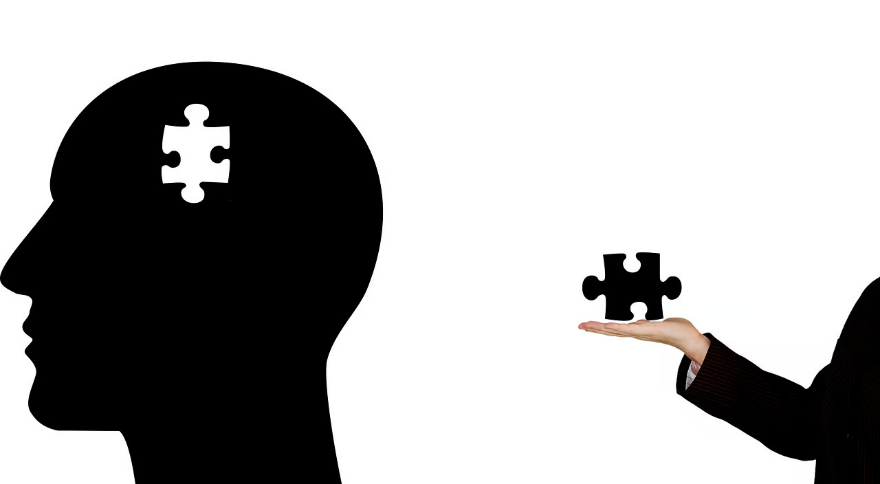
**Mental Health Risk Factors Among International Students in Japan – Data Analysis Project**



GITHUB

<https://github.com/AnesuorDavid>

**Contents**

[Introduction 3](#_Toc203766811)

[1. Defining the Project 3](#_Toc203766812)

[Project Objectives: 3](#_Toc203766813)

[Data Analysis Questions 3](#_Toc203766814)

[2. Knowing the Dataset 3](#_Toc203766815)

[3. Data Uploading in MySQL 5](#_Toc203766816)

[4. Data Cleaning (preprocess) 5](#_Toc203766817)

[5. Data Normalisation 5](#_Toc203766818)

[6. Data Aggregation 6](#_Toc203766819)

[7. Final Reporting and Documentation 9](#_Toc203766820)

# Introduction

This project investigates the mental health challenges faced by students in a multicultural university setting, with a specific focus on international students. Using survey data collected in 2018 from Ritsumeikan Asia Pacific University (APU) in Japan, the research explores the relationship between social connectedness, acculturative stress, and depressive symptoms. The study highlights the increased vulnerability of international students to mental health issues compared to domestic peers and emphasizes how factors such as cultural adaptation stress and perceived isolation can influence psychological well-being.

The aim of this data project is to preprocess, clean, and analyze the survey responses using MySQL to derive key insights and Tableau to visualise them. By doing so, we aim to support evidence-based approaches to student well-being policies and help shape campus mental health initiatives that are more inclusive and effective for culturally diverse student populations.

# 1. Defining the Project

This project focuses on analyzing mental health factors affecting international students at Ritsumeikan Asia Pacific University (APU) in Japan. Drawing on survey data collected in 2018 and published in 2019, the aim is to understand how **social connectedness** and **acculturative stress** influence **depression symptoms** among a diverse student population.

International students often face unique challenges when studying abroad, including cultural adaptation, language barriers, and social isolation. These factors can significantly increase the risk of mental health issues. This project uses SQL to explore these relationships through structured data analysis, looking for correlations, patterns, and risk factors

## Project Objectives:

* Explore the correlation between social connectedness and depression levels.
* Investigate how acculturative stress affects student mental health.
* Analyze demographic influences such as gender, nationality, and length of stay.
* Support data-driven recommendations for improving campus mental health support systems.

## Data Analysis Questions

1. **Do higher levels of social connectedness correlate with lower depression scores?**
2. **Does increased acculturative stress correlate with higher depression scores?**
3. **Are there significant differences in depression scores based on gender or length of stay?**
4. **Can we build a simple regression model to predict depression score?**

# 2. Knowing the Dataset

The dataset used in this project originates from a 2018 survey conducted among international students at Ritsumeikan Asia Pacific University (APU), Japan. The study was approved by multiple ethics boards and was published in 2019. It investigates how social and cultural factors affect students' mental health, specifically symptoms of depression.

Each row in the dataset represents a student respondent and includes responses to psychological and demographic questions. The survey items are aligned with validated psychological scales for measuring depression, acculturative stress, and social connectedness.

**These are the list of questions:**

**1. Depression (Over the last 2 weeks, how often have you been bothered by the following problems?)**

* Little interest or pleasure in doing things
* Feeling down, depressed, or hopeless
* Trouble falling or staying asleep, or sleeping too much
* Feeling tired or having little energy
* Poor appetite or overeating
* Feeling bad about yourself or that you are a failure or have let yourself or your family down
* Trouble concentrating on things, such as reading the newspaper or watching television
* Moving or speaking so slowly that other people could have noticed. Or the opposite—being so fidgety or restless that you were moving around more than usual
* Thoughts that you would be better off dead or of hurting yourself

**2. Social Connectedness (Choose the answer that shows how much you agree or disagree with each of the following statements)**

