



CASE STUDY

Data Analytics and Data Visualization Module



"ผลกระทบของการแพร่ระบาด

COVID-19

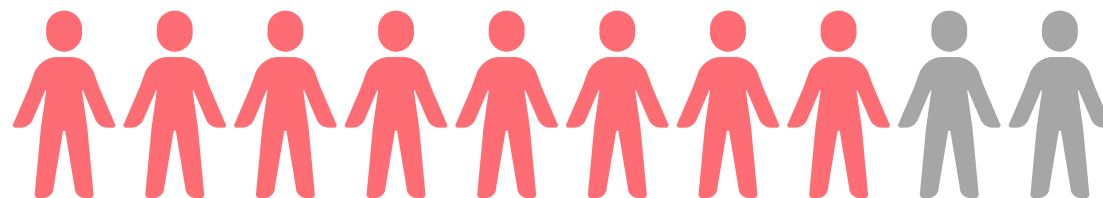
**ต่อปัญหาสุขภาพระหว่างช่วงล็อกดาวน์
ของนักเรียน"**



วัตถุประสงค์

1. เพื่อสำรวจอายุของคนในพื้นที่ ที่มีปัญหาสุขภาพระหว่างช่วงล็อกดาวน์ในสถานการณ์การแพร่ระบาดของ COVID-19
2. เพื่อศึกษาความสัมพันธ์ระหว่างพฤติกรรมการใช้เวลาว่างบนโซเชียลมีเดียที่ส่งผลต่อปัญหาสุขภาพของนักเรียนในช่วงล็อกดาวน์
3. เพื่อศึกษาปัจจัยทางสังคมและพฤติกรรมที่เกี่ยวข้องกับการเรียนออนไลน์ที่ส่งผลต่อปัญหาสุขภาพของนักเรียนในช่วงล็อกดาวน์

วิธีดำเนินการ



แหล่งที่มาของข้อมูล (SOURCES OF DATA)

ใช้โปรแกรม R ในการวิเคราะห์ข้อมูลจากไฟล์ COVID-19 Survey Student Responses.csv

ตัวแปรที่สนใจและใช้ในการวิเคราะห์ (Variables for Analysis)

ตัวแปรอิสระ		
ชื่อตัวแปร	คำอธิบายชื่อตัวแปร	ประเภทตัวแปร
Region of residence	ภูมิภาคที่อยู่อาศัย	ตัวแปรกลุ่ม
Age of Subject	อายุของนักเรียน	ตัวแปรต่อเนื่อง
Time spent on Online Class	เวลาที่ใช้เรียนออนไลน์ (ชั่วโมง)	ตัวแปรต่อเนื่อง
Time spent on social media	เวลาที่ใช้ในโซเชียลมีเดีย (ชั่วโมง)	ตัวแปรต่อเนื่อง
Number of meals per day	จำนวนมื้อต่อวัน	ตัวแปรต่อเนื่อง

ตัวแปรตาม		
ชื่อตัวแปร	คำอธิบายชื่อตัวแปร	ประเภทตัวแปร
Health issue during lockdown	ปัญหาสุขภาพระหว่างช่วงล็อกดาวน์ Yes = 1 , No = 0	ตัวแปรกลุ่ม

การทำความสะอาดข้อมูล (DATA CLEANSING)

```
summary(dt)
```

ตรวจสอบค่าขาดหาย (Missing values)

```
missing <- colSums(is.na(dt))
```

```
missing
```

#การตรวจสอบข้อมูลซ้ำ

```
dt <- dt[!duplicated(dt), ]
```

```
head(dt)
```

ตรวจสอบอีกครั้งว่าไม่มีค่าที่หายไปแล้ว

```
sum(is.na(dt$Rating.of.Online.Class.experience))
```

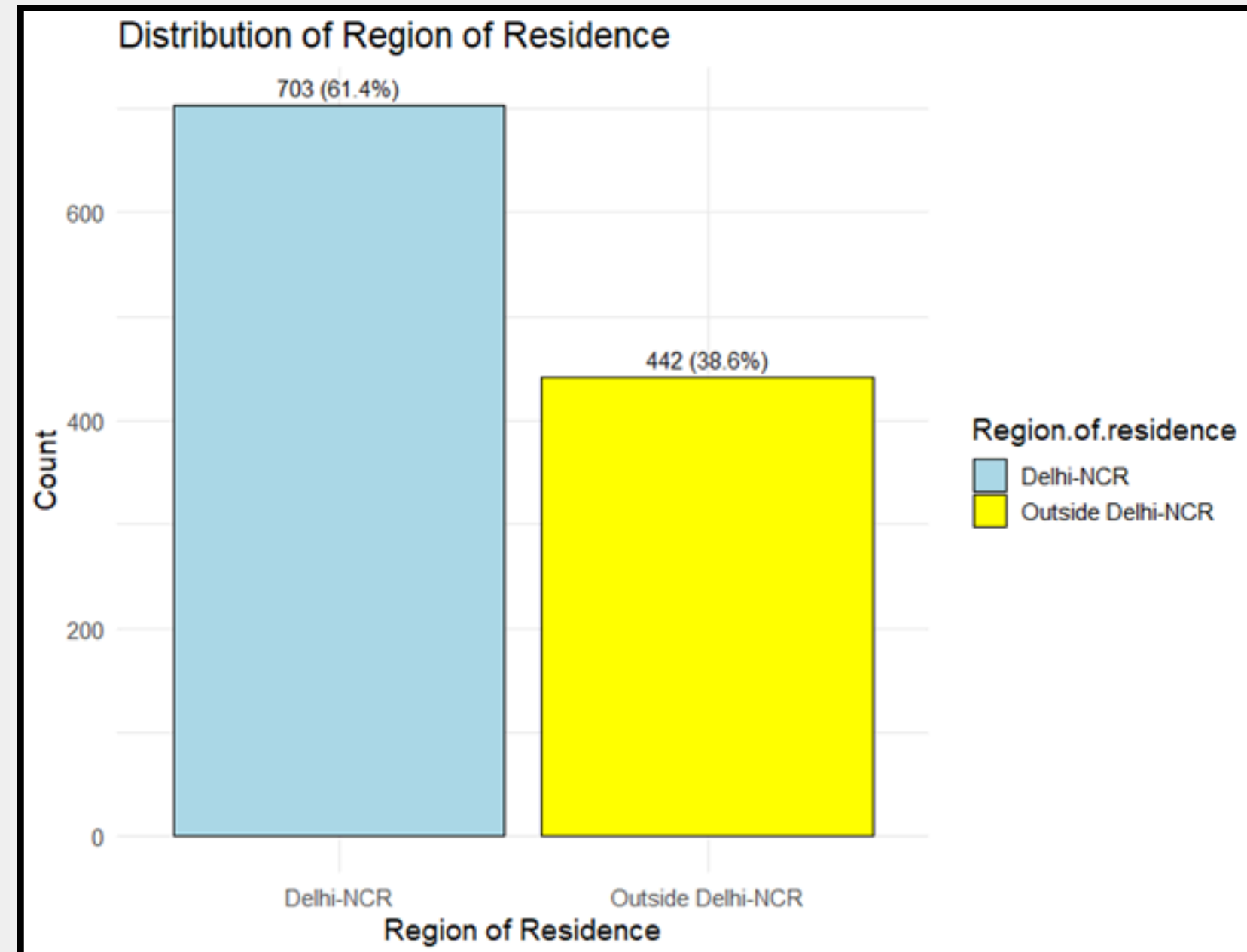
แปลงจากตัวแปรประเภทหมวดหมู่ (เช่น "0", "1") ซึ่งจะช่วยในการวิเคราะห์ทางสถิติที่เกี่ยวข้องกับ ตัวแปรหมวดหมู่

```
dt$Health.issue.during.lockdown <- as.factor(dt$Health.issue.during.lockdown)
```

