Population data search

Population Data Search is a SQL-aided website that provides geographical listings. Typical users of this site might be people with an academic interest in population data, such as students, or policy makers who need to be informed of population data, for example.

For this reason, we aimed to keep our website minimal, as the average user may not be particularly tech-savvy.

This report lays out the requirements we met, the features we implemented, future improvements and specific member contributions.

Meeting project requirements

The following section lays out our user stories/acceptance criteria, and direct comparison to how we met each.

Student user story

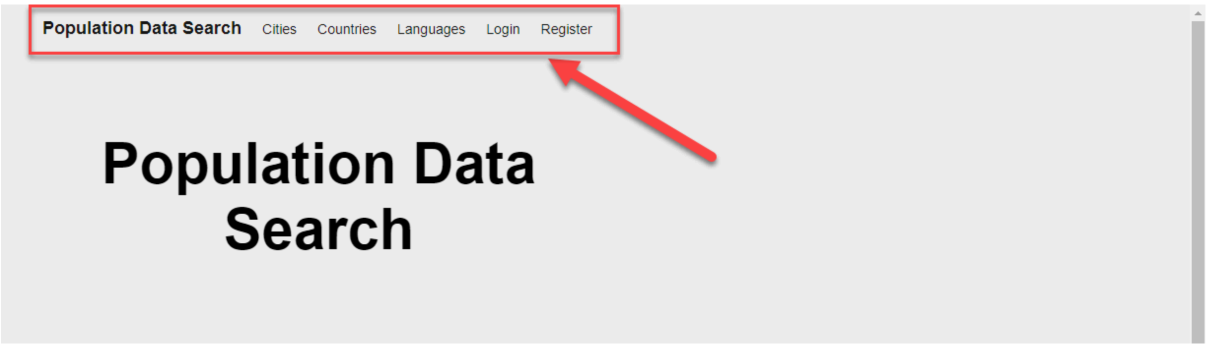
As a student, I can access hundreds of records of geographical data so that I can use accurate data in academic projects.

Student acceptance criteria 1

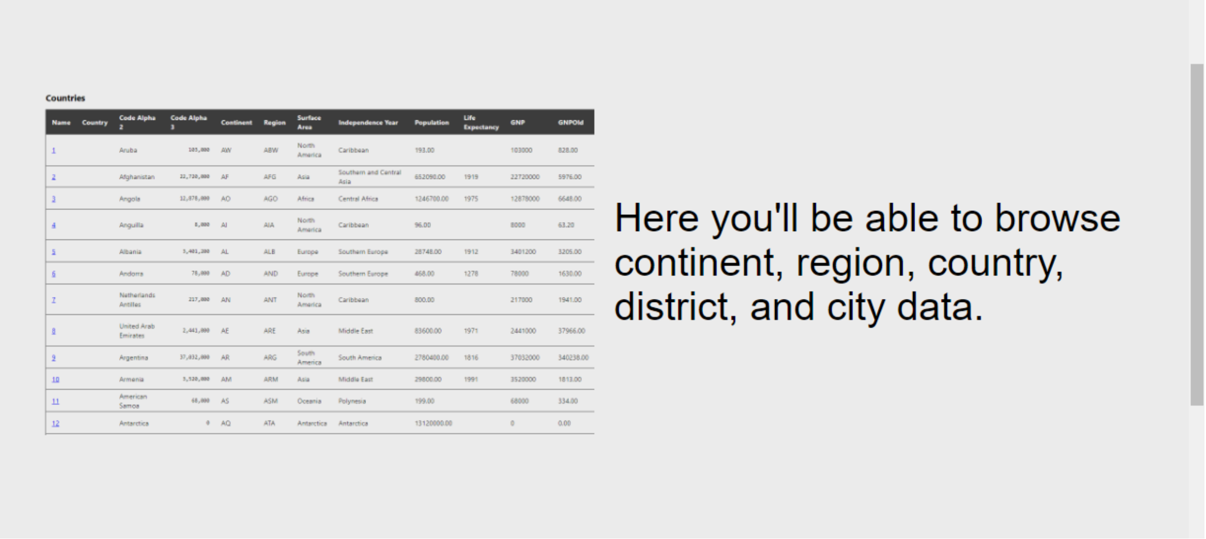
Given that I might not be very technically confident, when I use the population data system, then the interface must be clear and obvious.