* I feel disconnected from the world around me
* Even around people I know, I don’t feel that I really belong
* I feel so distant from people
* I have no sense of togetherness with my peers
* I don’t feel related to anyone
* I catch myself losing all sense of connectedness with society
* Even among my friends, there is no sense of brother/sisterhood
* I don’t feel that I participate with anyone or any group

**3. Acculturative Stress – Perceived Discrimination (Choose the answer that shows how much you agree or disagree with each statement)**

* I am treated differently in social situations
* Others are biased toward me
* Many opportunities are denied to me
* I feel that I receive unequal treatment
* I am denied what I deserve
* I feel that my people are discriminated against
* I am treated differently because of my race

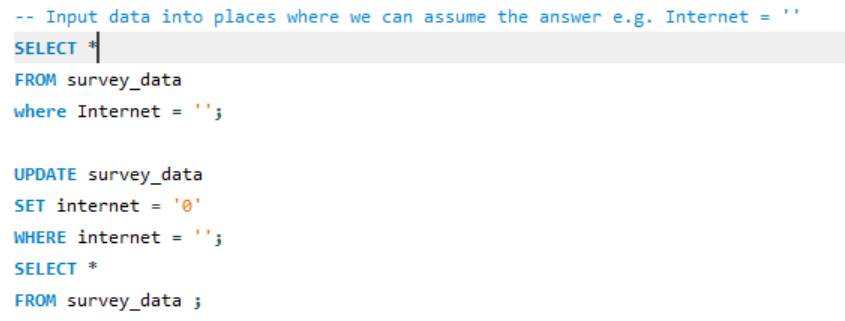
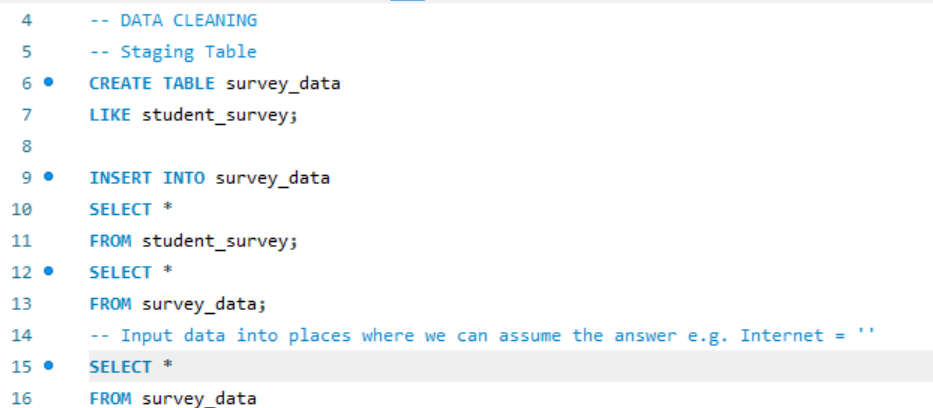
**4. Demographics**

* Do you agree to participate in the questionnaire?
* Are you a domestic student or international student?
* What is your country/region?
* What is your gender?
* What is your academic level?
* What is your age?
* How long have you been in Japan? *(International students)*
* How long have you been at APU? *(Domestic students)*
* What is your Japanese proficiency?
* What is your English proficiency?
* Do you have an intimate partner (e.g., girlfriend, boyfriend, spouse)?
* Do you consider yourself religious?

# 3. Data Uploading in MySQL

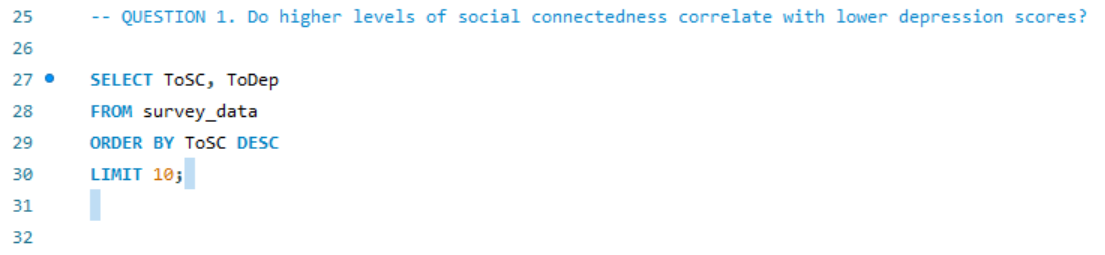
To prepare the dataset for SQL-based analysis, the survey data (provided as a .csv file) was imported into a MySQL database. This process involved creating a schema, defining an appropriate table structure, and ensuring that the data types for each column matched the expected format of the values in the CSV. I also uploaded the data.csv in Tableau as a text file.

# 4. Data Cleaning and Data Normalisation (preprocess)

I created a staging table in order to keep the integrity of original data. I also changed the answers in this question: ‘If you were having a personal or emotional problem, how likely is it that you would seek help from the following people?’ where the internet question is blank to ‘0’ which is the ‘no answer’ input. It has a risk.

# Data Aggregation

**Question 1. Do higher levels of social connectedness correlate with lower depression scores?**

Firstly we have to identify the data that we are meant to analyse. Social correctness is identified as Total Social Correctness score or as seen in data.csv, or ToSC. The total depression score is recorded as ToDep. By selecting and limiting the top ten ToSC and bottom 10 ToDep we can check there is a distribution of data.

