


Final Exam review

Probability simulation:

- 1) How to do the process once
 - how the game is played / how the data are simulated
 - how to represent the process in code
 - (e.g. create vector, sample from a vector, for loop / while loop, etc.)
- 2) what event you are interested in (what do we want the probability of, what is at stake of interest)
- 3) Repeat process many times, see how frequently the event occurs
 - for loop
 - store results for each iteration

Pivoting

df

id	group	value
1	a	5
1	b	6
2	a	7
2	b	8
3	a	9
3	b	10

↑ values in this column

pivot_wider



become column names after pivoting wider

id	a	b
1	5	6
2	7	8
3	9	10

df2

df2 |>

```
pivot_longer(-id,  
names_to = "group",  
values_to = "value")
```

df |>

pivot_longer

```
pivot_wider(id_cols = id,  
names_from = group,  
values_from = value)
```

(take column names,
make them entries in
a column of the data)

Class activity , October 18

Q3:

household data

family	dob-child1	dob-child2	name-child1	name-child2
1	1998-11-26	2000-01-24	Susan	Jose
2	1996-06-22	NA	Mark	NA
3				
4				
5				

want:

family	child	dob	name
1	child1		
	child2		
2	child1		
3	child1		
	child2		

household. melt (id_vars = 'family')

Exam review, Ex. 13:

```
strings <- c("George Washington: February 22, 1732",  
            "...")
```

Ex 13: want

"George Washington", "Thomas Jefferson", ...

```
str_extract(strings, ".+(?=:)" )
```

↑ anything appears at least once

Positive lookahead (?.=)

Negative lookahead (?!)

Positive lookbehind (?.<=)
↑ fix in pattern

Exercise 14 :

str_extract(strings, "(?L=:).+")

↑
everything after :_

Exercise 15 :

just want word before : (not full name)

e.g. "Washington"

str_extract(strings, "(?L=\\S).+(?=:)")