

STAT 3675Q Homework 1

Due date: **Thursday, September 4, at noon**

Zeshi Feng

Note:

- Ensure that your code is fully visible in the PDF and not cropped. If needed, break the code into multiple lines to fit.
- It is recommended to write descriptive answers outside of R code chunks (i.e., as text in the main body), while comments within the code chunks can be reserved for brief code annotations.
- In all homework questions, include a written explanation of any output to earn full credit. For example, explain what the histograms show and how they relate to the data.

Question 1 [20 points]

Install and load the R package `usdata`. Print the first six rows of the dataset `county_2019`.

Note: Install R packages in the Console using `install.packages("")` - this command cannot be run in the code chunks in an Rmd file. You only need to install a package once. However, you must load the package with the `library()` command each time you use it. Unlike `install.packages("")`, `library()` can be run in the code chunks in an Rmd file.

Answer:

```
library(usdata)
head(usdata::county_2019, 6)
```

##	state	name	fips	age_over_18	age_over_18_moe	age_over_65
## 1	Alabama	Autauga County	1001	76.2	0.1	15.0
## 2	Alabama	Baldwin County	1003	78.3	NA	20.0
## 3	Alabama	Barbour County	1005	79.1	0.1	18.6
## 4	Alabama	Bibb County	1007	79.4	0.2	15.9
## 5	Alabama	Blount County	1009	76.8	0.1	17.9
## 6	Alabama	Bullock County	1011	79.2	NA	16.0
##	age_over_65_moe	age_over_85	age_over_85_moe	age_under_5	age_under_5_moe	asian
## 1	0.2	1.6	0.3	5.8	0.2	1.0
## 2	0.1	1.9	0.3	5.5	0.1	0.9

## 3	0.1	1.6	0.4	5.3	0.1	0.5
## 4	0.3	2.0	0.7	5.8	0.8	0.1
## 5	0.2	1.8	0.3	5.9	0.1	0.4
## 6	NA	1.7	0.7	5.3	0.1	0.5
##	asian_moe	avg_family_size	avg_family_size_moe	bachelors	bachelors_moe	black
## 1	0.3	3.09	0.07	26.6	2.0	19.0
## 2	0.1	3.24	0.05	31.9	1.2	9.3
## 3	0.1	3.01	0.12	11.6	1.4	47.6
## 4	0.2	3.74	0.24	10.4	2.2	22.3
## 5	0.1	3.33	0.08	13.1	1.3	1.6
## 6	0.6	3.30	0.37	12.1	2.8	74.8
##	black_moe	hispanic	hispanic_moe	household_has_broadband		
## 1	0.5	2.8	NA	80.6		
## 2	0.2	4.6	NA	81.8		
## 3	0.6	4.4	NA	60.5		
## 4	0.5	2.6	NA	69.2		
## 5	0.3	9.3	NA	73.0		
## 6	3.2	2.6	3.1	60.1		
##	household_has_broadband_moe	household_has_computer	household_has_computer_moe			
## 1		1.9	73.0		2.1	
## 2		1.2	76.3		1.5	
## 3		2.6	51.9		3.0	
## 4		4.3	54.7		5.0	
## 5		2.5	63.5		2.3	
## 6		6.8	52.5		8.1	
##	household_has_smartphone	household_has_smartphone_moe	households			
## 1		78.4	1.8	21397		
## 2		81.7	1.1	80930		
## 3		64.2	2.8	9345		
## 4		66.6	4.2	6891		
## 5		70.1	2.5	20847		
## 6		70.6	5.2	3521		
##	households_moe	households_speak_asian_or_pac_isl				
## 1	325		1.8			
## 2	1127		0.6			
## 3	313		0.6			
## 4	333		0.0			
## 5	394		0.1			
## 6	255		0.4			
##	households_speak_asian_or_pac_isl_moe	households_speak_limited_english				
## 1		0.5			0.7	
## 2		0.2			1.2	
## 3		0.4			1.6	
## 4		0.5			0.6	
## 5		0.1			1.8	

