

# Homework#1

Neshma Simon

9/12/2020

#1: Study Group: Fareha, Stanley, Hertz

## R Markdown

### #3

"Can you get them to roll a 6 more or less often than would be expected? How would you know - what is "  
-It is difficult to detect what can be labeled as "more often" considering how hard it is to get the pr

"Can you write some code in R that will simulate a fair roll?"

```
dice = c("1", "2", "3", "4", "5", "6")
```

```
sample(x = dice, size = 20, replace = TRUE)"
```

```
> sample(x = dice, size = 20, replace = TRUE)
```

```
[1] "1" "6" "6" "1" "1" "5" "6" "3" "2" "6" "4" "5" "3" "6" "4" "4" "3" "4" "3"  
[20] "5"
```

```
> dice = c("1", "2", "3", "4", "5", "6")
```

```
> sample(x = dice, size = 20, replace = TRUE)
```

```
[1] "2" "3" "1" "5" "2" "4" "1" "6" "1" "3" "4" "1" "3" "5" "4" "2" "6" "6" "3"  
[20] "1"
```

```
> dice = c("1", "2", "3", "4", "5", "6")
```

```
> sample(x = dice, size = 20, replace = TRUE)
```

```
[1] "4" "4" "3" "6" "5" "6" "5" "3" "4" "6" "5" "6" "1" "6" "6" "5" "2" "3" "4"  
[20] "3"
```

```
> dice = c("1", "2", "3", "4", "5", "6")
```

```
> sample(x = dice, size = 20, replace = TRUE)
```

```
[1] "1" "6" "5" "4" "3" "4" "3" "5" "6" "4" "2" "4" "2" "5" "1" "1" "6" "6" "6"  
[20] "6"
```

### #4: PUMS DATA

```
load("acs2017_ny_data.RData")
```

```
acs2017_ny[1:10,1:7]
```

```
> acs2017_ny[1:10,1:7]
```

	AGE	female	educ_nohs	educ_hs	educ_somecoll	educ_college	educ_advdeg
1	72	1	0	0	0	0	1
2	72	0	0	0	0	0	1
3	31	0	0	0	0	1	0
4	28	1	0	0	0	1	0
5	54	0	0	0	0	0	1

6	45	1	0	1	0	0	0
7	84	1	0	0	1	0	0
8	71	0	0	0	0	1	0
9	68	1	0	0	1	0	0
10	37	1	1	0	0	0	0

attach(acs2017\_ny)

summary(acs2017\_ny)

AGE	female	educ_nohs	educ_hs
Min. : 0.00	Min. :0.0000	Min. :0.000	Min. :0.0000
1st Qu.:22.00	1st Qu.:0.0000	1st Qu.:0.000	1st Qu.:0.0000
Median :42.00	Median :1.0000	Median :0.000	Median :0.0000
Mean :41.57	Mean :0.5156	Mean :0.271	Mean :0.2804
3rd Qu.:60.00	3rd Qu.:1.0000	3rd Qu.:1.000	3rd Qu.:1.0000
Max. :95.00	Max. :1.0000	Max. :1.000	Max. :1.0000

educ_somecoll	educ_college	educ_advdeg	SCHOOL
Min. :0.000	Min. :0.0000	Min. :0.000	N/A : 5569
1st Qu.:0.000	1st Qu.:0.0000	1st Qu.:0.000	No, not in school:144968
Median :0.000	Median :0.0000	Median :0.000	Yes, in school : 46048
Mean :0.173	Mean :0.1567	Mean :0.119	Missing : 0
3rd Qu.:0.000	3rd Qu.:0.0000	3rd Qu.:0.000	
Max. :1.000	Max. :1.0000	Max. :1.000	

EDUC	
Grade 12	:55119
4 years of college	:30802
5+ years of college	:23385
1 year of college	:19947
Nursery school to grade 4	:14240
2 years of college	:14065
(Other)	:39027

EDUCD	
Regular high school diploma	:35689
Bachelor's degree	:30802
1 or more years of college credit, no degree	:19947
Master's degree	:17010
Associate's degree, type not specified	:14065
Some college, but less than 1 year	: 9086
(Other)	:69986