How we met it

We tried to make the interface clear and obvious by placing a simple navigation bar at the top of the site, as this is where a user would expect to see it.



On the home page, we provided text describing what users get from the site, alongside screenshots of the content.



Countries

Name	Country	Code Alpha 2	Code Alpha 3	Continent	Region	Surface Area	Independence Year	Population	Life Expectancy	GNP	GNPOld
1		Aruba	103,000	AW	ABW	North America	Caribbean	193.00		103000	828.00
2		Afghanistan	22,720,000	AF	AFG	Asia	Southern and Central Asia	652090.00	1919	22720000	5976.00
3		Angola	12,878,000	AO	AGO	Africa	Central Africa	1246700.00	1975	12878000	6648.00
4		Anguilla	8,000	AI	AIA	North America	Caribbean	96.00		8000	63.20
5		Albania	3,401,200	AL	ALB	Europe	Southern Europe	28748.00	1912	3401200	3205.00
6		Andorra	78,000	AD	AND	Europe	Southern Europe	468.00	1278	78000	1630.00
7		Netherlands Antilles	217,000	AN	ANT	North America	Caribbean	800.00		217000	1941.00
8		United Arab Emirates	2,441,000	AE	ARE	Asia	Middle East	83600.00	1971	2441000	37966.00
9		Argentina	37,032,000	AR	ARG	South America	South America	2780400.00	1816	37032000	340238.00
10		Armenia	3,520,000	AM	ARM	Asia	Middle East	29800.00	1991	3520000	1813.00
11		American Samoa	68,000	AS	ASM	Oceania	Polynesia	199.00		68000	334.00
12		Antarctica	0	AQ	ATA	Antarctica	Antarctica	13120000.00		0	0.00

Here you'll be able to browse continent, region, country, district, and city data.

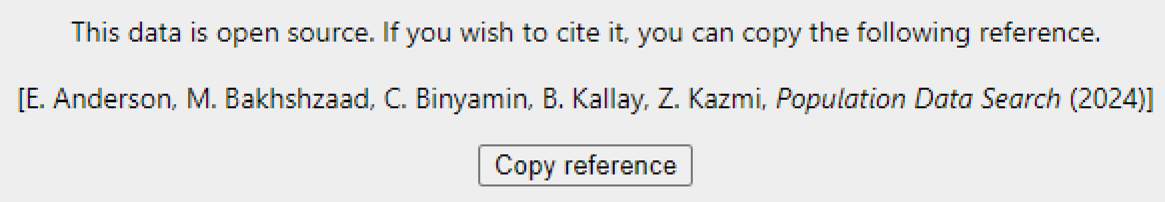
We maintained a minimal layout throughout the site, so that the user can focus on the data easily and not get lost in hundreds of colours, fonts, buttons, etc.

Student acceptance criteria 2

Given that I wish to use the data in projects, when I cite the source in academic work, then I need to be assured that the data is open source.

How we met it

It seemed that a good way to meet this criteria was to add a 'quick reference' feature which appears on all pages. So we implemented this idea, with a button to copy the reference to the clipboard.



This data is open source. If you wish to cite it, you can copy the following reference.

[E. Anderson, M. Bakhshzaad, C. Binyamin, B. Kallay, Z. Kazmi, *Population Data Search* (2024)]

Data entry professional user story

As a data entry professional, I can enter new data into the geographical database without having to worry too much about how accurately I enter the data because the system should be robust.