## 6	0.6	1.4		
## households_speak_limited_english_moe households_speak_other				
## 1	0.5	0.2		
## 2	0.4	0.0		
## 3	0.8	0.0		
## 4	0.7	0.0		
## 5	0.7	0.2		
## 6	1.6	0.0		
## households_speak_other_moe households_speak_other_indo_euro_lang				
## 1	0.3	0.3		
## 2	0.1	1.8		
## 3	0.1	1.1		
## 4	0.5	0.5		
## 5	0.2	0.9		
## 6	0.9	1.6		
## households_speak_other_indo_euro_lang_moe households_speak_spanish				
## 1	0.2	2.9		
## 2	0.3	4.6		
## 3	0.5	5.2		
## 4	0.5	1.9		
## 5	0.5	6.6		
## 6	1.3	3.6		
## households_speak_spanish_moe housing_mobile_homes housing_mobile_homes_moe				
## 1	0.8	26.7	2.2	
## 2	0.6	24.8	1.1	
## 3	1.3	39.1	2.6	
## 4	1.3	25.6	3.7	
## 5	0.7	21.2	2.0	
## 6	3.0	28.9	6.9	
## housing_one_unit_structures housing_one_unit_structures_moe				
## 1	17.3	1.9		
## 2	11.5	0.8		
## 3	26.1	2.4		
## 4	29.7	3.9		
## 5	24.0	2.0		
## 6	39.3	6.0		
## housing_two_unit_structures housing_two_unit_structures_moe hs_grad				
## 1	73.3	2.2	88.5	
## 2	75.2	1.1	90.8	
## 3	60.9	2.6	73.2	
## 4	74.4	3.7	79.1	
## 5	78.8	2.0	80.5	
## 6	71.1	6.9	74.7	
## hs_grad_moe mean_household_income mean_household_income_moe mean_work_travel				
## 1	1.4	75326	6004	24.4

## 2	0.7	80986	1930	NA	
## 3	1.8	47068	2424	NA	
## 4	3.2	60182	5709	NA	
## 5	1.6	65639	3912	NA	
## 6	4.9	48571	6108	27.2	
##	mean_work_travel_moe	median_age	median_age_moe	median_household_income	
## 1	1.3	38.2	0.6	58731	
## 2	NA	43.0	0.3	58320	
## 3	NA	40.4	0.5	32525	
## 4	NA	40.9	1.3	47542	
## 5	NA	40.7	0.3	49358	
## 6	3.5	40.2	2.3	37785	
##	median_household_income_moe	median_individual_income			
## 1		4410	29725		
## 2		1564	29802		
## 3		2291	17963		
## 4		5504	21958		
## 5		2136	26976		
## 6		12122	21571		
##	median_individual_income_moe	median_individual_income_age_25plus			
## 1		1643	40778		
## 2		905	37897		
## 3		1198	27434		
## 4		1400	28789		
## 5		969	39004		
## 6		1971	29516		
##	median_individual_income_age_25plus_moe	native	native_moe	other_single_race	
## 1		1343	0.3	0.1	0.7
## 2		1119	0.8	0.2	1.1
## 3		2812	0.3	0.2	3.6
## 4		3605	0.1	0.3	0.0
## 5		1955	0.1	0.1	0.9
## 6		2589	0.0	0.3	2.0
##	other_single_race_moe	pac_isl	pac_isl_moe	per_capita_income	
## 1	0.6	0	0.1	29819	
## 2	0.3	0	0.1	32626	
## 3	0.7	0	0.1	18473	
## 4	0.1	0	0.1	20778	
## 5	0.6	0	0.1	24747	
## 6	3.0	0	0.3	20877	
##	per_capita_income_moe	persons_per_household	persons_per_household_moe	pop	
## 1	2345		2.56	0.04	55380
## 2	758		2.59	0.04	212830
## 3	942		2.41	0.07	25361
## 4	1703		2.99	0.14	22493

