

## F20DV Data Visualisation and Analytics

# Coursework Brief – Group Project

**Note:** This coursework brief is a draft and subject to changes. Check the Canvas page for up-to-date announcements.

## I - Introduction

For this coursework, you are required to **design and develop an interactive data visualisation application for the dataset(s) of your choice** (approved by your lecturer). The aim of such application is to provide an audience with the means to explore and identify insights in the dataset. The coursework will enable you to demonstrate a variety of skills, such as, problem solving, communication skills and time-management. The coursework is structured to encourage a mix of guided and active learning, innovative development, critical thinking and knowledge-based expertise of the subject.

The core of the work will largely take place in the second half of the course (Week 6 – 12), with some preparation tasks in the late first half (Week 4 – 5).

This project will be carried out in **groups of 4 students**. This project accounts for 80% of your final course grade. While largely relying on group components, individual contributions will impact your final grade.

**Read this brief carefully. You might see yourself (and your group) losing marks if you do not pay attention to the instructions below. You should also look out for announcements on Canvas for any potential update and clarifications.**

With few exceptions, lab sessions in the second half of the course will be dedicated to assisting your group for this project. These are the best options for your group to clarify details, get some help with the coursework, and receive feedback on your progression (and feed forward for your final submission). You should work on this coursework in your own time, not just the scheduled weekly lab sessions.

## 2 - Group Work

This project will be conducted in groups of 4.

You should form your own groups. Groups should be from the same cohort, i.e., there cannot be a group with both F20DV and F21DV students. Individuals with no groups, or groups with a missing member, will be randomly allocated after the group registration deadline (see key dates below) by the teaching team.

While planning and carrying out group work it is important to allocate precise roles and tasks to each member of the group. Every group member is expected to contribute to the development of the application, which must be recorded in the Git logs of the group's project repository.

Most of the marks will be allocated as a group component; however, this group mark will be modified for each member according to their individual contribution (measured by the group's Git repository) and the peer review form. Your final individual grade will also comprise an individual report.

Some tips:

- It is the responsibility of every group member that everyone in the group participates and that the work is divided fairly amongst group members.
- It is the responsibility of every group member to ensure that everyone in the group respects academic integrity rules, does not plagiarise content and properly cites external sources.
- You should meet regularly, at a minimum of once a week.
- In your initial meeting, the group must designate a "group leader", who is responsible for communicating group queries to the teaching team and responsible to upload group submissions to Canvas. After that meeting, update the group on Canvas to reflect the selected group leader.
- In the initial meeting, it is essential to establish a primary mode of communication, to organise meetings and keep other group members up to date with your tasks.
- Assess each-other's strengths and weaknesses, to effectively divide tasks.
- During meetings, the group (pre-agreed group member) should allocate precise tasks or activities.
- Some group marks (in presentations and in the implementation) rely on your cooperation as a group.
- If one or two of your group members are not participating/unresponsive, you should notify your lecturer as early as possible. In the best case, this can be resolved quickly, in the worst case, there are evidence that will be considered during marking.
- Similarly, if you find yourself "cut" from the group and your group members are unresponsive to you, notify your lecturer so that a resolution might be achieved quickly.

### 3 - Instructions and Tasks

The stages below highlight key moments in the project, with associated deadlines.

#### Stage 0 – Group Formation

Your first task is to form your group. As instructed above, you should find 3 other group members to work with. Group members cannot be from the different cohorts, i.e., every group member should be from F20DV, not F21DV. Ask your lecturer if you are not sure which cohort you are part of.

#### Group Registration

By the end of week 5, you must register yourself to a group on Canvas (*People* section, then your respective cohort and campus *Groups* tab, e.g. F20DV Malaysia Campus).

The lecturer will allocate individuals randomly to form groups with 4 members if:

- You are not registered in a group; or
- Less than 4 students are registered to a group.

