WQD7004 Programming for Data Science (AA2 Sem 2 2022/2023)

* Indicates required question

Question 1

Which of the following is valid function for output? *

- ⦿ cat()
- ◯ cout()
- ◯ println()
- ◯ Option 4

Which function return an error message if the package cannot be loaded? *

- ⦿ library
- ◯ help
- ◯ install.package
- ◯ demo

Which of the following can extract the first name from "John Lim"? *

☑ substring

☑ substr

☐ substi

☐ subset

Which of the following set the path of the current working directory in R? *

○ setd("dir")

○ setd(dir)

○ setdir(dir)

◉ none of the above

Which of the following is true if |x| >= 4 *

○ x >= 4 & x <= -4

◉ x >= 4 | x <= -4

○ x >= 4 & x < -4

○ x >= 4 | x < -4

What is the value of X if Z is 128 after execute the program below? *

```
X<-?
Y<-12
Z<-Y%%5*4^X
```

○ 2

○ 4

○ 6

◉ none of the above

Which of the following are true? *

```
count<-0
while(count<100) {
  // Point A
  print("Welcome to R!")
  count<-count+1
  // Point B
}
// Point C
```

☑ count < 100 is always true at Point A

☑ count < 100 is always true at Point B

☐ count < 100 is always false at Point B

☐ count < 100 is always true at Point C

☑ count < 100 is always false at Point C

Which of the following are TRUE? *

☑ repeat will execute the loop when the condition is true

☐ next will execute the loop when the condition is true

☑ break is used to break the execution of a loop

☐ break is used to skip the current iteration of a loop

A Leap year is divisible by 4 but not by 100. But a Leap year is divisible by 400. *
Which of the following is used to determine leap year?

○ (year%%4==0 & year%%100!=0) & (year%%400==0)

○ (year%%4==0 & year%%100==0) | (year%%400==0)

◉ (year%%400==0) | (year%%4==0 & year%%100!=0)

○ (year%%400==0) | (year%%4!=0 & year%%100!=0)

Which of the following statements are true? *

☑ (x > 0 & x < 10) is same as ((x > 0) & (x < 10))

☑ (x > 0 | x < 10) is same as ((x > 0) | (x < 10))
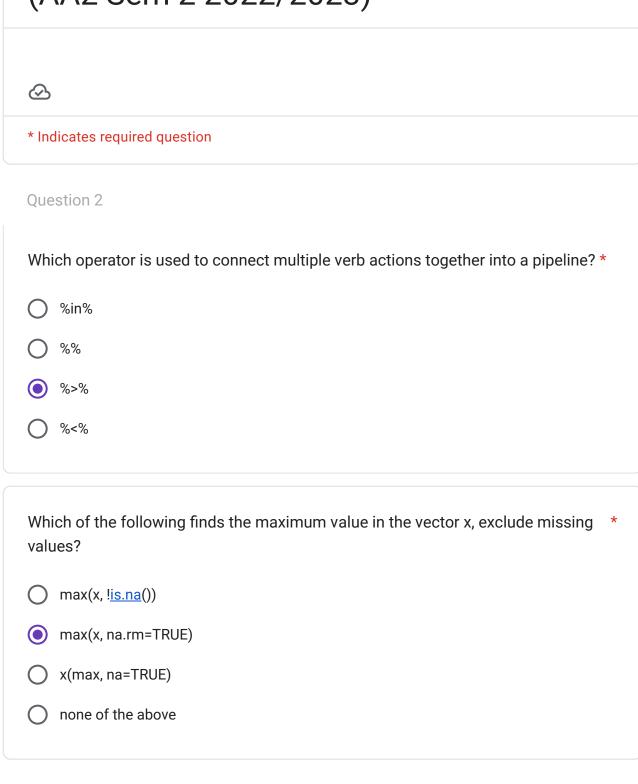
☑ (x > 0 | x < 10 & y < 0) is same as (x > 0 | (x < 10 & y < 0))

☐ (x > 0 | x < 10 & y < 0) is same as ((x > 0 | x < 10) & y < 0)

Back   Next   Clear form

Google Forms

WQD7004 Programming for Data Science (AA2 Sem 2 2022/2023)

* Indicates required question

Question 2

Which operator is used to connect multiple verb actions together into a pipeline? *

- ○ %in%
- ○ %%
- ● %>%
- ○ %<%

Which of the following finds the maximum value in the vector x, exclude missing *
values?

