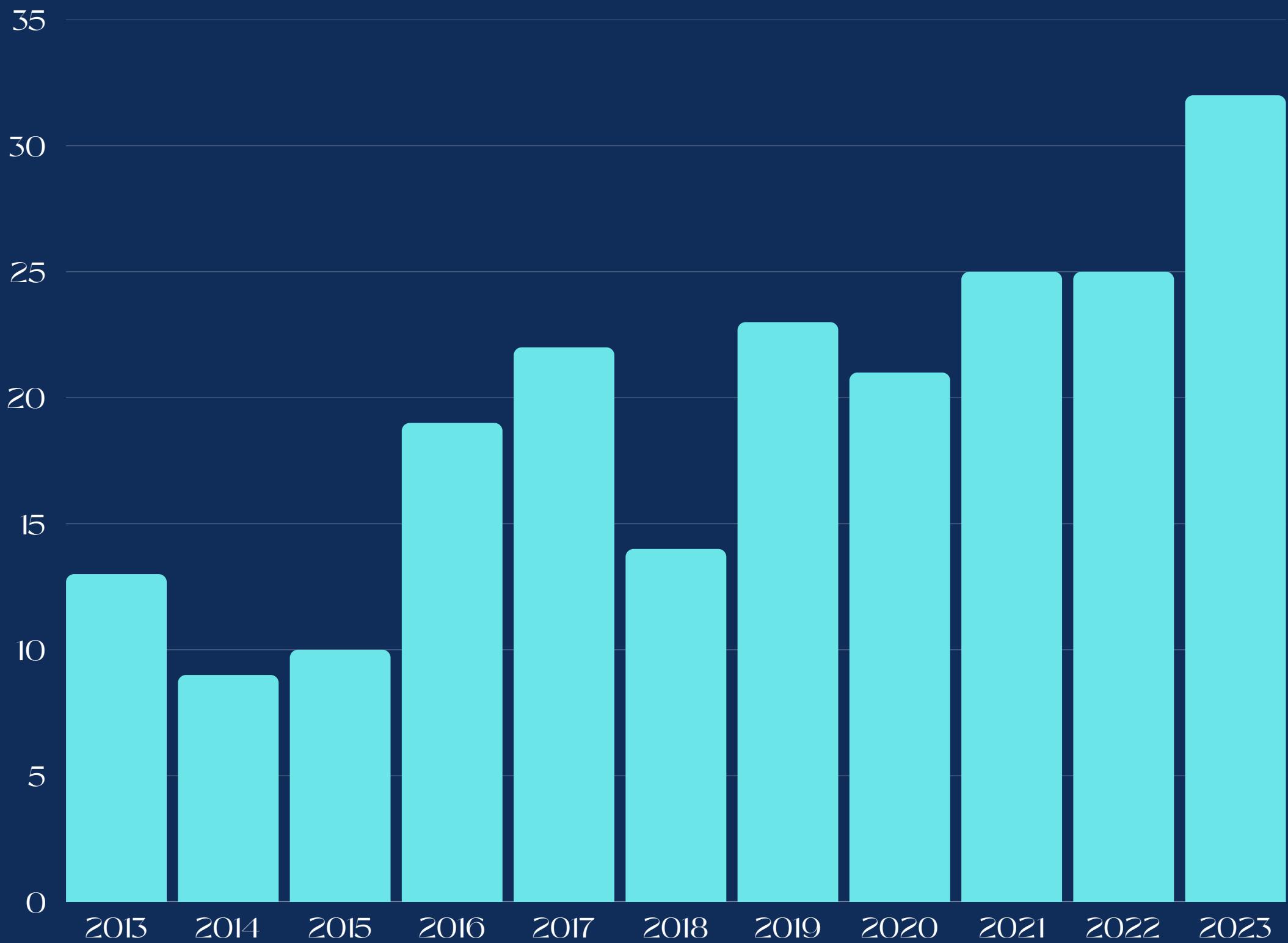


# INTRODUCTION

- Student suicide case become serious
- We want to investigate the reason behind those cases
- Relationship between suicide rate and student learning



THE NUMBER OF STUDENT SUICIDES IN HONG KONG HAS NEARLY DOUBLED OVER THE PAST 10 YEARS.



SOURCE FROM HK01

# DATASET

高處	22.462273126028425	113.9978353489257	非住宅	不詳	男	16	有	身亡
和順樓	22.38762100297017	114.19247957041216	公共屋邨	梁	男	50	沒有	身亡
龍達樓	22.3401108903606	114.195321938418	公共屋邨	陳	男	36	沒有	身亡
附近引水道	22.367670477403895	114.07534443800476	公共屋邨	程	男	40	沒有	身亡
一公廁	22.45513848813719	114.14108786239846	非住宅	盧	男	54	沒有	身亡
春葵樓	22.366345416015367	114.1306172378824	公共屋邨	郭	男	57	沒有	身亡
一單位	22.39256961771911	113.96571307675597	非住宅	張	男	46	有	身亡
秀山樓	22.37896914675205	114.13119036500726	公共屋邨	陳	男	47	有	身亡
玉成樓	22.33156112042987	114.16209393881215	洋樓	陳	男	50	有	迷糊
頌群樓	22.4240957424219	114.22679087791519	公共屋邨	鄭	男	32	沒有	身亡
美槐樓	22.379322239852904	114.17614721042985	公共屋邨	許	女	54	沒有	身亡
麗德樓	22.331095830820182	114.15806750185438	公共屋邨	歐	女	54	沒有	身亡
怡發花園	22.336591472499457	114.20827976276033	居屋	不詳	男	不詳	有	清醒
嘉匯大廈	22.328979614699982	114.16640258882741	洋樓	朱	男	41	有	身亡
1座	22.341972341612873	114.20762990341086	私人屋苑	吳	女	28	沒有	身亡
高瑞閣	22.296009373208523	114.24023816037308	居屋	不詳	男	不詳	有	清醒
一大廈高處	22.437288067278832	114.03269341962134	私人屋苑	郭	男	20	沒有	身亡
1座	22.50102251263429	114.14110863631815	私人屋苑	薛	男	57	沒有	身亡
富榮樓	22.3671323564113	114.124289609164	公共屋邨	關	女	25	沒有	身亡
洪昌樓	22.438528551037212	113.99694846635431	公共屋邨	梁	女	26	有	身亡
碧綠樓	22.2486026214298	114.157799156466	公共屋邨	李	女	69	沒有	身亡
對開海面	22.429721712145977	114.23579626396726	非住宅	張	男	22	沒有	身亡
翠榕樓	22.31493698840788	114.23075392417388	公共屋邨	陳	男	91	有	身亡
高處	22.41739115482632	114.20864911916888	非住宅	郭	女	23	有	昏迷
附近山坡	22.446451938725463	114.1704698253949	非住宅	何	男	64	沒有	身亡
元蘿樓	22.335895064183298	114.15634876587262	公共屋邨	周	男	66	有	身亡