In rare instances, it can mean that groups with already two or three members might be split and allocated to other groups.

## GitLab

Once your group has been confirmed, around the start of week 6, you should set up a GitLab repository for your group. You should have been added to the course GitLab group. If this not the case, please inform your lecturer to fix this.

The link to the group is: <https://gitlab-student.macs.hw.ac.uk/f21dv-2023-24>.

The group leader member of your group should fork the project repository template (f20dv-group-project) and then add the other members as "Developer" role members to the fork project.

Follow the separate *GitLab How To* document to set up GitLab on your computer.

## Stage 1 – Data Collection and Preparation

In parallel to your group formation, you should research datasets that you intend to use for this project.

You may select any datasets, if **you have the rights to use it**. You are expected to select **several related datasets** (unless the single dataset you choose displays enough complexity). The topic of these datasets is open to your own interests.

Some resources to start:

- The Gapminder project: <https://www.gapminder.org/>
- The Data Is Plural newsletter: <https://www.data-is-plural.com/>
- The UCI repository: <https://archive.ics.uci.edu/>
- The UK open data repository: <https://www.data.gov.uk/>
- Google's Dataset search: <https://datasetsearch.research.google.com/>

You are expected to engage and discuss your dataset selection with your respective teaching team, especially if you are unsure of its applicability for this project.

In the end, the data you will use for this project must contain:

- A variety of attribute types: categorical, numerical, etc.
- At least one complex structure, either hierarchical or geographical.
- A significant number of records (more than 50 per dataset; or 200 if you use one dataset) and attributes (more than 12).

You are allowed to use tools such as OpenRefine, MS Excel or Python to clean or restructure the datasets that you find into more suitable formats.

## Dataset Registration

**By the end of week 5**, you should submit details about the datasets you intend to use for this project, using the dedicated assignment activity on Canvas.

For each dataset, you must specify:

- A public URL where one can retrieve the dataset;
- Details about the license for the dataset;
- A high-level description of the dataset, including key attributes and complex structures;
- Any modifications that you have made (or intend to make) to the dataset, with a short justification and details of the methodology used.

Groups that did not register their datasets by week 5, might be provided with a default dataset to use for this project. However, this is the opportunity for you to investigate datasets that interest you and create your own theme for your project. The default dataset has a default theme that groups cannot deviate from.

## Stage 2 – Design

Once your group and dataset have been confirmed, you can start with the design stage.

In this stage, you will:

- Study and understand the aim and structure of your datasets: what are key purposes for these datasets? What insights or categories of insights do you think you will be able to find?
- Perform basic Exploratory Data Analysis (EDA) to confirm, disprove, or discover new insights.
- Decide on appropriate analytics, visualisations and interactions to support your audience in exploring the dataset and discovering the insights you have identified.
- Create mockups (hand drawn or using a drawing software) for your dashboard, showcasing its visualisations, their purpose and the interactions connecting them.
- Prepare a short presentation (maximum 7 minutes) highlighting all the above.

Your design must comprise **at least three different types of interactive visualisations**. You are strongly invited to propose design ideas beyond what you have seen in lectures and labs.

Your group will deliver your design presentation during your allocated session in Week 8. This presentation should be short, less than 7 minutes, so you must be concise. All members must attend the presentation and deliver some part of it. If one or more members do not attend the presentation without valid reasons, they will be awarded an 'ABS' and 0 marks for this section of the coursework.

The design presentation represents 20% of the group project's mark. You will be graded on the contents and clarity of the presentation.

As part of the presentation process, you will receive feedback on your design, that you may take on board for the next stage.

The group leader must submit their group's presentation materials (e.g. slides) on Canvas, using the appropriate assignment activity.

## Stage 3 – Development

From Week 8 to Week 11, you are expected to work on the implementation of your dashboard.

By that point, you would have received feedback on your design and explored more visualisations in lectures and labs. Reflecting and adapting your visualisation and interaction in consequence should be a priority.

