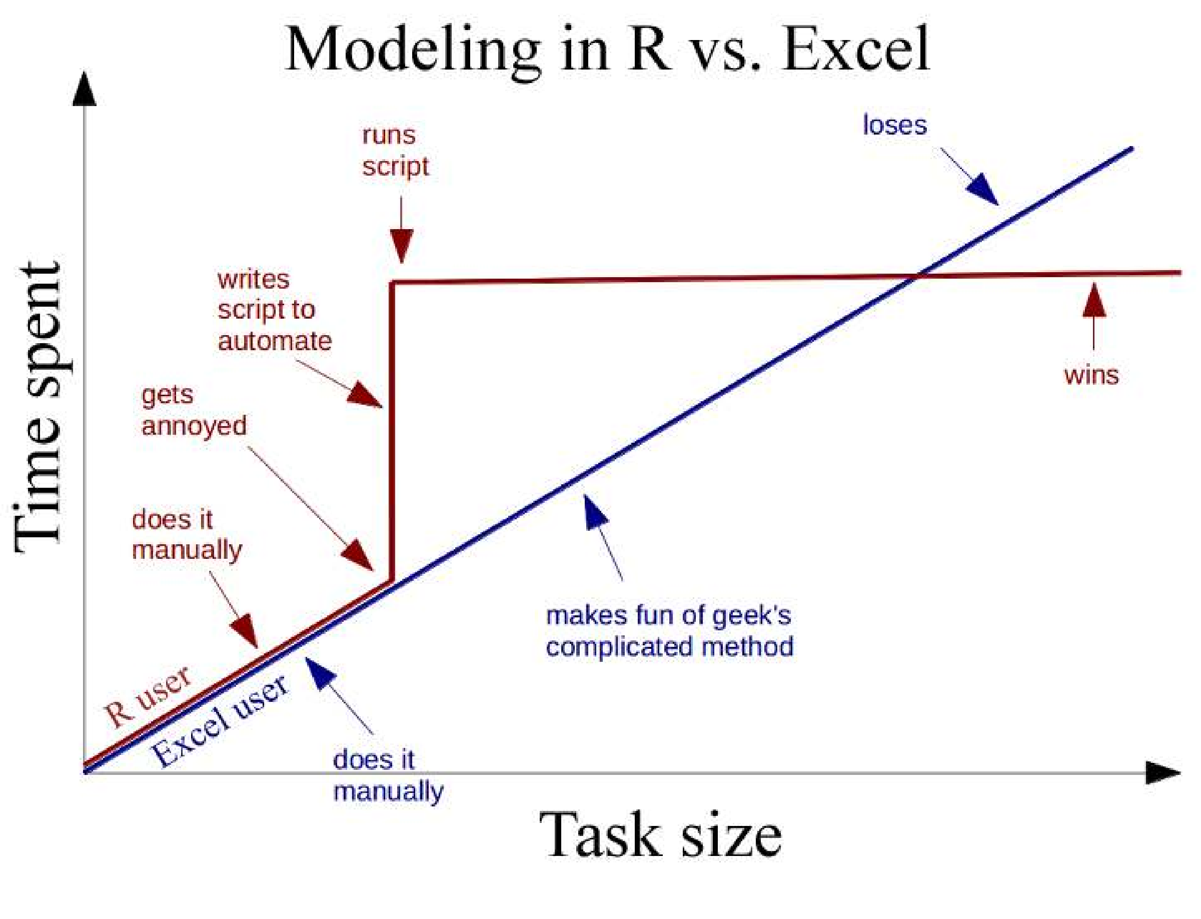
Tools

Course website: https://madd.seas.gwu.edu/2025-Fall/

Course slack: https://emse-madd-f25.slack.com

& RStudio: Course Software Page

Why R?


Modeling in R vs. Excel
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
Time spent
Task size

Wait, why aren't we using Python?

The field of choice modeling grew out of Statistics, not Machine Learning

- Python is a general purpose language developed by **Guido van Rossum**, a computer scientist.
- Unlike R, Python was not originally developed for data analysis.
- Both languages are extremely useful, and **you should probably learn python too**.

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
- …effectively communicate insights

Course prerequisites

This course requires prior exposure to:

- Probability theory
- Multivariable calculus
- Linear algebra
- Regression

Not sure?

Take this self assessment

Reflections (30% of grade)

Do some readings, recorded lectures, practice problems

Write a short reflection

📅 ~Every week (12 total)

🕓 Due 11:59pm Monday before class

✓ Graded for completion (looking for engagement)

Quizzes (10% of grade)

📅 At the start of class every other week-ish. Make ups only for excused absences (i.e. don't be late).

📅 5 total, lowest dropped

🕓 10 minutes

Why quiz at all? The "retrieval effect" - basically, you have to *practice* remembering things, otherwise your brain won't remember them (see 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 (45% of grade)