## 5		1318		2.74		0.05	57681
## 6		2064		2.79		0.20	10248
##	pop_moe	poverty	poverty_moe	poverty_65_and_over	poverty_65_and_over_moe		
## 1	NA	15.2	1.8	8.7		2.1	
## 2	NA	10.4	0.9	7.4		1.4	
## 3	NA	30.7	2.4	16.8		3.1	
## 4	NA	NA	NA	NA		NA	
## 5	NA	13.6	1.7	10.9		2.8	
## 6	NA	NA	NA	NA		NA	
##	poverty_under_18	poverty_under_18_moe	two_plus_races	two_plus_races_moe			
## 1		23.2	4.0	2.2		0.6	
## 2		13.4	2.0	1.7		0.3	
## 3		50.1	4.9	1.2		0.5	
## 4		NA	NA	0.6		0.4	
## 5		18.4	3.8	1.6		0.3	
## 6		NA	NA	0.8		1.0	
##	unemployment_rate	unemployment_rate_moe	uninsured	uninsured_moe			
## 1		3.5	1.0	7.1		1.0	
## 2		4.0	0.6	8.9		0.7	
## 3		9.4	1.9	11.3		1.5	
## 4		7.0	2.9	10.7		2.2	
## 5		3.1	0.9	10.8		1.4	
## 6		4.1	3.4	11.4		4.2	
##	uninsured_65_and_older	uninsured_65_and_older_moe	uninsured_under_19				
## 1		0.0	0.4			1.7	
## 2		0.3	0.3			3.8	
## 3		0.3	0.3			3.3	
## 4		0.0	0.9			2.0	
## 5		0.2	0.2			5.9	
## 6		0.0	2.1			1.0	
##	uninsured_under_19_moe	uninsured_under_6	uninsured_under_6_moe	veterans			
## 1		0.9	1.7	2.3		12.6	
## 2		1.1	2.2	1.3		11.8	
## 3		2.0	3.4	2.4		6.6	
## 4		1.7	4.5	4.4		8.0	
## 5		2.5	6.1	3.2		7.7	
## 6		1.2	2.1	3.6		3.3	
##	veterans_moe	white	white_moe	white_not_hispanic	white_not_hispanic_moe		
## 1		1.3	76.8	0.5	74.6		0.1
## 2		0.6	86.2	0.4	83.1		0.1
## 3		0.8	46.8	0.6	45.8		0.1
## 4		1.4	76.8	0.3	74.5		0.1
## 5		0.8	95.5	0.6	86.9		0.5
## 6		1.4	21.9	0.5	21.4		0.3

Question 2 [20 points]

Simulate 50 random numbers from a standard normal distribution and assign them to `x1`. Simulate another 50 random numbers from a standard uniform distribution and assign them to `x2`. Calculate the following quantities for each sequence:

- Sum
- Minimum
- Maximum
- Median

Hint: You can use functions - `rnorm`, `runif`, `sum`, `min`, `max`, `median`

Answer:

```
x1<-rnorm(50)
x2<-runif(50)
head(x1)

## [1] -0.6683842 -0.3841182 -0.7845324  0.3875331 -0.2329117 -1.3102332
sum(x1)

## [1] -4.321196
min(x1)

## [1] -2.521361
max(x1)

## [1]  1.914087
median(x1)

## [1] -0.004578741
head(x2)

## [1] 0.3016228 0.4845476 0.6422357 0.6946508 0.8766949 0.3350784
sum(x2)

## [1] 25.64169
min(x2)

## [1] 0.0006281382
max(x2)

## [1] 0.9394359
median(x2)
```

