



รายงาน

เรื่อง การศึกษาปัจจัยที่เกี่ยวข้องกับการเข้าดูพอดแคสต์ ของคนไทยด้วยหลักการทางสถิติ

จัดทำโดย

นายธนาณพ	กุลพันธ์	6530182121
นายวันพิชิต	ธรรมบัวชา	6530368621
นางสาวศิวพร	เพ็งผล	6532170021

เสนอ

รศ. ดร. อังศุมาลิน เสนจันทร์มิไชย

รายงานฉบับนี้เป็นส่วนหนึ่งของรายวิชา

2104254 สถิติสำหรับงานวิศวกรรม 2

ตอนเรียนที่ 1 ภาคการศึกษาปลาย ปีการศึกษา 2566

คณะวิศวกรรมศาสตร์ ภาควิชาวิศวกรรมอุตสาหการ

จุฬาลงกรณ์มหาวิทยาลัย



รายงาน

เรื่อง การศึกษาปัจจัยที่เกี่ยวข้องกับการเข้าดูพอดแคสต์ ของคนไทยด้วยหลักการทางสถิติ

จัดทำโดย

นายธนาณพ	กุลพันธ์	6530182121
นายวันพิชิต	ธรรมบัวชา	6530368621
นางสาวศิวพร	เพ็งผล	6532170021

เสนอ

รศ. ดร. อังศุมาลิน เสนจันทร์ติไชย

รายงานฉบับนี้เป็นส่วนหนึ่งของรายวิชา

2104254 สถิติสำหรับงานวิศวกรรม 2

ตอนเรียนที่ 1 ภาคการศึกษาปลาย ปีการศึกษา 2566

คณะวิศวกรรมศาสตร์ ภาควิชาวิศวกรรมอุตสาหการ

จุฬาลงกรณ์มหาวิทยาลัย

คำนำ

รายงานฉบับนี้เป็นส่วนหนึ่งของรายวิชา 2104254 สถิติสำหรับงานวิศวกรรม 2 ชั้นปีที่ 2 ภาคการศึกษาปลาย ปีการศึกษา 2566 ภาควิชาวิศวกรรมอุตสาหการ ซึ่งทางคณะผู้จัดทำได้ทำการศึกษาเรื่อง ปัจจัยที่เกี่ยวข้องกับการเข้าสู่ตลาดคลิปปอดแคสต์ของคนไทยโดยอาศัยหลักการทางสถิติบนแพลตฟอร์มยูทูบชื่อ The Secret Sauce โดยมีวัตถุประสงค์เพื่อศึกษาปัจจัยต่าง ๆ ที่เกี่ยวข้องสำหรับการเข้าสู่ตลาดคลิปปอดแคสต์ของคนไทย และการศึกษารูปแบบหรือพฤติกรรมการนำเสนอข้อมูล โดยแบ่งเป็นหมวดหมู่บนแพลตฟอร์มนั้น

คณะผู้จัดทำขอขอบพระคุณ รศ.ดร.อังศุมาลิน เสนจันทร์ติไชย และอาจารย์ท่านอื่น ๆ รวมไปถึงบุคคลที่เกี่ยวข้องที่มอบความรู้และให้แนวทางในการศึกษาตลอดช่วงเวลาในการทำรายงานการศึกษานี้

คณะผู้จัดทำหวังว่ารายงานฉบับนี้จะมอบความรู้ รวมถึงเป็นประโยชน์แก่ผู้เรียนและผู้อ่านทุกท่านเป็นอย่างยิ่ง ซึ่งหากมีข้อผิดพลาดประการใด คณะผู้จัดทำขอน้อมรับไว้และขออภัยมา ณ ที่นี้ด้วย

คณะผู้จัดทำ

สารบัญ

คำนำ.....	ก
สารบัญภาพ	ง
สารบัญตาราง	ฉ
บทที่ 1 บทนำ.....	1
1. แนวคิด ที่มา และความสำคัญ.....	1
2. วัตถุประสงค์ในการศึกษา.....	3
3. ขอบเขตของการศึกษาค้นคว้า.....	3
4. วัสดุ อุปกรณ์ หรือเครื่องมือที่ใช้งาน	3
5. ผลประโยชน์ที่คาดว่าจะได้รับ	4
บทที่ 2 วิธีการดำเนินงาน	5
1. สถานที่และระยะเวลา.....	5
2. ขั้นตอนการดำเนินงาน.....	6
2.1. ข้อมูลเบื้องต้นของแซนเนล The Secret Sauce	6
2.2. การวิเคราะห์ข้อมูล	7
บทที่ 3 ผลการดำเนินงาน.....	9
1. การเก็บข้อมูล	9
2. การจัดการข้อมูล.....	11
3. การออกแบบการวิเคราะห์ข้อมูล.....	13
บทที่ 4 การวิเคราะห์ทางสถิติ	16
1. การวิเคราะห์ด้วยสถิติเชิงพรรณนา	16
1.1. จำนวนคลิปที่ถูกอัปโหลดในแต่ละปีตลอดปี พ.ศ. 2561 ถึง 2566.....	16
1.2. จำนวนคลิปแต่ละหมวดหมู่ที่ถูกอัปโหลดในแต่ละปี	17
1.3. ค่าเฉลี่ยของจำนวนคนดูโดยแบ่งตามหมวดหมู่.....	18
1.4. ค่าเฉลี่ยของจำนวนคนดูโดยแบ่งตามประเภทสื่อ	19
1.5. การกระจายตัวของความยาววิดีโอตลอดปี พ.ศ. 2561 ถึง 2566	20
1.6. สัดส่วนความยาววิดีโอต่อไตรมาส	21
1.7. ค่าเฉลี่ยจำนวนคนดูโดย Keyword ที่พบบ่อย 9 อันดับแรก	22
1.8. การกระจายระหว่างจำนวนคนดูและจำนวนคนกดถูกใจ	23
1.9. การกระจายระหว่างจำนวนคนดูและสัดส่วนการกดถูกใจ	23
2. การวิเคราะห์ด้วยสถิติเชิงอนุมาน	24

2.1. การทดสอบสมมติฐานระหว่างค่าเฉลี่ยของจำนวนคนรับชมวิดีโอแบบมีพิธีกรรับเชิญและไม่มีพิธีกรรับเชิญในแต่ละวิดีโอ.....	24
2.2. การทดสอบความเป็นอิสระต่อกันระหว่างหมวดหมู่ Podcast และความยาวคลิปด้วยการทดสอบแบบไคสแควร์.....	27
2.3. การวิเคราะห์ความแปรปรวนของข้อมูล.....	31
2.4. การสร้างแบบจำลองสมการถดถอยเชิงเส้นอย่างง่าย.....	44
บทที่ 5 สรุป อภิปรายและข้อเสนอแนะ	53
1. วัตถุประสงค์ของการศึกษา	53
2. สรุป และอภิปรายผลของการศึกษา.....	53
3. ข้อเสนอแนะ	56
บรรณานุกรม.....	57
ภาคผนวก	58
1. รูปภาพเพิ่มเติม	59
1.1. Output ที่ได้จากสมการที่ 1 โดยใช้คำสั่งในโปรแกรม R – Studio	59
1.2. Output ช่วงแห่งความเชื่อมั่นโดยใช้คำสั่งในโปรแกรม R – Studio	59
1.3. หน้าปกสไลด์สำหรับการนำเสนอวันที่ 18 เมษายน 2567.....	60
2. ตารางเพิ่มเติม.....	60
2.1. ตารางแจกแจงหน้าที่ในการรับผิดชอบภาระงานแต่ละส่วน.....	60
2.2. ค่าสถิติพื้นฐานของจำนวนคนดูและจำนวนคนกดถูกใจ	61
2.3. ค่าสถิติพื้นฐานของความยาววิดีโอ	62
2.4. ตารางแจกแจงจำนวนคลิปในแต่ละช่วงเวลาสำหรับในแต่ละไตรมาส.....	63
2.5. ตารางแจกแจงจำนวนคลิปที่พบ Keyword 9 อันดับแรก	64
2.6. ตารางสถิติพื้นฐานแจกแจงจำนวนคนรับชมในแต่ละปี	64
2.7. ตารางแสดงจำนวนคนดูทั้งหมดในแต่ละประเภท.....	65
2.8. ตารางแจกแจงจำนวนวิดีโอที่อัปโหลดในแต่ละปี.....	66
2.9. ตารางแสดงการหาค่าไคสแควร์	66
3. เนื้อหา หรือข้อมูลเพิ่มเติม.....	67
3.1. ตัวแบบพยากรณ์สำหรับค่าเฉลี่ยของจำนวนคนดูคลิปแต่ละไตรมาส	67
3.2. การทำ Normality Test กับความยาววิดีโอด้วยการทดสอบแบบ Shapiro - Wilk.....	79
3.3. การวิเคราะห์ด้วยสถิติเชิงพรรณนา (เพิ่มเติม)	80

สารบัญภาพ

รูปภาพ 1 ช่องภาพแซนเชล The Secret Sauce.....	6
รูปภาพ 2 เคน-นครินทร์.....	6
รูปภาพ 3 ภาพรวมการดำเนินงาน	8
รูปภาพ 4 ลักษณะของข้อมูลที่เก็บมา (video_youtube)	9
รูปภาพ 5 ลักษณะของข้อมูลที่เก็บมา (playlist_youtube)	9
รูปภาพ 6 ภาพรวมข้อมูลหลังจากจัดการข้อมูลเพื่อนำไปใช้งาน	11
รูปภาพ 7 การจำแนกประเภทและหมวดหมู่ของวิดีโอ.....	12
รูปภาพ 8 แผนภาพแสดงลำดับการทำสถิติเชิงอนุमान	15
รูปภาพ 9 แผนภูมิแท่งแสดงจำนวนคลิปที่ถูกอัปโหลดในแต่ละปี	16
รูปภาพ 10 แผนภูมิแท่งแสดงจำนวนวิดีโอที่ถูกอัปโหลด	17
รูปภาพ 11 แผนภูมิแท่งแสดงจำนวนคนดูโดยเฉลี่ยแบ่งตามหมวดหมู่.....	18
รูปภาพ 12 แผนภูมิแท่งแสดงจำนวนคนดูโดยแยกเป็นประเภทในแต่ละปี.....	19
รูปภาพ 13 แผนภูมิฮีทแกรมแสดงการกระจายตัวของความยาววิดีโอ	20
รูปภาพ 14 แผนภูมิแท่งแสดงสัดส่วนความยาววิดีโอในแต่ละไตรมาส	21
รูปภาพ 15 ค่าเฉลี่ยจำนวนคนดูโดย Keyword ที่พบบ่อย 9 อันดับแรก	22
รูปภาพ 16 แผนภาพกระจายระหว่างจำนวนคนดู และจำนวนการกดถูกใจในแต่ละคลิป.....	23
รูปภาพ 17 แผนภาพกระจายระหว่างจำนวนคนดู.....	23
รูปภาพ 18 สรุปขั้นตอนการทดสอบสมมติฐานของจำนวนคนรับชมวิดีโอระหว่างแบบมีพิธีกรรับเชิญ และไม่มีพิธีกรรับเชิญ.....	26
รูปภาพ 19 สรุปขั้นตอนการทดสอบสมมติฐานด้วยการทดสอบแบบไคสแควร์.....	30
รูปภาพ 20 Normal Probability Plot of Residuals สำหรับ ANOVA (2.3.1).....	34
รูปภาพ 21 Plot of Residuals vs Fitted Value สำหรับ ANOVA (2.3.1)	34
รูปภาพ 22 แผนภาพแสดงความสัมพันธ์ระหว่างปัจจัยต่าง ๆ ที่เกี่ยวข้อง	39
รูปภาพ 23 Normal Probability Plot of Residuals สำหรับ ANOVA (2.3.2)	42
รูปภาพ 24 Plot of Residual vs Fitted value สำหรับ ANOVA (2.3.2).....	42
รูปภาพ 25 ภาพแสดงแนวโน้มของข้อมูลระหว่างจำนวนคนดูและจำนวนคนกดถูกใจ	44
รูปภาพ 26 Normal Probability Plot of Residuals สำหรับสมการที่ 1.....	48
รูปภาพ 27 Residuals VS Fitted Value plot สำหรับสมการที่ 1	48
รูปภาพ 28 Residuals VS Order สำหรับสมการที่ 1 (a).....	49
รูปภาพ 29 Residuals VS Order สำหรับสมการที่ 1 (b).....	49

รูปภาพ 30 ภาพแสดงระหว่างข้อมูลจริง.....	51
รูปภาพ 31 ช่วงแห่งความเชื่อมั่นของแบบจำลองที่ 1	52
รูปภาพ 32 Output ที่ได้ของสมการที่ 1 (ตาราง Regression) จากโปรแกรม R-Studio	59
รูปภาพ 33 Output ช่วงแห่งความเชื่อมั่น โดยใช้คำสั่งในโปรแกรม R - Studio.....	59
รูปภาพ 34 หน้าปกสไลด์สำหรับการนำเสนอวันที่ 18 เมษายน 2567.....	60
รูปภาพ 35 แผนภูมิเส้นแสดงการเปลี่ยนแปลงของค่าเฉลี่ยคนดูในแต่ละไตรมาส	68
รูปภาพ 36 แผนภูมิเส้นแสดงค่าพยากรณ์แบบ Holt's Exponential Smoothing	70
รูปภาพ 37 แผนภูมิเส้นแสดงค่าพยากรณ์แบบ Weight Moving Average.....	73
รูปภาพ 38 แผนภูมิกระจายระหว่างปีและค่าเฉลี่ยจำนวนคนดูในช่วง พ.ศ. 2561 ถึง 2565	74
รูปภาพ 39 แผนภูมิเส้นแสดงค่าพยากรณ์แบบ Linear Trend Line.....	76
รูปภาพ 40 แผนภาพกล่องแสดงการกระจายตัวของความยาววิดีโอ	79
รูปภาพ 41 Output ที่ได้จากการทำ Shapiro - Wilk Test จากโปรแกรม R	79
รูปภาพ 42 แผนภาพกระจายระหว่างจำนวนคนรับชมและจำนวนข้อความแสดงความคิดเห็น	80
รูปภาพ 43 แผนภาพกล่องแสดงจำนวนคนรับชมวิดีโอในแต่ละปี.....	81

สารบัญตาราง

ตาราง 1 ตารางแสดงลำดับการดำเนินโครงการ	5
ตาราง 2 ตารางอธิบายความหมายของคอลัมน์ที่เก็บข้อมูลมา	10
ตาราง 3 ตารางจำแนกความยาววิดีโอ	12
ตาราง 4 ตารางอธิบายการออกแบบการวิเคราะห์ข้อมูล (สถิติเชิงพรรณนา)	13
ตาราง 5 ตารางอธิบายการออกแบบการวิเคราะห์ข้อมูล (สถิติเชิงอนุมาน).....	14
ตาราง 6 ตารางแจกแจงความถี่ของความยาววิดีโอ	20
ตาราง 7 ตารางสรุปจำนวนคนรับชมวิดีโอเฉลี่ยโดยแบ่งตามการมีพิธีกรรับเชิญ	24
ตาราง 8 ตารางแสดงข้อมูลระหว่างหมวดหมู่และความยาวคลิป	27
ตาราง 9 ตารางแสดงค่าคาดหวังระหว่างหมวดหมู่และความยาวคลิป.....	28
ตาราง 10 ตารางแสดงจำนวนคนดูโดยเฉลี่ยในแต่ละประเภทและแต่ละหมวดหมู่.....	31
ตาราง 11 ตาราง ANOVA สำหรับค่าเฉลี่ยคนดูในแต่ละหมวดหมู่และแต่ละปี	32
ตาราง 12 ตารางสรุปข้อมูลที่ได้จากกราฟสำหรับการตรวจสอบสมมติฐาน ANOVA (2.2.1).....	35
ตาราง 13 ตารางสรุปข้อมูล (เบื้องต้น) ของปัจจัยต่าง ๆ	38
ตาราง 14 ตารางแสดงค่าสหสัมพันธ์ระหว่างปัจจัยต่าง ๆ	39
ตาราง 15 ตาราง ANOVA สำหรับแสดงปัจจัยต่าง ๆ กับค่าเฉลี่ยจำนวนคนดู.....	40
ตาราง 16 ตารางสรุปข้อมูลที่ได้จากกราฟจากการตรวจสอบสมมติฐานของ ANOVA (2.2.2).....	43
ตาราง 17 ตารางแสดง Regression ของสมการที่ 1	45
ตาราง 18 ตารางสรุปข้อมูลที่ได้จากกราฟสำหรับ	50
ตาราง 19 ตารางแจกแจงภาระหน้าที่ของคณะผู้จัดทำ	60
ตาราง 20 ตารางแสดงค่าสถิติพื้นฐานของจำนวนคนดูและจำนวนคนกดถูกใจ	61
ตาราง 21 ตารางแสดงค่าสถิติพื้นฐานของความยาววิดีโอ	62
ตาราง 22 ตารางแจกแจงความยาววิดีโอในแต่ละช่วงเวลาสำหรับแต่ละไตรมาส	63
ตาราง 23 ตารางแจกแจงจำนวนคลิปที่พบ Keyword 9 อันดับแรก	64
ตาราง 24 ตารางแจกแจงจำนวนคนรับชมในแต่ละปี.....	64
ตาราง 25 ตารางแสดงค่าเฉลี่ยของจำนวนคนดูในแต่ละประเภทวิดีโอ	65
ตาราง 26 ตารางแสดงค่าเฉลี่ยของจำนวนคนดูในแต่ละประเภทวิดีโอ (ต่อ).....	66
ตาราง 27 ตารางแจกแจงจำนวนวิดีโอที่อัปโหลดในแต่ละปี	66
ตาราง 28 ตารางแสดงการหาค่าไคสแควร์	66
ตาราง 29 ตารางแสดงการคำนวณการพยากรณ์แบบ Holt's.....	69
ตาราง 30 ตารางแสดงการคำนวณการพยากรณ์แบบ Moving Average.....	72

ตาราง 31 ตารางสรุปจำนวนคนดูโดยเฉลี่ย ในแต่ละปีตั้งแต่ พ.ศ. 2561 ถึง 2565	74
ตาราง 32 ตารางสรุปค่า MAPE จากการพยากรณ์ทั้ง 3 รูปแบบ	77
ตาราง 33 ตารางแสดงการหา MAPE ของการพยากรณ์ 3 รูปแบบ.....	78

บทที่ 1

บทนำ

1. แนวคิด ที่มา และความสำคัญ

ในปัจจุบัน เทคโนโลยีสารสนเทศและอินเทอร์เน็ตมีบทบาทสำคัญต่อการดำเนินชีวิตของเรา มากขึ้นอย่างยิ่งไม่แพ้กับสิ่งอื่นใด อินเทอร์เน็ตได้กลายมาเป็นปัจจัยสำคัญในการดำรงชีวิตในยุคของ ข้อมูล แม้ว่าบางครั้งเราอาจไม่รับรู้ถึงความสำคัญของมันได้ในทันที

หลังจากวิกฤตการณ์การแพร่ระบาดของอย่างหนักของไวรัสโคโรนา 2019 (COVID-19) รูปแบบ การใช้ชีวิตของคนทั่วโลกได้เปลี่ยนแปลงอย่างสิ้นเชิง รวมถึงการทำงาน แนวคิด พฤติกรรมต่าง ๆ รวมไปถึงการนำเทคโนโลยีและอุปกรณ์ต่าง ๆ มาใช้ เพื่ออำนวยความสะดวกในการใช้ชีวิตเพิ่มมากขึ้น โดยเฉพาะอย่างยิ่งการนำเทคโนโลยีเหล่านี้สำหรับใช้ในการเข้าถึงสื่อสารสนเทศ และข้อมูลใน อินเทอร์เน็ตได้สะดวก มากไปกว่านั้น ยังมีการเติบโตของสื่อที่นำเสนอข้อมูลผ่านการฟังเสียงเป็นหลัก นั่นก็คือ Podcast

Podcast หรือ พอดแคสต์ เริ่มมีมาตั้งแต่ปี พ.ศ. 2547 เป็นสื่อที่อาศัยการรับรู้ข้อมูลโดยการ ฟังเป็นหลัก มีความนิยมอย่างแพร่หลายในยุคปัจจุบันทั่วโลก เนื่องจากสามารถเข้าถึงได้ง่าย ใช้เวลา ในการฟังไม่นาน เป็นสื่อที่กระชับและมีข้อคิด ทำให้ที่มีคนสนใจเป็นอย่างมาก นอกจากนี้ มันยังสร้างความบันเทิงและดึงดูดผู้ฟังด้วยความสะดวกสบาย ซึ่งจากการสำรวจพบว่า ผู้คนมักฟัง Podcast เมื่อ อยู่บ้าน ทำกับข้าว ออกกำลังกาย หรือแม้กระทั่งขณะทำงาน นอกจากนี้ยังมีผู้ฟังที่ชอบฟัง Podcast ขณะเดินทางอีกด้วย ซึ่งด้วยเหตุผลเหล่านี้จึงสามารถตอบโจทย์คนในยุคสมัยนี้ที่มีพฤติกรรมการใช้ ชีวิตแบบเร่งรีบ

คนไทยส่วนใหญ่มักหันหน้าเข้าหาความบันเทิงเพื่อความผ่อนคลาย ไม่ว่าจะเป็นขณะที่กำลัง เดินทางอยู่บนการเส้นทางคมนาคมที่แน่นขนัดหรือจะเป็นช่วงเวลาว่าง เช่น นอนเล่น หรือการทำ กิจกรรมอื่น ๆ การเปิดสื่อบันเทิง ทั้งการฟังเพลง การดู YouTube การดูโทรทัศน์ หรือแม้กระทั่งการ ฟัง Podcast จึงเป็นรูปแบบหลักที่คนไทยให้ความสนใจเป็นอย่างมาก โดยที่ในประเทศไทยเอง การ ฟัง Podcast เองก็จัดเป็นกระแสนิยม ซึ่งสังเกตได้จากการมีช่องทางต่าง ๆ ในการเข้าถึง ตัวอย่างเช่น

บน YouTube ที่มีแขนงต่าง ๆ ซึ่งจัดทำรายการ Podcast อยู่มากมาย โดยมีจุดเด่นเป็นเอกลักษณ์ในแต่ละช่องเป็นของตนเอง บ้างก็เป็นการพูดถึงเรื่องเศรษฐศาสตร์และการเงิน บ้างพูดถึงเรื่องจิตวิทยา MBTI¹ นพลักษณ์² โหราศาสตร์ รวมถึงสิ่งที่เกี่ยวข้องกับการเมือง เป็นต้น

สำหรับการศึกษาในครั้งนี้ ทางคณะผู้จัดทำต้องการทราบพฤติกรรมการเข้าชม Podcast ของคนไทยบนแขนง The Secret Sauce ซึ่งเป็นหนึ่งในรายการย่อยของแขนง The Standard ที่ภายหลังได้จัดทำแขนงแยกเป็นรายการของตนเอง โดยมีผู้ติดตามมากถึง 1 ล้านคนซึ่งนับว่าเป็นแขนงส่วนน้อยของประเทศไทยที่มีจำนวนผู้ติดตามเกินกว่าหนึ่งล้านผู้ติดตาม โดยภายในแขนงได้มีการจัดแบ่งหมวดหมู่ของ Podcast ที่ได้อัปโหลดออกมาเป็นหลายหมวดหมู่ไม่ว่าจะเป็นทั้งการเมือง การเงิน การศึกษา ลอจิสติกส์ และอื่น ๆ ดังนั้นเราจึงตั้งใจที่จะศึกษาปัจจัยที่มีผลต่อการฟัง Podcast ของคนไทยด้วยหมวดหมู่เหล่านี้ โดยใช้เกณฑ์ต่าง ๆ เช่น ประเภทของ Podcast ความยาวคลิปของเนื้อหา วันเดือนปีที่อัปโหลด และอื่น ๆ เพื่อที่จะสามารถเข้าใจพฤติกรรมและลักษณะการดูของคนไทยในช่วงเวลาที่มีการเก็บข้อมูลได้มากขึ้น นอกจากนี้ข้อมูลที่ได้หลังผ่านการวิเคราะห์ ยังสามารถที่จะนำไปต่อยอดเพื่อใช้ในเชิงอุตสาหกรรมการผลิตสื่อในแวดวง Podcast ได้อีกเช่นกันเพื่อเป็นประโยชน์สำหรับทีมผู้จัดทำสื่อ ผู้ฟัง และคนอื่น ๆ ที่หันมาเริ่มสนใจในสื่อบันเทิง และอุตสาหกรรมประเภทนี้อีกด้วย

¹ MBTI (Myers-Briggs Type Indicator) ซึ่งเป็นแบบทดสอบที่ใช้จำแนกบุคลิกภาพของบุคคลออกเป็น 16 ประเภท

² นพลักษณ์ (Enneagram) คือ ศาสตร์เกี่ยวกับการเข้าใจผู้คนและบุคลิกภาพซึ่งจะแบ่งคนออกเป็น 9 ประเภท โดยที่แต่ละประเภคนั้นจะมีพฤติกรรม แรงจูงใจในการใช้ชีวิต หรือมุมมองที่มีต่อโลกแตกต่างกันออกไป

2. วัตถุประสงค์ในการศึกษา

1. เพื่อศึกษาพฤติกรรมและรูปแบบการเข้าชม Podcast ของคนไทยที่ส่งผลต่อจำนวนคนรับชม วิดีโอ Podcast บนแพลตฟอร์ม The Secret Sauce
2. เพื่อวิเคราะห์ปัจจัย 3 ปัจจัยที่ส่งผลต่อจำนวนคนดูในแต่ละวิดีโอได้แก่ หมวดหมู่ของวิดีโอ ความยาวของวิดีโอ และช่วงเวลาที่ยอดนิยมใน YouTube บนแพลตฟอร์ม The Secret Sauce

3. ขอบเขตของการศึกษาค้นคว้า

1. ข้อมูลที่มีการบันทึกเกี่ยวข้องกับการเข้าชม Podcast จาก YouTube บนแพลตฟอร์ม The Secret Sauce ตั้งแต่วันที่ 21 กุมภาพันธ์ พ.ศ. 2561 จนถึง 21 มีนาคม พ.ศ. 2566 เป็นจำนวนทั้งสิ้น 1039 วิดีโอ ระยะเวลา 5 ปี 28 วัน
2. องค์กรความรู้จากรายวิชา 2104253 สถิติเชิงวิศวกรรม 1 และ 2104254 สถิติเชิงวิศวกรรม 2 หลักสูตรวิศวกรรมอุตสาหการ ปีการศึกษา 2566
3. ระยะเวลาสำหรับการดำเนินการศึกษา เริ่มตั้งแต่วันที่ 11 มกราคม พ.ศ. 2567 จนถึงวันที่ 8 พฤษภาคม พ.ศ. 2567

4. วัสดุ อุปกรณ์ หรือเครื่องมือที่ใช้งาน

1. โปรแกรม Microsoft Excel, Microsoft Word สำหรับการดำเนินการศึกษาหลัก
2. Google Drive และ Google Sheet สำหรับการเก็บข้อมูล และทำการศึกษาย่อย
3. Discord และ Line สำหรับใช้ในการติดต่อสื่อสารกันภายในคณะผู้จัดทำ
4. Mycourseville เว็บไซต์สำหรับการศึกษาเนื้อหาสำหรับการทำงานหลัก
5. Python และ R – Studio³ สำหรับการวิเคราะห์ข้อมูลขนาดใหญ่

³ โปรแกรมสำหรับการวิเคราะห์ทางสถิติ โดยอยู่ในรูปแบบการทำงานผ่านการเขียนคำสั่งเป็นหลัก

5. ผลประโยชน์ที่คาดว่าจะได้รับ

1. เพื่อนำความรู้ทางสถิติที่ได้เรียนมาตลอด 1 ปีการศึกษา (จากทั้งรายวิชา 2104253 และรายวิชา 2104254) มาใช้ในออกแบบ และวิเคราะห์ข้อมูลในทางสถิติได้อย่างมีประสิทธิภาพ
2. เพื่อเข้าใจพฤติกรรมกรฟัง Podcast และปัจจัยต่าง ๆ การเข้าฟัง Podcast เหล่านั้นของคนไทยตลอดช่วงเวลาของข้อมูล
3. เพื่อนำเสนอข้อมูลผ่านการสรุปข้อมูลในแง่มุมต่าง ๆ โดยอาศัยการนำเสนอข้อมูลผ่านแผนภูมิ แผนภาพ และกราฟ
4. เพื่อเป็นกรณีศึกษาสำหรับการต่อยอดในเชิงอุตสาหกรรมการผลิตสื่อ รวมถึงนำข้อมูลที่ได้มาพัฒนาและปรับปรุงสื่อ Podcast ให้ดีขึ้น

บทที่ 2

วิธีการดำเนินงาน

คณะผู้จัดทำ ใช้สถานที่และระยะเวลา รวมถึงวิธีการดำเนินงานและมีเครื่องมือในการดำเนินการ ดังต่อไปนี้

1. สถานที่และระยะเวลา

สถานที่ในการจัดทำโครงการ

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ระยะเวลาในการจัดทำโครงการ

เริ่มตั้งแต่วันที่ 11 มกราคม 2567 สิ้นสุดวันที่ 10 พฤษภาคม 2567

ตาราง 1 ตารางแสดงลำดับการดำเนินโครงการ

ลำดับ	รายการ	ระยะเวลา	ผู้รับผิดชอบ
1	รวมกลุ่ม 3 คน และศึกษาวิธีทำโครงการ	11 มกราคม 2567	คณะผู้จัดทำ
2	ศึกษาเนื้อหาที่ต้องใช้ทำโครงการ	11 มกราคม – 4 เม.ย. 2567	คณะผู้จัดทำ
3	คิดหัวข้อโครงการ	14 กุมภาพันธ์ 2567	คณะผู้จัดทำ
4	รวบรวมข้อมูล และจัดการข้อมูล	16 กุมภาพันธ์ – 20 กุมภาพันธ์	ธนานพ
5	เสนอหัวข้อ หรือแนวทางการทำงาน	13 มีนาคม 2567	คณะผู้จัดทำ
6	วิเคราะห์สถิติเชิงพรรณนา	14 มีนาคม – 2 เมษายน 2567	วันพิชิต, ศิวพร
7	วิเคราะห์สถิติเชิงอนุมาน	25 มีนาคม – 5 เมษายน 2567	ธนานพ
8	เขียนโครงการ 5 บท	31 มีนาคม – 10 เมษายน 2567	ธนานพ, ศิวพร
9	ทำสไลด์ประกอบการนำเสนอ	7 เมษายน – 17 เมษายน 2567	วันพิชิต, ศิวพร
10	นำเสนอต่ออาจารย์ที่ปรึกษา	18 เมษายน 2567	คณะผู้จัดทำ
11	แก้ไขและปรับหัวข้อโครงการ	19 เมษายน – 3 พฤษภาคม 2567	คณะผู้จัดทำ
12	ส่งโครงการ	10 พฤษภาคม 2567	คณะผู้จัดทำ

2. ขั้นตอนการดำเนินงาน

2.1. ข้อมูลเบื้องต้นของแชนเนล The Secret Sauce



รูปภาพ 1 ช่องภาพแชนเนล The Secret Sauce

“ทำไมเราต้องฟังเรื่องแบรนด์ไทย ทั้งที่เราเองก็ไม่ได้ทำธุรกิจอะไรเสียหน่อย?

เคน-นครินทร์⁴ จะมาบอกคุณฟังว่า ถึงชื่อจะเหมือนรายการทำอาหาร แต่ที่จริงแล้ว เดอะซีเคร็ตซอส (The Secret Sauce) จะบอกเล่าเรื่องราวการก่อตั้ง ต่อสู้ ล้มลุก คลุกคลานสู่ความสำเร็จของธุรกิจที่เราคุ้นเคยดี แต่รู้เบื้องหลังเขาน้อยมาก ลองฟังตัวอย่างรายการเป็น ‘น้ำจิ้ม’ กันได้เลย”⁵

รายการ The Secret Sauce เป็นรายการที่ดำเนินโดย เคน-นครินทร์ วณิกไพบูลย์ ซึ่งเริ่มมีการลง Podcast ครั้งแรก วันที่ 6 สิงหาคม 2560 บนช่องแชนเนล The Standard ซึ่งภายหลังได้มีการสร้างช่อง YouTube รายการแยกเป็นตัวเอง ชื่อเดียวกับรายการ โดยเริ่มมีการลง Podcast ในช่องดังกล่าว เมื่อวันที่ 21 กุมภาพันธ์ 2561 โดยภายในรายการจะมุ่งเน้นการเล่าถึงความสำเร็จของแบรนด์ไทย นักธุรกิจไทยเป็นหลักในช่วงแรก และภายหลังได้มีการเพิ่มเนื้อหาไปสู่เรื่องอื่น ๆ เพิ่มเติม เช่น การศึกษา จิตวิทยา การเมือง และไลฟ์สไตล์ เป็นต้น ซึ่งได้มีช่องทางที่สามารถติดตามได้อย่างมากมายทั้งใน The Standard.co, iTunes Podcast, YouTube และ Podbean



รูปภาพ 2 เคน-นครินทร์

⁴ อดีตนักศึกษาคณะเภสัชศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ ผู้เป็นคอลัมนิสต์และบรรณาธิการการบริการ themomentum.co ซึ่งถูกจัดเป็นหนึ่งในคนรุ่นใหม่ที่น่าจับตามองในยุคสมัยปัจจุบัน

⁵ อ้างอิงจาก ‘แนะนำ ‘The Secret Sauce’ – THE STANDARD’ บนเว็บไซต์ <https://thestandard.co/>

2.2. การวิเคราะห์ข้อมูล

2.2.1. วางแผนลักษณะการเก็บข้อมูล

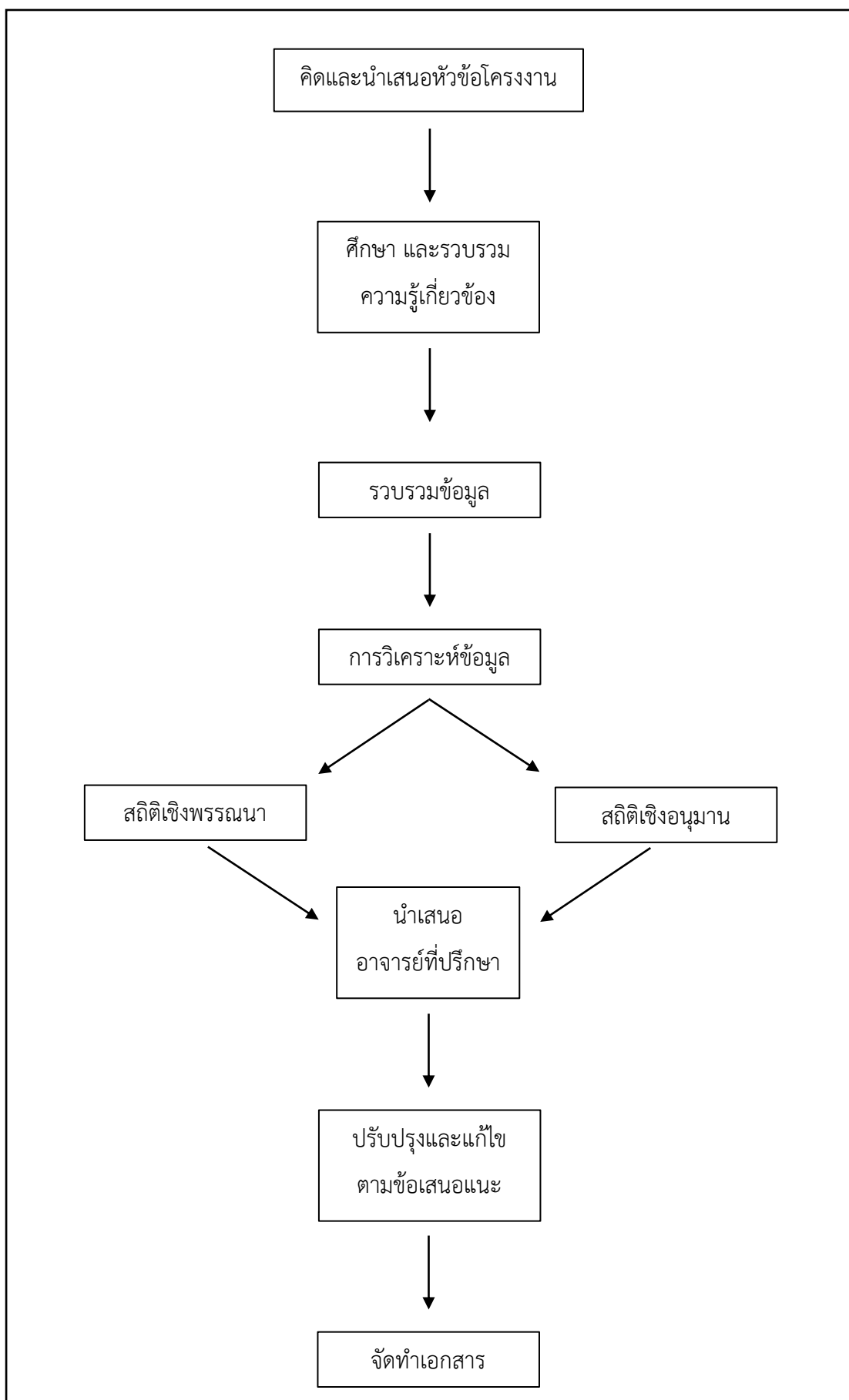
เริ่มต้นจะทำการใช้คำสั่งบน Google Sheet ที่ชื่อว่า IMPORTXML เพื่อสร้างโปรแกรมสำหรับการดึงข้อมูลจาก YouTube บนแชนเนล The Secret-Sauce โดยจะใช้ข้อมูลตั้งแต่ 21 กุมภาพันธ์ 2561 จนถึง 21 มีนาคม 2566 มีจำนวนทั้งสิ้น 1039 วิดีโอ เป็นเวลา 5 ปี 28 วัน มาเก็บไว้ในรูปแบบสเปรดชีต (Spreadsheet) ชนิดไฟล์.csv แล้วนำข้อมูลเหล่านั้นมาจัดการข้อมูล (Data Processing) ผ่านโปรแกรม Microsoft Excel และการเขียนคำสั่งด้วยภาษา Python โดยอาศัย Library ที่ชื่อ Pandas เป็นตัวช่วยหลัก เพื่อให้ข้อมูลที่ได้มานั้นเหมาะสมกับการนำไปใช้งานต่อในส่วนถัด ๆ ไป

2.2.2. การแบ่งส่วนการวิเคราะห์

เพื่อให้เป็นไปตามวัตถุประสงค์ที่วางไว้ ทางคณะผู้จัดทำจึงแบ่งการวิเคราะห์ด้วยหลักทางสถิติ 2 ส่วน ได้แก่

1. การวิเคราะห์โดยใช้สถิติเชิงพรรณนา (Descriptive Statistics) เพื่อใช้สำหรับการนำเสนอข้อมูลที่มีในการอธิบายลักษณะของข้อมูลในแง่มุมต่าง ๆ เช่น จำนวนวิดีโอที่อัปโหลดลงในแต่ละปี การแจกแจงฮิสโทแกรมของความยาววิดีโอ และการวิเคราะห์คำศัพท์บนชื่อวิดีโอ (Keyword) เพื่อดูความถี่ที่ค้นพบ เป็นต้น ซึ่งส่วนนี้จะใช้โปรแกรม Microsoft Excel ในการนำเสนอเป็นหลักเนื่องจากเป็นโปรแกรมที่สามารถออกแบบ และตกแต่งแผนภูมิต่าง ๆ ได้อย่างสวยงาม

2. การวิเคราะห์โดยใช้สถิติเชิงอนุมาน (Inferential Statistics) เพื่อสำหรับการอธิบายและวิเคราะห์ปัจจัยต่าง ๆ ที่ส่งผลต่อจำนวนคนดูในแต่ละวิดีโอ โดยจะใช้หลักการทางคณิตศาสตร์เพื่อสำหรับการวิเคราะห์ และเนื่องจากเป็นข้อมูลขนาดใหญ่ที่อาศัยการคำนวณหลาย ๆ ส่วนจึงใช้โปรแกรม R – Studio สำหรับการวิเคราะห์เป็นหลัก



รูปภาพ 3 ภาพรวมการดำเนินงาน

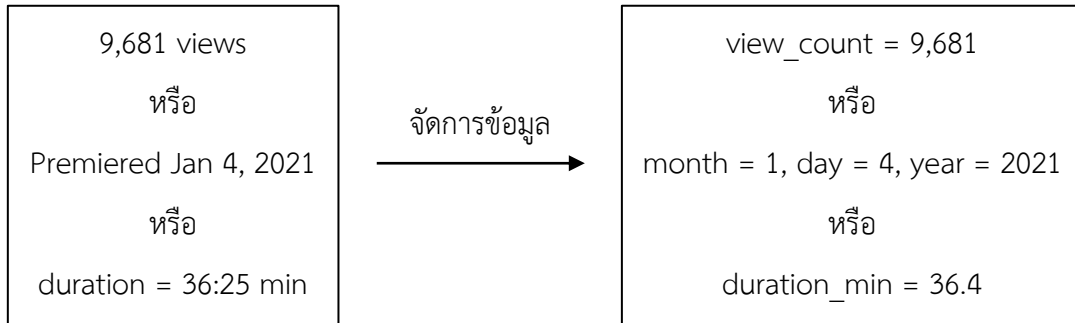
โดยลักษณะของข้อมูลที่เก็บมานั้น จะเก็บในรูปแบบของ Google spreadsheet บน Google Sheet ซึ่งแบ่งหมวดหมู่ตามคอลัมน์ดังนี้

ตาราง 2 ตารางอธิบายความหมายของคอลัมน์ที่เก็บข้อมูลมา

video_youtube	playlist_youtube
ข้อมูลรายละเอียดต่าง ๆ ของคลิปวิดีโอที่เผยแพร่ผ่านทาง YouTube ของช่อง THE STANDARD และ THE SECRET SAUCE ได้แก่	ข้อมูลรายชื่อ Playlist และวิดีโอ ที่ช่อง The Secret Sauce แบ่งเอาไว้ตามหมวดหมู่ธุรกิจ ได้แก่
ลิงก์วิดีโอ (video-href)	ชื่อ Playlist หมวดหมู่ธุรกิจ (playlist)
ชื่อตอน (title)	ลิงก์ Playlist (playlist-href)
ชื่อช่อง (channelTitle)	ลิงก์วิดีโอ (video-href)
วันที่อัปโหลด (publishedAt)	
ยอดการรับชม (viewCount)	
ยอดการถูกใจ (likeCount)	
ยอดการคอมเมนต์ (commentCount)	
ความยาววิดีโอ (duration)	
ยอดการดาวน์โหลด (download)	
รายละเอียดเพิ่มเติม (description)	

2. การจัดการข้อมูล

เนื่องจากข้อมูลที่เก็บมานั้นยังไม่ได้อยู่ในรูปแบบที่สามารถนำไปใช้ได้ จึงใช้โปรแกรม Microsoft Excel ซึ่งสามารถจัดการข้อมูลได้สะดวก ตัวอย่างเช่น



โดยผลลัพธ์จากการจัดการข้อมูลได้ดังต่อไปนี้

The screenshot shows a Microsoft Excel spreadsheet with a large table of video data. The table has columns for video ID, title, channel, view count, like count, comment count, and duration. The data is organized into a grid with multiple rows and columns, showing various video statistics. The spreadsheet is titled 'Project STAT P.B. Excel' and the user is 'Thananong Kullapan'.

รูปภาพ 6 ภาพรวมข้อมูลหลังจากจัดการข้อมูลเพื่อนำไปใช้งาน

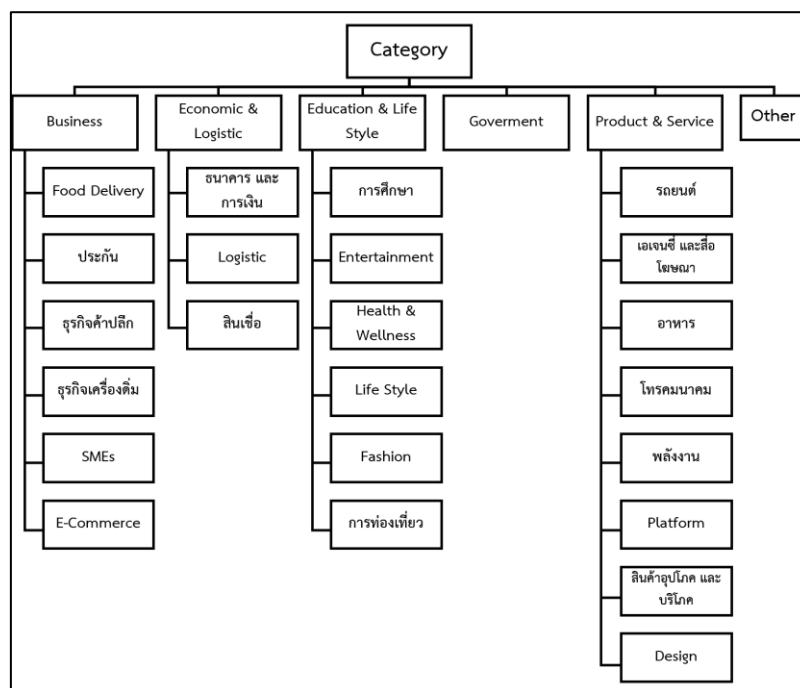
ขั้นตอนการจัดการข้อมูล

1. แยกข้อความวัน/เดือน/ปี ในแต่ละเซลล์ ออกมาเป็นตัวเลขในแต่ละคอลัมน์
2. เปลี่ยนค่าของความยาววิดีโอ เช่น 36:25 min เป็นตัวเลขปกติซึ่งจะได้เป็น 36.4 โดยการตัดข้อความ (String) ออก
3. ตัดข้อความบางส่วนเช่น likes และ views ในทุก ๆ คอลัมน์ viewCount และ likeCount
4. รวมข้อมูลเดือนบางส่วนแล้วสร้างคอลัมน์ใหม่เป็นไตรมาส (Quarter)
5. จัดประเภทความยาวของวิดีโอเป็น 4 หมวดหมู่ ได้แก่

ตาราง 3 ตารางจำแนกความยาววิดีโอ

หมวดหมู่	ความยาว
0 – 0.5 hours	0 – 30 นาที
0.5 – 1 hours	30 – 60 นาที
1 – 1.5 hours	60 – 90 นาที
1.5 – 2 hours	90 – 120 นาที

6. จัดประเภท (Industry) และหมวดหมู่ (Category) ของวิดีโอแต่ละวิดีโอ โดยเป็นได้ดังนี้



รูปภาพ 7 การจำแนกประเภทและหมวดหมู่ของวิดีโอ

3. การออกแบบการวิเคราะห์ข้อมูล

จากการเกริ่นในบทที่ 2 หัวข้อ 2.2. นั้น ทางคณะผู้จัดทำจึงได้แบ่งการวิเคราะห์ได้ตามหัวข้อทั้ง 2 ดังนี้

สถิติเชิงพรรณนา

ตาราง 4 ตารางอธิบายการออกแบบการวิเคราะห์ข้อมูล (สถิติเชิงพรรณนา)

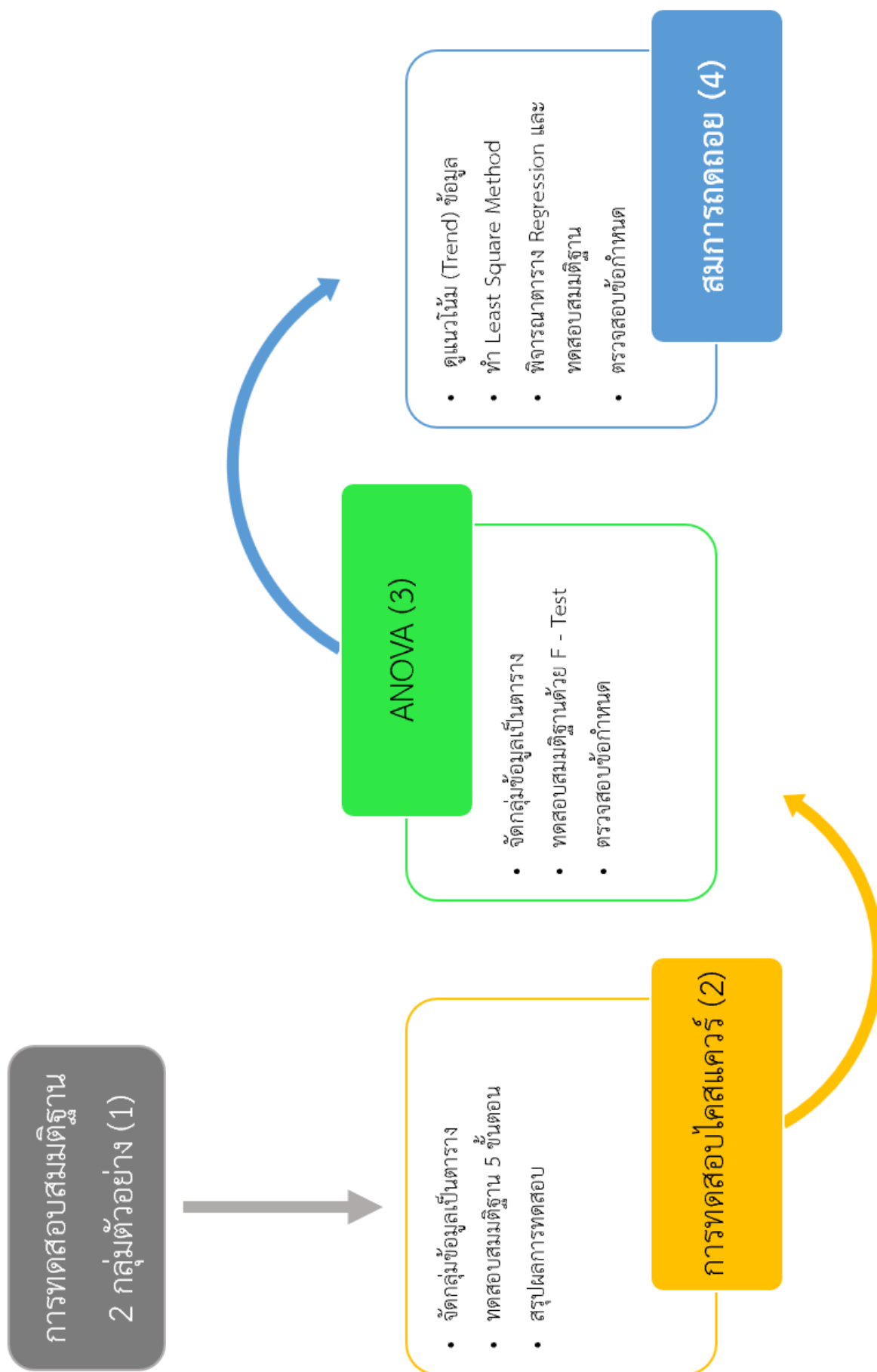
หัวข้อ	จุดประสงค์
1. จำนวนวิดีโอ	
1. จำนวนคลิปที่ถูกอัปโหลดในแต่ละปี ตลอดช่วงปี พ.ศ. 2561 ถึง 2566	เพื่อดูจำนวนของวิดีโอที่ถูกอัปโหลดในแต่ละปีว่า ในแต่ละปีมีจำนวนวิดีโอเท่าใด
2. จำนวนคลิปแต่ละหมวดหมู่ที่ถูก อัปโหลดในแต่ละปี	เพื่อเปรียบเทียบจำนวนวิดีโอในแต่ละหมวดหมู่ในแต่ละ ปี เพื่อดูลักษณะความนิยมของการนำเสนอสื่อ
2. จำนวนคนดู และความยาววิดีโอ	
3. ค่าเฉลี่ยของจำนวนคนดูโดย แบ่งตาม <u>หมวดหมู่</u>	เพื่อดูแนวโน้ม ลักษณะโดยรวมว่าคนไทยในแต่ละ ปีสนใจหรือมักฟัง Podcast ในหมวดหมู่ใด มากกว่ากัน
4. ค่าเฉลี่ยของจำนวนคนดูโดย แบ่งตาม <u>ประเภทสื่อ</u>	เพื่อดูความนิยมของการติดตาม Podcast ของคน ไทยในแต่ละประเภท
5. การกระจายตัวของความยาววิดีโอ ตลอด ปี พ.ศ. 2561 ถึง 2566	ดูการกระจายของตัวของความยาววิดีโอตลอด ช่วงเวลาที่เก็บข้อมูลว่ามีการกระจายตัวลักษณะ อย่างไร
6. สัดส่วนความยาววิดีโอต่อไตรมาส	ในแต่ละไตรมาสวิดีโอที่ถูกอัปโหลดโดยส่วนมาก แล้วจะมีความยาววิดีโอเท่าใด
7. ค่าเฉลี่ยจำนวนคนดูโดย Keyword ที่พบบ่อย 9 อันดับแรก	เพื่อตรวจหาคำหรือข้อความที่มักพบบ่อยในชื่อ ของวิดีโอ 9 อันดับแรก
3. การกระจายระหว่างจำนวนคนดูและการกดถูกใจ	
8. การกระจายระหว่างจำนวนคนดู และจำนวนคนกดถูกใจ	เพื่อดูการกระจายตัวระหว่างจำนวนคนดู และการ กดถูกใจ และเพื่อใช้ในการหาความสัมพันธ์ระหว่าง กันในส่วนข้อสถิติเชิงอนุมาน
9. การกระจายระหว่างจำนวนคนดู และสัดส่วนการกดถูกใจ	

