Peer Connections CRLA Level 3 Project

The Learning Challenges of Students In Different Demographic Groups

By Manh Tuong Nguyen

Table of Contents

Abstract

The goal of this project is to investigate how students from different demographic groups (e.g., year in college, major, gender, or ethnicity) vary in their unique learning challenges in their academic journey. By understanding these variations, the project aims to provide insights for developing tailored teaching strategies and support systems that focus on the specific needs of each group.

Background

As a Learning Assistant, I am partnered with a whole class throughout a semester. From my observations in 4 semesters, I noticed that each student group has different academic issues. For
example, fresh international students usually face language and cultural barriers as they just come to a new country. For minor ethnicity groups, they may find it difficult to collaborate with their classmates due to imposter syndrome. Or, freshmen will be more likely to struggle with the new way of learning in college. If we can synthesize their struggle, it will be easier for Learning Assistants to assist them with preparedness and also create a better learning environment.

Methodology

Planning

1. Data Collection: In this step, I will collect and format the data collected from the Learning Assistant team. Then, I will convert natural language data into numeric before beginning the analysis.
2. Exploratory Data Analysis: We will identify the differences between groups to find the best segmentation method. Then, we will identify the general trends among all students for each group. Finally, we will extract the critical challenges that all students face.
3. Statistical Analysis: In this step, we will investigate the correlation between the segmented group and the challenges that they face. We will find the frequency of each

issue in each demographic group. Moreover, since we have time data, we can explore whether the problem is seasonal or permanent using time series analysis.
4. Evaluation: In this step, I may need to interview Ashley, Lisa, or any pro-staff to make recommendations for the demographic group's challenges.
5. Conclusion: Finalize details, write and submit the final report.

Data Collection

The collected data is from the PRE and POST surveys distributed to all the students in classes served by the Learning Assistant team. It consists of 2302 entries with 20 variables (excluded irrelevant features):

Quantitative Variables	**Qualitative Variables**
Effective Study Strategies	Level (Year at SJSU)
Learning Understanding	Is Transfer Student
Breaking Complex Ideas	First Generation College
Time Management	International Student
Self Motivation	PELL Grant Eligible
Access Campus Resources	Primary Ethnic Racial Group
Comfort Asking Help	Belonging SJSU
	Instructors Care
	Weekly Work Hours

Pre-Processing and Exploratory Data Analysis

Among those variables, the quantitative variables will be used to tell the struggles of students. They are all measured by confidence level on a scale of 5. On the other hand, qualitative variables will be utilized to segment the demographic groups.

Since there are 9 categorical variables, using all the combinations will be very hectic. So, let's do a pre-processing feature selection to discard insignificant groups.

1/ Removing null (missing) values:

International Students, PELL Grant Eligible, and Primary Ethnic Racial Groups are the groups that have missing values. Particularly, International Students have 2281 entries, PELL Grant Eligible has 1387 entries, and Primary Ethnic Racial Group has 1734 entries. For demographic data, it is not safe to perform imputation as it will make the data biased. Furthermore, if performing techniques like K-Nearest-Neighbors to use the quantitative data to predict, we are committing a circular reasoning fallacy since we are using the objective result to prove it (i.e.: assume the confidence levels of groups to predict the demographic group). Hence, we should use removing entries with null values. With that, PELL Grant Eligible should be removed as it has the least availability. Primary Ethnic Racial Groups and International Students can also be removed but it is essential for segmentation so we should not ignore them. Also, we should ignore the data with invaluable segmentation factors such as "Prefer not to respond". Hence, the dataset will be shortened to 1607 rows x 18 features.

2/ Distribution of segmentation factors (qualitative variables):

