

# ECON 7201

## Applied Econometrics

### Emmanuel

#### Assignment 1

#### Due Date

September 18, 2025 at the start of class

#### Directions

Answer all questions. Submit both a PDF and Quarto file to the nexus assignment portal.

## 1. Git and GitHub

- (a) Create a GitHub repository called **econ\_3201** and connect it to RStudio.
- (b) Create a new R project in this newly created directory called **assignment\_1**. (Note, you do not have to click “Create git repository” as the directory is contained in a git enabled directory, i.e., **econ\_3201**).
- (c) Download the assignment PDF and Quarto file the **assignment\_1** folder.
- (d) Commit and push the changes to your **econ\_3201** repository on [GitHub.com](https://github.com).

**Ans.** I have created the repository and pushed all the changes to my repository on github.com ([https://github.com/Emmanuel-spec493/econ\\_3201](https://github.com/Emmanuel-spec493/econ_3201)).

## 2. LaTeX

LaTeX is useful for writing math equations and presents them in a neat and orderly way. To write in math mode, wrap your text in \$ for inline text use two \$s for display (i.e., centered on the page). Some very useful functions include:

- Fractions: \frac{}{}, e.g. \\$\frac{1}{2}\\$ gives  $\frac{1}{2}$  and \$\$\$\frac{1}{2}\$\$\$ gives:

$$\frac{1}{2}.$$

- Subscripts: `_` gives a subscript, e.g. `$x_1$` gives  $x_1$ . To include more than one term in the subscript, the items in the subscript must be enclosed by `{}`. E.g. `$x_{1,1}$` gives  $x_{1,1}$ . (Note that `$x_1,1$` gives  $x_{1,1}$ )
- Exponents: `^`, e.g. `$x^2$` gives  $x^2$ . `^` can also be used for superscripts in other math functions, including summations and integrals.
- Aligned: `aligned` neatly aligns multiple lines of an equation. `Align` is useful when writing multiple steps to solving an equation. To use it in Quarto, write `$$\begin{aligned} \dots \end{aligned}$$`. The `&` is used to mark the point where the lines should be aligned. Use `\backslash` at the end of each line E.g. `$$\begin{aligned} x &= 3 + 5 \\ &= 8 \end{aligned}$$`

gives

$$\begin{aligned} x &= 3 + 5 \\ &= 8 \end{aligned}$$

- Summation: `\sum` gives the summation sign, i.e.  $\sum$ . To include subscripts, use `_` and to use superscripts use `^`, e.g. `$\sum_{i=1}^n$` gives  $\sum_{i=1}^n$ , which reads as the sum of  $i$  equals 1 to  $n$ .
- Integral: `\int` gives an integral, i.e.  $\int$ . To place a lower limit use `_` and to place an upper limit, use `^`, e.g. `$\int_a^b$` gives  $\int_a^b$ .
- Greek letters: `\alpha, \beta, \gamma, \Gamma, \delta, \Delta, \epsilon, \varepsilon, \zeta, \eta, \sigma, \theta, \vartheta, \Theta, \iota, \kappa, \lambda, \Lambda, \mu`. (See [https://www.overleaf.com/learn/latex/List\\_of\\_Greek\\_letters\\_and\\_math\\_symbols](https://www.overleaf.com/learn/latex/List_of_Greek_letters_and_math_symbols))
- Accents: `\hat{Y}, \tilde{Y}, \bar{Y}` are examples of accents in math mode. E.g. `\hat{Y}, \tilde{Y}, \bar{Y}` gives  $\hat{Y}, \tilde{Y}, \bar{Y}$ , respectively.
- Text: To include text in your equation, i.e. non italicized text, use `\text{}`, e.g. `$x=2\text{ if } y=1$` gives  $x = 2$  if  $y = 1$ .
- Inequalities: Some mathematical expressions may be written as inequalities, rather than equations. For ‘less than’ and ‘greater than’, you can just use the symbol on your keyboard, i.e. `<` and `>`, respectively. For  $\leq$ , use `\leq` and for  $\geq$ , use `\geq`. An important note is that after writing a command, put a space after the command before writing the next term, otherwise you may get an error. E.g. To write  $a \leq b$ , write `$a\leq b$`, not `$a\leqb$`.

Re-write the following equations in LaTeX.

(a)  $E(Y) = y_1 p_1 + \dots + y_k p_k = \sum_{i=1}^k y_i p_i$

$$(b) \sigma_y = \text{Var}(Y) = E[(Y - \mu_y)^2] = \sum_{i=1}^k (y_i - \mu_y)^2 p_i$$

$$(c) \hat{\beta} = \frac{\sum_{i=1}^n (y_i - \mu_y)(x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

$$(d) P(a \leq Y \leq b) = \int_a^b f_Y(y) dy$$

$$(e) \hat{g}(x) = \frac{\frac{1}{nh} \sum_{i=1}^n y_i k(\frac{x_i - x}{h})}{\frac{1}{nh} \sum_{i=1}^n k(\frac{x_i - x}{h})}$$

### 3. R

#### 3.1. Assignment

Note: When creating variables based on equation, separate each element in the equation with the appropriate arithmetic symbol. E.g., to compute  $x(y - 2)$  in R, you would have to type `x*(y-2)`. `x(y-2)`, with not arithmetic symbol between `x` and the left bracket would result in an error.

