CEE110
Homework #1 Solution

Problem 1.

a.

Gender	Total number of data (n)	Mean	Standard Deviation	Minimum	Maximum
Female	83	23.72	44.95	0	256
Male	60	12.95	22.20	0	84

Mean was derived by adding all the data values divided by n
S.D. is a sample standard deviation, divided by (n-1)
Min and Max are the smallest and the largest values in the data set, respectively.

b.

Gender	Q1	Q2	Q3	f_s(=IQR)	Min whisker	Max whisker
Female	1	5	21.5	20.5	0	51
Male	0	3	13.5	13.5	0	25

Q2

- Female: Kate, 1985
- Male: Average of the two middle ones (Able 1952, Jerry 1989)

Q1

- Female: average of the two middle values (Dolly 2008 and Hermine 2016), including Q2 in the first half, (1+1)/2
- Male: average of the two middle values (Pable 2019 and Omar 2020), including one Q2 in the first half, (0+0)/2
-

Q3

- Female: average of the two middle values (Alicia 1983 and Gracie 1959), including Q2 in the last half, (21+22)/2
- Male: average of the two middle values (Juan, 1985 and David 1979), including one Q2 in the last half, (12+15)/2

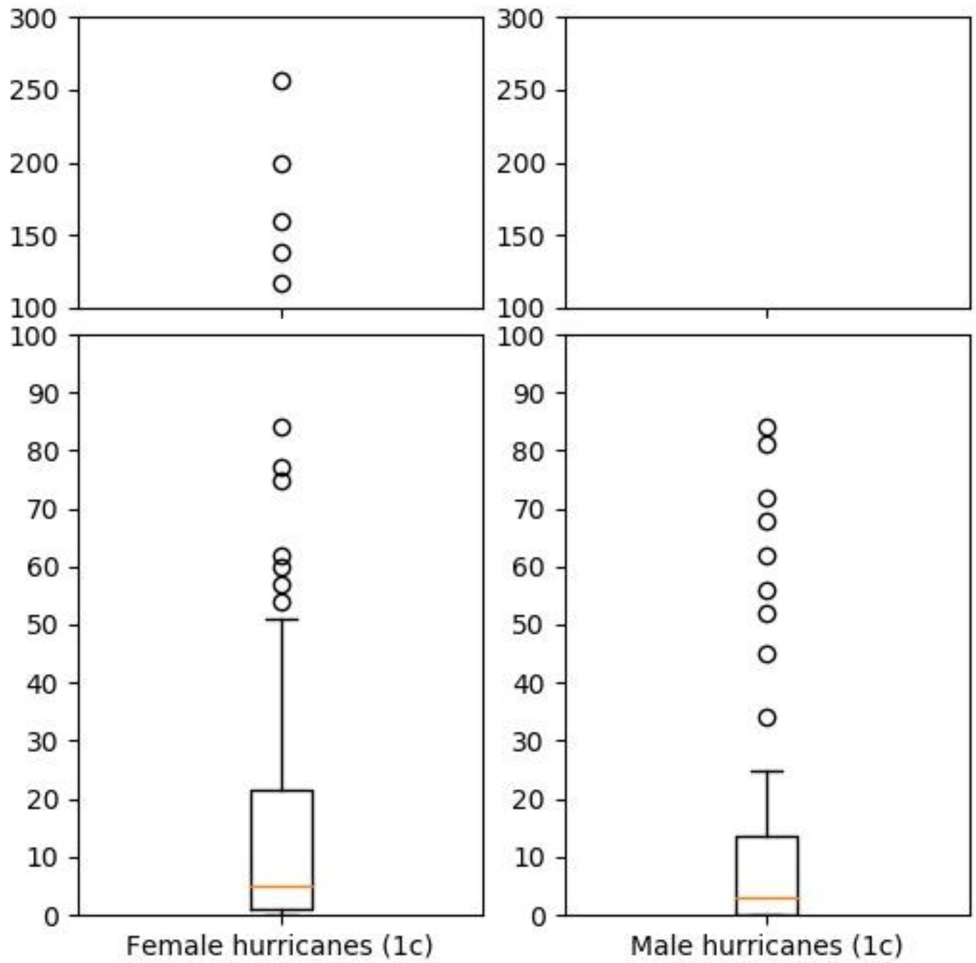
Min and Max whiskers

- Female: $f_s = 21.5 - 1 = 20.5$ and $1.5\times f_s = 1.5\times 20.5 = 30.75$
  Lower fence = Q1 - $1.5\times f_s = 1 – 30.75 = -29.75$ and min value in the data set within this range is 0.

