LOMA and DAVA ASSIGNMENT

PART 1

AY2021/2022 Apr Semester

DECLARATION

I declare that I am the originator of this work and that all other original sources used in this work have been appropriately acknowledged.

I understand that plagiarism is the act of taking and using the whole or any part of another person's work and presenting it as my own without proper acknowledgement.

I also understand that plagiarism is an academic offence and that disciplinary action will be taken for plagiarism."

☑ I Agree (Please Tick ✔)

My Information

Name (as in matriculation card)	REUBEN ANG D'SOUZA
Admin Number	2100463E
Class (E.g. T01)	T12

LOMA – Total Hotel Charges

1. You are assigned to work out the total charges (inclusive of GST and service charge) for one booking in the dataset.

Answer:

My matriculation number is : _____2100463E_____

Therefore, the booking number I work on is : _____1463_____

adr = __102.5__, stay_in_weekend_nights = ____1_____, stay_in_week_nights : _____1_____

The Length of stay is : _____2 days_____

Total **before** GST and service charges is : _____$205_____

Total **after** GST and service charges is : _____$241_____

My working is shown below:

Booking Number:

Last three digits of matriculation number + 1000

 = 463 + 1000

 = 1463

Length of Stay:

Stay in weekend nights + stay in week nights

 = 1 + 1

 = 2 days

Total before GST and service charges:

Average daily rate x length of stay

= 102.5 x 2

= $205

Total charges inclusive of GST:

205 x 1.07 = $219.35

Total charges inclusive of GST and service charges:

219.35 x 1.1 =$241.29 (2d.p)

=$241 (truncated to an integer)

2. Using the last two digits of your matriculation number to represent x and y respectively, find the value of **k** using a calculator (E.g. if your matric number is 2112345A, then use x = 4 and y = 5). Be sure to show your **step-by-step** working.

Re-compute the **total charges inclusive of GST and service charges** based on **k** (new GST rate). Truncate your answer to an integer.

Answer:

My matriculation number is: ____2100463E_____

Therefore, X = __6_____ and Y = _____3_____

The value of A is   : _____1_____

The value of B is   : _____0.615_____

The value of k is   : _____8_____

The total **after** GST and service charges is : _____$243_____

My working is shown below:

$$A(x) = \lceil e^{(3x+10)/10} \rceil \bmod 8$$
$$= \lceil e^{(3(6)+10)/10} \rceil \bmod 8$$
$$= \lceil e^{\frac{28}{10}} \rceil \bmod 8$$
$$= \lceil 16.444647 \rceil \bmod 8$$
$$= 17 \bmod 8$$
$$= 1$$

$$B(y) = \log_{200}(2y^2 + 8)$$
$$= \log_{200}(2(3)^2 + 8)$$
$$= \log_{200}(26)$$
$$= \frac{\ln(26)}{\ln(200)}$$
$$= 0.61493 \ (5 \ d.p.)$$
$$= 0.615 \ (3 s.f.)$$

$$k = \lfloor A(x) + B(y) + 7 \rfloor$$
$$= \lfloor 1 + 0.61493 + 7 \rfloor$$
$$= \lfloor 8.61493 \rfloor$$
$$= 8$$

new GST rate = 8% = 0.08

total charges inclusive of GST
= 205 × 108%
= $221.40

total charges inclusive of GST and service charges
= 221.40 × 1.1
= $243.54
= $243 (truncated to an integer)

3a.   Is the function one-to-one?

Answer: The function is not one-to-one as elements -4 and 6 share the same image of 1 and elements 15 and 5 share the same image of 5. Therefore, the function is not one-to-one.

3 (a) $A(x) = \lceil e^{(3x+10)/10} \rceil \mod 8$

Let $x = 15$
$A(x) = \lceil e^{\frac{5.5}{}} \rceil$
$A(x) = \lceil e^{[(3)(15)+10]/10} \rceil \mod 8$
$A(x) = \lceil e^{5.5} \rceil \mod 8$
$A(x) = \lceil 244.692 \rceil \mod 8$
$A(x) = 245 \mod 8$
$A(x) = 5$

