GEOGRAPHIC INFORMATION SYSTEM (GIS) PIPELINE INSPECTION AND VISUALIZATION

by

NURAREESYA AMALIN BINTI ADNAN

17002970

DISSERTATION OF FINAL YEAR PROJECT REPORT

BACHELOR OF INFORMATION SYSTEMS (HONS.)

Universiti Teknologi PETRONAS

Bandar Seri Iskandar

32610 Tronoh

Perak Darul Ridzuan

SEPTEMBER 2021

GEOGRAPHIC INFORMATION SYSTEM PIPELINE INSPECTION AND VISUALIZATION

by

NURAREESYA AMALIN BINTI ADNAN

17002970

Dissertation submitted in partial fulfilment of

the requirement for the

BACHELOR OF INFORMATION SYSTEMS (Hons.)

SEPTEMBER 2021

Universiti Teknologi PETRONAS

Bandar Seri Iskandar

32610 Tronoh

Perak Darul Ridzuan

CERTIFICATION OF APPROVAL

GEOGRAPHIC INFORMATION SYSTEM PIPELINE INSPECTION AND VISUALIZATION

by

NURAREESYA AMALIN BINTI ADNAN

17002970

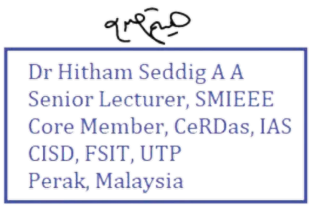
A project dissertation submitted to the Information Systems Programme

Universiti Teknologi PETRONAS

In a partial fulfilment of the requirement for the

BACHELOR OF INFORMATION SYSTEMS (Hons)

Approved by



(**Dr Hitham Seddig Alhassan Alhussian**)

UNIVERSITI TEKNOLOGI PETRONAS

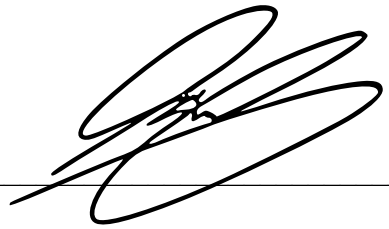
BANDAR SERI ISKANDAR

32610 TRONOH

PERAK DARUL RIDZUAN

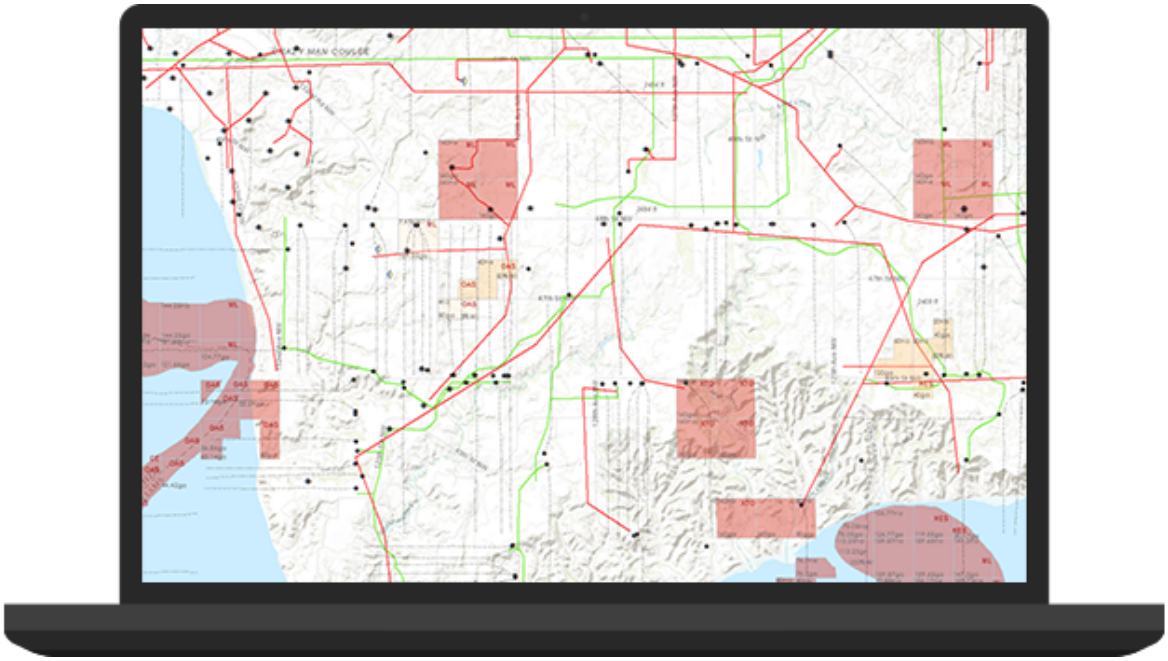
CERTIFICATION OF ORIGINALITY

This is to certify that I am responsible for the work submitted in the report project, that the original work is my own except as specified in the references and acknowledgement and the original work contained herein have not been undertaken or done by unspecified sources or persons.



NURAREESYA AMALIN BINTI ADNAN

ABSTRACT



Geographic Information System or GIS is a conceptual framework and a spatial system that specifically functions to develop, manages, inspects, and maps all types of data. Creating foundation of mapping used by GIS experts, this system helps to integrate locations with numerous of information for analysis in various industries. This report studies an in-depth knowledge about GIS pipeline inspections and visualization. Inspections involves controlling the utilities on mapping GIS coordinates to execute findings of the analysis process. Furthermore, GIS visualizations involves a broader range on modelling the data sources into an understanding manner. This project concerns on GIS inspections and visualization specifically for the usability of the external users that need to analyse and evaluate GIS information to elucidate their judgements on the pipeline's mapping output. To reach the target output, the suitability of following CRISP- DM as the methodology helps to solve the project problems and use case. Moreover, with the advancement of open-source packages and libraries and full stack approach helps this project to shape a better output.

ACKNOWLEDGMENTS

Alhamdulillah, praise to Allah, the Almighty, on whom ultimately we depend on for sustenance and guidance. The One that gives me so much strength and chances to be a responsible, hardworking and good student throughout my final year in UTP.

For the past 6 Months starting from 3rd May 2021 truthfully, I consider myself as a lucky person to receive this GIS Inspection & Visualization project. I would like to express my upmost gratitude for His blessings and kindness for allowing me to strive on approaching the completion of this ambitious project especially during the tough times. Honestly, I am grateful to have grabbed this chance that gave me something to gain back in return.

Firstly, would like to express a deepest sense of gratitude and love to my best supporters which is my family that always support, motivate and advices me throughout this year. I will always make this project as a comprehend value for my work development. Skills and knowledge obtained throughout this six -months that will always be appreciated and ventured. I am hoping that I will continue to normalize this culture for my future as well as for my own benefits.

Not to forget , I would like to give a special thank you and gratitude to Universiti Teknologi PETRONAS (UTP) for making this happened. Deepest thanks to both of my supervisors, Dr Hitham Seddig Alhassan Alhussian as my Final Year Project supervisor and Dr Nordin as the assist supervisor that also assist me on trainings, giving most of the advices and guidance in order for me to do well and better throughout this final year project period. In addition, would like to give a special thanks to my Graduate Assistant, Mr. Mohamad Gamal Ragab that also assist me on training throughout the project deliverables.

I am overwhelmed in all humbleness and gratefulness of having great friends around throughout the development of this course and this project deliverables. I am honoured to have such a talented friends which in a heartfelt, would like to thank Fatin, Khairina, Yasrani, Tengku Sya, Tia, Haikal, Qayyum, Suhaimi, Azizul, Ilham, Arsh, Aiman and Syafiqah for the knowledge and guidance shared, foods especially during lunches, joke and laughs that we shared together during stressful moments focusing our minds, hearts and should on completing this project. Would also want express my gratitude to the people who has helped me indirectly on my research and development of this project.

TABLE OF CONTENTS

TABLES AND FIGURE CONTENTS

CHAPTER 1 INTRODUCTION

1.1 Background of Study

In Malaysia, oil, and gas acts as a vital necessity since it was firstly discovered in 1910. Even the whole world needs oils and gases for their everyday life's products such as for cooking, cars to drive, machines to reduce workloads, oil refining to produce plastics, lubricants tars even asphalts for roads. Geography however plays a significant role for the oil and gas industry to operate plus acts as an integral part on economy and infrastructure in most countries. Oil and gas industries will initiate the work from identifying the location, extraction, field of upstream and downstream management to resource transportation that requires geographical elements to reach to its output.

This is the reason why we need Geographical Information System (GIS) which acts as power of data manipulation and analysis plus aimed to map or image reproduction of socioeconomic or population related data. Moreover, GIS focuses are based on the data accuracy, availability, accessibility, and realistic model of the real globe. This system currently does not follow the traditional spatial analysis mapping such as cartographic documents. Based on Figure 1, Cartography is the study of maps and practice to of combining various techniques of sciences of the real- life imagination or reality. It also gives an effective communication on applying information towards the audiences.

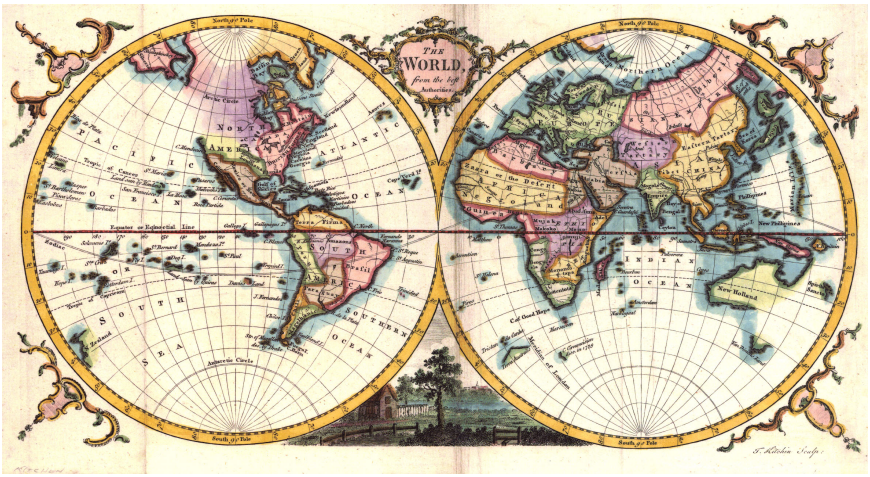


Figure 1 Example of Cartography in ancient days.

GIS visualization can be represented into **two (2)** features such as spatial and non-spatial feature. Spatial is an analytical technique that studies the location of base mapping that distribute the relationship between the geographical phenomena. It allows numerical values to detect the locations of the Earth. One the functionality of GIS Pipeline in Oil and Gas Industry is to ease GIS expertise and its organizations on conducting or visualizing their GIS projects effectively and efficiently. For example, to make data creation, analysis, for maintenance and storage in the pipeline industry.

GIS have its own coordinate systems, projections, and transformations. In geographical perspectives, data is defined in both horizontal and vertical coordinates (ArcGISPro, n.d.). GIS consist of 2 types of coordinates system. Firstly, is GCS or Geographic Coordinate System acts as a framework measured in globe-spherical shape. GCS consist of longitude and latitude lines hence also acts as a measurement numerical unit to locate locations of the map. Secondly, PCS or Projected Coordinate System are based on flat map surface which uses Universal Transverse Mercator (UTM) as its unit in the form of Easting or Northing. Transformation happens when data are converted into another coordinate systems. In this project, some conversion will be involved between UTM to Longitude and Latitude. GIS thus is very important technology for evaluation of the potential for oil and gas in promising location.

Both coordinates depend on Ellipsoid and Datum. Ellipsoid designated similarly depending to the alteration of a circle or sphere. Taking Earth as an example, its shape depends by a smooth Ellipsoid surface such as Figure (2) below. Scientists and mathematicians use the ellipsoid model to more accurately define the shape of the Earth and other planets (DJI Enterprise, 2021).It is best described as one of the GIS fundamentals for dimensional scaling. However, in GIS, it is quite uneven to use only Ellipsoid without its Datum. A Datum is a mathematical model that approximates the shape of the Earth to enable accurate position, length, and area calculations (GIS & GPS Tips and Techniques , 2014). Every each of the pipeline that needs to be identifiable, inspect and locate are based on Datum such as WGS84 represented in dot lines as Figure (3). Also, different areas or country have different Datum as describe in Figure (3) such European Datum or North American Datum which identifies the centre mass of the earth. Placing the wrong Datum during the visualization the pipeline may risk the calculations and analysations of developing the right GIS Model.

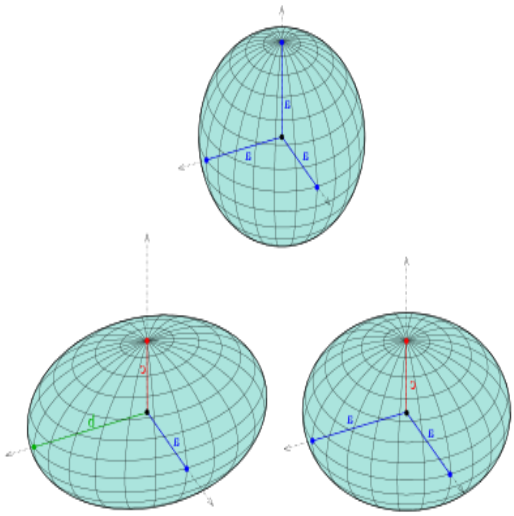


Figure 2: Examples of Ellipsoid

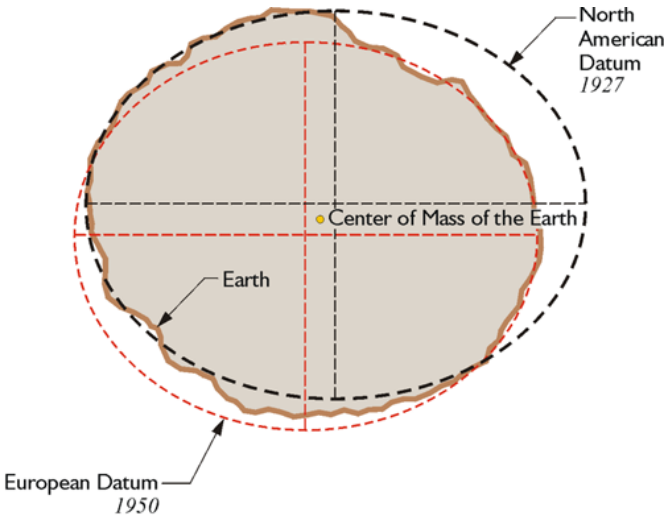


Figure 3: Example of different Datums.

