

## 02 - Experimental Design

HCI/PSYCH 522  
Iowa State University

January 20, 2022

# Workday Course Registration Case Study

## ← Select Sections for Schedule

Academic Period 2020 Spring Semester

Schedule Name ★ Plan A Schedule - Spring 2020

Viewing:

Click here to sort

**FINC 351 - Investments**  
3 Units

**FINC 307 - Survey of Finance**  
3 Units

**ECON 495 - Capstone in Economics**  
3 Units

**ECON 330 - Money and Credit**  
3 Units

**ANTH 340 - Anthropology of Food**  
3 Units

**Course Listing** ECON 495 - Capstone in Economics

**Description** Students review selected economic literature and complete written assignments that relate readings to their area of interest in economics. Class meetings provide opportunities for constructive critiques from the instructor, feedback from other students, in-class writing, and oral presentations.

**Units** 3 Units

**Lecture**

1 item

Select	Section	Eligible	Section Status	Start Date	End Date	Meeting Patterns	Instructor	Section Capacity	Seats Available
<input type="checkbox"/>	ECON 495-1 - Capstone in Economics	<input checked="" type="checkbox"/>	Open	01/07/2020	05/17/2020		Sarjeet Jain	6	3

# Hypothesis

Scientific question: Does the type of registration system (A vs B) affect mean registration time?

Null hypothesis: No

Alternative hypothesis: Yes

Click here to search

Academic Period: 2020 Spring Semester  
Schedule Name: Plan A Schedule - Spring 2020

Viewing: 10

Course Listing: ECON 495 - Capstone in Economics  
Description: Students review selected economic literature and current events that relate readings to their area of interest in economics. Class meetings provide opportunities for constructive critiques from the instructor, feedback from peers, in-class writing, and oral presentations.

Units: 3 Units

**Lecture**

Select	Section	English	Section Status	Start Date	End Date	Meeting Pattern	Instructor	Section Capacity	Seats Available
<input type="checkbox"/>	ECON 495 - Capstone in Economics		Open	01/01/2020	05/11/2020		Subject Jahn	6	3

OK Cancel

Click here to search

Academic Period: 2020 Spring Semester  
Schedule Name: Plan A Schedule - Spring 2020

Viewing: 10

Course Listing: ECON 495 - Capstone in Economics  
Description: Students review selected economic literature and current events that relate readings to their area of interest in economics. Class meetings provide opportunities for constructive critiques from the instructor, feedback from peers, in-class writing, and oral presentations.

Units: 3 Units

**Lecture**

Select	Section	English	Section Status	Start Date	End Date	Meeting Pattern	Instructor	Section Capacity	Seats Available
<input type="checkbox"/>	ECON 495 - Capstone in Economics		Open	01/01/2020	05/11/2020		Subject Jahn	6	3

OK Cancel

# Recruitment



# Randomization

```
registration <- data.frame(subjectID = 1:30,  
                           system = sample(rep(c("A","B"), times = 15)))
```

registration

##	subjectID	system
## 1	1	B
## 2	2	A
## 3	3	A
## 4	4	A
## 5	5	A
## 6	6	A
## 7	7	A
## 8	8	B
## 9	9	A

# Balanced experiment

```
table(registration$system)
```

```
##
```

```
##  A  B
```

```
## 15 15
```

# Data Collection



# Exploratory statistics

```
summary(registration)
```

##	subjectID	system	time
##	Min. : 1.00	Length:30	Min. :0.8155
##	1st Qu.: 8.25	Class :character	1st Qu.:2.0621
##	Median :15.50	Mode :character	Median :2.7388
##	Mean :15.50		Mean :2.8362
##	3rd Qu.:22.75		3rd Qu.:3.5307
##	Max. :30.00		Max. :4.2267
##			NA's :1

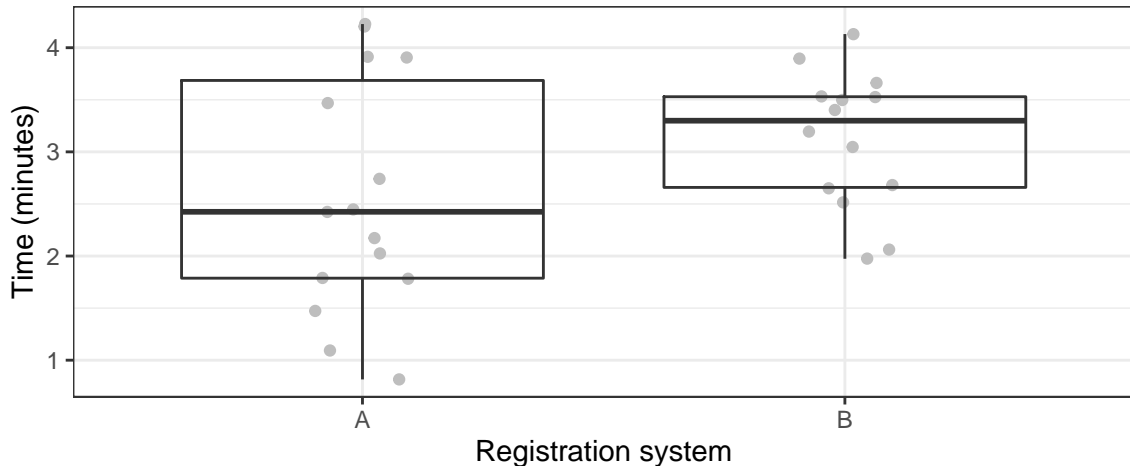


# Exploratory statistics

```
registration %>%  
  group_by(system) %>%  
  summarize(n = sum(!is.na(time)),  
            mean = mean(time, na.rm = TRUE),  
            sd = sd(time, na.rm = TRUE))  
  
## # A tibble: 2 x 4  
##   system      n mean   sd  
##   <chr> <int> <dbl> <dbl>  
## 1 A      15  2.57  1.13  
## 2 B      14  3.13  0.661
```

