Matlab, Python & R

Assessment Objectives

|  |  |  |
| --- | --- | --- |
| **Author** | **Date** | **Note** |
| Patrick Upson | June 4, 2014 | Document Creation |
|  |  |  |

Contents

[Purpose 1](#_Toc389638116)

[Project 1](#_Toc389638117)

[Objectives 1](#_Toc389638118)

[Programming Languages 1](#_Toc389638119)

[Python 1](#_Toc389638120)

[Tools 1](#_Toc389638121)

[Learning Resources 2](#_Toc389638122)

[R 2](#_Toc389638123)

[Tools 2](#_Toc389638124)

[Learning Resources 2](#_Toc389638125)

[Appendix A. Terms 3](#_Toc389638126)

# Purpose

The purpose of this document is to outline the requirements and objectives for the Matlab, Python, R Assessment project. This document should contain defined objectives and intended actions to reach the stated objectives.

# Project

The purpose of this project is to assess the process of migrating away from Matlab as the primary scientific programming language. It is very heavily used by scientist in Fisheries and Oceans to:

1. Automate Data Processing
2. Analyze and Publish Data
3. Generate Graphs and Charts

There may be cases where Matlab is a required tool, but the overall goal of the project is to determine if and when Matlab is required, whether Python or R could be suitable replacements and how much effort may be required to update Matlab code to one of the possible replacement languages.

It should be noted that one of the factors to take into consideration when determining whether Python or R are suitable replacements will be the owner of the projects preference.

# Objectives

1. Rewrite the Matlab ODF reader tool in both Python and R
2. Determine strengths and weaknesses of Python and R against Matlab
3. Make note of when it might be appropriate to use Matlab or when it would be beneficial to migrate code to Python or R
4. Document tools, Learning resources and APIs required to work with the code
5. Unit Testing APIs and techniques should be used when developing code if possible

# Programming Languages

The primary difference between Python and R appears to be that Python is a heavily used programing language with well supported APIs and a strong user community that also has the ability to do heavy data processing; it is a programing language first. R is a statistical language with heavy programing elements and also a strong, although more scientific, user community. It is the primary choice for researchers because of its powerful math and data crunching abilities; it is a statistical language first.

## Python

Tools:

1. IDE – Eclipse
2. Unit Testing – PyUnit

Learning Resources:

## R

Tools:

1. IDE – Eclipse
2. Unit Testing – testthat

Learning Resources:

# Appendix A. Terms

**API – Application Program Interface**

A collection of libraries serving a specific function used to speed up development of applications