Coursework 2 – Data Visualization Website

1. Summary

Version: 1.0

Individual project

Weighting: 50% of overall mark

Deadline for project proposal: 16:00 Friday 21st January 2022 (end of Week 14)

Deadline for final submission: 16:00 Friday 1st April 2022 (end of Week 24)

2. Key Points

- You create a website that displays numerical data, predictions about the numerical data and the results of sentiment analysis.
- The numerical data will be obtained from web services. It cannot be obtained from web scraping. For example, it could be product price data from web services, stock prices, exchange rate prices, weather, football results, etc.
- The text data for sentiment analysis will be obtained from web services, such as the Twitter API or Facebook Graph.
- Machine learning will be used to make predictions about future values of the data.
- You will also display synthetic data that we will provide to check your data visualization and machine learning.
- All third party data will be stored in the cloud.
- The front end of the website only has to display visualizations of the data, predictions about the data and the results of the sentiment analysis. No other functionality is required.
- The code that downloads data from web services and uploads it to the cloud must be written in TypeScript.
- Code running in the cloud must be implemented in JavaScript/TypeScript and run as Lambda functions (not a virtual machine, such as an EC2 instance).
- Your website will be hosted on the cloud using serverless technology.
- The front end of your website can use ordinary JavaScript or a JavaScript framework.
- WebSockets will push new data items to subscribed clients.
- The coursework and teaching materials will be based on Amazon Web Services (AWS). You must use the AWS Educate account that is provided for this coursework.
- The final submission of your project will only receive a mark if your submission includes a video demonstration.

3. What Needs to be Submitted

3.1 Project Proposal (Deadline: 16:00 Friday 21st January 2022)

A document that contains:

- Brief description of the proposed website.
- Mock-ups of front end of website showing proposed data visualization. These can be pen and paper, Word, Inkscape etc.
- List of source(s) of numerical data.
- List of source(s) of data for sentiment analysis.
- Screenshot and URL of static website hosted on cloud.

Submit Word or PDF version of project proposal using the link in the Coursework 2 section of the course website. **Your proposal must be in Word or PDF format.**

We will use the project proposal to give you feedback about your idea and help you realize it in the time available. You can reuse material from the project proposal in the final project submission.

The project proposal will be marked online. It is worth 10% of the mark for Coursework 2.

3.2 Final Submission (Deadline: 16:00 Friday 1st April 2022)

Submit a zip file that contains:

1. *Project report*. This must include:
   - Screenshot(s) of the front end of your website.
   - Screenshots of all the data visualization. The pictures of the data visualization must be high resolution so that we can check your predictions about the synthetic data.
   - Architecture diagram showing the relationships between Lambda functions, API Gateway, database, etc.
   - Description of the website. You should explain the machine learning and sentiment analysis and how your lambda functions and database work.
   - **Do not include screenshots of code.**
   - **This must be a Word or PDF document.**

2. *Source code*. Your source code folder should contain the following files:
   - TypeScript source files.
   - Source code for Lambda functions.
   - Source files for front end of website, for example, HTML, JS, Vue.js files etc.
   - **Please do not include the node_modules folder in your submission. It is likely to make your submission too large to upload!**

3. *Machine learning files*.
   - Include all of the files that you used for training and testing the machine learning in a separate folder.

4. *5-minute video demonstration*. **Video demonstrations are mandatory for the final submission.** I strongly recommend that you watch the talk on recording video demonstrations on the course website.

Upload the zip file using the link in the Coursework 2 section of the course website.

Marks will only be allocated for functionality that exists in your submitted files, so make sure that you upload all of the files for your project.

The final submission is worth 90% of the mark for Coursework 2.

4. Technology

This project must be implemented using the following technology:

- The code for uploading numerical and text data must be written in TypeScript.
- Code in the cloud must be implemented in JavaScript/TypeScript.
- Code in the cloud must run as Lambda functions (not on a virtual machine, such as an EC2 instance).
- Machine learning must be done with AWS SageMaker.
- Sentiment analysis must be done using AWS Comprehend.
- WebSockets must be running in the cloud using AWS API Gateway.
- No marks will be given for functionality implemented in Python.
- You must use the AWS Educate account that is provided for this coursework.