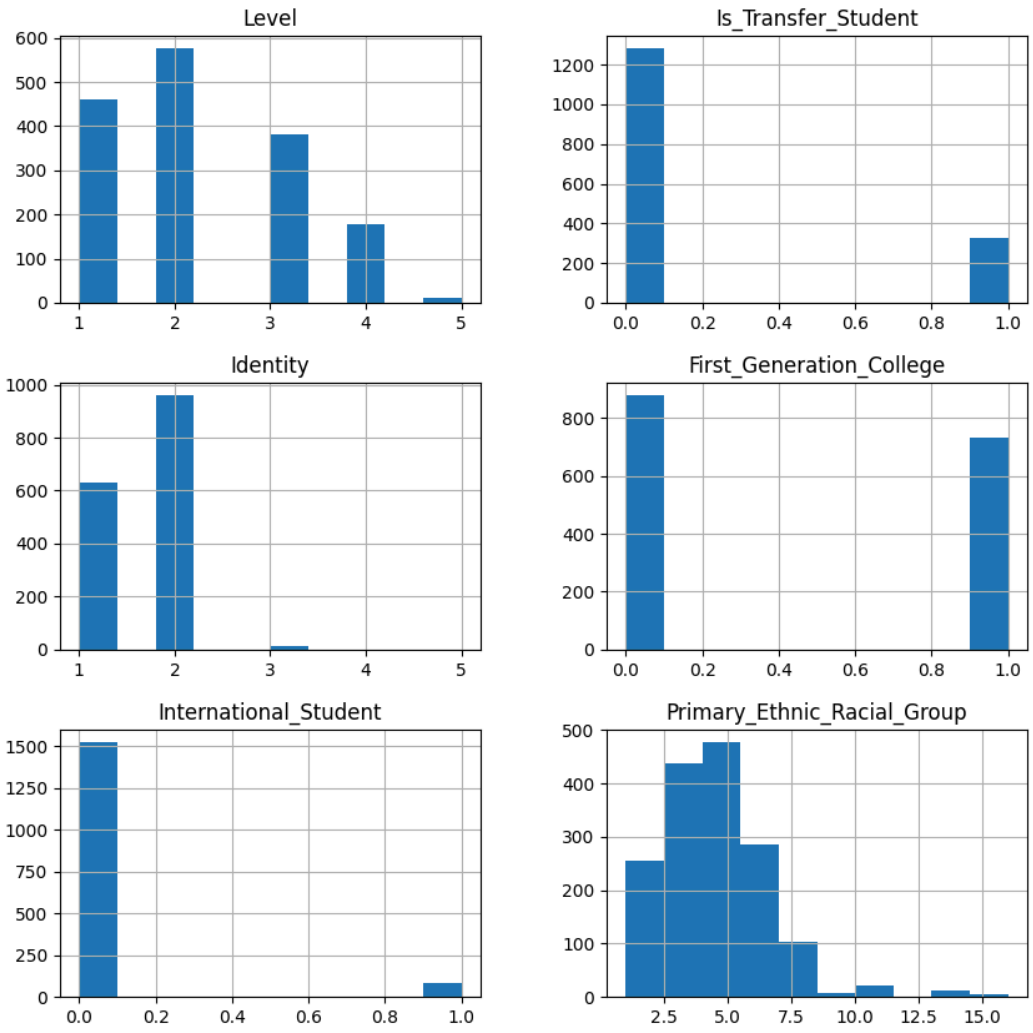

Figure 3.1.1a: Histograms of qualitative groups

From the histograms, we can see that students are international students, Central Asian (Tajikistan, Uzbekistan Kazakhstan Turkmenistan, Kyrgyzstan), Indigenous Mexican, Pacific Islander (Samoa, Tahiti, Tonga, Guam, Fiji, Papua New Guinea, Solomon Islands, etc.), Native Hawaiian, North African (Morocco, Algeria, Tunisia, Libya, Egypt, Sudan, Mauritania, Mali, Niger, Chad), Indigenous Central American, Native American/Native Alaskan, Indigenous South American, transgenders, bi-sexuals, other genders, and in graduated school are under-represented in the data. Hence, we should also exclude those entries since the available data are insignificant to draw any

conclusion. Note that even if International Student is an important factor in segmenting demographic groups, if we include them, the conclusion is not accurate it will not represent the whole population. Hence, we need more data from international students to investigate. But for now, we have to remove them.

After this step, our data has 1535 rows and 17 features.

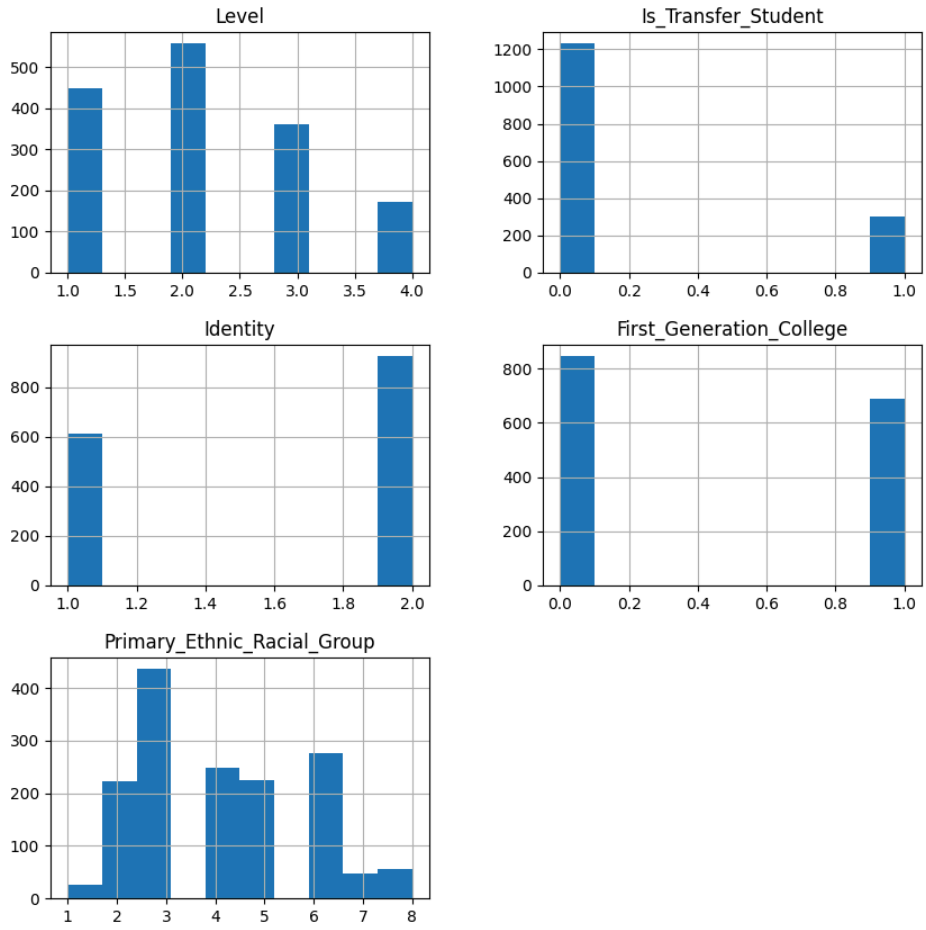

Figure 3.1.1b: Histograms of qualitative groups after the exclusions

Analytical Analysis

As of this step, we have 4 x 2 x 2 x 2 x 8 = 192 demographic combinations. Performing analytical analysis on this many groups is possible but time-consuming. So, the plan is to compare each categorical group to each other, which is 5 tests.