Data entry professional acceptance criteria 1

Given that I am an administrator in the system, when I input new data into the geographical database, then the system should validate the data for basic accuracy and integrity. It should be impossible to enter duplicate records into the database, for example.

Data entry professional acceptance criteria 2

Given that I may make mistakes during data entry, when I add data to the system, then the system should provide feedback on any errors or inconsistencies found in the data.

How we met it

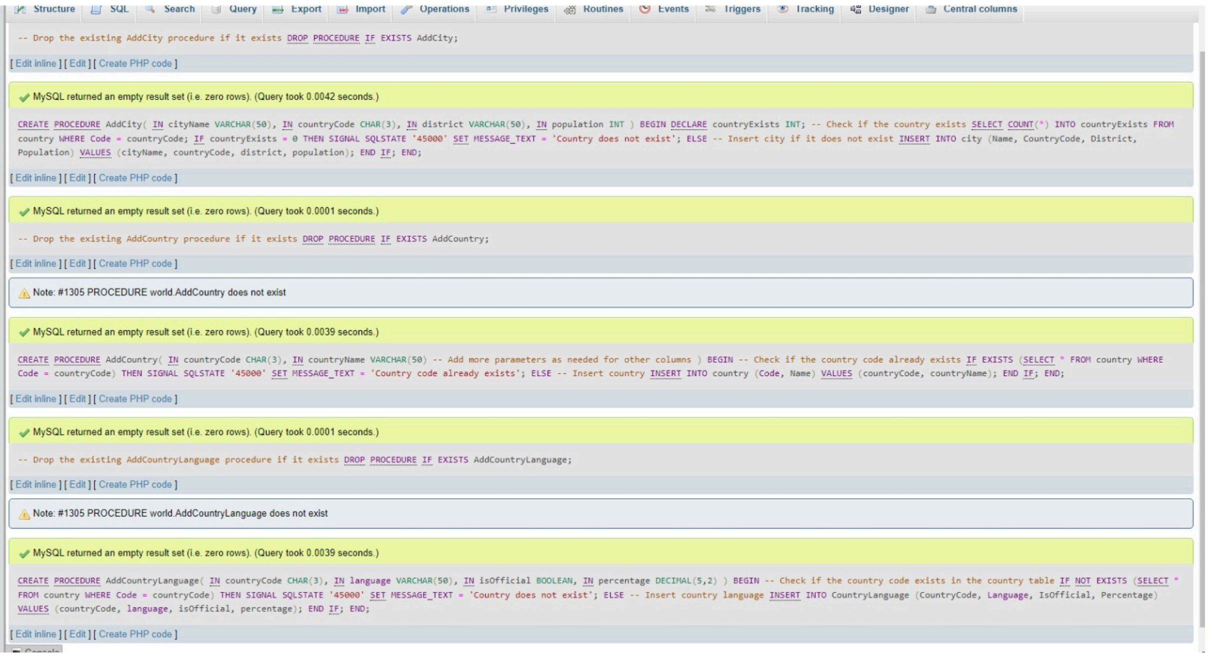
Constraints needed to be placed on the data in the SQL world database file to meet these criteria.

To do this, we added error checking procedures to the SQL file.

Additionally, we provided message text to explain to admins exactly why they can’t enter particular data, for example, “Duplicate country record” displays when somebody in the admin role attempts to enter a country twice.