ผลการวิเคราะห์จากการจัดการข้อมูล

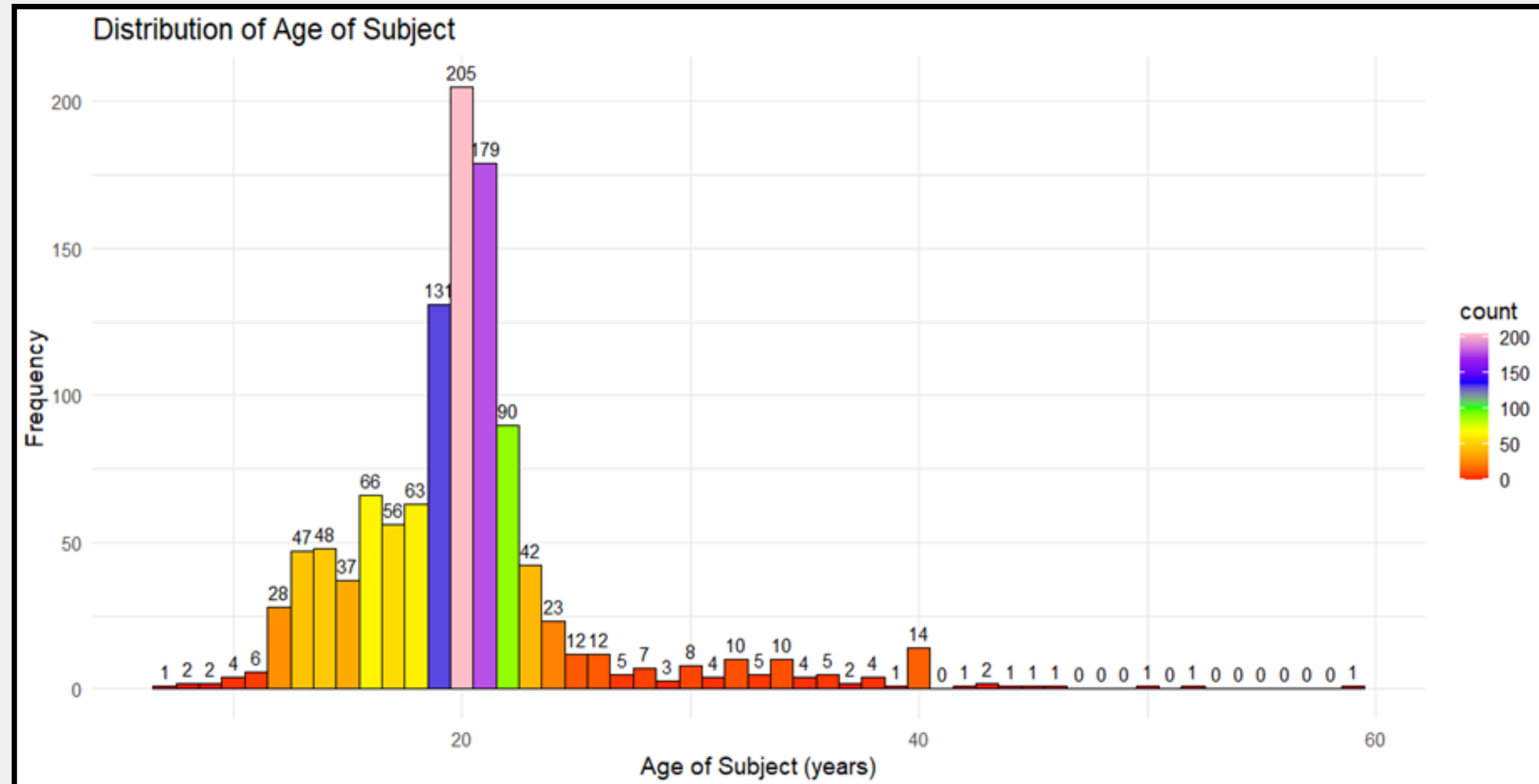
```
> str(dt)
'data.frame': 1145 obs. of 21 variables:
 $ ID                                     : chr  "R1" "R2" "R3" "R4" ...
 $ Region.of.residence                   : Factor w/ 2 levels "Delhi-NCR","Outside Delhi-$
 $ Age.of.Subject                         : num  21 21 20 20 21 21 19 19 21 20 ...
 $ Time.spent.on.Online.Class             : num  2 0 7 3 3 0 2 2 3 0 ...
 $ Rating.of.Online.Class.experience      : chr   "Good" "Excellent" "Very poor" "Very poor$
 $ Medium.for.online.class                : chr   "Laptop/Desktop" "Smartphone" "Laptop/Des$
 $ Time.spent.on.self.study               : num  4 0 3 2 3 6 2 1 4 1 ...
 $ Time.spent.on.fitness                  : num  0 2 0 1 1 0 1 1 1 0.5 ...
 $ Time.spent.on.sleep                    : num  7 10 6 6 8 5 5 10 8 8 ...
 $ Time.spent.on.social.media             : num  3 3 2 5 3 1 4 5 2 5 ...
 $ Preferred.social.media.platform        : chr   "Linkedin" "Youtube" "Linkedin" "Instagra$
 $ Time.spent.on.TV                       : chr   "1" "0" "0" "0" ...
 $ Number.of.meals.per.day                : int   4 3 3 3 4 1 3 3 3 3 ...
 $ Change.in.your.weight                  : chr   "Increased" "Decreased" "Remain Constant"$
 $ Health.issue.during.lockdown           : Factor w/ 2 levels "0","1": 1 1 1 1 1 2 1 2 1 $
 $ Stress.busters                         : chr   "Cooking" "Scrolling through social media$
 $ Time.utilized                          : chr   "YES" "YES" "NO" "NO" ...
 $ Do.you.find.yourself.more.connected.with.your.family..close.friends...relatives...: chr   "YES" "NO" "YES" "NO" ...
 $ What.you.miss.the.most                 : chr   "School/college" "Roaming around freely" $
 $ SocialMediaTime_binned                 : Factor w/ 6 levels "0-1","1-2","2-3",...: 3 3 2$
 $ Age.group                              : Factor w/ 5 levels "15-18","19-21",...: 2 2 2 2$
```