- (a) In statistics,  $n$  is often used to denote the sample size. Set the number of observations  $n = 1000$ .

```
n <- 1000
```

- (b) Generate two random variables,  $u_1 \sim U(0, 1)$  and  $u_2 \sim U(0, 1)$  with  $n/2 = 500$  observations. That is, create two variables that follow a uniform distribution between 0 and 1 that each have 500 observations. In R, we can create random uniform variables using the `runif(k, min, max)` function, where `k` is number of observations, `min` is the minimum value, and `max` is the maximum value. The default values for `min` and `max` are 0 and 1, respectively. Type `?runif` into your console to learn more.

```
k <- n/2
u1 <- runif(k, min = 0, max = 1)
u2 <- runif(k, min = 0, max = 1)
```

- (c) Generate two variables  $z_1$  and  $z_2$  that take on the following values:

$$z_1 = \sqrt{-2 \ln(u_1)} \times \cos(2\pi u_2)$$

and

$$z_2 = \sqrt{-2 \ln(u_1)} \times \sin(2\pi u_2).$$

In R,  $\sqrt{-2 \ln(u_1)}$  is computed using `sqrt()`,  $\ln$  is computed using `log()`,  $\cos$  is computed using `cos()`, and  $\sin$  is computed using `sin()`.

```
z1 <- sqrt(-2*log(u1))*cos(2*pi*u2)
z2 <- sqrt(-2*log(u1))*sin(2*pi*u2)
```

(d) Generate a vector  $z = [z_1, z_2]$

```
z <- c(z1, z2)
```

(e) Generate two variables  $\mu$  (spelled mu) and  $\sigma$  (spelled sigma). Set  $\mu = 5$  and  $\sigma = 2$ .

```
mu    <- 5
sigma <- 2
```

(f) Generate a variable  $x = \mu + \sigma \times z$

```
x <- mu + sigma * z
```

(g) Calculate the mean of  $x$ , using `mean()` and the standard deviation of  $x$  using `sd()`

```
mean(x)
```

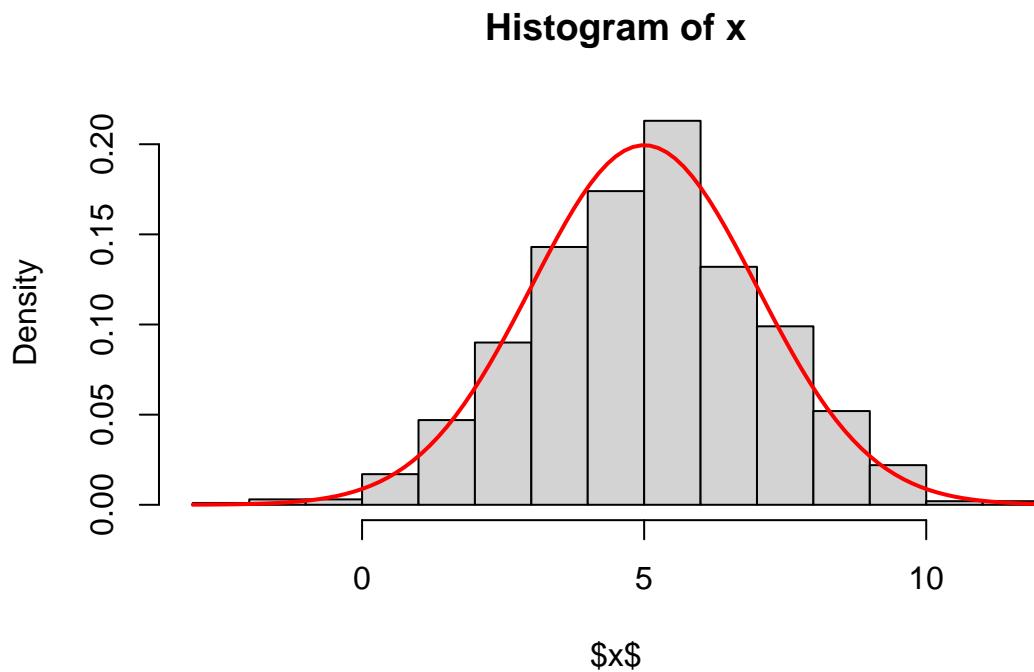
[1] 5.040638

```
sd(x)
```

[1] 2.056363

. (h) Use the following code to plot a histogram of  $x$  with the normal distribution curve.

```
hist(x,freq = FALSE,ylab = "Density", xlab = "$x$")
curve(dnorm(x, mean = mu , sd = sigma ), col = "red" , lwd = 2 , add = TRUE)
```



```
hist(x,
  freq = FALSE,
  ylab = "Density",
  xlab = "$x$")

curve(dnorm(x, mean = mu, sd = sigma),
  col = "red", lwd = 2, add = TRUE)
```

### 3.2. Data frames and Indexing

A data frame in R is a table-like data structure used to store data in rows and columns, similar to a spreadsheet or a database table. It is one of the most commonly used structures for storing datasets in R.

Table 1 displays the total health expenditure by use of funds in Canada from 1975 to 2022. The data is stored in the `data.frame` called `df`.

Table 1: Total health expenditure by use of funds, in millions of current dollars, Canada, 1975 to 2022 (Source: CIHI National Health Expenditure Trends)