Now, the learning ability-related confidence metrics include:

Effective Study Strategies
Learning Understanding
Breaking Complex Ideas
Time Management
Self Motivation
Access Campus Resources
Comfort Asking Help

The main method to compare here is using bar plots to rank the confidence level of each criterion of interest (each quantitative variable). Then, we will pick the group with the lowest rank in each criterion and classify such criteria as the challenges for those groups. Before comparing the using bar plots, we have to first perform the Kruskal-Wallis (KW) test. KW test is a test that compares the mean of each feature (the confidence level category) for each group (racial/ethnic group or racial/ethnic + second categorical group) and justifies if they are different than each other. The assumptions for the KW test are the data is ordinal and independent, which fits all the confidence levels data that we have.

1/ Correlation between the confidence categories:

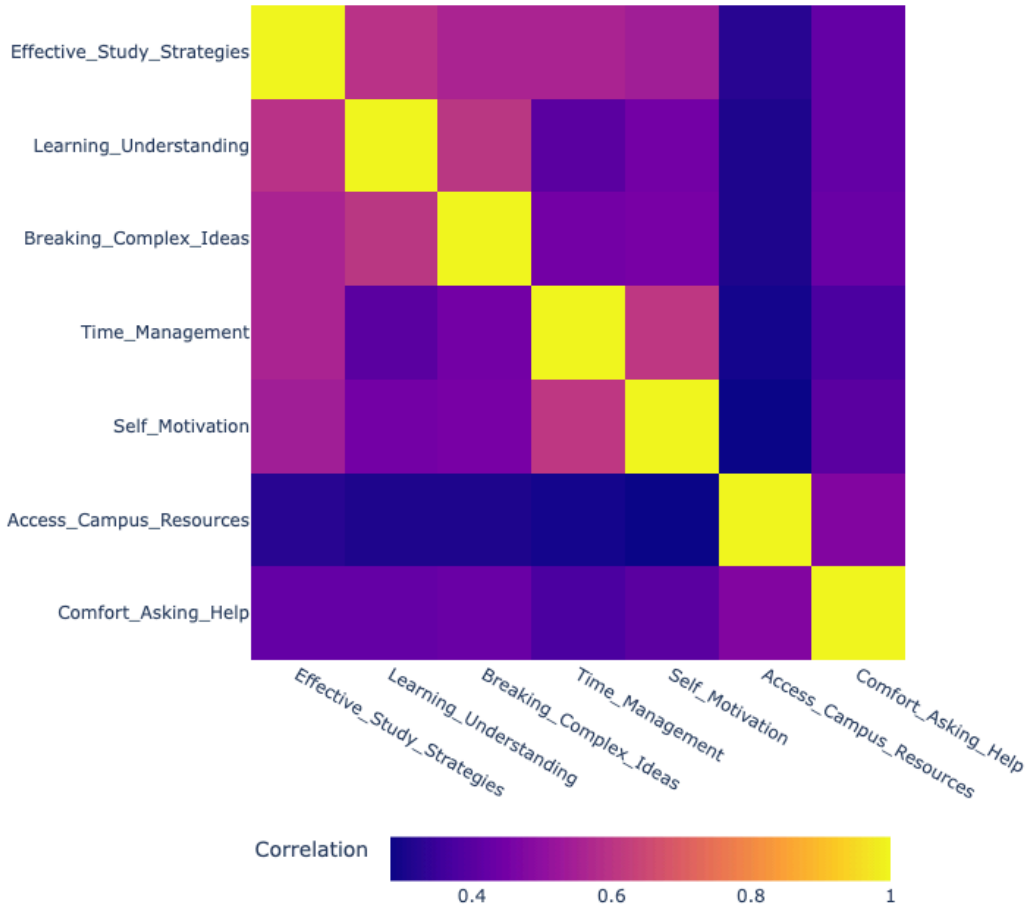

Figure 4.1.1: Heatmap of correlation between confidence categories

Firstly, we want to explore the interactions between the metrics. The technique here is pairwise correlation. The interpretation is that the change (positive or negative) of a variable will make the other changes accordingly. Positive correlation means the second variable will change in the direction of the first variable and negative correlation is the reverse of it. The closer to one (or negative one), the stronger the correlation.

From Figure 2.1, we can see that most of the variables have interactions with each other. However, Access Campus Resources seem to have none to very little impact on other metrics. It only influences the comfort of asking for help from the student, which makes sense as students with high knowledge of resources will know more where to ask the question. Another

insight we can get from this plot is that students who are comfortable asking questions do not improve their other learning abilities. It may be because the questions they usually ask do not help them develop their learning strategies or other educational aspects. A way that I think can improve this is to encourage students to attend workshops that discuss how to ask questions properly in a way that they can learn more than just “Is my question correct?”.

On the other hand, students with high confidence in Effective Study Strategies are very confident in other learning abilities (excluding access to campus resources and comfort in asking for help). It could be explained that those students have high self-motivation and can find the answers themselves. In other words, students with low confidence in effective study strategies have less confidence in other learning abilities. So, the leading mission of Peer Educators is to help students develop good learning strategies.

For other variables, only time management and self-motivation highly correlated to each other. Since self-motivation is a passive ability (depends on mental ability, not trainable), this means that students with good time-management skills will have better self-motivation.