You should interact with the teaching team as much as possible during lab sessions to gather feedback on your progress.

Your group must use its forked GitLab repository to contribute to the development of your dashboard. Logs of commits made throughout the duration of this project will be reviewed and constitute part of your individual marks. You might lose marks (individually) if:

- You have made significantly less contributions than your peers;
- Your contributions have been significantly smaller than your peers;
- Your contributions have been irregular, e.g., all within a few days before the deadline.

You are not allowed to use any other library or programming language in your development other than:

- **Standard JavaScript, HTML and CSS features**
- **D3.js version 7+**

Your project must run on common web browsers found in the Computer Laboratory on your respective campus: Google Chrome / Edge v120; Mozilla Firefox v120. You can use the website [caniuse.com](https://caniuse.com) to check if language features are supported. Your dashboard must be contained within a standard browser window (around 1900 x 900 pixels).

Importantly you should maintain a *README* markdown file in your project that will contain and describe the structure of your project: what are the files? What is their purpose? How do they relate to/depend on each other? Someone accessing your project for maintenance or improvements should get a comprehensive understanding of your application's architecture from this file. Each script should also be appropriately commented.

The actual code project represents 40% of the group project's mark. You will be graded on whether you have met requirements and the quality of your code. Peer-reviews and Git logs might affect your individual marks.

Once completed, your project's code should be archived in a **.zip** file and uploaded to Canvas, using the dedicated assignment activity. You must exclude your datasets from this submission, only your code is required.

## Stage 4 – Demonstration and Report

### Demonstration

Your final project will be demonstrated during Week 12.

Like the design presentation, this should be a short demonstration (less than 8 minutes). Within this time frame, you should present your dashboard:

- The visualisations implemented, and the data they show.
- The interactions implemented, and how they impact/update the visualisations.
- Key insights about the datasets that result from examining and interacting with the visualisations.

You should also present, in high-level, the overall architecture of your application.

While you shouldn't use slides for your demonstration, you must build a storyboard (key screenshots with brief comments) of your demonstration using slides and submit these on Canvas using the appropriate assignment activity.

The demonstration represents 20% of the group project's marks. You will be graded on the content and clarity of your demonstration.

## Report

You should also write a short **individual** report (maximum of one A4 page, not including references) about your involvement in this project:

- A **personal** reflection about your contribution to this project and how you approached tasks in your group, highlighting key challenges (technical issues, knowledge gap, group work problems, etc.), how you addressed them, and what you have learned from them. You should also reflect on the positive elements and how you would reimplement them or improve them in the future. This report is not a group work. You must write and submit it on your own. It represents 20% of the group project's mark. You must submit in on Canvas, using the dedicated assignment activity.

## 4 – Requirements

### Dataset

- You must select several related datasets to work on during this project.
- In case a single dataset shows enough complexity, you work with it on its own.
- In total, your data should comprise a mix of 12 or more categorical, ordinal and numerical attributes.
- Your data should have more than 50 records per dataset, or more than 200 if you are working with a single dataset. Discuss this requirement with your lecturer.
- Your data should include a complex structure, geographical or hierarchical.
- You must have the rights to use this dataset.

You will be provided with an alternative dataset if your suggestion does not adhere to the above.

### Application Core Requirements

You must meet these requirements to get a passing grade (40%).

- Your design should be informed by broad questions/assumptions about the data, supported by basic EDA insights.
- Your design must include interactions to highlight and filter data items within and between visualisations.
- You must develop and collaborate on your project with the provided Gitlab repositories.



- Your application must consist of a single HTML page, with scripts and styles implemented that load, transform, display and interact with data dynamically.
- Only standard HTML, CSS, JavaScript, and D3.js (version 7) features may be used for this project. You cannot use server-side code or CSS / JS supersets (e.g., Sass or TypeScript).

