GROUP PROJECT DESCRIPTION (WIA2005 - Algorithm Analysis and Design)

University/Programme/Course: University of Malaya/Bachelor of Computer Science/Algorithm Analysis and Design**Year: 2nd** year / 4th semester

Pedagogical Approach: Project-based Learning

Learning Outcome:

Experience analysing and designing algorithms for problem-solving with other teammates.

a. Utilise all of the chosen tools

b. Apply algorithms that have been learned in this course to solve the given problems.

c. Explore at least one new algorithm to solve the given problem.

d. Integrate all of the algorithms within a computer program.

e. Execute the computer program while explaining the relation between steps in algorithms with the behaviour/output of the computer program.

f. Analyse the complexity of the main algorithms that solve the given problem.

g. Function effectively as a team member.

h. Communicate effectively through report and presentation.

Objective:

This project requires you and your teammates to analyse, design, and code a computer program using Python and the chosen tools to solve the given problems.

Project Scope:

To meet the project requirements, you will need to:

✔ Form a work team of 6 members (must be within the same tutorial session).

✔ Elect a team leader, write a contract item and sign using the group contract template in Appendix A.

✔ Identify clear roles and responsibilities, distributing and coordinating various tasks appropriately, and able to operate as a high performing team. You must clearly communicate how you have worked as a team. Refer to FILA form in Appendix B.

✔ Analyse, design, and code a computer program using Python and the chosen tools to solve the given problems as the following:-

One of the essences of computer science and information technology is to solve problems faced by humankind. As the outcome of this project, you are required to develop a computer program that is able to resolve the following problems:-

Problem 1: Customer who needs to make a delivery needs to know which courier company can does it fast (assuming that the shorter the distance, the quicker is the delivery). The application will analyse five (5) local courier companies which have their delivery hubs located in various locations in West Malaysia. The details of the courier companies and their delivery hubs are given in Table 1.

Courier Company	Delivery Hub	Coordinate
City-link Express	Port Klang	3.0319924887507144, 101.37344116244806
Pos Laju	Petaling Jaya	3.112924170027219, 101.63982650389863
GDEX	Batu Caves	3.265154613796736, 101.68024844550233
J&T	Kajang	2.9441205329488325, 101.7901521759029
DHL	Sungai Buloh	3.2127230893650065, 101.57467295692778

*Hub locations are fictional – created for the purposes of this project.

Table 1: Details of courier companies and their hub locations.

1. Get and mark locations of all hub locations given in Table 1.
    a. Guide 1: you can use Python Geocoding Toolbox
       Lookup: https://pypi.python.org/pypi/geopy#downloads
    b. Guide 2: you can use gmplot
       Lookup: https://github.com/vgm64/gmplot
2. Let say, three (3) customers would like to make a door-to-door parcel delivery using any one of the courier companies, details as given in Table 2. Get the distances between origin and destination.

Customer	Origin	Destination
Customer 1	Rawang (3.3615395462207878, 101.56318183511695)	Bukit Jelutong (3.1000170516638885, 101.53071480907951)
Customer 2	Subang Jaya (3.049398375759954, 101.58546611160301)	Puncak Alam (3.227994355250716, 101.42730357605375)
Customer 3	Ampang	Cyberjaya

	(3.141855957281073, 101.76158583424586)	(2.9188704151716256, 101.65251821655471)	

Table 2: Example customer delivery requests.

    a. Guide 1: you can use Python Geocoding Toolbox

    b. Guide 2: you should use Google Distance Matrix API

        i. Login to the google developer's website and follow through the examples. It is important that you know how to use the API key given to you within the code that you are going to use. Refer to this link: https://developers.google.com/maps/documentation/distance-matrix/start

3. Assuming that all deliveries must go through their delivery hub. For example, a delivery from Kuala Lumpur (origin) will go through Pos Laju Hub in Petaling Jaya to Putrajaya (destination). Suggest the least distance that the parcel has to travel for each customer using every courier company.

