UTM Computing Proceedings

Innovations in Computing Technology and Applications

Volume: 3  **|** Year: 2018  **|** ISBN No. :

*The statistic of students activities during weekdays night*

Nur Asnani binti Hassan

A18CS0186

School of Computing,

Universiti Teknologi Malaysia

Johor, Malaysia

Tamilarasi a/p Subrumaniam

B18CS0027

School of Computing,

Universiti Teknologi Malaysia

Johor, Malaysia

Izzat Marzuqi bin Ghali

A18CS0080

School of Computing,

Universiti Teknologi Malaysia

Johor, Malaysia

Arif Raffiq bin Mustaff Kamaruzzaman

A18CS0036

School of Computing,

Universiti Teknologi Malaysia

Johor, Malaysia

**Abstract**— This short paper had conducted the statistic of student’s activities during weekdays night in UTM. The student’s activities during weekday’s night is to study about the average time spent per night for students to do that activities using hypothesis testing. Then, we used correlation analysis for duration spend on activities in hours and activities that student’s do during weekday’s night. To show the linear relationship, we used regression analysis between linear relationship between the frequency of DurationSpendOnActivity and WeekdayActivity. Lastly, for chi square we used data from the survey.

# Introduction

Students at Universiti Teknologi Malaysia conduct many activities during weekdays night. We used previous collected data from project 1 to perform and inferential statistics. The data involved about what the students usual activities, time students go to sleep during weekday, duration of sleep they get, what kind of activities they priorities on, amount of time that students spent for the activities and with whom they do the activities.

Muhammad Amin bin Izam Hasri

B18CS0012

School of Computing,

Universiti Teknologi Malaysia

Johor, Malaysia

# methodology

The population in this project are students which are studying in Universiti Teknologi Malaysia from different ages, faculties, and level of study. The sample size for this project is 63 students. This study adopted a quantitative method through the distribution of questionnaire. We have created a google form survey which asks the students’ information about what are their usual activities during weekday nights, such as the duration of sleep they get, what kind of activities they priorities on, how much time they spent studying and so on. This survey is then shared to students throughout a variety of media social applications, primarily through WhatsApp and Facebook. After a few weeks, we have obtained the primary data for the project. Then we used SPSS Statistics 24 to generate representation and analysis the data that we collected from the survey into informative data.

# data analysis

## Hypothesis Testing-2 Samples

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Sample Size** | **Standard Deviation** | **Average Time Spent Per Night (Mean)** |
| **Study** | 18 | 1.150 | 3.50 |
| **Playing Video Game** | 18 | 0.924 | 3.17 |

Table 1 standard and mean for time spent per night by students

Conduct a hypothesis test at the 5% level of significant to test the claim.

Solution:

Test of hypothesis:

H0: µ1 - µ2 = 0;

H1: µ1 > µ2. We want to reject H0 if the average amount of student spend Time on Study each night is greater than Playing Video Game.

Given, α = 0.05. The test statistic is

Z0 = x̅1 - x̅2

δ21 + δ22

 n1 n2

Where δ21 =(1.150)2 = 1.3225 and

 δ22 = (0.924)2 = 0.8538

 n1 = 18 and n2 = 18 ,

Z0.05 = 1.645

Reject H0 if Z0 > Z0.05 = 1.645

Computations: Since x̅1 = 3.50 and x̅2 = 3.17, w

Z0 = 3.50 – 3.17

 1.3225+0.8538

 18 18

 = 0.9491

Conclusion: Since Z0 = 0.9491 < 1.645, we fail to reject H0 at the level 0.05 level and conclude that there is sufficient evidence to support that average amount of student spend time on Study each night is lower than Playing Video Game

0.05

1.645

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Sample Size** | **Standard Deviation** | **Average Time Spent Per Night (Mean)** |
| **Hang out with friends** | 18 | 0.984 | 2.65 |
| **Sports** | 9 | 1.067 | 3.11 |

Conduct a hypothesis test at the 5% level of significant to test the claim.

Solution:

Test of hypothesis:

H0: µ1 - µ2 = 0;

H1: µ1 > µ2. We want to reject H0 if the average amount of student spend Time Hang out with friends each night is greater than doing Sports activity.

Given, α = 0.05. The test statistic is

T0 = x̅1 - x̅2

δ21 + δ22

 n1 n2

Where δ21 =(0.984)2 = 0.9683 and

 δ22 = (1.067)2 = 1.1385

 n1 = 18 and n2 = 9,

T0.05 = 1.645

Reject H0 if T0 > T0.05 = 1.645

Computations: Since x̅1 = 2.65 and x̅2 = 3.11, w

T0 = 2.65 – 3.11

 0.9683 +1.1385

 18 9

 = -1.083

Conclusion: Since T0 = 0.9491 < 1.645, we reject H0 at the level 0.05 level and conclude that there is sufficient evidence to support that average amount of student spend time on Hang out with friends each night is lower than Sports activity.