21	Medical	Female	3.35	2	4	3	Good	Moderate	Good	High
18	Law	Male	3.65	2	2	5	Good	Moderate	Average	Moderate
21	Business	Female	3.4	0	3	3	Average	Low	Average	Low
24	Medical	Male	3.8	3	2	1	Poor	Low	Average	Moderate
19	Engineering	Female	3.05	2	5	0	Average	Moderate	Good	Low
23	Law	Female	3.74	3	2	4	Average	Low	Good	Moderate
28	Engineering	Female		3	0	3	Average	Moderate	Average	Moderate
22	Computer Science	Male	3.19	1	1	3	Average	Moderate	Average	Moderate
27	Medical	Male	3.26	3	2	2	Average	Moderate	Average	High
24	Medical	Female	3.2	3	0	3	Average	Low	Poor	Moderate
25	Law	Male	3.61	3	1	5	Good	Low	Average	Moderate
18	Medical	Female	3.85	4	1	3	Good	Low	Average	Moderate
19	Medical	Male	3.26	5	1	1	Good	Low	Average	High
22	Computer Science	Male	3.46	3	1	0	Good	Moderate	Average	Moderate
20	Medical	Male	3.43	2	2	2	Good	High	Average	Low
20	Business	Male	3.38	5	3	2	Average	Moderate	Average	Low
27	Computer Science	Male	3.56	1	3	2	Good	High	Good	High
31	Medical	Female	3.5	4	5	2	Average	Moderate	Average	High
18	Computer Science	Female	3.6	2	0	5	Good	Low	Average	Low

# Suicide case in Hong Kong & Mental Health Survey of University Student



## HYPOTHESIS

Poor academic performance or  
lower CGPA is correlated with  
higher levels of stress



# METHODOLOGY - I

## Import CSV & DAT dataset

### SUICIDE CASE OF HONG KONG

```
PROC IMPORT DATAFILE="/home/u64020952/sasuser.v94/香港自殺報道資料庫 - RAW DATA_Variable Name (變數名稱).dat"
  OUT=suiside_data
  DBMS=CSV
  REPLACE;
  GETNAMES=YES;
RUN;
```

### MENTAL HEALTH SURVEY OF UNIVERSITY STUDENT

```
data mental_health;
  infile '/home/u64020914/SAS 366F/students_mental_health_survey.dat' dlm=',' firstobs=2;
  input Age Course $ Gender $ CGPA Stress_Level Depression_Score Anxiety_Score Sleep_Physical_Activity $ Diet_Quality $ Social_Support $ Relationship_Status $ Substance_Use $ Counseling_Service_Use $ Family_History $ Chronic_Illness $ Financial_Stress Extracurricular_Involvement Semester_Credit_Load Residence_T;
run;
```



## METHODOLOGY - 2

# PROC Freq

```
proc freq data=mental_health;
  tables Gender*(Depression_Score Anxiety_Score Stress_Level) / chisq plots=mosaicplot;
run;
```

```
PROC IMPORT DATAFILE="/home/u64020952/sasuser.v94/香港自殺報道資料庫 - RAW DATA_Variable Name (數據每小時更新).csv"
OUT=suicide_data
DBMS=CSV
REPLACE;
GETNAMES=YES;
RUN;
```



# METHODOLOGY - 3

## PROC MEANS

```
proc means data=mental_health n mean min max;  
  var Age CGPA Stress_Level Depression_Score Anxiety_Score;  
run;  
  
/* Frequency distributions for categorical variables */
```



# METHODOLOGY - 4

## PROC SGPLOT

### HISTOGRAM

```
PROC SGPLOT DATA=mental_health;
  histogram CGPA;
  density Stress_Level;
  xaxis min=2.75 max=4; /* Set the x-axis scale */
  yaxis min=1 max=5; /* Set the y-axis scale */

  TITLE "Depression and Anxiety Rates Over the Years";
RUN;
```

