

## Homework 1

**Due by: February 19, 2023, 11:59 PM**

---

*All assignments must be submitted through **eLearning**. Alternative submission methods are not acceptable. **Submissions after the deadline will not be accepted**, and accordingly, a grade of zero will be automatically applied for a missing submission after the deadline.*

*When you submit, please make your file a **docx version** and specify the **name of the file** as follows. "[Homework 1] Your first name" (e.g., [Homework 1] Andrew)*

---

**Q1. (1 point) [short answer]** Fill in the blanks.

“There are three types of business analytics to drive decision making. ( a ) analytics tell us what has already happened. ( b ) analytics show us what could happen in the future based on previous trends and patterns. Lastly, prescriptive analytics inform us what should happen in the future.”

**Answer)** (a)

(b)

**Q2. (1 point) [multiple choice]** Which kind of machine learning algorithm is most required for the following task?

“To write a program to filter out spam emails, a computer programmer can train a machine learning algorithm with a set of spam-like emails labeled as spam and regular emails labeled as not-spam. The idea is to make an algorithm that can learn the characteristics of spam emails from this training set to filter out spam emails when it encounters new emails.”

- (1) Convolutional Neural Networks
- (2) Collaborative Filtering
- (3) Classification
- (4) Clustering
- (5) Unsupervised Learning

**Answer)**

**Q3. (1 point) [short answer]** Write the code.

“R packages are collections of functions and data sets developed by the community. If you want to install the package **"ggplot2"**, what should you type the code on the console window?”



**Answer)**

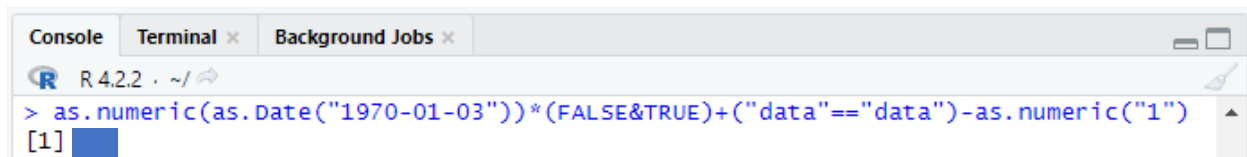
**Q4. (1 point) [multiple choice]** Which variable name is invalid in R?

- (1) Variable\_2023.01.01
- (2) 2023\_Variable
- (3) VARIABLE.2023.
- (4) v\_a\_r\_i\_a\_b\_l\_e\_2023
- (5) VARIABLE.\_.

**Answer)**

**Q5. (1 point) [short answer]** What is the answer to the following command?

```
> as.numeric(as.Date("1970-01-03")) * (FALSE&TRUE) + ("data"=="data") - as.numeric("1")
```

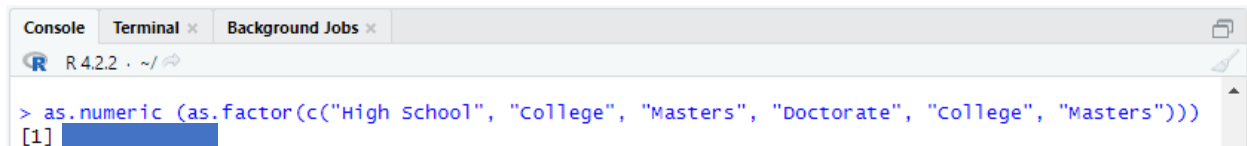


```
R 4.2.2 . ~/
> as.numeric(as.Date("1970-01-03"))*(FALSE&TRUE)+("data"=="data")-as.numeric("1")
[1] 1
```

**Answer)**

**Q6. (1 point) [short answer]** What is the answer to the following command?

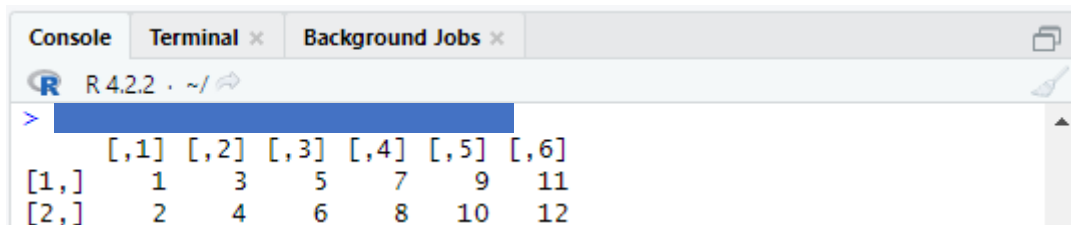
```
> as.numeric(as.factor(c("High School", "College", "Masters", "Doctorate", "College", "Masters")))
```



```
R 4.2.2 . ~/
> as.numeric(as.factor(c("High School", "College", "Masters", "Doctorate", "College", "Masters")))
[1] 1 2 3 4 2 3
```

**Answer)**

**Q7. (1 point) [short answer]** What should you type if you want to make the following matrix?

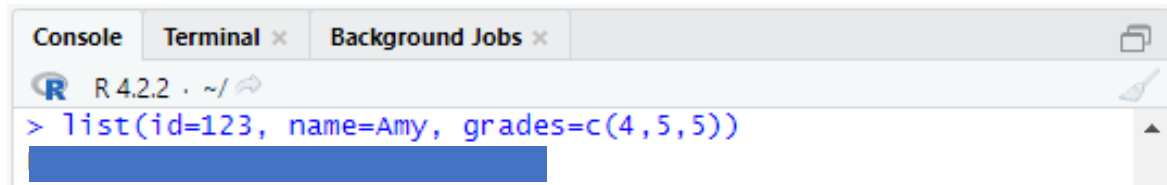


```
R 4.2.2 . ~/
> matrix(1:12, nrow=2, byrow=TRUE)
      [,1] [,2] [,3] [,4] [,5] [,6]
[1,]    1    3    5    7    9   11
[2,]    2    4    6    8   10   12
```

**Answer)**

**Q8. (1 point) [short answer]** What is the answer to the following command? Please briefly describe the reason for the answer.

```
> list(id=123, name=Amy, grades=c(4,5,5))
```

A screenshot of an R console window. The window has three tabs: 'Console', 'Terminal', and 'Background Jobs'. The 'Console' tab is active. The console shows the R prompt '>' followed by the command 'list(id=123, name=Amy, grades=c(4,5,5))'. The output of the command is a list object, displayed as 'list(id=123, name=Amy, grades=c(4,5,5))'. The console window also shows the R logo and version 'R 4.2.2' in the top left corner.

**Answer)**