4. Plot line between the destinations before and after the algorithm (defined in 4) is chosen.
    a. Guide1: you can use **google.maps.Polyline**. You can refer to this link:

https://www.sitepoint.com/create-a-polyline-using-the-geolocation-and-the-google-maps-api/

Problem 2: Shortest distance travelled does not mean that the courier company is the most recommended option for customer to use. A sentiment analysis of the related news articles about the courier company must be conducted.

5. Extract information from three (3) articles from online news websites that have published stories related to each courier company.
    a. Sometimes a webpage must be converted to the text version before it can be done.
        i. Guide 1: You may refer to this website to extract word from a website
          https://www.textise.net/
    b. Guide 2: You may refer to this website on how to count word frequency in a website https://programminghistorian.org/lessons/counting-frequencies
    c. You can also filter stops words from the text you found.
        i. Guide 3: Stops words such as conjunctions and prepositions. You may refer to this link: https://www.ranks.nl/stopwords
        ii. Program using a string matching algorithm to find and delete the stop words.

6. Plot line/scatter/histogram graphs related to the word count using Plotly (Word count, stop words)
    d. Guide 3: You may refer to this link on how to install Plotly and how to use the API keys

http://www.instructables.com/id/Plotly-with-Python/

https://plot.ly/python/getting-started/

7. Compare words in the web pages with the positive, negative and neutral English words using a String-Matching algorithm
    e. Guide 4: You may use the following word as positive and negative English words
        http://positivewordsresearch.com/list-of-positive-words/
        http://positivewordsresearch.com/list-of-negative-words/

    f. Put these words in a text file for you to access them in your algorithm.

    g. Words that are not on the list can be considered as 'neutral'.

8. Plot histogram graphs of positive and negative words found in the web pages.

    h. Guide 5: Use Plotly

9. Give an algorithmic conclusion regarding the sentiment of those articles

    i. Guide 6: If there are more positive words, conclude that the article is giving positive sentiment, if there are more negative words, conclude that the article is giving negative sentiment.

    j. You may try to conclude in different perspectives such as whether the list of positive and negative words above is accurate to be used in the context of the article you extracted the text.

    k. Based on the conclusion, you may say which courier company have the best sentiment.

Problem 3: Customers need to be able to choose the best courier company based on the distance as well as the result of sentiment analysis of related online articles.

10. Calculate the total probability distribution of possible routes. Then, write the summary of the courier companies, ranking from the most recommended to the least recommended based on distance and sentiment.

Problem 4: Assuming that video or audio from the news or customer feedback will be used to provide sentiment insights in the future, Dynamic Time Warping (DTW) is one of the algorithms that can potentially be used. Explain the concept of DTW and demonstrate the implementation of DTW in analysing a video or an audio. For example, given the following video (https://www.youtube.com/watch?v=ZwVFj8CfFeE), identify some words, for example "J&T", "memohon" and "maaf".

✔ Week 6: Only the group leader has to submit the following in a .zip file through Spectrum:
    ▪   **Group Contract**
    ▪   **1st FILA form (compulsory to all members to develop the program)**

✔ Week 6-12: Analyse, design, and code a computer program using Python and the chosen tools to solve the given problems.

✔ Week 12: Each student needs to individually submit the peer evaluation form through Google Form (Link will be provided later). Your peer evaluation is confidential and will not be exposed to other team members.

✔ Week 13-14: Each group has to do 30 minutes of system demonstration and presentation using the Powerpoint slides. All group members must present and have the camera switched on.

✔ Week 14: One final submission for each group. Only the group leader has to submit the following in a .zip file through Spectrum:
    ▪   **Source code: raw python files (All programming codes must use python 2 or 3)**
    ▪   **2nd FILA form (compulsory to all members to develop the program)**
    ▪   **Report Content:**
    1. Introduction
    2. Description
     - Elaborate how tools and algorithms resolve each of the given problems.
     - Highlight at least one new algorithm that has NOT been taught in this course to solve the given problems.