## Additional Requirements

- Your design must include 4 different types of charts.
- One of the charts you implement should demonstrate complexity and creativity, and differ from the designs shown in lectures or labs.
- Your dashboard design (visualisation and interactions) should be intuitive to use.
- Interactions between visualisations should be bidirectional.
- Animations should be used when updating visualisations.
- Your application should be consistent with its style.
- Your dashboard should fit within a standard browser window (1900\*900).
- The visualisations should show different facets of the same data.
- Your code must be documented, including a top-level README file outlining the application architecture.
- Your application and its documentation (including comments) should facilitate maintenance: high-level description of functions, details about expected parameters' format, clean code...

## 5 - Deliverables – Key Dates

### Summary of Schedule

Note that this schedule is only a suggestion. For instance, you can start pre-analysing your data while researching suitable datasets; your dashboard implementation might start while you perform the exploratory data analysis ...

Week	Activity	Deadline
1	N/A	
2		
3	Group formation and dataset research	Group project start
4		
5		Group and Dataset Registration (0%)
6	Dataset analysis and dashboard design	
7		
8		Design Presentation (20%)
9	Dashboard implementation	
10		

11		
12		Final Project (40%) Presentation (20%) Report (20%) Peer Review (Modifier)

## Group/Dataset Registration (0%)

Form your groups on Canvas (*People* section, *Groups* tab). You should agree on an empty group to join, and individually add yourselves to it.

Using the dedicated assignment on Canvas, submit details about the data you will use for this project: URLs, licenses, descriptions, suitability for the project, modifications made, etc. Details about the datasets should be short and concise.

**Deadline: Thursday of week 5 – 15<sup>th</sup> February 2024, by 23:59 (campus local time)**

## Design Presentation (20%)

Along with the presentation you will give in Week 8, you must submit your presentation slides on Canvas, using the appropriate assessment drop box.

Your slides must include:

- Group information: members' names and HWU id.
- Dataset presentation and analysis.
- Design choices targeting an identified audience.
- Mockups of your data visualisation dashboards.

The time and date of your presentation will be communicated in advance. Slides should be submitted by the deadline below.

Note that your presentation only will be marked. Elements added to the slides after the presentation will not be marked.

While the mark for this presentation will be given to the group, the markers reserve the right to deduct marks for individuals if they are not participating to the presentation.

This is a group submission, meaning that one member can submit for the whole group.

**Deadline: Thursday of week 8 – 7<sup>th</sup> March 2024, by 23.59 (campus local time)**

## Final Demonstration (20%)

With the demonstration you will give in Week 12, you must submit your demonstration's storyboard on Canvas, using the appropriate assessment drop box.

Your storyboard should be in the format of slides and must include (you should use this storyboard to prepare for your demonstration):

- Group information: members' names and HWU id.



- Annotated screenshots of your application for key demonstration elements: e.g., interactions, insights, etc. Annotations should be short and concise.
- Diagram of your application's architecture.

The time and date of your demonstration will be communicated in advance. Slides should be submitted by the deadline below.

Note that your demonstration only will be marked. Elements added to the slides after the presentation will not be marked.

While the mark for this demonstration will be given to the group, the markers reserve the right to deduct marks for individuals if they are not participating to the demonstration.

This is a group submission, meaning that one member can submit for the whole group.

**Deadline: Thursday of week 12 – 4<sup>th</sup> April 2024, by 23.59 (campus local time)**

## Final Project (40%)

The dashboard's development should be incremental and organised via the Gitlab repository for your group.

As highlighted above, Git logs will be used to assess your contributions' regularity and significance. A penalty of up to -5 marks may be applied. If a group member does not contribute at all, they may receive 0 for this component.

On completion of your project, you should create a **.zip** archive of your code (not datasets) and upload it to Canvas, using the appropriate assessment drop box.

This is a group submission, meaning that one member can submit for the whole group.