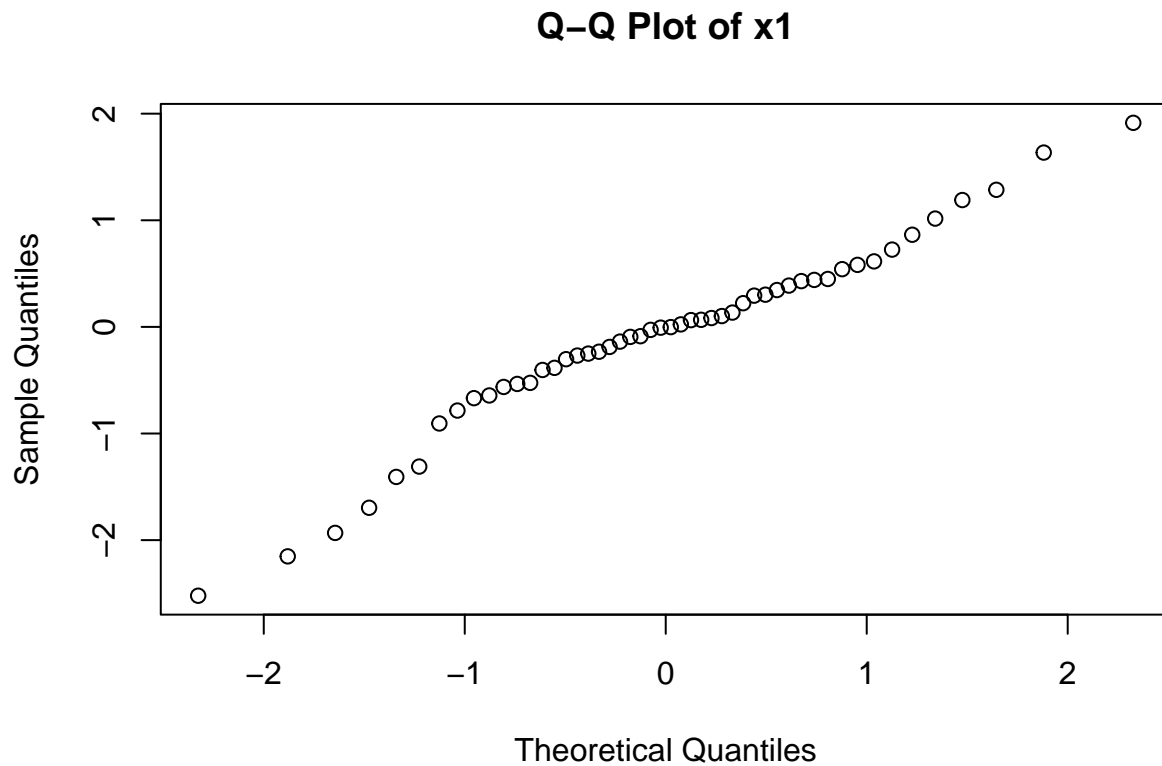
```
## [1] 0.5645845
```

Question 3 [20 points]

Visually assess the normality of x_1 and x_2 in the previous question (e.g. plot histogram).

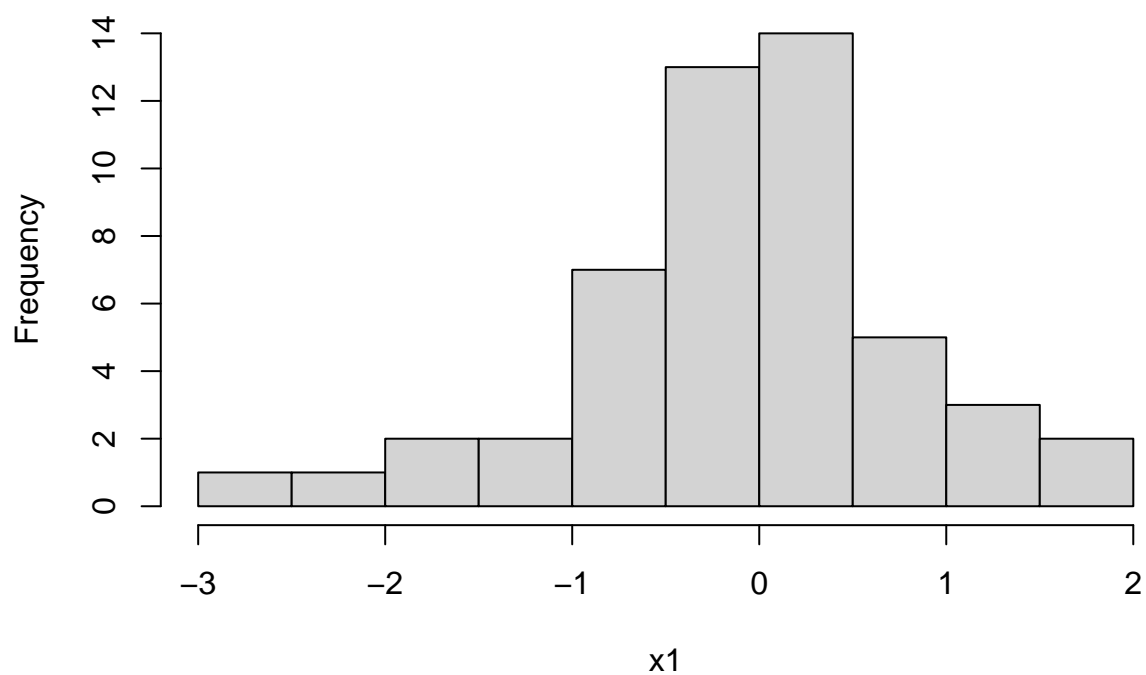
Answer:

```
qqnorm(x1, main = "Q-Q Plot of x1")
```



