**New Course – Biology 2XXX**

**Data skills for biologists**

**RESOURCE IMPLICATIONS:**

This course will use teaching resources currently available in the Department of Biology.

**Instructional Costs**

No additional instructional costs will be required: the course will be taught by faculty members appointed in Biology (Amy Hurford, ??).

**RESOURCE IMPLICATIONS: Library Holdings and/or Other Resources Required**

The costs associated with new course can be met from within the existing budget allocation of Biology.

Signature of Unit Head (if appropriate): \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Date: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Signature of Dean/Associate Vice-President (Academic)/Vice-President:

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Date: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Course Number and Title**

BIOL 2XXX – Data skills for biologists

**Abbreviated Course Title**

Data in Biology

**Calendar Description**

To be drafted when the content of the course is clearer.

LH: 3 hours (lab only course)

PR: None

RECOMMENDED:

**Secondary Changes (if applicable)**

This course should be added to the list of recommended courses for the following majors and honours in Biology. The course should be a PR for BIOL 3295, 4360, 4605, and 4607, ???

**Rationale**

The ability to work with data using computers is valued by employers in the biology sciences and applies widely across all sub-disciplines of biology. Undergraduate students in biology are aware of a missing curriculum component in this area, and we have proposed this new course in part response to the Academic Unit Program report recommending a review of the biology curriculum.

A main suggestion to improve quantitative training in ecology-related disciplines is ‘to relate theoretical and statistical modeling to applied ecological problems’ (Barraquand et al. 2014). In this lab-only course, students will practice data archiving, loading, processing, analysis, and visualization using data from a variety of areas within the biological sciences. The proposed course will develop student’s data skills, while communicating the relevance and need for these skills in biology.

This course will provide fundamental training across all areas in biology, and the skills learned during this course will be built upon during upper-level courses in biology.

**Consultations**

Grenfell Campus:

St. John’s Campus: All Faculty of Science Departments

Marine Institute:

MUN Library

**Sample Course Outline and Method of Evaluation**

**Data skills for Biologists – Course Outline**

**Instructors:** ??

**Textbook:** None.

**Description:**

**Evaluation:** Labs (12) – 100%

**Labs:**

1. Data entry and storage
   1. Improving messy data
   2. Data entry and validation in Excel
   3. More from BES ref
2. Introduction to R and RStudio
   1. R & RStudio
   2. Data Structures (factors, strings, data frames, vectors, variables)
   3. Paths, loading data
   4. Basic Reproducibility
3. Introduction to R and RStudio
   1. Loading Packages
   2. Built-in functions
   3. Errors, Debugging, Getting Help
4. Working with Data
   1. Data cleaning with the R package: dplyr
5. Data Visualization
   1. ggplot2
6. Data Visualization 2
   1. Survey of data visualizations commonly used in biology
   2. Logarithmic (and other) transformations
7. Writing scripts
   1. Good coding style
   2. Debugging
8. More advanced R basics
   1. Writing functions
   2. Conditionals
   3. Loops
9. Writing hypotheses as equations
   1. Choosing notation
   2. Units
   3. General linear models in R
10. Uncertainty and hypothesis testing
    1. Type I and Type II error
    2. ASA’s statement on p-values
    3. Residuals and error distributions
    4. Nonlinear models

**Recommended Resources:**

Carey, M. A. and J. A. Papin. 2018. Ten simple rules for biologists learning to program. PLoS Comput Biol. 14(1): e1005871.

Harrison, K. Ed. 2018. Data management. British Ecological Society Guide to Better Science Series. <https://www.britishecologicalsociety.org/wp-content/uploads/2019/06/BES-Guide-Data-Management-2019.pdf>

Cumming, G., Fidler, R. and D. L. Vaux. 2007. Error bars in experimental biology. J Cell Biol 177(1):7-11. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2064100/>

Weissgerber, T. L., Milic, N. M., Winham, S. J., and V. D. Garovic. 2015. Beyond bar and line graphs: time for a new data presentation paradigm. PLOS Biology 13(4): e1002128 <https://doi.org/10.1371/journal.pbio.1002128>

Wickham, H. Style guide. http://r-pkgs.had.co.nz/style.html

Yaddanapudi, L. N. 2016. The American Statistical Association statement on P-values explained. J Anaesthesiol Clin Parmacol. 32(4): 421-423.

Zuur, A. F., E. N. leno, and C. S. Elphick. 2010. A protocol for data exploration to avoid statistical problems. Methods in Ecology and Evolution 1: 3-14.

**Instructor(s)**

Dr. Amy Hurford, Math & Biology, MUN

Dr. Shawn Leroux, Biology, MUN

**References**

Barraquand, F.

Lai, J., C. J. Lortie, R. A. Muenchen, J. Yang, K. Ma. 2019. Evaluating the popularity of R in ecology. Ecosphere e02567 h[ttps://doi.org/10.1002/ecs2.2567](https://doi.org/10.1002/ecs2.2567)

**SUMMARY PAGE FOR SENATE**

**Approval Form**

**Course Title and Number:**

**Abbreviated Course Title:**

**Calendar Description:**

LH:

PR:

RECOMMENDED:

**Rationale** :

**Consultations Sought From Comments Received**

***Provide a consultation list and indicate whether or not comments were received***

**Library Report Received** Yes/No

**Approved by Dean, Associate Vice-President (Academic) or Vice-President**

Yes/No

Name

**-------------------------------------------------------------------------------------------------------------------------------**

**FOR OFFICE USE ONLY**

**APPROVAL GRANTED BY SENATE COMMITTEE ON UNDERGRADUATE STUDIES**

Chair:

Secretary:

Date: