

รายงาน

การวิเคราะห์ข้อมูลของการนอนหลับในตอนกลางคืน

จัดทำโดย

นายสิทธิกร	เฉลิมกิตติชัย	รหัส	640315
นายชิษณุชา	อัศรกุลพิชา	รหัส	640549
นายณัฐวีร์	วัฒนัยสมสุข	รหัส	640556
นายวรกานต์	สุเมธเสนีย์	รหัส	640586
นายชินวุธ	ตั้งชูสกุลชาติ	รหัส	641797

เสนอ

อาจารย์ นพมาศ อัศรจันทโชติ

รายงานฉบับนี้เป็นส่วนหนึ่งของรายวิชาคณิตศาสตร์และสถิติสำหรับปัญญาประดิษฐ์ AI2433

คณะวิทยาศาสตร์และเทคโนโลยี หลักสูตรปัญญาประดิษฐ์ ภาคเรียนที่ 1 ปีการศึกษา 2565

มหาวิทยาลัยหัวเฉียวเฉลิมพระเกียรติ

คำนำ

รายงานฉบับนี้เป็นส่วนหนึ่งของวิชาคณิตศาสตร์และสถิติสำหรับปัญญาประดิษฐ์โดยมีจุดประสงค์เพื่อนำความรู้ที่ได้จากการเรียนเรื่องสถิติพรรณามาใช้ และเพื่อวิเคราะห์การถดถอยแบบหลายตัวแปร ในรายงานฉบับนี้มีเนื้อหาซึ่งประกอบด้วยความรู้เกี่ยวกับกราฟตารางสถิติพรรณนาและสมการถดถอยเชิงเส้น จากการนำข้อมูล ในเว็บไซต์

คณะผู้จัดทำได้เลือกหัวข้อนี้ในการทำรายงานเนื่องจากเป็นเรื่องที่น่าสนใจนอกจากนี้การนอนหลับยังเป็นเรื่องสำคัญของมนุษย์ในการพักผ่อนชีวิตประจำวัน ทั้งนี้ผู้จัดทำต้อง ขอขอบคุณพระคุณอาจารย์ นพมาศ อัครจันทโชติ ผู้ให้ความรู้และแนวทางการศึกษาคณะผู้จัดทำหวังว่ารายงานฉบับนี้จะให้ความรู้และเป็นประโยชน์แก่ผู้อ่านทุกท่าน หากมีข้อผิดพลาดประการใด คณะผู้จัดทำต้องขออภัยมา ณ ที่นี้ด้วย

คณะผู้จัดทำ

7 พฤศจิกายน 2565

สารบัญ

เรื่อง	หน้า
คำนำ.....	ก
เนื้อเรื่อง.....	1-5
วัตถุประสงค์ของการศึกษา.....	1
การนำเสนอข้อมูลโดยกราฟและตาราง.....	1-2
การนำเสนอข้อมูลโดยสถิติพรรณนา.....	2-3
การนำเสนอข้อมูลโดยสมการถดถอยเชิงเส้น.....	3-4
สรุปผลการศึกษา.....	4-6
ภาคผนวก.....	7
อ้างอิง.....	8