DEGFIELD	
N/A	:142398
Business	: 9802
Education Administration and Teaching	: 6708
Social Sciences	: 4836
Medical and Health Sciences and Services	: 3919
Fine Arts	: 3491
(Other)	: 25431

DEGFIELD	
N/A	:142398
Psychology	: 2926
Business Management and Administration	: 2501
Accounting	: 2284
General Education	: 2238

English Language and Literature : 2202  
(Other) : 42036

#### DEGFIELD2

N/A :190425  
Business : 972  
Social Sciences : 853  
Education Administration and Teaching: 611  
Fine Arts : 465  
Communications : 352  
(Other) : 2907

#### DEGFIELD2D

N/A :190425  
Psychology : 284  
Economics : 260  
Political Science and Government : 243  
Business Management and Administration : 217  
French, German, Latin and Other Common Foreign Language Studies: 205  
(Other) : 4951

PUMA		GQ	OWNERSHP	OWNERSHPD	MORTGAGE
Min. : 100	Min. :1.000	Min. :0.000	Min. : 0.00	Min. :0.000	
1st Qu.:1500	1st Qu.:1.000	1st Qu.:1.000	1st Qu.:12.00	1st Qu.:0.000	
Median :3201	Median :1.000	Median :1.000	Median :13.00	Median :1.000	
Mean :2713	Mean :1.148	Mean :1.266	Mean :14.95	Mean :1.453	
3rd Qu.:3902	3rd Qu.:1.000	3rd Qu.:2.000	3rd Qu.:22.00	3rd Qu.:3.000	
Max. :4114	Max. :5.000	Max. :2.000	Max. :22.00	Max. :4.000	

OWNCOST		RENT	COSTELEC	COSTGAS	COSTWATR
Min. : 0	Min. : 0	Min. : 0	Min. : 0	Min. : 0	
1st Qu.: 1208	1st Qu.: 0	1st Qu.: 960	1st Qu.: 840	1st Qu.: 320	
Median : 2891	Median : 0	Median :1560	Median :2400	Median :1400	
Mean :38582	Mean : 393	Mean :2311	Mean :5032	Mean :4836	
3rd Qu.:99999	3rd Qu.: 630	3rd Qu.:2520	3rd Qu.:9993	3rd Qu.:9993	
Max. :99999	Max. :3800	Max. :9997	Max. :9997	Max. :9997	

COSTFUEL		HHINCOME	FOODSTMP	LINGISOL
Min. : 0	Min. : -11800	Min. :1.000	Min. :0.000	
1st Qu.:9993	1st Qu.: 41600	1st Qu.:1.000	1st Qu.:1.000	
Median :9993	Median : 81700	Median :1.000	Median :1.000	
Mean :7935	Mean : 114902	Mean :1.147	Mean :1.002	
3rd Qu.:9993	3rd Qu.: 140900	3rd Qu.:1.000	3rd Qu.:1.000	
Max. :9997	Max. :2030000	Max. :2.000	Max. :2.000	

NA's :10630

ROOMS		BUILTYR2	UNITSSTR	FUELHEAT
Min. : 0.000	Min. : 0.000	Min. : 0.00	Min. :0.000	
1st Qu.: 4.000	1st Qu.: 1.000	1st Qu.: 3.00	1st Qu.:2.000	
Median : 6.000	Median : 3.000	Median : 3.00	Median :2.000	
Mean : 5.887	Mean : 3.711	Mean : 4.39	Mean :2.959	
3rd Qu.: 8.000	3rd Qu.: 5.000	3rd Qu.: 6.00	3rd Qu.:4.000	
Max. :16.000	Max. :22.000	Max. :10.00	Max. :9.000	

SSMC		FAMSIZE	NCHILD	NCHLT5
Min. :0.00000	Min. : 1.000	Min. :0.0000	Min. :0.00000	
1st Qu.:0.00000	1st Qu.: 2.000	1st Qu.:0.0000	1st Qu.:0.00000	
Median :0.00000	Median : 3.000	Median :0.0000	Median :0.00000	

