**Effects of University Social Life on a Student’s GPA**

**Introduction:**

The research aims to look at the cumulative effect of the different parameters of social life like involvement in extracurricular and academic clubs, participation in sports, socializing with friends, and other recreational activities on the GPA of students. This is done by using surveys as a medium of data collection which collects information on all the different parameters of social life that we have included as being a part of the x variable. This is then followed by using the Ordinary Least Squares (OLS) regression to drive meaning out of the collected data. This data is then analyzed, and a conclusion is given. The initial assumption was that higher involvement in one’s social life would lead to a decline in the GPA of that person. This paper tries to look into this assumption and check the effect of the x variable on the y variable, with the x variable being involved in social life and the y variable being GPA.

**Literature Review:**

The aim was to find relevant and recent studies with similar demographics of students that will help us finds gaps in what has not been explored and give us direction on how to proceed with working on said gaps. Most of the studies here have taken place in North America, so the paper tries to recognize socio-economic differences, cultural differences, etc.

**‘What Affects Student Outcomes More: GPA or Participation in co-curricular activities?’-Dr. Joanna Mirecki Millunchick and Yixian Zhou (2020)**

This study had similar variables but took a different approach to analyzing them. Here, they explored the difference in GPA as well as social and professional outcomes between students who participated in extracurricular activities and those who did not. The study is limited to a single disciple, which is engineering, with a sample size of 8,734,022 students across two batch cohorts who were invited to be surveyed, out of which researchers were able to extract 873 complete responses. They found that there was a positive relationship between participation in extracurricular activities and GPA, as well as interpersonal relationships. The authors describe that people who participate in extracurricular activities are likely to bond with people and make study groups, which attributes to their higher GPA. The authors define this as Social Bonding Capital. They also suggested that participation in co-curricular activities contributes to a well-rounded education by fostering skills such as leadership, teamwork, time management, and social interaction. One difference between the extracurricular activities that the paper is looking at and the ones under the purview of this study is that the said extracurricular activities are limited to engineering-related co-curricular, sports, community service, etc. I did not consider community service in our survey; however, I took a diverse range of activities under consideration for the study such as involvement in sports, student organizations, an active social life, and even the relationship status of students.

**‘The Life Cycle Pattern of Collegiate GPA: Longitudinal Cohort Analysis and Grade Inflation’- Wayne A. Grove and Tim Wasserman (2004)**

This study did a time series analysis of the GPA of students enrolled in an 8-semester undergraduate program. They had a much bigger sample size, as they were studying a large cohort of batches ranging from years 1998-2002. Their aim was to see trends in semester-wise GPA, along with how specifically sororities and fraternities affect GPA. As JGU does not have sororities or fraternities, our variables for time spent in student organizations will play a similar role. Further, they showed an immediate dip in the GPA of students who joined a sorority or fraternity than their counterparts that did not.

This given study was conducted over a time period of 5 years, which gives them more data points to work with in terms of the evolution of GPA with respect to the progression of studies. However, this study also homogenizes a very heterogenous group, as students range from multiple disciplines and levels of study. My paper only considers students of one batch i.e. The 2020 batch, to maintain homogeneity in the research. This given study also utilized secondary data from the university, whereas we collected data from the students themselves.

Moreover, this study had additional focus areas, such as the problem of grade inflation, emphasizing the importance of maintaining grading standards to preserve the integrity of GPA as a meaningful measure of student achievement.

**‘Productive procrastination: academic procrastination style predicts academic and alcohol outcomes’- Erin C. Westgate, Stephanie V. Wormington, Kathryn C. Oleson, and Kristen P. Lindgren (2016)**

This study explores how procrastination styles can predict alcohol intake or alcohol-related issues. It conducted a survey among a large, randomized group of undergraduate students, with 1106 respondents. They also provided a monetary compensation of $15 to respondents, which we had not done.

Their questionnaires had a range of situational-type questions, for which there were multiple options of responses from which the respondents could choose. They were also able to verify the self-reported grades of the students, and they found that there was little discrepancy between the two. They also made use of predefined alcohol-related questionnaires from relevant sources in their series of questionnaires to analyze the drinking patterns of students and further categorize and correlate them to a procrastination style. This creates an indirect pathway of the relation between GPA and alcohol intake, but we have attempted to make a more direct correlation between the consumption of alcohol and GPA.

They used GPA as a center point to relate it to procrastination and alcohol consumption, whereas we are only looking at alcohol intake.

In its findings, the study showed that active procrastination, characterized by engaging in alternative productive tasks, may have some benefits in terms of academic performance; however, passive procrastination, involving non-productive activities, can have negative consequences for both academic achievement and alcohol consumption. This study hinted at the beginnings of our second research question, which involves the study of “non-productive” activities on GPA, we do not, however, analyze the relation between non-productive activities and alcohol intake.

**‘Comparison of GPA between Students Involved and Uninvolved in a Relationship’- Jessica M. Vaeth (2009)**

This study examines exclusively the relationship between the GPA of students that are involved in romantic relationships and those who are not. This is one of the factors that I have considered to have an effect on student performance measured by GPA.

There was a survey conducted among 92 students that had been chosen by classroom and were of similar ages, the average being 21 years. They were asked questions relating to information regarding their GPA, relationship status, duration of relationship, and seriousness of relationships. This study found that there was no significant relation between GPA and a student’s relationship status. The author feels other factors, such as time management skills, individual motivation, and study habits, may have a more significant influence on academic performance than relationship status alone.

A challenge that this author faced was inflated GPAs, as they saw a discrepancy between self-reported GPAs and actual GPAs.

‘**Exploring the Association between Campus Co-Curricular Involvement and Academic Achievement’- Dessa Bergen-Cico and Joe Viscomi (2013)**

The study had over 3,000 subjects across two batches of students, who were examined from their entry to an undergraduate program till the completion of 8 semesters. They did not have any direct contact with the students, as they took all the required data from the university itself. They took data for attendance at co-curricular events on the university campus through the database the university had by scanning student ID cards at the entrance of the event. The students under consideration had been informed at the time of admission that their student ID numbers will be shared for research purposes. Additional information like gender, academic year, cumulative GPA, and current GPA were also pulled up.

Researchers found extremely significant results, where the GPA of students who attended a higher number of extracurricular, on-campus activities showed better academic performance as measured by GPA than those who attended fewer events.

### **Research Objective:**

The trade-off between social activities and academic performance.

### **Research Questions:**

### This study aims to examine how extracurricular activities affect the GPA of undergraduate students of the 2020 batch enrolled in O. P. Jindal Global University, currently in their 3rd year of undergraduate education, namely involvement in sports, and student organizations. The paper also explores how social, “non-productive” activities, namely (1) Smoking, (2) Drinking, (3) Going out, and (4) Involvement in romantic relationships, affect a student’s GPA.

### **Data:**

### **Data Collection-**

The data collection process was composed of primary data which was taken in the form of online surveys. This was done so as to minimize the possible social desirability bias that may arise when asking the people being surveyed what their GPA for a particular semester was, whether they smoke and/or drink, and how frequently. The social desirability bias may also contribute to people exaggerating how much time they contribute in clubs/ societies as more time contributed may make them seem dedicated to what they do. In online forms, anonymity prevents this from taking place as much as it would in person, however, there still may be a deviation in what is reported from the actual truth. The online surveys were only floated to third-year students of OP Jindal Global University to preserve the uniformity of data. This was done to take the semester 4 and semester 5 GPAs to seek the link between extra-curricular activities and GPA. I took semester 4 and semester 5 in particular as this was right after the covid lockdown and semester 6 (the final semester for the three-year course students) may include skewed data as more people would focus on studying instead of extracurricular activities. Additionally, it would be tough to gather the GPAs for semester 6 as it was ongoing at the time of data collection.

The data collection process proved to be more difficult than initially anticipated. I had floated the survey forms on class groups, however seeing the minimal responses received, I floated the survey forms on smaller groups on WhatsApp. This yielded better results, gaining a total of 65 responses.

### **Data Structure-**

With the survey, I got a total of 65 responses. However, out of the 65 responses, 3 were taken for the initial pilot study, and 12 responses had to be discarded due to incomplete information thus, I was left with a total of 50 valid responses to work with and include in the model.