เนื้อหา

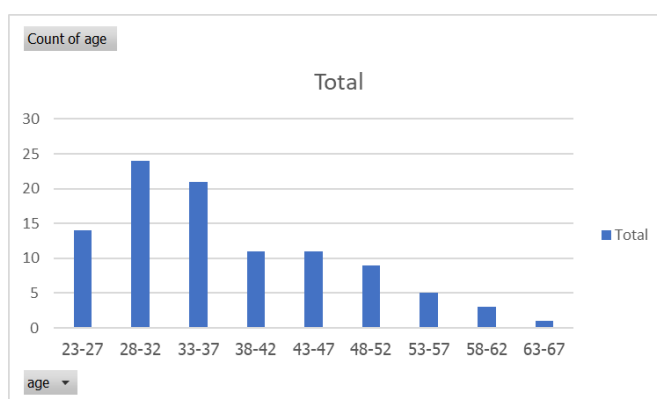
วัตถุประสงค์ของการศึกษา

- เพื่อศึกษาอายุของบุคคลมีผลต่อการนอนหลับในตอนกลางคืน
- เพื่อศึกษารายได้ต่อชั่วโมงของแต่ละบุคคลมีผลต่อการนอนหลับในตอนกลางคืน
- เพื่อศึกษาระดับการศึกษาของแต่ละบุคคลมีผลต่อการนอนหลับในตอนกลางคืน
- เพื่อศึกษาการนอนหลับระหว่างวันมีผลต่อการนอนหลับในตอนกลางคืน

การนำเสนอข้อมูลโดยกราฟและตาราง

1.อายุ

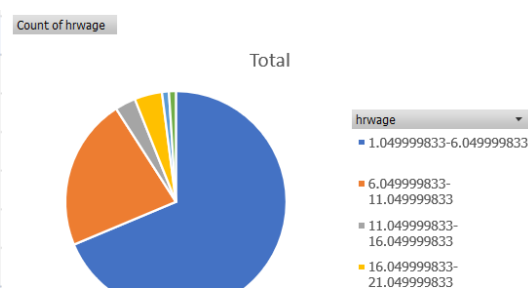
Row Labels	Count of age
23-27	14
28-32	24
33-37	21
38-42	11
43-47	11
48-52	9
53-57	5
58-62	3
63-67	1
Grand Total	99



จากตารางและกราฟของอายุ แสดงให้เห็นว่า มีจำนวนช่วงอายุ 28-32 ปี สูงที่สุด 24 คน และมีจำนวนช่วงอายุ 63-67 ปี ต่ำที่สุด 1 คน

2.รายได้ต่อชั่วโมงของแต่ละบุคคล

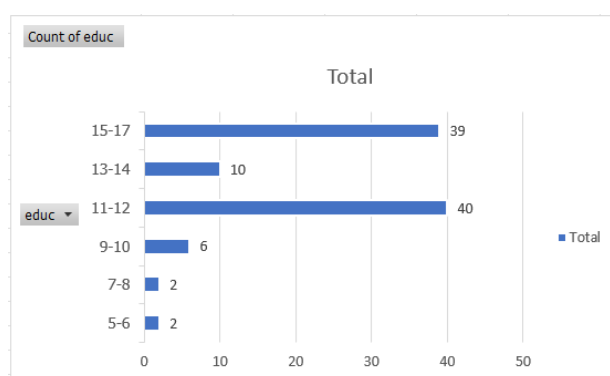
Row Labels	Count of hrwage
1.049999833-6.049999833	68
6.049999833-11.049999833	22
11.049999833-16.049999833	3
16.049999833-21.049999833	4
21.049999833-26.049999833	1
31.049999833-36.049999833	1
Grand Total	99



จากตารางและกราฟของรายได้ต่อชั่วโมงของแต่ละบุคคล แสดงให้เห็นว่าคนส่วนใหญ่ทำงานได้รายได้ต่อชั่วโมงอยู่ในช่วง 1.05 - 6.05 ดอลลาร์สหรัฐ และมีจำนวนเท่ากับ 68 คนจากทั้งหมด 99 คน

3.ระดับการศึกษาของแต่ละบุคคล

Row Labels	Count of educ
5-6	2
7-8	2
9-10	6
11-12	40
13-14	10
15-17	39
Grand Total	99