Year	Hospitals	Physicians	Other Services	Dental	Vision	Other Professionals
1,975	5,136.77	1,813.15	796.62	56.40	35.86	46.72
1,976	5,977.68	2,041.52	999.08	69.81	40.65	53.92
1,977	6,372.73	2,252.12	1,175.16	83.70	44.86	60.54
1,978	6,861.92	2,528.34	1,367.51	103.96	51.91	75.52
1,979	7,487.62	2,804.48	1,581.37	143.83	57.99	88.88
1,980	8,585.16	3,235.98	1,821.48	194.94	67.23	104.90
1,981	10,127.35	3,775.12	2,146.66	278.44	78.74	126.67
1,982	12,001.93	4,353.14	2,531.36	270.04	91.13	143.01
1,983	13,174.55	4,973.30	2,794.37	260.66	105.68	163.99
1,984	13,936.30	5,444.58	2,923.26	266.74	117.66	181.02
1,985	14,737.75	5,962.06	3,066.46	275.52	130.42	214.58
1,986	15,937.05	6,597.89	2,982.43	287.16	146.05	260.66
1,987	17,154.21	7,266.23	3,132.08	286.27	157.30	276.36
1,988	18,497.17	7,862.51	3,468.29	311.35	180.78	296.02
1,989	20,268.98	8,422.71	3,828.51	350.27	205.62	341.53
1,990	20,528.15	9,090.92	5,100.45	371.70	235.89	379.81
1,991	21,783.23	10,014.44	5,868.30	387.93	265.51	442.89
1,992	22,652.40	10,249.61	6,253.82	394.80	262.22	470.54
1,993	22,619.06	10,306.29	6,190.38	407.31	229.69	460.64
1,994	22,096.82	10,533.27	6,266.36	418.63	221.20	429.23
1,995	21,849.46	10,506.52	6,498.12	408.13	197.12	427.63
1,996	21,997.29	10,651.80	6,591.26	373.98	196.90	426.18
1,997	22,307.52	11,103.52	6,834.19	365.18	215.12	448.14
1,998	23,530.41	11,627.85	7,172.47	352.30	204.66	481.07
1,999	24,751.97	12,255.39	7,578.69	380.04	219.28	523.72
2,000	26,950.76	13,045.53	8,170.94	397.63	230.47	577.24
2,001	28,606.54	14,001.53	8,784.35	406.72	247.80	559.25
2,002	30,683.55	14,939.47	9,308.19	421.57	239.86	521.36
2,003	32,903.18	16,084.37	9,841.96	409.33	244.00	526.93
2,004	35,269.82	17,084.00	10,629.24	425.19	250.30	530.73
2,005	37,112.35	18,302.66	11,064.58	450.38	223.05	469.67
2,006	39,704.71	19,743.14	11,593.52	504.41	231.54	482.76
2,007	42,376.77	21,308.72	12,192.52	541.84	239.84	541.96
2,008	45,362.04	23,370.83	12,809.06	586.77	264.34	619.50
2,009	47,996.52	25,249.61	13,578.95	664.37	295.77	671.40
2,010	50,947.81	27,107.23	14,316.45	714.70	311.87	692.20
2,011	52,126.35	28,813.05	15,324.80	721.61	332.69	734.94

Table 1: Total health expenditure by use of funds, in millions of current dollars, Canada, 1975 to 2022 (Source: CIHI National Health Expenditure Trends)

Year	Hospitals	Physicians	Other Services	Dental	Vision	Other Professionals
2,012	53,299.96	29,801.63	15,923.80	759.13	353.62	782.67
2,013	54,954.28	31,202.28	16,386.15	762.36	358.08	730.08
2,014	56,123.22	32,490.79	16,966.03	782.00	389.71	685.88
2,015	57,352.33	33,886.08	18,313.73	821.42	430.46	1,179.18
2,016	58,168.97	35,283.98	18,809.91	875.86	461.42	1,355.90
2,017	60,356.12	36,490.87	19,665.65	918.62	484.33	1,491.51
2,018	62,896.86	37,494.64	20,548.31	961.17	517.89	1,614.12
2,019	65,034.33	38,914.04	21,446.58	1,018.36	557.19	1,729.01
2,020	67,221.53	37,288.46	23,675.08	896.76	513.22	1,711.94
2,021	69,663.71	41,479.50	25,678.66	922.86	559.07	1,906.92
2,022	73,778.17	44,195.30	28,095.86	991.82	584.06	2,047.50

- (a) Determine if there are any missing values for the variable Hospitals.

Import and set dataframe

```
read.csv("hlthexp.csv")
```

	Year	Hospitals	Other.Institutions	Physicians
1	1975	5136.77	796.62	1813.15
2	1976	5977.68	999.08	2041.52
3	1977	6372.73	1175.16	2252.12
4	1978	6861.92	1367.51	2528.34
5	1979	7487.62	1581.37	2804.48
6	1980	8585.16	1821.48	3235.98
7	1981	10127.35	2146.66	3775.12
8	1982	12001.93	2531.36	4353.14
9	1983	13174.55	2794.37	4973.30
10	1984	13936.30	2923.26	5444.58
11	1985	14737.75	3066.46	5962.06
12	1986	15937.05	2982.43	6597.89
13	1987	17154.21	3132.08	7266.23
14	1988	18497.17	3468.29	7862.51
15	1989	20268.98	3828.51	8422.71
16	1990	20528.15	5100.45	9090.92
17	1991	21783.23	5868.30	10014.44
18	1992	22652.40	6253.82	10249.61
19	1993	22619.06	6190.38	10306.29

20	1994	22096.82	6266.36	10533.27
21	1995	21849.46	6498.12	10506.52
22	1996	21997.29	6591.26	10651.80
23	1997	22307.52	6834.19	11103.52
24	1998	23530.41	7172.47	11627.85
25	1999	24751.97	7578.69	12255.39
26	2000	26950.76	8170.94	13045.53
27	2001	28606.54	8784.35	14001.53
28	2002	30683.55	9308.19	14939.47
29	2003	32903.18	9841.96	16084.37
30	2004	35269.82	10629.24	17084.00
31	2005	37112.35	11064.58	18302.66
32	2006	39704.71	11593.52	19743.14
33	2007	42376.77	12192.52	21308.72
34	2008	45362.04	12809.06	23370.83
35	2009	47996.52	13578.95	25249.61
36	2010	50947.81	14316.45	27107.23
37	2011	52126.35	15324.80	28813.05
38	2012	53299.96	15923.80	29801.63
39	2013	54954.28	16386.15	31202.28
40	2014	56123.22	16966.03	32490.79
41	2015	57352.33	18313.73	33886.08
42	2016	58168.97	18809.91	35283.98
43	2017	60356.12	19665.65	36490.87
44	2018	62896.86	20548.31	37494.64
45	2019	65034.33	21446.58	38914.04
46	2020	67221.53	23675.08	37288.46
47	2021	69663.71	25678.66	41479.50
48	2022	73778.17	28095.86	44195.30

