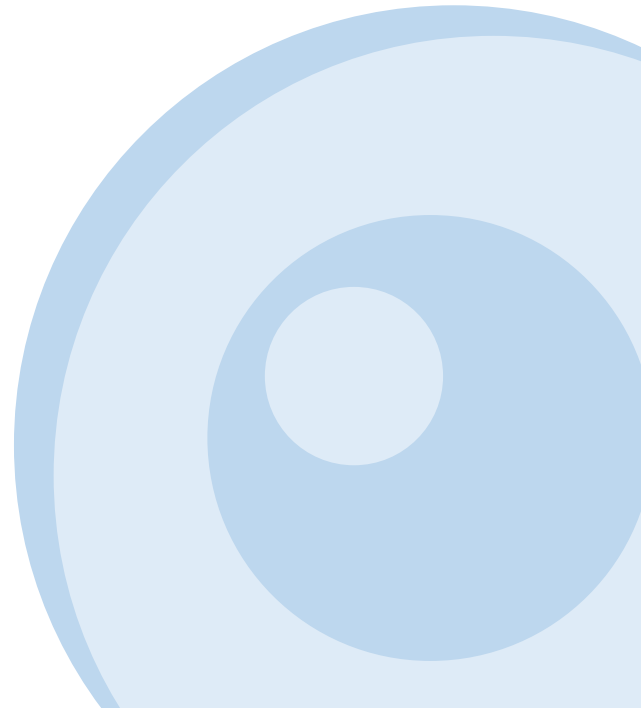


PSYC 193R:

Data Analysis & Visualization with R

Lecture 8: Paired-samples t-test

Jonas S.-H. Lau [silau -et- ucsd.edu]
psyc193r.ucsd.edu



Blind-tasting: The Design

- What was the research question?
- What was the target population?
- What were the IV(s) and DV(s)?
- How were the IV(s) and DV(s) operationalized?
(also, scale of measurement?)
- Was it a between / within-subjects design?
- What are the expected results?
If the manipulation does not work:
If the manipulation works:



Import Data

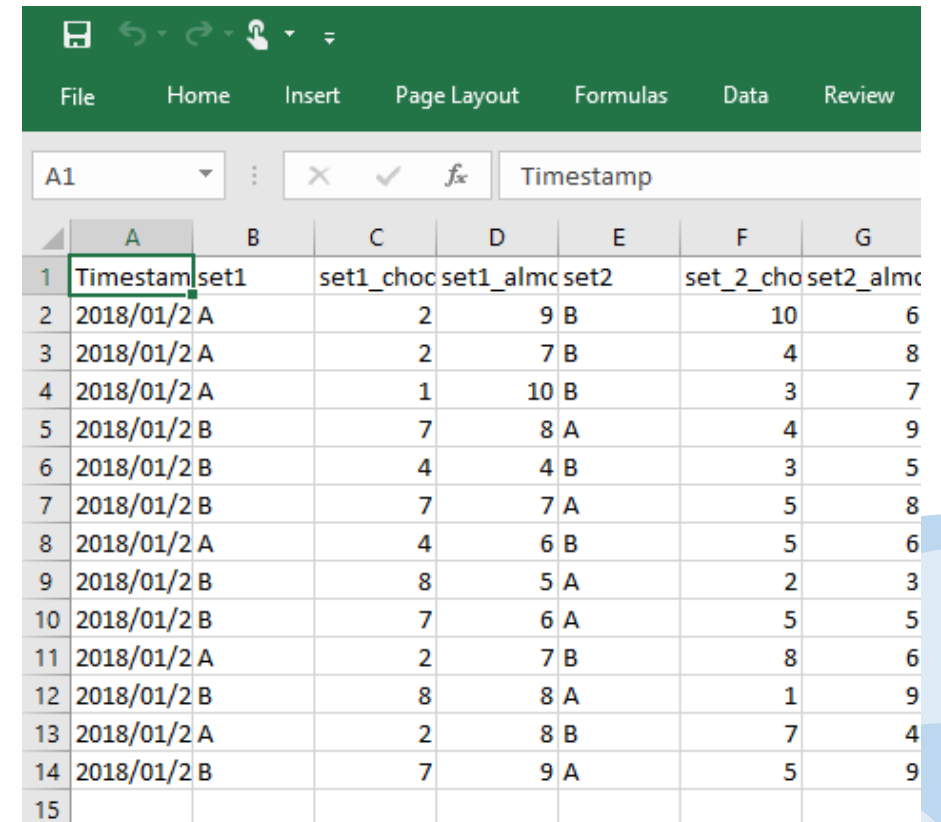
- Download the data file:
https://psyc193r.ucsd.edu/data/psyc193r_blindTaste_data.csv
- Tell R where your data file is:
Session → Set Working Directory → Choose Directory
`setwd()`
- Check the *comma separated values* (csv) file with MS Excel to ensure correct format
- Import the file into R as a data frame, saving it to “blindTasteData” variable
`... = read.csv("file_name.csv")`

Data format

The data set has a "wide" format:
multiple responses in the same row

The variables that are relevant:

- `set1`: the first set to evaluate
- `set1_chocolate`:
quality of the chocolate of the first set
- `set1_almond`:
quality of the almond of the first set
- `set2`: <same> but for 2nd set
- `set2_chocolate`: <same> but for 2nd set
- `set2_almond`: <same> but for 2nd set



	A	B	C	D	E	F	G
1	Timestamp	set1	set1_choc	set1_almc	set2	set_2_choc	set2_almc
2	2018/01/2	A	2	9	B	10	6
3	2018/01/2	A	2	7	B	4	8
4	2018/01/2	A	1	10	B	3	7
5	2018/01/2	B	7	8	A	4	9
6	2018/01/2	B	4	4	B	3	5
7	2018/01/2	B	7	7	A	5	8
8	2018/01/2	A	4	6	B	5	6
9	2018/01/2	B	8	5	A	2	3
10	2018/01/2	B	7	6	A	5	5
11	2018/01/2	A	2	7	B	8	6
12	2018/01/2	B	8	8	A	1	9
13	2018/01/2	A	2	8	B	7	4
14	2018/01/2	B	7	9	A	5	9
15							

The Ideal Data Frame: setA / setB columns

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins

psyc193r_blindTaste.R x psyc193r_I08_paired-sample-t_jonas.... x blindTasteData x

Filter

chocolate	set1_almond	set2	set2_chocolate	set2_almond	gender	like_chocolate	like_almond	set1_score	set2_score	setA	setB	diff
2	9	B	10	6	Female	10	8	11	16	11	16	5
2	7	B	4	8	Male	9	10	9	12	9	12	3
1	10	B	3	7	Female	10	10	11	10	11	10	-1
7	8	A	4	9	Male	7	8	15	13	13	15	2
4	4	B	3	5	Female	4	4	8	8	8	8	0
7	7	A	5	8	Male	9	7	14	13	13	14	1
4	6	B	5	6	Female	7	7	10	11	10	11	1
8	5	A	2	3	Female	6	5	13	5	5	13	8
7	6	A	5	5	Female	8	7	13	10	10	13	3
2	7	B	8	6	Female	5	5	9	14	9	14	5
8	8	A	1	9	Female	10	8	16	10	10	16	6
2	8	B	7	4	Male	8	8	10	11	10	11	1
7	9	A	6	9	Male	9	9	16	14	14	16	2

Showing 1 to 13 of 13 entries

Data Cleaning / Processing Plan

- `factorize()` / `as.numeric()` the relevant variables

- Remove invalid entries

- Participants evaluating only A or B

```
blindTasteData = blindTasteData[blindTasteData$set1 != __ ,]
```

- Make a composite score for each set

- Combining each participant's set 1 chocolate + almond scores (`set1_score`)
 - Do the same for set 2 (`set2_score`)

```
blindTasteData$set1_score = blindTasteData$set1_chocolate + ...
```

Data Processing Plan

- Make 2 new columns, one for each set
 - setA
 - setB

```
blindTasteData$setA = NA  
blindTasteData$setB = NA
```

- Make a for() loop to go over each row,
if set1 == "A", setA == set1_score, etc

```
for (i in 1:nrow(blindTasteData)){  
  if (blindTasteData$set1[i] == "A"){  
    blindTasteData$setA[i] = NA  
    ...  
  } else {  
    blindTasteData$setA[i] = NA  
    ...  
  }  
}
```

The Ideal Data Frame: setA / setB columns

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins

psyc193r_blindTaste.R x psyc193r_I08_paired-sample-t_jonas.... x blindTasteData x

Filter

chocolate	set1_almond	set2	set2_chocolate	set2_almond	gender	like_chocolate	like_almond	set1_score	set2_score	setA	setB	diff
2	9	B	10	6	Female	10	8	11	16	11	16	5
2	7	B	4	8	Male	9	10	9	12	9	12	3
1	10	B	3	7	Female	10	10	11	10	11	10	-1
7	8	A	4	9	Male	7	8	15	13	13	15	2
4	4	B	3	5	Female	4	4	8	8	8	8	0
7	7	A	5	8	Male	9	7	14	13	13	14	1
4	6	B	5	6	Female	7	7	10	11	10	11	1
8	5	A	2	3	Female	6	5	13	5	5	13	8
7	6	A	5	5	Female	8	7	13	10	10	13	3
2	7	B	8	6	Female	5	5	9	14	9	14	5
8	8	A	1	9	Female	10	8	16	10	10	16	6
2	8	B	7	4	Male	8	8	10	11	10	11	1
7	9	A	6	9	Male	9	9	15	14	14	15	1

Showing 1 to 13 of 13 entries

Assumptions / Outlier Checking

- Make a “difference” score variable: $\text{diff} = \text{setB} - \text{setA}$
- Plot a histogram of the difference scores

```
hist(...)
```