We added SQL statements such as the following…

```
DELIMITER $$
CREATE TRIGGER `PreventDuplicateCountry` BEFORE INSERT ON `country` FOR EACH
ROW BEGIN
    DECLARE country_count INT;
    SET country_count = (SELECT COUNT(*) FROM country WHERE Code = NEW.Code);
    IF country_count > 0 THEN
        SIGNAL SQLSTATE '45000' SET MESSAGE_TEXT = 'Duplicate country
record';
    END IF;
END
$$
DELIMITER ;
```



```
-- Drop the existing AddCity procedure if it exists DROP PROCEDURE IF EXISTS AddCity;
[ Edit inline ] [ Edit ] [ Create PHP code ]

✓ MySQL returned an empty result set (i.e. zero rows). (Query took 0.0042 seconds.)
CREATE PROCEDURE AddCity( IN cityName VARCHAR(50), IN countryCode CHAR(3), IN district VARCHAR(50), IN population INT ) BEGIN DECLARE countryExists INT; -- Check if the country exists SELECT COUNT(*) INTO countryExists FROM country WHERE Code = countryCode; IF countryExists = 0 THEN SIGNAL SQLSTATE '45000' SET MESSAGE_TEXT = 'Country does not exist'; ELSE -- Insert city if it does not exist INSERT INTO city (Name, CountryCode, District, Population) VALUES (cityName, countryCode, district, population); END IF; END;
[ Edit inline ] [ Edit ] [ Create PHP code ]

✓ MySQL returned an empty result set (i.e. zero rows). (Query took 0.0001 seconds.)
-- Drop the existing AddCountry procedure if it exists DROP PROCEDURE IF EXISTS AddCountry;
[ Edit inline ] [ Edit ] [ Create PHP code ]

⚠ Note: #1305 PROCEDURE world.AddCountry does not exist

✓ MySQL returned an empty result set (i.e. zero rows). (Query took 0.0039 seconds.)
CREATE PROCEDURE AddCountry( IN countryCode CHAR(3), IN countryName VARCHAR(50) -- Add more parameters as needed for other columns ) BEGIN -- Check if the country code already exists IF EXISTS (SELECT * FROM country WHERE Code = countryCode) THEN SIGNAL SQLSTATE '45000' SET MESSAGE_TEXT = 'Country code already exists'; ELSE -- Insert country INSERT INTO country (Code, Name) VALUES (countryCode, countryName); END IF; END;
[ Edit inline ] [ Edit ] [ Create PHP code ]

✓ MySQL returned an empty result set (i.e. zero rows). (Query took 0.0001 seconds.)
-- Drop the existing AddCountryLanguage procedure if it exists DROP PROCEDURE IF EXISTS AddCountryLanguage;
[ Edit inline ] [ Edit ] [ Create PHP code ]

⚠ Note: #1305 PROCEDURE world.AddCountryLanguage does not exist

✓ MySQL returned an empty result set (i.e. zero rows). (Query took 0.0039 seconds.)
CREATE PROCEDURE AddCountryLanguage( IN countryCode CHAR(3), IN language VARCHAR(50), IN isOfficial BOOLEAN, IN percentage DECIMAL(5,2) ) BEGIN -- Check if the country code exists in the country table IF NOT EXISTS (SELECT * FROM country WHERE Code = countryCode) THEN SIGNAL SQLSTATE '45000' SET MESSAGE_TEXT = 'Country does not exist'; ELSE -- Insert country language INSERT INTO CountryLanguage (CountryCode, Language, IsOfficial, Percentage) VALUES (countryCode, language, isOfficial, percentage); END IF; END;
[ Edit inline ] [ Edit ] [ Create PHP code ]
```

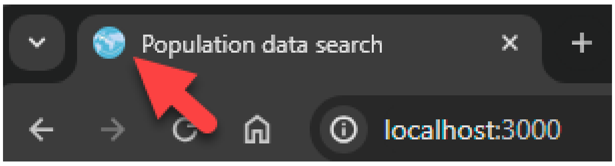
Features

Favicon

For our website, we created a logo to represent world data.

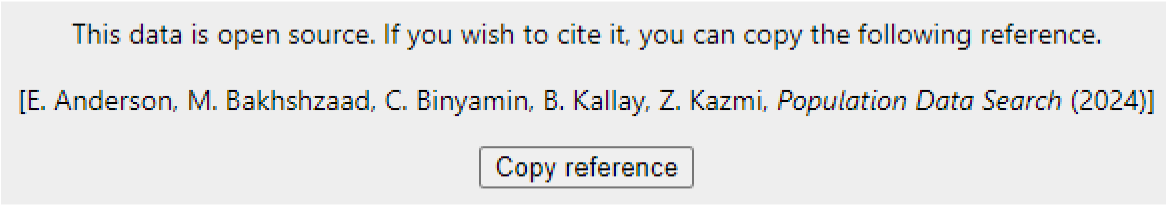
We used copyright free images in our logo.