In general, we should focus our analysis on comparing the groups to see which one faces challenges with study strategies.

2/ Comparing confidence levels among 8 ethnic groups:

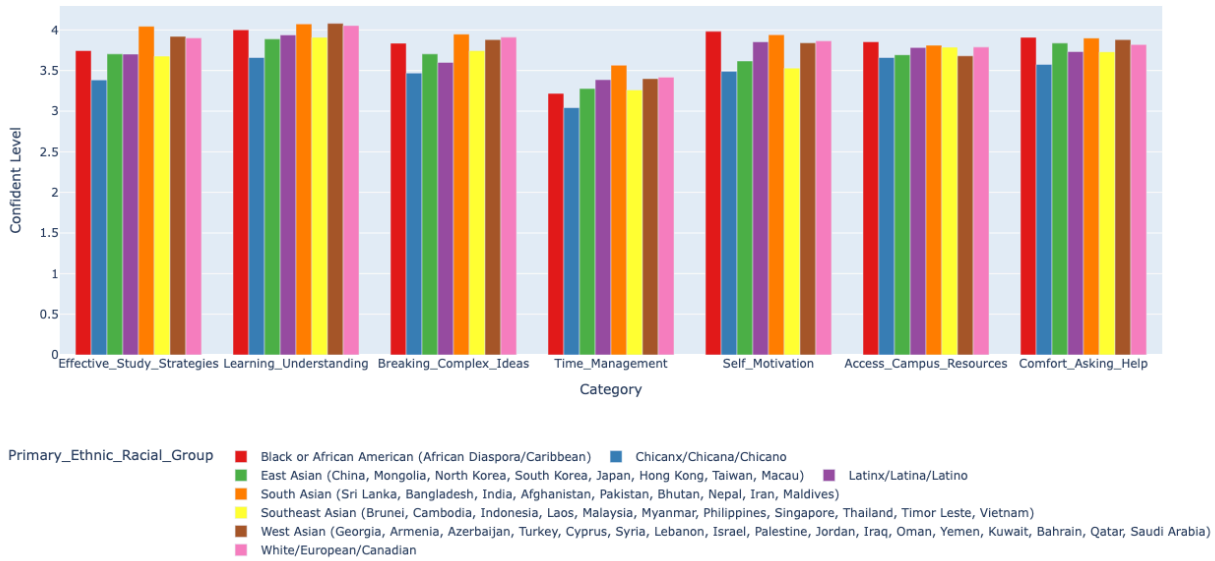

Figure 4.2.1: Confidence Level of Confidence Category by Racial/Ethnic group

Firstly, the time management ability has the lowest confidence level among the groups. The other groups have similar shapes so we will perform KW and Dunn's test to find out which student ethnic group struggles the most with each confidence category.

1/ Hypothesis:

- **Null Hypothesis (H_0)**: There is no significant difference in the median scores of the metric (e.g., "Effective Study Strategies") across racial/ethnic groups.
- **Alternative Hypothesis (H_A)**: There is a significant difference in the median scores of the metric across racial/ethnic groups.

2/ Test statistic and p-values (α=0.05):

Feature	**Test statistic**	**p-values**	**Decision**
Effective_Study_Strategies	58.4848	0.0000	Reject H_0
Learning_Understanding	22.5008	0.0021	Reject H_0
Breaking_Complex_Ideas	39.3256	0.0000	Reject H_0
Time_Management	21.0208	0.0037	Reject H_0
Self_Motivation	53.4236	0.0000	Reject H_0
Access_Campus_Resources	6.4939,	0.4834	Fail to reject H_0
Comfort_Asking_Help	15.4368	0.0308	Reject H_0

Table 4.2.1: Kruskal Wallis Test Result for Ethnic Groups Comparison

3/ Results

For all the metrics, there are significant differences between the ethnic groups. This conclusion means that the mean of the metrics is not identical for all the student ethnic groups. Hence, we can safely proceed to extract the ranking of confidence categories for each ethnic group.

Only for Access Campus Resources, all the groups have similar confidence. So, we can exclude the Access Campus Resources from further exploration with ethnic groups.

4/ Ranking confidence categories for ethnic groups:

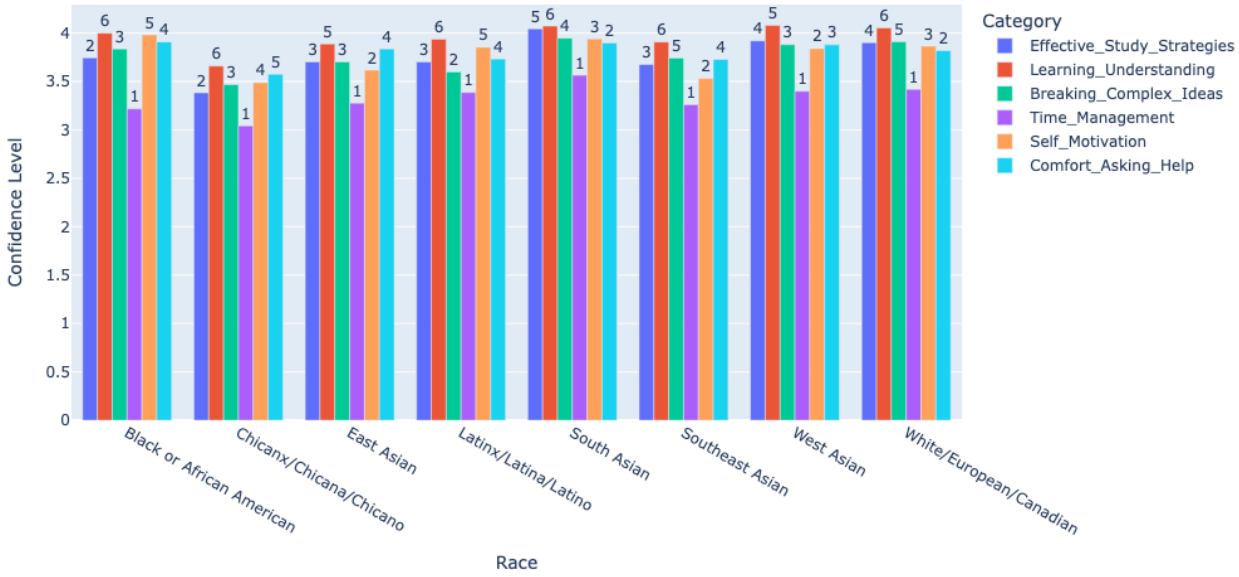

Figure 4.2.2: Confidence Level ranking of each Racial/Ethnic group by Confidence Category

From the plot, we can form a table of ranking of challenges:

Race	**Top 1**	**Top 2**	**Top 3**
Black or African American	Effective Study Strategies	Breaking Complex Ideas	Comfort Asking Help
Chicanx/Chicana/ Chicano	Effective Study Strategies	Breaking Complex Ideas	Self Motivation
East Asian	Self-motivation	Effective Study Strategies / Breaking Complex Ideas	Comfort Asking Help
Latinx/Latina/	Breaking Complex	Effective Study	Comfort Asking Help

Latino	Ideas	Strategies	
South Asian	Comfort Asking Help	Self Motivation	Breaking Complex Ideas
Southeast Asian	Self Motivation	Effective Study Strategies	Comfort Asking Help
West Asian	Self Motivation	Breaking Complex Ideas	Effective Study Strategies
White/European/ Canadian	Comfort Asking Help	Breaking Complex Ideas	Effective Study Strategies

Table 4.2.2: Top 3 Most Challenging ability by Ethnic Group

As we know from Figure 2.1, students in all ethnic groups struggle with time management. Peer Connections did a very good job in making tutorials, training, and workshops about this topic to help students. Come in second is Breaking Complex Ideas with 6 appearances in the ethnic group's top 3 challenging abilities. Third place is Effective Learning Strategies with 5 appearances then Self Motivation at fourth and Comfort Asking Help is last. Among the skills we assess, all student ethnic groups feel confident with their Learning Understanding ability.

3/ Comparing confidence levels among 4-year levels at SJSU

The procedure of comparison will be the same as when we compare ethnic groups. Firstly, we will take a look at what the general trend looks like in those 4-year levels.

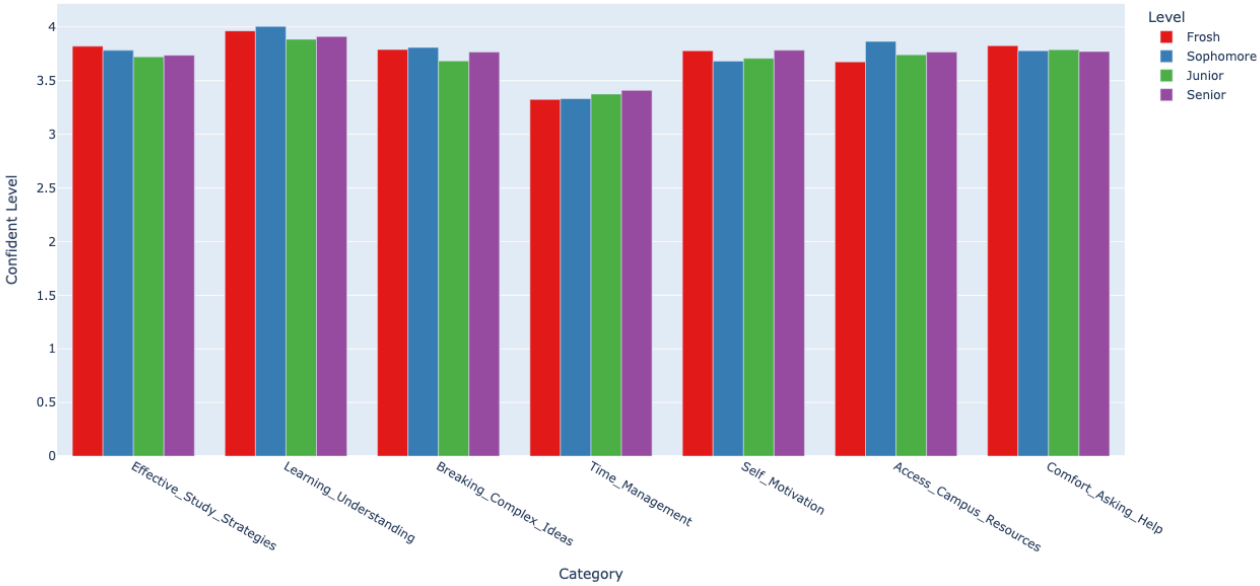

Figure 4.3.1: Confident Level of Confidence Category by Year at SJSU

Because this group has a time-nature in it, we need to incorporate it into our interpretation. However, similar to ethnic groups, Time Management is also the most challenging skill for all levels of students at SJSU. It gets better from time to time as the students start to get used to the college schedule. However, the Effective Learning Strategies and Comfort Asking Help confidence tend to decrease over four years. So, let's perform the KW test and see how all these metrics differ among levels.