1.1.1 History of GIS

The history of GIS started with spatial analysis that was map by Charles Piquet on the case of Cholera Outbreak in Paris in 1832. In 1854, John Snow has used the same concept to illustrate the concept of Cholera death in London. As an improvement of Spatial analysis made by Charles piquet in London, John snow has put an argument on his method of analysis. In the early of 20th century, the were separated into layers without any amend on the analysation. A printing technique called photozincography was introduced, which allowed users to separate layers from a map (Tate, 2018). In 1960, GIS was introduced by Roger Tomlinson as a method of enhancing digital maps in an understanding manner. Forward to 1970, few countries national agencies adopted the new concept of GIS. Hence, it was brough into present adding more functionality for the method on emphasizing urban area efficiently including organizing cities.

1.2 Problem Statement

In Oil and Gas Industry such as PETRONAS, the current local Geographic Information System does not par with the international GIS coordinates. Furthermore, pipeline inspection vendors need the international GIS coordinates to perform their inspection process. PETRONAS in Malaysia, however, it is suitably to use Universe Transverse Mercator (UTM) coordinates to inspect and analyse their pipeline due to differential of datum. Datums is a reference for measuring the locations on the surface of the Earth depending on locations. In local GIS we're using different datums that fits with different types of coordinates system called UTM, however, the international GIS uses different datum with GCS (Long Lat) coordinates.

Therefore, they will ned a platform that can convert the geographic coordinate system to continue pipeline analysation, projection, and transformation. For example, from GIS Universal Transverse Mercator (UTM: 2- Dimensional) to GIS Longitude and Latitude (LNG/LAT: 3-Dimensional).

Besides, PETRONAS have an issue on giving the access of Arc GIS, (act as a software for geographic information system expertise) on their permission to the external parties such as to the person who handle external projects that involve geographical analytics and inspection of data. This problem may consume longer time for users and vendors to analyse the data. Example of pipeline vendors are as below:

- PlugCo. Distributor in Miri, **MALAYSIA**. ...
- ADSI Industries Sdn Bhd. based in Ampang, **MALAYSIA**. ...
- Vastas. Distributor in Ampang, **MALAYSIA**. ...
- Solmax. Office in Pelabuhan Klang, **MALAYSIA**. ...
- NACE International. Office in Kuala Lumpur, **MALAYSIA**. ...
- AGV Environment Sdn. Bhd. ...
- Rotork plc.

1.3 Objectives

As a solution for this project, it will be best to develop a dashboard to visualize the converted coordinates to ensure the data are consistent with the pipeline GIS coordinates. In addition, external users that collaborate projects with PETRONAS may reduce cost and time to overcome and progress their work. Another solution is by using an interactive web map that can easily lend them a platform to encounter errors from test data sets. Data and information involved on the coordinate conversion, projection and inspection would be more accurate and precise. Upon questioning, there are 4 relatable objectives to enhance this problem such as:

- To study various coordinate systems such as UTM and Latitude and Longitude. Understanding both concepts help users and developers on the basic education on geographical spatial analysis and its usability.
- To design a web-based tool to convert or project the UTM coordinate systems into Latitude & Longitude coordinates.
- To visualize a dashboard to test the converted GIS coordinates using front- end and back-end tool.
- To involve data profiling and error inspection of the coordinates in this project. Data profiling is the process evaluating the source data into an understanding structure. This may help improving the content of the data visualizations. Error inspection, however, will help users to identify potential improvements on the data correlations from the issues on modelling the algorithms or from data samples itself.

1.4 Scope of Study

The study of the work is to visualize and understand the knowledge being applied towards this GIS Pipeline and visualisation project. The scope prioritizes Data Visualization. Data Visualization gives users the opportunity uses charts, maps, and graphs to patterns huge amount of data into an organized and understanding manner thus improving most of the business entities these days. Likewise, knowledge in data visualization help users on giving comparisons between data whether it consist of errors or differential in magnitude levels. However, for the progress of this project mostly use CRISP -DM model to pattern out my understanding, research and finding on the project output. CRISP- DM are usually used in Machine Learning project but somehow, it is also usable for visualization that involves building data algorithms and manipulate the data. Based on the data sciences definition, machine learning is a method of data analysis that uses algorithms to learn from your data (Boost LABS, 2017). Not to forget, this study provides a scope on introducing MVT Patterns based on language framework used. Throughout the study, MVT Patterns are the concept used to build the model. This exposed a person into various ways on the knowledge of web-development.

1.5 Significance of Study

- Values
- Time
- Creativity & Quality

There are 3 major significance that can be pointed out regarding GIS Pipeline Inspection and Visualization as the 3 bullet points above. One of the significant of the studies are on values gained towards the output, my-self, and the end users. Learning and doing research about GIS basic helps widen a person in depth -knowledges about the correlations between geography and data visualizations. Relating with IT -based, it gives a person a broader understanding to find, prepare and manipulate data while ongoing this project. Having a basic own understanding on the output trying to be implied, will also help to understand others as well which it may attract users who will soon and interested using this platform to help themselves represent their information, solutions, and prominent decisions.

Next, “Time” are also relatable on the significance of the study. Together it brings better accuracy of describing the data with prove analysis hence the visualization reduces the workload for decisionmakers and GIS experts or external parties to utilize their analysis and visualization projects.

Lastly, are on Creativity and Quality points of view. The significance gained from this study involves various types of charts and method of visualizations. Data visualization demands creativity to both developer and end -users. Presenting and transforming huge complex information for business using suitable and adaptable tool make the visualization to be more creative plus full of quality. Not to forget, this project involves Web- API to help users visualize the pipelines and analytics in the web browser.

CHAPTER 2: LITERATURE REVIEW

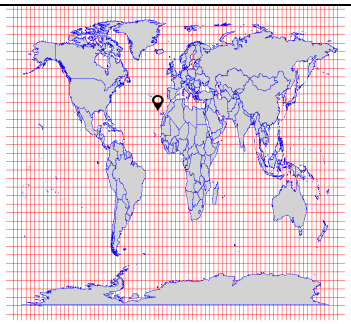
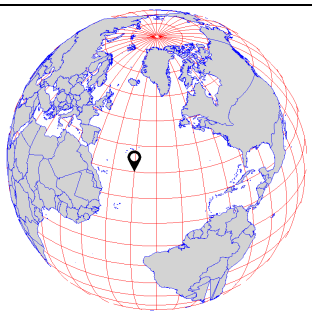
2.1 Comparative Studies Between UTM and Latitude/ Longitude

In geographic information systems, there are two types of coordinates systems that can easily be differentiated. To differentiate both PCS and GEO, firstly will be on the dimensionality. As mentioned in the introduction, the GIS coordinates can be categorized in two such as Projected Coordinate System (PCS) and Geographic Coordinate Systems (GCS).

Project coordinate system describes the geographical Earth surface of the map in 2-Dimensional shape or so called as a flat map or local GIS for industries in Malaysia. Based on (Hussain, 2017), a map projection cannot be a perfect representation, because it is not possible to show a curved surface on a flat map without creating distortions and discontinuities. In certain industries and countries preferably to use GCS or geographic coordinates system uses 3- Dimensional shape since it consists of Datums that determine better accuracies and approximation on coordinating the pipeline used in international Oil and Gas Pipeline Vendors. GCS uses longitude and latitude coordinate measure to map its data while PCS use Universal Transverse Mercator (UTM). He also described that (Hussain, 2017), a datum defines the position of the spheroid relative to the centre of the earth.

Next, in terms of projection and visualization on PCS, paper maps, web maps and mapping applications represents the Earth in two dimensions and therefore require spatial features on its surface to be projected (Thomas, 2017). I can support this, because 2D is visualized as a flat map surface which hardly for the analytics to identify the spatial data layers to reach this project objectives. (Thomas, 2017) also described, a perfect three-dimensional scale model of the Earth is the only way to eliminate all four distortion types. For example, by area, shape, scale, or angle. To help support this literature review, Table 1, shows the structured comparisons between UTM and Longitude or Latitude.

Table 1 Difference between Universal Transverse Mercator and Longitude/ Latitude

Universal Transverse Mercator UTM (Easting /Northing)	**Differences**	**Latitude / Longitude (GCS)**
• A map projection cannot be a perfect representation. • 2-D Projections.	**Dimensionality** (Hussain, 2017)	• A datum defines the position of the spheroid relative to the centre of the earth. • 3-D Projections.
• Paper maps, web maps and mapping applications represent the Earth in two dimensions and therefore require spatial features on its surface to be projected.	**Functionality** (Thomas, 2017)	• A perfect three-dimensional scale model of the Earth is the only way to eliminate all four distortion types • Such as by area, shape, scale, or angle.
	**Example Figures**	

CHAPTER 3: METHODOLOGY

3.1 System Architecture Using CRISP – DM

The project system architecture will be related with geographic spatial analysis. It is a type of analysis that explains patterns of the human's behaviours towards each layer of geographical or locational layers in terms of mathematic and geometry. So, mapping the GIS coordinates requires understanding the concept of urban systems such as Street Data, Building Data, Vegetation, and Integrated Data such as. Street Data displays land- use patterns such as the location of parks and housing complexes. Buildings Data however, functions detect location of streams such as pipeline, houses, and mall. Not to forget, vegetation data detects different kinds of soils and tress on the Earth surface. The last layer will be the integrated which data will be collected summing up street data, building and vegetation such as population, income of the population and education level.

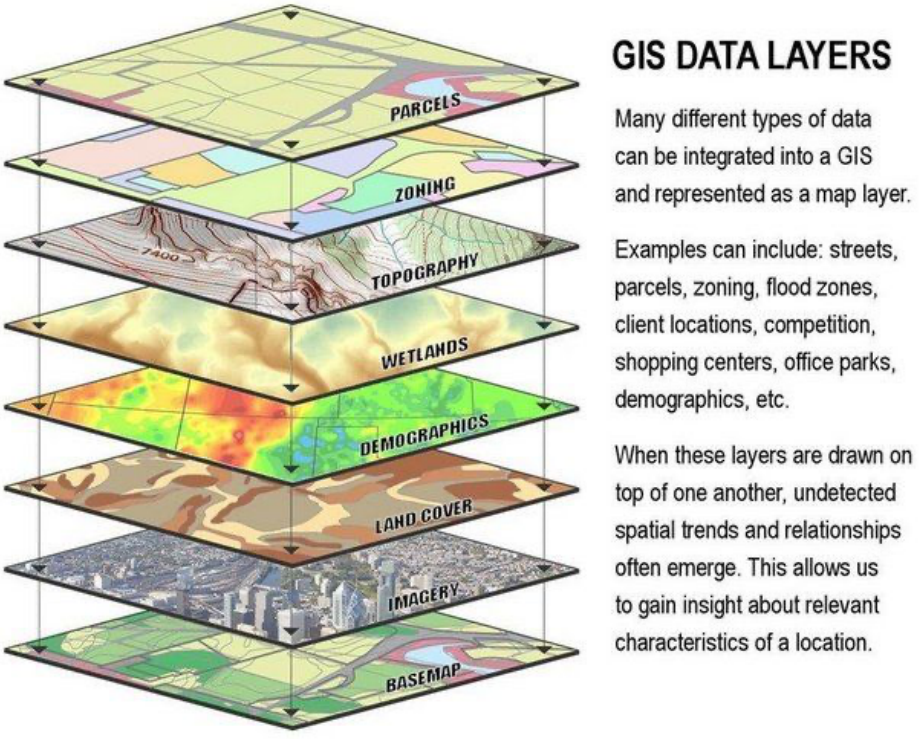


Figure 4 GIS Data Layers

The concept of building or visualizing this GIS coordinates are on the web development that are interconnected between operations and user. The first element that needs to be created is the web- based GUI. The content or types of modelling in the web- based GUI involves line and scatter charts. The front- end language will be HTML and the back- end will be Python Django.

In this thesis, my method structure would be based on the system architecture to achieve this topic's objectives. The methodology will be based on CRISP DM Model. What is CRISP DM Model? CRISP or Cross Industry Standard Process for Data Mining is a framework for building machine learning systems. Usually, it has always been describing as a data science life cycle where this model helps to plan, organize, and implement machine learning projects. As you may refer to Figure 5 below, CRISP -DM consist of 6 phases which is Business Understanding, Data Understanding, Data Preparation, Modelling, Evaluation and Deployment. Not to forget, CRISP can be both Agile or Waterfall depending on the flexibility and how people make use plus implement it.

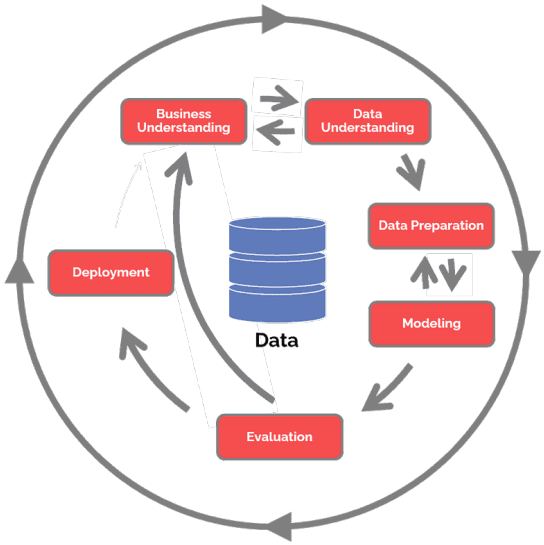


Figure 5: CRISP- DM Model Phases

3.1.1 Business Understanding

3.1.1.1Business Objectives

The business objective focuses on reaching the aims of the GIS usability, functionality, and efficiencies. In corporate entities, employees may use this GIS visualization to understand further data and information given and to easily make great business decisions rather than making random assumptions. The benefits of this project that can be accomplish is when user gets a wide understanding on the 3D visualization. Therefore, the business objectives are as below:

- To investigate the conversion or projections of the pipeline coordinates.