### SCATTERPLOT

```
proc sgplot data=mental_health;
  scatter x=CGPA y=Stress_Level;
run;
```



# METHODOLOGY-5

## MERGE DATASET

```
/* Step 1: Import the first dataset (香港自殺報道資料庫) */
data suicide_data;
  infile '/home/u64020952/sasuser.v94/香港自殺報道資料庫 - RAW DATA_Variable Name (數據每小時更新).csv' dlm=',' firstobs=2;
  input caseID $ newsDate : yymmdd10. newsTime $ newsHelp $ newsURL $ Suicide_Case;
  format newsDate yymmdd10.;

run;

/* Step 2: Import the second dataset (students_mental_health_survey.dat) */
data mental_health_data;
  infile '/home/u64020952/sasuser.v94/students_mental_health_survey.dat' dlm=',' firstobs=2;
  input Age Course $ Gender $ CGPA Stress_Level Depression_Score
    Anxiety_Score Sleep_Quality $ Physical_Activity $ Diet_Quality $
    Social_Support $ Relationship_Status $ Substance_Use $
    Counseling_Service_Use $ Family_History $ Chronic_Illness $
    Financial_Stress Extracurricular_Involvement Semester_Credit_Load
    Residence_Type $;
  run;

proc sql;
  create table merged_data as
  select a.Stress_Level, b.Suicide_Case, a.CGPA
  from mental_health_data as a
  inner join suicide_data as b
  on a.CGPA= b.Suicide_Case;
quit;

/* Step 3: Derive a Suicide_Case variable in mental_health_data */
data mental_health_data;
  set mental_health_data;

  /* Define Suicide_Case based on thresholds for mental health indicators */
  if Depression_Score >= 4 or Anxiety_Score >= 4 or Stress_Level >= 4 then Suicide_Case = 1;
  else Suicide_Case = 0;
run;

/* Step 4: Perform a T-Test to compare CGPA based on the derived Suicide_Case */
proc ttest data=mental_health_data;
  class Suicide_Case; /* 0 = No, 1 = Yes */
  var CGPA;
run;
```



## METHODOLOGY-6

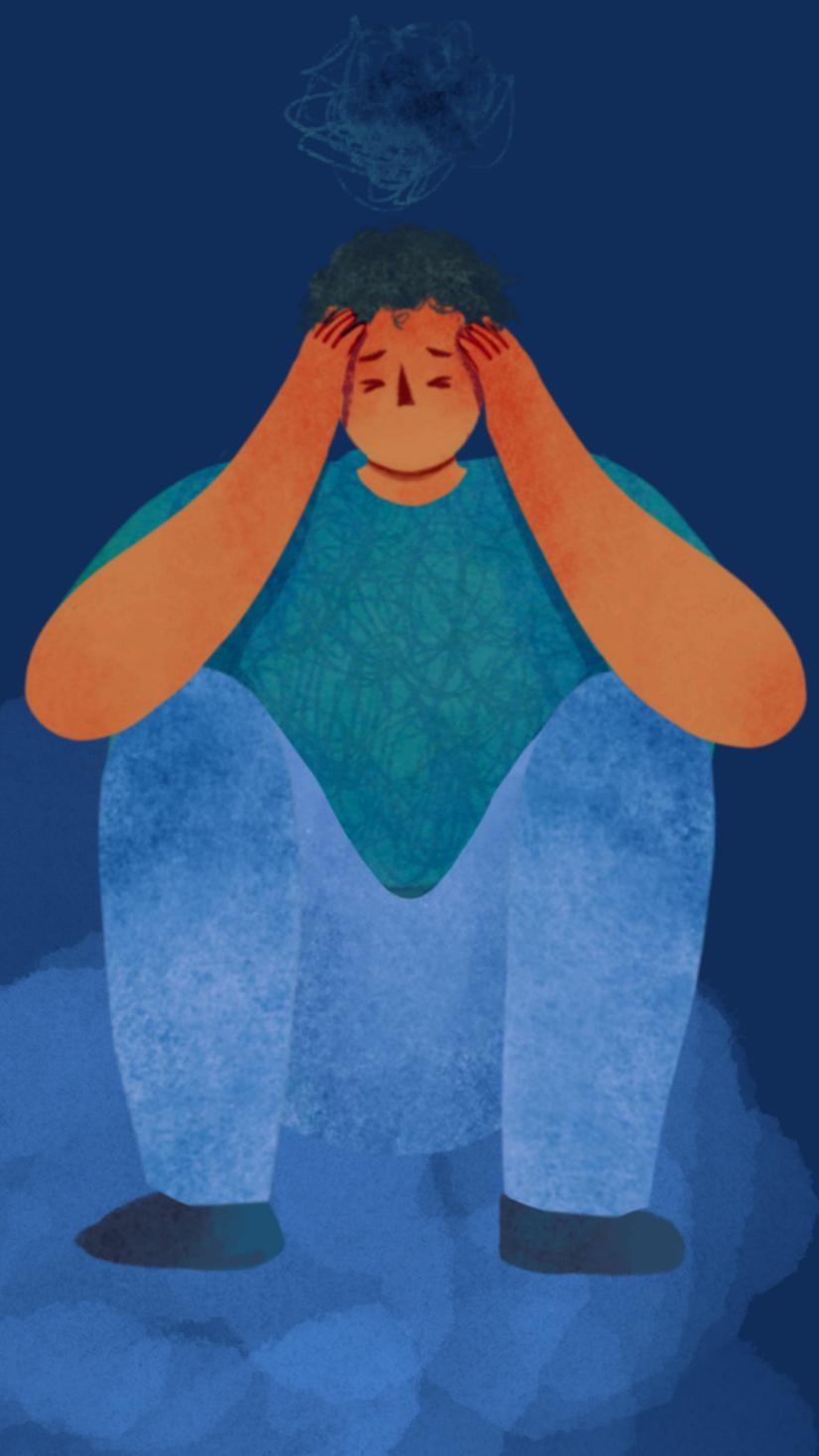
### PROC TTEST

```
29 proc ttest data=mental_health_data;  
30   class Suicide_Case; /* 0 = No, 1 = Yes */  
31   var CGPA;  
32 run;
```