The Pilot study had us change certain values to only numerical or to limit the possible answers as people were not only giving answers that varied in metric but also non-serious responses. While the sample is considered small in general, I believe that it represents the population of the environment in which the study was conducted (OP Jindal Global University) as it takes into account a wide range of students from different degree domains and different interests. Since the population was limited and due to time constraints, I believe the sample was enough to conduct the research.

### **Ethics-**

In the data collection process, priority was given to the ethical aspect of the process than to the number of samples attained. The survey instructions made sure to notify the person being surveyed that the survey was completely voluntary. Additionally, due to the survey being online, the person being surveyed had the complete liberty to stop the survey at whatever point they felt uncomfortable. I had also assured anonymity and that the data provided would only be used in the context of the research and not on an individual basis.

### **The Survey Questions and the reasons behind choosing them-**

The survey questions were mainly conditional questions such as “Are you part of any sports teams,” if the answer was yes there was a follow-up question of “How many hours a week do you dedicate to sports.” The same question was asked with respect to clubs. Another question asked was “How many days a week do you go out on a recreational basis”. This was to provide an alternative option in case people didn't spend their time on sports or clubs but still could not dedicate time to academics. This recreation was also captured by the questions of “Do you drink?” and “Do you smoke?” and how frequently for both individual questions. This was to see if drinking or smoking impacts people's GPA in either a positive or negative way. There was also a question “Were you in a relationship for the period of Semester 4 or/and 5” and if yes then “Did it end badly?” This was to gauge whether time was spent on a relationship instead of academics or if there was a positive or negative deviation of Semester 5 GPA from Semester 4 GPA, it could be explained by the relationship. Finally, the last two questions were “What was your Semester 4 GPA” and “What was your Semester 5 GPA”. These questions were to obtain information necessary for the Dependent Variable.

I decided on these questions in particular in order to see what is the relation between time spent on clubs and sports on GPA, also collected information on recreational activities such as smoking and drinking as this may have a correlation to a student's GPA. This correlation may be negative or positive and may impact the GPA in addition to time spent on activities concerning clubs and/or sports.

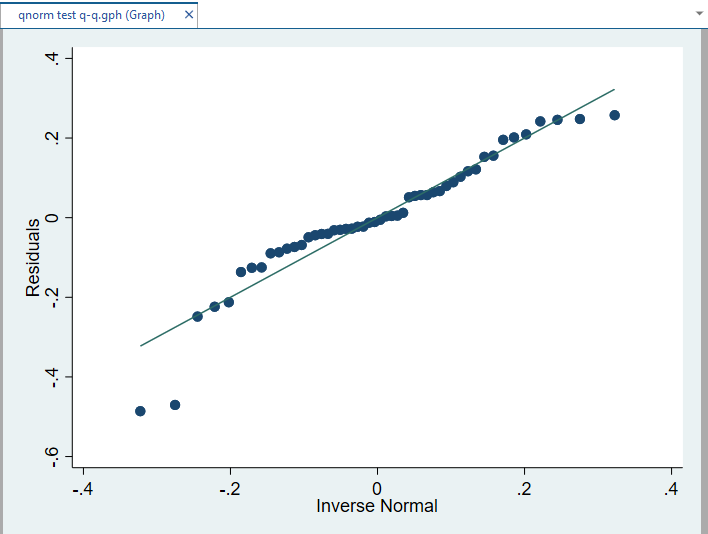
**Model:**

To perform an Ordinary Least Squares (OLS) regression, I first need to make sure that two of the fundamental Gauss-Markov assumptions were not being violated-

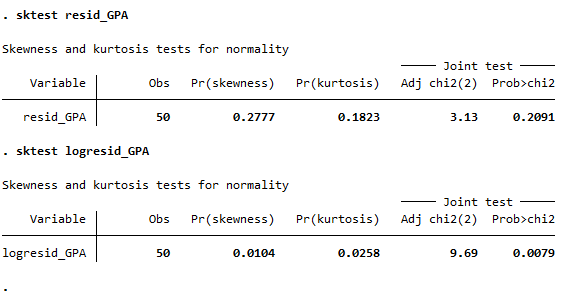
1. The residuals or the errors of the regression line are approximately normally distributed.
2. Data is homoskedastic- This is an assumption of equal or similar variances in the variables. Homoscedasticity means “having the same scatter”, i.e., the difference between predicted and observed values of an experiment for the variables is constant.