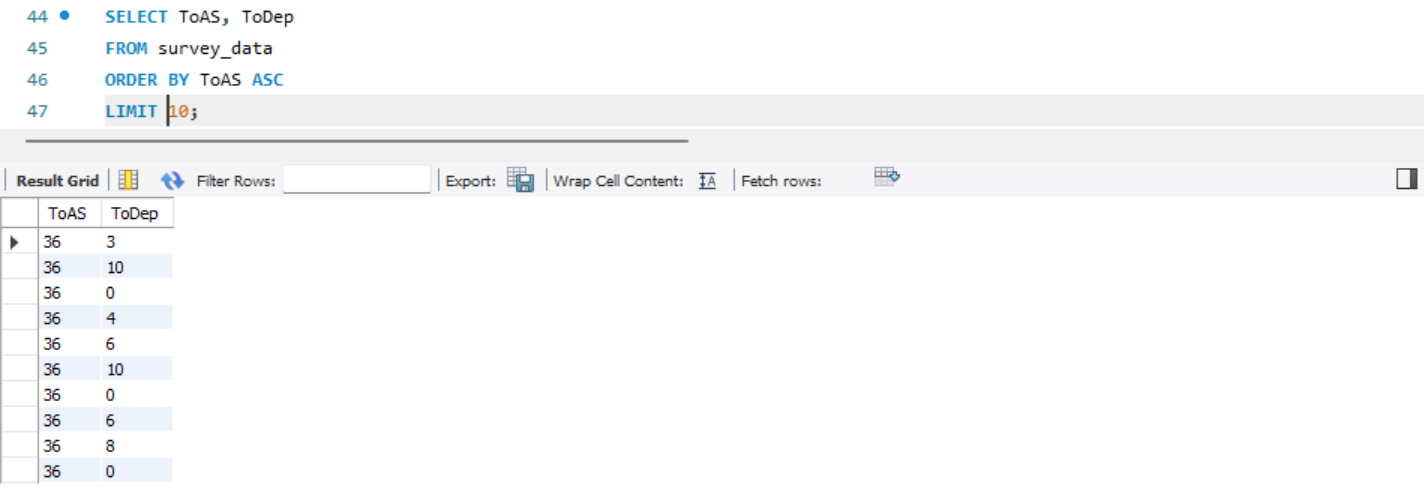
|  |
| --- |
|  |

I then transferred the data to Tableau to analyse it where I placed ToSC on the column and ToDep on the rows. Tableau software was then used to draw the trend line and calculate R-squared and P value.

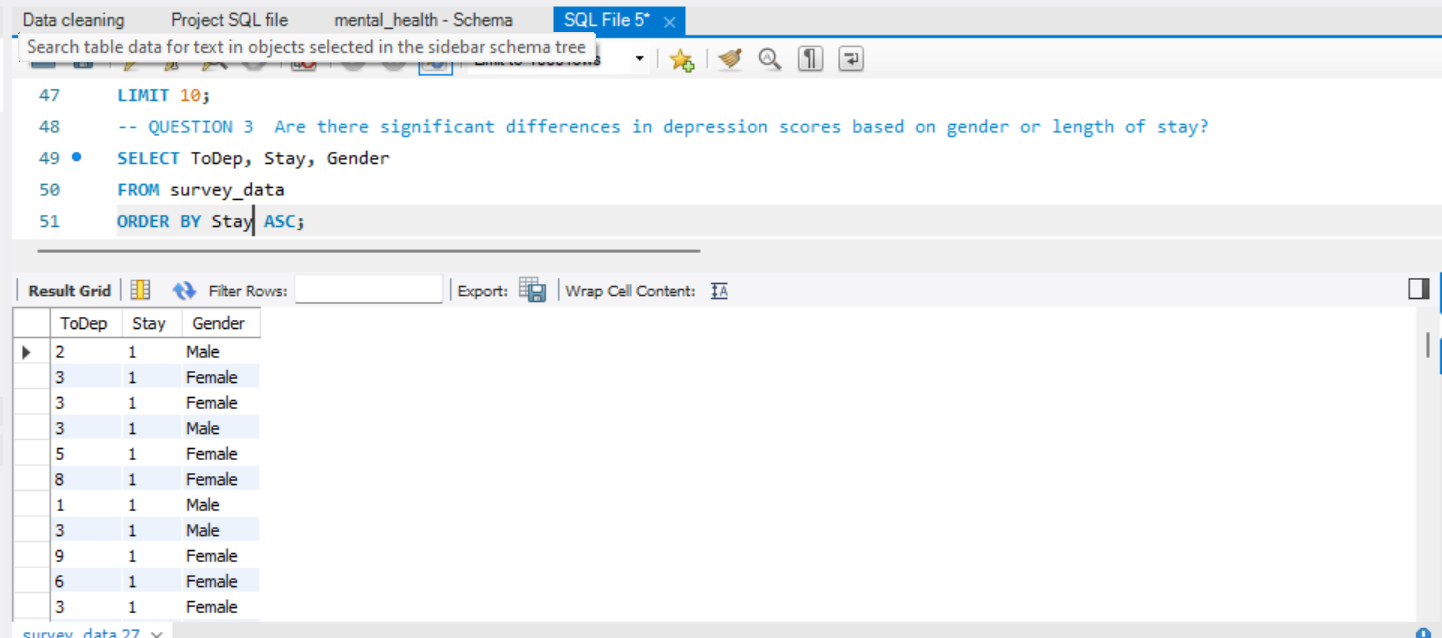
**Question 2. Does increased acculturative stress correlate with higher depression scores?**

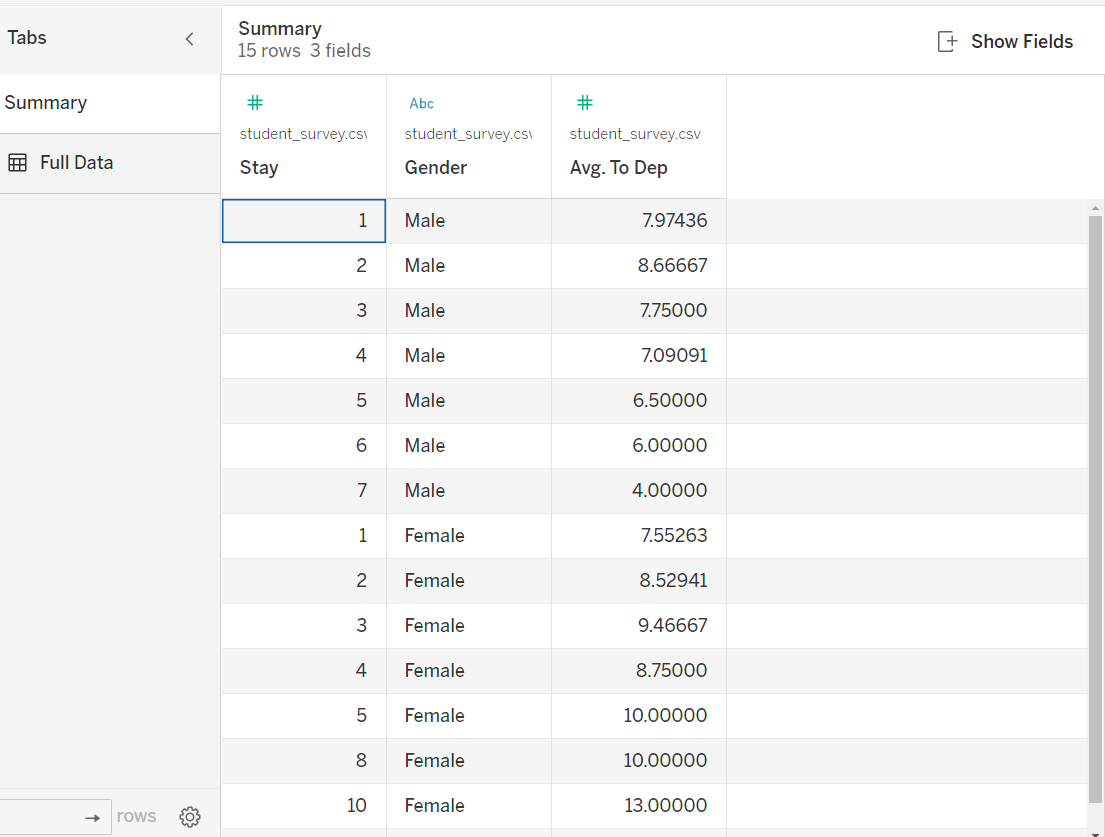
****To do this we must identify the acculturative stress values. These are defined by the ToAS column. Without visualising we can aggregate the values in MySQL to see the general comparison between the 2 columns, one containing ToAS and the other containing ToDep.

The data shows that higher **acculturative** stress actually results in generally higher rates of depression. In contrast, lower rates of **acculturative** results in higher rates of depression