**Deadline: Thursday of week 12 – 4<sup>th</sup> April 2024, by 23.59 (campus local time)**

## Peer Review (modifier)

Along with your final project submission, you are instructed to complete a Peer Review form, using the MS Form linked on Canvas.

The aim of this form is to gather your view on the contribution of your team members.

From this review, marks for each member in the group may vary between +5 and -5 marks. If it is made apparent that one group member did very little or no contribution to the group work, the marker reserves the right to assign them a 0.

For example:

- Group members A and B report that they each contributed equally, while member C did a bit less. Member C reports that they contributed less than A and B. A and B would get +2 points each, while member C would get -4 points.
- Group member A reports that B did a bit more than they did, and C did a lot more. B and C concur. A would get -3 points, B -2 points, C +5 points.
- Group members A and C report that B did not contribute at all to the work, while they worked equally. A and C would not see a change, B would get 0.

This modification will be applied on the Final Project mark and checked against the Git log of your project repository.

This is an individual submission, meaning that you must submit your own review. 5 marks may be deducted from your grade if you do not submit a review.

**Deadline: Thursday of week 12 – 4<sup>th</sup> April 2024, by 23.59 (campus local time)**

## Individual Report (20%)

You must submit a PDF of your report on Canvas, using the appropriate drop box.

This is an individual submission, meaning that you must submit your own report.

**Deadline: Thursday of week 12 – 4<sup>th</sup> April 2024, by 23.59 (campus local time)**

## Late Submissions

Late submissions will be subject to the normal penalties as defined in the late coursework policy.

The University recognises that, on occasion, students may be unable to submit coursework on the submission date or be unable to present their work on the submission date. In these cases, the University's Submission of Coursework Policy outlines are:

- No individual extensions are permitted under any circumstances.
- Standard 30% deduction from the mark awarded (maximum of five working days).
- In the case where a student submits coursework up to five working days late, and the student has valid mitigating circumstances, the mitigating circumstances policy will apply, and appropriate mitigation will be applied.
- Any coursework submitted after five working days of the set submission date shall be automatically awarded a no grade with no formative feedback provided.

Please contact your Personal Tutor or Counsellor if you are unable to meet the deadlines or need information for Mitigating Circumstances or Temporal Suspensions of Studies.

## Notes

- Coursework reports must be written in your own words. If some text or diagrams in the coursework has been taken from other sources, these sources must be properly referenced and should be open for use. The link to the University's policy on Plagiarism is:

<https://www.hw.ac.uk/students/studies/examinations/plagiarism.htm>

- Failure to reference work that has been obtained from other sources or to copy the words of another student is plagiarism, and if detected, this will be reported to the School's Discipline Committee. If a student is found guilty of plagiarism, the penalty could involve voiding the course.
- You must never give hard or soft copies of your coursework reports to another student. You must always refuse any request from another student for a copy of your report.

- Sharing a coursework report with another student is collusion, and if detected, this will be reported to the School's Discipline Committee. If found guilty of collusion, the penalty could involve voiding the course.
- You can expect to receive formative feedback on the coursework within 15 working days after the submission deadline.

DRAFT

## Marking Rubrics

### Design Presentation

Each criterion – apart from <i>Presentation quality</i> – is marked along the following scale: 0 – Missing: the criterion is not included in the presentation 1 – Insufficient: the criterion is missing several elements, or severely lacking in quality 2 – Poor: the criterion is incomplete in few aspects 3 – Good: the criterion is complete 4 – Excellent: claims/decisions are justified appropriately 5 – Outstanding: creativity/out-of-the box thinking is displayed in the analysis/design		
Criterion	Description	Marks
Dataset presentation	The datasets' provenance and purpose are outlined. Assumptions and interests are explained.	/5
Dataset analysis	Data preparation and/or EDA steps and key insights are outlined.	/5
Design choices	Key visualisations and interactions are outlined and explained in mockups. Choices are appropriate.	/5
Presentation quality	1 mark per element: the presentation is well organised; the progression is fluid; the presenters are clear; the presenters are engaging with the audience; the slides are polished.	/5
<b>Group Total</b>		<b>/20</b>
Penalties are applied to members who did not contribute to the presentation: -5 marks for not actively presenting (out of 20)      Total of 0 for not attending the presentation		
Student:		<b>/20</b>
Student:		<b>/20</b>
Student:		<b>/20</b>
Student:		<b>/20</b>