0.05

1.703

## Correlation

The two coefficients are Pearson’s product-moment correlation coefficient and spearman’s rho rank correlation coefficient. Pearson’s product-moment correlation are used which is duration spend on activity and the weekday activity are both interval data type. Since the Pearson’s correlation are used then the purpose for this test is to test the strength of the linear relationship between, time spend on activity and weekday activity.



Hypothesis:

*H0:* p = 0 (no linear correlation)

*H1:* p ≠ 0 (linear correlation exists)

****

Figure 1 Simple scatter plot with fit line of DurationSpendOnActivity by SleepTime

Since -0.239 < 0.59 means we reject *H0*. There is sufficient evidence to show that there is no linear correlation between duration spend on activity and the weekday activities.

## Regression

Regression analysis is about prediction the value of a dependent variable based on the value of at least one independent variable. So, in this regression test will we try to discuss the relationship between duration spend on activity and weekly activity using inferential statistics.

**Inferential statistics**

****

Figure 1Histogram of DurationSpendOnActivity

Hypothesis test for this relation:

*H0*: There is no linear relationship between the frequency of DurationSpendOnActivity and WeekdayActivity ($β\_{1}=0$)

*H1*: There is a linear relationship between the frequency of DurationSpendOnActivity and WeekdayActivity ($β\_{1}\ne 0$)



|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ***Week day Activities (y)*** | ***Duration Spend On Activities (x)*** | ***(x \* y)*** | ***x2*** | $$\hat{y}$$ | ***(***$y\_{i}-\hat{y}$***)2*** |
| 2 | 3 | 6 | 9 | 0.0486 | 1.9514 |
| 1 | 3 | 3 | 9 | 0.0486 | 0.9514 |
| 3 | 3 | 9 | 9 | 0.0486 | 2.9514 |
| 3 | 3 | 9 | 9 | 0.0486 | 2.9514 |
| 1 | 2 | 2 | 4 | 0.3901 | 0.6099 |
| 1 | 4 | 4 | 16 | -0.2929 | 1.2929 |
| 4 | 4 | 16 | 16 | -0.2929 | 4.2929 |
| 3 | 4 | 12 | 16 | -0.2929 | 3.2929 |
| 3 | 3 | 9 | 9 | 0.0486 | 2.9514 |
| 1 | 4 | 4 | 16 | -0.2929 | 1.2929 |
| **∑y = 22** | **∑x = 33** | **∑xy = 74** | **∑x2 = 113** | **∑**$\hat{y}$ **= -0.5385** | **∑(**$y\_{i}-\hat{y}$**)2 =****22.5385**  |

**Calculation**

|  |  |
| --- | --- |
| $$b\_{1}= \frac{Σxy- \frac{Σx\*Σy}{n}}{Σx^{2}- \frac{(Σx)^{2}}{n}}$$$$b\_{1}= \frac{74- \frac{33\*22}{10}}{113- \frac{(33)^{2}}{10}}$$$$b\_{1}= \frac{1.4}{4.1}$$$$b\_{1}= 0.3415$$ | $$b\_{0}= \overline{y}-(b\_{1}\*\overline{x})$$$$b\_{0}=(\frac{22}{10})-(0.3415\*\frac{33}{10})$$$$b\_{0}= 2.2-(1.1268)$$$$b\_{0}= 1.0731$$$$\hat{y}=1.0731-0.3415x$$ |
| Sε = $\sqrt{\frac{SSE}{n-k-1}}$ = $\sqrt{\frac{22.5385 }{10-1-1}}$ = 1.6785$S\_{b\_{1}}=\frac{1.6785}{\sqrt{113- \frac{(33)^{2}}{10}}}$  = 0.4094 | $t=\frac{b\_{1}-β\_{1}}{S\_{b\_{1}}}$  = (0.3415- 0) / 0.4094 = 0.8341v = n – l  = 10 – 1 = 9 α = 0.05, α/2 = 0.025, *tα/2*= 2.262 |

From the calculation above, since$ t<t\_{\frac{α}{2}}$, so we reject *H0*. There is insufficient evidence to show that there is a linear relationship between the night weekdays activity and duration time spent for doing the activities, which is$ β\_{1}\ne 0$.

## Chi-Square test of independent

Chi Square test of independent is used to test whether relationship is exists between the activities of students during weekday nights and the time spend on that activities. 63 data sample is used and the activities of students do during weekday nights is divided into group study, playing games, Hangout with friends and sports while time spend on that activities is distributed into 4 categories which is 1 hour, 2hour, 3 hour and more than 4hours. We claim that at 0.10 significant level, the time spend on that activities is independent to the activities of students do during weekday nights. The null hypothesis and alternative hypothesis are as follow:

Hypothesis:

*H0:* time spend on that activities is independent to the activities of students do during weekday nights

*H1:*  time spend on that activities is dependent to the activities of students do during weekday nights