Mean :0.01102	Mean : 3.087	Mean :0.5009	Mean :0.08441
3rd Qu.:0.00000	3rd Qu.: 4.000	3rd Qu.:1.0000	3rd Qu.:0.00000
Max. :2.00000	Max. :19.000	Max. :9.0000	Max. :5.00000

RELATE	RELATED	MARST	RACE	RACED
Min. : 1.000	Min. : 101.0	Min. :1.000	Min. :1.00	Min. :100
1st Qu.: 1.000	1st Qu.: 101.0	1st Qu.:1.000	1st Qu.:1.00	1st Qu.:100
Median : 2.000	Median : 201.0	Median :5.000	Median :1.00	Median :100
Mean : 3.307	Mean : 335.6	Mean :3.742	Mean :2.03	Mean :205
3rd Qu.: 3.000	3rd Qu.: 301.0	3rd Qu.:6.000	3rd Qu.:2.00	3rd Qu.:200
Max. :13.000	Max. :1301.0	Max. :6.000	Max. :9.00	Max. :990

HISPAN	HISPAND	BPL
Min. :0.0000	Min. : 0.00	New York :128517
1st Qu.:0.0000	1st Qu.: 0.00	West Indies : 8481
Median :0.0000	Median : 0.00	China : 4964
Mean :0.4153	Mean : 44.75	SOUTH AMERICA: 4957
3rd Qu.:0.0000	3rd Qu.: 0.00	India : 3476
Max. :4.0000	Max. :498.00	Pennsylvania : 3303
		(Other) : 42887

BPLD	ANCESTR1
New York :128517	Not Reported :32021
China : 4116	Italian :20577
Dominican Republic: 3517	Irish, various subheads,:16388
Pennsylvania : 3303	German :12781
New Jersey : 3127	African-American : 9559
Puerto Rico : 2272	United States : 8209
(Other) : 51733	(Other) :97050

ANCESTR1D	ANCESTR2
Not Reported :32021	Not Reported:141487
Italian (1990-2000, ACS, PRCS) :20577	German : 9476
Irish :15651	Irish : 9238
German (1990-2000, ACS/PRCS) :12605	English : 4895
African-American (1990-2000, ACS, PRCS): 9559	Italian : 4531
United States : 8209	Polish : 3113
(Other) :97963	(Other) : 23845

ANCESTR2D	CITIZEN	YRSUSA1
Not Reported :141487	Min. :0.0000	Min. : 0.000
German (1990-2000, ACS, PRCS) : 9441	1st Qu.:0.0000	1st Qu.: 0.000
Irish : 8809	Median :0.0000	Median : 0.000
English : 4895	Mean :0.4793	Mean : 5.377
Italian (1990-2000, ACS, PRCS): 4531	3rd Qu.:0.0000	3rd Qu.: 0.000
Polish : 3113	Max. :3.0000	Max. :92.000
(Other) : 24309		

HCOVANY	HCOVPRIV	SEX	EMPSTAT
Min. :1.000	Min. :1.000	Male : 95222	Min. :0.000
1st Qu.:2.000	1st Qu.:1.000	Female:101363	1st Qu.:1.000
Median :2.000	Median :2.000		Median :1.000
Mean :1.951	Mean :1.691		Mean :1.514
3rd Qu.:2.000	3rd Qu.:2.000		3rd Qu.:3.000
Max. :2.000	Max. :2.000		Max. :3.000

EMPSTATD	LABFORCE	OCC	IND
Min. : 0.00	Min. :0.000	0 : 79987	0 :79987

1st Qu.:10.00	1st Qu.:1.000	2310	:	3494	7860	:	9025
Median :10.00	Median :2.000	5700	:	3235	8680	:	6354
Mean :15.16	Mean :1.331	430	:	3025	770	:	6279
3rd Qu.:30.00	3rd Qu.:2.000	4720	:	2666	8190	:	5873
Max. :30.00	Max. :2.000	4760	:	2563	7870	:	4041
		(Other):101615			(Other):85026		