Residuals are basically the differences between the predicted and the actual value of the dependent variable. To check whether the residuals in the dataset are **normally distributed** or not, we perform a Q-Q test. A Q-Q plot, also known as a "quantile-quantile" plot shows the residuals to be normally distributed if they follow an approximate straight line at a 45-degree angle.

This is done by using the Stata command ‘qnorm’ of the residuals found after the regression.



According to the Q-Q plot, we see that on average the residuals are fairly normally distributed, apart from a few deviations along the tail ends.

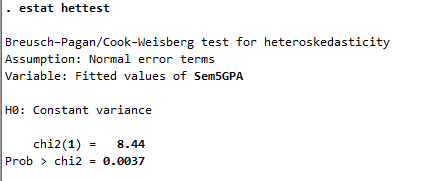


The skewness and the Kurtosis test for normalcy ‘sktest’ also show that the null hypothesis of normal data is not rejected as chi^2 value>0.05.

Thus, I validate the checks for the assumption of the residuals following a normal distribution and proceed ahead.

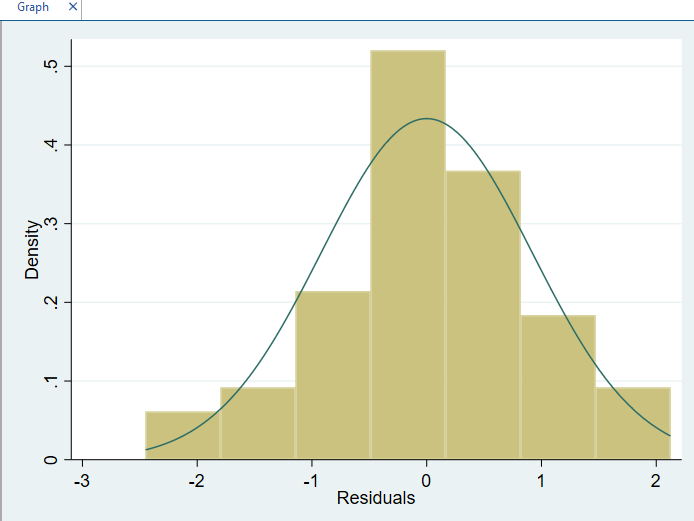
**Homoskedasticty** can be checked by the Breusch-Pagan and Cook-Weisberg test. However, it is important to note that this test for heteroskedasticity assumes that the errors in the linear regression are normally distributed. If the residuals in your model are not normally distributed, then this test for homoskedasticity will be biased.

We check this by the command ‘estat hettest’ in Stata.



It was found that as 0.0037<0.05, I reject the null that the data is homoskedastic. To correct this heteroskedasticity, I conducted a log transformation on the dependent y-variable. Log transforming the dependent variable makes the heteroskedastic or the skewed data closer to having a normal distribution. It improves the linearity between the dependent and independent variables.

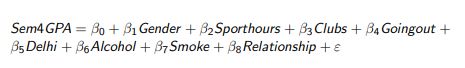
After taking the natural log value of the dependent variable Sem5GPA, the residuals were found to be closer to a normal distribution as seen plotted in the histogram below, and thus the violation of the homoskedasticity assumption was now corrected. I performed similar tests and corrections for Sem4GPA as well.

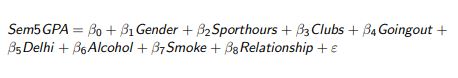


After checking that the data meet the assumptions of an Ordinary Least Squares Regression, I then started with the OLS Regression. According to Gauss Markov, OLS regression produces the most unbiased estimates that have the smallest variance of all possible linear estimators and thus has been named as BLUE- Best Linear Unbiased Estimator.

**Empirical Strategy:**

I have used two regression models to tease out the relationship between social activities and semester GPA.