Teams of 3-4 students

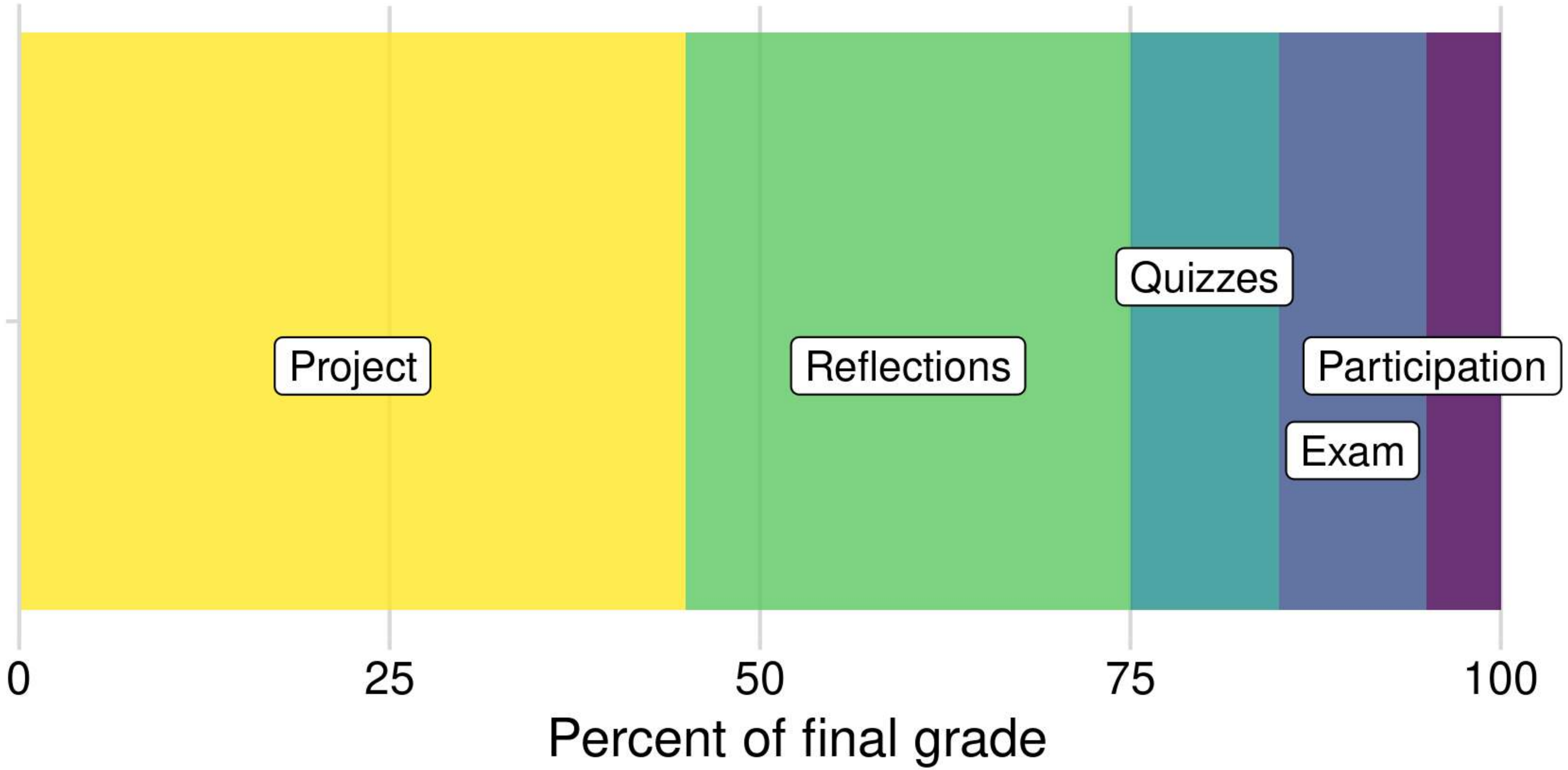
Goals:

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

Key deliverables:

Item	Weight	Due
Project Proposal	5 %	Sep. 21
Survey Plan	5 %	Sep. 30
Pilot Survey	5 %	Oct. 14
Pilot Analysis	5 %	Nov. 02
Final Survey	5 %	Nov. 16
Final Analysis Report	15 %	Dec. 07
Final Presentation	5 %	Dec. 09

Grades


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

Grades

Item	Weight	Notes
Participation / Attendance	5%	(Yes, I take attendance)
Reflections	30 %	Weekly assignment (10 × 3%, lowest dropped)
Quizzes	10 %	5 quizzes, lowest dropped
Final Exam	10 %	Take home exam
Project Proposal	5 %	Teams of 3-4 students
Survey Plan	5 %	
Pilot Survey	5 %	
Pilot Analysis	5 %	
Final Survey	5 %	
Final Analysis Report	15 %	
Final Presentation	5 %	