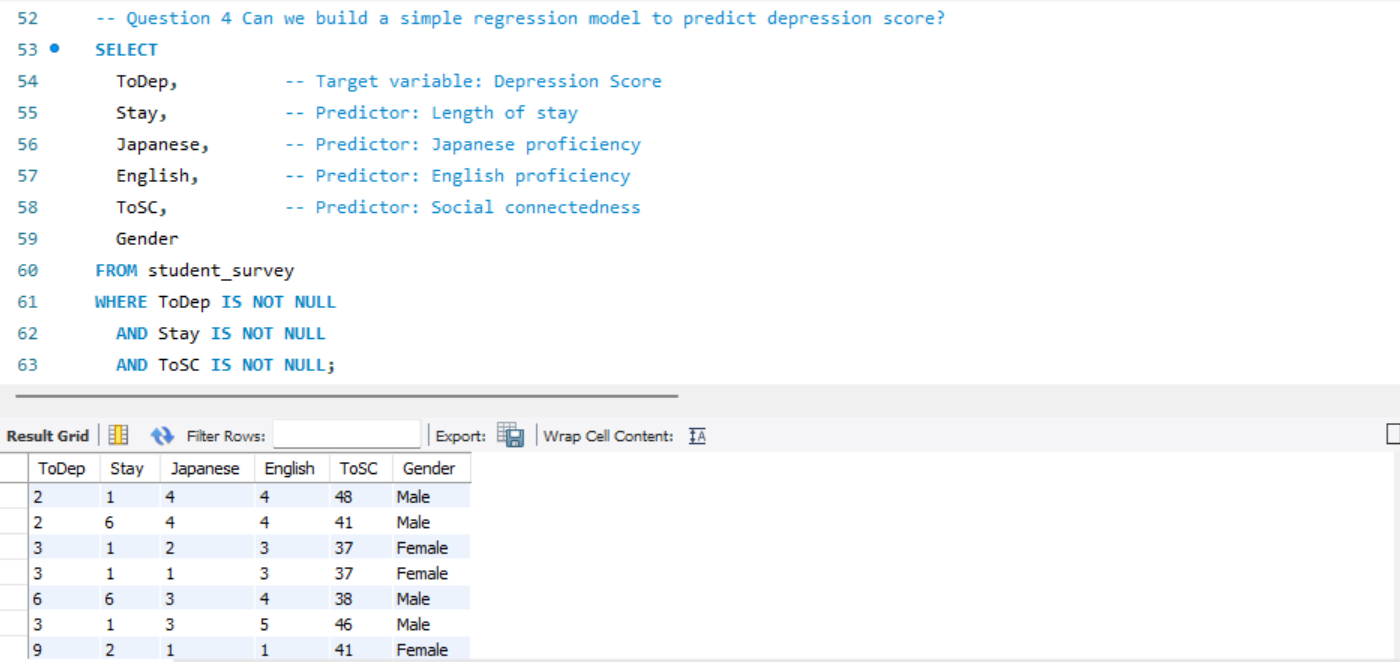


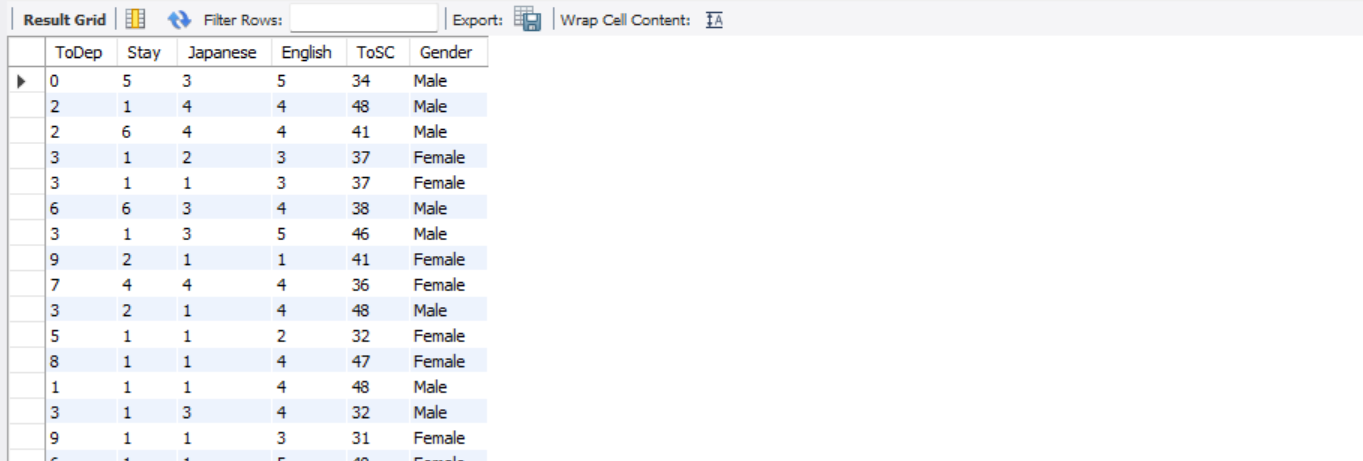
**Question 3. *Are there significant differences in depression scores based on gender or length of stay?***

In order to highlight the differences in Depression scores based on gender and/or length we must use 3 values; Total Depression Score, Gender, Stay. They are to be aggregated for comparison, then analyzed on a visualization in Tableau.

In Tableau I plotted Average ToDep against Length of stay.

**Question 4. *Can we build a simple regression model to predict depression score?***

I believe it can be done but, as a visual one only in Tableau. Excel To do this, we’ll first aggregate the data to take a look at what we have.



After seeing the data is clean and simple to use, I used Tableau to do a single predictor. This can theoretically be repeated. The point of making this model **higher social connectedness scores** predict **lower depression scores** using **a regression line**, and possibly compare this pattern **by gender**.