ผลการวิเคราะห์



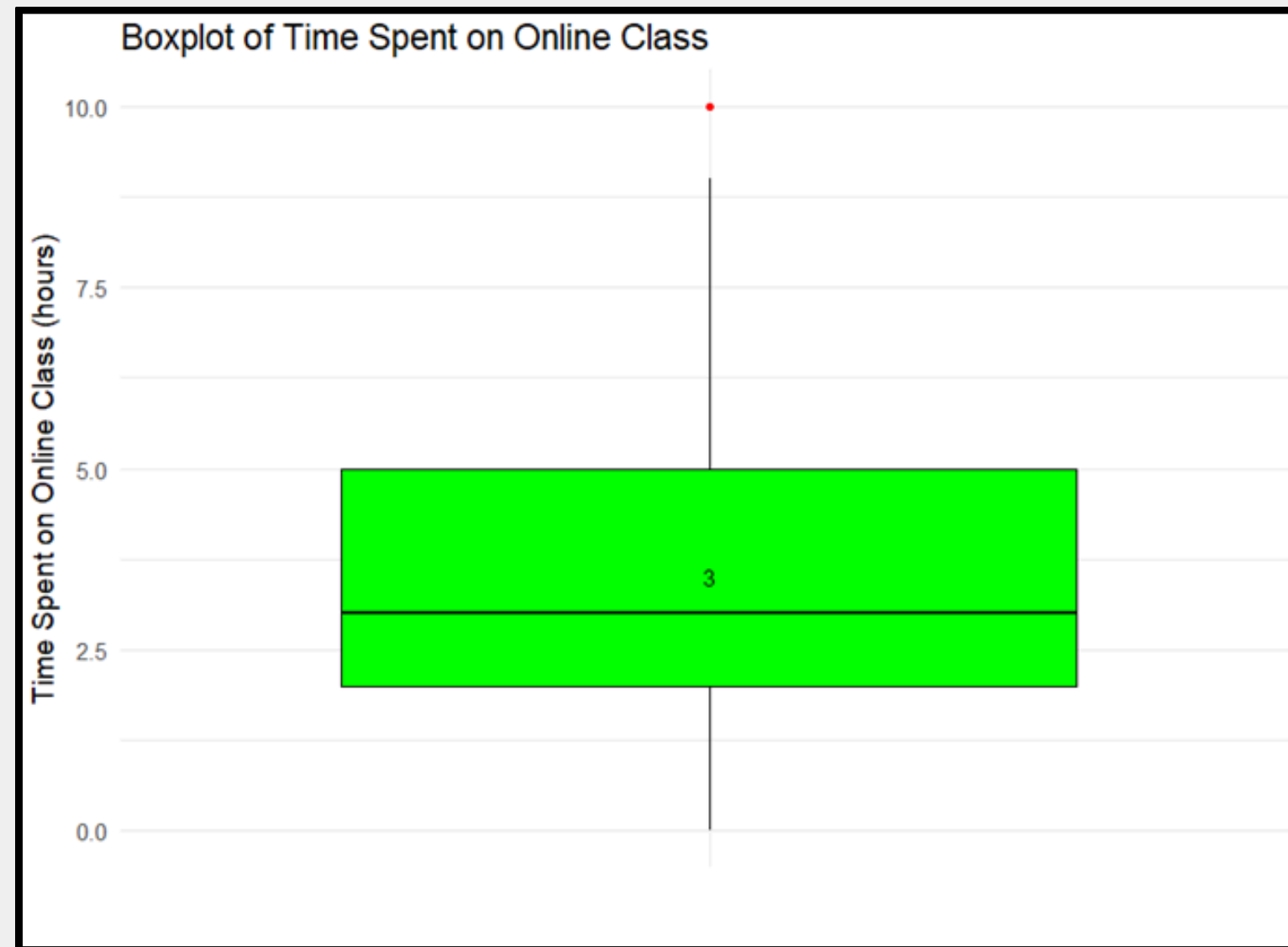
จากกราฟ ผู้ที่อาศัยอยู่ใน Delhi-NCR มีจำนวนมากกว่า คิดเป็น 61.4% และผู้ที่อาศัยอยู่นอก Delhi-NCR คิดเป็น 38.6% ของจำนวนทั้งหมด

ผลการวิเคราะห์



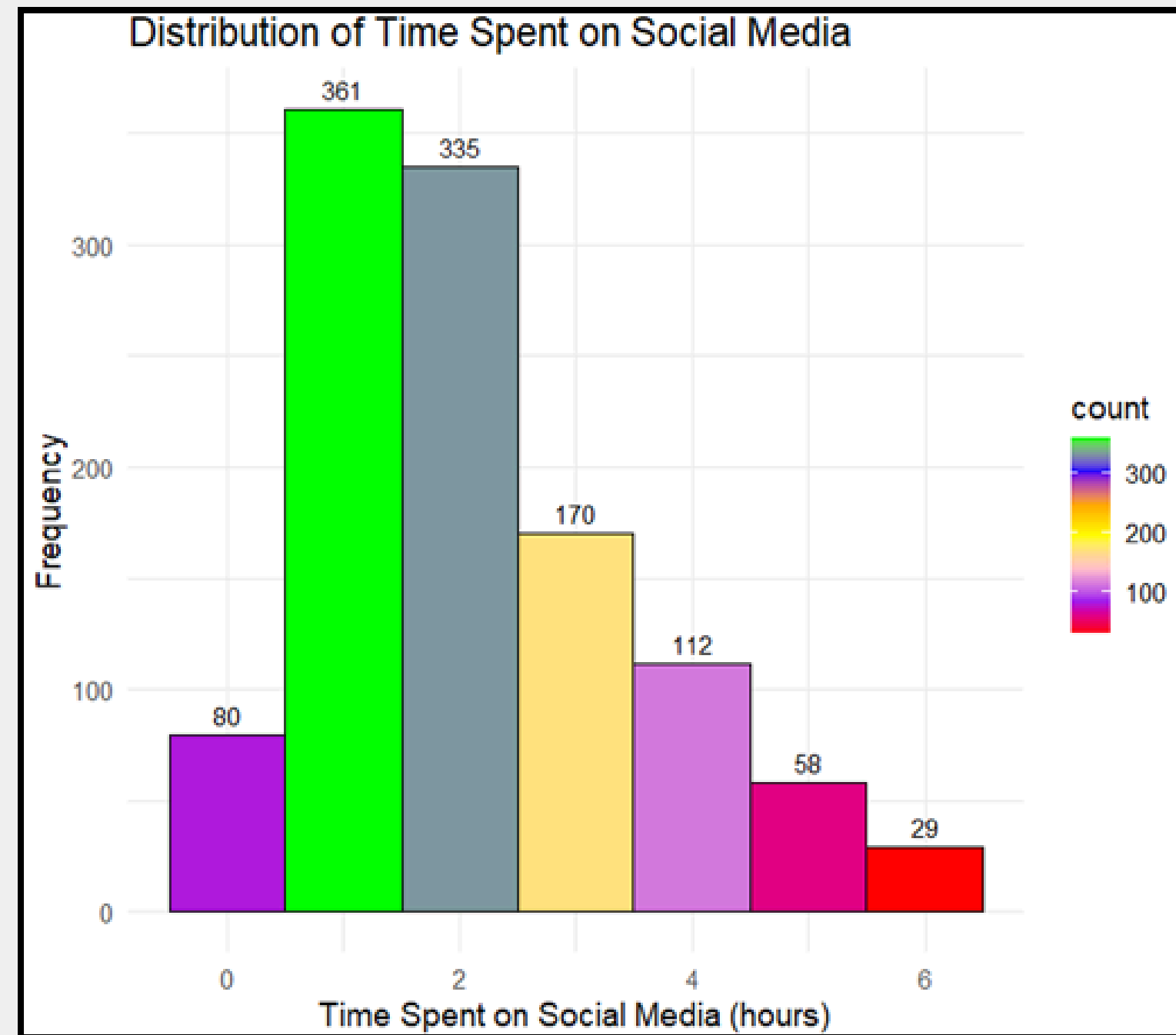
จากกราฟ มีผู้ที่มีอายุอยู่ในช่วง 15–25 ปี มีความสูงที่สุด และจำนวนของผู้ที่มีอายุเกิน 40 ปีมีจำนวนน้อยมากเมื่อเทียบกับกลุ่มอายุอื่น

ผลการวิเคราะห์



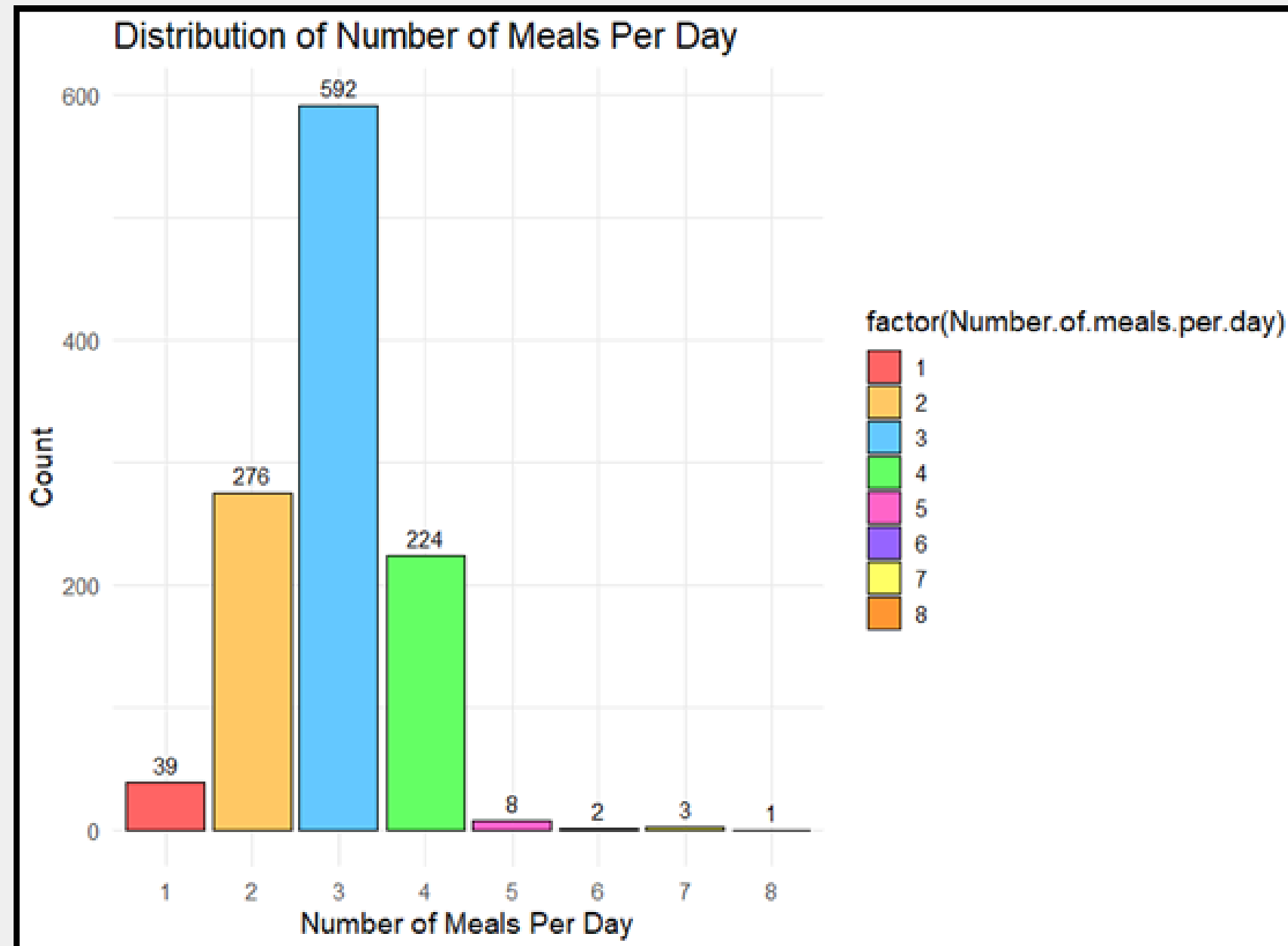
จากกราฟ นักเรียนส่วนใหญ่ใช้เวลาเรียนออนไลน์ประมาณ 2.5 – 5 ชั่วโมง และมึนักเรียนบางคนที่ใช้เวลาเรียนออนไลน์มากกว่า 7.5 ชั่วโมง

ผลการวิเคราะห์



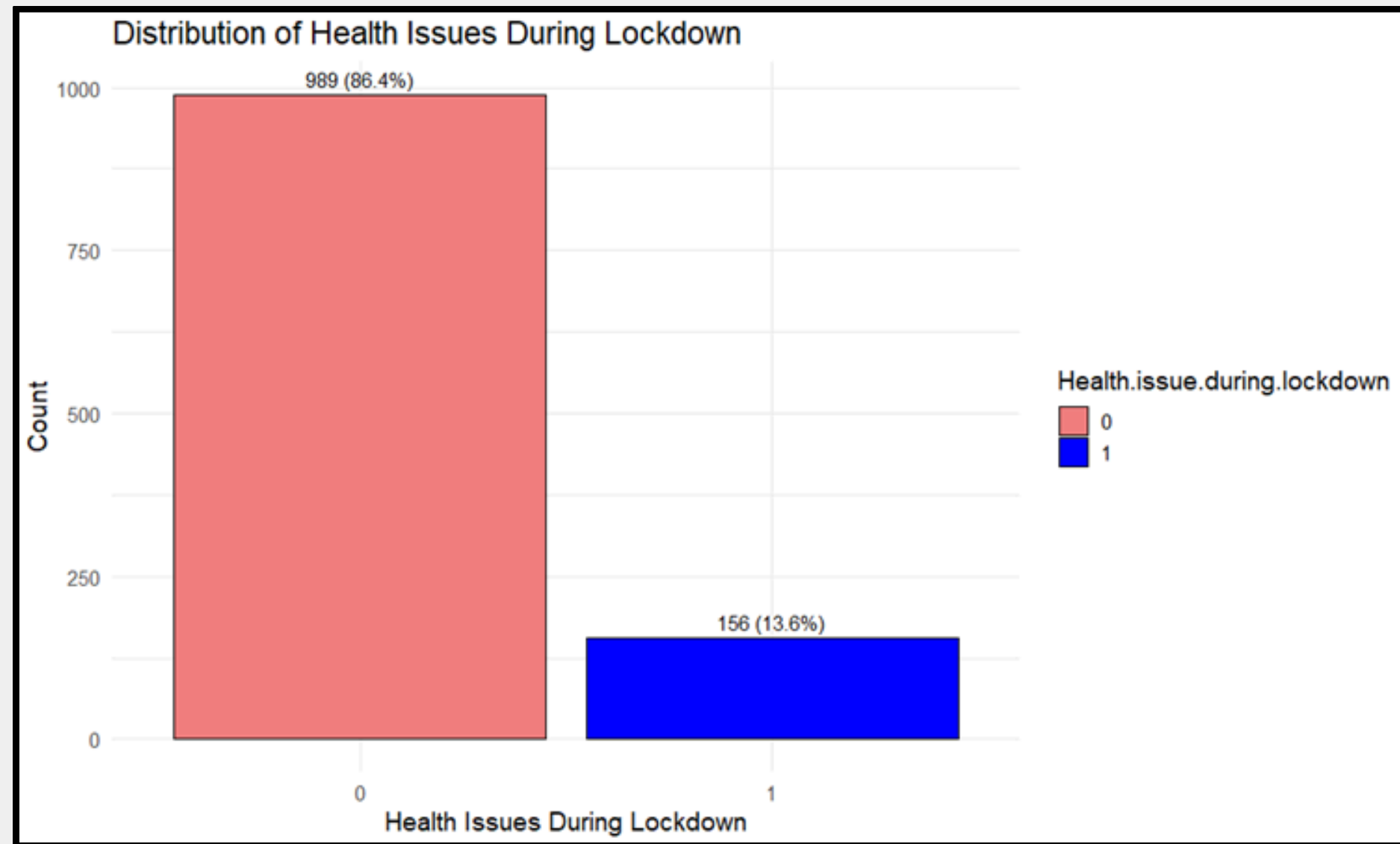
จากกราฟ ส่วนใหญ่ใช้เวลาบนโซเชียลมีเดียในช่วงเวลาประมาณ 2 ชั่วโมงและมีผู้คนบางส่วนที่ใช้เวลาบนโซเชียลมีเดียมากกว่า 4 ชั่วโมงต่อวัน แต่ก็ยังมีผู้คนบางส่วนที่ใช้เวลาบนโซเชียลมีเดียน้อยกว่า 1 ชั่วโมงต่อวัน

ผลการวิเคราะห์



จากกราฟ จะเห็นได้ว่ามีผู้คนจำนวนมากที่สุดที่รับประทานอาหาร 3 มื้อต่อวัน และมีบางส่วนที่รับประทานอาหารน้อยกว่า 3 มื้อ หรือมากกว่า 3 มื้อ แต่มีจำนวนน้อยมากที่รับประทานอาหารมากกว่า 5 มื้อต่อวัน

ผลการวิเคราะห์



จากกราฟ ส่วนใหญ่ไม่มีปัญหาสุขภาพในช่วงล็อกดาวน์คิดเป็น 86.4% และส่วนน้อยประสบปัญหาสุขภาพในช่วงล็อกดาวน์ คิดเป็น 13.6% ของจำนวนทั้งหมด

ผลวิเคราะห์ความสัมพันธ์ระหว่างตัวแปร อิสระและตัวแปรตาม

ตัวแปร (Variable)	NO (ไม่มี)	YES (มี)	Test Stat.	P-value
Total	989	156	-	-
Region of residence	Delhi-NCR 619 (62.6)	Delhi-NCR 84 (53.8)	Chisq. (1 df) =4.35	0.037
	Outside Delhi-NCR 370 (37.4)	Outside Delhi-NCR 72 (46.2)		
Age of Subject	median(IQR) = 20 (17,21)	median(IQR) = 20 (19,21)	Ranksum test	< 0.001
Time spent on Online Class	median(IQR) = 3 (2,5)	median(IQR) = 2 (1,4)	Ranksum test	< 0.001
Time spent on social media	median(IQR) = 2 (1,3)	median(IQR) = 2 (1,3)	Ranksum test	0.885
Number of meals per day	median(IQR) = 3 (2,3)	median(IQR) = 3 (2,3)	Ranksum test	0.001



สรุปผลวิเคราะห์ความสัมพันธ์ระหว่างตัวแปร อิสระและตัวแปรตาม

ตัวแปรอิสระ	ค่า P-value	ความสัมพันธ์
Region of residence	0.037	มีนัยสำคัญ
Age of Subject	< 0.001	มีนัยสำคัญ
Time spent on Online Class	< 0.001	มีนัยสำคัญ
Time spent on social media	0.885	ไม่มีนัยสำคัญ
Number of meals per day	0.001	มีนัยสำคัญ

จากการวิเคราะห์ ตัวแปรมีความสัมพันธ์กับ ตัวแปรตาม (Health issue during lockdown) 4 ตัว คือ Region of residence , Age of Subject , Time spent on Online Class , Number of meals per day เพราะว่าค่า p-value <0.05

การวิเคราะห์การถดถอยพหุลอจิสติก (MULTIPLE LOGISTIC REGRESSION)

จากผลการวิเคราะห์ จำนวนข้อมูลทั้งหมด 1145 ค่า จำนวนตัวแปร 6 ตัว

ตัวแปรต้น x คือ: Region of residence , Age of Subject , Time spent on Online Class ,
Time spent on social media , Number of meals per day ตัวแปรตาม y ที่สนใจ คือ Health
issue during lockdown

ตารางแสดงชื่อของตัวแปร ค่า frequency และ percent

Health issue during lockdown	Frequency	Percent
Yes	156	13.62
No	989	86.38
Total	1145	100

การคัดเลือกตัวแปรไว้ในตัวแบบ

การคัดเลือกตัวแบบวิธีถอยหลัง (backward selection)

ผลการวิเคราะห์ข้อมูล

ขั้นตอนการคัดเลือก	ตัวแปรที่ถูกตัดออก	ค่า AIC
1	Age of Subject	891.40
2	Time spent on social media	892.94

จากการวิเคราะห์ข้อมูล โดยใช้วิธีการคัดเลือกตัวแปรแบบถอยหลัง (backward selection) เราพบว่าตัวแปรอิสระบางตัวควรตัดออกจากตัวแบบเนื่องจากไม่มีนัยสำคัญทางสถิติในการทำนายปัญหาสุขภาพในช่วงล็อกดาวน์ (Health issue during lockdown) ตัวแปรที่ควรตัดออกได้แก่ Age of Subject , Time spent on social mediaสามารถตัดตัวแปร and ออกจากตัวแบบ โดยพิจารณาจากค่า AIC จากตัวแปร พบว่าหากตัดตัวแปรข้างต้นนี้ออก จะทำให้ค่า AIC ลดลง



การสร้างตัวแบบสุดท้าย (FINAL MODEL)

การแปลผลตัวแบบสุดท้าย (final model)

ผลการวิเคราะห์ข้อมูล

ตัวแปรอิสระ	Estimate (β)	Std. Error	z-value	p-value
(Intercept)	-0.740219	0.487415	-1.519	0.128847
Region of residence	0.309762	0.175873	1.761	0.078190
Age of Subject	0.010324	0.015025	0.687	0.491996
Time spent on Online Class	-0.152246	0.045470	-3.348	0.000813
Time spent on social media	-0.003673	0.063847	0.058	0.954129
Number of meals per day	-0.348371	0.109901	-3.170	0.001525

สมการสุดท้ายของโมเดลถดถอยพหุอสมติที่ทำนายปัญหาสุขภาพระหว่างล็อกดาวน์ คือ

$$\text{Logit}(\pi) = (-0.152) (\text{Time spent on Online Class}) + (-0.348) (\text{Number of meals per day})$$

การแปลผลตัวแบบสุดท้าย (FINAL MODEL)

การแปลผลตัวแบบสุดท้าย (final model)

ตัวแปร	ค่า Crude OR (95% CI)	ค่า Adjusted OR (95% CI)	ค่า P (Wald's Test)	ค่า P (LR Test)
Region of Residence: Outside Delhi-NCR vs Delhi-NCR	1.43 (1.02, 2.01)	1.36 (0.97, 1.92)	0.078	0.079
Age of Subject	1.02 (0.99, 1.05)	1.01 (0.98, 1.04)	0.492	0.498
Time Spent on Online Class	0.85 (0.78, 0.93)	0.86 (0.79, 0.94)	< 0.001	< 0.001
Time Spent on Social Media	1.0183 (0.9017, 1.1499)	0.9963 (0.8791, 1.1291)	0.954	0.954
Number of Meals per Day	0.7 (0.56, 0.86)	0.71 (0.57, 0.88)	0.002	0.001

ผลสรุปการแปลผลตัวแบบสุดท้าย (FINAL MODEL)

ข้อมูลของโมเดล :

Log-Likelihood: -441.468

จำนวนตัวอย่าง: 1145

ค่า AIC: 894.9375

อธิบาย :

ภูมิภาคที่อยู่อาศัย แสดงความสำคัญที่ใกล้เคียงกับนัยสำคัญทางสถิติ (ค่า $p \sim 0.078-0.079$) โดยผู้ที่อาศัยอยู่นอก Delhi-NCR มีโอกาสพบปัญหาสุขภาพสูงกว่าผู้ที่อยู่ใน Delhi-NCR เล็กน้อย

เวลาที่ใช้ในคลาสเรียนออนไลน์ มีความสัมพันธ์เชิงลบอย่างมีนัยสำคัญกับปัญหาสุขภาพ (ค่า $p < 0.001$) แสดงว่าเวลาที่ใช้ในการเรียนออนไลน์มากขึ้นสัมพันธ์กับปัญหาสุขภาพที่ลดลง

จำนวนมื้อต่อวัน มีความสัมพันธ์กับปัญหาสุขภาพอย่างมีนัยสำคัญ (ค่า $p = 0.002$) โดยการรับประทานอาหารน้อยมื้อต่อวันสัมพันธ์กับการเกิดปัญหาสุขภาพที่สูงขึ้น