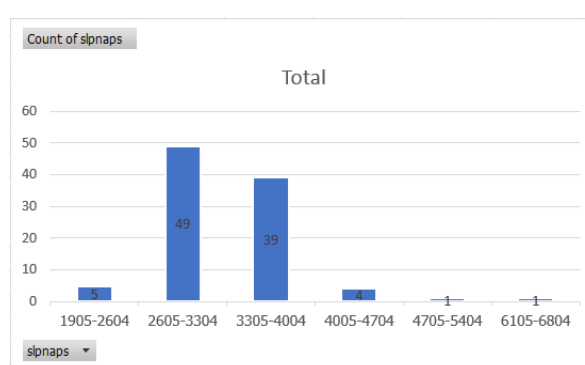


จากตารางและกราฟของระดับการศึกษาของแต่ละบุคคล แสดงให้เห็นว่าคนส่วนใหญ่

มีจำนวนปีการศึกษาอยู่ที่ 11-12 ปี และมีจำนวนใกล้เคียงกับ 15-17 ปี

4.การนอนหลับระหว่างวัน

Row Labels	Count of slnaps
1905-2604	5
2605-3304	49
3305-4004	39
4005-4704	4
4705-5404	1
6105-6804	1
Grand Total	99



จากตารางและกราฟของการนอนหลับระหว่างวัน แสดงให้เห็นว่าคนที่นอนหลับระหว่างวัน

มีจำนวน 49 คนเป็นจำนวนมากที่สุด อยู่ในช่วง 2,605-3,304 นาฬิกาต่อสัปดาห์ หรือ 6.2024 - 7.8667 ชั่วโมงต่อวัน

การนำเสนอข้อมูลโดยสถิติพรรณนา

sleep		age		educ		slnaps		hrwage	
Mean	3177.737374	Mean	37.23232323	Mean	13.47474747	Mean	3298.616162	Mean	6.134949617
Standard Error	45.79666906	Standard Error	1.005635356	Standard Error	0.271410888	Standard Error	55.22516401	Standard Error	0.518328164
Median	3173	Median	36	Median	12	Median	3248	Median	5.060002327
Mode	3105	Mode	23	Mode	12	Mode	2760	Mode	5.769999504
Standard Deviation	455.6711038	Standard Deviation	10.00594546	Standard Deviation	2.70050424	Standard Deviation	549.483444	Standard Deviation	5.157300115
Sample Variance	207636.1548	Sample Variance	100.1189445	Sample Variance	7.29272315	Sample Variance	301932.0552	Sample Variance	26.59774448
Kurtosis	1.600597388	Kurtosis	-0.29225919	Kurtosis	0.021273046	Kurtosis	7.307147807	Kurtosis	12.15088815
Skewness	-0.19772614	Skewness	0.630734396	Skewness	-0.458735654	Skewness	1.530037554	Skewness	3.044196758
Range	2738	Range	41	Range	12	Range	4205	Range	34.45998705
Minimum	1575	Minimum	23	Minimum	5	Minimum	1905	Minimum	1.049999833
Maximum	4313	Maximum	64	Maximum	17	Maximum	6110	Maximum	35.50998688
Sum	314596	Sum	3686	Sum	1334	Sum	326563	Sum	607.3600121
Count	99	Count	99	Count	99	Count	99	Count	99

1.อายุ

มีการแจกแจงปกติ เพราะมีความเบ้และความโด่งไม่เกิน ± 1 จึงเลือกใช้ค่าสถิติกลาง คือ ค่าเฉลี่ย มีค่า 37.2323 และมีเลือกใช้ค่าความแปรปรวน คือ ส่วนส่วนเบี่ยงเบนมาตรฐาน มีค่า 10.0059

2.รายได้ต่อชั่วโมงของแต่ละบุคคล

มีการแจกแจงไม่ปกติ เพราะมีความเบ้และความโด่งเกิน ± 1 จึงเลือกใช้ค่าสถิติกลาง คือ ค่ามัธยฐาน มีค่า 5.06 และมีเลือกใช้ค่าความแปรปรวน คือ IQR มีค่า 26.5977

3.ระดับการศึกษาของแต่ละบุคคล

มีการแจกแจงปกติ เพราะมีความเบ้และความโด่งไม่เกิน ± 1 จึงเลือกใช้ค่าสถิติกลาง คือ ค่าเฉลี่ย มีค่า 13.4747 และมีเลือกใช้ค่าความแปรปรวน คือ ส่วนส่วนเบี่ยงเบนมาตรฐาน มีค่า 2.7005