CLASSWKR	CLASSWKRD	WKSWORK2	UHRSWORK
Min. :0.000	Min. : 0.00	Min. :0.000	Min. : 0.00
1st Qu.:0.000	1st Qu.: 0.00	1st Qu.:0.000	1st Qu.: 0.00
Median :2.000	Median :22.00	Median :1.000	Median :12.00
Mean :1.116	Mean :13.03	Mean :2.701	Mean :19.77
3rd Qu.:2.000	3rd Qu.:22.00	3rd Qu.:6.000	3rd Qu.:40.00
Max. :2.000	Max. :29.00	Max. :6.000	Max. :99.00

INCTOT	FTOTINC	INCWAGE	POVERTY
Min. : -7300	Min. : -11800	Min. : 0	Min. : 0.0
1st Qu.: 8000	1st Qu.: 35550	1st Qu.: 0	1st Qu.:159.0
Median : 25000	Median : 74000	Median : 10000	Median :351.0
Mean : 45245	Mean : 107110	Mean : 33796	Mean :318.7
3rd Qu.: 56500	3rd Qu.: 132438	3rd Qu.: 47000	3rd Qu.:501.0
Max. :1563000	Max. :2030000	Max. :638000	Max. :501.0
NA's :31129	NA's :10817	NA's :33427	

MIGRATE1	MIGRATE1D	MIGPLAC1	MIGCOUNTY1
Min. :0.000	Min. : 0.00	Min. : 0.000	Min. : 0.000
1st Qu.:1.000	1st Qu.:10.00	1st Qu.: 0.000	1st Qu.: 0.000
Median :1.000	Median :10.00	Median : 0.000	Median : 0.000
Mean :1.122	Mean :11.51	Mean : 6.184	Mean : 4.117
3rd Qu.:1.000	3rd Qu.:10.00	3rd Qu.: 0.000	3rd Qu.: 0.000
Max. :4.000	Max. :40.00	Max. :900.000	Max. :810.000

MIGPUMA1	VETSTAT	VETSTATD	PWPUMA00
Min. : 0	Min. :0.0000	Min. : 0.000	Min. : 0
1st Qu.: 0	1st Qu.:1.0000	1st Qu.:11.000	1st Qu.: 0
Median : 0	Median :1.0000	Median :11.000	Median : 0
Mean : 277	Mean :0.8621	Mean : 9.412	Mean : 1255
3rd Qu.: 0	3rd Qu.:1.0000	3rd Qu.:11.000	3rd Qu.: 3100
Max. :70100	Max. :2.0000	Max. :20.000	Max. :59300

TRANWORK	TRANTIME	DEPARTS	in_NYC
Min. : 0.000	Min. : 0.00	Min. : 0.0	Min. :0.0000
1st Qu.: 0.000	1st Qu.: 0.00	1st Qu.: 0.0	1st Qu.:0.0000
Median : 0.000	Median : 0.00	Median : 0.0	Median :0.0000
Mean : 9.725	Mean : 14.75	Mean : 373.3	Mean :0.3615
3rd Qu.:10.000	3rd Qu.: 20.00	3rd Qu.: 732.0	3rd Qu.:1.0000
Max. :70.000	Max. :138.00	Max. :2345.0	Max. :1.0000

in_Bronx	in_Manhattan	in_StatenI	in_Brooklyn
Min. :0.0000	Min. :0.00000	Min. :0.00000	Min. :0.000
1st Qu.:0.0000	1st Qu.:0.00000	1st Qu.:0.00000	1st Qu.:0.000
Median :0.0000	Median :0.00000	Median :0.00000	Median :0.000
Mean :0.0538	Mean :0.04981	Mean :0.02084	Mean :0.126
3rd Qu.:0.0000	3rd Qu.:0.00000	3rd Qu.:0.00000	3rd Qu.:0.000
Max. :1.0000	Max. :1.00000	Max. :1.00000	Max. :1.000

in_Queens	in_Westchester	in_Nassau	Hispanic
Min. :0.0000	Min. :0.00000	Min. :0.00000	Min. :0.0000
1st Qu.:0.0000	1st Qu.:0.00000	1st Qu.:0.00000	1st Qu.:0.0000
Median :0.0000	Median :0.00000	Median :0.00000	Median :0.0000
Mean :0.1111	Mean :0.04413	Mean :0.07032	Mean :0.1387
3rd Qu.:0.0000	3rd Qu.:0.00000	3rd Qu.:0.00000	3rd Qu.:0.0000
Max. :1.0000	Max. :1.00000	Max. :1.00000	Max. :1.0000