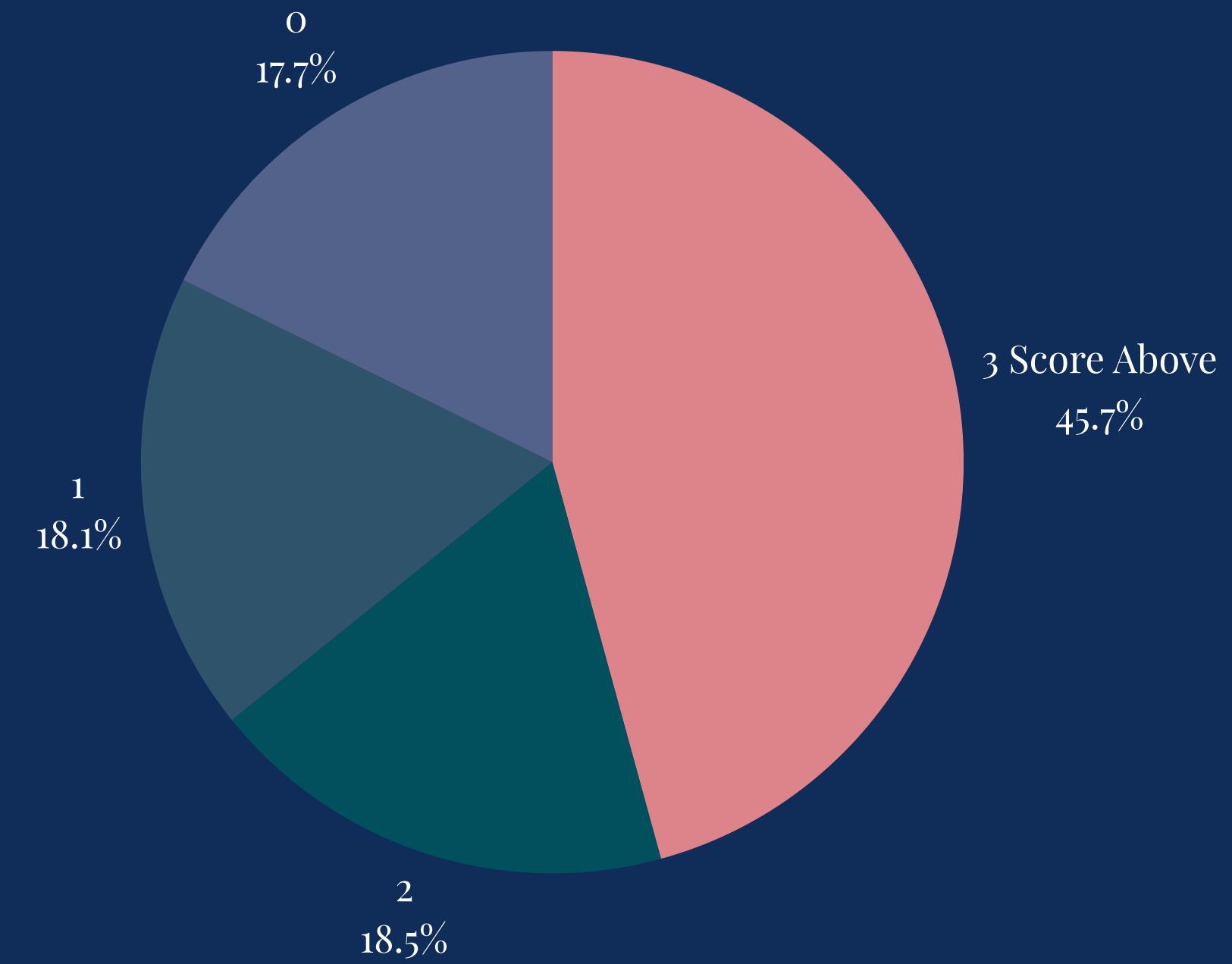


# ANALYTICS

# Anxiety\_Score Of University Student