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

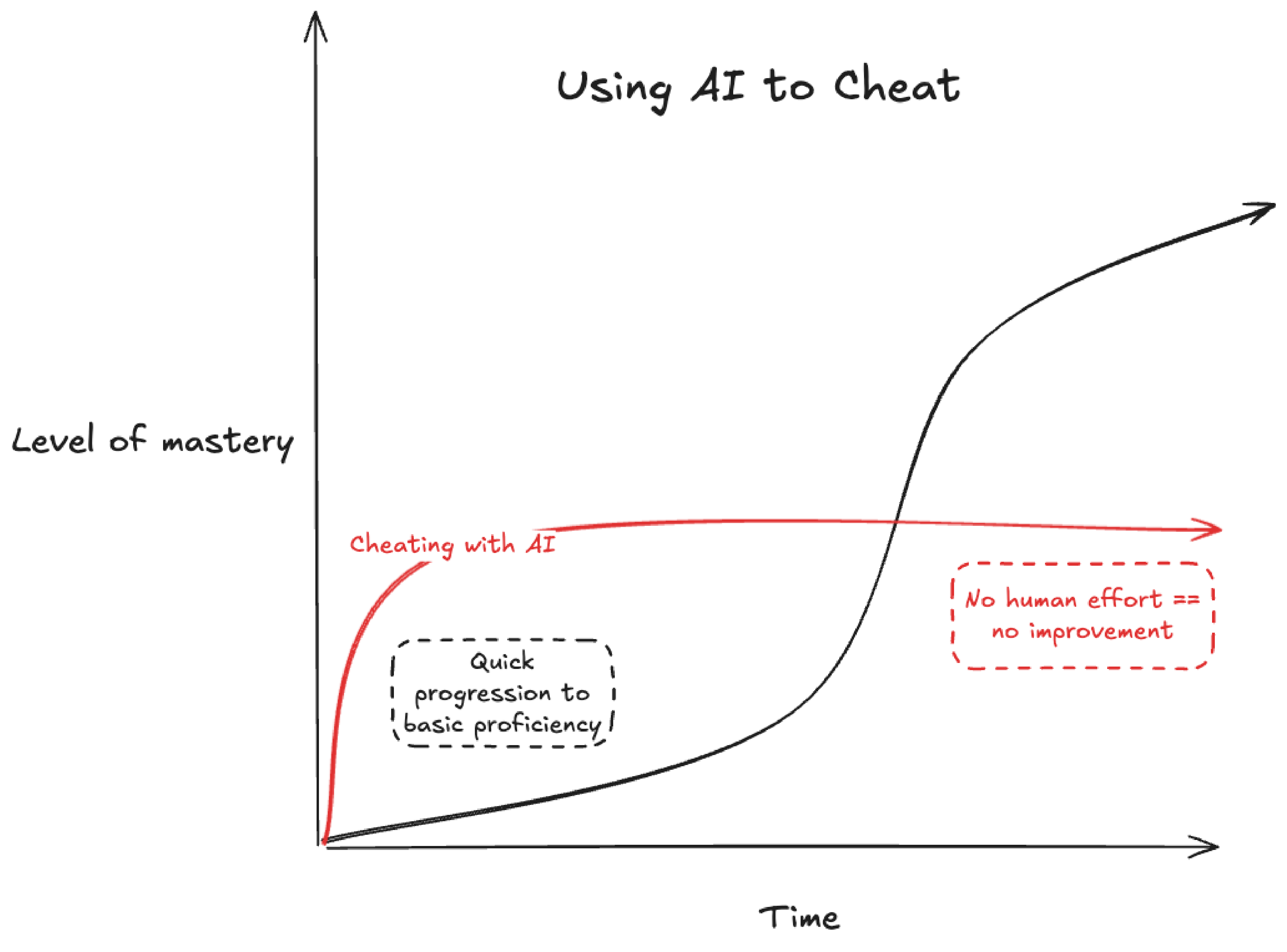
Use of chatGPT and other AI tools

- Large language models (LLMs) are pretty good
- Sometimes they suck.
- I will grade whatever you submit. **It should not suck.**

Ways to not have your work suck:

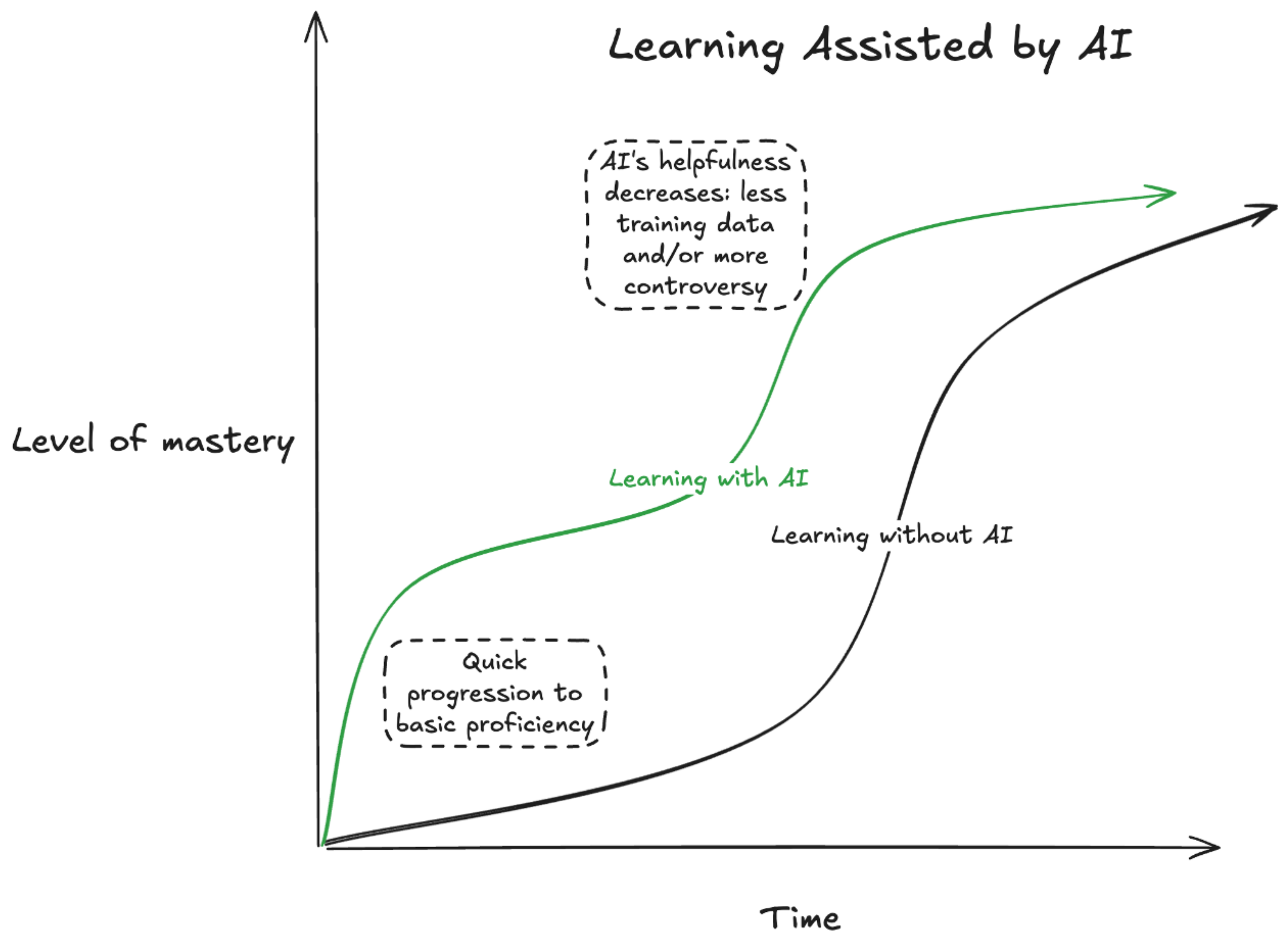
- Don't submit code that doesn't run (actually run it before submitting it).
- Actually read what the AI generates, and **don't submit something you don't understand**. (Ask the LLM how it works)
- There are dozens of ways to do things - **you should use the approach I teach**.
- Ask yourself **"what would I ask Prof. Helveston?"**

