Programming for Data - Level 8

Learner name:

|  |
| --- |
| Eleni Sakkoula |

Assessment date:

|  |
| --- |
| 31st March 2022 |

The work in this portfolio contains the following:

* Work that shows understanding of the processes and technologies used in a data programming environment.
* Work that shows evidence that the main concepts, processes and practices have been put into practice to create clean, efficient code for data analysis and is documented to show a good understanding.
* A set of functions and programs to perform data analysis written to solve a range of problems.
* Evidence of testing these functions and programs for data analysis and modifying them so that they produce the desired output.
* A reflection on learning

All evidence is contained in Colab Notebooks on Github.

Links to Github repository projects and worksheets

|  |
| --- |
| **Movies project -**[https://github.com/EllenSakkla/Python/blob/main/projects/Movies\_Project.ipynb](https://www.google.com/url?q=https://github.com/EllenSakkla/Python/blob/main/projects/Movies_Project.ipynb&sa=D&source=editors&ust=1649875262513825&usg=AOvVaw2Q4-ESHGtlMyQLwekjweK6)  **Decision tree project -**  [https://github.com/EllenSakkla/Python/blob/main/projects/Decision\_Tree\_coded\_model.ipynb](https://www.google.com/url?q=https://github.com/EllenSakkla/Python/blob/main/projects/Decision_Tree_coded_model.ipynb&sa=D&source=editors&ust=1649875262514566&usg=AOvVaw058jyoEky3qsgrMT0kKR8u)  **Page views project -**  [https://github.com/EllenSakkla/Python/blob/main/projects/Page\_Views\_Project.ipynb](https://www.google.com/url?q=https://github.com/EllenSakkla/Python/blob/main/projects/Page_Views_Project.ipynb&sa=D&source=editors&ust=1649875262515144&usg=AOvVaw1EMRr90W1co4Ys3IqYYrON)  **Sea Level project -**  [https://github.com/EllenSakkla/Python/blob/main/projects/Sea\_Level\_Project.ipynb](https://www.google.com/url?q=https://github.com/EllenSakkla/Python/blob/main/projects/Sea_Level_Project.ipynb&sa=D&source=editors&ust=1649875262515733&usg=AOvVaw2U9DxfNaUYdfRpnK61d-if)  **Tone analysis project -**  [https://github.com/EllenSakkla/Python/blob/main/projects/WatsonToneAnalyser.ipynb](https://www.google.com/url?q=https://github.com/EllenSakkla/Python/blob/main/projects/WatsonToneAnalyser.ipynb&sa=D&source=editors&ust=1649875262516294&usg=AOvVaw1JDoevk7lbIQalOXYPzvI_)  **R workshee**t - [https://github.com/EllenSakkla/Python/blob/main/worksheets/6\_R\_worksheet.ipynb](https://www.google.com/url?q=https://github.com/EllenSakkla/Python/blob/main/worksheets/6_R_worksheet.ipynb&sa=D&source=editors&ust=1649875262516880&usg=AOvVaw2bFwEQ-tMjGJMZh0m0mL6l)  **SQL worksheet** - [https://github.com/EllenSakkla/Python/blob/main/worksheets/10\_1\_SQL\_databases\_worksheet.ipynb](https://www.google.com/url?q=https://github.com/EllenSakkla/Python/blob/main/worksheets/10_1_SQL_databases_worksheet.ipynb&sa=D&source=editors&ust=1649875262517558&usg=AOvVaw3W0A3tNhCdFgmB1uhVYWh1)  **Error handling and testing worksheet** - [https://github.com/EllenSakkla/Python/blob/main/worksheets/13\_Error\_handling\_and\_testing.ipynb](https://www.google.com/url?q=https://github.com/EllenSakkla/Python/blob/main/worksheets/13_Error_handling_and_testing.ipynb&sa=D&source=editors&ust=1649875262518219&usg=AOvVaw0HFr9exgzSd-tqPWyzHd_y)  **Encoding and dummy coding worksheet** - [https://github.com/EllenSakkla/Python/blob/main/worksheets/15\_1\_Encoding\_and\_Dummy\_Coding.ipynb](https://www.google.com/url?q=https://github.com/EllenSakkla/Python/blob/main/worksheets/15_1_Encoding_and_Dummy_Coding.ipynb&sa=D&source=editors&ust=1649875262518865&usg=AOvVaw0EPT4fEIoQ08XSXn0jI4x5) |