- To point the converted coordinates into the base map which displays information about new location of the pipeline.

As mentioned above, this project referred based on geographic spatial analysis that functions to combine geographical information from various resources. By using this concept, users can predict new sets of information or decisions or answering complex GIS questions such as the best location for building oil pipeline. Moreover, this concept helps to detect and quantify problems including patterns of the visualization such as error precisions.

3.1.2 Data Understanding

Data understanding acts as an initial part of this project progress. Without data understanding, this would not help solve any of the problems hence, it is hardly to model the solutions. On this phase, the first step is to collect the resources before any sort of website or cleaning being done. It is best to collect raw data set as it best described in terms of its originality. The output depends on the quality of data we are using. Developers usually use this method to find several unexpected values or missing values depending on logical perspectives and comparison between one data set with another. Not to forget, the type of formats involves in data collection will be in the form of CSV and XML- based file format. In definition, Comma Separated Values (CSV) is just a plain text file containing lists of data often used by separation of commas while Microsoft Excel Open XML Spreadsheet (XLSX) is an extension of Microsoft Excel. This type of file usually been compressed into a ZIP file whenever the data file is large. It functions as a data set to contain, organize, and store data either text or numerical including mathematical formulas that can be manipulated.

Table 2 Metocean_PL325 Data Set

MetoceanPoints_PL325

	Easting	Northing
0	688653.687	481027.2948
1	694241.7978	464451.8388
2	699806.0084	458934.9245
3	705370.9192	453418.0658
4	705396.1483	442358.7621
5	710962.1147	436841.6433
6	716554.032	420264.6509
7	722121.6373	414747.2411
8	722146.5619	403687.0478
9	722170.8136	392626.876
10	727739.5949	387108.793
11	727763.4196	376048.3115

Next, is to understand the data. In Table 2, shown a sample data set receive as an initial start of the project progress. Easting and Northing are the current UTM coordinates received before it needs to be converted into GCS format. To reach the objectives of this project, there are 2 sorts of PETRONAS data sets received and important such as Metocean Pipeline data set, and GIS Pipeline Data set. 2 of these different data sets are all in CSV form. Also, each of these data sets or file categorized by pipeline numbering or id such as below:

Table 3 Types of Data Sets containing its own ID.

Metocean Data Set	GIS Dataset
Definition: Subsea cables pipeline.	**Definition:** Pipeline that provides tools for maintenance, analyzation, and creation of the GIS Map.
Functionality: Deliver oceanographic and environmental information such as water depth, waves,	**Functionality:** Acts as monitor and maintenance of pipeline honesty and analyse such

and site-specific marine such as oil spill trajectory.	errors might occur in the urban mapping areas. Involves Longitude and Latitude measurement based on Datum.
Data Set Received: • MetoceanPoints_PL325 • MetoceanPoints_PL327 • MetoceanPoints_PL393,	**Data Set Received:** • PL315_GIS • PL393_GIS • PL327_GIS

Let's look on Table 3 that represent Metocean data set. This is the sample of raw data set with the absence of longitude and latitude. After the conversion process executed (described in modelling phase), the output document will add another 2 rows besides Easting and Northing, making the dataset having both GCS (Latitude, Longitude) and UTM (Easting, Northing) coordinates together. The rows are provided with numbers of attributes to present the visualization.

3.1.3 Data Preparation

Step 1: Create a folder for the data sets in an organize manner.

Before cleaning the data, it is best to start a structure step hence create a folder and placing it in an organize spot will not confuse the developers and users. Figure 6, shown the datasets collected in an organized manner. Developer should recognize the location of the folder such as /Users/areesyaadnan/Desktop/GISFYP/Data for an input files and /Users/areesyaadnan/Desktop/GISFYP/Data/Pipeline for output files after coordinate conversions.

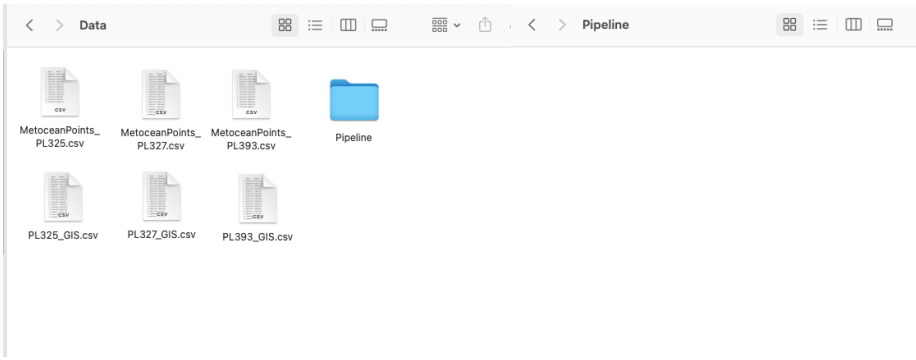


Figure 6 Data Set Files for Input and output folder renamed as Data and Pipeline

.

Step 2: Remove Complications in the data set.

This is the phase where developer should start cleaning the data so data will be more consistent and much easier to visualize it. However, the data set does not have any complications.

3.1.4 Modelling

In building a good output, the cleaned data does not only help to build the presentation, but the data also functions to be manipulated in the visualization system. Without user interfaces and data manipulations, the presentation of the visualization and explorations of its insights will never be achieved. This dissertation introduces three concepts of the important process that acts as a solution to achieve the target output, which is Front- End Development, Back End Development and Full Stack Development. These concepts are important to create beautiful view and interactive web – application framework. Not to forget, all these concepts create effective and efficient interactive dashboard fulfilling the site with a drop-down bar, zoom in and out features that are functioning. Furthermore, acquiring this process leads to some insights of the target output throughout displaying this presentation to the end- user.

The Concept of Front- End, Back-End and Full Stack Development

Based on Figure 7, shown the 2 important component that plays significant role on the process of Web Programming. Firstly, user will use the browser built in the desktop machine to send requests to a specific web domain in the server. The server that responsible to store, manipulate and fetch data from the Excel spreadsheet or database, is where the website runs on. The response process continues where the server will return a response page back to the browser by presenting the page back to the end user following the information requested.

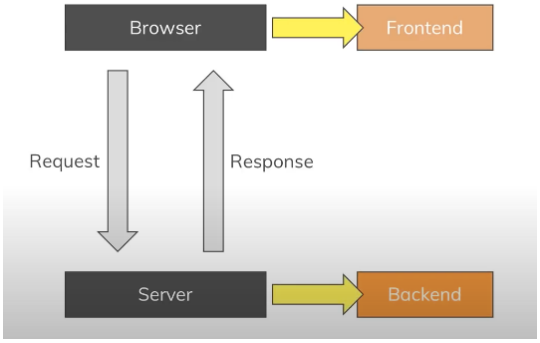


Figure 7 Process of Back- end and Front-End Development and how it become full stack approach.

Since the browser help user's views, their information needed, browser referred as the Front- End approach which user interface (UI) will be rendered to the end user while server works as the Back-End Approach responsible on data operations returning the page to the browser. According to (Alexa Goins, 2020), front- end development is a style of computer programming that focuses on the coding and creation of elements and features of a website that will be seen by the user. The front- end represents the interior design that are built from the back- end developers. Moreover, front- end works where user may interact with the website designs such as fonts, page, images, maps, buttons, and links. All these designs are to produce a beautiful user interface by the combination of HTML, CSS and JavaScript languages, codes, or algorithms.

However, back-end development stays out-front of the seen, where end- users indirectly interact with the back end by using front-end application that the developers have worked on. A back-end developer will help manage this database and the site contents stored on it (Alexa Goins, 2020). Usual operations of the back end include building codes to manipulate data, troubleshooting and debugging, managing the database, and utilizing framework.

Front- end and back- end is a concept that have been used to create this wonderful dashboard. Front end and back-end development requires codes. In building the model, the language used are JavaScript applied for Front- End Development while Python language are used for the Back-End Development. To combine these two languages is called as full stack development. This project succeeded by the setup of

the web- based framework called as Python Django that acts as a full stack development.

Introduction to Python Django as Web Development Framework

According to (Web Applications with Django , 2010), Django is a modern and open- source Python web framework that redefined web development in the python world. Django consist of pragmatic design of the creation of the website and full stack approach (using MVT Architecture) that make users to easily connects its operations between front- end and back- end codes shown in Figure 8 and 9. The benefits of using Phyton Django are as below:

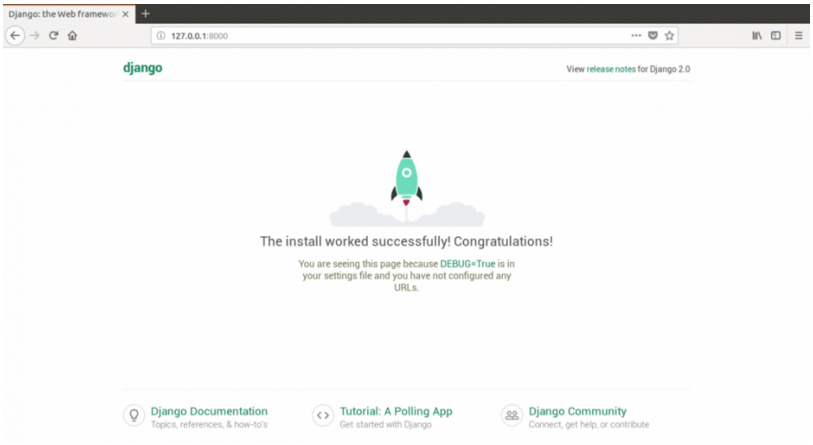


Figure 8 Pragmatic design of web application.

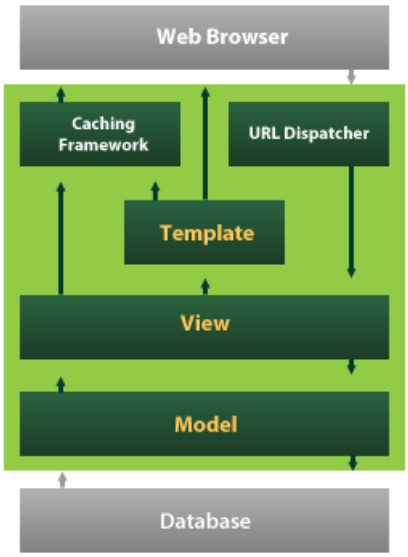


Figure 9 Python Django MVT Architecture.

Understanding Django Architecture

Surely people will ask, “What is MVT Architecture in Django?” or “How does it work?”. Model- View- Template or MVT describes the patterns and software design that Django applies towards developers who needs an easier solution to design and develop their web-application. Based on Figure 9 shown the whole set of Django MVT Architecture contains various components (coloured in green) of code template built in together with its modern full stack approach and functionality. One of the benefits of Django Architecture is that it does not require re-structing process whenever to place big or complex codes in its suitable locations due to the built in MVT architecture prepared in the software. Figure 9 shows a green bar which presents the importance of MVT architecture components to ensure the web application runs smoothly, effectively, and efficiently towards presenting the dashboard inside the browser at the lowest amount of time. Detailed briefs of MVT are as below:

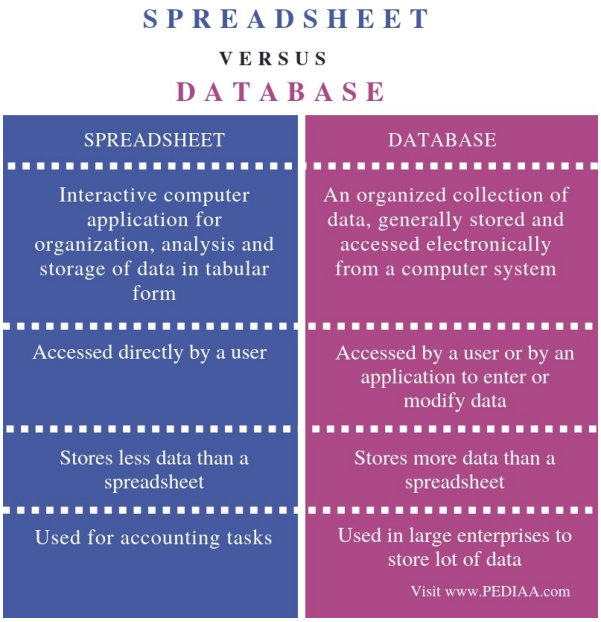
1. **Model as the back-end environment**

   Model acts as the component of data interface. Model or data model is a conceptual process of classifying dataset descriptions and semantics of the data. Moreover, it is responsible to manage the logical data structure of the back- end environment towards the whole application. For example, storing and handling big data usually involves databases (back-end) which organizes and manipulates the coding deliverables such as by using MySQL or Postgres. Unfortunately, this project does not have to involve database since the data volume of each pipeline are quite small.

   Therefore, this project provides an easier and user- friendly solution that they could just store the data in their local PC or hard-disk by a specific folder of spreadsheets before starting to explore the sites for visualizations and inspections. The difference between database and spreadsheet basically describes by different meaning and tools. Moreover, it differs based on data dimension such as data volume and kinds of data being collected. Database has no limitations while spreadsheet have limitation on handling volumes of information that are unmanageable like database.

However, using database provides unlimited data storage usually suitable for large amount of data. Since this project only requires a less than 40 rows of data, it will be best to use spreadsheet. Additionally, spreadsheet is suitable for small volume of data depending on the hard- disk of the users. In today's work environment, Excel Spreadsheet can be accessible to multiple users within the office of the same document. Nevertheless, will take only a user to amend data one user at a time. Table 4 points the difference between database and spreadsheet or Excel Workbook.