Let $x = 5$
$A(x) = \lceil e^{[(3)(5)+10]/10} \rceil \mod 8$
$A(x) = \lceil e^{2.5} \rceil \mod 8$
$A(x) = \lceil 12.18244 \rceil \mod 8$
$A(x) = 13 \mod 8$
$A(x) = 5$

Let $x = 6$
$A(x) = \lceil e^{[(3)(6)+10]/10} \rceil \mod 8$
$A(x) = \lceil e^{2.8} \rceil \mod 8$
$A(x) = \lceil 16.44464 \rceil \mod 8$
$A(x) = 17 \mod 8$
$A(x) = 1$

Let $x = -4$
$A(x) = \lceil e^{[(3)(-4)+10]/10} \rceil \mod 8$
$A(x) = \lceil e^{-0.2} \rceil \mod 8$
$A(x) = \lceil 0.818731 \rceil \mod 8$
$A(x) = 1 \mod 8$
$A(x) = 1$

3b.   Is the function onto?

Answer: Each element in R has a pre-image in Z. Therefore, the function is onto.

3 (b) $B(y) = \log_{200}(2y^2 + 8)$

When $y = 2$
$B(y) = \log_{200}[(2)(2)^2 + 8]$
$= \log_{200}(16)$
$= \frac{\ln(16)}{\ln(200)}$
$= 0.52330$ (5 s.f.)

When $y = -2$
$B(y) = \log_{200}[(2)(-2)^2 + 8]$
$= \log_{200}(16)$
$= \frac{\ln(16)}{\ln(200)}$
$= 0.52330$ (5 s.f.)

When $y = 1$
$B(y) = \log_{200}[(2)(1)^2 + 8]$
$= \log_{200}(10)$
$= \frac{\ln(10)}{\ln(200)}$
$= 0.43459$ (5 s.f.)

When $y = -1$
$B(y) = \log_{200}[(2)(-1)^2 + 8]$
$= \log_{200}(10)$
$= \frac{\ln(10)}{\ln(200)}$
$= 0.43459$ (5 s.f.)

When $y = 0$
$B(y) = \log_{200}[(2)(0)^2 + 8]$
$B(y) = \log_{200}(8)$
$= \frac{\ln(8)}{\ln(200)}$
$= 0.39247$ (5 s.f.)

4. Explain whether X or Y has a greater influence.

Answer: X has a greater influence as when X is higher than Y, K is bigger. Thus, X has a greater influence than Y.

4.

$$k = \lfloor (\lceil e^{(3x+10)/10} \rceil \mod 8) + \log_{200}(2y^2 + 8) + 7 \rfloor$$

let $x = 1$ and $y = 1$

$k = \lfloor (\lceil e^{(3(1)+10)/10} \rceil \mod 8) + \log_{200}[(2)(1)^2 + 8] + 7 \rfloor$

$k = \lfloor (\lceil 3.66930 \rceil \mod 8) + \log_{200}(10) + 7 \rfloor$

$k = \lfloor (4 \mod 8) + \frac{\ln(10)}{\ln(200)} + 7 \rfloor$

$k = \lfloor (4 + 0.4346 + 7) \rfloor$

$k = \lfloor 11.4346 \rfloor$

$k = 11$

now let $x = 5$ and $y = 1$

~~$k = \lfloor (\lceil e^{(3)(5)/10} \rceil$~~

$k = \lfloor (\lceil e^{(3)(5)+10/10} \rceil \mod 8) + \log_{200}[(2)(1)^2 + 8] + 7 \rfloor$

$k = \lfloor (\lceil 12.182494 \rceil \mod 8) + \log_{200}(10) + 7 \rfloor$

$k = \lfloor (13 \mod 8) + \frac{\ln(10)}{\ln(200)} + 7 \rfloor$

$k = \lfloor (5 + 0.4346 + 7) \rfloor$

$k = \lfloor 12.4346 \rfloor$

$k = 12$

now let $x = 1$ and $y = 5$

$k = \lfloor (\lceil e^{(3)(1)+10/10} \rceil \mod 8) + \log_{200}[(2)(5)^2 + 8] + 7 \rfloor$

$k = \lfloor (\lceil 3.66930 \rceil \mod 8) + \log_{200}(58) + 7 \rfloor$

$k = \lfloor (4 \mod 8) + \frac{\ln(58)}{\ln(200)} + 7 \rfloor$

$k = \lfloor (4 + 0.7664 + 7 \rfloor$

$k = \lfloor 11.7664 \rfloor$

$k = 11$

DAVA – Truth or Myth?