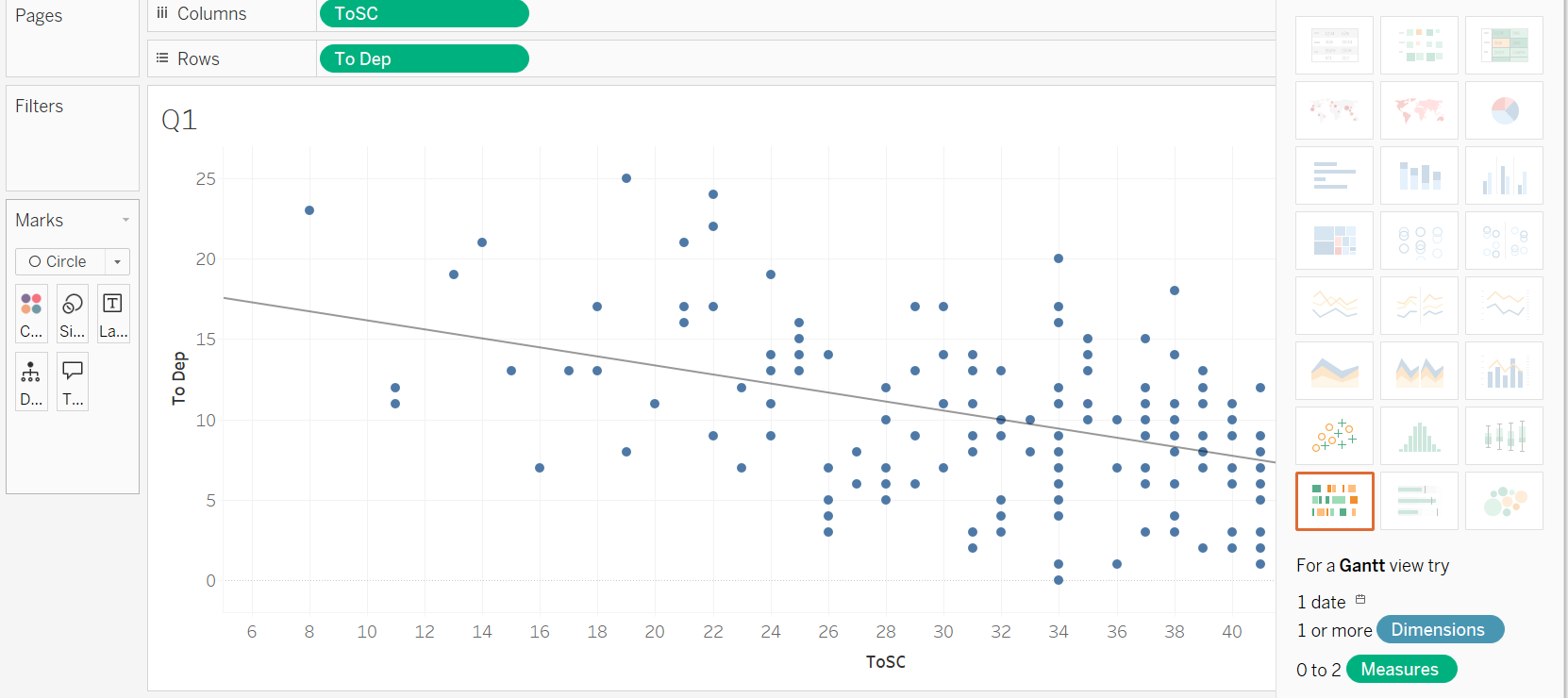
# 7. Final Reporting and Documentation

**Question 1. *Do higher levels of social connectedness correlate with lower depression scores?***

To answer this question the respondents of the survey where asked the following series of questions

* I feel disconnected from the world around me
* Even around people I know, I don’t feel that I really belong
* I feel so distant from people
* I have no sense of togetherness with my peers
* I don’t feel related to anyone
* I catch myself losing all sense of connectedness with society
* Even among my friends, there is no sense of brother/sisterhood
* I don’t feel that I participate with anyone or any group

Without visualizing the data and by Limiting the top ten ToSC and bottom 10 ToSC we saw there is a trend is where high ToSC correlates to low ToDep, and low ToSC correlates to higher ToDep. This is confirmed by the Scatter plot diagram. It is a downward slope trend suggesting that higher social connectedness leads to lower levels of depression.

It must be noted that the R2 value I >.3 and the p value is <.0001.

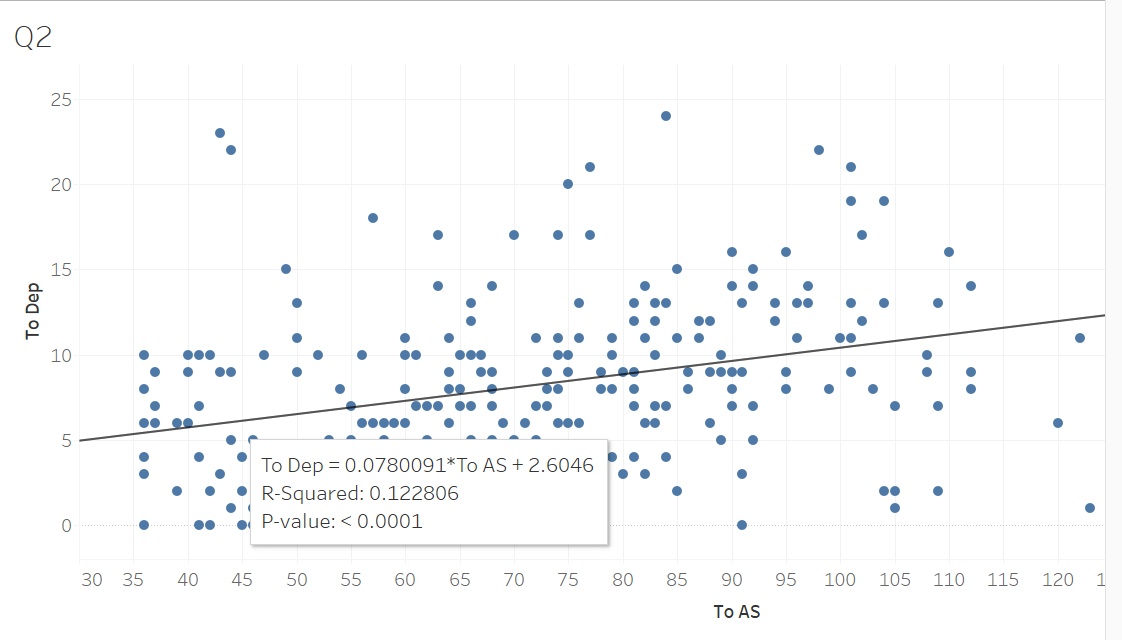
The data then shows, how be it inconclusively that the less connected someone is the more likely they are to be depressed or score higher depression scores. It shoud be noted that the data is scattered and so there maybe variables that can be accounted for like peoples preferences and personality.

