

# The University of British Columbia

## Data Science 570 Predictive Modelling

### Lab Assignment 4

You are expected to submit answers to exercise 3. Instructions: Please use a png, pdf or html file for submission.

1. We'll analyze the marketing data set to predict sales, based on advertising budgets spent on youtube and facebook. First, let's take a look at the data.

```
install.packages('datarium', repos='http://cran.us.r-project.org')

## Installing package into 'C:/Users/xshi/Documents/R/win-library/4.0'
## (as 'lib' is unspecified)

## package 'datarium' successfully unpacked and MD5 sums checked
##
## The downloaded binary packages are in
## C:\Users\xshi\AppData\Local\Temp\RtmpusACop\downloaded_packages

library(datarium)

## Warning: package 'datarium' was built under R version 4.0.5

data("marketing", package = "datarium")
head(marketing)

##   youtube facebook newspaper sales
## 1  276.12    45.36    83.04  26.52
## 2   53.40    47.16    54.12  12.48
## 3   20.64    55.08    83.16  11.16
## 4  181.80    49.56    70.20  22.20
## 5  216.96    12.96    70.08  15.48
## 6   10.44    58.68    90.00   8.64
```

- 1.1 Build the model:  $\text{sales} = b_0 + b_1 \cdot \text{youtube} + b_2 \cdot \text{facebook}$  and call summary of `lm()` (2 marks).
  - 1.2 a) Build the model with interaction:  $\text{sales} = b_0 + b_1 \cdot \text{youtube} + b_2 \cdot \text{facebook} + b_3 \cdot \text{youtube} \cdot \text{facebook}$  and call summary of `lm()` (2 marks). b) When advertising on facebook costs \$1,000, how much would a \$1,000 increase in advertising on youtube increase sales? (1 mark) c) When advertising on youtube costs \$1,000, how much would a \$1,000 increase in advertising on facebook increase sales? (1 mark)
  - 1.3 Based on the adjusted  $R^2$  values, could we remove the interaction term? (2 marks)
2. (3 marks) Simulate 100 Poisson random numbers using `rpois(100,  $\lambda$ )`, where  $\lambda = e^{\beta_0 + \beta_1 x + \beta_2 x^2}$ ,  $\beta_0 = 1$ ,  $\beta_1 = -1$ ,  $\beta_2 = -2$ , and  $x$  is from a standard normal distribution. Use `glm()` function to fit Poisson regression and print the coefficients.
3. (3 marks) Simulate 100 Bernoulli random numbers using `rbinom(100, 1, prob = p)`, where  $p = \frac{e^{\beta_0 + \beta_1 x + \beta_2 x^2}}{1 + e^{\beta_0 + \beta_1 x + \beta_2 x^2}}$ ,  $\beta_0 = 1$ ,  $\beta_1 = -1$ ,  $\beta_2 = -2$ , and  $x$  is from a standard normal distribution. Use `glm()` function to fit logistic regression and print the coefficients.