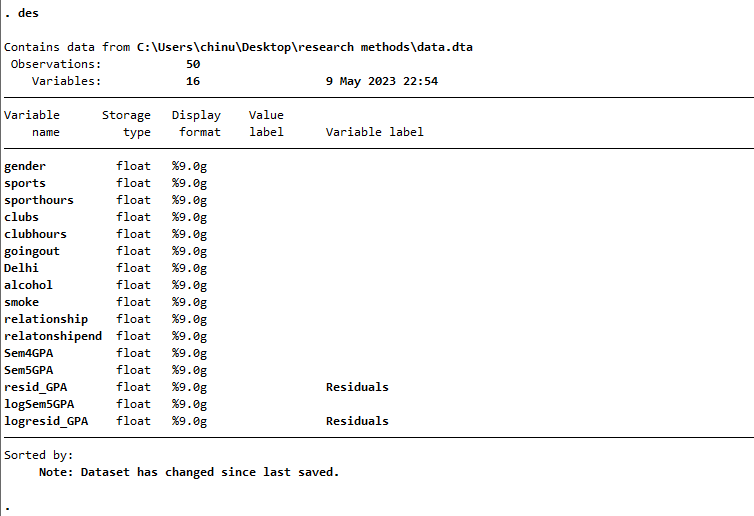


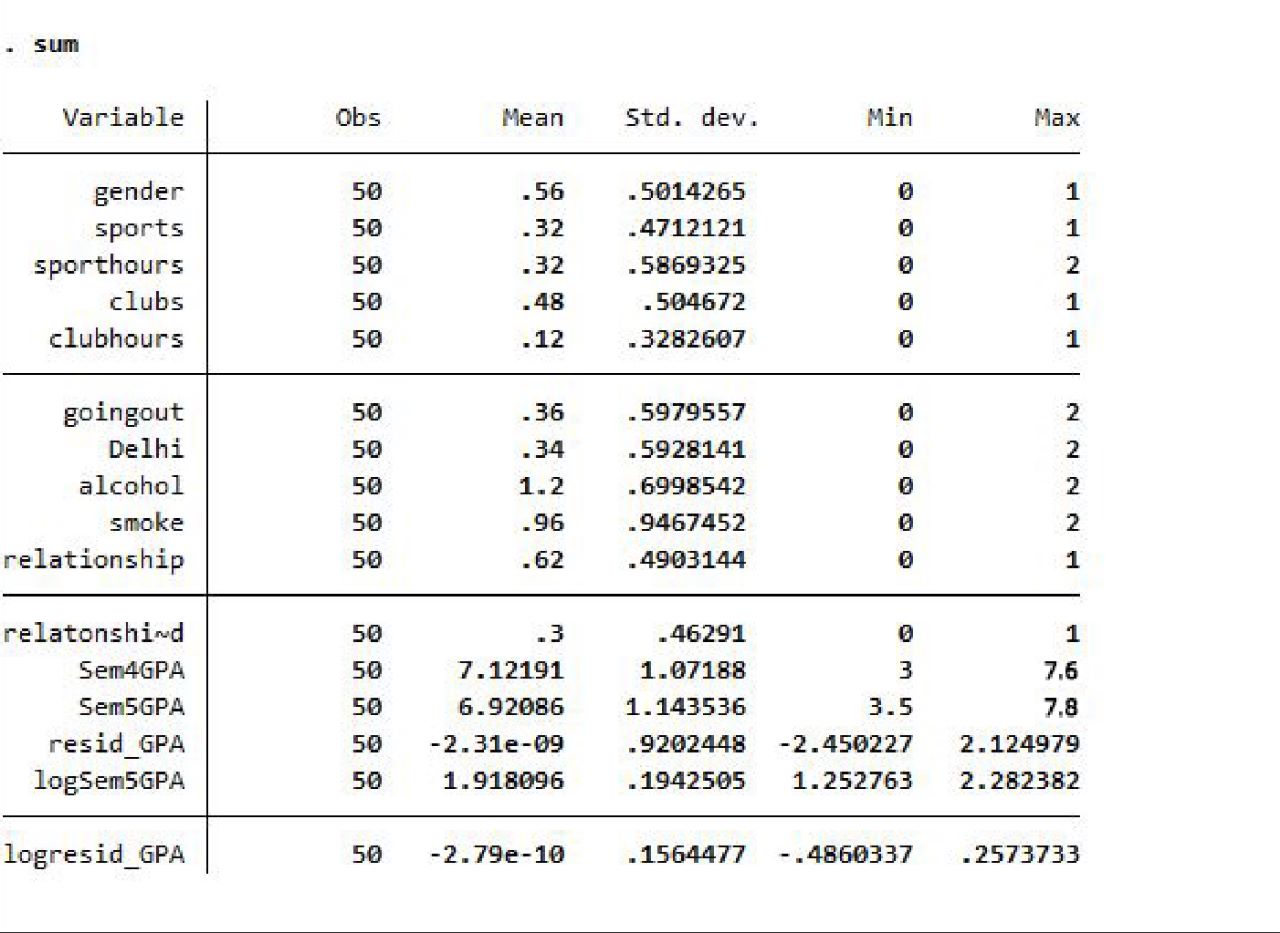
Where Sem4GPA and Sem5GPA are the two dependent variables