We are happy to discuss alternative ways of implementing your project, but you must get our agreement first. **Functionality that is implemented using different technology without the permission of the module leader will receive zero marks.**

5. Formative Feedback

Formative assessments do not directly contribute to the overall module mark but they do provide an important opportunity to receive feedback on your learning. They provide an opportunity to evaluate and reflect on your understanding of what you have learnt. They also help your tutors identify what further support and guidance can be given to improve your grade.

We are happy to give you feedback about Coursework 2 in the labs and can also give feedback about drafts of Coursework 2 that are sent to us more than one week prior to the deadline.

6. Extenuating Circumstances

If you have personal problems that interfere with your studies, you can apply for extra time to complete coursework without a mark penalty. You have to provide appropriate documentary evidence.

More information here: https://unihub.mdx.ac.uk/your-study/assessment-and-regulations/extenuating-circumstances.

You must let the module leader know if you have been granted an extension.

7. Late Submission

We are very unlikely to give extensions to coursework and very unlikely to accept excuses. So we strongly recommend that you hand coursework in on time.

Contact the module leader before the deadline if you run into problems. **Zero marks will be awarded for coursework that is received more than six hours after the deadline.**

8. Plagiarism

Plagiarism is a serious academic offence. Students that submit identical projects will be reported to the university. If they are found guilty, they will have to resubmit their work, their marks could be capped or they could fail the module.

We recognize that there is often a blurry line between copying and collaboration. People work together and help each other to solve problems and apply the solutions to their own projects. We strongly encourage this kind of collaboration. But it is not acceptable for students to collaborate on a project which they submit as individual work. To penalize this, the mark for near-identical projects will be divided between the projects. So suppose a project gets a mark of 60% and near-identical versions are handed in by 3 people. Each person will get 20%, instead of 60%. This only applies to the marks for the parts of the project that are identical.

We are not going to police this and make detailed investigations. So if you allow your project to be copied, you will be as liable for plagiarism as the person who submits it as their own work. Both the original and the copy will receive zero or reduced marks.

Links to the relevant University regulations and additional support resources can be found here:

- Academic Integrity Awareness Course: https://mdx.mrooms.net/mod/lesson/view.php?id=877307. (You will have to log into to MyUniHub and then MyLearning to access the course.)
- Section F: Infringement of Assessment Regulations/Academic Misconduct: https://www.mdx.ac.uk/about-us/policies/university-regulations.
- Referencing & Plagiarism: Suspected of plagiarism?: http://libguides.mdx.ac.uk/c.php?g=322119&p=2155601.
- Referencing and avoiding plagiarism: http://unihub.mdx.ac.uk/your-study/learning-enhancement-team/online-resources/referencing-and-avoiding-plagiarism.

The MDXSU Advice Service offers free and independent support face-to-face in making an appeal, complaint or responding to any allegations of academic or non-academic misconduct. https://www.mdxsu.com/advice.

9. Assessment Methods

9.1 Project Proposal

We will read your project proposal and give you feedback online and in the labs.

9.2 Final Submission

We will look at the code, read your report and view up to 5 minutes of your video demonstration. Your video demonstrations should not be significantly longer than 5 minutes. **Zero marks will be awarded for a final submission of Coursework 2 without a video demonstration.** I strongly recommend that you watch the talk on recording video demonstrations on the course website.

The project will be given a mark out of 100. This will be scaled down to a mark between 0 and 50 that corresponds to 50% of the overall mark for the module.