  

Hisp_Mex	Hisp_PR	Hisp_Cuban	Hisp_DomR
Min. :0.00000	Min. :0.0000	Min. :0.000000	Min. :0.00000
1st Qu.:0.00000	1st Qu.:0.0000	1st Qu.:0.000000	1st Qu.:0.00000
Median :0.00000	Median :0.0000	Median :0.000000	Median :0.00000
Mean :0.01626	Mean :0.0436	Mean :0.003403	Mean :0.02827
3rd Qu.:0.00000	3rd Qu.:0.0000	3rd Qu.:0.000000	3rd Qu.:0.00000
Max. :1.00000	Max. :1.0000	Max. :1.000000	Max. :1.00000

  

white	AfAm	Amindian	Asian
Min. :0.0000	Min. :0.000	Min. :0.000000	Min. :0.00000
1st Qu.:0.0000	1st Qu.:0.000	1st Qu.:0.000000	1st Qu.:0.00000
Median :1.0000	Median :0.000	Median :0.000000	Median :0.00000
Mean :0.6997	Mean :0.125	Mean :0.003779	Mean :0.08656
3rd Qu.:1.0000	3rd Qu.:0.000	3rd Qu.:0.000000	3rd Qu.:0.00000
Max. :1.0000	Max. :1.000	Max. :1.000000	Max. :1.00000

  

race_oth	unmarried	veteran	has_AnyHealthIns
Min. :0.0000	Min. :0.00	Min. :0.00000	Min. :0.0000
1st Qu.:0.0000	1st Qu.:0.00	1st Qu.:0.00000	1st Qu.:1.0000
Median :0.0000	Median :0.00	Median :0.00000	Median :1.0000
Mean :0.1324	Mean :0.45	Mean :0.04443	Mean :0.9513
3rd Qu.:0.0000	3rd Qu.:1.00	3rd Qu.:0.00000	3rd Qu.:1.0000
Max. :1.0000	Max. :1.00	Max. :1.00000	Max. :1.0000

  

has_PvtHealthIns	Commute_car	Commute_bus	Commute_subway
Min. :0.0000	Min. :0.0000	Min. :0.00000	Min. :0.00000
1st Qu.:0.0000	1st Qu.:0.0000	1st Qu.:0.00000	1st Qu.:0.00000
Median :1.0000	Median :0.0000	Median :0.00000	Median :0.00000
Mean :0.6906	Mean :0.2997	Mean :0.02162	Mean :0.07468
3rd Qu.:1.0000	3rd Qu.:1.0000	3rd Qu.:0.00000	3rd Qu.:0.00000
Max. :1.0000	Max. :1.0000	Max. :1.00000	Max. :1.00000

  

Commute_rail	Commute_other	below_povertyline	below_150poverty
Min. :0.00000	Min. :0.00000	Min. :0.000	Min. :0.0000
1st Qu.:0.00000	1st Qu.:0.00000	1st Qu.:0.000	1st Qu.:0.0000
Median :0.00000	Median :0.00000	Median :0.000	Median :0.0000
Mean :0.01332	Mean :0.05506	Mean :0.122	Mean :0.1965
3rd Qu.:0.00000	3rd Qu.:0.00000	3rd Qu.:0.000	3rd Qu.:0.0000
Max. :1.00000	Max. :1.00000	Max. :1.000	Max. :1.0000

  

below_200poverty	foodstamps
Min. :0.0000	Min. :0.0000
1st Qu.:0.0000	1st Qu.:0.0000
Median :0.0000	Median :0.0000
Mean :0.2676	Mean :0.1465
3rd Qu.:1.0000	3rd Qu.:0.0000

```

Max.      :1.0000    Max.      :1.0000

print(NN_obs <- length(AGE))
[1] 196585

summary(AGE[female == 1])
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  0.00  23.00  44.00  42.72  61.00  95.00
summary(AGE[!female])
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  0.00  21.00  40.00  40.35  59.00  95.00
mean(AGE[female == 1])
[1] 42.71629
sd(AGE[female == 1])
[1] 23.72012
mean(AGE[!female])
[1] 40.35398
sd(AGE[!female])
[1] 23.1098

```

"Tell me something else interesting, that you learned from the data, for example about educational attainment. I found that each neighborhood has a direct correlation to the educational attainments within the city."