Gender, Sporthours, Clubs, Goingout, Delhi, Alcohol, Smoke, and Relationship are dependent variables

E is the error term that captures all the unobservable factors affecting the above relationship

**Descriptive Statistics-**





The dataset consists of 16 variables and 50 observations. For the regression, I have used the relevant 8 independent and 2 dependent variables.

Gender is balanced in our sample with 56% of the 50 observations consisting of females. Students who play sports are fewer as compared to students who are part of clubs in the university. However, the time dedicated to playing sports per week exceeds the time dedicated to clubs.

The average of the sample goes out to Delhi or to other recreational places nearby the University like Oasis, Yellow Box, ABCD, etc. 0-1 times per week. When asked if students have never smoked, smoke rarely, or frequently, on average people (according to our limited sample size) have chosen rarely, which signifies that students in Jindal smoke as a recreational activity and that the average do not smoke often or on a daily basis.

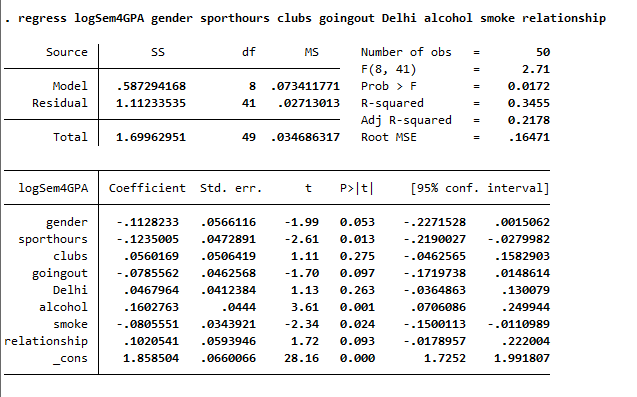
A majority (62%) of the students were in a relationship last year but only 3% reported to their relationship has ended on a bad note.

Drinking alcohol is more common among the student body and on average students have answered that most drink alcohol rarely or frequently per week.

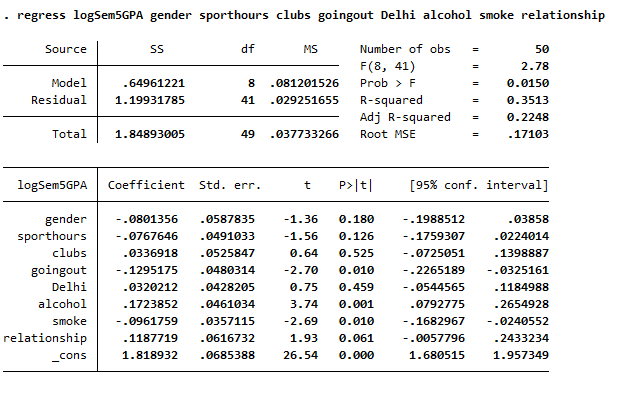
Students fared better in their 5th semester as compared to their 4th semester. The 4th semester GPAs in our sample ranged from 3 to 7.6 out of 8 while that in semester 5 ranged from 3.5 to 7.8 out of 8.