สถิติเชิงอนุมาน

ตาราง 5 ตารางอธิบายการออกแบบการวิเคราะห์ข้อมูล (สถิติเชิงอนุมาน)

หัวข้อ	จุดประสงค์
1. การทดสอบสมมติฐานสำหรับ 2 กลุ่มตัวอย่าง (Hypothesis Testing with 2 Samples)	
1. การทดสอบค่าเฉลี่ยจำนวนคนรับชมวิดีโอระหว่างมีพิธีกรรับเชิญ และไม่มีพิธีกรรับเชิญ	เพื่อวิเคราะห์ว่าคนรับชมวิดีโอส่วนใหญ่มักจะดูแบบมีพิธีกรรับเชิญ หรือไม่มีพิธีกรรับเชิญมากกว่ากัน
2. การทดสอบด้วยไคสแควร์ (Chi – square Test)	
1. การทดสอบความเป็นอิสระต่อกันระหว่างหมวดหมู่ Podcast และความยาวคลิปด้วย การทดสอบแบบไคสแควร์	เพื่อวิเคราะห์ว่าหมวดหมู่ของวิดีโอและความยาวของวิดีโอมีความเกี่ยวข้องกันหรือไม่
3. การวิเคราะห์ความแปรปรวน (Analysis of Variance; ANOVA)⁶	
1. การวิเคราะห์ความแปรปรวนระหว่างหมวดหมู่ Podcast และในแต่ละปี	เพื่อวิเคราะห์ปัจจัยหรือการมีผลในจำนวนคนดูเฉลี่ยสำหรับแต่ละหมวดหมู่ ว่าหมวดหมู่มีปัจจัยส่งผลต่อจำนวนคนดูในแต่ละปีหรือไม่ โดยเป็นตัวอย่างการวิเคราะห์แบบพื้นฐาน
2. การวิเคราะห์ความแปรปรวน 3 ปัจจัย ได้แก่หมวดหมู่ Podcast, ความยาววิดีโอ และไตรมาสที่อัปโหลด	เพื่อวิเคราะห์ปัจจัยต่าง ๆ ในการศึกษาว่าปัจจัยทั้ง 3 ส่งผลต่อจำนวนคนดูหรือไม่ โดยจะพิจารณาในส่วนของความยาววิดีโอ
4. การสร้างสมการถดถอยเชิงเส้น (Linear Regression)	
1. สมการถดถอยเชิงเส้นแบบง่ายเพื่อแสดงความสัมพันธ์ระหว่างจำนวนคนดู และคนกดถูกใจ	เพื่อแสดงความสัมพันธ์ระหว่างจำนวนคนดูและคนกดถูกใจ ซึ่งจะมีการนำแผนภูมิการกระจายในสถิติเชิงพรรณนามาอ้างอิง

⁶ ในการทำ ANOVA หัวข้อที่ 1 นั้นจะมีการแสดงตัวอย่างผลลัพธ์ที่ได้จาก Microsoft Excel ในขณะที่หัวข้อที่ 2 นั้นจะใช้โปรแกรม R – studio ในการทดสอบ



รูปภาพ 8 แผนภาพแสดงลำดับการทำสถิติเชิงอนุมาน

บทที่ 4

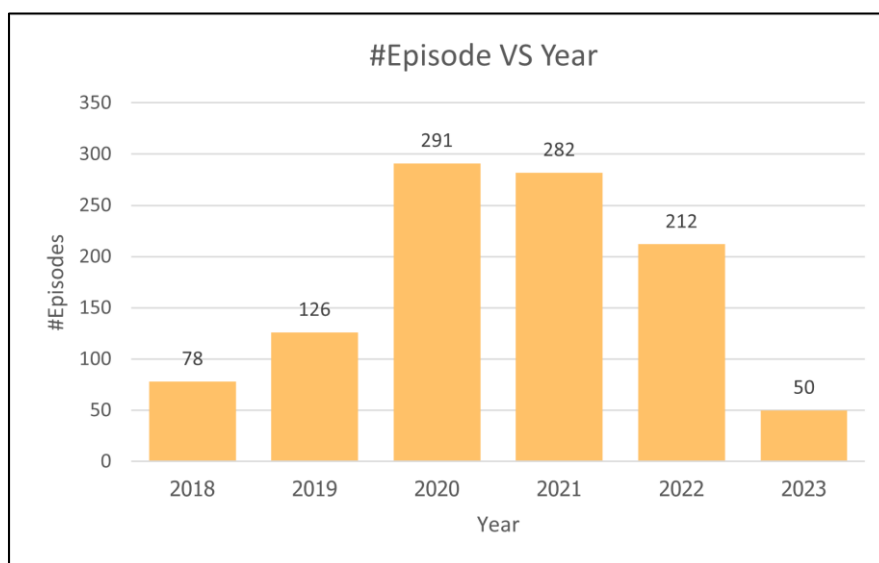
การวิเคราะห์ทางสถิติ

จากการนำข้อมูลที่รวบรวมของแซนเนล The Secret Sauce เพื่อวิเคราะห์ข้อมูลในหลายแง่มุมตามวัตถุประสงค์ที่ได้เสนอไว้ในบทที่ 1 และได้อธิบายไว้ในบทที่ 2 สามารถจำแนกผลการวิเคราะห์ได้ 2 ส่วน ดังนี้

1. การวิเคราะห์ด้วยสถิติเชิงพรรณนา (Descriptive Statistics)
2. การวิเคราะห์ด้วยสถิติเชิงอนุมาน (Inferential Statistics)

1. การวิเคราะห์ด้วยสถิติเชิงพรรณนา

1.1. จำนวนคลิปที่ถูกอัปโหลดในแต่ละปีตลอดปี พ.ศ. 2561 ถึง 2566



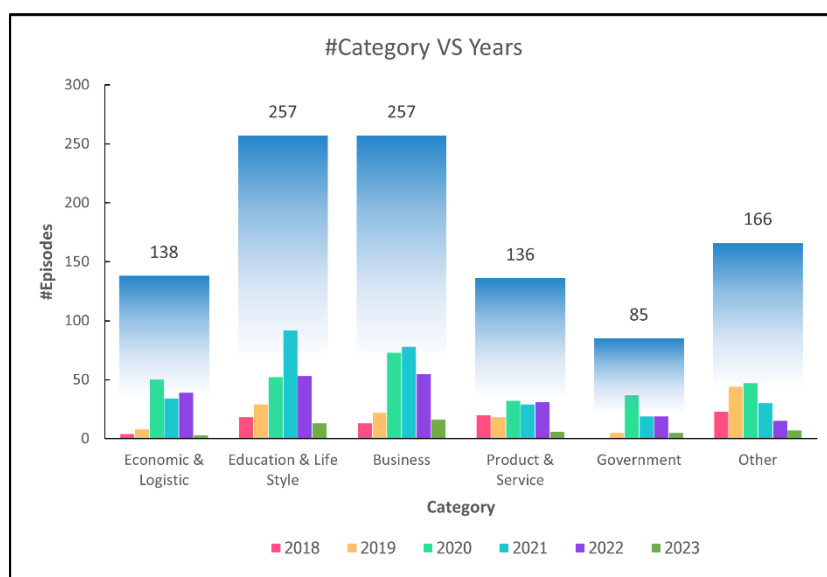
รูปภาพ 9 แผนภูมิแท่งแสดงจำนวนคลิปที่ถูกอัปโหลดในแต่ละปี

จากแผนภูมิแท่งแสดงให้เห็นว่าในช่วงตลอด 5 ปี 28 วันมานี้ มีการอัปโหลดวิดีโอของแซนเนล ซึ่งในปีแรกนั้นมีจำนวนที่ถูกอัปโหลดน้อย⁷ และเพิ่มมากขึ้นในช่วงหลัง ทั้งนี้เนื่องจากปี พ.ศ. 2561 เป็นปีแรกเริ่มที่มีการสร้างแซนเนล จึงมีจำนวนที่ถูก

⁷ ทั้งนี้เนื่องจากในช่วงแรกแซนเนล The Secret Sauce แต่เดิมที่เป็นรายการย่อยของ The Standard ซึ่งมีวิดีโอบางส่วนได้อัปโหลดภายในแซนเนล The Standard ภายใต้รายการ The Secret Sauce

อัปโหลดค่อนข้างน้อย ในขณะเดียวกันช่วง พ.ศ. 2563 ถึง พ.ศ. 2564 เป็นช่วงที่เป็น การแพร่ระบาดของไวรัสโควิดซึ่งถือเป็นช่วงที่มีนวัตกรรมแปลกใหม่และสิ่งต่าง ๆ เกิด ขึ้นมาเป็นจำนวนมาก รวมไปถึงข่าวสารทางธุรกิจและการลงทุน ทำให้มีการอัปโหลด คลิปเพื่อนำเสนอข้อมูลมีมากขึ้น