**Other.Professionals..Dental.Services**

1		56.40
2		69.81
3		83.70
4		103.96
5		143.83
6		194.94
7		278.44
8		270.04
9		260.66
10		266.74
11		275.52
12		287.16
13		286.27

14	311.35
15	350.27
16	371.70
17	387.93
18	394.80
19	407.31
20	418.63
21	408.13
22	373.98
23	365.18
24	352.30
25	380.04
26	397.63
27	406.72
28	421.57
29	409.33
30	425.19
31	450.38
32	504.41
33	541.84
34	586.77
35	664.37
36	714.70
37	721.61
38	759.13
39	762.36
40	782.00
41	821.42
42	875.86
43	918.62
44	961.17
45	1018.36
46	896.76
47	922.86
48	991.82

**Other.Professionals..Vision.Care.Services**

1	35.86
2	40.65
3	44.86
4	51.91
5	57.99
6	67.23
7	78.74

8	91.13
9	105.68
10	117.66
11	130.42
12	146.05
13	157.30
14	180.78
15	205.62
16	235.89
17	265.51
18	262.22
19	229.69
20	221.20
21	197.12
22	196.90
23	215.12
24	204.66
25	219.28
26	230.47
27	247.80
28	239.86
29	244.00
30	250.30
31	223.05
32	231.54
33	239.84
34	264.34
35	295.77
36	311.87
37	332.69
38	353.62
39	358.08
40	389.71
41	430.46
42	461.42
43	484.33
44	517.89
45	557.19
46	513.22
47	559.07
48	584.06
Other.Professionals..Other.Services	Total.Other.Professionals
1	46.72
	138.97

2	53.92	164.38
3	60.54	189.10
4	75.52	231.39
5	88.88	290.70
6	104.90	367.07
7	126.67	483.86
8	143.01	504.18
9	163.99	530.33
10	181.02	565.43
11	214.58	620.52
12	260.66	693.87
13	276.36	719.93
14	296.02	788.15
15	341.53	897.42
16	379.81	987.39
17	442.89	1096.32
18	470.54	1127.56
19	460.64	1097.64
20	429.23	1069.06
21	427.63	1032.89
22	426.18	997.07
23	448.14	1028.44
24	481.07	1038.02
25	523.72	1123.04
26	577.24	1205.34
27	559.25	1213.77
28	521.36	1182.80
29	526.93	1180.26
30	530.73	1206.22
31	469.67	1143.09
32	482.76	1218.70
33	541.96	1323.63
34	619.50	1470.60
35	671.40	1631.54
36	692.20	1718.77
37	734.94	1789.25
38	782.67	1895.41
39	730.08	1850.53
40	685.88	1857.59
41	1179.18	2431.06
42	1355.90	2693.18
43	1491.51	2894.46
44	1614.12	3093.17

45		1729.01		3304.56
46		1711.94		3121.92
47		1906.92		3388.85
48		2047.50		3623.38
	Prescribed.Drugs	Non.Prescribed.Drugs	Total.Drugs	Public.Health
1	158.56	-	158.56	406.82
2	215.84	-	215.84	510.61
3	266.56	-	266.56	598.51
4	327.94	-	327.94	615.40
5	386.41	-	386.41	706.99
6	465.01	-	465.01	845.37
7	566.86	-	566.86	1006.11
8	683.59	-	683.59	1149.70
9	817.37	-	817.37	1241.42
10	942.85	-	942.85	1376.49
11	1118.71	-	1118.71	1515.72
12	1319.21	-	1319.21	1570.16
13	1492.70	-	1492.70	1660.98
14	1702.82	-	1702.82	1786.57
15	1970.86	-	1970.86	1953.59
16	2277.77	-	2277.77	2231.40
17	2604.52	-	2604.52	2415.81
18	2904.88	-	2904.88	2578.79
19	3041.61	-	3041.61	2691.54
20	3086.62	-	3086.62	2970.96
21	3366.17	-	3366.17	3136.73
22	3331.27	-	3331.27	3277.75
23	3598.78	-	3598.78	3420.91
24	4007.53	-	4007.53	4295.33
25	4551.32	-	4551.32	4419.15
26	5294.63	-	5294.63	4896.32
27	6070.27	-	6070.27	5601.75
28	6815.61	-	6815.61	5814.63
29	7574.10	-	7574.10	6871.51
30	8271.07	-	8271.07	6980.61
31	8922.40	-	8922.40	7818.47
32	9594.93	-	9594.93	8470.97
33	10145.06	-	10145.06	9058.93
34	10739.76	-	10739.76	9489.80
35	11492.08	-	11492.08	9883.11
36	11734.46	-	11734.46	10108.76
37	11986.16	-	11986.16	10566.04
38	12114.49	-	12114.49	10750.86