Table 4 Difference between Spreadsheet and Database



The back-end environment used are under Excel Spreadsheet which the pipeline locations are stored under GISFYP/Data folder in the local PC. Figure 10 shown pipelines data are placed under the GISFYP project folder, subset in Data folder.

Phase 1: Creation of sub folders

Organizing folders are important for every projects especially for project that does not use database as their back- end development. It requires developers to set up manually from their desktop machines. In this phase, a folder has been created that contains input files including a sub folder for output files once Django execute its operations of coordinate conversion. Input files contains

coordinates before conversion. While output files contain pipelines coordinates after conversion from UTM to Long Lat coordinates. The output files will pop out once the pipeline drop down menu are clicked and after the execution out conversion process that convert UTM to GIS coordinates. Figure 10 and 11 shown input folder (named Data) and output folder (named Pipeline)

Reason to separate the files are as below:

- To avoid confusion on searching the input and output files used by both end-users (to convert data) and developers (for development, maintenance, and testing).
- To enforce data consistency without any duplications of pipeline files.

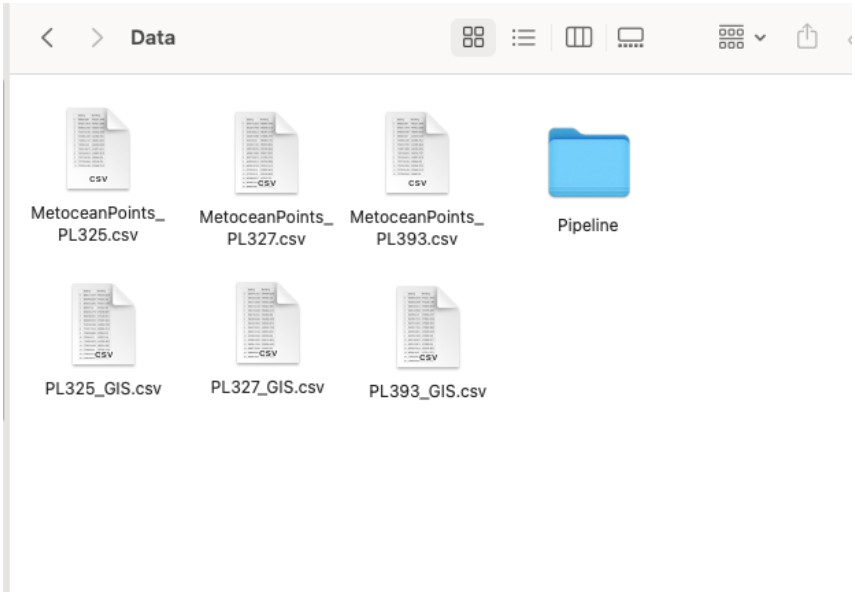


Figure 10 Input Files under Data Folder

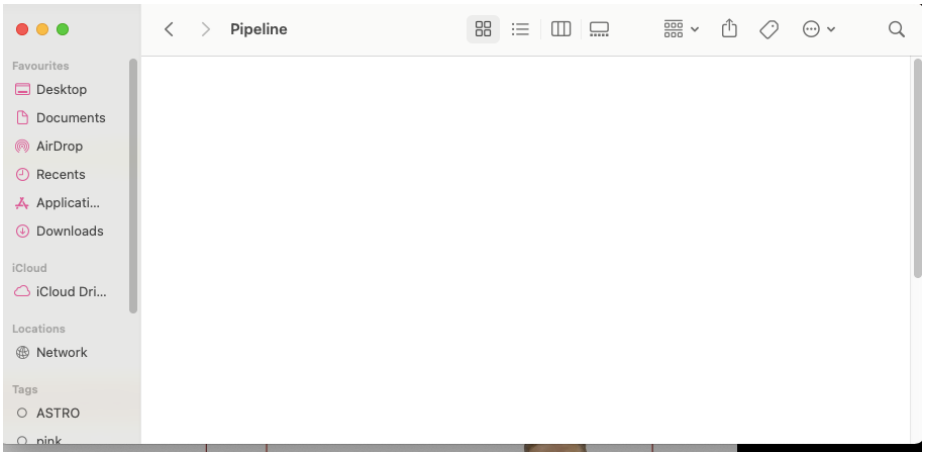


Figure 11 Output Files before conversion button executed

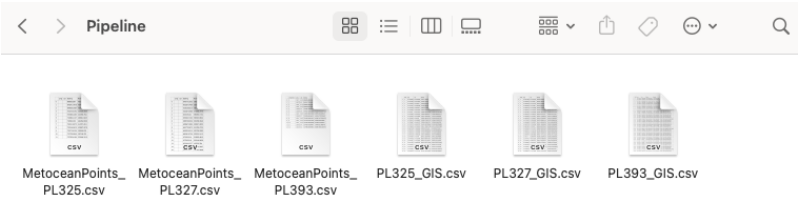


Figure 12 Files expected after execution of conversion button (change updated picture.

Phase 1: Django Setup - Installation

The Django setup can be done using the terminal. Since this development uses Mac as its operating system, Django should be installed under Python 3 as the latest version. The reason to use Python 3 is because it has a modern syntax and great demand on the typing system for the development of data science, artificial intelligence, and machine learning. Next, operations require the installation of condo and pip as the package management system in python3. Conda (Anaconda) and pip (Python) are responsible to manage the installation and uninstallation of libraries and package in python to help developers on using these packages on building Django operations. Whenever new packages and libraries are installed using pip, it will search it in a packages repository called as Python Package Index (PyPI).

Moreover, both package management system helps to create virtual environment. Virtual environment needs to be installed that separates python environment for different projects, determining the usage of different version of Django. Not to forget, virtual environment separates other Django projects by its own dependencies and packages needed. Therefore, I have named the virtual environment as GISFYP. Continuing, verification of the existence of virtual environment also been done.

Phase 2: Django Setup – Creation of Project and Application in Django

As the last part of this phase is where Django Framework installation involved using pip. After installation being made, a project needs to be created. Therefore, project named created as "fyp_project" in the FYPGIS environment. Django framework provides a simple, flexible, simpler framework design such as MVT and reliable framework that protect the web-application from security issues. App and project simplify the distribution and organizes the structure of the codes and modules. Therefore, it is crucial that Django projects should be included in this framework for the usage of MVT web application design framework as well as for future work. Shown in Figure 13, the highlighted yellow box is the project named "fyp_project" while the red box is the application created for the GIS visualization and inspection website. Finishing the setups, it is time to activate and start the server to display the built-in website that Django provides such Figure 14. Once the server activated, the operations will connect with the browser's local host by localhost:8000.

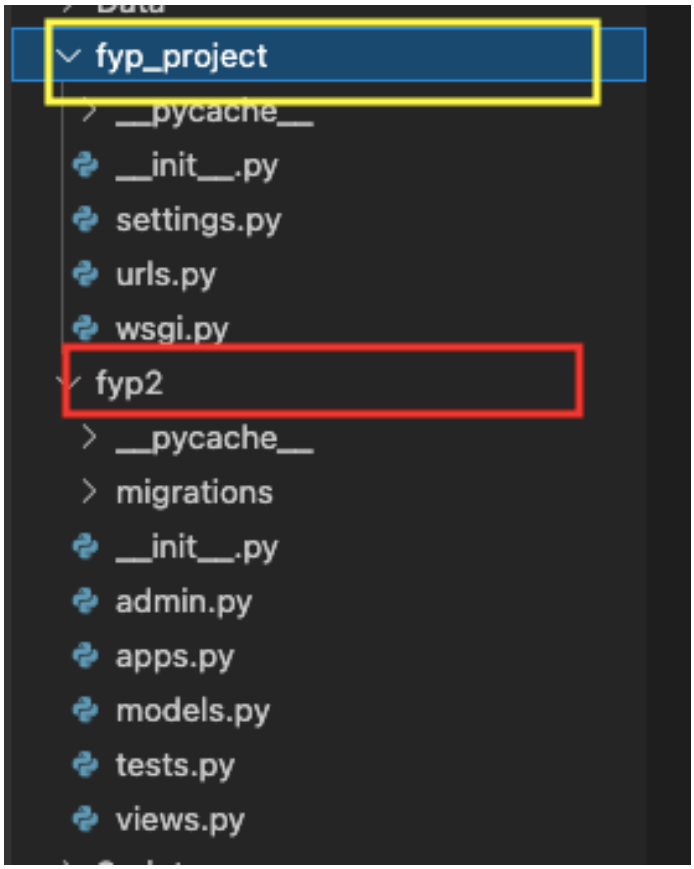


Figure 13 The highlights shown Project and App created.

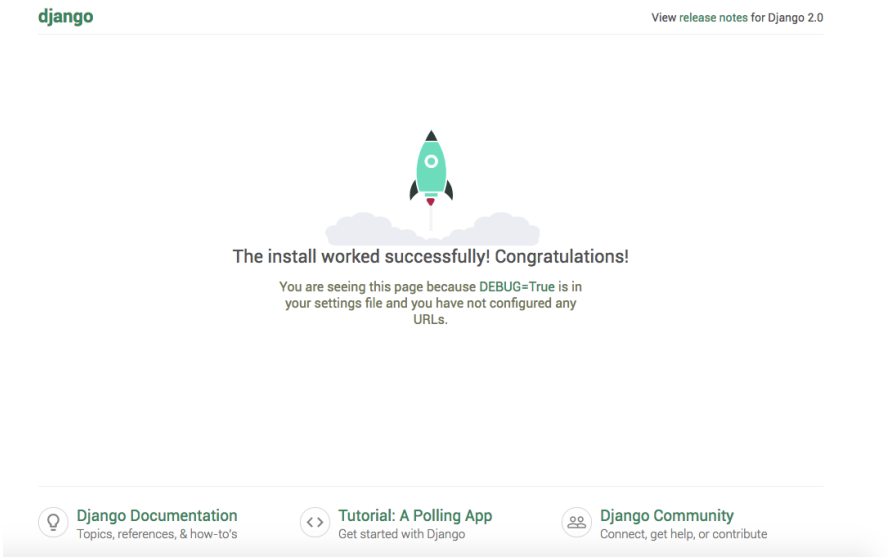


Figure 14 Django built in web application page

2. View as the Middle- End (Business Logic)

The View is the middle- end operation that involves the interactivity operation of web pages, user interface layers, and presentations such as functional drop-down list of pipelines and display the base map to the website. View is responsible on interactivity operations which involves work whenever end-users click example on a button. Surely the users expect the website to understand the request and do the work by returning with specific responses.

For example, User A would like to display UTM coordinates in the map from the data files provided. User A then clicked the pipeline from the drop-down menu with the expectation that Django view framework will execute the conversion operations and plotting from the systems as a response. Therefore, view is the hardest part on building this web application. Every work for each interactivity requires its own codes and functions before connecting it to templates and model. To make sure the execution runs smoothly, each functions requires repeatable testing before developers compile all deliverable together in the view python file.

Phase 4: Connecting View and URL Patterns

As described in the Phyton Django Architecture, URL plays a crucial role to request the page from the browser and the server to reach Django's operations in order server response back to the front-end of the website. Therefore, URL needs to connect with view's patterns so that the output pops out its interactivities. Figure 15, shown a specific Views.py file which requires only for view functions and other interactivity work that needs to be executed as a response back to the website. Before developing the views' function, Views.py file needs to be attached with URL file under the same application. The best way is by invoking the view functions in url.py of the project such as Figure 16. The way is to import views.py and rename as map_views to be used in the URL patterns invoking map functions by function names.

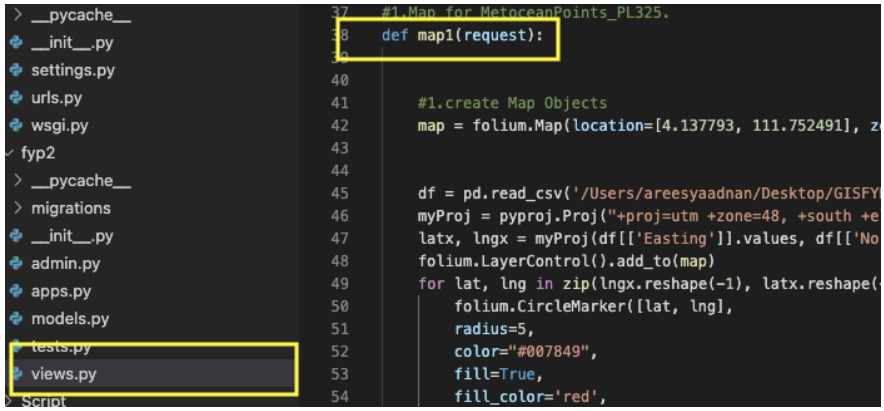


Figure 15 Views.py coding page that include views functions

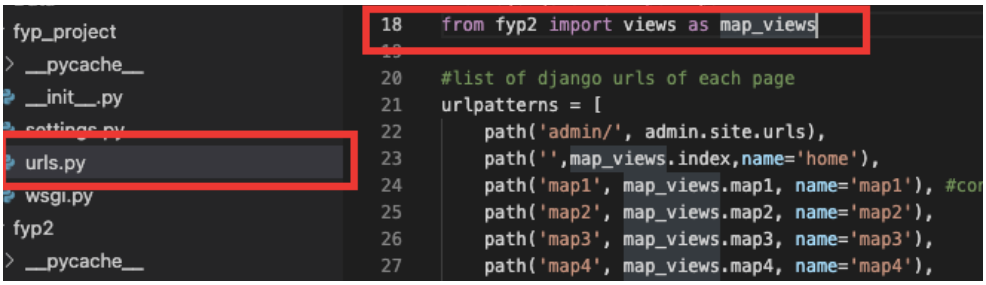


Figure 16 Importing views.py in urls.py in fyp_project

Phase 5: Set up View Functions/Algorithms to make the dashboard interactive

Function 1: Algorithm to display base map using Folium

Firstly, it to display the base map. Based on (Dempsey, 2011), the term base map is see often in GIS and refers to a collection of GIS data and/or orthorectified imagery that form the background setting for a map. Orthorectified imagery is an image based on satellite imagery angle which plays a significant role to display and evaluate the GIS coordinates for this project. Since GIS maps are based on spatial concept and perspective, orthorectification helps to display spatial information gathered inside the GIS mapping. Without orthorectification process, user would not be able to angle the positions accurately. To start with the algorithm, is to firstly import the software packages called as Folium. Folium is a python module operates to display geospatial data in the web page extending its features from an interactive Leaflet.js library. In this GIS Visualization and Inspection project, every page of the map is assigned with folium module with based on different types of pipelines. There are 7 web pages declared in every function by the name of index (as the first default operation in default page), map1, 2,3 until 6 following the structure of pipeline names. The structure of the map is as followed:

Table 5 Pipeline names and respective functions names for views.py

Pipeline Name	Functions Name
Home Page	index
Metoceans_PL325	map1
Metoceans_PL327	map2
Metoceans_PL393	map3
GIS_PL325	map4
GIS_PL327	map5
GIS_PL393	map6

In each of the map displayed, I would want to present the HTML map in default coordinates whenever execution happens. The location set are nearby Bintulu, Sarawak, Malaysia in its coastal areas shown in Figure 18. Also, I have set the height and width that contains only quarter part of the website.

```
#1.create Map Objects
map = folium.Map(location=[4.137793,111.752491], zoom_start=8, width=900,height=500)
```

Figure 17 Folium map and its configuration included in map variable

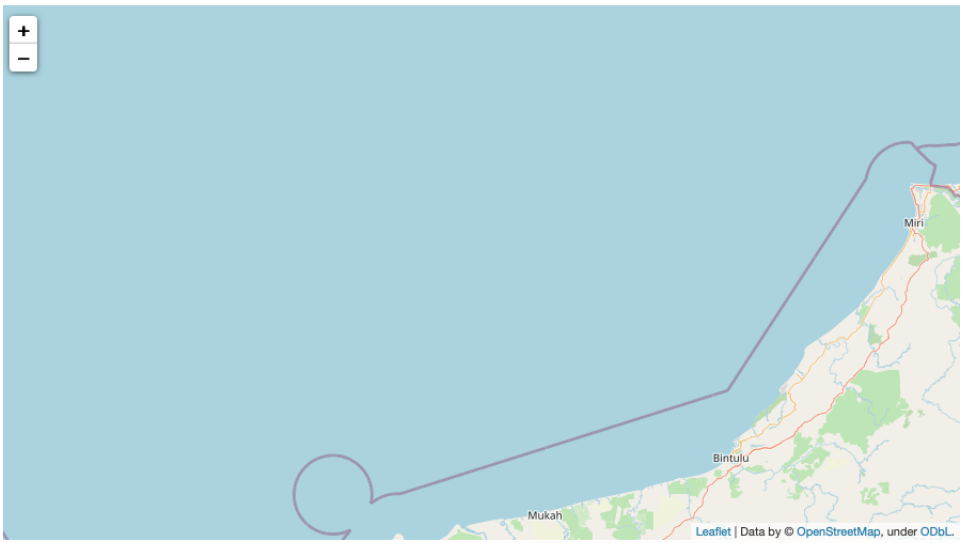


Figure 18 Base Map using Folium, Leaflet.js located nearby Bintulu, Sarawak

Function 2: Algorithm for conversion function

Conversion functions helps the dataset imported in data frame format transform it from UTM to GIS projection. The first step is to import the library packages called as Pyproj. The pyproj is a Python package that performs cartographic transformations and geodetic computations (Paul Crickard, April 2018). The conversion process read the input data set in csv format (Data Folder) and convert the UTM coordinates using pyproj library before printing it into a new csv file in the output folder called as Pipeline. Thirdly, to describe the specific current UTM projection location of the pipeline before conversion happens including latitude, longitude, zone systems, map locations, ellipsoid and datum towards the operations parameter.

Table 6 List of parameters involved on data conversions.

Parameters	Explanations
Geographic Coordinates Systems : Easting and Northing	Easting as the x- axis. Northing as the y-axis.
Projection Notation : Universal Transverse Mercator (UTM)	Describes the local GIS before conversion into geographic coordinate system.
UTM Zone: 49 (Malaysia Location)	The UTM coordinate system divides the earth into 60 zones. Zones describes as the reference of the latitude and longitude being fetched from the data set into the correct UTM location to place inside the map.
Map Locations : South of Malaysia	Default local direction the target area.
Ellipsoid:WGS84	Define the size and shape of the Earth.
Datum: WGS84	Network of benchmarks that knows longitude and latitude using mathematical model to identify accuracy, length and area of the locations. In Malaysia, datum used are WGS84.

```
df = pd.read_csv('/Users/areesyaadnan/Desktop/GISFYP/Data/MetoceanPoints_PL325.csv')
```

```
myProj = pyproj.Proj("+proj=utm +zone=49, +south +ellps=WGS84 +datum=WGS84 +units=m +no_defs")
latx, lngx = myProj(df[['Easting']].values, df[['Northing']].values, inverse=True)
```

Figure 19 Algorithm for UTM transformations to GIS.

Next, is to assign 2 variables, “latx” and “lngx” as longitude and latitude after conversion happened by turning inverse parameter into True. If it executes the correct coordinates, the algorithm will be applied towards other 6 pipeline maps based on map’s numbering shown in Table 7. By then, the conversion executed will run these codes and write the converted data in the Pipeline output folder.

Table 7 Example of expected output after the conversion from UTM to Longitude Latitude of Metocean Pipeline 325.

MetoceanPoints_PL325

	Easting	Northing	Longitude	Latitude
0	688653.687	481027.2948	4.349999999773370	112.70000000006600
1	694241.7978	464451.8388	4.200000000365530	112.74999999981100
2	699806.0084	458934.9245	4.149999999637930	112.80000000008000
3	705370.9192	453418.0658	4.09999999990451	112.85000000018400
4	705396.1483	442358.7621	3.999999999886170	112.85000000032000
5	710962.1147	436841.6433	3.9500000000794900	112.89999999997900
6	716554.032	420264.6509	3.7999999999782300	112.9499999997410
7	722121.6373	414747.2411	3.7500000000109500	112.9999999997840
8	722146.5619	403687.0478	3.6500000003693100	113.00000000028700
9	722170.8136	392626.876	3.549999999649530	112.99999999974900
10	727739.5949	387108.793	3.5000000002799800	113.05000000003700
11	727763.4196	376048.3115	3.4000000000744800	113.0499999996030

Function 3: Algorithm to Plot and Visualize the coordinates to base map

In pointing out Point x (Longitude) & Point x (Latitude) for each of the pipeline, the library packages used are also based on folium module. Nevertheless, to point out the coordinates requires marker features that iteratively reads the data frame of longitude and latitude from the output folders whenever end- user presses pipeline from the drop- down menu. Below are the operations designed:

```
folium.LayerControl().add_to(map)
for lat, lng in zip(lngx.reshape(-1), latx.reshape(-1)):
    folium.CircleMarker([lat, lng],
    radius=5,
    color="#007849",
    fill=True,
    fill_color='red',
    fill_opacity=0.7).add_to(map)
```

Figure 20 Pointing out the coordinates at the base map.

Using for loop, I have assigned converted of GIS coordinates into latitude and longitude variables plotted in a Circle Marker. Circle Marker is a representation that displays the GIS coordinates points inside the base map. Based on yellow box highlighted in Figure (20), zip represent a function that helps to merge two lists together. In return of the output will structure each list (longitude and latitude) in a structured manner. In addition, the folium marker can be design following the radius or

size, colour, availability of the filling and its colour and the opacity of the marker. Polylines are also included in the map. This is to visualization to flow line of the pipelines. Also, another line marker is added to present the flow of the pipeline that interconnect each point in between the GIS coordinates. Lastly, is to declare add_to() function that helps to view the marker and points it into the base map.

All 3 functions concepts are applied to all the 6 pipelines by declaring the operations on each of the 6 maps functions so that every pipeline being clicked received the same concept as a response.

3. Templates as the Layout

According to (Templates, n.d.), a template contains the static parts of the desired HTML output as well as some special syntax describing how dynamic content will be inserted. Templates acts as the layout in the design pattern architecture that requires HTML, JavaScript, and CSS as the language to build the layout output in the Django framework.

This project uses an open-source CSS framework to design the web page layout called as Bootstrap. Bootstrap allows developers to put their custom built in styling to the website as the design of the layout page. Bootstraps are meant for an interactive and responsive web application that provides custom and built-in navigation features such as drop- down menu, search bar, buttons that requires the load of stylesheets links and scripts links inside the templates. The easy part is, there is no need to download the stylesheet but only uses a specific link to load it inside the template which makes the custom design are built much faster and can be customized quickly. Therefore, this framework is suitable for short- term projects. This project development used the templates for further styling, margin, and position of the web pages. At the head of the html document, are which meta tag must be declared. The meta tags functions to inform the operations to transfer data from machine codes into readable text.

Phase 5: Styling the web pages

Layout 1: Navigation Bar

Figure 21 presents the navigation bars contains a drop- down menu (highlighted in red box) that lists all the Metocean and GIS Pipelines. Also,

there is one menu button (highlighted in purple box) that will guide user back to the home page. In terms of the styling and colours are maintained in grey.

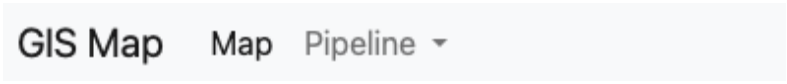


Figure 21 Navigation Bar consist of drop-down menus that are clickable.

- Layout 2: Map

  In this layout, where the folium map will be rendered into the layout. The position of the map will be located on the right side of the page for home button while map will be on the left side.

- Layout 3: GIS Insights

  GIS insight contains results and discussions of the presented pipeline.

- Layout 4: Haversine Distance Result

  Haversine is a mathematic and geographical concept of calculating the shortest distance between two coordinate points on the spherical Earth surface. In oil and gas, this mathematical algorithm motives to search for the real distance from the initial point to last point of the pipeline thar relates with the navigations of the pipeline.

Phase 6: Checking MVT connections

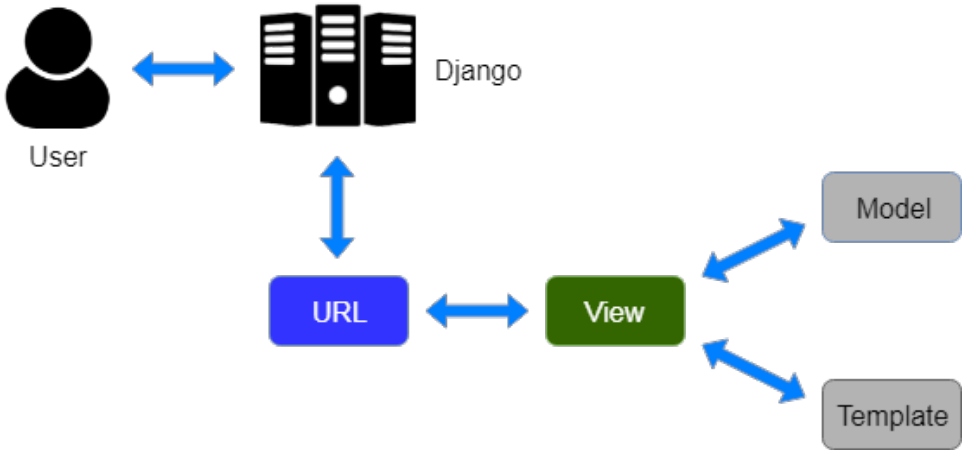


Figure 22: MVT Patterns

How does the end users interact with this product? End users that want to use this pipeline visualizations and inspection will need to use Django web application as their resource. Firstly, user's needs to prioritize their pipeline datasets. Pipeline datasets

consist of pipeline locations displayed in various map projections (UTM or GIS locations). In our case, the original dataset contains locations formatted in UTM coordinates that needs to be displayed in GIS map projection using this web-application tool. The data set will need be in local GIS folder on the user's PC. Placing the dataset in the right folder that is aligned with the algorithm or codes of this project will make the visualization and inspection process efficient.

Secondly, end- users will need to access Django application or website via a request from the URL link provided. Django that functions as the controllers will investigate the availability of the resource (website templates) from the back- end codes. If the back-end codes contain MVT patterns that interconnected with each other and aligned with the URL link, the Django(controller) will response by rendering the templates (html pages) as requested for the end- users to view.

3.1.5 Evaluation

In this evaluation phase is where we evaluate the visualized data. Developers usually questions themselves, "does the models satisfy the original use case?" which look through back on the business understanding. In data visualization, not all the data are accurate as expected. Data outliers can sometimes occur if the coordinates or points presented area away than the majorities. This can cause deviations and faulty for the future analytics that involves the wrong data.

However, this project does not aim to develop an outlier detection but only to evaluate the anomalies seen after the modelling process being done. This is the reason to use scatter plots in the line graph to detect anomalies of the line drawn. Figure 23, shown scatter plots and its pattern with anomalies given as an example. Referring to it, you may see the red dots or points that are far away from the majority points. This can be called as outlier or anomaly.

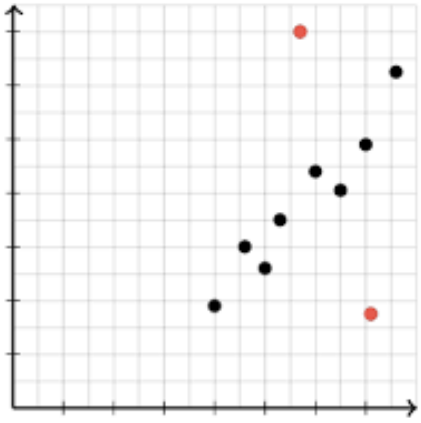


Figure 23 Example of scatter plots and its error.

The importance of evaluating the graph is crucial for this project to succeed. Accurate and an efficient data help both users and organization to use the data for other and future useful analyzation projects such as involving innovations, machine learning or artificial intelligence, hence identifying its suitable and definite location. Wrong coordinates and spatial references might ruin the whole GIS pipelining process. Therefore, differentiate the scatter plots by its position on the graph and identify rare element on the charts helps users to differentiate the distance between the data by using haversine formula.

