

**R**

# Exercise Book

**TRAINING MATERIALS - EXERCISE BOOK**

---

Contacts

**Thomas.Knowles@qa.com**

*team.qac.all.trainers@qa.com*

[www.consulting.qa.com](http://www.consulting.qa.com)

# Contents

Introduction	– 3
Basic	– 3
Hello World!	– 3
Assignment	– 3
Parameters	– 3
Parameters/Operators	– 3
Conditionals	– 3
Conditionals 2	– 4
Recursion	– 4
Anonymous Functions	– 4
Intermediate	– 5
Blackjack	– 5
Unique Sum	– 5
Too hot?	– 5
Leap Year	– 5
Working with Files	– 6
Plotting	– 6
CO2 Data	– 7
Orchard Sprays Data	– 7
Chick Data	– 7
Advanced	– 8
Prime Numbers 1	– 8
Prime Numbers 2	– 8
Salary Predictions	– 8

# Introduction

This document contains a range of exercises in order to help you get to grips with R. You are not expected to complete all tasks, but you should aim to complete as many as you can to maximise your R ability.

## Basic

### HELLO WORLD!

Output `"Hello World!"` to the console.

### ASSIGNMENT

Store `"Hello World!"` in a variable, then output it to the console.

### PARAMETERS

Create a function that accepts a string as a parameter, and then outputs that string to the console.

### PARAMETERS/OPERATORS

Create a function that accepts two integers as an input, then returns an integer that is a sum of the two integers given.

### CONDITIONALS

Modify your method from the Parameters/Operators task to accept another parameter, a Boolean, which if it is true, the method will return a sum of the two numbers, and if it is false it will return the multiplication of the two numbers.

## R

### CONDITIONALS 2

Modify your function from the Conditionals task to have another if statement that checks if one of the numbers is 0, if this is true then return the other non-0 number.

```
Input -> 1, 0 Return 1  
Input -> 1, 2 Return 3
```

### RECURSION

Create a for loop that will call and output the result of your method from the Conditionals 2 task 10 times, using the current iteration as one of the parameters you pass to it.

### ANONYMOUS FUNCTIONS

Create an anonymous function to return the first value in a vector.

# Intermediate

## BLACKJACK

Given 2 numeric values greater than 0, return whichever value is closest to 21 without going over 21. If they both go over 21 then return 0

```
Input (18, 21) -> 21
Input (20, 18) -> 20
Input (22, 22) -> 0
```

## UNIQUE SUM

Given 3 numeric values, return their sum. If one value is the same as another value, they do not count towards the sum. Aka only return the sum of unique numbers given.

```
Input (1, 2, 3) -> 6
Input (3, 3, 3) -> 0
Input (1, 1, 2) -> 2
```

## TOO HOT?

Given a numeric value and a Boolean value, **temperature** and **isSummer**, if temperature is between 60 and 90 (inclusive), unless it's summer where the upper limit is 100 instead of 90. Return true if the temperature falls within the range, false otherwise.

## LEAP YEAR

Given a year work out if it is a leap year or not.