The expected count Eij is calculated using this formula:

$$E\_{ij}= \frac{(i^{th}Row total)(j^{th}Column total)}{Total sample size}$$

The test statistic x2 is calculated using this SPSS:

$$Test statistic, x^{2}= \frac{∑(O\_{ij}-E\_{ij})^{2}}{E\_{ij}}$$

|  |  |  |
| --- | --- | --- |
|  |  Time spend |  |
| Activity | 1hour | 2hour | 3hour | More than 4hour | Total |
| Study  | 1 (1.7143) | 6(5.4286) | 3 (5.4286) | 8 (5.4286) | 18 |
| Playing games | 1 (1.7143)) | 5 (5.4286) | 6 (5.4286) | 6(5.4286) | 18 |
| Hangout with friends | 0(1.7143) | 5(5.4286) | 8(5.4286) | 5(5.4286) | 18 |
| Sports  | 4 (0.8571) | 3(2.7143) | 2(2.7143) | 0(2.7143) | 9 |
| Total | 6 | 19 | 19 | 19 | 63 |

Table 2 Expected frequency of independence test table

|  |  |  |  |
| --- | --- | --- | --- |
| Cell,ij | Observed count, Oij | ExpectedCount , Eij | [(Oij - Eij)²Eij |
| 1, 1 | 1 | 1.7143 | 0.2976 |
| 1, 2 | 6 | 5.4286 | 0.0601 |
| 1, 3 | 3 | 5.4286 | 1.0865 |
| 1, 4 | 8 | 5.4286 | 1.218 |
| 2,1 | 1 | 1.7143 | 0.2975 |
| 2,2 | 5 | 5.4286 | 0.0338 |
| 2,3 | 6 | 5.4286 | 0.0601 |
| 2,4 | 6 | 5.4286 | 0.0601 |
| 3,1 | 0 | 1.7143 | 1.7143 |
| 3,2 | 5 | 5.4286 | 0.0338 |
| 3,3 | 8 | 5.4286 | 1.218 |
| 3,4 | 5 | 5.4286 | 0.0338 |
| 4,1 | 4 | 0.8571 | 11.5247 |
| 4,2 | 3 | 2.7143 | 0.0301 |
| 4,3 | 2 | 2.7143 | 0.188 |
| 4,4 | 0 | 2.7143 | 2.7143 |
| x² = | 20.5707 |

Table 3 Expected count for each cell

The x2 = 20.5707, the degree of freedom: (r-1) (c-1) = 9. The critical region could be finds in the Chi-Square Distribution table with degree of freedom equal to 2 and significant level of 0.10. The critical values are found to be approximately 16.919.



16.919

20.5707

Since the test statistic, x2 > critical value (20.5707 > 16.919), the test statistic, x2 fall within the critical region, therefore we reject the null hypothesis. There is insufficient evidence to show that the time spend on that activities is independent to the activities of students do during weekday nights.

## Anova

Anova method of testing is used to analyze for the significant differences between means. One-way ANOVA with equal sample size method is used and assumption likes the populations have normal distribution, same variance and samples are random and independent of each other is concerned. We used 0.05 significance level to test that the different types of activity have the same average time spent per night.

Hypothesis:

* H0: µ1 = µ2 = µ3
* H1: at least one mean is different

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Sample Size** | **Standard Deviation** | **Average Time Spent Per Night (Mean)** |
| **Study** | 18 | 1.150 | 3.50 |
| **Playing Video Game** | 18 | 0.924 | 3.17 |
| **Hang out with friends** | 18 | 0.984 | 2.65 |

Table 4 Mean number of population and standard deviation

Mean between sample= 3.50 + 3.17 + 2.65

3

 = 3.107

Variance between samples

= 18( (3.5-3.107) )² + (3.17-3.107) )² + (2.65-3.107) )² )

3-1

 = 3.305

Variance within sample

= (1.150) )² + (0.924) )² + (0.984)²

 3

= 1.048

Test statistic:

F **=** 3.305

 1.048

 = 3.154

Numerator and denominator:

* Numerator = 3 - 1= 2
* Denominator = 3(18-1) = 51

Critical value of F with a=0.05:

F critical value= 3.2389

Conclusion:

 Since F test statistic < F critical value (3.154 < 3.2389) , we fail to reject the null hypothesis, H0. There is sufficient evidence to claim that the different types of activity have the same mean of time spent per night by the students.

# conclusion

Based on the statistical data we can conclude that the time spent per night by students is dependent to the type of activity. Moreover we also can observe that every student have their different activity during weekday nights and sometime student can do more activities in a time. Besides that, the mean of duration of student spent for different activity is not the same has been proved in hypothesis testing. The relationship between type of activity and time spent per night can be proved from the chi-square analysis.

# Acknowledgment

We would like to express special thanks of gratitude for our lecturer (Assoc. Prof Dr Azlan bin Mohd Zain) who gave us an opportunity do this wonderful project on the topic (Student activities during weekday nights). Dr. also helped us a lot in doing Research and guide our group to complete this project.

# References

1. (n.d.). Retrieved from http://web.mnstate.edu/malonech/Psy%20230/Notes/Chi%20Square%20GW17.htm