3.1.6 Deployment

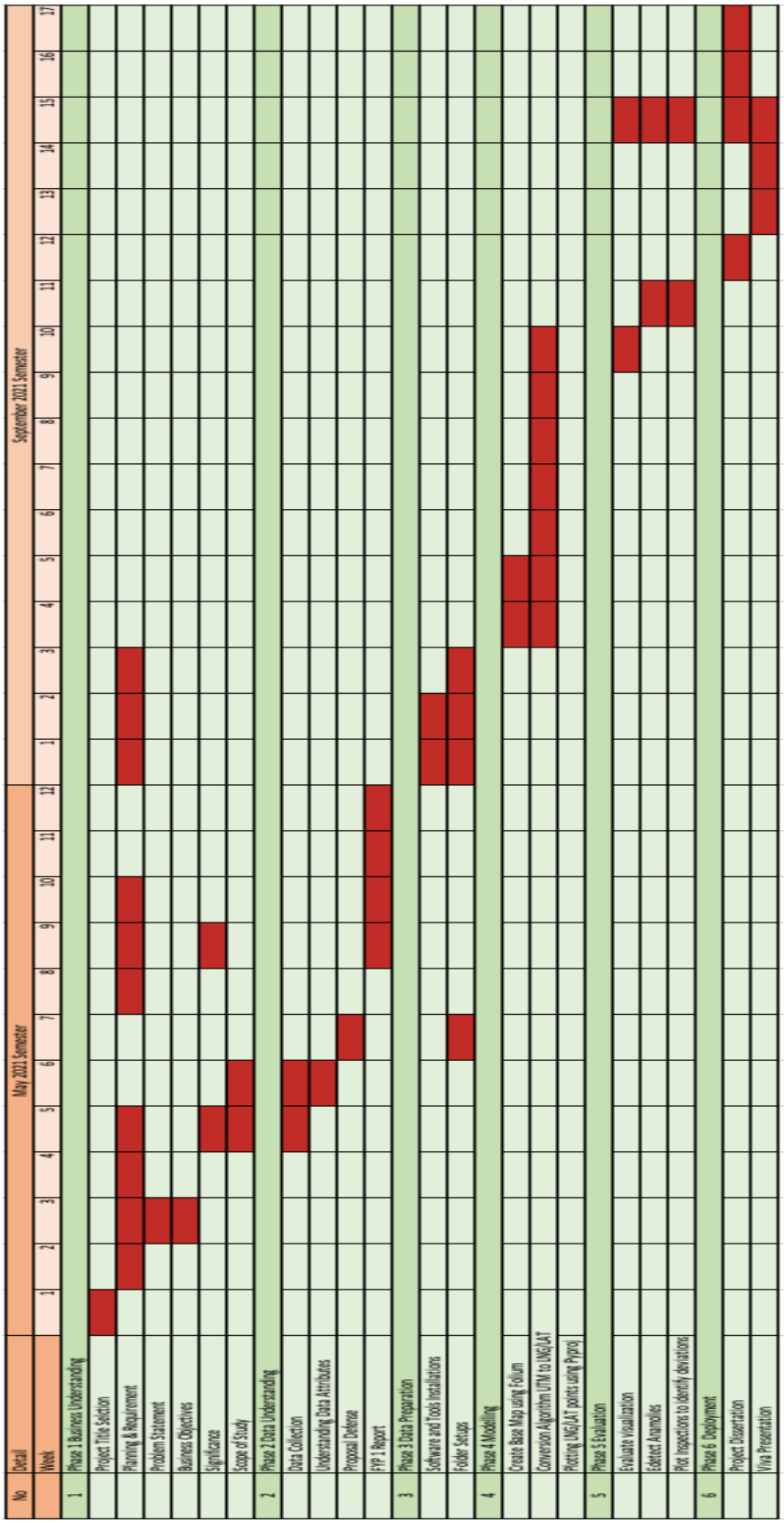
Planning and Presenting the Deployment

Deployment is the stage where the dashboard should be put into real use so companies could realize the functionalities of the model in real life plus to make a judgement about it. Firstly, is to plan the ways to present the information to the users. Capturing on this project's objective which is to provide a dashboard that is interactive and based on mapping charts will help the external users and business entities understands better on the business and decisions- making being applied. In addition, the benefits of having interactive charts assist a quick engage to the end- users as well as giving the experience of better insights on the dashboard. In terms of design and colouring, the dashboard surely must have the best visualizations that fits with enriching the business objectives such as Line Charts, Mapping Charts, Scatter Plots and Table Charts (if needed). It all depends on the relationship between element or

variable of the data source received either numeric to numeric values or numeric to categorical. Also, having a clear navigation and with simple functionality colours ease the end user's journey to interact with the dashboard. Lastly, the hovering and selecting information inside the interactive charts must be intuitive for users such as zooming function that assist users to make in depth observations on the charts presented. In terms of communication and presentation wise, presenter should firstly understand the information that must be applied before the presentation started. Furthermore, presenter should explain more on the various perspectives of analysis rather than just corelate it in one's points of view. Making sure the information delivered are compatible is the responsibility of the presenter on making sure the progress of presenting is effective and efficient.

3.2 Gantt Chart

Below is the Gantt Chart that outlines the proposed timeline for GIS Inspection and Visualization Project.



CHAPTER 4: RESULTS AND DISCUSSIONS

4.1 Result

4.1.1Conversion Output

As the result of the conversion function, copies of data set file in CSV format are written back in the pipeline folder (output) from the Django view operations. Figure 24 shows the output of data after the transformation from projection coordinate system (UTM) to Geographic Coordinate System (Long/Lat).

MetoceanPoints_PL325

New Variables Are Added in the output folder after the conversion

	Easting	Northing	Longitude	Latitude
0	688653.687	481027.2948	4.349999999773370	112.70000000006600
1	694241.7978	464451.8388	4.200000000365530	112.74999999981100
2	699806.0084	458934.9245	4.149999999637930	112.80000000008000
3	705370.9192	453418.0658	4.09999999990451	112.85000000018400
4	705396.1483	442358.7621	3.999999999886170	112.85000000032000
5	710962.1147	436841.6433	3.9500000000794900	112.89999999997900
6	716554.032	420264.6509	3.7999999999782300	112.9499999997410
7	722121.6373	414747.2411	3.7500000000109500	112.9999999997840
8	722146.5619	403687.0478	3.6500000003693100	113.00000000028700
9	722170.8136	392626.876	3.549999999649530	112.99999999974900
10	727739.5949	387108.793	3.5000000002799800	113.05000000003700
11	727763.4196	376048.3115	3.4000000000744800	113.0499999996030

Figure 24 Metocean_PL325 Pipeline on the output folder after coordinate conversion executed

4.1.2 Dashboard as Project Prototype

As a result of the dashboard development, combining information, interactive functions, and geospatial data proves the visualization project success. Figure 26 shows the prototype dashboard showing the visualization of the pipeline including several discussions regards to local Geographical Information System (UTM) and modern GIS (Geographic Coordinate System or longitude latitude). The results shown are based on the website configured using Django Framework. In terms of the content provided in the web page, it is obvious that web page contains designs and multiple features that are coloured and structured into organized manner of the webpage.

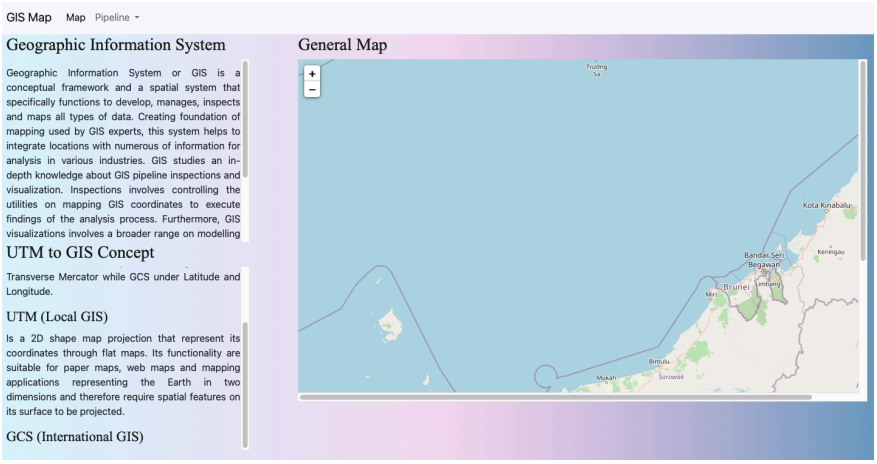


Figure 25 GIS Home Page

In this GIS Dashboard, there are total of 7 web pages including the general page and 6 GIS pipeline visualization pages that contains introduction of GIS, brief issue on the local GIS and differences between coordinate system such as UTM and GIS. The pipeline web pages will be chosen in the drop- down menu without any sequential steps.

The first page introduces the users on Geographic Information Systems together with understanding of UTM and process. The second pages views users the plotting of the Metocean_PL325 including the brief on the Types of pipelines such as Figure 26. Third map introduced the users into GIS data layers the 4th layout explains scatter plot as the chosen visualization such as Figure 27. Lastly, discussion on inspections.

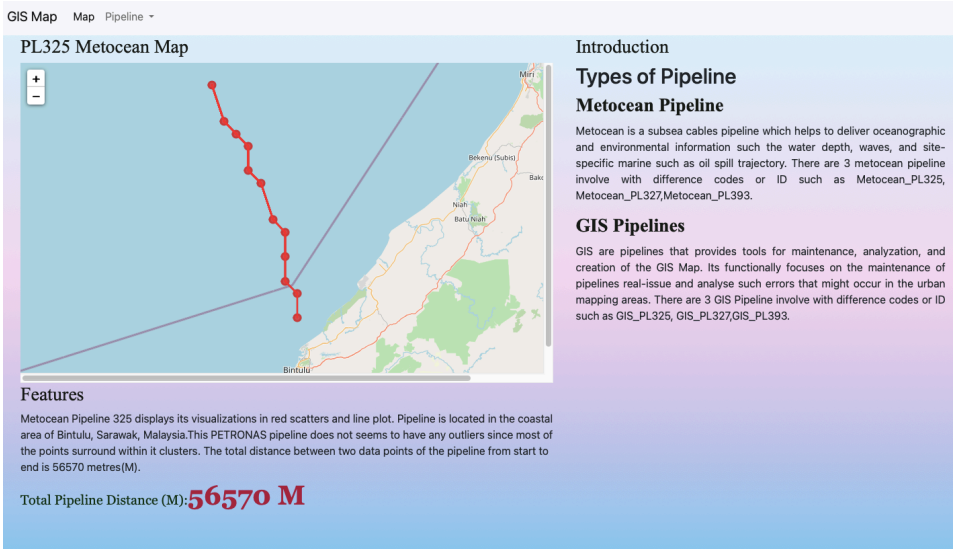


Figure 26 PL325 Metocean Pipeline Dashboard User Interface

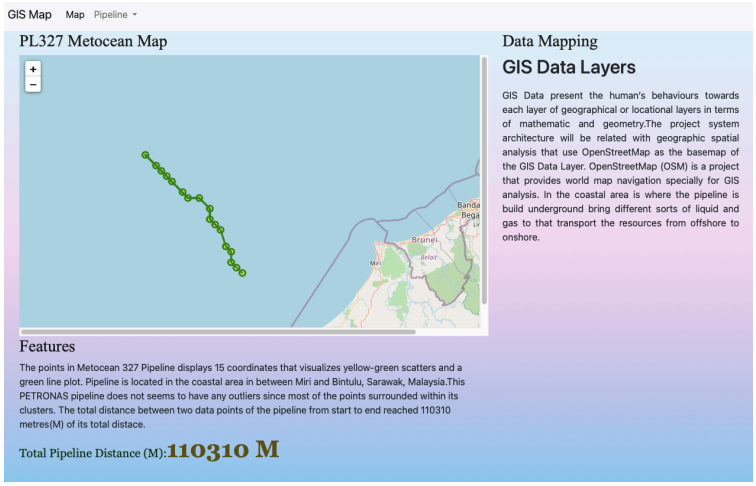


Figure 27 PL327 Metocean Pipeline Dashboard User Interface

Scatter Plot provides relationship between multiple numerical attributes. Another reason to use scatter plot is because it identifies anomalies of the streaming data during inspection process. For example, there are deviation of data far away from its group making the pipeline look unusual from the previous inspection records. The scatter plots provided represents the longitude and latitude following from the data set. Flow of the scatter plot differentiates the pipeline by its colour The scatter displayed uses Circle Marker declared in the view design pattern. Having longitude as the y variable and latitude as x variable that uses Polyline to connect with each other builds visualisations of the flow of the pipeline.