**Regression Results-**

**Semester 4 GPA-**



**Semester 5 GPA results-**



**Analysis and Implications:**

I now discuss the results obtained from the transformed model. Since the dependent variable is taken in its log form, the interpretation of the coefficient will be the exponentiate of the coefficient. To get the exponentiated values, I used STATA command exp () on the already statistically significant coefficients. This gives the multiplicative factor for every one-unit increase in the independent variable. The following variables mentioned are all the statistically significant variables at the confidence level of 90%

Coming to the results of the model for Semester 4:

**1)** **Gender**: The coefficient obtained for the variable is -0.11282. This indicates that for every one unit increase in gender, sem4GPA will decrease by a factor of 0.8933. GPA of females is lower by a factor of 0.893 than that of the male sample.

**2)** **Sport Hours**: The coefficient of -0.1235 indicates that for every one hour increase in sport hours, GPA decreases by a factor of 0.884. This effect can be very well understood intuitively as people who spend more hours playing a sport don’t necessarily get ample amounts of time to study and thus fail to maintain their GPA.

**3)** **Going Out:** We get a coefficient of -0.0785 which again tells us that for every unit increase in individuals going out their expected GPA falls by a factor of 0.9244. Individuals who are very regular in going out of campus are possibly irregular at studying and hence their GPA reflects it.

**4)** **Alcohol:** The coefficient obtained for alcohol is 0.16027. This indicates that for every individual who drinks alcohol, sem4GPA will increase by a factor of 1.1738. This moves opposite to our hypothesis as we expected an increase in alcohol consumption to be negatively related to GPA, however, our obtained results show us that people who consume alcohol tend to have a better GPA.

**5)** **Smoke:** We get a coefficient of -0.08055, meaning a fall of GPA by a factor of 0.922 for every individual smoking. We get a negative relationship between smoking and GPA. This result debunks the common hypothesis that smoking helps with stress and in turn can help one get better grades. But as per the results, GPA falls for every smoker in Jindal. Another possible interpretation of this coefficient could be the existence of simultaneity bias in the model; a person who is already bad at academics shifts to instant dopamine resources like smoking.

**6)** **Relationship:** The coefficient of 0.10205 indicates that for every person who was/is in a relationship, their GPA increased by a factor of 1.107. This proves that being in a relationship (given, it’s healthy) can affect your mental health by reducing stress and help achieve better grades.

Further, interpreting the results of the regression performed in semester 5, I get similar outcomes compared to semester 4. The interpretations are as follows:

**1)** **Going Out:** I get a coefficient of -0.12951 which tells us that for every unit increase in individuals going out their expected GPA falls by a factor of 0.8785. Comparing this coefficient to the sem4GPA model’s coefficient I observed that going out during semester 5 has a bigger impact on GPA compared to going out in semester 4. Semester 4 was considered a hybrid semester and students’ exams were mostly done through online modes, which honestly did not require any considerable effort. However, semester 5 was a full-fledged offline semester and all the exams were conducted in an offline mode, this could have intimidated students and the ones who went out regularly during this semester faced the consequences heavily (a factor of 0.8785 to be precise).

**2)** **Alcohol:** The coefficient obtained for alcohol is 0.17238. This indicates that for every individual who drinks alcohol, sem5GPA will increase by a factor of 1.188. I again don’t get a negative relationship as we expected and rather a very strong positive relationship as the p-value for this variable is 0.001 which is significant at possibly all levels of confidence.

**3)** **Smoke:** The coefficient of -0.09617 means a fall in GPA by a factor of 0.908 for every individual smoking. I similarly get a negative relationship between smoking and GPA.

**4)** **Relationship:** The coefficient of 0.11877 indicates that for every person who was/is in a relationship, their GPA increased by a factor of 1.1261. This acknowledges the positive relationship between a healthy relationship and a greater GPA.

In the analysis, most of the variables had credible outcomes which were intuitively very well explained. However, the alcohol variable kept me unsatisfied with the results as it did not make any logical sense. To delve deeper into understanding the impact of alcohol consumption accurately, I interviewed some students from the sample (people who have filled out the survey) and asked them questions about their drinking habits.