**Question 2. Does increased acculturative stress correlate with higher depression scores**?

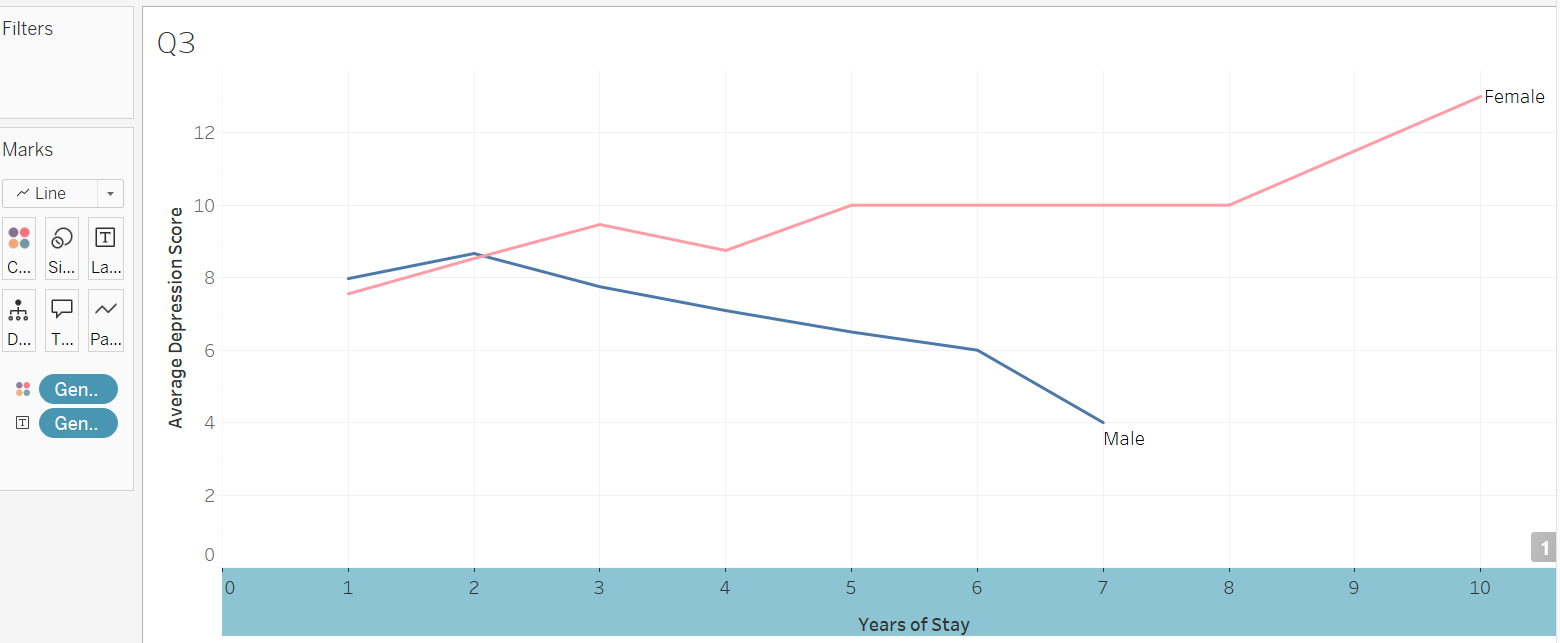
To answer this question the respondents of the survey where asked the following series of questions

* I am treated differently in social situations
* Others are biased toward me
* Many opportunities are denied to me
* I feel that I receive unequal treatment
* I am denied what I deserve
* I feel that my people are discriminated against
* I am treated differently because of my race

They all gave a score for this and the value was added together to make Total for both values – Total Depression Score (ToDep) and Total Acculturative Stress score (ToAS).