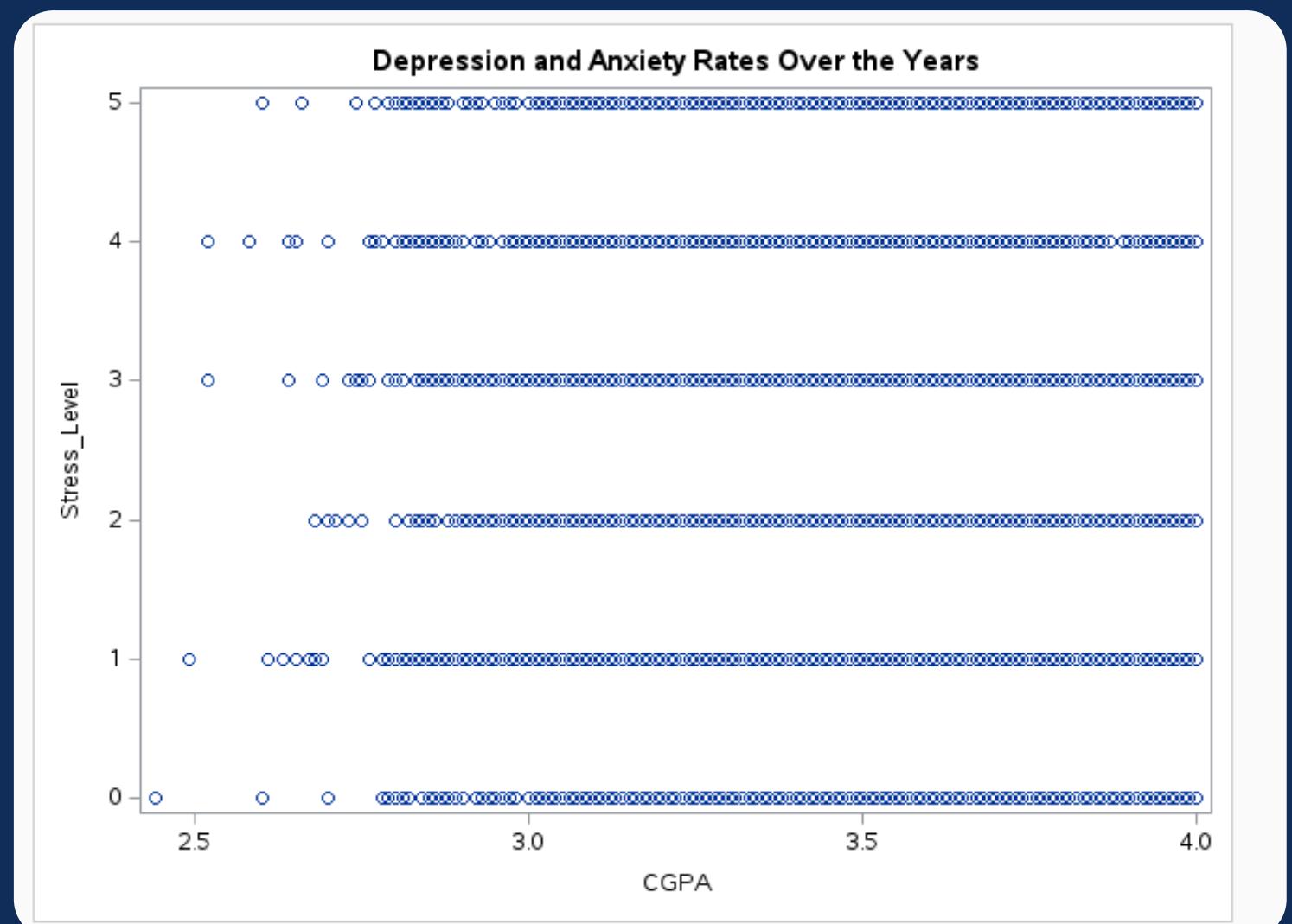
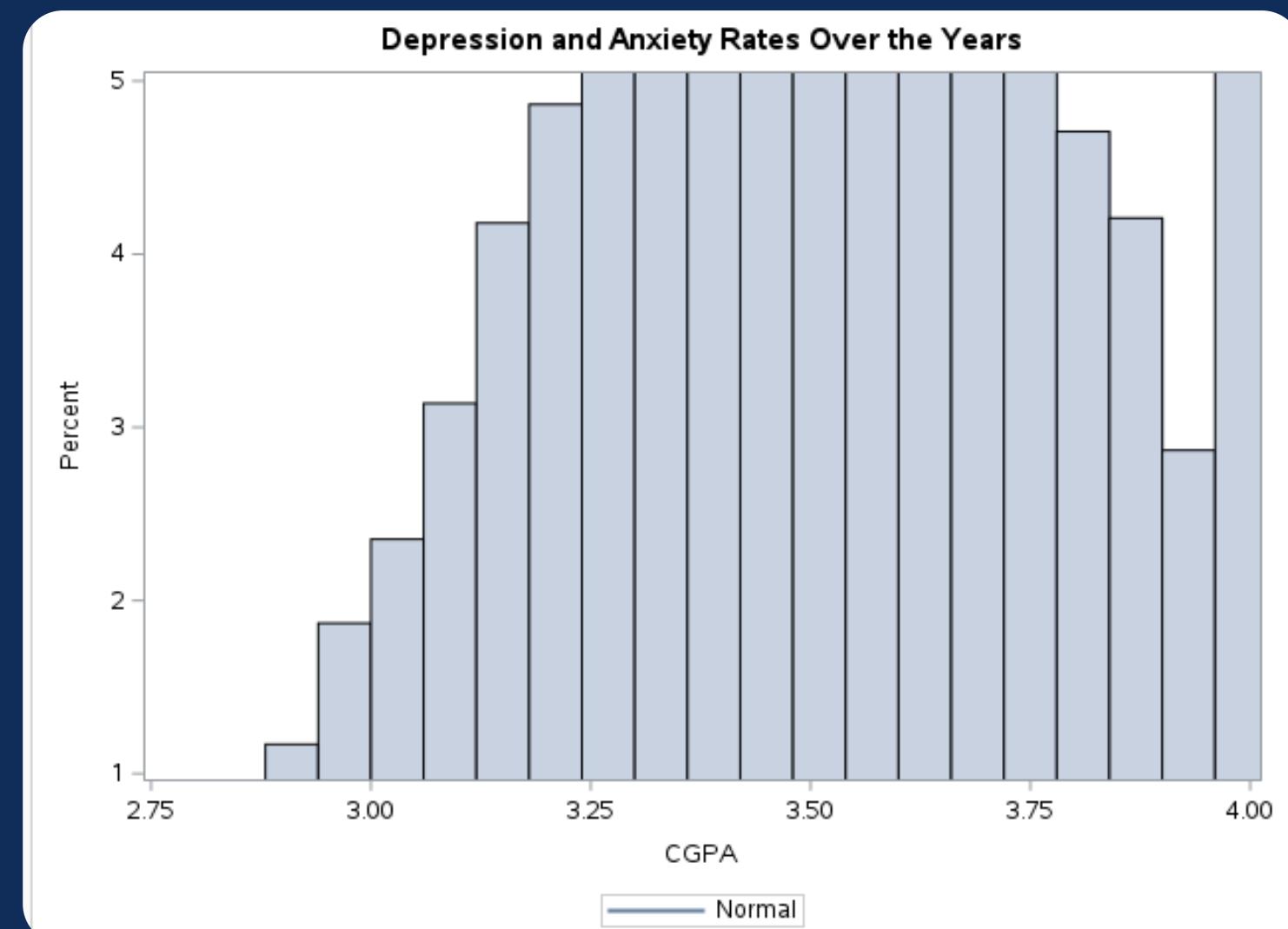
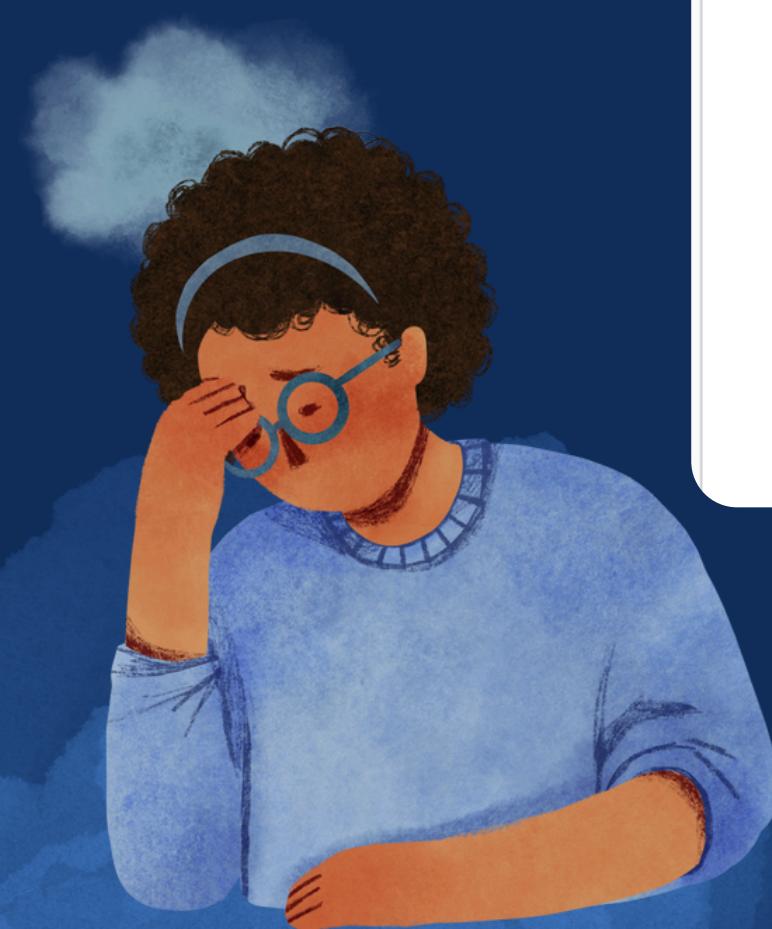