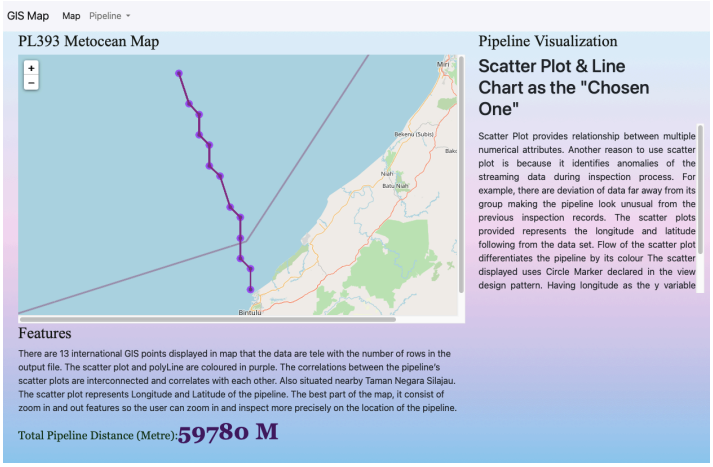


Figure 28 PL393 Metocean Pipeline Dashboard User Interface

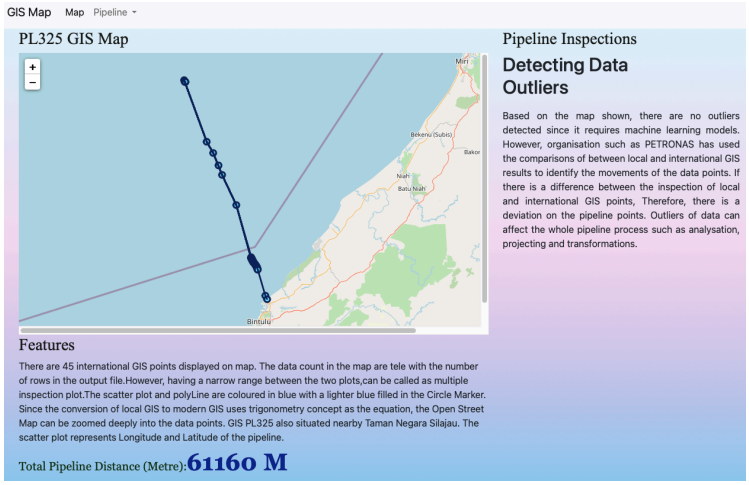


Figure 29 PL325 GIS Pipeline Dashboard User Interface

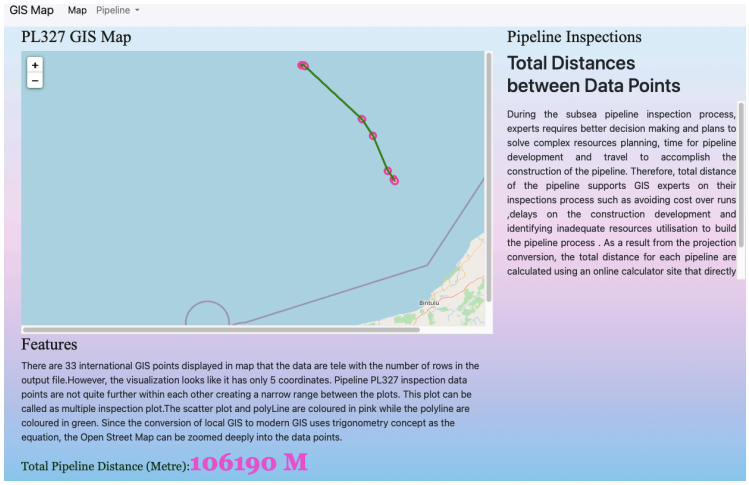


Figure 30 PL327 GIS Pipeline Dashboard User Interface

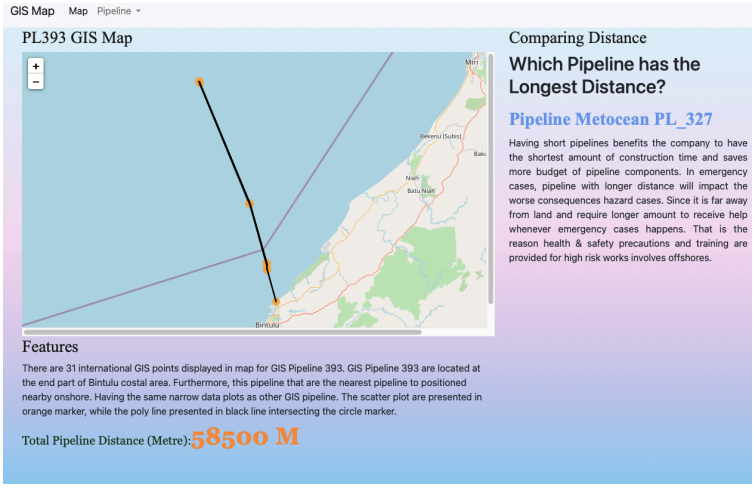


Figure 31 PL393 GIS Pipeline Dashboard User Interface

4.2 Discussions and Inspections

4.2.1 Transformation of Information into Graphical Context

It is proven that it is possible to transform geographical data into graphics such as the above figures in the results. Before data analytics, people want to understand the corelations between data by using mock picture or paint or anything involves graphical context. However, depending only on data tables could bring several cons towards making the audience understand which will then pending the time for decision making process. According to (Opila, 2018), acquiring and transferring knowledge is a complex and continuous process involving many techniques of expression, including oral communication, writing as well as numerous forms of visualization. This proves that visualization drives the transportation of acquiring and transferring knowledge towards the end-users. For example, Power BI that has been used by many organizations to pitch their new ideas, business process and problems in sense to make the audience understands it via dynamic map or real- time statistics. Understanding the input and output data, project use case and project functions can smoothen the visualizations and knowledge transfer in external and internal entities. Visualizations via Web – Applications avoids manual work especially for big data analytics.

4.2.2 Justification on the software and tool used that fits with MVT Pattern Concept on Web- Application Development

The execution results using MVT components justifies its collaboration could build the data visualization on the website. Phyton is the best language for Dashboard and Inspection development. As stated by (Carbonnelle, 2021), Python is the most popular language, Python grew the most in the last 5 years (14.8%) and Java lost the most (-6.1%). Figure 32, proves phyton maintained conquering the highest ranking in Data Visualization development for worldwide based on its latest updates of November 2021.

Worldwide, Nov 2021 compared to a year ago:

Rank	Change	Language	Share	Trend
1		Python	29.47 %	-1.5 %
2		Java	17.27 %	+0.8 %
3		JavaScript	8.97 %	+0.5 %
4		C#	7.17 %	+0.8 %
5	↑	C/C++	6.6 %	+0.6 %
6	↓	PHP	5.52 %	-0.5 %
7		R	3.88 %	-0.0 %
8		Objective-C	2.21 %	-1.4 %
9		Swift	2.13 %	-0.2 %
10		TypeScript	2.09 %	+0.2 %
11	↑	Kotlin	1.74 %	+0.1 %
12	↓	Matlab	1.68 %	-0.1 %
13		Go	1.39 %	+0.0 %
14	↑↑	VBA	1.38 %	+0.3 %
15	↓	Rust	1.2 %	+0.1 %

Figure 32 Worldwide Ranking records of the best data visualization tools

In terms of coordinate conversions, Django has a module package contains built in function on the calculations of converting Projection Coordinate System to Geographic Coordinate System (Longitude and Latitude) called as Pyproj. This can be the easiest way to convert the coordinate system plus it reserves more time. Manual calculations involve trigonometry formula and equations that are complex, Therefore, Python provide pyproj that can solve the transformation of coordinates just in a few seconds.

4.2.3 Inspection on the errors of the data

There is no outlier being found as it requires machine learning model that can builds and apply the outlier detection. However, the method of identifying outlier in this dashboard can be detected by inspecting the obvious error. The obvious error is by inspecting deviations. Organization like PETRONAS usually keep track on the daily record of the pipelines. If the GCS converted does not tele with local GIS (UTM), this can be called as deviation or shift pipeline coordinates causes the data to be wrong. Wrong data and inspection process will affect the whole process of data analytics causing the sharing of false information that might give a major downturn for the company.

4.2.4 Total Distance of the Pipelines

During the subsea pipeline inspection process, experts require better decision making and plans to solve complex resources planning, time for pipeline development and travel to accomplish the construction of the pipeline. Therefore, total distance of the pipeline supports GIS experts on their inspections process such as avoiding cost over runs, delays on the construction development and identifying inadequate resources utilizations to build the pipeline process. As a result, from the projection conversion, the total distance for each pipeline is calculated using an online calculator site called as "GeoDataSource" that directly calculates location from the first and last point of latitude and longitude. Figure 33 represent the online calculations to calculate the total distance of the pipeline that uses trigonometry calculations based on Spherical Law of Cosines that resulted the total distance of pipeline in metres(M). Spherical Law of Cosines is a formula that calculates the bearings between 2 location (coordinate) points. By my estimate, with this precision, the simple spherical law of cosines formula ($\cos c = \cos a \cos b + \sin a \sin b \cos C$) gives well-conditioned results down to distances as small as a few metres on the earth's surface (Calculate distance, bearing and more between Latitude/Longitude points, n.d.) .

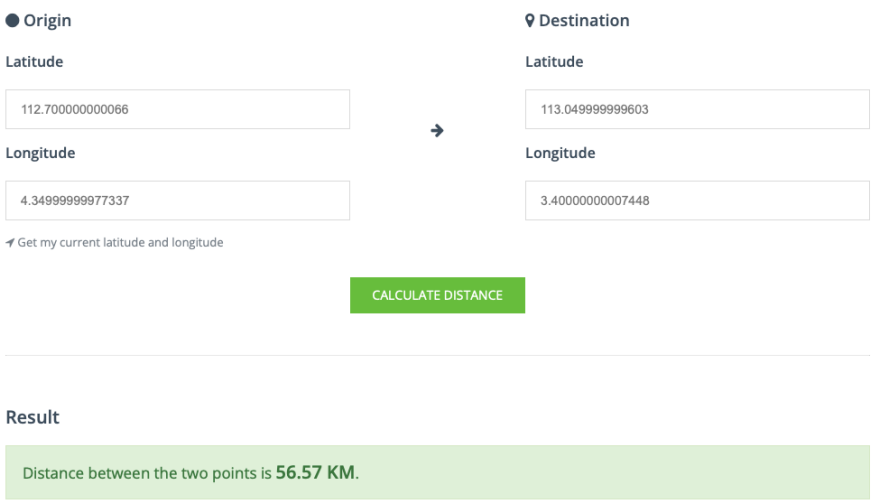


Figure 33 A website used to calculate the total distance of each pipeline by the start and end latitude and longitude coordinates

CHAPTER 5: CONCLUSION & RECOMMENDATIONS

5.1 Conclusion

In conclusion, this dissertation provides the understanding and techniques on way to produce an operative and efficient output for the external users to use this product for their decision making or any problem- solving of evaluating and inspecting GIS Pipelines. The successful history of GIS, given my-self the confidence to continue implementing this project. To get the perfect decision- making output it must come from a creative and simple understanding of visualizations. Also, this study introduced me into various of tools that are usable ad efficient to create different sorts of charts while the process of research and development the product happened.

Not to forget, the dissertation provides a full description of visualizing the information on inspection process enriching the quality of the project's objective. As discussed on the above methodologies depending on the system architecture used which is CRISP- DM modelling, without identifying the outliers, the information provided to continue the pipelining process will be inaccurate. Hence, will make this project progress as a waste for GIS experts on making their decisions to enhance the pipeline progress.

This project introduced into various coordinates and various geographical map learnings. Even, the opportunity to be in depth learn about Datum and different sorts of spatial mapping that existed such as UTM and Latitude/ Longitude plus to contrast its functionality representing societies and Earth layers. While in modelling phase, various concept and libraries are developed that are convenient and able to help continuing the development process to future developers to make evaluation on the proposed product itself by testing. Thus, using the overall research and development of the system architecture, the output can be presented to the end user's or business or organizational unit to present their report, creation, and solutions to carry on with business strategies and problem solving.

5.2 Recommendations

Future works is the segment which guides the results of this report for improvements. In my opinion, the project must be carefully and provenly executed for future upgrades. In addition, in every development there must be some errors that might occur. It is best to think of the consequences and solutions to adapt the issues. Thirdly, cleaning and transforming data sets plays a vital role to initiate the process, hence developers should reluctantly focus on the development. Lastly, with the output received by a smooth working process, end- users should be able to understand the inspection decision receive by the product.

Outlier Detection or anomaly detection is the identification of rare data points, events or observations which may raise suspicions by differing on data accuracy than any other majority data. Outlier is a data point that is not similar with other data points. Usually in companies or organisations, outlier detection used to detect any plotting, structures and deviation errors that occur while visualizing GIS coordinates.

REFERENCES

Hussain, R. Y. (2017). COMPATIBILITY BETWEEN MOBILE GEOPOSITIONS AND LOCAL MAPS IN IRAQ. *Journal of Engineering and Sustainable Development* , 27.

ArcGISPro. (n.d.). Retrieved from esri : https://pro.arcgis.com/en/pro-app/latest/help/mapping/properties/coordinate-systems-and-projections.htm

Tate, L. (2018, March 21). *An Overview of GIS History* . Retrieved from Geospatial World : https://www.geospatialworld.net/blogs/overview-of-gis-history/

Boost LABS. (2017, August 9). Retrieved from How Machine Learning Benefits Data Visualizations : https://boostlabs.com/blog/machine-learning-benefits-data-visualization/

Durham, K. (2020, August 20). *Towards Data Science* . Retrieved from Machine Learning Model Implementation: Assessing Variable Importance Across Models: https://towardsdatascience.com/machine-learning-model-implementation-assessing-variable-importance-across-models-30c282d143ec

Dempsey, C. (2011, December 16). *GIS LOUNGE*. Retrieved from Basemaps Defined : https://www.gislounge.com/basemaps-defined/

DJI Enterprise. (2021, March 2). *DJI Enterprise* . Retrieved from Geoid vs Ellipsoid: What's the Difference and Why Does it Matter?: https://enterprise-insights.dji.com/blog/geoid-vs-ellipsoid

GIS & GPS Tips and Techniques . (2014, November 11). *Youtube*. Retrieved from A Simple Exaplanation of Datum : https://www.youtube.com/watch?v=xKGlMp__jog

Thomas, A. (2017, March). *RECORDER*. Retrieved from Datums, Projections and Coordinate Systems: https://csegrecorder.com/articles/view/datums-projections-and-coordinate-systems

Alexa Goins. (2020, August 19). *Front End vs. Back End: What's the Difference?* Retrieved from KENZIE ACADEMY from Southern New Hampshire University: https://www.kenzie.academy/blog/front-end-vs-back-end-whats-the-difference/

Web Applications with Django . (2010). 281.

Paul Crickard, E. v. (April 2018). *Mastering Geospatial Analysis with Python* . Packt.