1/ Hypothesis:

- **Null Hypothesis (H_0)**: There is no significant difference in the median scores of the metric (e.g., "Effective Study Strategies") across level groups.
- **Alternative Hypothesis (H_A)**: There is a significant difference in the median scores of the metric across level groups.

2/ Test statistic and p-values (α=0.05):

Feature	Test statistic	p-values	Decision
Effective_Study_Strategies	3.5413	0.3154	Fail to reject H_0
Learning_Understanding	5.3232	0.1496	Fail to reject H_0
Breaking_Complex_Ideas	5.6204	0.1316	Fail to reject H_0
Time_Management	0.9165	0.8215	Fail to reject H_0
Self_Motivation	4.0434	0.2568	Fail to reject H_0
Access_Campus_Resources	12.8676	0.0049	Reject H_0
Comfort_Asking_Help	0.3813	0.9441	Fail to reject H_0

Table 4.3.1: Kruskal Wallis Test result for Level groups

3/ Results

It is very interesting that even though the graphs show some inequalities across the confident categories, in fact, they do not differ between the year-level groups. The only difference among the level groups is Access Campus Resources. So, we will examine how it differs among the level groups.

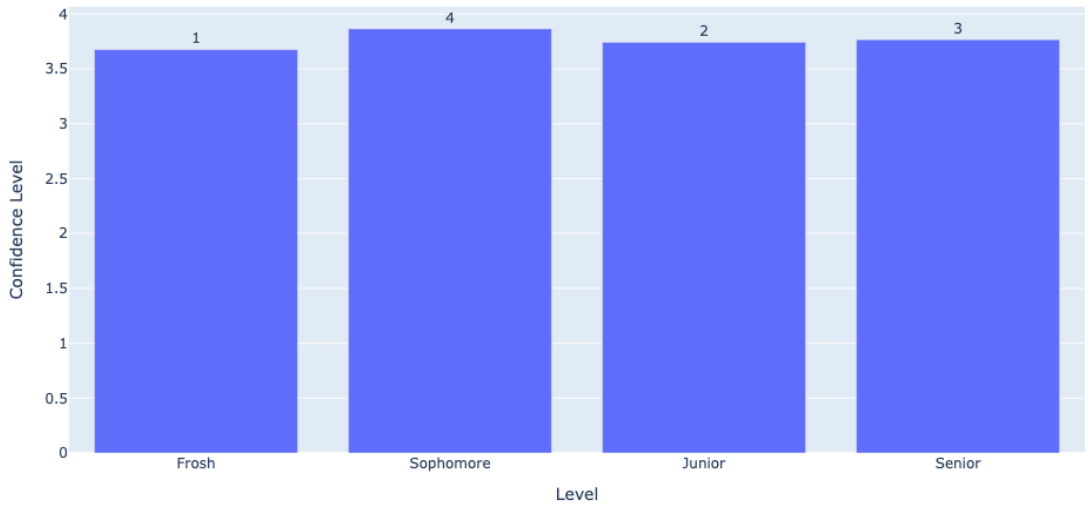

Figure 4.3.2: Confidence Level ranking of Access Campus Resources by Racial/Ethnic groups

Although there are differences, we see that the confidence in Access Campus Resources is generally very close to each other ~3.77/5. Explaining the ranking, we see that Frosh has limited knowledge about Campus Resources since it is their first time at SJSU. Sophomores are the ones who know about Campus Resources the best as they received a lot of guidance and orientation from the school and other organizations. Junior and Senior stays in 2nd and 3rd but they are not different. It may be because they know the resources already but do not update the new ones. So, for Level groups, we don't have any notable challenges among the groups besides Time Management.

4/ Comparing confidence levels between genders - transfer students - first-gen college students.

Both genders (identity), transfer students, and first-gen college students are binary groups (2 possible values). So, we will group them and perform the Kruskal Wallis test once to save time. First, let's look at the general trend.

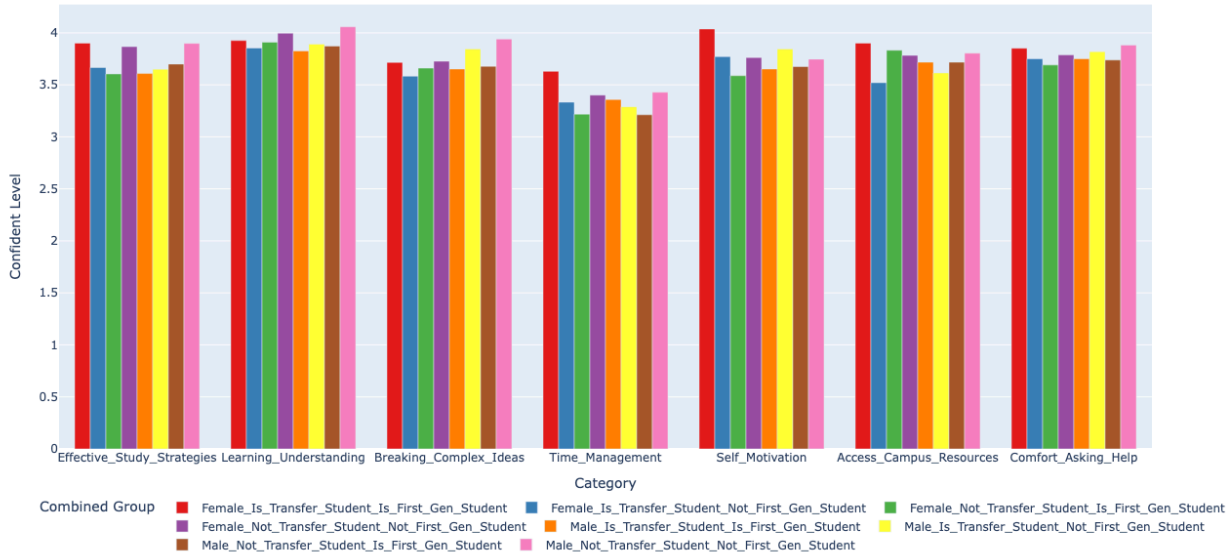

Figure 4.4.1: Confidence Level ranking of each Confidence Category by (genders - transfer students - first-gen college students) group