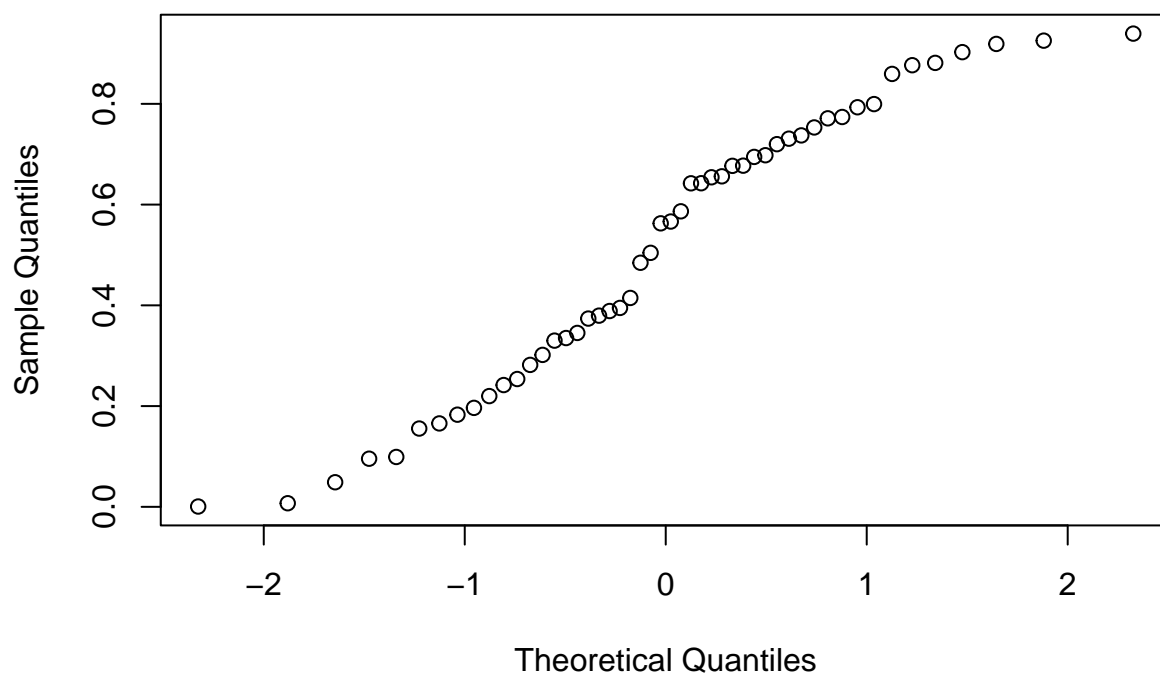
```
hist(x1, main = "Histogram of x1")
```