Anxiety_Score	Frequency	Percent
0	1243	17.70
1	1270	18.09
2	1297	18.47
3	1429	20.35
4	912	12.99
5	871	12.40



# ANALYTICS

University students with higher CGPAs often experience greater levels of stress



# DATA ANALYTICS

p-value > 0.05 indicates  
the difference in group means  
is not statistically significant.

The TTEST Procedure							
Variable: CGPA							
Suicide_Case	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
0		2748	3.4903	0.3123	0.00596	0	4.0000
1		4251	3.4859	0.2964	0.00455	1.0000	4.0000
Diff (1-2)	Pooled		0.00448	0.3027	0.00741		
Diff (1-2)	Satterthwaite		0.00448		0.00750		

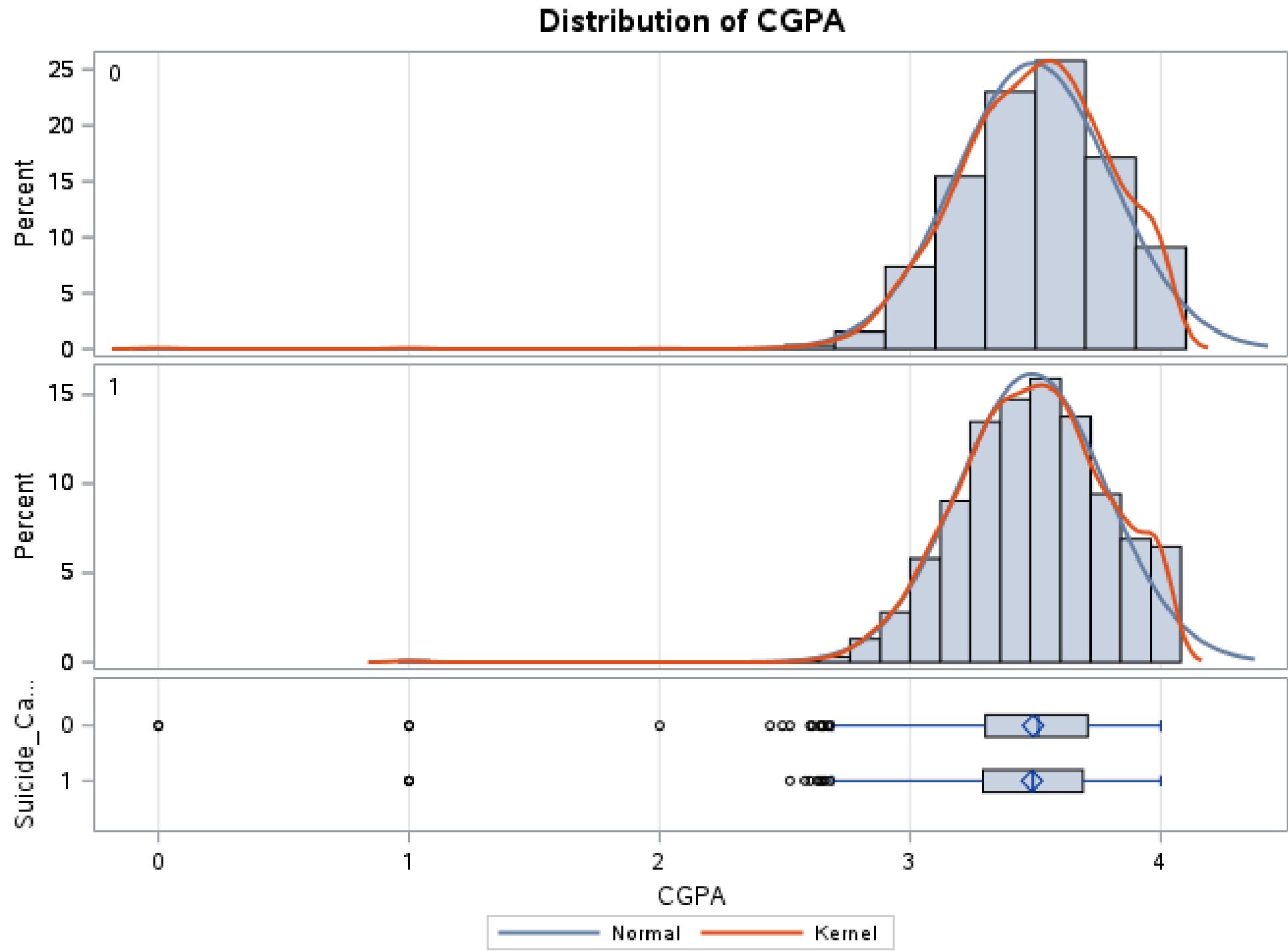
Suicide_Case	Method	Mean	95% CL Mean	Std Dev	95% CL Std Dev
0		3.4903	3.4788	3.5020	0.3123
1		3.4859	3.4769	3.4948	0.2964
Diff (1-2)	Pooled	0.00448	-0.0101	0.0190	0.3027
Diff (1-2)	Satterthwaite	0.00448	-0.0102	0.0192	