## Final Demonstration

Each criterion – apart from <i>Demonstration quality</i> – is marked along the following scale: 0 – Missing: the criterion is not included in the presentation 1 – Insufficient: the criterion is missing several elements, or severely lacking in quality 2 – Poor: the criterion is incomplete in few aspects 3 – Good: the criterion is complete 4 – Excellent: claims/decisions are justified appropriately 5 – Outstanding: creativity/out-of-the box thinking is displayed in the analysis/design/implementation		
Criterion	Description	Marks
Visualisations and Interactions	The visualisations' content and purpose are outlined. The interactions/links between visualisations are shown. If necessary, improvements/future work are explained.	/5
Insights	Data insights are highlighted using the visualisations and interactions implemented.	/5
Code structure	The dashboards' code architecture is presented and reflected on.	/5
Demonstration quality	1 mark per element: the demonstration is well organised; the demonstration is fluid; the presenters are clear; the presenters are familiar with the material; the presenters are engaging with the audience.	/5
<b>Group Total</b>		<b>/20</b>
Penalties are applied to members who did not contribute to the demonstration: -5 marks for not actively demonstrating (out of 20)      Total of 0 for not attending the demonstration		
Student:		<b>/20</b>
Student:		<b>/20</b>
Student:		<b>/20</b>
Student:		<b>/20</b>

## Final Project

Each criterion – apart from <i>Code quality</i> – is marked along the following scale:			
0 – Missing: the criterion is not included in the code			
1 – Insufficient: the criterion is missing many elements/core requirements, or severely lacks in quality			
2 – Poor: the criterion is incomplete in few aspects but meets core requirements			
3 – Good: the criterion is complete and meets most requirements			
4 – Excellent: elements beyond the taught material have been successfully incorporated			
5 – Outstanding: creativity/out-of-the box thinking is displayed in the methodology/problem solving			
Criterion	Description		Marks
Data processing	Datasets are dynamically loaded and parsed within the application. Appropriate data processing methodologies are used efficiently and allow responsive data queries.		/10
Visualisations	Data-appropriate visualisations are implemented. Appropriate visual mapping methodologies are employed to make the visualisation scalable and reusable.		/10
Interactions	Meaningful and useful interactions are implemented within and between visualisations.		/5
Design and storytelling	The dashboards design is intuitive, serves the discovery of insights and supports storytelling in the data presentation.		/5
Code quality	2 marks per element: reusability: components are cohesive and decoupled. scalability: the implementation can adapt to new data. readability: the code is organised, commented and clean. maintainability: issues can be identified and fixed quickly. documentation: a README file outlines the code's structure.		/10
Group Total			/40
Peer-review: the peer-review can modify the student's grade within a +5 to -5 range.			
Git logs: penalties (up to -5 marks) are applied if there are no evidence of regular and significant contribution to the project.			
If there are no evidence of involvement at all, marks of 0 may be applied.			
Student	Peer Review	Git logs	Marks
			/40
			/40
			/40
			/40



## Individual Report

Criterion	Description	Marks
Personal reflection	Individual contributions are clearly outlined, along with a description of the student's role within the group. Key challenges or successes are well presented in their scope and (potential) impacts. Similarly, mitigations are described. The report shows reflection, outlining lessons learned and/or possible future work.	/15
Report quality	The report is well-presented with little grammatical/typological mistakes. References are appropriate and correctly incorporated.	/5
Total		/20