4.การนอนหลับระหว่างวัน

มีการแจกแจงไม่ปกติ เพราะมีความเบ้และความโด่งเกิน ± 1 จึงเลือกใช้ค่าสถิติกลาง คือ ค่ามัธยฐาน มีค่า 3,248 และมีเลือกใช้ค่าความแปรปรวน คือ IQR มีค่า 301,932.05

การนำเสนอข้อมูลโดยสมการถดถอยเชิงเส้น

SUMMARY OUTPUT								
Regression Statistics								
Multiple R	0.860011297							
R Square	0.739619431							
Adjusted R Square	0.728539407							
Standard Error	237.4132131							
Observations	99							
ANOVA								
	df	SS	MS	F	Significance F			
Regression	4	15050030	3762507	66.75251034	1.22187E-26			
Residual	94	5298313.174	56365.03					
Total	98	20348343.17						
	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	707.9257104	218.1271002	3.245473	0.00162505	274.8292397	1141.022181	274.8292397	1141.022181
age	-0.519158195	2.585385269	-0.2008	0.841284928	-5.652501252	4.614184863	-5.652501252	4.614184863
educ	6.819898791	9.747799655	0.699635	0.485883253	-12.53458593	26.17438351	-12.53458593	26.17438351
slpnaps	0.718146121	0.044564736	16.11467	8.54762E-29	0.629661794	0.806630449	0.629661794	0.806630449
hrwage	4.62209379	5.138793017	0.899451	0.370710745	-5.581100208	14.82528779	-5.581100208	14.82528779

$$\hat{Y} = 707.9257 - 0.5192x_1 + 6.8199x_2 + 0.7181x_3 + 4.6221x_4$$

สรุปผลการศึกษา

ทดสอบทั้งสมการ

$$H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0 ;$$

ตัวแปรไม่สามารถทำนายการนอนหลับในตอนกลางคืนได้

$$H_1 : \beta_i \neq 0 ; i=1,2,3,4;$$

มีอย่างน้อย 1 ตัวที่สามารถทำนายการนอนหลับในตอนกลางคืนได้

$$\alpha = 0.05$$

$$F = 66.7525$$

$$p\text{-value} = 1.22187E^{-26} < \alpha$$

ดังนั้น ปฏิเสธ H_0 สรุปได้ว่า มีอย่างน้อย 1 ตัวที่สามารถทำนายการนอนหลับในตอนกลางคืนได้

ทดสอบ β_1

$$H_0 : \beta_1 = 0 ;$$

อายุไม่สามารถทำนายการนอนหลับในตอนกลางคืนได้

$$H_1 : \beta_1 \neq 0 ;$$

อายุสามารถทำนายการนอนหลับในตอนกลางคืนได้

$$\alpha = 0.05$$

$$t = -0.2008$$

$$p\text{-value} = 0.8413 > \alpha$$

ดังนั้น ยังไม่ปฏิเสธ H_0 สรุปได้ว่า อายุไม่สามารถทำนายระยะเวลาการนอนหลับในตอนกลางคืนได้

ทดสอบ β_2

$$H_0 : \beta_2 = 0 ;$$

การศึกษาไม่สามารถทำนายการนอนหลับในตอนกลางคืนได้

$$H_1 : \beta_2 \neq 0 ;$$

การศึกษาสามารถทำนายการนอนหลับในตอนกลางคืนได้

$$\alpha = 0.05$$

$$t = 0.6996$$

$$p\text{-value} = 0.4859 > \alpha$$

ดังนั้น ยังไม่ปฏิเสธ H_0 สรุปได้ว่า การศึกษาไม่สามารถทำนายระยะเวลาการนอนหลับในตอนกลางคืนได้