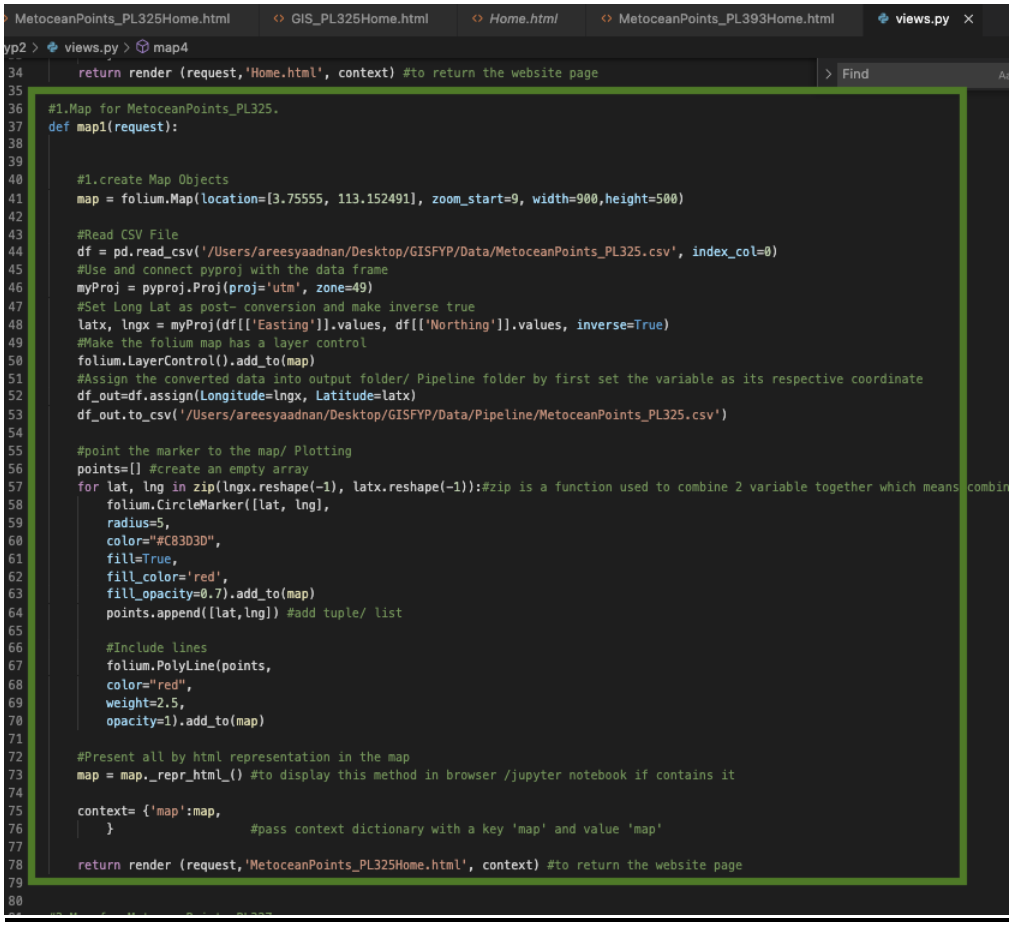
Templates. (n.d.). Retrieved from Django: https://docs.djangoproject.com/en/3.2/topics/templates/

Opila, J. (2018). Visualization in a Knowledge Transfer Process. *IRENET - Society for Advancing Innovation and Research in Economy, Zagreb* (pp. 485-493). ENTRENOVA - ENTerprise REsearch InNOVAtion Conference.

Carbonnelle, P. (2021). *PYPL PopularitY of Programming Language*. Retrieved from PYPL Index: https://pypl.github.io/PYPL.html

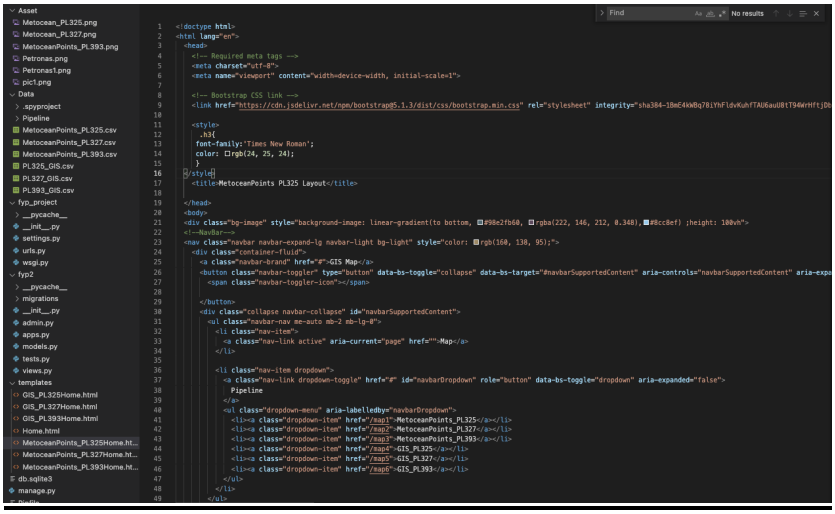
APPENDIX

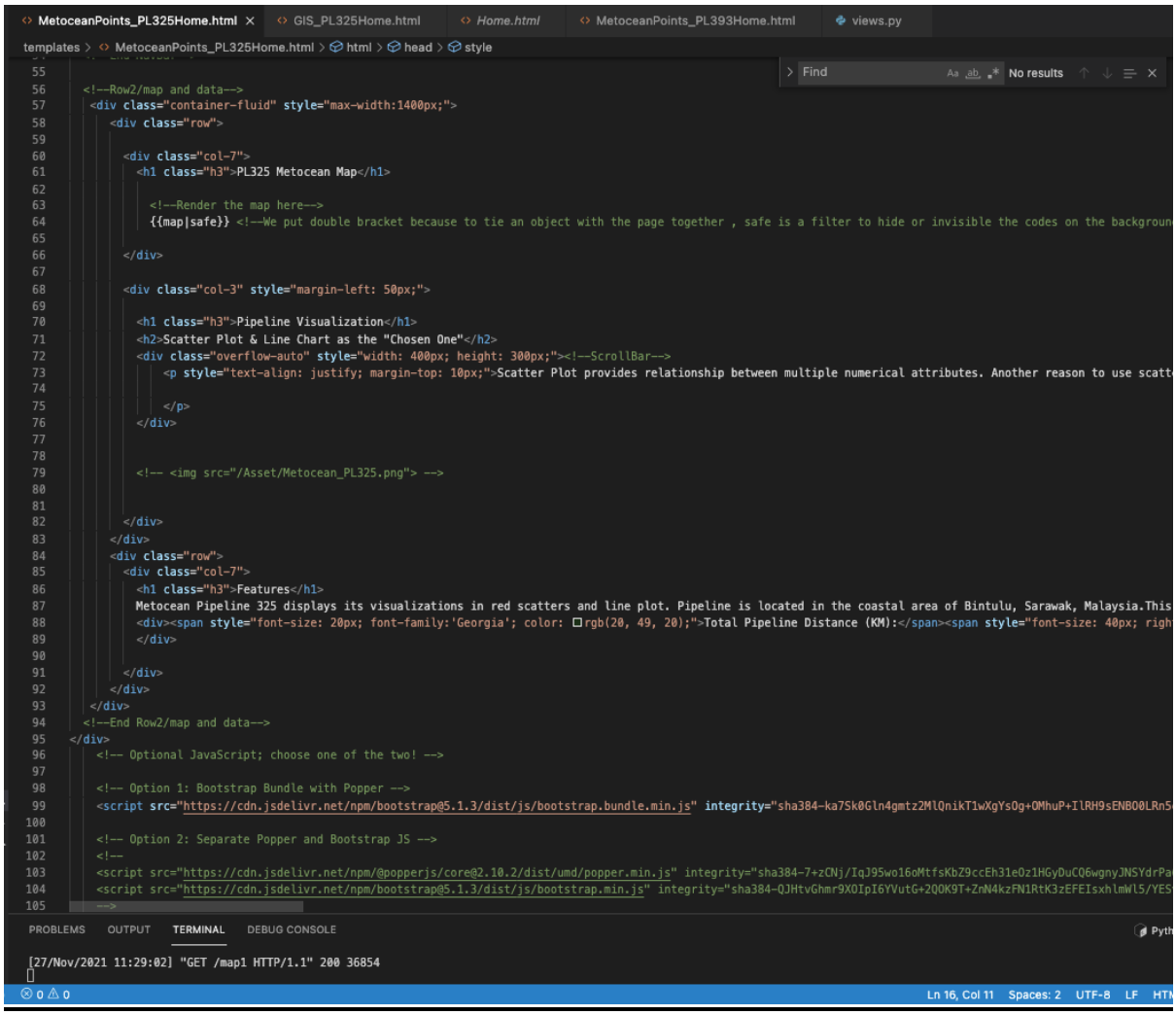
Views for Map 1: Metocean_PL325 Pipeline




```
    return render (request,'Home.html', context) #to return the website page

#1.Map for MetoceanPoints_PL325.
def map1(request):


    #1.create Map Objects
    map = folium.Map(location=[3.75555, 113.152491], zoom_start=9, width=900,height=500)

    #Read CSV File
    df = pd.read_csv('/Users/areesyaadnan/Desktop/GISFYP/Data/MetoceanPoints_PL325.csv', index_col=0)
    #Use and connect pyproj with the data frame
    myProj = pyproj.Proj(proj='utm', zone=49)
    #Set Long Lat as post- conversion and make inverse true
    latx, lngx = myProj(df[['Easting']].values, df[['Northing']].values, inverse=True)
    #Make the folium map has a layer control
    folium.LayerControl().add_to(map)
    #Assign the converted data into output folder/ Pipeline folder by first set the variable as its respective coordinate
    df_out=df.assign(Longitude=lngx, Latitude=latx)
    df_out.to_csv('/Users/areesyaadnan/Desktop/GISFYP/Data/Pipeline/MetoceanPoints_PL325.csv')

    #point the marker to the map/ Plotting
    points=[] #create an empty array
    for lat, lng in zip(lngx.reshape(-1), latx.reshape(-1)):#zip is a function used to combine 2 variable together which means combi
        folium.CircleMarker([lat, lng],
        radius=5,
        color="#C83D3D",
        fill=True,
        fill_color='red',
        fill_opacity=0.7).add_to(map)
        points.append([lat,lng]) #add tuple/ list

        #Include lines
        folium.PolyLine(points,
        color="red",
        weight=2.5,
        opacity=1).add_to(map)

    #Present all by html representation in the map
    map = map._repr_html_() #to display this method in browser /jupyter notebook if contains it

    context= {'map':map,
        }                   #pass context dictionary with a key 'map' and value 'map'

    return render (request,'MetoceanPoints_PL325Home.html', context) #to return the website page
```

Template codes for Metocean_PL325 Pipeline




MetoceanPoints_PL325Home.html
GIS_PL325Home.html
Home.html
MetoceanPoints_PL393Home.html
views.py
templates > MetoceanPoints_PL325Home.html > html > head > style
Find
No results
<!--Row2/map and data-->

<h1 class="h3">PL325 Metocean Map</h1>
<!--Render the map here-->
{{map|safe}} <!--We put double bracket because to tie an object with the page together , safe is a filter to hide or invisible the codes on the backgroun

<h1 class="h3">Pipeline Visualization</h1>
<h2>Scatter Plot & Line Chart as the "Chosen One"</h2>
<!--ScrollBar-->
Scatter Plot provides relationship between multiple numerical attributes. Another reason to use scatt

<!-- <img src="/Asset/Metocean_PL325.png"> -->

<h1 class="h3">Features</h1>
Metocean Pipeline 325 displays its visualizations in red scatters and line plot. Pipeline is located in the coastal area of Bintulu, Sarawak, Malaysia.This
Total Pipeline Distance (KM):<span style="font-size: 40px; righ

<!--End Row2/map and data-->

<!-- Optional JavaScript; choose one of the two! -->
<!-- Option 1: Bootstrap Bundle with Popper -->
<script src="https://cdn.jsdelivr.net/npm/bootstrap@5.1.3/dist/js/bootstrap.bundle.min.js" integrity="sha384-ka7Sk0Gln4gmtz2MlQnikT1wXgYsOg+OMhuP+IlRH9sENBO0LRn5
<!-- Option 2: Separate Popper and Bootstrap JS -->
<!--
<script src="https://cdn.jsdelivr.net/npm/@popperjs/core@2.10.2/dist/umd/popper.min.js" integrity="sha384-7+zCNj/IqJ95wo16oMtfsKbZ9ccEh31eOz1HGyDuCQ6wgnyJNSYdrPa
<script src="https://cdn.jsdelivr.net/npm/bootstrap@5.1.3/dist/js/bootstrap.min.js" integrity="sha384-QJHtvGhmr9XOIpI6YVutG+2QOK9T+ZnN4kzFN1RtK3zEFEIsxhlmWl5/YES
-->
PROBLEMS
OUTPUT
TERMINAL
DEBUG CONSOLE
[27/Nov/2021 11:29:02] "GET /map1 HTTP/1.1" 200 36854
Ln 16, Col 11
Spaces: 2
UTF-8
LF

URL

```python
#list of django urls of each page
urlpatterns = [
    path('admin/', admin.site.urls),
    path('',map_views.index,name='Home'),
    path('map1', map_views.map1, name='map1'), #connect url with views
    path('map2', map_views.map2, name='map2'),
    path('map3', map_views.map3, name='map3'),
    path('map4', map_views.map4, name='map4'),
    path('map5', map_views.map5, name='map5'),
    path('map6', map_views.map6, name='map6'),
    ]
```

Settings

```python
TEMPLATES = [ #directory where Django look for templates
    {
        'BACKEND': 'django.template.backends.django.DjangoTemplates',
        'DIRS': [os.path.join(BASE_DIR, 'templates')],
        'APP_DIRS': True,
        'OPTIONS': {
            'context_processors': [
                'django.template.context_processors.debug',
                'django.template.context_processors.request',
                'django.contrib.auth.context_processors.auth',
                'django.contrib.messages.context_processors.messages',
            ],
        },
    },
]
```