Use AI as an assistant, not a solutions manual



The **wrong** ways to use LLMs

Don't copy-paste past the basics - it'll rob you of mastery



The **right** way to use LLMs

- Asking to explain error messages
- Asking to explain *why* something works or doesn't work

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 8:00-4:30pm
- Tuesdays from 8:00-4:30pm
- Fridays from 8:00-4:00pm

GW Coders

Week 1: *Getting Started*

1. Course orientation

2. Intro to conjoint analysis

3. Introductions

BREAK: Teaming

4. Getting started with R & RStudio

Engineers often design things nobody wants!

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

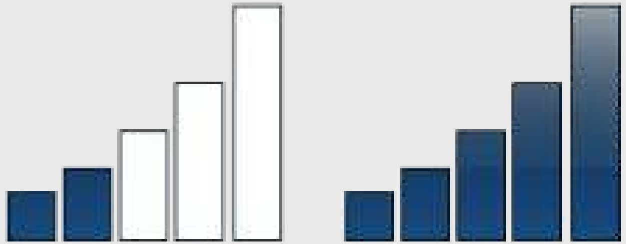
(People want everything)

Which feature do you care more about?

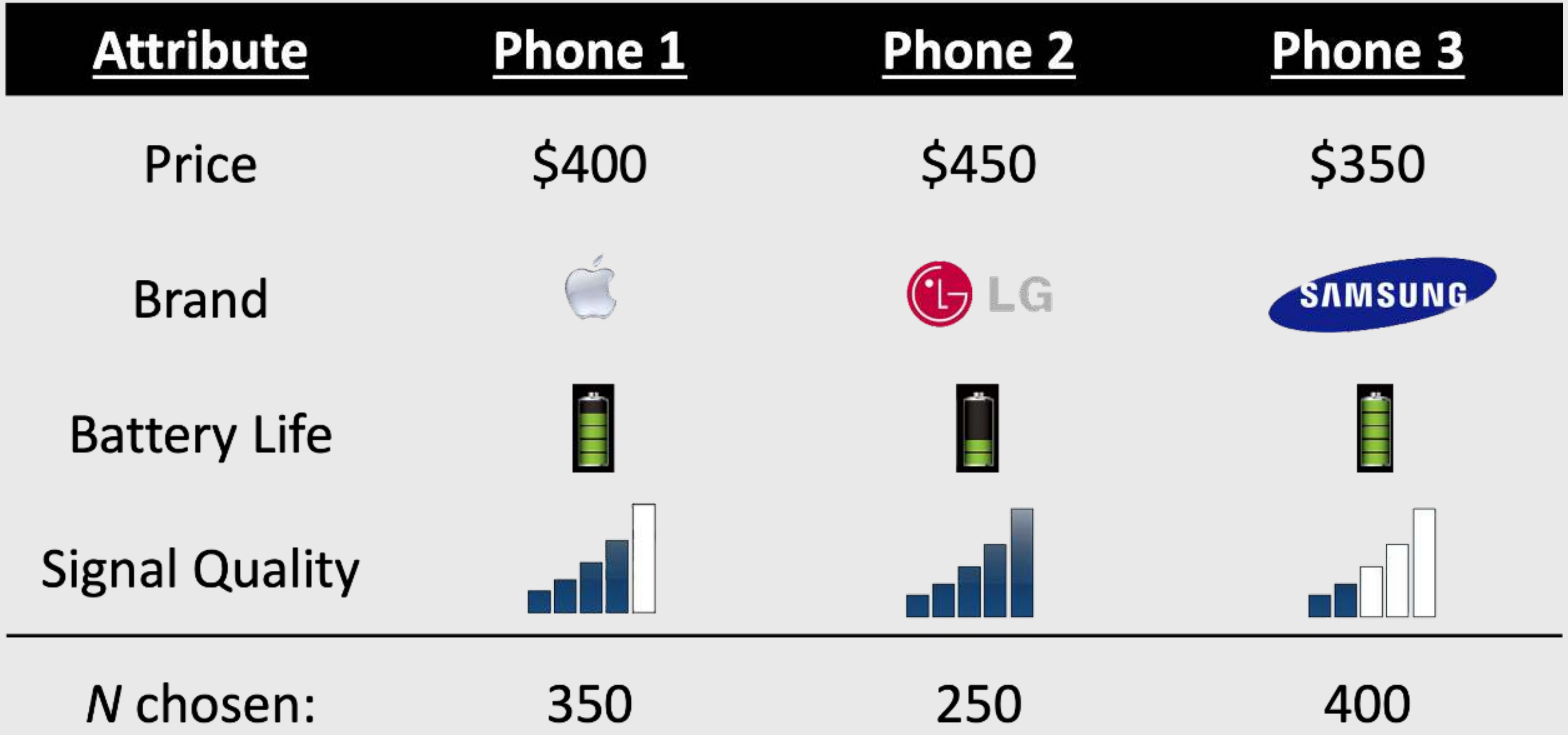
Battery Life?

Brand?

Signal quality?