ทดสอบ β_3

$$H_0 : \beta_3 = 0 ;$$

การนอนหลับระหว่างวันไม่สามารถทำนายการนอนหลับในตอนกลางคืนได้

$$H_1 : \beta_3 \neq 0 ;$$

การนอนหลับระหว่างวันสามารถทำนายการนอนหลับในตอนกลางคืนได้

$$\alpha = 0.05$$

$$t = 16.1146$$

$$p\text{-value} = 8.547 > \alpha$$

ดังนั้น ยังไม่ปฏิเสธ H_0 สรุปได้ว่า

การนอนหลับระหว่างวันไม่สามารถทำนายระยะเวลาการนอนหลับในตอนกลางคืนได้

ทดสอบ β_4

$$H_0 : \beta_4 = 0 ;$$

รายได้ต่อชั่วโมงไม่สามารถทำนายนอนหลับในตอนกลางคืนได้

$$H_1 : \beta_4 \neq 0 ;$$

รายได้ต่อชั่วโมงสามารถทำนายนอนหลับในตอนกลางคืนได้

$$\alpha = 0.05$$

$$t = 0.8994$$

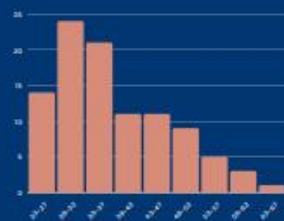
$$p\text{-value} = 0.3707 > \alpha$$

ดังนั้น ยังไม่ปฏิเสธ H_0 สรุปได้ว่า

รายได้ต่อชั่วโมงไม่สามารถทำนายระยะเวลาการนอนหลับในตอนกลางคืนได้

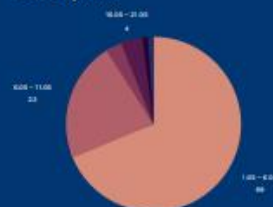
การวิเคราะห์ข้อมูลของการนอนหลับในตอนกลางคืน

อายุ



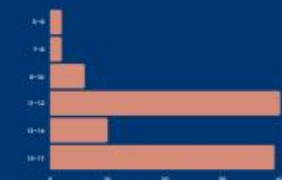
กราฟของอายุ แสดงให้เห็นว่า มีจำนวนช่วงอายุ 28-32 ปี สูงที่สุด 24 คน และมีจำนวนช่วงอายุ 3-8 ปี ต่ำที่สุด 1 คน

รายได้ต่อชั่วโมงของแต่ละบุคคล



ส่วนใหญ่ทำงานได้รายได้ต่อชั่วโมงอยู่ในช่วง 1.05 - 8.05 ดอลลาร์สหรัฐ และมีจำนวนเท่ากับ 68 จากทั้งหมด 99 คน

ระดับการศึกษาของแต่ละบุคคล



ส่วนใหญ่มีจำนวนปีการศึกษาอยู่ที่ 11-12 และมีจำนวนใกล้เคียงกับ 13-17 ปี

การนอนหลับระหว่างวัน



คนที่นอนหลับระหว่างวันมีจำนวน 49 คนเป็นจำนวนมากที่สุด อยู่ในช่วง 2805 - 3304 นาทีต่อสัปดาห์ หรือ 6.2024 - 7.8867 ชั่วโมงต่อวัน

ข้อมูลโดยสมการถดถอยเชิงเส้น

SUMMARY OUTPUT									
Regression Statistics									
Multiple R	0.890012297								
R Square	0.796029431								
Adjusted R Square	0.726539407								
Standard Error	237.4132131								
Observations	99								
ANOVA									
	df	SS	MS	F	Significance F				
Regression	4	15050230	3762557	66.75251034	1.22187E-26				
Residual	94	5298313.174	56365.03						
Total	98	20348543.17							
	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%	
Intercept	707.6057104	238.1271902	2.945473	0.00162505	274.8203397	1141.022181	274.8203397	1141.022181	
age	-0.519158195	2.505385289	-0.2008	0.841284028	-5.052501252	4.014194063	-5.052501252	4.014194063	
รายได้	6.818698791	5.747799615	0.899635	0.485803253	-12.53458593	26.17438351	-12.53458593	26.17438351	
การนอน	0.716146121	0.044644736	16.111467	0.547620E-39	0.620660794	0.806030449	0.620660794	0.806030449	
Income	4.62208370	5.130793517	0.899451	0.372710745	-5.501106208	14.82528779	-5.501106208	14.82528779	

สรุปผลการศึกษา

มีเพียงตัวแปรของการนอนหลับระหว่างวันที่สามารถทำนายการนอนหลับในตอนกลางคืนได้ เมื่อมีตัวแปรอื่นๆ มีค่าคงที่

ภาคผนวก

wooldridge	openness	openness	114	12	2	0	0	0	12	CSV DOC
wooldridge	pension	pension	194	19	13	0	0	0	19	CSV DOC
wooldridge	phillips	phillips	56	7	0	0	0	0	7	CSV DOC
wooldridge	pntrsprd	pntrsprd	553	12	6	0	0	0	12	CSV DOC
wooldridge	prison	prison	714	45	16	0	0	0	45	CSV DOC
wooldridge	prminwge	prminwge	38	25	0	0	0	0	25	CSV DOC
wooldridge	rdchem	rdchem	32	8	0	0	0	0	8	CSV DOC
wooldridge	rdtelec	rdtelec	29	6	0	0	0	0	6	CSV DOC
wooldridge	recid	recid	1445	18	10	0	0	0	18	CSV DOC
wooldridge	rental	rental	128	23	2	0	0	0	23	CSV DOC
wooldridge	return	return	142	12	0	0	0	0	12	CSV DOC
wooldridge	saving	saving	100	7	1	0	0	0	7	CSV DOC
wooldridge	school93_98	school93_98	10668	18	6	0	0	0	18	CSV DOC
wooldridge	sleep75	sleep75	706	34	12	0	0	0	34	CSV DOC
wooldridge	slp75_81	slp75_81	239	20	7	0	0	0	20	CSV DOC
wooldridge	smoke	smoke	807	10	2	0	0	0	10	CSV DOC
wooldridge	traffic1	traffic1	51	13	7	1	0	0	12	CSV DOC
wooldridge	traffic2	traffic2	108	48	13	0	0	0	48	CSV DOC
wooldridge	twoyear	twoyear	6763	23	15	0	0	0	23	CSV DOC
wooldridge	volat	volat	558	17	0	0	0	0	17	CSV DOC
wooldridge	vote1	vote1	173	10	1	1	0	0	9	CSV DOC
wooldridge	vote2	vote2	186	26	5	1	0	0	25	CSV DOC
wooldridge	voucher	voucher	990	19	13	0	0	0	19	CSV DOC
wooldridge	wage1	wage1	526	24	16	0	0	0	24	CSV DOC
wooldridge	wage2	wage2	935	17	4	0	0	0	17	CSV DOC

Excel PivotTable Fields task pane showing the 'age' field selected for the 'Sum of age' calculation.

PivotTable Fields

Choose fields to add to report:

- ☐ sleep
- ☒ age
- ☐ educ
- ☐ slpnaps
- ☐ hrwage

More Tables...

Drag fields between areas below:

Filters

Columns

Rows: age

Σ Values: Sum of age

☐ Defer Layout Update

Update

Excel Regression dialog box showing the input ranges and output options.

Regression

Input

Input Y Range: \$A\$1:\$A\$100

Input X Range: \$B\$1:\$E\$100

☒ Labels

☐ Constant is Zero

Confidence Level: 95 %

Output options

☐ Output Range:

☒ New Worksheet Ply:

☐ New Workbook

Residuals

☐ Residuals

☐ Standardized Residuals

☐ Residual Plots

☐ Line Fit Plots

Normal Probability

☐ Normal Probability Plots

OK

Cancel

Help

อ้างอิง

<https://vincentarelbundock.github.io/Rdatasets/datasets.html>