39	12199.19	-	12199.19	10567.88
40	12668.45	-	12668.45	10886.02
41	13298.98	-	13298.98	11082.10
42	13616.80	-	13616.80	11455.54
43	13957.25	-	13957.25	11936.20
44	14442.70	-	14442.70	12766.36
45	14939.93	-	14939.93	13974.93
46	15435.35	-	15435.35	15719.83
47	16034.55	-	16034.55	15636.37
48	17094.52	-	17094.52	17379.34
Administration Other.Health.Spending..Health.Research..HR.				
1	271.37			70.32
2	299.90			75.93
3	318.15			93.08
4	320.57			107.24
5	351.71			123.41
6	395.20			139.52
7	445.02			164.74
8	546.61			184.22
9	590.93			215.15
10	647.34			245.66
11	696.81			278.98
12	734.18			334.01
13	779.65			316.88
14	799.25			343.86
15	864.79			393.08
16	1023.89			455.52
17	1043.13			466.84
18	1058.37			557.06
19	1072.31			530.37
20	1163.46			518.37
21	1173.19			478.67
22	1154.56			467.28
23	1217.02			716.09
24	1344.01			733.90
25	1450.94			693.29
26	1596.61			873.49
27	1738.57			1266.20
28	1856.43			1186.34
29	2022.55			1325.15
30	2144.89			1397.72
31	2350.11			1452.07
32	2412.03			1718.76

33	2573.03	1918.43
34	2453.61	2403.01
35	2464.38	2335.09
36	2428.08	2147.92
37	2438.49	2086.25
38	2554.53	2139.34
39	2593.25	2084.83
40	2578.07	2210.43
41	2543.93	2132.24
42	2623.96	2231.28
43	2916.68	2341.75
44	3069.68	2324.61
45	3128.33	2142.36
46	3556.53	2685.25
47	3842.35	3124.68
48	3824.95	3244.05
<b>Other..Health..Spending..Net.of.HR Total..Other..Health..Spending</b>		
1	131.25	201.57
2	165.28	241.21
3	193.73	286.80
4	225.38	332.62
5	271.77	395.18
6	351.77	491.29
7	494.97	659.70
8	586.42	770.64
9	677.37	892.51
10	735.10	980.76
11	861.62	1140.60
12	1009.08	1343.09
13	1141.98	1458.87
14	1364.34	1708.20
15	1609.53	2002.61
16	2080.01	2535.53
17	2552.12	3018.96
18	2664.33	3221.38
19	2741.61	3271.98
20	2937.31	3455.67
21	3001.57	3480.24
22	3017.59	3484.87
23	3224.33	3940.42
24	3545.88	4279.78
25	3837.33	4530.62
26	4107.23	4980.72

27		4283.17		5549.36
28		4422.01		5608.35
29		4581.51		5906.66
30		4658.77		6056.48
31		5003.80		6455.87
32		5340.10		7058.85
33		5808.31		7726.74
34		6471.01		8874.02
35		7190.15		9525.25
36		7676.59		9824.51
37		8086.68		10172.93
38		11050.84		13190.18
39		12462.07		14546.89
40		13006.36		15216.79
41		14310.55		16442.79
42		15233.60		17464.87
43		16100.30		18442.05
44		17073.89		19398.50
45		18361.56		20503.92
46		19109.42		21794.67
47		19686.90		22811.58
48		21416.88		24660.93
	COVID.19.Response.Funding	Sub.Total	Capital	Total
1		- 8923.83	376.43	9300.26
2		- 10450.22	366.97	10817.19
3		- 11459.12	385.44	11844.56
4		- 12585.69	454.82	13040.51
5		- 14004.45	547.82	14552.28
6		- 16206.55	635.29	16841.85
7		- 19210.68	731.91	19942.59
8		- 22541.15	905.70	23446.84
9		- 25014.77	1065.24	26080.01
10		- 26817.00	1139.87	27956.88
11		- 28858.63	1236.88	30095.51
12		- 31177.88	1351.54	32529.43
13		- 33664.64	1391.06	35055.70
14		- 36612.96	1550.70	38163.67
15		- 40209.47	1702.41	41911.88
16		- 43775.50	1740.30	45515.80
17		- 47844.72	1685.79	49530.51
18		- 50046.81	1694.87	51741.69
19		- 50290.81	1650.22	51941.03
20		- 50642.23	1943.17	52585.41

```

21          - 51043.31 1823.98 52867.29
22          - 51485.86 1685.94 53171.80
23          - 53450.81 1757.65 55208.46
24          - 57295.40 1873.99 59169.39
25          - 60661.12 2626.82 63287.94
26          - 66140.85 3134.92 69275.77
27          - 71566.15 3419.91 74986.05
28          - 76209.02 3700.54 79909.56
29          - 82384.59 4181.48 86566.07
30          - 87642.34 4412.08 92054.42
31          - 93169.54 5116.89 98286.43
32          - 99796.87 5185.97 104982.84
33          - 106705.40 7604.44 114309.83
34          - 114569.74 6497.34 121067.08
35          - 121821.46 7120.53 128941.99
36          - 128186.07 7842.85 136028.92
37          - 133217.06 7859.91 141076.97
38          - 139530.84 8047.84 147578.68
39          - 144300.44 7616.70 151917.14
40          - 148786.96 7041.13 155828.09
41          - 155350.99 7062.94 162413.93
42          - 160117.21 6639.63 166756.84
43          - 166659.28 6402.95 173062.23
44          - 173710.23 6721.83 180432.06
45          - 181246.61 7152.76 188399.37
46      29271.69 217085.05 8803.58 225888.63
47      32546.92 231082.49 9514.69 240597.18
48      14549.53 227201.99 10415.66 237617.65

```

```
df <- (read.csv("hlthexp.csv"))
```

```
any(is.na(df$Hospitals))
```

[1] FALSE

(b) Add a variable called “Total Other Services” to the data frame df, where

Total Other Services = Dental + Vision + Other Professional.