Conjoint approach:
Use consumer choice data to model preferences



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

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

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

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 β_1, β_2, β_3, β_4 by minimizing $-L = -\sum_n^N \sum_j^J y_{nj} \ln P_{nj}$

Willingness to Pay

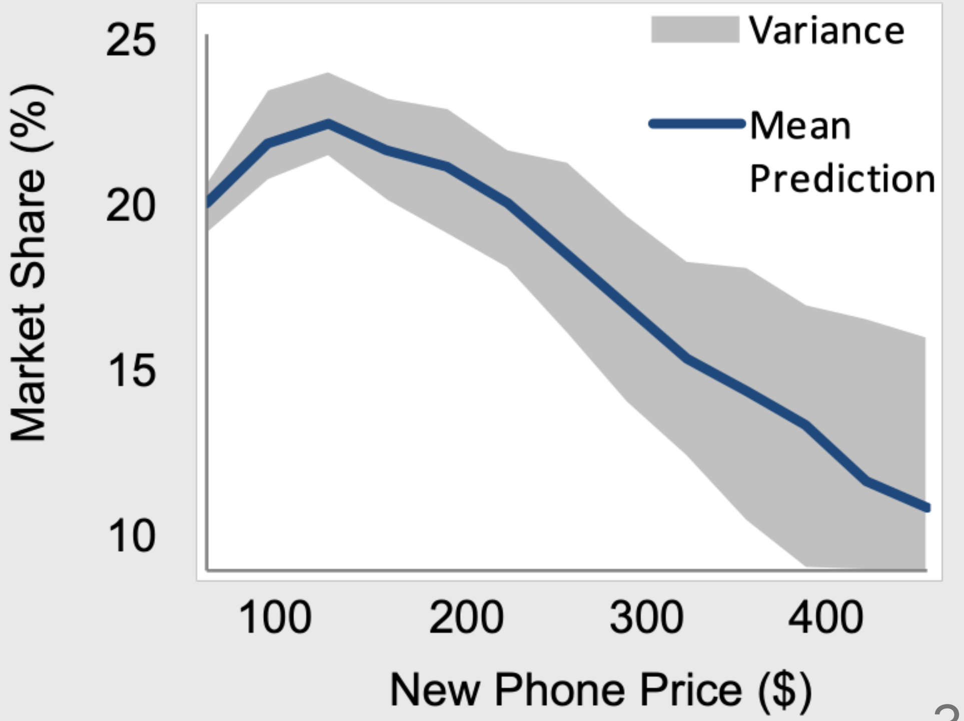
$$u_j = \beta' x_j + \alpha p_j + \varepsilon_j$$

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

"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

Attribute Units

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

Decision Variable Units

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

Durability

Degradation Protections

Portability

Power Output

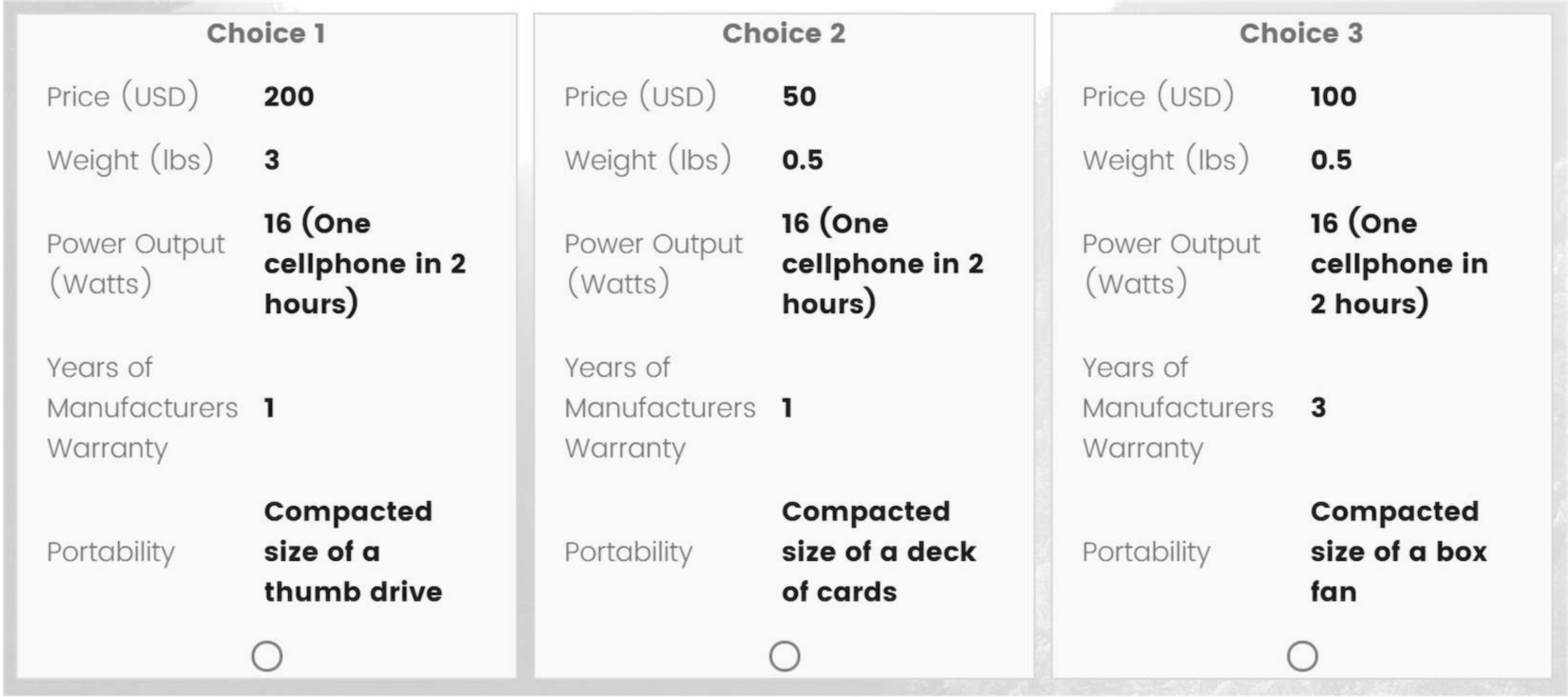
Power Density

Packing Design

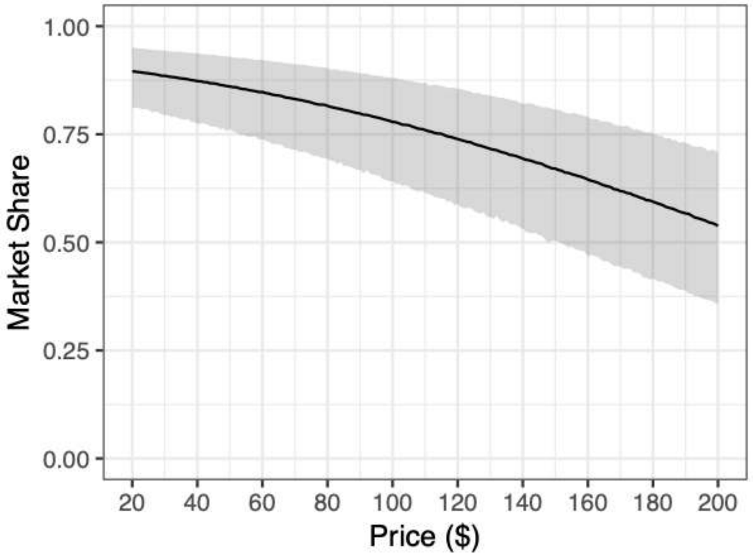
Weight

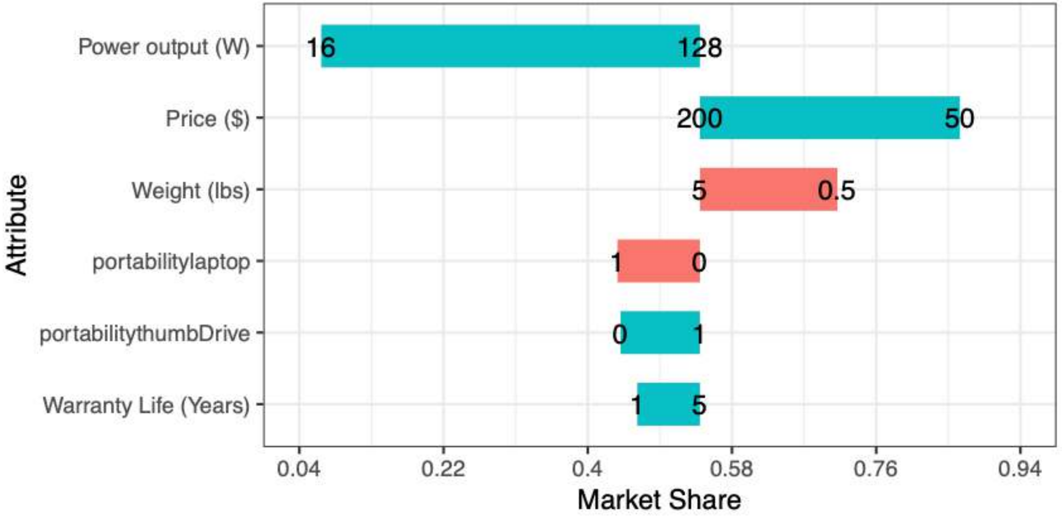
Image Sources:
1.https://www.deviantart.com/kota3214/art/Rainy-Sun-244479413
2.https://www.pennscale.com/product/ntep-certified-scales/label-printing-[illegible]
3. https://www.thisiswhyimbro[illegible]/portable-foldable-solar-panel-charger/

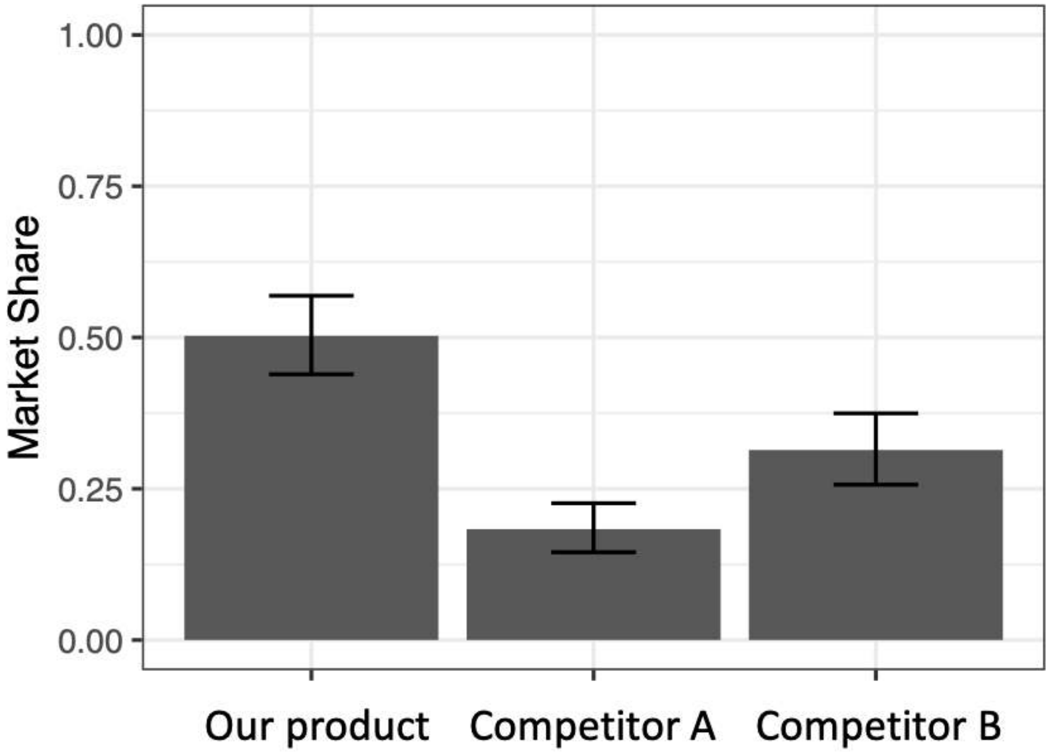
Example survey choice question



	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**
	○	○	○


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


Attribute
Power output (W)
Price ($)
Weight (lbs)
portabilitylaptop
portabilitythumbDrive
Warranty Life (Years)
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


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

Your project starts now!