Upper fence = Q3 + 1.5×f_s =21.5 + 30.75 = 52.25 and max value in the data set within this range is 51.

- Male: f_s = 13.5 - 0 = 13.5 and 1.5×f_s = 1.5×13.5 = 20.25
  Lower fence = Q1 - 1.5×f_s = 0 - 20.25 = -20.25 and min value within this range is 0.
  Upper fence = Q3 + 1.5×f_s = 13.5+20.25=33.75 and max value within this range is 25.
  Note that Min and Max whisker should be the values within Q1-1.5×f_s or Q3+1.5×f_s. Outliers are the values beyond these ranges.

c. Comparitive Boxplot



12 outliers of female hurricanes:

Ophelia	2017	Female	54
Florence	2018	Female	57
Carol	1954	Female	60
Rita	2005	Female	62
Betsy	1965	Female	75
Laura	2020	Female	77
Dorian	2019	Female	84
Agnes	1972	Female	117
Irma	2017	Female	138
Sandy	2012	Female	159
Diane	1955	Female	200
Camille	1969	Female	256

9 outliers of male hurricanes::

Joaquin	2015	Male	34
Nate	2017	Male	45
Gustav	2008	Male	52
Floyd	1999	Male	56
Andrew	1992	Male	62
Harvey	2017	Male	68
Michael	2018	Male	72
Earl	2016	Male	81
Ike	2008	Male	84

Considering the mean values, the Female named hurricanes seem to be more deadly. However, considering the median values, there is little difference. Since the median represent a typical number of deaths from a hurricane, we can conclude the female named hurricanes are NOT more deadly than male named hurricanes.

d.

Gender	Total number of data (n)	Mean	Standard Deviation	Minimum	Maximum
Female	86	84.59	383.10	0	3057
Male	61	22.33	76.48	0	585

Mean was derived by adding all the data values divided by n
S.D. is a sample standard deviation, divided by (n-1)
Min and Max are the smallest and the largest values in the data set, respectively.

e.

Gender	Q1	Q2	Q3	f_s(=IQR)	Min whisker	Max whisker
Female	1	5	26	25	0	62
Male	0	3	15	15	0	34

Q2
- Female: Average of the two middle ones (Jeanne 2004 and Wilma 2005)
- Male: the middle value, Jerry 1989

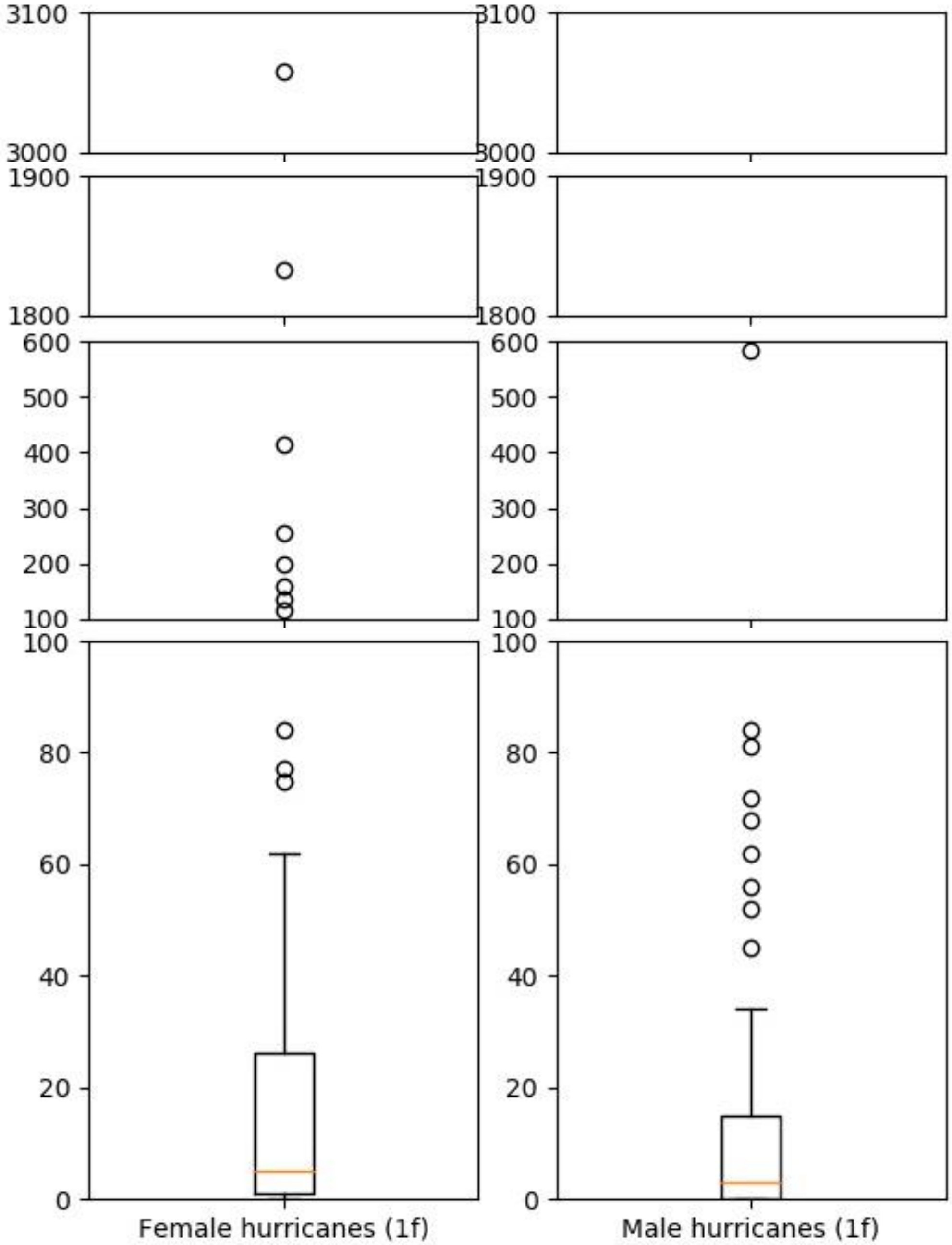
Q1
- Female: the middle value (Hermine 2016), including one Q2 in the first half
- Male: Average if two middle ones (Omar 2020), including Q2 in the first half

-

Q3
- Female: the middle value (Fran 1996), including Q2 in the last half
- Male: the middle value (David 1979), including Q2 in the last half

Min and Max whiskers
- Female: $f_s = 26 - 1 = 25$ and $1.5 \times f_s = 1.5 \times 25 = 37.5$
  Lower fence = Q1 - $1.5 \times f_s = 1 – 37.5 = -36.5$ and min value in the data set within this range is 0.
  Upper fence = Q3 + $1.5 \times f_s = 26 + 37.5 = 63.5$ and max value in the data set within this range is 62.
- Male: $f_s = 15 - 0 = 15$ and $1.5 \times f_s = 1.5 \times 15 = 22.5$
  Lower fence = Q1 - $1.5 \times f_s = 0 – 22.5 = -22.5$ and min value within this range is 0.
  Upper fence = Q3 + $1.5 \times f_s = 15 + 22.5 = 37.5$ and max value within this range is 34.

f. Comparative box plot



11 outliers of female hurricanes:

Betsy	1965	Female	75
Laura	2020	Female	77
Dorian	2019	Female	84
Agnes	1972	Female	117
Irma	2017	Female	138
Sandy	2012	Female	159
Diane	1955	Female	200
Camille	1969	Female	256
Audrey	1957	Female	416
Katrina	2005	Female	1833
Maria	2017	Female	3057

9 outliers of male hurricanes:

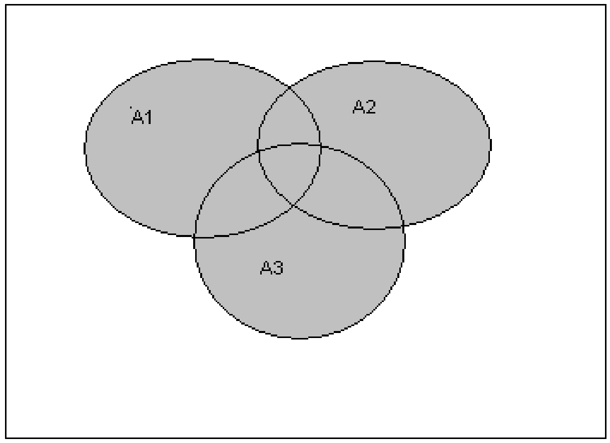
Nate	2017	Male	45
Gustav	2008	Male	52
Floyd	1999	Male	56
Andrew	1992	Male	62
Harvey	2017	Male	68
Michael	2018	Male	72
Earl	2016	Male	81
Ike	2008	Male	84
Matthew	2016	Male	585

g. When the death totals for Audrey, Katrina, and Maria are added to the data set, the mean increases from about 24 to about 85. Additionally, the standard deviation experiences a vast increase from about 45 to about 383. When extreme outliers are part of a data set, the mean and standard deviation values could be strongly affected. On the other hand the median value remains unchanged at 5.

h. The median better represent a typical number of deaths from a hurricane. While mean and standard deviation greatly affected by extreme values, the median number of deaths remains unchanged. Also note that the value of the Interquartile Range (IQR) changes by only 4.5 deaths. The quartiles are not affected by extreme data values and are therefore resistant (robust) measures of center and spread.

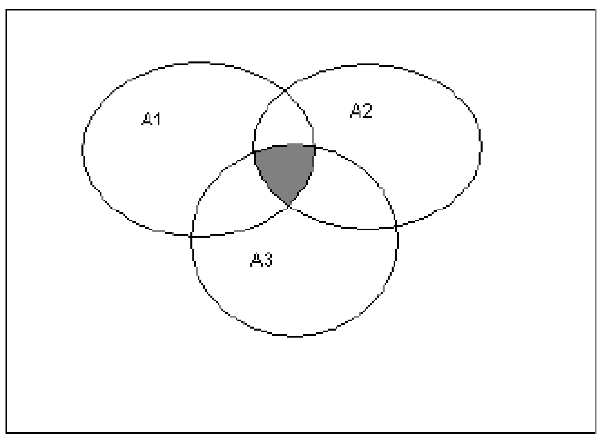
Problem 2.

a.



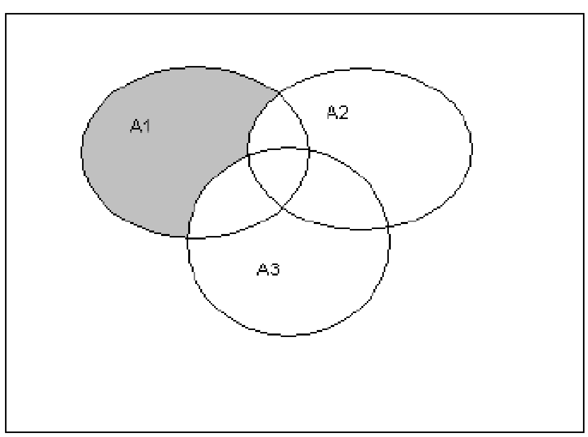
$A_1 \cup A_2 \cup A_3$

b.



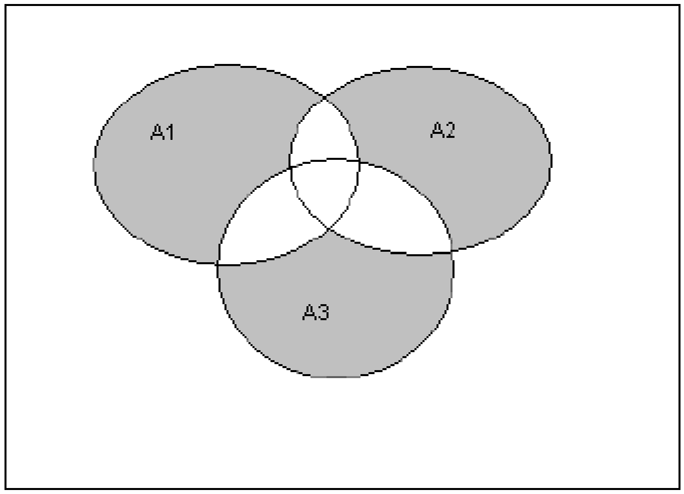
$A_1 \cap A_2 \cap A_3$

c.