# Visualization

## Registration system comparison



## Two-sample t-test

```
t.test(time ~ system, data = registration)

##
##  Welch Two Sample t-test
##
## data:  time by system
## t = -1.6405, df = 22.796, p-value = 0.1146
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -1.2695283  0.1468389
## sample estimates:
## mean in group A mean in group B
##           2.565219           3.126563
```

# Regression

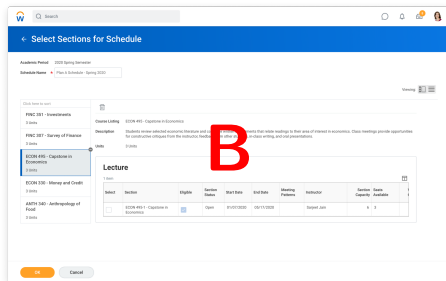
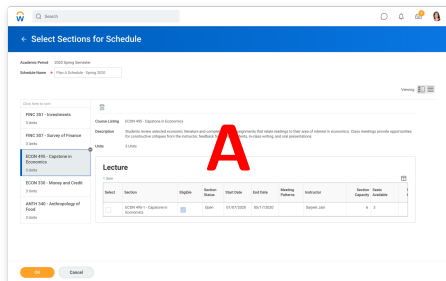
```
summary(lm(time ~ system, data = registration))

##
## Call:
## lm(formula = time ~ system, data = registration)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.74973 -0.60921 -0.07862  0.53688  1.66149
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2.5652     0.2420  10.602   4e-11 ***
## systemB       0.5613     0.3482   1.612   0.119
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.9371 on 27 degrees of freedom
## (1 observation deleted due to missingness)
## Multiple R-squared:  0.08778, Adjusted R-squared:  0.054
```

# Conclusion and Decision

Conclusion: Since  $p\text{-value} > 0.05$ , there is insufficient evidence to indicate any difference in mean registration time between the two systems.

Decision: The decision of which system to use may depend on many factors including cost.



# What was done well in this study?

# What was done poorly in this study?

# What else could have been done?



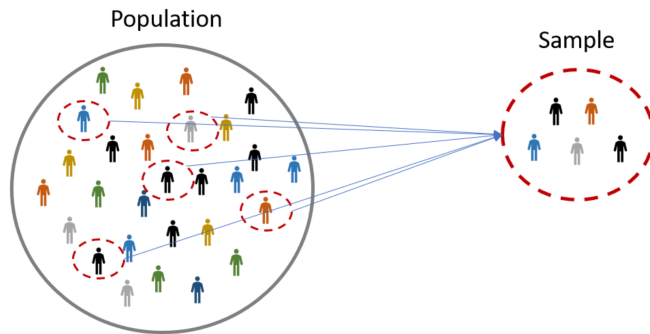
# Data Collection Methods

# Population vs Sample

from <https://www.scribbr.com/methodology/population-vs-sample/>

A **population** is the entire group that you want to draw conclusions about.

A **sample** is the specific group that you will collect data from.



# Sample

# Data vs Information

## DIFFERENCE BETWEEN DATA AND INFORMATION



### DATA

Data is raw, unorganized facts that need to be processed. Data can be something simple and seemingly random and useless until it is organized.



### INFORMATION

When data is processed, organized, structured or presented in a given context so as to make it useful, it is called information.

# Information spectrum

Data quality (least informative to most informative)

- Survey
- Observational study
- Randomized experiment
  - paired/blocked

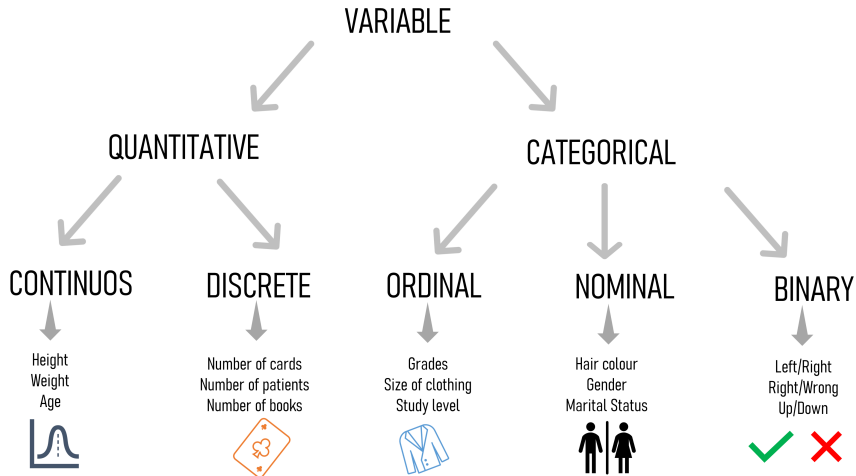
Sample (least informative to most informative)

- You
- Select students
- Randomly selected students
- All students

# Data Types

- Qualitative: categorical or non-numeric
  - Binary: only two options
  - Nominal: order doesn't matter
  - Ordinal: order matters
- Quantitative: numerical
  - Discrete: count
    - upper maximum does not exist
    - upper maximum exists
  - Continuous

# Data Types Graphic



# Statistical Methods for Continuous Dependent Variables