View projects

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

1. If you haven't already, install everything on the software page

2. Stand up, meet each other, (maybe form teams?...use this sheet)

05:00

Week 1: *Getting Started*

1. Course orientation

2. Intro to conjoint analysis

3. Introductions

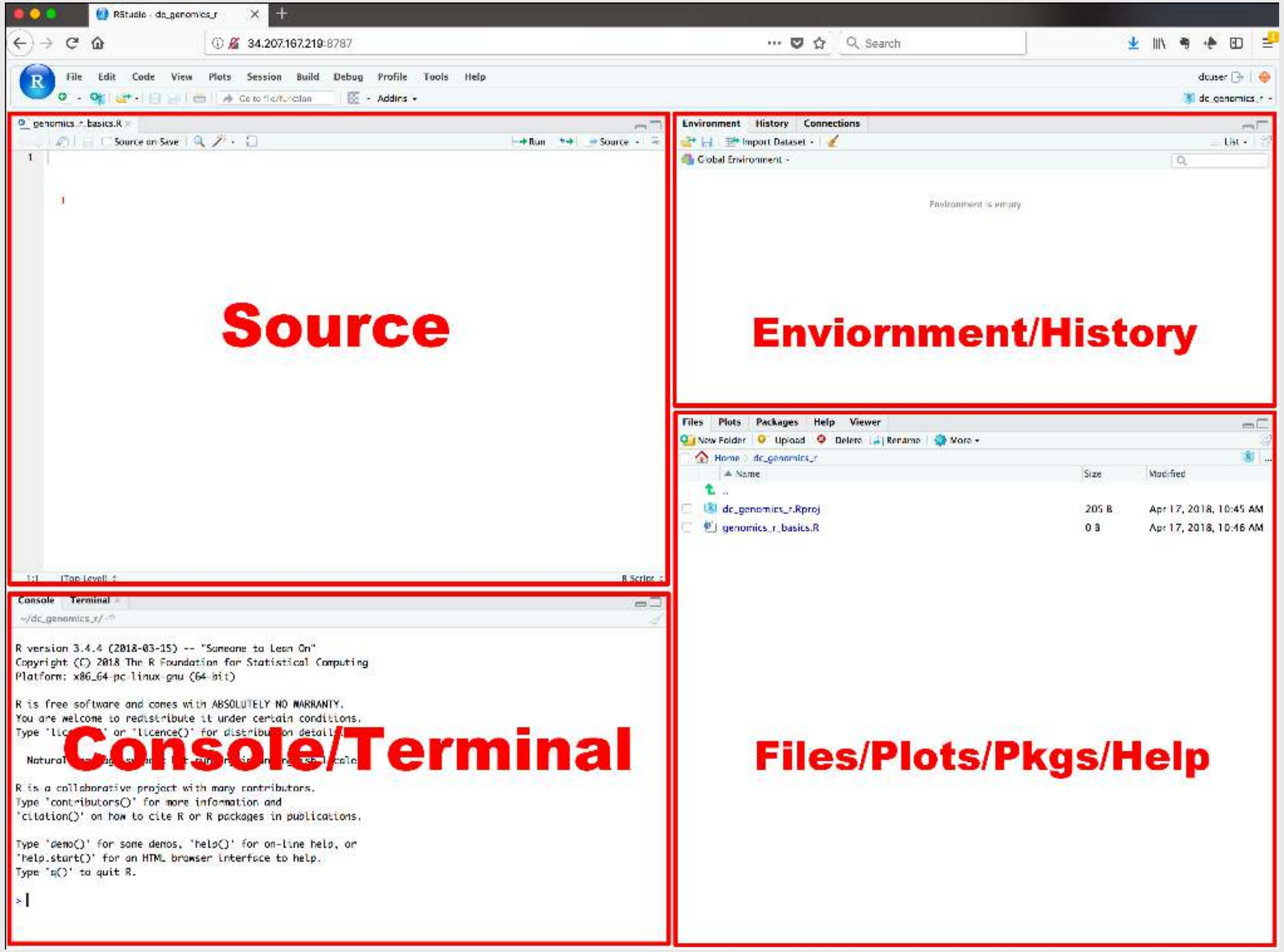
BREAK: Teaming

4. Getting started with R & RStudio

If you're familiar with R and RStudio,

download and try out Positron

RStudio Orientation



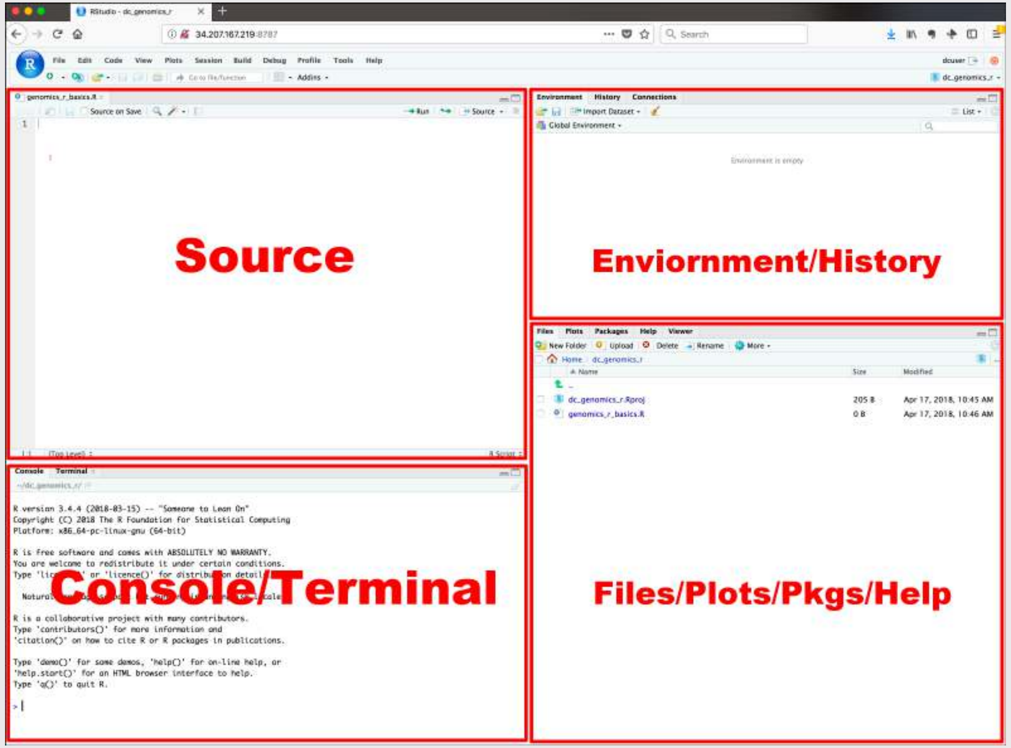
Name	Size	Modified
..		
dc_genomics_r.Rproj	205 B	Apr 17, 2018, 10:45 AM
genomics_r_basics.R	0 B	Apr 17, 2018, 10:46 AM

```
R version 3.4.4 (2018-03-15) -- "Someone to Lean On"
Copyright (C) 2018 The R Foundation for Statistical Computing
Platform: x86_64-pc-linux-gnu (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

  Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

> 
```

- 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



Name	Size	Modified
..		
dc_genomics_r.Rproj	205 B	Apr 17, 2018, 10:45 AM
genomics_r_basics.R	0 B	Apr 17, 2018, 10:46 AM

```
R version 3.4.4 (2018-03-15) -- "Someone to Lean On"
Copyright (C) 2018 The R Foundation for Statistical Computing
Platform: x86_64-pc-linux-gnu (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

  Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

> |
```

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

