

ICA Project

Module: Programming Fundamentals (FT/BL)	
Programme: BSc (Hons) Computing	Level: 4
Teacher: Mohamed Bettaz	Semester: 2104

Project Details

Title: Analyse data, design and implement visualisation	Deadline: 24. 1. 2022 - 23:59
--	--------------------------------------

Learning Outcomes

Students are required to:

1. Develop a project based on creating a program for the given problem. (PTS1)
2. Identify and articulate the correct data structures and syntactical notions required to develop a programming code. (RKC6)
3. Identify basic project stages relevant for the development of the programming code, including time-management and selection of appropriate tools. (PS2)

Assessment Criteria and Indicative Weighting

#	Description	Points
1	Proper data cleaning: listing and briefly commenting the steps used for data cleaning. The choice of each step has to be well-argued. The methods, eventually the algorithms, used for data cleaning have to be clearly pointed out, and their use in your project clearly referenced.	20
2	Proper application design: First, describe what is expected from your application. Specify the input, the output and the process relating them without referring to implementation details. Then describe the overall structure of your application, in a way that meets the upper specifications. Concretely, the choice of variables from selected data sets (with pertinent relationships between the chosen variables), the choice of "appropriate" visualisations (created graphs, plots, maps, etc.), and the process relating them should be well-described. That is, the process of creating the application should be broken down into analysis, specification, design, implementation and Testing stages.	30
3	Proper application implementation and deployment: Appropriate exploitation of Python ecosystem, with due justification of used modules and libraries. Drawing insights from data visualisation, interpreting data, clearly communicating results. Arguing why the obtained visualisations are the most suitable for your project. The interpretation of the data and the insights drawn from their visualisation have to be well-argued.	20

4	Selection of appropriate tools and their justification: The selection of the tools is well-discussed and justified in the report.	15
5	Graded by Patrick: points will be awarded on the following basis: <ul style="list-style-type: none"> • 5 points for appropriate structure, voice, concise exposition, and formal language usage • 5 points for a correctly structured academic paper with LaTeX • 5 points for referencing at least five academically relevant sources by using the Harvard Referencing Style (1 point for each correct in-text and end-text reference pair) 	15

The total is **100** points accounting for **60%** of the final module grade.

Introduction

This second assignment is designed to utilise all topics covered throughout the entire course to create a web application using Dash python library.

This second assignment is designed to be solved individually and covers 60% of the final module grade.

This assignment assesses the student's ability to:

- Understand analysed data.
- Produce code/use method to clean data and transform them to a tidy format.
- Select suitable visualisations to trend data and data characteristics.
- Design and create a web application to show visualised data.
- Produce a report documenting the process of application creation and describing visualised data.

Tasks

The student is a freelance data scientist helping various companies with their data visualisation. The student has been contacted by a company and has been asked to help the company fully understand data related to COVID-19 pandemic. They might be interested by analysing their data for a country, a region or a continent. They would like to create a web application containing a dashboard with some graphs and a little information about the data. This project assumes that the student can manipulate data frames using appropriate Python libraries such as pandas, matplotlib, and others. The teacher will give a detailed tutorial on how to work with Python libraries. The student is requested to select pertinent datasets from those published by trusted institutions on the net (consult the terms of use and acknowledge properly the use of the selected data sets). Example datasets might be found at <https://github.com/CSSEGISandData>. In each case the student is requested to provide the URL (https) of the selected datasets.

The tasks are provided in chronological order, and to facilitate the completion of the tasks, it is recommended that the student follows them as a checklist. After the completion of each task, it is strongly suggested to review it with the tutor.

1. Select a dataset and transform it to a tidy format. An Example will be provided by the teacher through a relevant case study. (PS2)
2. Design dash application, i.e., select a suitable visualisation for the data.(PTS1, PS2)
3. Code and deploy application. (PTS1, PS2)

4. Write a technical report, maxim 1000 words (*2.5 Times New Roman single-spaced pages, size 12*), documenting steps 1-3, the selection of appropriate tools, and the basic project stages relevant for the development of the code. The report must address the relevant audience and have a correct structure, voice, and referencing. (RKC6)

Deliverables

The final outcome of the assignment is a Jupyter notebook containing code to clean data, an application file containing python code, and a technical report describing both the data and application, maxim 1000 words (*2.5 Times New Roman single-spaced pages, size 12*), which will be submitted to Moodle (documents as a pdf). The deadline of the submission will be communicated at the launch of the assignment.

To summarise, the deliverables are:

- Jupyter notebook containing a description of data cleaning, the produced code/used method, etc.
- Python file containing Dash application (*dashboards could eventually be built directly from Jupyter notebooks*).
- Link to deployed application (Heroku).
- Technical report - pdf.

Submission info

The cover sheet to your project must include the following information:

1. Your name
2. Your student id
3. semester code
4. Module

important:

Please make sure that your submission includes all the required materials and has been saved in the correct format. Failure to include all the files, or using incorrect format will result in rejection of the ICA submission.