Histogram of x1

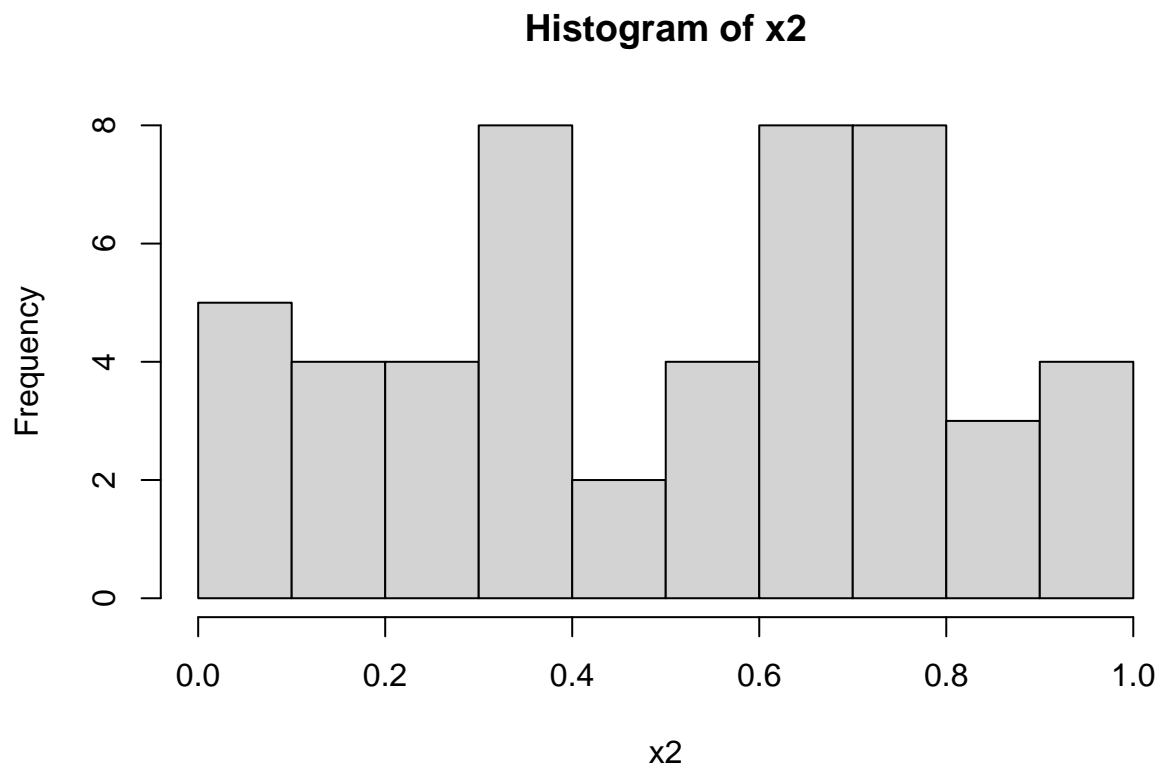


```
qqnorm(x2, main = "Q-Q Plot of x2")
```

Q-Q Plot of x2



```
hist(x2, main = "Histogram of x2")
```

Question 4 [20 points]

Use the `seq` function to define a new variable, `z`, which contains 50 numbers between -5 and 5, and round these numbers to 3 decimal points. (hint: you can use functions - `seq`, `round`)

Answer:

```
z <- seq(-5,5,length=50)
z <- round(z,3)
z
```

```
## [1] -5.000 -4.796 -4.592 -4.388 -4.184 -3.980 -3.776 -3.571 -3.367 -3.163
## [11] -2.959 -2.755 -2.551 -2.347 -2.143 -1.939 -1.735 -1.531 -1.327 -1.122
## [21] -0.918 -0.714 -0.510 -0.306 -0.102  0.102  0.306  0.510  0.714  0.918
## [31]  1.122  1.327  1.531  1.735  1.939  2.143  2.347  2.551  2.755  2.959
## [41]  3.163  3.367  3.571  3.776  3.980  4.184  4.388  4.592  4.796  5.000
```

Question 5 [20 points]

Create a scatterplot with `x1` on the horizontal axis and `z` on the vertical axis.

Answer:

```
plot(x1,z,main="scatterplot of x1 and z",
     xlab="x1",ylab="z")
```

scatterplot of x1 and z