Method	Variances	DF	t Value	Pr >  t
Pooled	Equal	6995	0.60	0.5457
Satterthwaite	Unequal	5636.4	0.60	0.5502

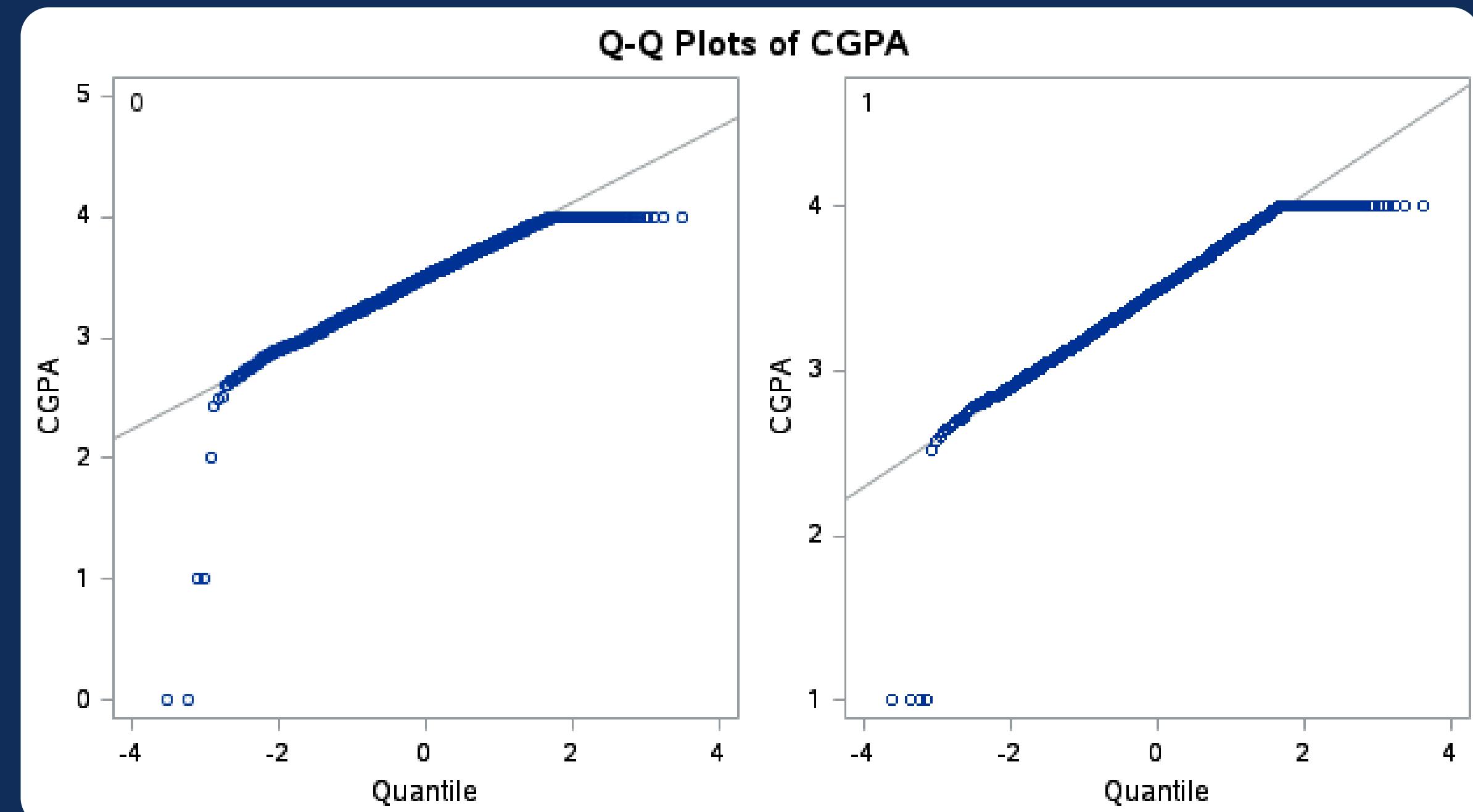
Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	2745	4250	1.11	0.0024