We then applied it as our favicon to give the website a more professional feel.



Quick reference/citation

One feature we're happy with is our 'quick citation' feature. We provide an IEEE style reference, with a copy to clipboard button to make it simple for users to cite the data in academic work.



This data is open source. If you wish to cite it, you can copy the following reference.

[E. Anderson, M. Bakhshzaad, C. Binyamin, B. Kallay, Z. Kazmi, *Population Data Search* (2024)]

What we would do better next time

Privileges from being logged in

Creating an account at the moment doesn't provide any more access to the data tables. We aimed originally to give away access to meet the acceptance criteria "Given that I am a citizen of a country, when I visit the website and I'm not logged in, then I should be able to see my country's stats 'for free' (in the sense that I'm not yet logged in)."

Graphs

Graphs cannot be created dynamically from population data.

We aimed originally to give away access to meet the acceptance criteria “Given that I need to analyse trends, when I retrieve data from the database, then there should be an option to present data in a format I can use for my analysis, such as tables or graphs.”

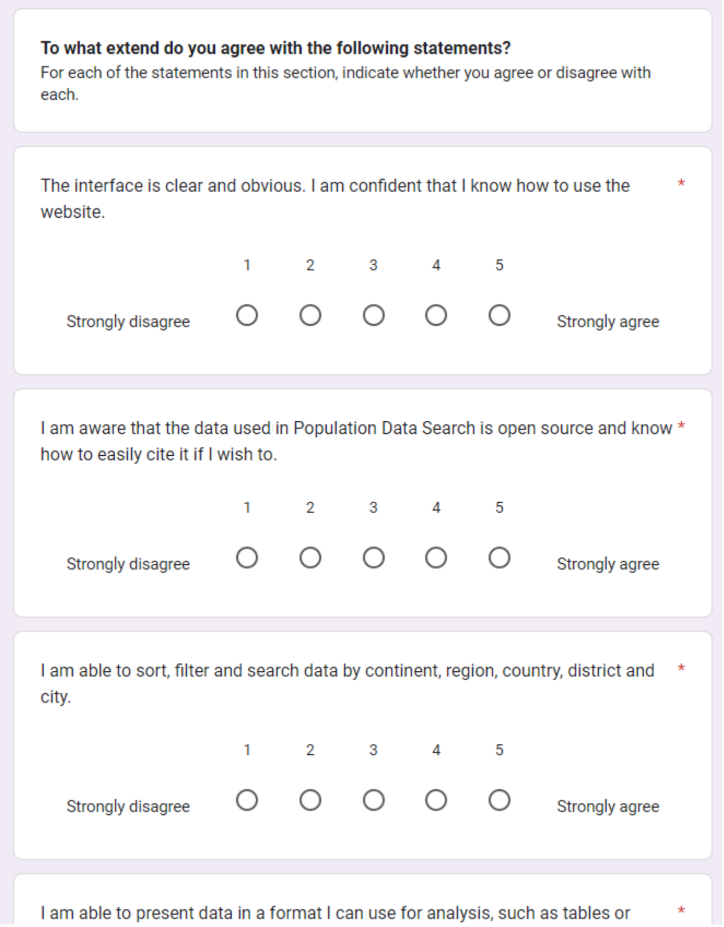
Visual comparison of data

There’s no way to visually compare data.

We aimed originally to give away access to meet the acceptance criteria “Given that I need to compare trends of mine and other countries to inform decision making, when I analyse population data, then the database should provide tools for comparing and contrasting trends between different countries and regions.”