To add a newly created variable to a data frame use the syntax `dataframe$varname <- expression`.

```
df$TotalOtherServices <- df$"Other.Professionals..Dental.Services" +
  df$"Other.Professionals..Vision.Care.Services" + df$"Total.Other.Professionals"
```

- (a) Are there any years for which Total Other Professionals
- (b) Another way to add a variable to a data frame is to simply create a new data frame and append the new variable to it. Note: we can use the same data frame name. I.e., `df<-data.frame(df,newvarname = newvar)`. Add the variable “Prescription Drugs” to the `df` data frame using the append method, where prescription drugs is named “Prescribed.Drugs” in the `cihi` data.frame.

```
df <- data.frame(df, Prescription.Drugs = df$Prescribed.Drugs)
```

- (a) Using a single R command, determine the expenditure on hospitals in 1983.

```
df$Hospitals[df$Year == 1983]
```

[1] 13174.55

- (a) Using a singe R command, list the expenditures by year for 2012-2022.

```
df[df$Year %in% 2012:2022, ]
```

	Year	Hospitals	Other.Institutions	Physicians
38	2012	53299.96	15923.80	29801.63
39	2013	54954.28	16386.15	31202.28
40	2014	56123.22	16966.03	32490.79
41	2015	57352.33	18313.73	33886.08
42	2016	58168.97	18809.91	35283.98
43	2017	60356.12	19665.65	36490.87
44	2018	62896.86	20548.31	37494.64
45	2019	65034.33	21446.58	38914.04
46	2020	67221.53	23675.08	37288.46
47	2021	69663.71	25678.66	41479.50
48	2022	73778.17	28095.86	44195.30
			Other.Professionals..Dental.Services	
38			759.13	
39			762.36	
40			782.00	
41			821.42	
42			875.86	
43			918.62	

44		961.17		
45		1018.36		
46		896.76		
47		922.86		
48		991.82		
	Other.Professionals..Vision.Care.Services			
38		353.62		
39		358.08		
40		389.71		
41		430.46		
42		461.42		
43		484.33		
44		517.89		
45		557.19		
46		513.22		
47		559.07		
48		584.06		
	Other.Professionals..Other.Services	Total.Other.Professionals		
38		782.67	1895.41	
39		730.08	1850.53	
40		685.88	1857.59	
41		1179.18	2431.06	
42		1355.90	2693.18	
43		1491.51	2894.46	
44		1614.12	3093.17	
45		1729.01	3304.56	
46		1711.94	3121.92	
47		1906.92	3388.85	
48		2047.50	3623.38	
	Prescribed.Drugs	Non.Prescribed.Drugs	Total.Drugs	Public.Health
38	12114.49	-	12114.49	10750.86
39	12199.19	-	12199.19	10567.88
40	12668.45	-	12668.45	10886.02
41	13298.98	-	13298.98	11082.10
42	13616.80	-	13616.80	11455.54
43	13957.25	-	13957.25	11936.20
44	14442.70	-	14442.70	12766.36
45	14939.93	-	14939.93	13974.93
46	15435.35	-	15435.35	15719.83
47	16034.55	-	16034.55	15636.37
48	17094.52	-	17094.52	17379.34
	Administration	Other.Health.Spending..Health.Research..HR.		
38	2554.53		2139.34	

39	2593.25	2084.83
40	2578.07	2210.43
41	2543.93	2132.24
42	2623.96	2231.28
43	2916.68	2341.75
44	3069.68	2324.61
45	3128.33	2142.36
46	3556.53	2685.25
47	3842.35	3124.68
48	3824.95	3244.05

	Other..Health..Spending	Net..of..HR	Total..Other..Health..Spending
38		11050.84	13190.18
39		12462.07	14546.89
40		13006.36	15216.79
41		14310.55	16442.79
42		15233.60	17464.87
43		16100.30	18442.05
44		17073.89	19398.50
45		18361.56	20503.92
46		19109.42	21794.67
47		19686.90	22811.58
48		21416.88	24660.93

	COVID.19.Response.Funding	Sub.Total	Capital	Total	TotalOtherServices
38	-	139530.8	8047.84	147578.7	3008.16
39	-	144300.4	7616.70	151917.1	2970.97
40	-	148787.0	7041.13	155828.1	3029.30
41	-	155351.0	7062.94	162413.9	3682.94
42	-	160117.2	6639.63	166756.8	4030.46
43	-	166659.3	6402.95	173062.2	4297.41
44	-	173710.2	6721.83	180432.1	4572.23
45	-	181246.6	7152.76	188399.4	4880.11
46		29271.69	217085.0	8803.58	225888.6
47		32546.92	231082.5	9514.69	240597.2
48		14549.53	227202.0	10415.66	237617.6

	Prescription.Drugs
38	12114.49
39	12199.19
40	12668.45
41	13298.98
42	13616.80
43	13957.25
44	14442.70
45	14939.93

```
46      15435.35
47      16034.55
48      17094.52
```

### 3.3 Other useful R commands.

Load the `mpg` dataset from the `ggplot2` package using `mpg <- ggplot2::mpg`. (Be sure to install the `ggplot2` package before you start.)

```
library(ggplot2)
mpg <- ggplot2 ::mpg
```

Subset the data to include only observations from 2008. Search `?subset` in the console.

```
mpg_2008 <- subset(mpg, year == 2008)
```

- (a) Calculate the maximum and minimum miles per gallon in city limits (`cty`). Search `?min` in the console.

```
min(mpg_2008$cty, na.rm = TRUE)
```

```
[1] 9
```

```
max(mpg_2008$cty, na.rm = TRUE)
```

```
[1] 28
```