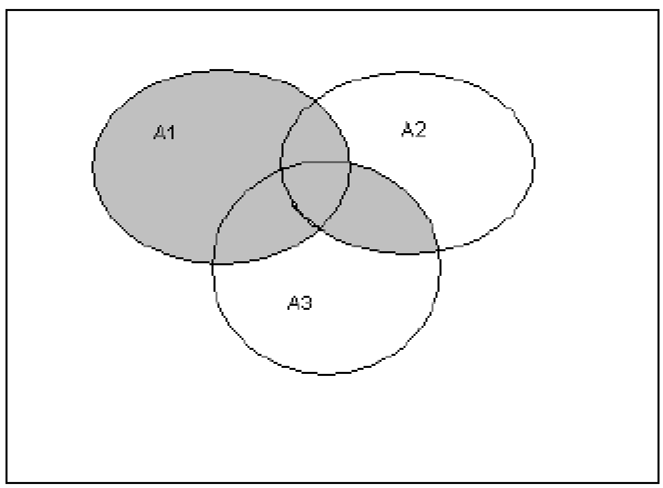
$A_1 \cap A_2' \cap A_3'$

d.



$$(A_1 \cap A_2' \cap A_3') \cup (A_1' \cap A_2 \cap A_3') \cup (A_1' \cap A_2' \cap A_3)$$

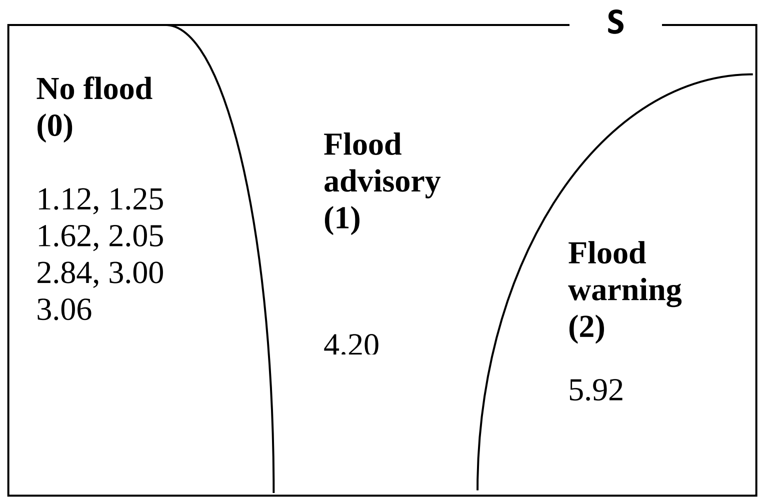
e.



$$A_1 \cup (A_2 \cap A_3)$$

Problem 3.

a.



b. L0, L1 and L2 are mutually exclusive because their intersection is ∅

c. L0, L1 and L2 are mutually exclusive and collectively exhaustive because $0 \cup 1 \cup 2 = S$.

Problem 4.

a.

Inflow	Outflow
6’	5’
6’	6’
6’	7’
7’	5’
7’	6’
7’	7’
8’	5’
8’	6’
8’	7’

Hence possible combinations of inflow and outflow are
(6’,5’), (6’,6’), (6’,7’), (7’,5’), (7’,6’), (7’,7’), (8’,5’), (8’,6’) and (8’,7’).

b. The possible water levels in the tank are 6’, 7’, 8’, 9’ and 10’.

c. Let E = at least 9 ft of water remains in the tank at the end of day.
Sample points (7’,5’), (8’,5’) and (8’,6’) are favorable to the event E.
So P(E) = 3/9 = 1/3.

Problem 5.

a. $P(A_1') = 1 - P(A_1) = 1 - .12 = .88$

b. $P(A_1 \cap A_2) = P(A_1) + P(A_2) - P(A_1 \cup A_2) = .12 + .07 - .13 = .06$

c. $P(A_1 \cap A_2 \cap A_3') = P(A_1 \cap A_2) - P(A_1 \cap A_2 \cap A_3) = .06 - .01 = .05$

d. P(at most two errors) = 1 – P(all three types)
$= 1 - P(A_1 \cap A_2 \cap A_3)$
$= 1 - .01 = .99$