I asked questions in a vague/communication form to get authentic responses and we established that there is a division in the types of drinking preferences which couldn’t reflect in our model. One section of students goes out drinking at night, usually a Thursday, and tend to miss a few classes or stay the night out due to the fright of strict security policies of Jindal, leading to a shabby routine and hence a bad GPA. Another section of students usually prefers to go out for drinking in the day and avoid all the issues that the other section goes through, leading to a better GPA for them. What we concurred from this interview was that our data for alcohol consumption could possibly be skewed towards people who day-drink rather than the other half. The bias in our sample towards day drinkers explains the unusual positive relationship between GPA and alcohol consumption.

**Limitations:**

The study carried out has the following possible limitations:

* Self-selection bias/Nonresponse bias:

Self-selection bias is when respondents who have strong opinions on a topic are more likely to reply to survey questions than respondents who have indifferent opinions. Self-selection bias occurs as the survey was floated online and it was up to the discretion of people whether they would fill the form or not. Following the nonresponse error are the statistical distinctions between a survey that only includes respondents and a faultless survey that would also include nonrespondents. However, in the survey carried out a particular group of people like for example those with low GPA’s could have refrained from filling the form and therefore we would experience the non-response bias.

* Social desirability bias-

The respondent's desire, whether conscious or implicit, to acquire prestige or appear in a different social role can cause a bias from the side of the respondent known as the social desirability bias. In the study carried out there was an attempt to minimize the social desirability bias by keeping the forms anonymous, however, there cannot be a full guarantee of the eradication of the bias as a whole.

* Low external validity-

If the external validity is low, then the results are not generalizable to a wider population. In this study, since the sample size was restricted and not too huge, therefore, we can say that the external validity of the obtained results would be low.

**Future research areas:**

1) The effect of specific social activities: While the initial research may have looked at the overall influence of social life on semester GPA, subsequent studies might concentrate on the effects of individual social activities. For instance, researchers may investigate the impact of socializing with friends, visiting events, and engaging in extracurricular activities on GPA. A researcher may, for instance, compare the GPA of students engaged in various activities, such as sports, student government, or participation in an artistic or musical organization. Upon examining the statistics, one may discover that students who are involved in groups with an academic focus often have GPAs that are higher than those of their peers who participate in extracurricular activities that are non-academic.

2. The impact of social networks: Upcoming research may look at how social networks, such as friendship groups and peer groups, influence GPA. Researchers might look at whether students with strong social networks are more or less likely to have their GPA decline, as well as if the strength of these social networks influences GPA. Additionally, by comparing students who have friends with great academic accomplishments to those who have friends with inferior performance, the researcher may investigate how peer influence affects GPA. By drawing a comparison between the two groups, one may find out if kids who are surrounded by academically strong peers are likely to have better GPAs because of good peer impact and academic drive or not.

3. The importance of time management abilities: Studies may also look at whether students with good time management abilities are less likely to have their social lives have a detrimental effect on their GPA. This can include looking at the connection between efficient time management, social interaction, and academic achievement. For example, researchers can ask them how they manage their time between academics and social obligations like going to parties, hanging out with friends, or utilizing the internet for communication. Analyzing the data may reveal that students who prioritize studying and successfully manage their time tend to get better grades even when they participate in extracurricular activities.

4. The effect of mental health: Future research might examine the relationship between social life, GPA, and students' mental health. Researchers can investigate if students who are more involved in their social lives but also have mental health problems are more likely to see their GPA drop. Future researchers can consider doing research in which they examine how students' social engagement and mental health relate to their overall well-being. For example, the researcher may investigate whether students who actively engage in social activities like joining clubs or attending social gatherings report greater levels of well-being, lower levels of stress, and better mental health. Additionally, he/she may investigate how beneficial social experiences contribute to students' general success by looking at the relationship between social engagement, well-being, and academic achievement.

5. The effect of student support services: Lastly, future research might look at how student support services can lessen the negative effects of social life on grades. If students employ services like academic advising, tutoring, or counseling, are they better able to juggle their social obligations and academic responsibilities? The researcher, for instance, can compare the academic performance of students who actively participate in social activities and get support services to those of students who do not use such services. This study could aid in clarifying how supportive services can help to offset any detrimental effects of social participation on academic performance.

The issue of how university social life influences semester GPA has a wide range of prospective study topics to investigate, and the selection of these areas should be informed by the gaps in the body of current knowledge as well as the research question.