- (b) Estimate the average miles per gallon within city limits for cars produced in 2008 using the formula

$$\text{Average mpg} = \frac{\sum_{i=1}^n \text{cty}_i}{n}$$

Recall that  $n$  is the number of observations. Search `?length` in the console.

```
sum(mpg_2008$cty) / length(mpg_2008$cty)
```

```
[1] 16.70085
```

- (c) Estimate the average miles per gallon within city limits for cars produced in 2008 using the `mean()` function.

```
mean(mpg_2008$cty)
```

[1] 16.70085

- (d) Create a variable called `compact`, which takes a value of 1 if the vehicle is a compact and 0 otherwise. Search `?ifelse` in the console.

```
library(ggplot2)
df <- mpg
df_2008 <- subset(df, year == 2008)
df$compact <- ifelse(df$class == "compact", 1, 0)
```

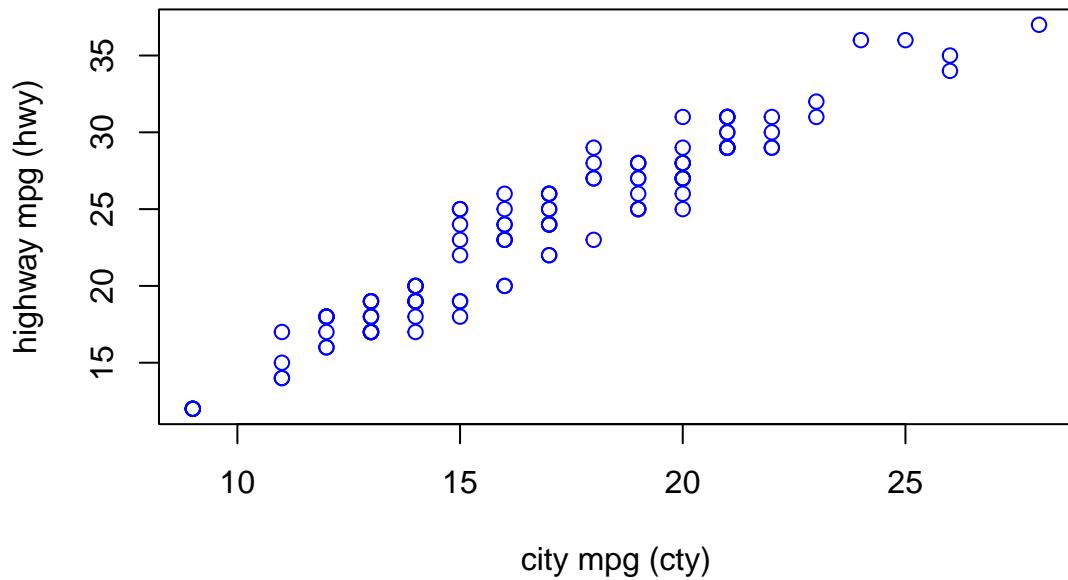
- (e) Estimate the average miles per gallon within city limits for compact cars. (You may use whichever method you prefer).

```
mean(df$compact)
```

[1] 0.2008547

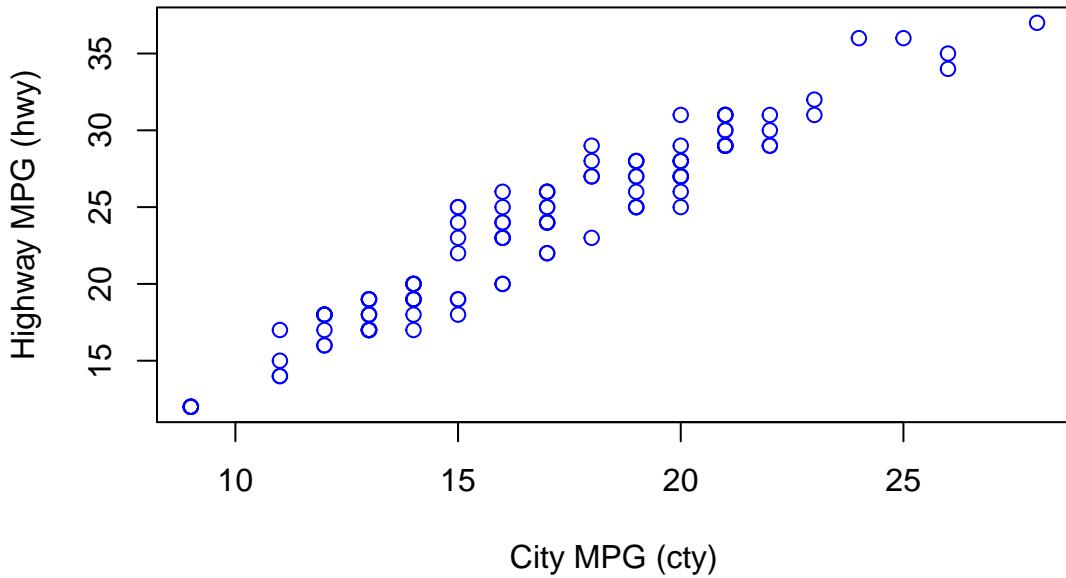
- (f) Create a simple scatter plot with city mpg (`cty`) on the x-axis and highway mpg (`hwy`) on the y-axis. Search `?plot` and choose “Generic X-Y Plotting”.

```
plot(df_2008$cty, df_2008$hwy, xlab = "city mpg (cty)", ylab = "highway mpg (hwy)", col = "black")
```