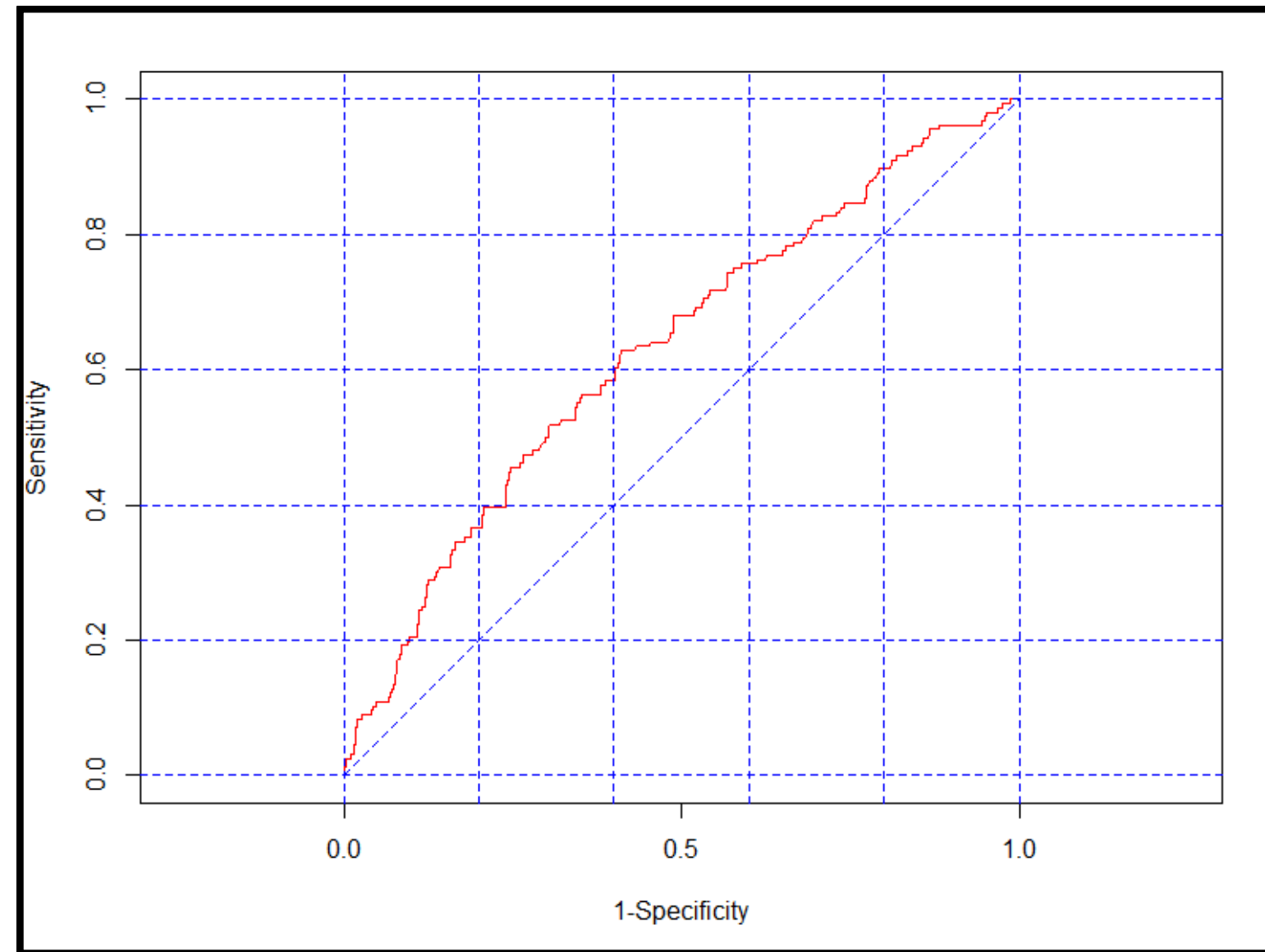
อายุ และ เวลาที่ใช้ในโซเชียลมีเดีย ไม่มีความสัมพันธ์อย่างมีนัยสำคัญกับปัญหาสุขภาพในโมเดลที่ปรับค่าแล้ว

ผลการวิเคราะห์พบว่า

จากตัวแปร Time spent on Online Class การใช้เวลาเรียนออนไลน์ มีความสัมพันธ์กับการเกิดปัญหาสุขภาพในช่วงล็อกดาวน์ โดยพบว่าผู้ที่ใช้เวลาเรียนออนไลน์มากขึ้นมีความเสี่ยงต่อการเกิดปัญหาสุขภาพลดลง 0.86 เท่า เมื่อเทียบกับผู้ที่ใช้เวลาเรียนน้อยลง

นอกจากนี้ Number of meals per day จำนวนมื้อต่อวัน ก็มีความสัมพันธ์กับการลดความเสี่ยงต่อการเกิดปัญหาสุขภาพเช่นกัน โดยการรับประทานอาหารมากขึ้นสัมพันธ์กับการลดความเสี่ยงต่อปัญหาสุขภาพในช่วงล็อกดาวน์ 0.71 เท่า เมื่อเทียบกับผู้ที่รับประทานอาหารน้อยลง

การทดสอบความกลมกลืนของตัวแบบ



จากกราฟได้ค่า $AUC = 0.629$ แปลผล คือ

ตัวแบบที่ใช้ในการทำนายปัญหาสุขภาพระหว่างช่วงล็อกดาวน์มีความสามารถในการทำนายได้พอสมควร แต่ยังไม่ถือว่าดีมากเมื่อเทียบกับแบบจำลองที่มีประสิทธิภาพสูง ค่า AUC ดังกล่าวบ่งชี้ว่าแบบจำลองนี้สามารถแยกแยะผู้ที่มีปัญหาสุขภาพออกจากผู้ที่ไม่มีปัญหาได้ในระดับที่ปานกลาง (ประมาณ 62.9% ของเวลาที่แบบจำลองสามารถทำนายได้ถูกต้อง)

ผลการศึกษา

การศึกษานี้มุ่งเน้นวิเคราะห์ผลกระทบของการแพร่ระบาด COVID-19 ต่อสุขภาพของนักเรียนในช่วงล็อกดาวน์ โดยใช้ตัวแปรต่างๆ เช่น เวลาที่ใช้ในการเรียนออนไลน์, การใช้โซเชียลมีเดีย, จำนวนมื้ออาหารในแต่ละวัน รวมถึงภูมิภาคที่นักเรียนพักอาศัย เพื่อประเมินความเสี่ยงในการเกิดปัญหาสุขภาพ จากผลการวิเคราะห์ logistic regression พบว่า:

- เวลาที่ใช้ในการเรียนออนไลน์ (Time spent on Online Class) มีความสัมพันธ์กับการลดความเสี่ยงต่อปัญหาสุขภาพ โดยนักเรียนที่ใช้เวลาเรียนออนไลน์มากขึ้นมีโอกาสเกิดปัญหาสุขภาพลดลง 0.86 เท่า (OR = 0.86, 95% CI = 0.79, 0.94, $p < 0.001$)
- จำนวนมื้ออาหารต่อวัน (Number of meals per day) มีความสัมพันธ์เชิงบวกกับการลดความเสี่ยงต่อการเกิดปัญหาสุขภาพ โดยนักเรียนที่รับประทานอาหารมากขึ้นมีความเสี่ยงลดลง 0.71 เท่า (OR = 0.71, 95% CI = 0.57, 0.88, $p = 0.001$)
- ปัจจัยอื่นๆ เช่น ภูมิภาคที่พำนักอาศัย และการใช้เวลาบนโซเชียลมีเดีย ไม่มีนัยสำคัญทางสถิติในการทำนายปัญหาสุขภาพ

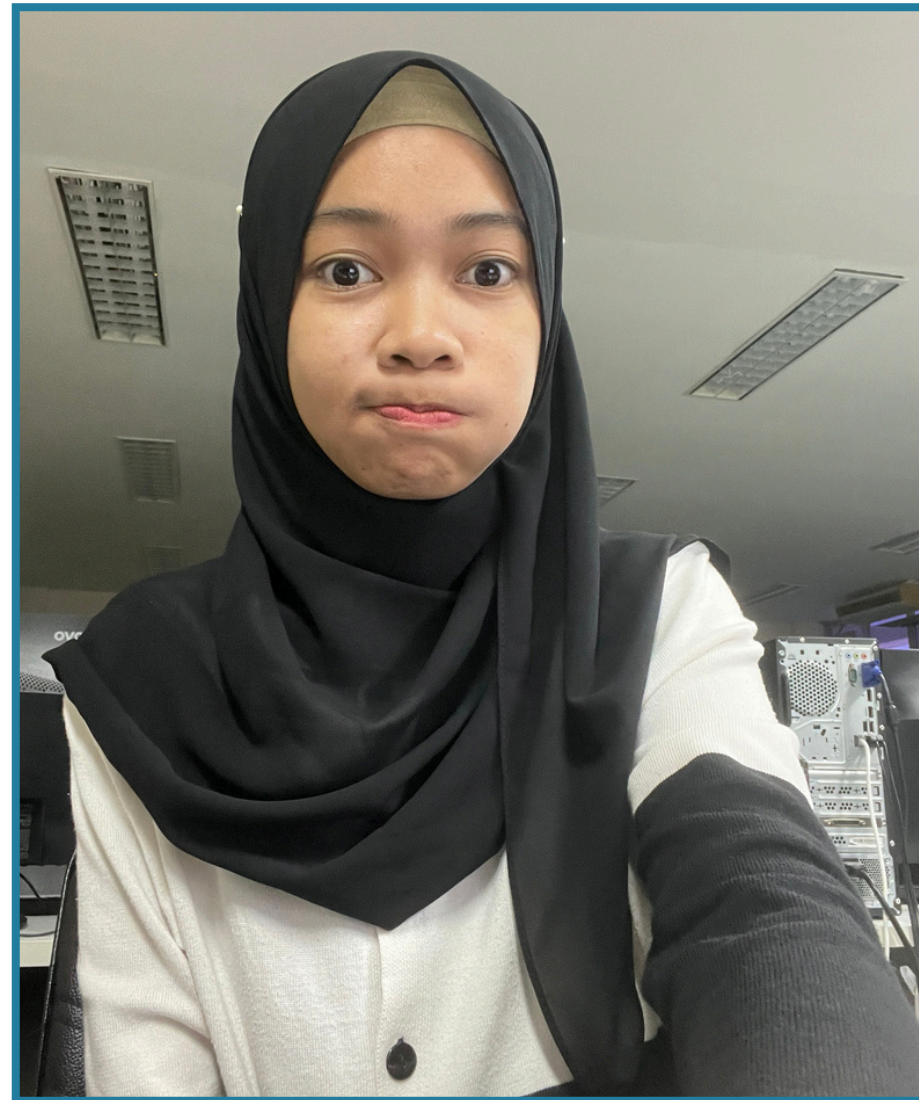
สรุป

การศึกษานี้แสดงให้เห็นว่าปัจจัยการใช้เวลาในการเรียนออนไลน์และจำนวนมื้ออาหารในแต่ละวันมีผลกระทบสำคัญต่อสุขภาพของนักเรียนในช่วงล็อกดาวน์ ในขณะที่ปัจจัยอื่นๆ เช่น การใช้เวลาบนโซเชียลมีเดียและภูมิภาคที่พักอาศัย ไม่ได้ส่งผลกระทบที่มีนัยสำคัญทางสถิติ การลดความเสี่ยงจากปัญหาสุขภาพในกลุ่มนักเรียนจึงอาจต้องให้ความสำคัญกับพฤติกรรมกรเรียนออนไลน์และโภชนาการ

เอกสารอ้างอิง

- ธนโชติ เทียมแสง. (2023). วารสารราชานุกูล, 35(1).
จาก <https://he03.tci-thaijo.org/index.php/RJJ/article/view/1257>
- สิริสิทธิ์ ศักดิ์ภิรมย์, & กาญจนา บุญส่ง. (2023). วารสาร มจร อุบลปริทรศน์.
จาก <https://so06.tci-thaijo.org/index.php/mcjou/article/view/265625>
- จำพร ศรียาภัย, et al. (2023). วารสารสุขภาพและการศึกษาพยาบาล.
จาก <https://he02.tci-thaijo.org/index.php/Jolbcnm/article/view/259643>
- พิระพัฒน์ วิศิษฐ์ศรี. (2023). Kalasin University Journal of Science Technology and Innovation, 2(2).
จาก https://li01.tci-thaijo.org/index.php/sci_01/article/view/258130
- Phutsang, P., et al. (2024). Vajira Nursing Journal, 26(1).
จาก <https://he02.tci-thaijo.org/index.php/vnj/article/view/266629>

จัดทำโดย



นางสาว ธิลณี สังกัณณกุล
รหัส 6520310018



Mr. Chhery Chorn
No. 6520310203





THANK YOU

Sorry for any mistakes in our presentation.