Use our questionnaire to get user feedback

We created a user feedback questionnaire with the intention of obtaining useful feedback for how functional our website is.



To what extend do you agree with the following statements?
For each of the statements in this section, indicate whether you agree or disagree with each.

The interface is clear and obvious. I am confident that I know how to use the website. *

	1	2	3	4	5	
Strongly disagree	○	○	○	○	○	Strongly agree

I am aware that the data used in Population Data Search is open source and know how to easily cite it if I wish to. *

	1	2	3	4	5	
Strongly disagree	○	○	○	○	○	Strongly agree

I am able to sort, filter and search data by continent, region, country, district and city. *

	1	2	3	4	5	
Strongly disagree	○	○	○	○	○	Strongly agree

I am able to present data in a format I can use for analysis, such as tables or *

Unfortunately, we were unable to send this out in time because we faced problems along the way.

In future we'd want to manage our time better and use the form we'd created to give ourselves points to focus on and from this improve our website.

Member's contributions

Each Member of the team was given tasks to apply to the coursework, based on time and skill. Attempts were made to ensure everyone had an equal amount of work to do within the time given. Many tasks were not done due to miscommunication, timing and personal private matters. Considering that our group merged at the sprint 1 presentation, partial data on who did what was lost and unfair to document due to the sudden nature of the merge.

Betty

For sprint 3, Betty managed to pull dynamic content from database to templates via PUG. This was necessary as other members of the team were quite unsure on how to start this task. She also managed to pull content generated HTML pages from PUG which helped progress to the next sprint.

In sprint 4, the website wireframes should be quite similar if not the same to the website. Betty went through the designs and communicated her thoughts about the differences between the two. When there was issues on the database connecting using execute, Betty attempted to fix with trial and error, while unsuccessful at that point, efforts were made.

Cat

For sprint 3, With Evie, helped implement the database to the templates via the PUG files. Again with Evie, database implementation with pre-filled data was successfully attempted which was then followed by merging the conflicts from the develop branch to the main branch. With the help of everyone else combined, a name for the website was decided.

For sprint 4, I partially checked the website wireframes and ensured and questioned if they matched the design and styling. There were conclusions made and Betty noted the notes into a .doc file in the documents folder on the Git repository. A logo was also created and was sent to Evie for quality checks. Cat, took part in writing, checking and creating the report.

Evie

For sprint 3, Evie implemented the database with pre-filled data and moved the database creation logic files to the services folder in the Git repository. Evie also had the main idea of the name for the website database.

For sprint 4, Evie had a lot to do in sprint 4 and managed to do the majority of her tasks and help with others as well. She created a 'quick citation' referencing feature on the website which gives a user the ability to cite the source at the home page. The styling process was majorly flawed, and Evie did her best to correct the errors within the time she was given. A user feedback questionnaire was created to help the team deal with site errors and handle advice and criticism to aid the database. Evie checked the logo after Cat and used it as a favicon. All data was constantly ensured that it was up to date. Evie added a login and register system feature to the navbar to help users create a login and register if no login is found. The code was roughly commented to ensure readability as well as massively improved the styling throughout the site. And lastly, the report, which was helped with Cat.

Micail

For sprint 3, Micail updated all user stories and then created a task board in the documents folder , as well as helped the team pick a group name.