Signposting:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **LO** |  | **Knowledge or Skill** | **Evidence of learning ( worksheet)** | **Evidence of skills and knowledge (project)** |
| 1 | 1.1 | Software types for data analysis | Low code (e.g. Excel, Power BI)  No code ( e.g. Google Data Studio)  Coded (e.g. python, R, etc) | Decision tree  Movies  Page Views  Sea Levels  watson Tone Analyser |
| 1.2 | High level programming languages for data analysis including functional languages | Python worksheets  R  Decision tree  List and dictionaries challenges | Decision tree  Movies  Page Views  Sea Levels  watson Tone Analyser |
| 1.3 | Database languages including SQL | SQL | Movies |
| 1.4 | Low code software development tools including:   * notebooks * Integrated Development Environments * Command Line Interfaces * Application Programming Interfaces * Software libraries | Data Wrangling  Encoding and dummy coding  GitBash  ToneAnalysis  Numpy  R | Decision tree  Movies  Page Views  Sea Levels  watson Tone Analyser |
| 1.5 | Programming process | Decision tree  PROJECTS | Decision tree |
| 1.6 | Source/version control | Github and gitBash | Decision tree  Movies  Page Views  Sea Levels  watson Tone Analyser |
| 2 | 2.1 | Setting up a project including software installation | Github and gitBash | Decision tree  Movies  Page Views  Sea Levels  watson Tone Analyser |
| 2.2 | Programming techniques | Data Wrangling  Encoding and dummy coding  GitBash  ToneAnalysis  Numpy | Decision tree |
| 2.3 | Data sources | SQL  Sorting and cleaning  Data Wrangling  ToneAnalysis  Numpy | Decision tree  Movies  Page Views  Sea Levels  watson Tone Analyser |
| 2.4 | Data types and data structures | Data structures  ToneAnalysis | Decision tree  Movies  Page Views  Sea Levels  watson Tone Analyser |
| 2.5 | Library functions for data analysis | Working with datetime  Numpy  Pandas | Decision tree  Movies  Page Views  Sea Levels  watson Tone Analyser |
| 2.6 | Data interfaces and data flow | API  SQL to CSV | Movies  watson Tone Analyser |
| 2.7 | Data modelling | SQL to CSV | Decision tree  Movies  Page Views  Sea Levels  watson Tone Analyser |
| 2.8 | Data security/secret management | API key management  SSH keys |  |
| 2.9 | Code reuse | User-defined functions | Decision tree |
| 2.10 | Algorithmic bias and data bias | TO BE COVERED | N/A |
| 3 | 3.1 | Writing function/program to process data from an external file | Sorting and cleaning  Data wrangling | Decision tree  Movies  Page Views  Sea Levels  watson Tone Analyser |
| 3.2 | Writing function/program to process data from a database | SQL | Movies |
| 3.3 | Writing function/program to clean data | Sorting and cleaning | Decision tree  Movies |
| 3.4 | Writing function/program to wrangle data | Data wrangling | Decision tree  Movies  Page Views  Sea Levels  watson Tone Analyser |
| 3.5 | Writing function/program to perform statistical analysis | Describing and interrogating data  Correlation and regression | Decision tree  Page Views  Sea Levels |
| 3.6 | Writing function/program to create visualisations | Visualisation 7.1 and 7.2 | Decision tree  Movies  Page Views  Sea Levels  watson Tone Analyser |
| 3.7 | Writing documentation for function/program | DOCUMENTING PROJECTS | Sea Levels |
| 4 | 4.1 | Test plans | Error handling and testing | Error handling and testing worksheet |
| 4.2 | Testing and training datasets | Encoding and dummy data | Encoding and dummy data worksheet  Decision tree |
| 4.3 | Debugging | ALL (reflect on this) | Decision tree  Movies  Page Views  Sea Levels  watson Tone Analyser |
| 4.4 | Code review | TO BE COVERED | Decision tree  Movies  Page Views  Sea Levels  watson Tone Analyser |
| 4.5 | Performing test and revising code | ALL | Decision tree  Movies  Page Views  Sea Levels  watson Tone Analyser |

|  |
| --- |
| **General feedback** |
| **This set of projects demonstrates a very good understanding of the concepts of code and data.  Good use of functions, good level of problem solving.  Reflections show a good level of understanding and learning.**  . |
| **Recommendations** |
|  |