#5

"link: <https://finance.yahoo.com/quote/%5EGSPC/history?period1=1514851200&period2=1517529600&interval=1>"

Date	Open	High	Low	Close	Adj Close	Volume
1/2/18	2683.72998	2695.889893	2682.360107	2695.810059	2695.810059	3367250000
1/3/18	2697.850098	2714.370117	2697.77002	2713.060059	2713.060059	3538660000
1/4/18	2719.310059	2729.290039	2719.070068	2723.98999	2723.98999	3695260000
1/5/18	2731.330078	2743.449951	2727.919922	2743.149902	2743.149902	3236620000
1/8/18	2742.669922	2748.51001	2737.600098	2747.709961	2747.709961	3242650000
1/9/18	2751.149902	2759.139893	2747.860107	2751.290039	2751.290039	3453480000
1/10/18	2745.550049	2750.800049	2736.060059	2748.22998	2748.22998	3576350000
1/11/18	2752.969971	2767.560059	2752.780029	2767.560059	2767.560059	3641320000
1/12/18	2770.179932	2787.850098	2769.639893	2786.23999	2786.23999	3573970000
1/16/18	2798.959961	2807.540039	2768.639893	2776.419922	2776.419922	4325970000
1/17/18	2784.98999	2807.040039	2778.379883	2802.560059	2802.560059	3778050000
1/18/18	2802.399902	2805.830078	2792.560059	2798.030029	2798.030029	3681470000
1/19/18	2802.600098	2810.330078	2798.080078	2810.300049	2810.300049	3639430000
1/22/18	2809.159912	2833.030029	2808.120117	2832.969971	2832.969971	3471780000
1/23/18	2835.050049	2842.23999	2830.590088	2839.129883	2839.129883	3519650000
1/24/18	2845.419922	2852.969971	2824.810059	2837.540039	2837.540039	4014070000
1/25/18	2846.23999	2848.560059	2830.939941	2839.25	2839.25	3835150000
1/26/18	2847.47998	2872.870117	2846.179932	2872.870117	2872.870117	3443230000
1/29/18	2867.22998	2870.620117	2851.47998	2853.530029	2853.530029	3573830000
1/30/18	2832.73999	2837.75	2818.27002	2822.429932	2822.429932	3990650000
1/31/18	2832.409912	2839.26001	2813.040039	2823.810059	2823.810059	4261280000
2/1/18	2816.449951	2835.959961	2812.699951	2821.97998	2821.97998	3938450000

Using Excel:

Daily Return	Date
-0.004501227	1/2/18
-0.005637808	1/3/18

-0.001720999	1/4/18
-0.004327497	1/5/18
-0.00183764	1/8/18
-5.09376E-05	1/9/18
-0.0009761	1/10/18
-0.005299763	1/11/18
-0.005797478	1/12/18
0.008053005	1/16/18
-0.006308845	1/17/18
0.001559332	1/18/18
-0.002747431	1/19/18
-0.008475865	1/22/18
-0.001439069	1/23/18
0.002769322	1/24/18
0.002455868	1/25/18
-0.008916704	1/26/18
0.004778114	1/29/18
0.003639606	1/30/18
0.003036232	1/31/18
-0.001963475	2/1/18

Daily Return Mean: -0.001532244

"What is the mean return on days when the previous day's return was positive? When the previous 2 days were positive?"

- The mean return on days where the previous day's return was positive was -0.006308845 on 1/17/18. When the previous 2 days were positive was -0.001963475 on 2/1/18.