3.   Explain the time complexity of each of the algorithm that resolves the given problem.
4.   The program code: source code and snapshots of input/output.
5.   Conclusion
6.   References

Assessment Criteria:

The assessment for this group project is divided into two categories:
- Assessment criteria for soft skill as described in Table 1
- Assessment criteria for algorithms in solving the given problems as described in Table 2

Table 1: Assessment criteria for soft skill (Individual Assessment)

Skill level	Score Description	Partially meets 1	Meets 2	Exceeds 3-4	Exemplary 5
CS1 (KIM) **Presentation (2%)**	The ability to present ideas clearly, effectively and confidently, in both oral, written forms Oral Parameters: ● delivery, ● projection (pace, volume, enunciation) ● appearance (attire and demeanor)	Either one parameter is acceptable.	All parameters are acceptable.	Some parameters are exceptional.	All parameters are exceptional.
TS4 **FILA form (3%)**	**The ability to contribute towards:** ● **planning,** ● **coordination** **of the team's efforts**	The student is able to contribute towards any one task	The student is able to feasibly contribute towards both tasks.	The student is able to contribute towards both tasks well.	The student is consistently able to contribute towards both tasks excellently.

Table 2: Assessment criteria for algorithms in solving the given problems (Group Assessment)

Criteria	Scoring			
Accuracy / Content Knowledge	**5**	**4**	**3**	**2-1**
	All algorithms and the usage of tools are presented, executed without error and output appears to be accurate.	Almost all algorithms and the usage of tools are presented, executed without error and output appears to be accurate.	Most algorithms and the usage of tools are presented, without error but output appears to be less accurate.	Some algorithms and the usage of tools are presented, executed with minor/major error, resolved with hard-codes, output appears to be accurate.
-Algorithm to resolve Problem 1 (5%)				
-Algorithm to resolve Problem 2 (5%)				
-Algorithm to resolve Problem 3 (5%)				
Integration between items (5%)	**5**	**4**	**3-2**	**1-0**
	The entire system appears to be integrated	Most system appears to be integrated	Some system appears to be integrated with some hard-coding	Minor integration between items 1-10 with hard-coding
Include new algorithm that is not taught in the course (5%)	**5-4**		**3-1**	**0**
	Adopt or created at least one new algorithm that resolves Problem 1, Problem 2, or Problem 3.		At least one new algorithm that resolves the sub-problem.	None
Problem 4 (5%)	**5-4**		**3-1**	**0**
	Able to implement Dynamic Time Warping to analyse video or audio.		Able to explain the concept of Dynamic Time Warping	None

Appendix A

| Semester 2 2020/2021 | **WIA2005: ALGORITHMS ANALYSIS AND DESIGN** |
| | **GROUP CONTRACT** |

A team of at most 6 students to

●   Declare and identify individual strength

●   Identify individual role in the team

●   Agreed on meeting time, venue, communication means and approaches to arrives at any decisions

●   Develop team social contact

Deliverable / To submit

Each member role and contract

Group Leader:

Contract Item: As a team, we agree to	
● Participation	
● Communication	
● Meetings	

● Conduct	
● Deadlines	
● Conflict	

Clause In any violation of the above, we agree	

Please ensure that the items in the clause is effective and feasible.

No	Matric No	Name	Team Role	Signature

(Assessor: Date Received:)

Appendix B

FILA FORM – University of Malaya				
FACTS	**IDEAS**	**LEARNING ISSUES**	**ACTION**	**DATELINE**
What we know about the task		What do we need to find out?	Who is going to do it?	
- The phases or process to develop or model the system - Integrate theory and practical	- To analyze and translated the theory into practical or real scenario - To transfer the knowledge and suited with various of industries and provide the solution and improvement		- Activities and group discussion - Find research information through internet or library	immediately