**Rule:** A year is a leap year if it is divisible by 4, and either divisible by 400 or not divisible by 100.

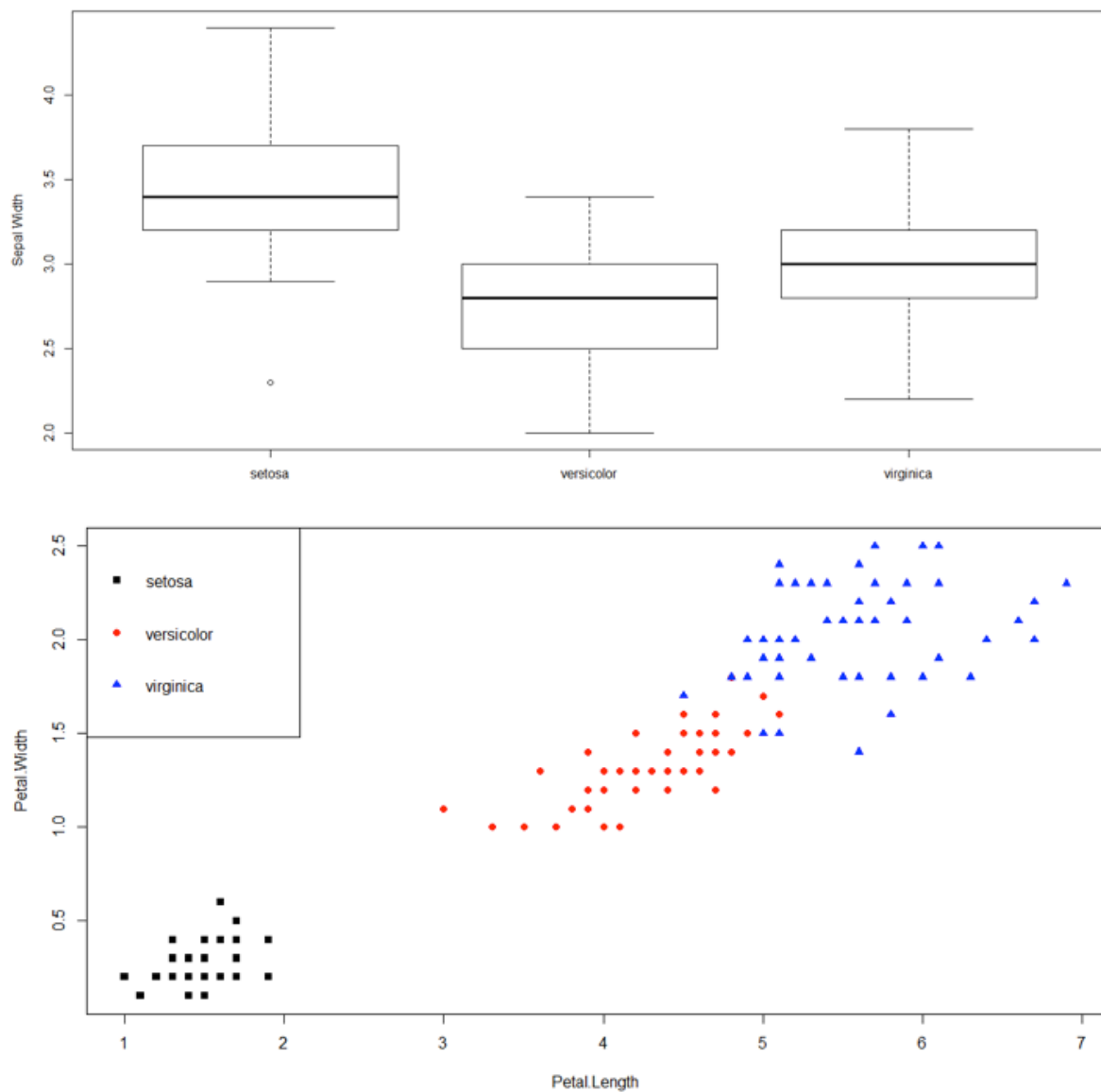
## R

### WORKING WITH FILES

Create a vector of even numbers. Write this out to a csv file called `evens.csv`. Read this file back in, add one to each element, then write this new vector out to a file called `odds.csv`.

### PLOTTING

Using the inbuilt Iris data (type the command `data(iris)` to access), recreate the two plots below using `boxplot()` and `plot()` (these should be two separate commands)



## R

### CO2 DATA

Using the inbuilt CO2 dataset, perform the following tasks

Find out the class type of the Plant column

Store the mean of the uptake column in a variable called mean\_uptake

Create a boxplot showing the uptake for Mississippi and Quebec, on the same graph but with a plot of the data for each subset

Create a new data frame called quebec\_CO2 that only contains rows from CO2 where the Type is Quebec. Create another one for Mississippi.

Create a function called mean\_checker that will take in 2 vectors, give you the mean of both, then which has the higher mean. Run it to find out whether Mississippi or Quebec has the higher uptake.

### ORCHARD SPRAYS DATA

Using the inbuilt OrchardSprays data, complete the following tasks.

Find the treatment that is responsible for the maximum individual decrease. Store this in a variable called max\_decrease.

Create a boxplot to show the decrease for each treatment.

### CHICK DATA

Using the inbuilt ChickWeight data, explore the data and create at least 3 clearly different graphs. Use this to outline an argument for which diet is 'best'.

# Advanced

## PRIME NUMBERS 1

Create an algorithm that determines how many prime numbers are between 1 and 3 million.

## PRIME NUMBERS 2

Create an algorithm that determines how many prime numbers are between 1 and 2 billion.

## SALARY PREDICTIONS

Using the `censusdata_train.csv` file, predict the salaries (either above or below 50K, binary answer) of the people in the `censusdata_test.csv` file. Submit your prediction as often as you like to your trainer for a score. Don't forget to clean your data before making predictions.

You must create at least one graph (type of your choosing).

Otherwise, explore the data however you see fit. Create a document of any visuals you create. The data is explained below.

No. of rows in train data: **32561**

No. of rows in test data: **16281**

### Rows:

1. Age – continuous
2. Workclass – type of work – Private, Self-emp-not-inc, Self-emp-inc, Federal-gov, Local-gov, State-gov, Without-pay, Never-worked
3. Final weight determined by Census – continuous
4. Education – level of education of individual – Bachelors, Some-college, 11th, HS-grad, Prof-school, Assoc-acdm, Assoc-voc, 9th, 7th-8th, 12th, Masters, 1st-4th, 10th, Doctorate, 5th-6th, Preschool



## R

5. Education-num – number of years of education – continuous
6. Marital-status – categorical – Married-civ-spouse, Divorced, Never-married, Separated, Widowed, Married-spouse-absent, Married-AF-spouse
7. Occupation – categorical – Tech-support, Craft-repair, Other-service, Sales, Exec-managerial, Prof-specialty, Handlers-cleaners, Machine-op-inspct, Adm-clerical, Farming-fishing, Transport-moving, Priv-house-serv, Protective-serv, Armed-Forces
8. Relationship – categorical – Wife, Own-child, Husband, Not-in-family, Other-relative, Unmarried
9. Race – categorical – White, Asian-Pac-Islander, Amer-Indian-Eskimo, Other, Black
10. Sex – categorical – Female, Male
11. Capital-gain – continuous, gain made by individual
12. Capital-loss – continuous, loss made by individual
13. Hours-per-week – continuous, no. of hours spent on work by the individual
14. Native-country – categorical - United-States, Cambodia, England, Puerto-Rico, Canada, Germany, Outlying-US(Guam-USVI-etc), India, Japan, Greece, South, China, Cuba, Iran, Honduras, Philippines, Italy, Poland, Jamaica, Vietnam, Mexico, Portugal, Ireland, France, Dominican-Republic, Laos, Ecuador, Taiwan, Haiti, Columbia, Hungary, Guatemala, Nicaragua, Scotland, Thailand, Yugoslavia, El-Salvador, Trinidad&Tobago, Peru, Hong, Holand-Netherlands