We will be using sqlite3, which is a serverless database. This comes preinstalled on a macbook.

We will also be using mssql since it has the PIVOT keyword while sqlite3 does not. To download mssql on mac, use Docker:
Follow these tutorials :
https://www.youtube.com/watch?v=glxE7w4D8v8
https://www.youtube.com/watch?v=XETohkMN6TQ

We will be operating on a macbook, which means we will have to install dotnet 7.0 from https://dotnet.microsoft.com/en-us/download

If errors persist where it says dotnet command is not found, try
ln -s /usr/local/share/dotnet/x64/dotnet /usr/local/bin/
This creates a symbolic link that allows you to access a file or directory from a different location in the file system without duplicating the actual data. Symbolic links are references to files or directories that act as pointers to the original file or directory.

After that, create a new C# project:
dotnet new console -o app # create new C# project called app

To download the correct packages for C#: Use Nuget Package Manager
Follow: https://www.youtube.com/watch?v=e1GhVs85Rss

After unzipping the file, go to the directory using cd in terminal and
Do dotnet run in terminal to run application

You will see the zip file downloaded from the web, a folder created with the csv extracted from zip, and a database file that contains the table with the csv contents. The db file is compatible with sqlite3.

References:
https://www.youtube.com/watch?v=DAsyjpqhDp4
https://www.youtube.com/watch?v=e1GhVs85Rss
https://dotnetcodr.com/2015/01/28/compressing-and-decompressing-files-with-bzip2-in-net-c/
https://stackoverflow.com/questions/12783007/how-to-decompress-bz2-file-in-c
https://stackoverflow.com/questions/61446785/unzipping-a-gz-file-in-c-sharp-system-io-invalidda taexception-the-archive-en
https://www.youtube.com/watch?v=IyBWmYAV9eY
https://www.youtube.com/watch?v=WrSZrGoPYyc
https://stackoverflow.com/questions/38562777/csv-reading-as-empty
https://www.youtube.com/watch?v=BH9-H-b41Ys&t=80s
https://dotnetcodr.com/2015/01/28/compressing-and-decompressing-files-with-bzip2-in-net-c/
https://stackoverflow.com/questions/45711428/download-file-with-webclient-or-httpclient

https://learn.microsoft.com/en-us/dotnet/api/system.io.compression.zipfile.openread?view=net-7.0
https://stackoverflow.com/questions/38562777/csv-reading-as-empty
https://stackoverflow.com/questions/61949887/how-to-download-directory-from-github
https://stackoverflow.com/questions/53030531/dotnet-command-not-found-in-mac
https://www.youtube.com/watch?v=PciPCh906OM
https://www.youtube.com/watch?v=BH9-H-b41Ys&t=80s

Step 1: Create Table SQL command

```
CREATE TABLE IF NOT EXISTS Users(
                                    UserIndex INT NOT NULL,
                                    UserId VARCHAR(100) PRIMARY KEY NOT NULL,
                                    FirstName VARCHAR(100) NOT NULL,
                                    LastName VARCHAR(100) NOT NULL,
                                    Sex VARCHAR(100) NOT NULL,
                                    Email VARCHAR(100) NOT NULL,
                                    Phone VARCHAR(100) NOT NULL,
                                    DateOfBirth DATE NOT NULL,
                                    JobTitle VARCHAR(100) NOT NULL
                                    );
```

Here are the column names: UserIndex, UserId, FirstName, LastName, Sex, Email, Phone, DateOfBirth, JobTitle

References for SQL Table:
https://stackoverflow.com/questions/14660666/inserting-a-string-with-double-quotes-into-a-table

Step 3: SQL commands

MSSQL:

```
// Step 3 Part 1 MS SQL: USING PIVOT KEYWORD [dbo].[people-1000] is Users Table in the
sqlite expression
SELECT
   'Age By Gender' AS 'Category',
   [Male],
   [Female]
FROM
(
   SELECT
       [Male],
       [Female]
   FROM
```

```
    (
        SELECT
            CASE
                WHEN Sex = 'Male' THEN 'Male'
                WHEN Sex = 'Female' THEN 'Female'
            END AS Category,
            CAST(DATEDIFF(YEAR, Date_Of_birth, '2023-01-01') AS DECIMAL(10, 2)) AS Age
        FROM [dbo].[people-1000]
    ) AS AgeData
    PIVOT
    (
        AVG(Age)
        FOR Category IN ([Male], [Female])
    ) AS PivotTable
) AS Result;

Result:
Category        Male        Female
Age By Gender   56.250988   60.427125


// Step 3 Part 2 MS SQL: [dbo].[people-1000] is Users Table in the sqlite expression
WITH AverageAgeByGender (Category, MaleAverageAge, FemaleAverageAge) AS (
    SELECT
        'Age By Gender' AS 'Category',
        [Male],
        [Female]
    FROM
    (
        SELECT
            [Male],
            [Female]
        FROM
        (
            SELECT
                CASE
                    WHEN Sex = 'Male' THEN 'Male'
                    WHEN Sex = 'Female' THEN 'Female'
                END AS Category,
                CAST(DATEDIFF(YEAR, Date_Of_birth, '2023-01-01') AS DECIMAL(10, 2)) AS
Age
            FROM [dbo].[people-1000]
```

```
      ) AS AgeData
      PIVOT
      (
          AVG(Age)
          FOR Category IN ([Male], [Female])
      ) AS PivotTable
   ) AS Result
),
SexToGenderAvg AS (
   SELECT 'Male' AS Sex, MaleAverageAge AS AverageAge
   FROM AverageAgeByGender
   UNION ALL
   SELECT 'Female' AS Sex, FemaleAverageAge AS AverageAge
   FROM AverageAgeByGender
)
SELECT User_Id, Age, Age - AverageAge AS AgeMinusGenderAverage
FROM (
   SELECT User_Id, DATEDIFF(YEAR, Date_Of_birth, '2023-01-01') AS Age, Sex
   FROM [dbo].[people-1000]
) AS Subquery
LEFT JOIN SexToGenderAvg ON Subquery.Sex = SexToGenderAvg.Sex
```