Industry Project Week – Software Development Bootcamp

Python Photo DeDuplicator Utility

A software company released a duplicate file finder which has named it "File DeDuplicator". Duplicate files are files that are the same (bit for bit identical). Space is often wasted on computers by having multiple versions of the same file. Duplicate file finders are programs that find and identify duplicate files on a hard drive so that they can be removed. The software team that produced the File DeDuplicator decides to make the Photo DeDuplicator. This new version will be able to detect duplicated images of multiple formats – as the previous version could not- and should be able to run on a desktop computer (PC or Mac) using Python.

Tasks	Essential	Desirable
PROJECT MANAGEMENT METHODOLOGY		
1. Some members of the team suggest agile programming whilst others would prefer to use the waterfall lifecycle. The team must decide which methodology they will use for the project.	●	
2. In a presentation, discuss the advantages and disadvantages of the two methodologies and justify which you would implement for this project.	●	
IMPLEMENTATION TEAM		
3. Due to a limited budget allocation the team needs to operate without sacrificing timelines and product quality. Create three job roles essential to the implementation team to cover all aspect of the delivery.	●	
4. Outline the JD set for each role which should include up to 10 duties and responsibilities. Each JD set should have a blend of distinct and/or overlapping soft and hard skills.		●
PROJECT PLAN		
5. Create a grid consisting of job roles and development stages for the methodology selected. Allocate the responsibilities for each job role throughout each of the development stages for the methodology selected.	●	
6. Using Blue Sky Thinking, recommend all possible features that could be developed to satisfy the wider spectrum of business objectives.		●
7. Outline the software tools and skills currently available within the team that would be required to implement each of these features.		●
8. Commit to specific features and business objectives for this implementation.	●	
STANDARD PROJECT DELIVERY		
9. Create a slide presentation for the prototype. Figma or any design tool (including Microsoft Office) can be used for this presentation.	●	
10. Include a flowchart for the Python code implementation.		●
11. Write the code for the application and apply good coding practices. Outline and explain the good coding practices used.	●	
12. Produce a video presentation of the working model.		●
13. Produce a unit test table and include a minimum of 5 tests.	●	
BLUE SKY DELIVERY OPTIONS		
14. The Python application should allow the user to store the results (files to keep and files to delete) in a SQL table. Produce a formatted report reading from the SQL tables.		●
15. Offer the user an option to export the report to a .csv file		●
16. Automate the production of a single index.html file with links to Style, JavaScript or links to online resources (if needed) that produces a web page with thumbnails of the files to delete and files to keep organised in separate sections.		●

Software development tools:
- Python (any version and any libraries)
- HTML
- CSS
- JavaScript (Vanilla only)
- SQL
- Microsoft Office Suite
- Visual Studio Community (any version)
- Visual Studio Code (any version)
- SSMS Studio
- Figma
- Mind Meister
- Any collaboration tool that conforms to 'The Skills Network' safeguarding policy.

Provided Support documents:
- PowerPoint Presentation Template
- Zipped folder with test images.

Presentation:
- Maximum 20 minutes for the presentation will be allocated to every team.
- A provision of 10 minutes for questions at the end of the presentation.
- All members of the team are required to take part in the delivery of the presentation.

Deliverables:
- One zipped folder with all resources named SD_Team[X]_[ddmmyy] will be emailed to the tutor john.miaris@theskillsnetwork.com by each group.
- An upload of the PowerPoint presentation by each learner to the EQUAL platform.
- The folder will contain the below structure of subfolders and resources:
  - (folder) Presentation
    - (file) The PowerPoint presentation.
  - (folder) PhotoDedup
    - (file) PhotoDedup.py
    - (file) Prototype in any format
    - (file) Flowchart in any format
    - Other files
    - (folder) Images
  - (folder) Automation
    - (folder) HTML
      - (file) index.html
      - (file) main.css
      - (file) main.js
      - (file) Other external link files
    - (folder) BlueSky
      - (file) CreateTables.sql