- ○ max(x, !is.na())
- ● max(x, na.rm=TRUE)
- ○ x(max, na=TRUE)
- ○ none of the above

Which of following used ; as the field separator? *

○ write.csv

○ writecsv

◉ write.csv2

○ writecsv2

Which of the following add new variable from the existing variable? *

○ filtering

○ selecting

◉ mutate

○ arrange

Which of the following return the number of missing values in each column? *

☑ colSums(is.na(df1))

☑ apply(is.na(df1), 2, sum)

☑ table(is.na(df1))

☐ sapply(df1, function(y) sum(is.na(x)))

Which of the following can be used in the summarise() function? *

☐ head()

☐ maximum()

☐ arrange()

☑ group_by()

Which of the following are TRUE for .rds file? *

☑ save.RDS() function is used to export a data object into the .rds file.

☑ the .rds file preserved data types and classes of variable.

☐ the .rds file required a lot of storage space.

☑ readRDS() function is used to read the contents of .rds file

Which of the following is used to display columns in descending order? *

○ descarrange()

○ descsort()

○ descending()

◉ none of the above

Which of the following represent the missing values, impossible values and undefined value in R? *

- ☑ NA
- ☑ NaN
- ☑ null
- ☐ " "

Which function finds the intersection between two different sets of data? *

- ○ Converge
- ○ Delegate
- ⦿ Merge
- ○ Swap

Back   Next   Clear form

Never submit passwords through Google Forms.

This form was created inside of University of Malaya. Report Abuse

Google Forms

WQD7004 Programming for Data Science (AA2 Sem 2 2022/2023)

* Indicates required question

Question 3

Which function calculates the count of each category of a categorical variable? *

○ category

○ group

◉ table

○ none of the above

Select one of the following statement that is not or least relevant to the rest of the *
statements.

○ feature scaling

○ resize to same range

○ normalize

◉ over sampling

Given the following output of confusion matrix, what is the value of accuracy in 3 decimal place? *

n=200	Predicted No	Predicted Yes
Actual No	20	8
Actual Yes	32	140

0.800

Given the following output of confusion matrix, what is the value of sensitivity in 3 decimal place? *

n=200	Predicted No	Predicted Yes
Actual No	20	8
Actual Yes	32	140

0.814

Which of the following is very sensitive to outliers? *

○ mode

○ median

● mean

○ trimmed mean

A scatterplot of "news" and "pricing" based on PCReview dataset is created. *
Which of the following add a smoother to the scatterplot?

○ ggplot(news, pricing, data = PCReview) + stats_smooth()

◉ ggplot(news, pricing, data = PCReview) + geom_smooth()

○ ggplot(news, pricing, data = PCReview, smoother = TRUE)

○ ggplot(news, pricing, data = PCReview) +geom_line()

Given the following output of confusion matrix, what is the value of specificity in 3 *
decimal place?

n=200	Predicted No	Predicted Yes
Actual No	20	8
Actual Yes	32	140

0.714

Which of the following are TRUE about ggplot2? *

☐ ggplot2 is a data visualization package used for data cleaning.

☑ ggplot2 supports data in a data frame.

☑ the geom is the plotting object in ggplot2.

☑ the geom is used to make conditioning plots.

To use a t-test, the dependent variable must have *

○ Nominal or interval data

○ Ordinal or ratio data

● Interval or ratio data

○ Ordinal or interval data

Which of the following statements are TRUE about overfitting? *

☑ unable to generalize

☑ unable to capture full complexity of data

☑ holds no predictive power because model is perfectly suited to the sample data

☐ high error rate is good indicators of overfitting

Back   Next   Clear form

Never submit passwords through Google Forms.

Google Forms