1. The **first** belief I am investigating is: Resort hotels are cheaper than city hotels.

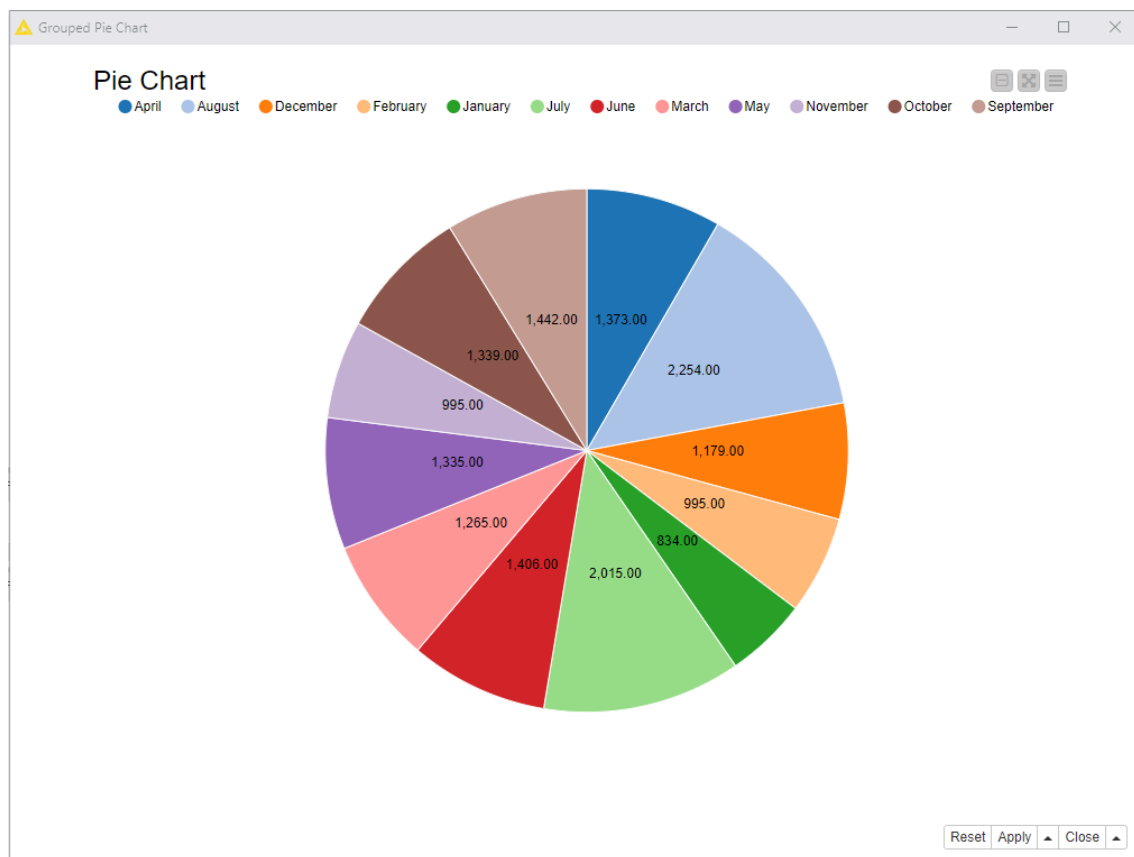   Conclusion: Based on data, this conventional belief is true

    Justification: The histogram below shows the binning column as hotel and aggregation column as adr or the average daily rate. The histogram shows that the adr for Resort hotels are lower than City hotels. City hotels have an average daily rate of 45050 while resort hotels have an average daily rate of 28194 as shown in the other picture of the histogram. This proves that Resort hotels are cheaper than City hotels.

2. The **second** belief I am investigating is: <u>The best time of the year to get the cheapest average daily rate (adr) is in January.</u>

Conclusion: Based on data, this conventional belief is <u>true</u>

Justification: As shown in the pie chart, it shows that in January (the dark green portion), it has the cheapest and lowest average daily rate (adr). Compared to the other months where they are all higher than in January. I used a GroupBy node and grouped arrival_date_month and adr. I then linked it to a Pie Chart node where I selected arrival_date_month as my category column. Finally, it produced the data I wanted, showing specifically what the average daily rate is every month. Allowing me to find that the best time of the year to get the cheapest average daily rate is in January.

3. The **third** belief I am investigating is: <u>The higher the number of children, the higher the average for total_of_special_requests. This applies to all customer types.</u>
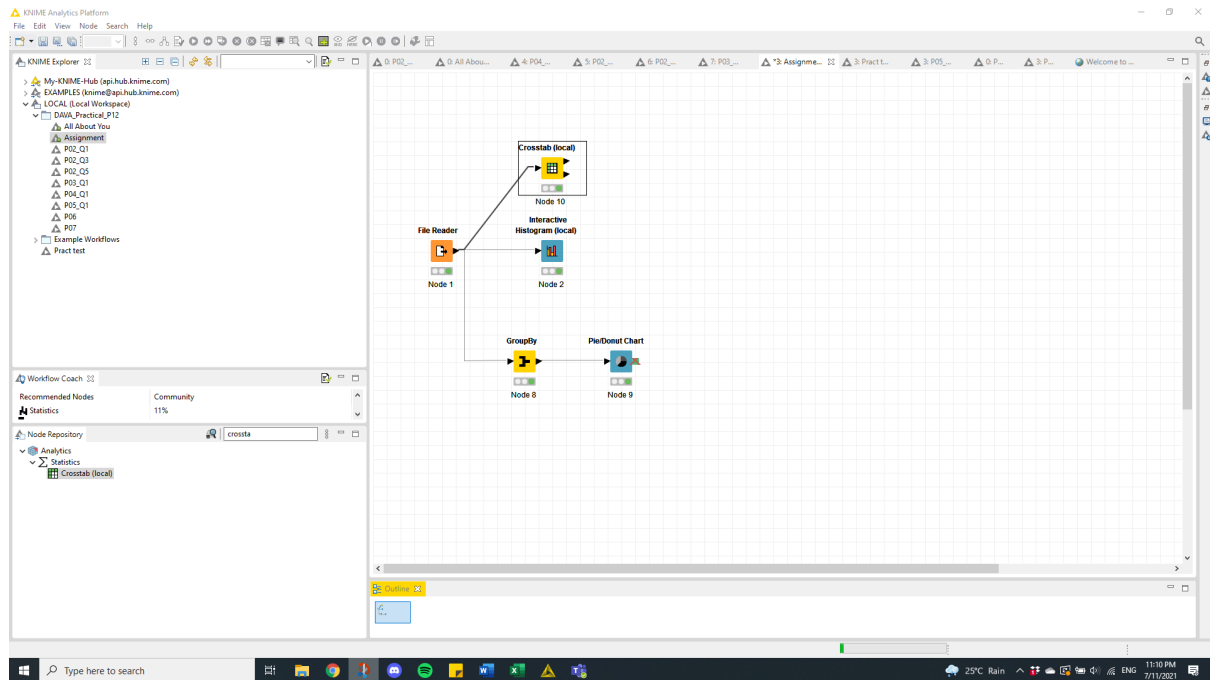
    Conclusion: Based on data, this conventional belief is <u>false</u>

    Justification: I have used the Crosstab node and set the row variable as children and column variable as the total_of_special_requests. This allowed me to make a table as shown below where on the left, it shows the number of children increasing, and on the right the total number of special requests. As we focus on the total, it shows how when there are 0 children, the total is 67,886, and when there is 1 child, the total is 3,244, when there are 2 children, the total is 2,060 and finally when there are 3 children, the total is 54. So, it shows that as the number of children increase, the total number of special requests keep decreasing. Therefore, the conventional belief that the higher the number of children, the higher the average for total of special requests is false.

Cross Tabulation of children by total_of_special_requests

Frequency	0	1	2	3	4	5	Total
0	33,858	23,483	8,651	1,622	240	32	67,886
1	911	1,164	857	273	35	4	3,244
2	844	734	374	92	14	2	2,060
3	22	16	11	5			54
Total	35,635	25,397	9,893	1,992	289	38	73,244

Screenshot of my overall KNIME workflow is shown below.

***** END OF ASSIGNMENT PART 1 *****