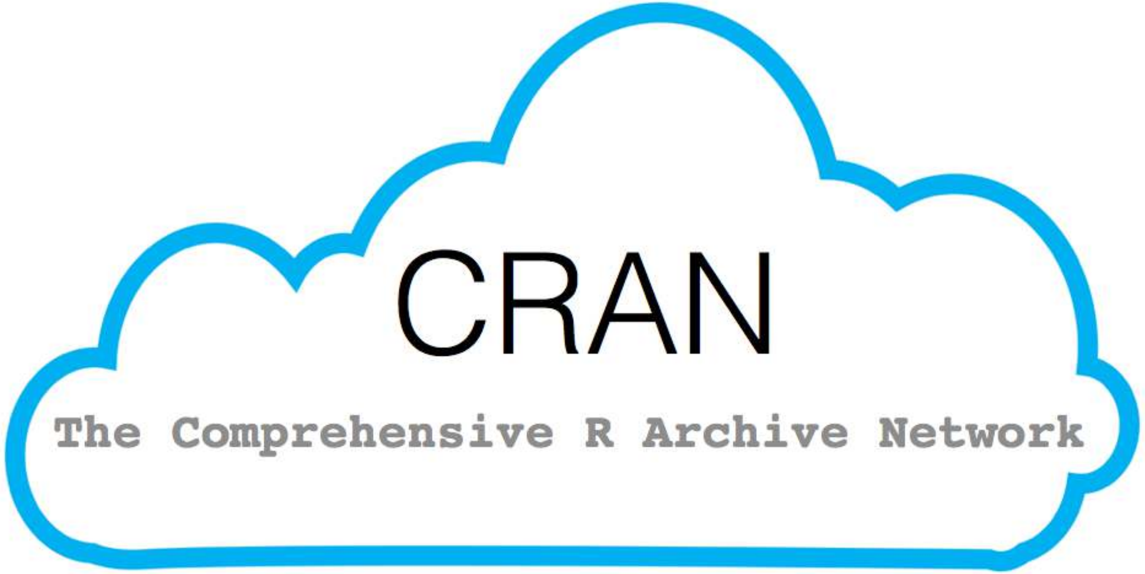
1). Create objects to store the values in this table:

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)

```r
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)

```r
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")
```

wikiHow to Change a Light Bulb

USE MANY TIMES:

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

wikiHow to Change a Light Bulb

Example: **wikifacts**

Install the Wikifacts package, by Keith McNulty:

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

Load the package:

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

Use one of the package functions

```r
wiki_randomfact()
```

```
#> [1] "Did you know that on October 12 in 1799 – Jeanne Geneviève Garnerin became the
first woman to make a parachute descent, falling 900 metres (3,000&nbsp;ft) in the
gondola of a hot air balloon. (Courtesy of Wikipedia)"
```

Example: **wikifacts**

Now, restart your RStudio session:

> Session -> Restart R

Try using the package function again:

```r
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] "Did you know that on April 26 in 2007 – Controversy surrounding the relocation
of the Bronze Soldier of Tallinn, a Soviet Red Army World War&nbsp;II memorial in
Tallinn, Estonia, erupted into mass protests and riots. (Courtesy of Wikipedia)"
```

If you haven't yet, install these packages

Back `intro-to-R.R` for the rest of class!