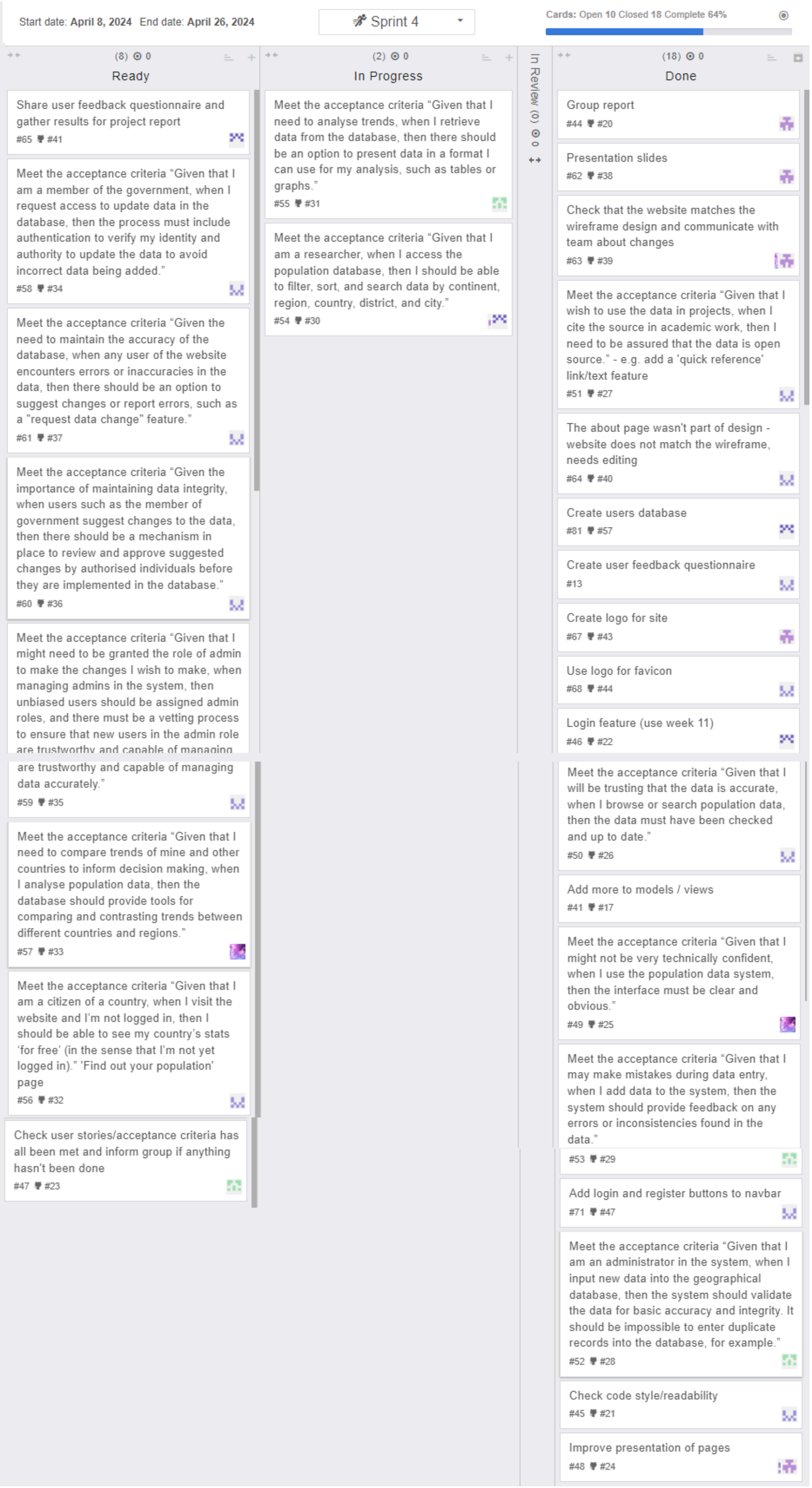
For sprint 4, Micail ensured that When data is added to the system, the system should give feedback on any errors in the data. This was tested and was considered successful. When new info is entered the system should validate the data for basic accuracy meaning it should be impossible to enter duplicates. A graph on population was also attempted but it didn't go successfully so was left out.

Zee

For sprint 3, Zee managed to perform many tests running such as circle CI and plenty unit tests, which determined the usability of the site at the time of testing. As well as everyone else, a name was picked and decided on.