10. Assessment Criteria

Feature	Deadline	Marks
Project proposal. Must include: • A short description of the proposed website. • List of web services that will be accessed to obtain data for the machine learning and sentiment analysis. • Mock-ups of front end of website showing proposed data visualization. • Screenshot and URL of static website hosted on the cloud.	**16:00 21/1/22**	**1 mark**. Brief description of proposed website. **1 mark**. List of source(s) of numerical data. **1 mark**. List of source(s) of data for sentiment analysis. **1 marks**. Mock-ups of front end of website showing proposed data visualization. These can be pen and paper, Word, Inkscape etc. **2 marks**. Proposal quality. Is it clearly written? Are the data sources sensible? Do the wireframes clearly show the website design?
Static website hosted on cloud. A static HTML page containing at least one image hosted on the cloud.	**16:00 21/1/22**	**2 marks.** Static 'Hello World' HTML page hosted on cloud. **2 marks**. Static website contains an image that is also hosted on the cloud.

Cloud storage of data from third party web service(s). The code for uploading the data must be written in TypeScript – there are no marks for using Java or ordinary JavaScript. Aim for at least 1000 data points.	**16:00 1/4/22**	**2.5 marks.** Download of numerical data from third party web service(s). **2.5 marks**. Download of text data for sentiment analysis from third party web service(s). **4 marks**. Storage of data from third party web service(s) in cloud database.
Machine learning. Application of machine learning to numerical data stored in the cloud. The machine learning must be done using AWS SageMaker.	**16:00 1/4/22**	**5 marks**. Use of machine learning to generate correct predictions about synthetic data. **5 marks**. Use of machine learning to generate predictions about numerical data from third party web service(s).
Sentiment analysis. Application of sentiment analysis to text data stored in the cloud. The sentiment analysis must be done using AWS Comprehend.	**16:00 1/4/22**	**10 marks**. Sentiment analysis of text data stored in the cloud.
Serverless. Back end of website runs in the cloud using serverless technology.	**16:00 1/4/22**	**5 marks**. Back end of website runs in cloud with serverless technology. **5 marks**. Events connect Lambda functions together for data processing. For example, database updates trigger lambda functions, which update the sentiment analysis and push new data to the client. 2.5 marks per database trigger used for data processing or sending new data to client.
Cloud hosting of website. All of the front end files for the final website are hosted on the cloud. Website can be accessed through a public URL.	**16:00 1/4/22**	**5 marks**. Website entirely hosted on the cloud, for example, using Amazon S3.
Data visualization. Visualization of data on website using graphs and other appropriate techniques.	**16:00 1/4/22**	**2.5 marks**. Visualization of third party numerical data stored in the cloud. **2.5 marks**. Visualization of synthetic data. **2.5 marks**. Visualization of predictions about third party numerical data. These predictions must have been generated by machine learning. **2.5 marks**. Visualization of predictions about synthetic data. These predictions must have been generated by machine learning. **5 marks**. Display of the results of sentiment analysis.
Website design. Design of the front end of the website.	**16:00 1/4/22**	**5 marks.** How close is the design of the website to professional quality?

Data quality. How much data is shown? How good is the data visualization?	**16:00 1/4/22**	**5 marks.** Amount of data and quality of data visualization.
WebSockets. The WebSocket back end must be running in the cloud using AWS API Gateway.	**16:00 1/4/22**	**5 marks.** WebSockets send data to a single client when it connects. **5 marks**. WebSockets broadcast new data to all connected clients when the database changes.
Code quality. For example, comments, layout, organization, etc.	**16:00 1/4/22**	**2.5 marks**. Typescript code quality. **2.5 marks**. JavaScript code quality.
Project report. Briefly describes the project and includes high resolution screenshots of the data visualization. **Do not include screenshots of code.**	**16:00 1/4/22**	**2 marks**. Screenshot(s) of the front end of website and the data visualization. The pictures of the data visualization must be high resolution so that we can check your predictions about the synthetic data. **2 marks**. Architecture diagram showing the relationships between Lambda functions, API Gateway, database, etc. This must be accurate. **2 marks**. Content of report. Does it clearly describe the project? You should explain the machine learning and sentiment analysis, describe how your lambda functions work and document your database.