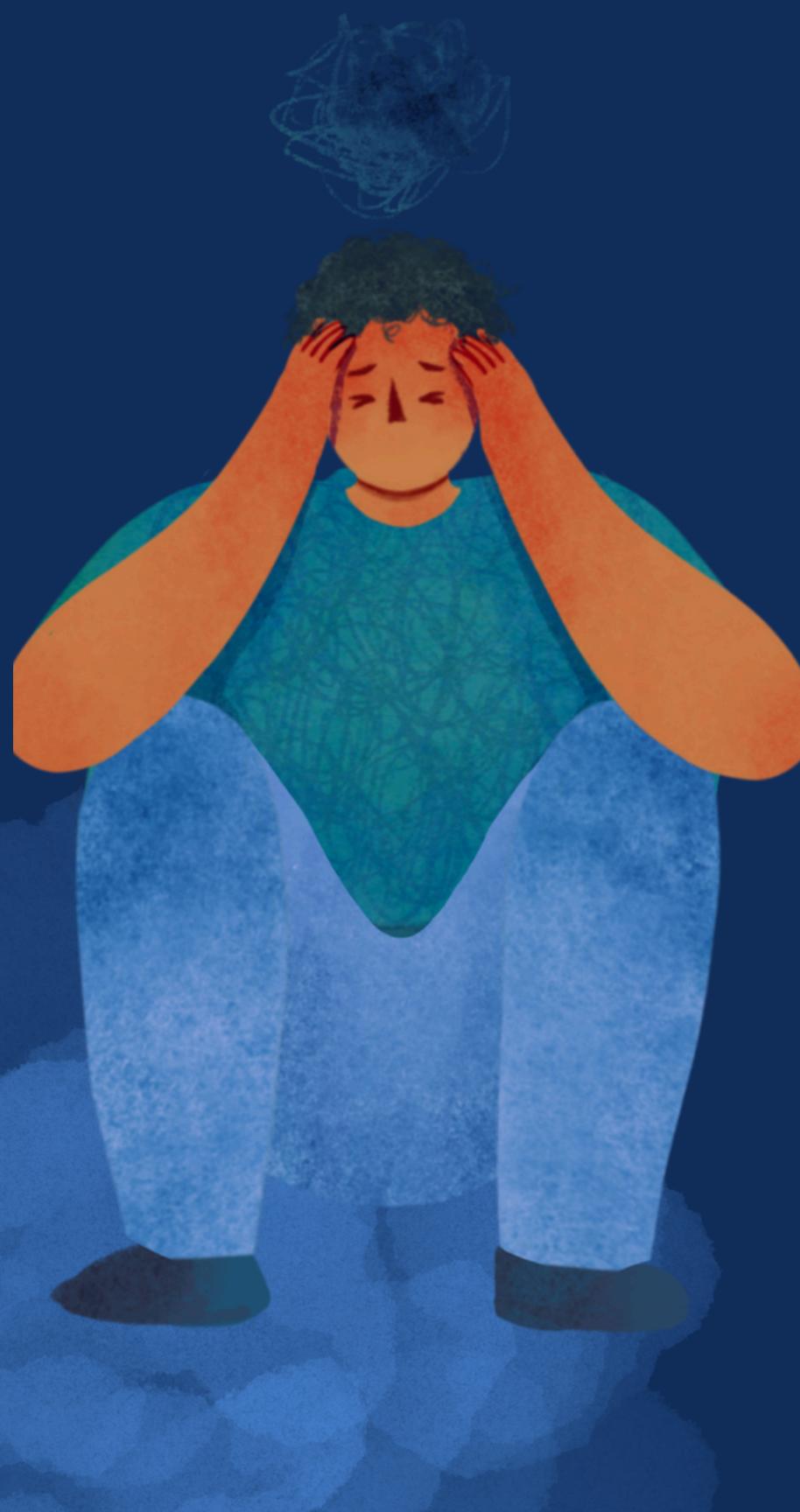


# DATA ANALYTICS



The Q-Q plots demonstrate that the CGPA data for both groups ('Suicide\_Case = 0' and 'Suicide\_Case = 1') approximates normality in the central range, with slight deviations observed in the tails, indicating reasonable validity for parametric analysis.  
not statistically significant.





# Is there a correlation between academic study and suicide?

## Frequency of Suicide by Gender (18-24)

The FREQ Procedure

Gender	Frequency	Percent
女	133	35.75
男	239	64.25

## Frequency of Suicide by Area (18-24)

The FREQ Procedure

reason3	Frequency	Percent
Frequency Missing = 364		

Area	Frequency	Percent
九龍	120	32.26
新界	185	49.73
港島	67	18.01

## Frequency of Reasons for Suicide (18-24)

The FREQ Procedure

type2	Frequency	Percent
Frequency Missing = 358		

reason1	Frequency	Percent
不詳	116	31.18
健康	74	19.89
學業	27	7.26
家庭	24	6.45
工作	9	2.42
感情	66	17.74
生活	29	7.80
畏罪	5	1.34
財政	19	5.11
酒精	3	0.81

# CONCLUSION

- 
1. We merge two dataset together.
  2. we found that CGPA does not appear to be a strong indicator of suicide risk in this dataset.
  3. Although there is no significant correlation between academic studies and suicide, it still has a certain degree of influence.
  4. we have proved that it is insufficient evidence to reject null Hypothesis since we got the  $0.6 \text{ P-Value} > 0.05$
  5. The mental health of students is worth to be noticed