Foer sprint 4, using week 11, Zee created a login and register feature to aid in adding more users for the site. Last minute styling was applied to the site, fixing smaller issues.
Partial final styling. Styling on home page like fixing the countries table and sorting the redundant database,.was also applied. Zee also did a lot of unit testing for sprint 4, and fixed the main issue - establishing a database connection.

Final state of the Kanban board



Start date: **April 8, 2024** End date: **April 26, 2024**

Sprint 4

Cards: Open 10 Closed 18 Complete 64%

(8) ⊙ 0 Ready

- Share user feedback questionnaire and gather results for project report
  #65 #41
- Meet the acceptance criteria “Given that I am a member of the government, when I request access to update data in the database, then the process must include authentication to verify my identity and authority to update the data to avoid incorrect data being added.”
  #58 #34
- Meet the acceptance criteria “Given the need to maintain the accuracy of the database, when any user of the website encounters errors or inaccuracies in the data, then there should be an option to suggest changes or report errors, such as a "request data change" feature.”
  #61 #37
- Meet the acceptance criteria “Given the importance of maintaining data integrity, when users such as the member of government suggest changes to the data, then there should be a mechanism in place to review and approve suggested changes by authorised individuals before they are implemented in the database.”
  #60 #36
- Meet the acceptance criteria “Given that I might need to be granted the role of admin to make the changes I wish to make, when managing admins in the system, then unbiased users should be assigned admin roles, and there must be a vetting process to ensure that new users in the admin role are trustworthy and capable of managing data accurately.”
  #59 #35
- Meet the acceptance criteria “Given that I need to compare trends of mine and other countries to inform decision making, when I analyse population data, then the database should provide tools for comparing and contrasting trends between different countries and regions.”
  #57 #33
- Meet the acceptance criteria “Given that I am a citizen of a country, when I visit the website and I'm not logged in, then I should be able to see my country's stats ‘for free’ (in the sense that I'm not yet logged in).” ‘Find out your population’ page
  #56 #32
- Check user stories/acceptance criteria has all been met and inform group if anything hasn't been done
  #47 #23

(2) ⊙ 0 In Progress

- Meet the acceptance criteria “Given that I need to analyse trends, when I retrieve data from the database, then there should be an option to present data in a format I can use for my analysis, such as tables or graphs.”
  #55 #31
- Meet the acceptance criteria “Given that I am a researcher, when I access the population database, then I should be able to filter, sort, and search data by continent, region, country, district, and city.”
  #54 #30

In Review (0) ⊙ 0

(18) ⊙ 0 Done

- Group report
  #44 #20
- Presentation slides
  #62 #38
- Check that the website matches the wireframe design and communicate with team about changes
  #63 #39
- Meet the acceptance criteria “Given that I wish to use the data in projects, when I cite the source in academic work, then I need to be assured that the data is open source.” - e.g. add a 'quick reference' link/text feature
  #51 #27
- The about page wasn't part of design - website does not match the wireframe, needs editing
  #64 #40
- Create users database
  #81 #57
- Create user feedback questionnaire
  #13
- Create logo for site
  #67 #43
- Use logo for favicon
  #68 #44
- Login feature (use week 11)
  #46 #22
- Meet the acceptance criteria “Given that I will be trusting that the data is accurate, when I browse or search population data, then the data must have been checked and up to date.”
  #50 #26
- Add more to models / views
  #41 #17
- Meet the acceptance criteria “Given that I might not be very technically confident, when I use the population data system, then the interface must be clear and obvious.”
  #49 #25
- Meet the acceptance criteria “Given that I may make mistakes during data entry, when I add data to the system, then the system should provide feedback on any errors or inconsistencies found in the data.”
  #53 #29
- Add login and register buttons to navbar
  #71 #47
- Meet the acceptance criteria “Given that I am an administrator in the system, when I input new data into the geographical database, then the system should validate the data for basic accuracy and integrity. It should be impossible to enter duplicate records into the database, for example.”
  #52 #28
- Check code style/readability
  #45 #21
- Improve presentation of pages
  #48 #24