From the plot, there is no stand-out combined group among all the confidence categories. However, the most challenging skill is still Time Management. So, let's dive into the Kruskal Wallis test.

1/ Hypothesis:

- **Null Hypothesis (H_0)**: There is no significant difference in the median scores of the metric across (genders - transfer students - first-gen college students) groups.
- **Alternative Hypothesis (H_A)**: There is a significant difference in the median scores of the metric across (genders - transfer students - first-gen college students) groups.

2/ Test statistic and p-values (α=0.05):

Feature	**Test statistic**	**p-values**	**Decision**
Effective_Study_Strategies	31.1132	0.0001	Reject H_0
Learning_Understanding	13.4730	0.0614	Fail to reject H_0
Breaking_Complex_Ideas	29.9446	0.0001	Reject H_0
Time_Management	15.7161	0.0278	Reject H_0
Self_Motivation	15.4330	0.0308	Reject H_0
Access_Campus_Resources	12.0416	0.0992	Fail to reject H_0
Comfort_Asking_Help	9.4047	0.2249	Fail to reject H_0

Table 4.4.1: Kruskal Wallis Test Result for (genders - transfer students - first-gen college students) Groups Comparison

3/ Results:

From the result of the Kruskal Wallis test, while there are significant differences among the combined groups in Effective Study Strategies, Breaking Complex Ideas, Time Management, and

Self Motivation, those groups do not exhibit any significant differences in Learning Understanding, Access Campus Resources, and Comfort Asking Help. So, we will explore and perform challenge ranking.

4/ Ranking confidence categories for (genders - transfer students - first-gen college students) groups:

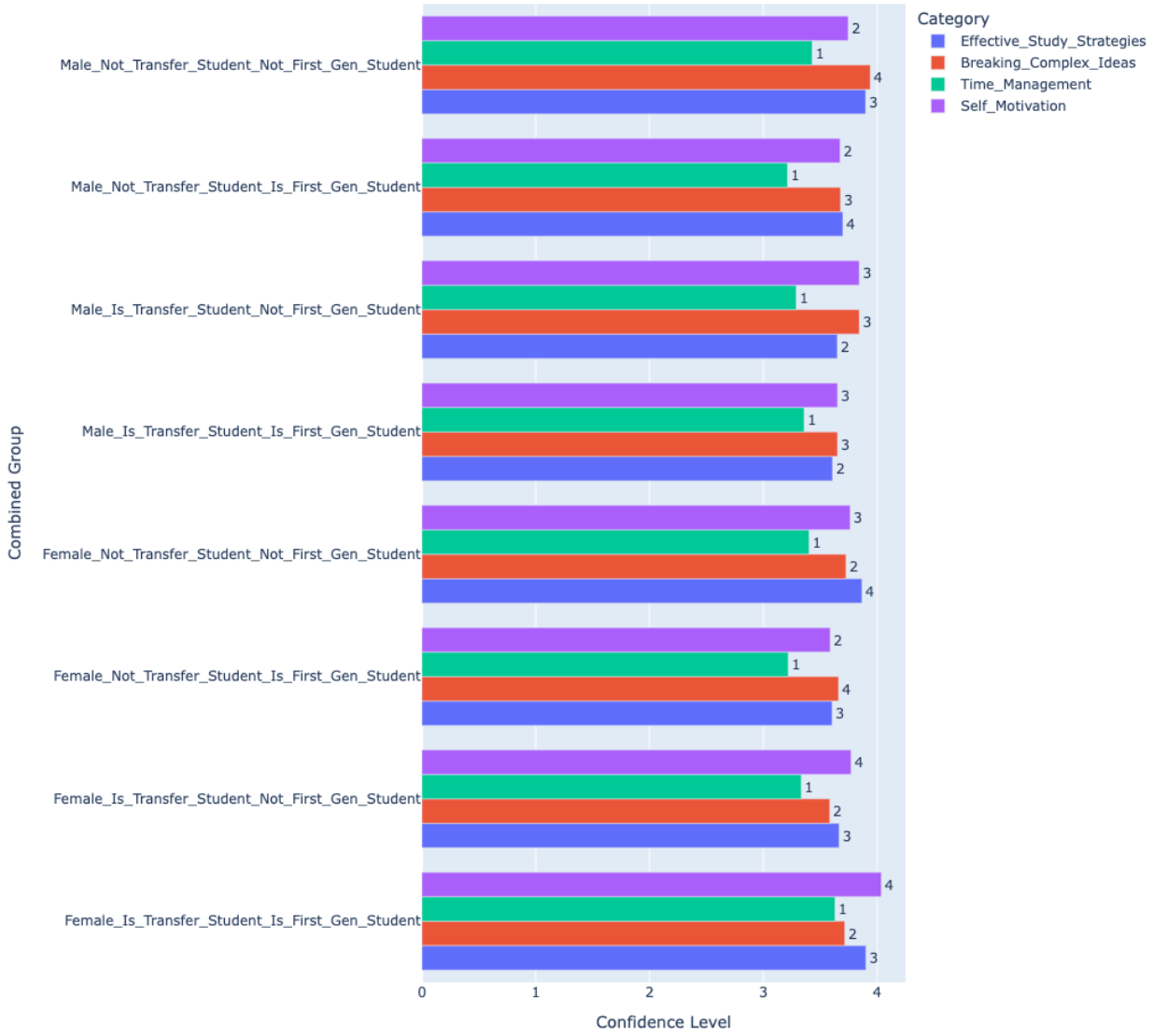

Figure 4.4.2: Confidence Level ranking of Confidence Categories by (genders - transfer students - first-gen college students) groups

