Assignment #1: Introduction to Data Reading, Fitting, and Plotting (total 10 points), due by 11:59 pm Friday, 29 September 2023

Question 1: Please, use this space to introduce yourself. Provide *briefly* information about yourself in the following categories:

(a) Your experience (educational and otherwise) that you think may be relevant to data analysis. **(1 point)**

(b) Your computer skills (e.g., familiarity with different operating systems and programming languages). **(1 point)**

(c) Your expectations for this course (e.g., specific topics or concepts). **(1 point)**

(d) Describe briefly the kind of research you are or will be doing at UCalgary and tell if it involves data analysis. If your project is still undefined, then describe your research aspirations and how they align with the research carried out in the group you are part of. **(1 point)**

For this question, answer all four points above within a minimum of 1 page (not too short answers) and a maximum of 2 pages (not too long answers). You can use any text editor of your preference, but, please, set the document layout with the following features:

- Page orientation: portrait;
- Page size: 8.5" × 11" (letter);
- Margins: 2.54 cm (top, bottom, left, right);
- Line spacing: between 1.0 and 1.15 pt;
- No double-column format. Write your answers as a 1-column text;
- Paragraph alignment can be 'align left' or 'justify';
- Font type and size: Times New Roman, 12 pt, non-italic, non-bold;
- Use *italic*/**bold**/<u>underline</u> only to emphasize brief pieces of text if needed.
- ONLY `pdf` FILES ARE ACCEPTED. Do not submit editable files such as `.docx`.

*****

Question 2: Photon-counting experiment. Consider the following two-dimensional data (also available in the `photon_exp.dat` file), in which X is the independent variable, and Y is the dependent variable assumed to be derived from a photon-counting experiment:

x_i	y_i
0.0	25
1.0	36
2.0	64
3.0	49
4.0	81

Note that the file has an additional column n, but that is only for indexing the rows. Assume that the error associated with each dependent datapoint Y_i is the square root of the measurement (Poisson-like statistic).

(a) Find the best-fit parameters a, b of the linear regression curve $y(x) = a + bx$, the errors in the best-fit parameters and the correlation coefficient between them. The correlation coefficient is defined as $r = \sigma_{ab}^2/(\sigma_a \sigma_b)$. **(2 points)**

(b) Calculate the minimum χ^2 of the fit. Use the 'rule of thumb (practical guide) to determine goodness-of-fit' introduced in class to determine if the fitting is acceptable. **(2 points)**

(c) Present a code that computes the quantities requested in the items above. Explain your code and its outputs using comment lines, markdowns, and/or docstrings – depending on your chosen integrated development environment (IDE) software – and show a plot of the best fitting with the data points. If you are working with an IDE that does not save figures on their platform, save your figures in `.jpg`. **(2 points)**

IMPORTANT: Your code needs to output all required numbers of the exercise and they should be printed clearly. The printed values should be properly identified with text saying what is being printed. We cannot debug your code and/or search the answers inside your code for you. This is a graded component of the course, therefore, we can ONLY evaluate what is uploaded and see printed on the screen. We will test other parameter variations and settings in your code, but we will not alter its structure and logistics. Interpret your code as any other graded component as 'in paper', meaning what you submitted is your final answer and what is not provided (or printed) in terms of answers, we cannot consider for marks.

Submission Information

Name the `pdf` file you generated for question 1 as `report_assignment_1.pdf` and the code for question 2 `code_assignment_1.ext` (since this is assignment #1) in which `ext` is the file extension associated to your coding platform. You will upload the `pdf` file of your report, your code, figures, and ALL raw data files you worked on in the assignment to the Gradescope platform. Your code needs to be documented, i.e., introduce comment lines to explain its main command procedures. Pure code lines without explanatory comments will have reduced marks. **It is important that you log in to Gradescope.ca and not Gradescope.com.** You can upload multiple files to the Gradescope platform and you can resubmit your work until the due date. We will test-run your submitted code for syntax errors and check if it generates the requested figures. We will also check if the results/figures presented in your assignment are 'paper-like' quality and that the quantitative predictions are scientifically/mathematically founded.

ONLY `pdf` (for the written report), `jpg/jpeg` (for figures), and `.py`, `.ipynb`, or `.m` (for codes) FILES ARE ACCEPTED. This means we are accepting Python, Python Notebook, or Matlab codes. If you are more familiar with another programming language, please, contact the instructor in advance.

IMPORTANT: In your code include instructions on how to run it and, if applicable, include testing values for the initial conditions or settings that you attempted so we can reproduce the results. We will test other initial conditions and parameter variations to evaluate the robustness of your code. But we need a place to start. Include also information about the version of your coding platform. If this information is not included, points will be deducted.

*****