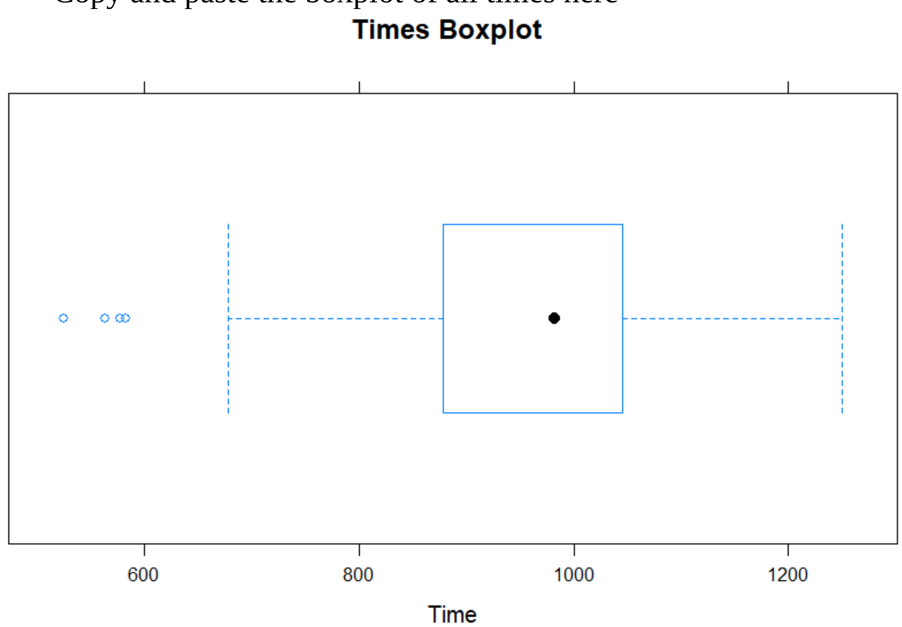
NAME: ________________________ Set: _________

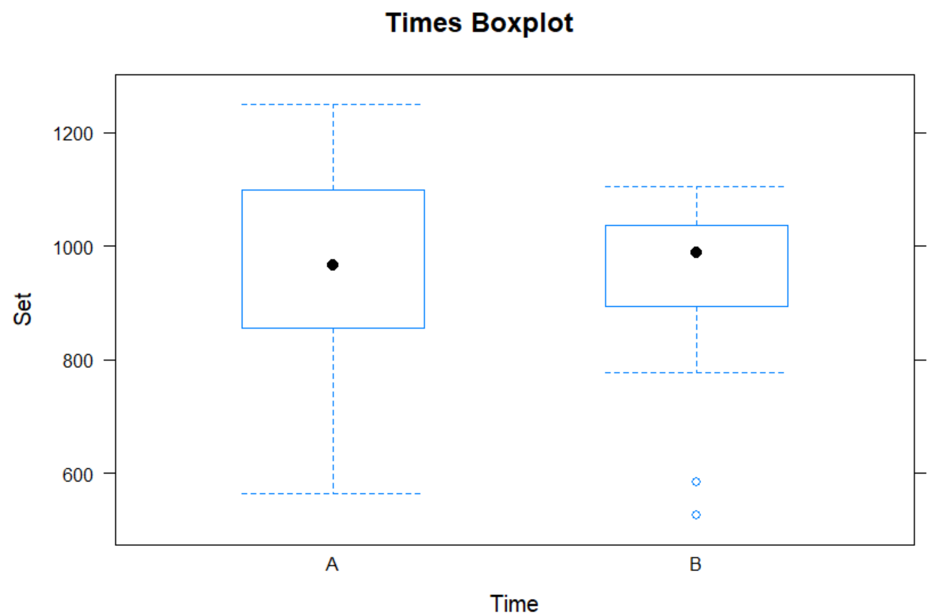
MATH 1350 **Statistics for Information Technology**

Lab # 3 – Boxplots and Probability Basics
Answer/Grading Sheet

Step:	**Answer (if requested)**	**Mark**	
5	• Load the mosaic library		/ 1
7	• Import commands for the data file		/ 1
10-11	• Modified bwplot command that includes graph title option • Copy and paste the boxplot of all times here  		/ 1 / 1
13-14	• Modified command for horizontally or vertically stacked boxplots, WITH axis labels and graph title • Copy and paste the boxplot of times by set here		/ 2 / 1

NAME & Set: Markus Afonso, Set C

Step:	Answer (if requested)	Mark	



Step:	Answer (if requested)	Mark	
15	• Which set do you feel was faster? EXPLAIN. By examination of the box plots, I would say B is only slightly faster. They both seems to be close, but B has more of their times at or above ~900 whereas Set A has a wider range. However, with the favstats cmd, it is shown that Set A has a higher average.		/ 2
17	• Paste the mean payout value for the cointoss game here: `mean(payout)` `[1] 0.504`		/ 1
18	(code below)		/ 4

```
n <- 1:20000
cards <- sample(c(1,2,3,4,5,6,7,8,9,10,11,12,13),
length(n),replace=TRUE)
dierolls <- sample(c(1,2,3,4,5,6), length(n),replace=TRUE)

payout <- 0
#let a 0 represent a heads up coin, a 1 represent a tails
up coin
for (i in n) {
  if (cards[i] < 5) {
```

NAME & Set: Markus Afonso, Set C

<table>
<tr><th>Step:</th><th>Answer (if requested)</th><th>Mark</th><th></th></tr>
<tr><td></td><td><pre>   payout[i]<- 5*dierolls[i]
  }
  else {
    payout [i] <- dierolls[i]
  }
}</pre></td><td></td><td></td></tr>
<tr><td>19</td><td>Paste the mean payout value for the card & dieroll game here:
<pre>> mean(payout)
[1] 7.807135</pre>
What is your estimate of the “break even” price to play my new game?
My estimate break even price would be around $7.8.</td><td></td><td>/ 1</td></tr>
</table>

R script

Paste your R script here. Make sure that it contains ALL of the elements worth points listed above.

```
# Lab 3
# Markus Afonso

library(mosaic)

gametimes <- read.delim("C:/Users/Markus/OneDrive -
BCIT/Desktop/Term2/MATH 1350 Statistics for IT/Week3/lab3.txt",
comment.char="#")

bwplot(~Time,data=gametimes, main='Times Boxplot')

bwplot(Time~Set,data=gametimes, main='Times Boxplot', ylab='Set',
xlab='Time')

favstats(Time~Set,data=gametimes)

#Stage 2 - Probability Experiments in R
n <- 1:20000
cointosses <- sample(c(0,1), length(n),replace=TRUE)
payout <- 0*n
#let a 0 represent a heads up coin, a 1 represent a tails up coin
for (i in n) {
  if (cointosses [i]==1) {
```

NAME & Set: Markus Afonso, Set C

Step:	Answer (if requested)	Mark	
	`` payout [i] <- 1 } else { payout [i] <- 0 } } mean(payout) ## Cards n <- 1:20000 cards <- sample(c(1,2,3,4,5,6,7,8,9,10,11,12,13), length(n),replace=TRUE) dierolls <- sample(c(1,2,3,4,5,6), length(n),replace=TRUE) payout <- 0 #let a 0 represent a heads up coin, a 1 represent a tails up coin for (i in n) { if (cards[i] < 5) { payout[i]<- 5*dierolls[i] } else { payout [i] <- dierolls[i] } } mean(payout) ```		
	Paper and Pencil problem #1 (this is just a space for your marks)		/ 3

NAME & Set: Markus Afonso, Set C

Step:	Answer (if requested)	Mark	
	mean $3.9+4.0+4.1+4.2+4.5+4.7+4.8+5.0+5.1+5.1=45.4$ $\frac{45.4}{10}=4.54$ $1.6+2.8+3.2+3.6+4.1+5.1+5.1+5.9+6.7+7.7=45.8$ $\frac{45.8}{10}=4.58$ Trump Trudeau $4.58-4.54=0.04$ median $\frac{4.5+4.7}{2}=4.6$ } same median $\frac{4.1+5.1}{2}=4.6$ Standard deviation Trudeau $s=\sqrt{\frac{(3.9-4.54)^2+(4.0-4.54)^2+(4.1-4.54)^2+(4.2-4.54)^2+(4.5-4.54)^2+(4.7-4.54)^2+(4.8-4.54)^2+(5.0-4.54)^2+(5.1-4.54)^2+(5.1-4.54)^2}{10-1}}$ $s=\sqrt{\frac{1.944}{9}}$ $s=0.46425$ Trump $s=\sqrt{\frac{(1.6-4.58)^2+(2.8-4.58)^2+(3.2-4.58)^2+(3.6-4.58)^2+(4.1-4.58)^2+(5.1-4.58)^2+(5.1-4.58)^2+(5.9-4.58)^2+(6.7-4.58)^2+(7.7-4.58)^2}{10-1}}$ $s=\sqrt{\frac{31.656}{9}}$ $s=1.8755$ Trump has a much larger sd.		

NAME & Set: Markus Afonso, Set C

Step:	Answer (if requested)	Mark	
	deviation coefficient $CV = \frac{s}{\bar{x}} \times 100$ trudeau $CV = \frac{0.46475}{4.54} \times 100$ $CV = 10.237\%$ Trump $CV = \frac{1.8755}{4.58} \times 100$ $CV = 40.95\%$ By calculating the difference between the average times we found that Trudeau has a 0.04s faster line, but Trump has a much larger variance. But, with a large enough sample, I believe that the times would be basically the same.		
	Paper and Pencil problem #2 2) By using the equation $CV = \frac{s}{\bar{x}} \times 100$ we can find the coefficiant of variance. Sandi $CV = \frac{1.91}{23} \times 100$ $CV = 8.3043\%$ > cameron $CV = \frac{2.12}{41} \times 100$ $CV = 5.1707\%$ ∴ cameron has a more consistent reading speed.		/ 2
	Paper and Pencil problem #3		/ 3

NAME & Set: Markus Afonso, Set C

Step:	Answer (if requested)	Mark	
	3) a) using combination $n = 52$ $r = 2$ $\quad \frac{n!}{(n-r)!\,r!}$ $\frac{52!}{(52-2)!\,2!}$ $\frac{52 \times 51 \times \cancel{50!}}{\cancel{50!}\,2}$ $\frac{52 \times 51}{2}$ $= 1326$ b) # of Ace pairs $= {}_4C_2$ $\frac{4!}{(4-2)!\,2!}$ $\frac{4 \times 3 \times \cancel{2!}}{\cancel{2!} \cdot 2!}$ $\frac{12}{2!}$ $= 6$ different pairs		

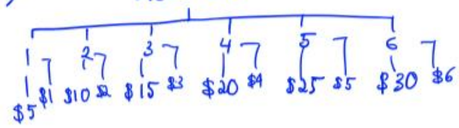
NAME & Set: Markus Afonso, Set C

Step:	Answer (if requested)	Mark	
	c) # of possible pairs $_{52}C_2$ $\frac{52!}{(52-2)!\,2!}$ $\frac{52 \times 51 \times \cancel{50!}}{\cancel{50!}\ 2!}$ $\frac{2652}{2!}$ $= 1326$ $P(\text{\# of ace pairs}) = \frac{\text{\# of ace pairs}}{\text{\# of possible pairs}}$ $= \frac{6}{1326}$ $= 0.00452$ or 0.452%		
	Paper and Pencil problem #4		/ 3

NAME & Set: Markus Afonso, Set C

Step:	Answer (if requested)	Mark	

4) Dice Rolls



b) possible times × amount

→ all cards under 5 → all cards above 5

$4(5 + 10 + 15 + 20 + 25 + 30) + 9(1 + 2 + 3 + 4 + 5 + 6)$ → the values of dice for each event

$420 + 189$

$= \$609$

c) $\frac{\text{total amount}}{\text{possible tries}}$

$= \frac{609}{78}$

$= 7.8077$

it is the same or very close to the number found in stage 2.

Total /26