Since we already know Time Management is the biggest challenge, we will skip it and look at how it is for other categories.

Group	Top 1	Top 2	Top 3
Male **Not Transfer Student** **Not First Gen Student**	Self Motivation	Effective Study Strategy	Breaking Complex Idea
Male **Not Transfer Student** **Is First Gen Student**	Self Motivation	Breaking Complex Idea	Effective Study Strategy
Male **Is Transfer Student** **Not First Gen Student**	Breaking Complex Ideas / Self Motivation	Effective Study Strategy	
Male **Is Transfer Student** **Is First Gen Student**	Breaking Complex Ideas / Self Motivation	Effective Study Strategy	
Female **Not Transfer Student** **Not First Gen Student**	Breaking Complex Idea	Self Motivation	Effective Study Strategy
Female **Not Transfer Student** **Is First Gen Student**	Self Motivation	Effective Study Strategy	Breaking Complex Idea
Female **Is Transfer Student** **Not First Gen Student**	Breaking Complex Idea	Effective Study Strategy	Self Motivation
Female **Is Transfer Student** **Is First Gen Student**	Breaking Complex Idea	Effective Study Strategy	Self Motivation

Table 4.4.2: Top 3 Most Challenging Categories by (genders - transfer students - first-gen college students) groups

Overall, only the Male Transfer Not First Gen, Male Transfer First Gen, Female Transfer Not First Gen, and Female Transfer First Gen groups have the same top 3 challenges, other groups have different top 3. The main component that makes the most difference is gender, as we can see a very large difference in the top 1 challenging category between the first and last four groups. But overall, we finally reviewed all the significant demographic groups in the dataset.

Outcomes

The analysis brought up a lot of good insights about the interactions between the confidence categories. Among the categories, Effective Study Strategy influences mostly all other learning-related confidence categories. Also, students with good time-management skills will have better self-motivation. By exploring the general trends of all confidence categories, time management is the skill that students have the least confidence in. When segmenting the dataset by demographic data, the analysis found that there are no significant differences in confidence categories between levels at SJSU groups. Also, we came up with the top 3 challenging skills by group:

Group	**Top 1**	**Top 2**	**Top 3**
Black or African American	Effective Study Strategies	Breaking Complex Ideas	Comfort Asking Help
Chicanx/Chicana/ Chicano	Effective Study Strategies	Breaking Complex Ideas	Self Motivation
East Asian	Self-motivation	Effective Study Strategies / Breaking Complex Ideas	Comfort Asking Help
Latinx/Latina/ Latino	Breaking Complex Ideas	Effective Study Strategies	Comfort Asking Help

South Asian	Comfort Asking Help	Self Motivation	Breaking Complex Ideas
Southeast Asian	Self Motivation	Effective Study Strategies	Comfort Asking Help
West Asian	Self Motivation	Breaking Complex Ideas	Effective Study Strategies
White/European/ Canadian	Comfort Asking Help	Breaking Complex Ideas	Effective Study Strategies
Male **Not Transfer Student** **Not First Gen Student**	Self Motivation	Effective Study Strategy	Breaking Complex Idea
Male **Not Transfer Student** **Is First Gen Student**	Self Motivation	Breaking Complex Idea	Effective Study Strategy
Male **Is Transfer Student** **Not First Gen Student**	Breaking Complex Ideas / Self Motivation	Effective Study Strategy	
Male **Is Transfer Student** **Is First Gen Student**	Breaking Complex Ideas / Self Motivation	Effective Study Strategy	
Female **Not Transfer Student**	Breaking Complex Idea	Self Motivation	Effective Study Strategy

Not First Gen Student			
Female **Not Transfer Student** **Is First Gen Student**	Self Motivation	Effective Study Strategy	Breaking Complex Idea
Female **Is Transfer Student** **Not First Gen Student**	Breaking Complex Idea	Effective Study Strategy	Self Motivation
Female **Is Transfer Student** **Is First Gen Student**	Breaking Complex Idea	Effective Study Strategy	Self Motivation

Table 5.1.1: Summary of Top 3 challenged skills among investigated groups

This table is a suggestion about the challenges that students in each specific group may face. Peer Educators can utilize this information to plan the approach and give students personalized assistance.

Suggestions

To better utilize this information, I suggest incorporating research papers/ discussions on how to tailor the solutions to help students in each specific group tackle the challenges that they face. Moreover, we can repeat this analysis on a 4-year-long frequency to assess the challenges of the current set of students. Noted that in this analysis, I discarded a lot of

variables that can be utilized as the demographic segmenting factor due to the under-representation of input. If we could gather more input from a diverse set of people, we could expand the size of the guidelines to capture more in-depth groups. Furthermore, we can try different grouping combinations to isolate the groups from each other.