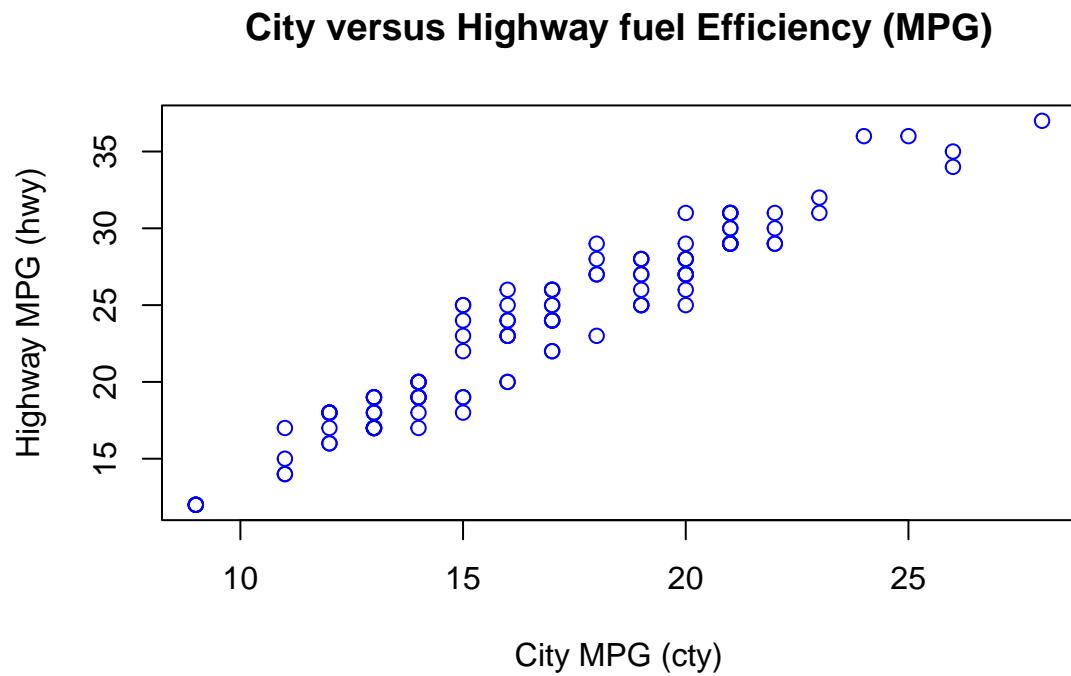
- (i) Change the x-axis label using the option `xlab = "City MPG"` and change the y-axis label using the option `ylab = "Highway MPG"`.

```
plot(df_2008$cty, df_2008$hwy, xlab = "City MPG (cty)" , ylab = "Highway MPG (hwy)" , col =
```



- (ii) Add the caption “City Versus Highway Fuel Efficiency (MPG)”

```
plot(df_2008$cty, df_2008$hwy,  
      xlab = "City MPG (cty)",  
      ylab = "Highway MPG (hwy)",  
      main = "City versus Highway fuel Efficiency (MPG)", ,col = "blue")
```



- (iii) Cross reference the figure and add the text “Figure 1 shows the fuel efficiency for city driving versus highway driving”

```
plot(df_2008$cty, df_2008$hwy,
      xlab = "City MPG (cty)",
      ylab = "Highway MPG (hwy)",
      main = "City versus Highway Fuel Efficiency (MPG)",
      col = "blue")
```

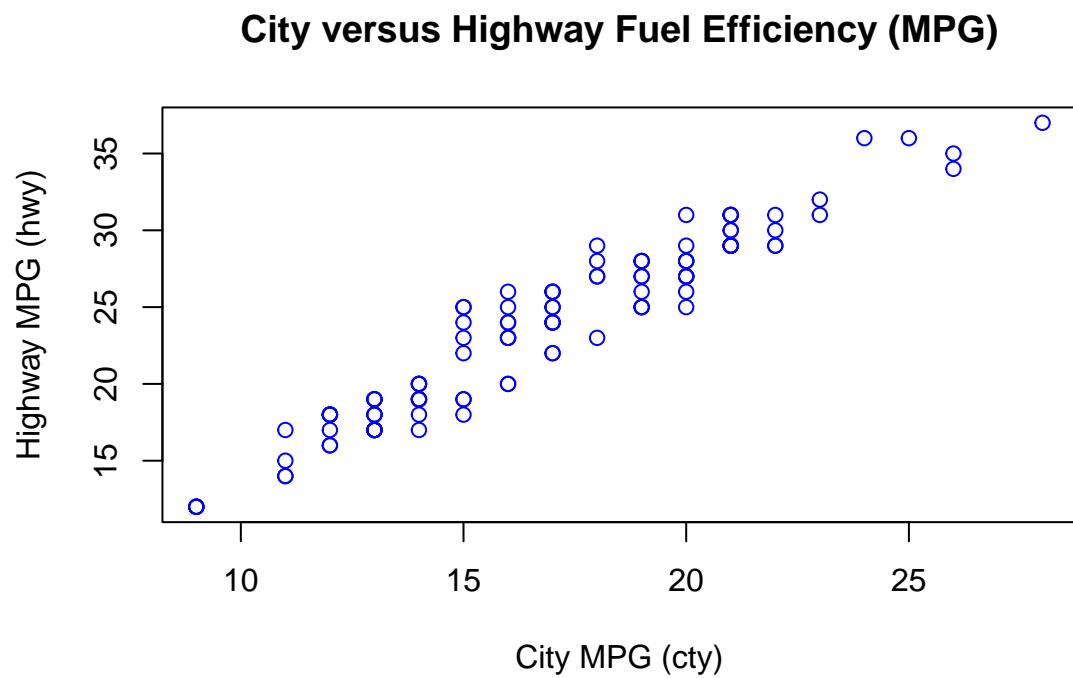


Figure 1

Figure 1 shows the fuel efficiency for city driving versus highway driving”