1.2. จำนวนคลิปแต่ละหมวดหมู่ที่ถูกอัปโหลดในแต่ละปี

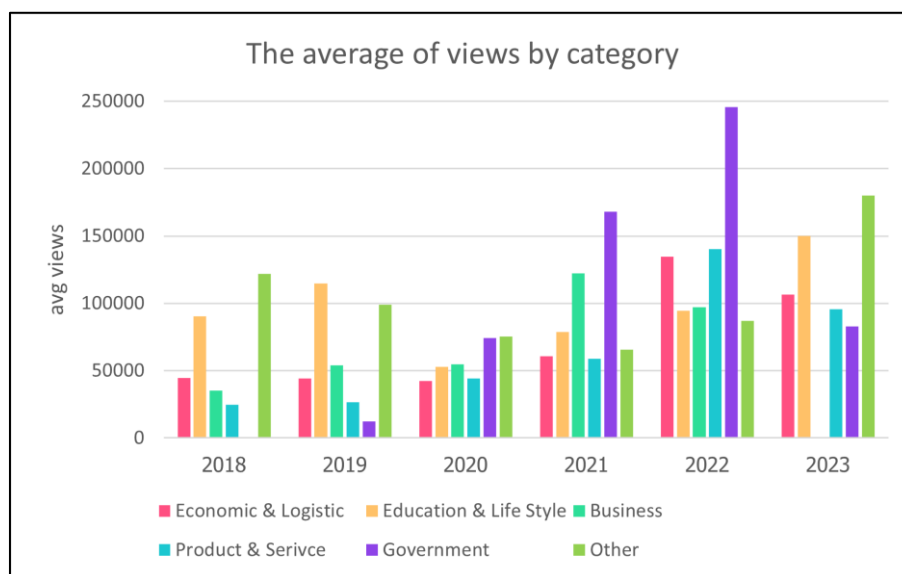


รูปภาพ 10 แผนภูมิแท่งแสดงจำนวนวิดีโอที่ถูกอัปโหลด
โดยแยกเป็นหมวดหมู่ในแต่ละปี

จากแผนภูมิแท่งดังกล่าวแสดงให้เห็นว่า ในแต่ละปี Category หรือหมวดหมู่ของ วิดีโอทั้งหมดมีการเปลี่ยนแปลงตลอดในทุกปี เนื่องจากความนิยม หรือกระแสของ เนื้อหาในช่วงนั้น รวมถึงความน่าสนใจของการนำเสนอข้อมูลในช่วงเวลาต่าง ๆ โดย จากแผนภูมียังแสดงให้เห็นอีกว่าหมวดหมู่จำพวก Education & Life Style และ จำพวก Business จะมีการอัปโหลดมากเป็นอันดับหนึ่ง ในขณะที่หมวดหมู่จำพวก Government จะมีการอัปโหลดน้อยที่สุด

นอกจากนี้จะเห็นได้ว่าภายในแผนภูมินั้น จะเน้นการนำเสนอในหมวดหมู่ Economic & Life Style และ Business เป็นหลักซึ่งสังเกตได้จากผลรวมจำนวน วิดีโอที่อัปโหลดนั้นมีมากที่สุด 2 อันดับแรกซึ่งสอดคล้องกับวัตถุประสงค์ของการสร้าง รายการ The Secret Sauce ในช่วงแรกและมีการเพิ่มเติมในหมวดหมู่อื่นด้วยเช่นกัน

1.3. ค่าเฉลี่ยของจำนวนคนดูโดยแบ่งตามหมวดหมู่

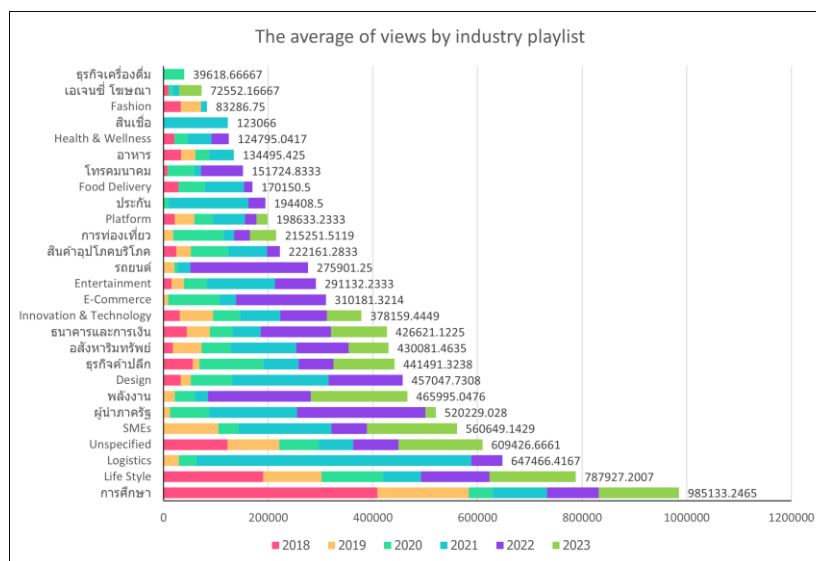


รูปภาพ 11 แผนภูมิแท่งแสดง

จำนวนคนดูโดยเฉลี่ยแบ่งตามหมวดหมู่

เมื่อพิจารณาจากแผนภูมิดังกล่าว พบว่าจำนวนคนรับชมวิดีโอโดยเฉลี่ยในแต่ละหมวดหมู่ในแต่ละปีมีแนวโน้มไปในทิศทางเดียวกัน โดยที่ปีใดที่มีการเพิ่มขึ้นของค่าเฉลี่ยรวมของคนดู จะทำให้หมวดหมู่ในแต่ละหมวดหมู่นั้นมีการเพิ่มขึ้นด้วย เป็นต้น ทั้งนี้อาจเป็นเพราะสาเหตุหลักมาจากในปีหลัง ๆ จำนวนคนเข้าฟัง Podcast มีมากขึ้นกว่าแต่ก่อน และแซนเนล The Secret Sauce ได้จัดทำ Podcast ที่มีเนื้อหาที่ทันสมัยตามกระแสในช่วงนั้น และยังกระจ่ายการทำหมวดหมู่ต่าง ๆ อีกด้วยทำให้ค่าเฉลี่ยโดยรวมของคนดูมีทิศทางไปในทางเดียวกัน

1.4. ค่าเฉลี่ยของจำนวนคนดูโดยแบ่งตามประเภทสื่อ



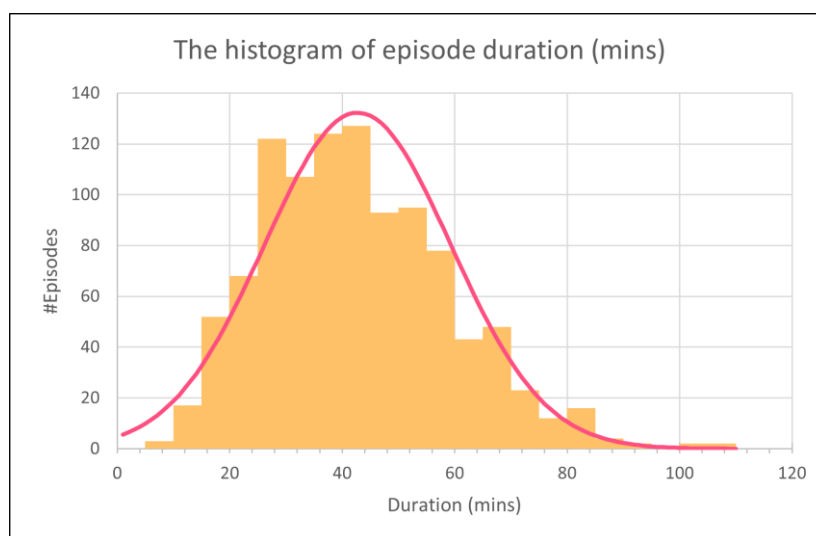
รูปภาพ 12 แผนภูมิแท่งแสดง

จำนวนคนดูโดยแยกเป็นประเภทในแต่ละปี

จากแผนภูมิแท่งดังกล่าว เมื่อพิจารณาค่าเฉลี่ยรวมของคนดูสื่อแต่ละประเภททั้ง 27 ประเภทในตลอดเวลาที่มีข้อมูลนั้น พบว่าหมวดหมู่ที่มีจำนวนค่าเฉลี่ยรวมของคนดูมากที่สุดคือ “การศึกษา” โดยที่ส่วนมากจะมาจากในปี พ.ศ. 2561 ที่มีหลายวิดีโอที่มีความน่าสนใจต่อผู้ฟังเป็นพิเศษ เช่น “ทักษะแห่งอนาคต จาก ริวิส หาญอุตสาหกรรม และ กวีรุณี เต็มภูวภัทร | The Secret Sauce EP.49” ที่มีคนดูถึง 121,164 คนดูเป็นต้น และอันดับถัดลงมาคือ “Life Style” เป็นต้น

นอกจากนี้ข้อมูลดังกล่าว ยังแสดงให้เห็นว่าในแต่ละปีจำนวนเฉลี่ยของคนดูในแต่ละหมวดหมู่นั้นไม่ได้มีลักษณะที่คงที่เสมอไป เนื่องจากเป็นเพราะกระแส หรือความนิยมในช่วงเวลานั้นด้วย เช่น ในปี พ.ศ. 2564 จะเห็นว่าหมวดหมู่ “Logistics” มีจำนวนที่มากกว่าปกติเป็นอย่างมาก เมื่อเทียบกับจำนวนในปีอื่น ๆ ของหมวดหมู่เดียวกัน และหมวดหมู่อื่น ๆ ในปีเดียวกัน

1.5. การกระจายตัวของความยาววิดีโอตลอดปี พ.ศ. 2561 ถึง 2566



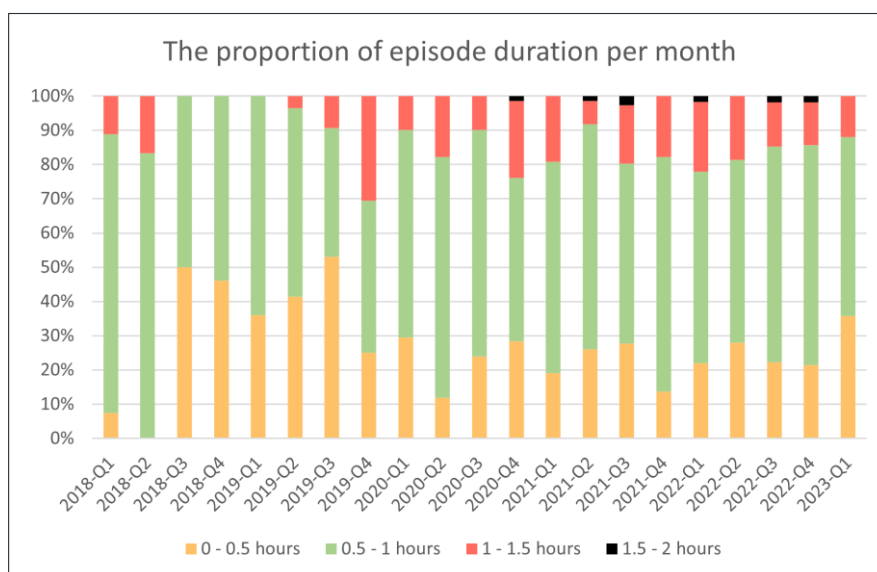
รูปภาพ 13 แผนภูมิฮิสโตแกรมแสดงการกระจายตัวของความยาววิดีโอ