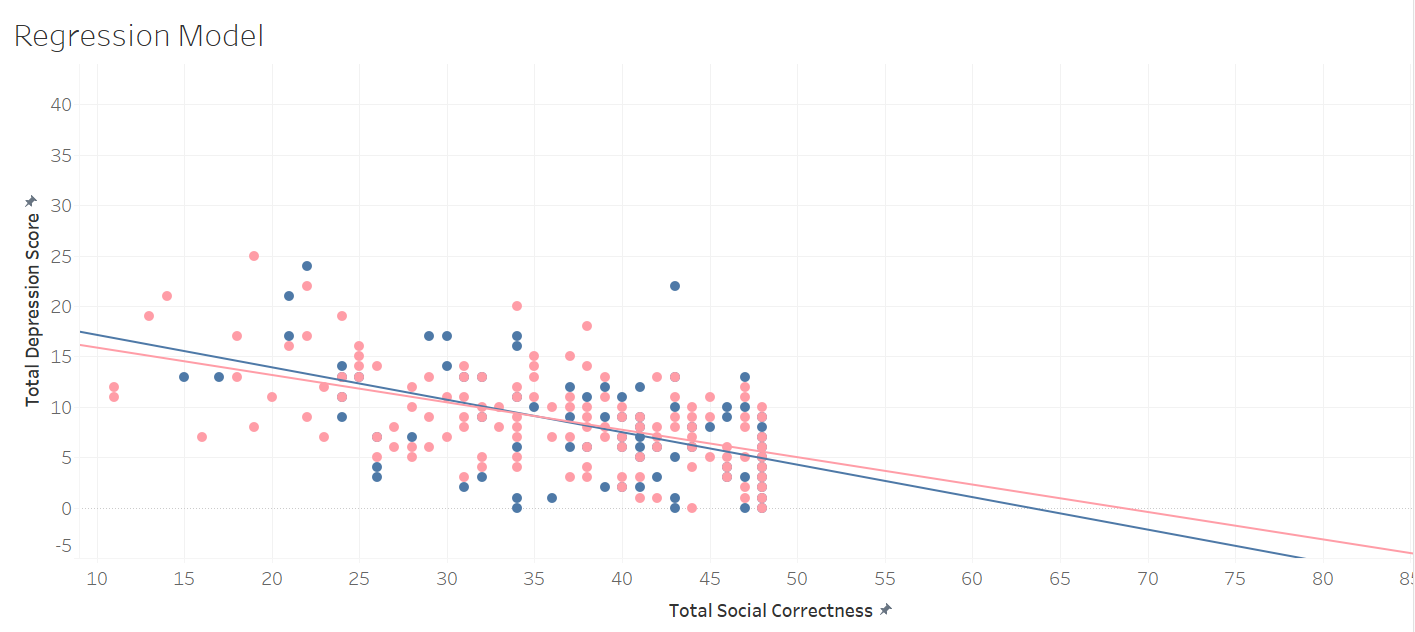
The above scatter plotline shows the relationship between Total Depression Score and Total Acculturative Stress score. The plotline slopes upwards however small the degree is and the R-squared is . a122806nd the P-value is <.0001.

Based on this scatter plot diagram and its sloping upwards trendline it seems that there is a low correlation between **acculturative stress** and higher depression scores. However as is seen on the above image the values are concentrated on the low depression side regardless of Acculturative stress levels. Acculturative stress refers to the psychological and physical strain experienced by individuals when adapting to a new culture. The data above implies that although acculturative stress may be able to affect your level of depression, it may not matter too much how locals treat the students. This could be for a myriad of reasons, some being the awareness that students have to the reclusive nature Japanese are known to have towards foreigners, and even amongst themselves.

**Question 3. *Are there significant differences in depression scores based on gender or length of stay?***

When looking at the graph above, it is clear to see the trends in the graph. By plotting the average Depression score against the years of stay, we can see the trends in the data. Observing the male average depression score, it seems men have a higher depression score in their first 2 years with a steep decrease in depression as their stay increases. However, it seems they stay is also much shorter in Japan than their female counter parts. Women who study in Japan, in contrast, initially have lower average depression scores. However, theirs steadily increase as time continues. This could be a result of women being more social, or, that is needing more social interactions to decrease the risk of falling into depression. It could also be an adverse effect of the loneliness epidemic in Japan, and the lack of dating options in Japan which has resulted in a decline in population. There are a varying number of reasons why this could be so, and those variables make it difficult for one to pinpoint an accurate root of this cause.

**Question 4. *Can we build a simple regression model to predict depression score?***

******

To explore the predictive capability of the dataset, a simple regression model was constructed to assess whether depression scores could be forecasted using one or more related variables—specifically social connectedness (ToSC). The hypothesis was that a strong inverse relationship between social connectedness and depression may allow for a rudimentary predictive model.

The regression was built using Tableau's built-in trend line features. Social connectedness (ToSC) was placed on the X-axis, and total depression score (ToDep) was plotted along the Y-axis. Tableau then generated a regression line that visually represented the strength and direction of the relationship. The line trended downward, indicating that as social connectedness increased, depression scores decreased—just as observed in earlier scatter plots.

Statistical outputs, such as the R-squared and p-value, confirmed the relationship’s significance. The R-squared value explained a moderate portion of the variation in depression based on social connectedness, while the p-value (< 0.0001) affirmed that the relationship was not due to random chance. However, as this was a simple linear regression, it is important to note that it only considered one variable at a time. In future research, multiple regression models incorporating several predictors like gender, stay duration, or acculturative stress could offer more accuracy.

Despite its limitations, this visual regression served as a helpful tool to better understand how certain factors might influence student mental health and where intervention might be most impactful.

# Conclusion

This project demonstrates the real-world value of data analysis in addressing student mental health in multicultural academic settings. By leveraging survey data from Ritsumeikan Asia Pacific University, the study uncovered key trends and relationships between depression, social connectedness, and acculturative stress. The results show that students who feel socially isolated or experience cultural adaptation stress tend to report higher levels of depression. Furthermore, gender and duration of stay in Japan also appear to influence how students experience and report depressive symptoms.

Through the use of MySQL for data preprocessing and Tableau for visualization, we were able to create a structured, interpretable, and insightful view of the challenges faced by international students. Regression analysis confirmed that social connectedness plays a pivotal role in shaping mental health outcomes, potentially guiding future university mental health initiatives toward strengthening social bonds among students.

While limitations remain—particularly in terms of sample size, diversity of predictive variables, and cultural nuance—the findings underscore the importance of creating culturally aware, inclusive support systems for international students. As global education continues to expand, so too must our commitment to understanding and addressing the emotional and psychological needs of diverse student populations.