จากแผนภูมิฮิสโตแกรมแสดงการกระจายตัวดังกล่าว พบว่าข้อมูลความยาวคลิปมีลักษณะเบ้ขวา (Skewed right) พบจำนวนวิดีโอในช่วงเวลา 30 – 45 นาทีต่อวิดีโอเป็นจำนวนมาก โดยข้อมูลจำนวนวิดีโอสามารถทำตารางแจกแจงความถี่ได้ดังนี้

ตาราง 6 ตารางแจกแจงความถี่ของความยาววิดีโอ

<i>Bin</i>	<i>Frequency</i>	<i>Bin</i>	<i>Frequency</i>
5	0	60	78
10	3	65	43
15	17	70	48
20	52	75	23
25	68	80	12
30	122	85	16
35	107	90	4
40	124	95	2
45	127	100	1
50	93	105	2
55	95	110	2

1.6. สัดส่วนความยาววิดีโอต่อไตรมาส



รูปภาพ 14 แผนภูมิแท่งแสดงสัดส่วนความยาววิดีโอในแต่ละไตรมาส

จากแผนภูมิแท่ง แสดงให้เห็นว่าวิดีโอส่วนใหญ่ในแต่ละไตรมาสภายในแซนเนล จะมีความยาววิดีโอในช่วงความยาว 30 – 60 นาที และมีช่วงความยาว 0 – 30 นาที เป็นลำดับถัดมา ในขณะที่มีความยาวตั้งแต่ 60 นาทีเป็นต้นไปค่อนข้างน้อย ซึ่งจากการพิจารณาจากค่าเฉลี่ยคิดเป็น 42 นาที 27 วินาที

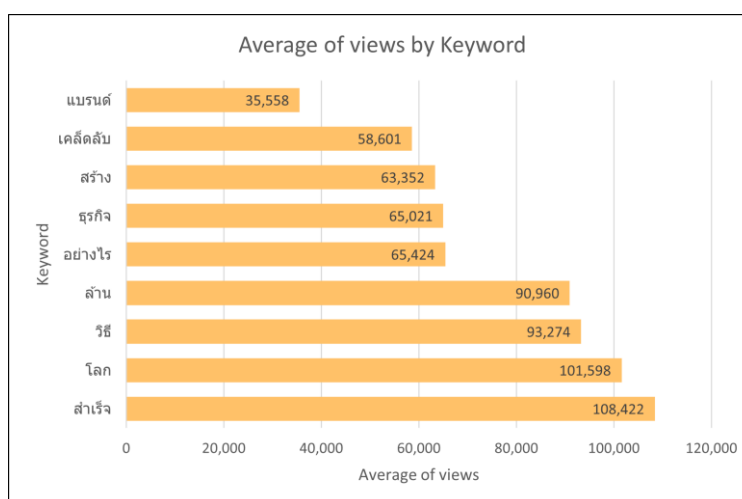
โดย Episode ที่มีความยาววิดีโอสั้นที่สุด 3 อันดับแรก ได้แก่

1. 8 นาที 35 วินาที - ต้องทำงานในวันที่ทุกคนไปเที่ยวกันหมด ทำอย่างไรให้ยังมีกำลังใจและประสิทธิภาพ The Secret Sauce EP.101 (13 เมษายน 2562)
2. 8 นาที 22 วินาที - วิธีแก้ Post-vacation blues เพราะหลังพักร้อนคือความเจ็บปวด | The Secret Sauce EP.104 (16 เมษายน 2562)
3. 5 นาที 47 วินาที - เคน นครินทร์ แนะนำรายการ The Secret Sauce Podcast | The Secret Sauce EP.0 (21 กุมภาพันธ์ 2561)

ขณะที่ Episode ที่มีความยาววิดีโอยาวที่สุด 3 อันดับแรก ได้แก่

1. 1 ชั่วโมง 47 นาที - สรุบบาทดิจิทัล มาแนปี่หน้า เตรียมพร้อมอย่างไร? | Executive Espresso EP.263 (2 กันยายน 2564)
2. 1 ชั่วโมง 46 นาที - ดีเบตคริปโต รัฐ vs. เอกชน ทำไมคริปโตชำระสินค้าไม่ได้? | Executive Espresso EP.318 (31 มกราคม 2565)
3. 1 ชั่วโมง 42 นาที - เคน นครินทร์ คุศลดี เอียววิทย์ เศรษฐกิจปี 2023 | The Secret Sauce [SPECIAL LIVE!] (31 ตุลาคม 2565)

1.7. ค่าเฉลี่ยจำนวนคนดูโดย Keyword ที่พบบ่อย 9 อันดับแรก

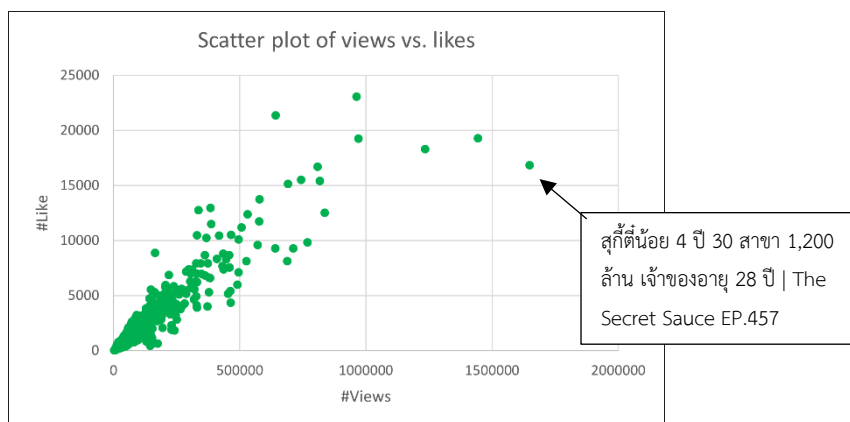


รูปภาพ 15 ค่าเฉลี่ยจำนวนคนดูโดย Keyword ที่พบบ่อย 9 อันดับแรก

จากแผนภูมิดังกล่าวแสดงให้เห็นว่าแชนเนล The Secret Sauce มีการนำข้อความ หรือคำต่าง ๆ มาใช้สร้างเป็นชื่อวิดีโอสำหรับการดึงดูดและสร้างแรงจูงใจสำหรับการเข้ามาชมวิดีโอเหล่านั้น โดยจากแผนภูมิแท่งคำว่า “สำเร็จ” และ “โลก” เป็นคำที่สามารถสร้างแรงจูงใจในการรับชมได้ค่อนข้างมาก เนื่องจากมีค่าเฉลี่ยของจำนวนคนดูที่สูงเป็นอันดับต้น ๆ และนอกจากนี้วิดีโอที่ปรากฏคำเหล่านี้มักเป็นวิดีโอที่เกี่ยวข้องกับธุรกิจ และแนวทางการศึกษาเป็นหลัก เช่น “เศรษฐกิจวงจรสมบูรณ์ของจีนเปลี่ยนโลก กระบไทยอย่างไร | Executive Espresso EP.188 และ “ศิลปะของการทำสิ่งสำคัญให้สำเร็จ | The Secret Sauce EP.347”⁸

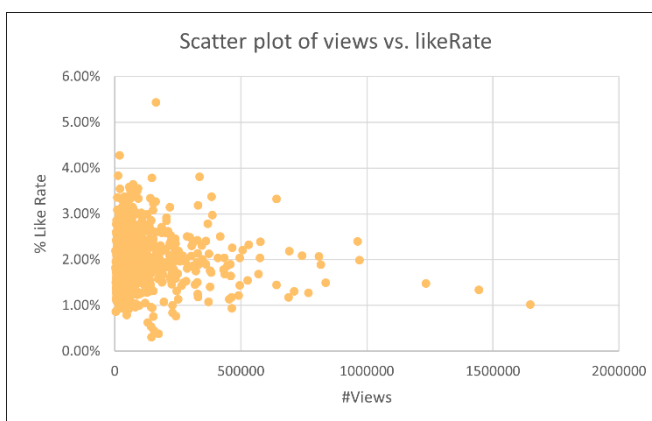
⁸ สามารถอ่านตารางแจกแจงข้อมูลแต่ละ Keyword ได้ในภาคผนวก

1.8. การกระจายระหว่างจำนวนคนดูและจำนวนคนกดถูกใจ



รูปภาพ 16 แผนภาพกระจายระหว่างจำนวนคนดู
และจำนวนการกดถูกใจในแต่ละคลิป

1.9. การกระจายระหว่างจำนวนคนดูและสัดส่วนการกดถูกใจ



รูปภาพ 17 แผนภาพกระจายระหว่างจำนวนคนดู
และสัดส่วนการกดถูกใจ

จากทั้งสองแผนภาพการกระจาย จะเห็นว่าจำนวนคนดู และจำนวนคนกดถูกใจ ในแต่ละคลิปมีความสัมพันธ์กันในทางบวก และในช่วง 0 – 1,000,000 คนรับชม จะมีความสัมพันธ์เชิงเส้นตรงเป็นส่วนมาก ในขณะที่เกินกว่า 1,000,000 คนรับชมจะมีการกระจายที่ไม่ใช่เส้นตรง ในขณะที่เดียวกันหากพิจารณาแผนภาพการกระจายระหว่างคนดู และสัดส่วนคนกดถูกใจ จะเห็นได้ว่าสัดส่วนคนกดถูกใจมีค่อนข้างน้อยมาก โดยคิดเป็น 2% ถึง 3% จากคนดูทั้งหมด

2. การวิเคราะห์ด้วยสถิติเชิงอนุมาน

2.1. การทดสอบสมมติฐานระหว่างค่าเฉลี่ยของจำนวนคนรับชมวิดีโอแบบมีพิธีกรรับเชิญ และไม่มีพิธีกรรับเชิญในแต่ละวิดีโอ

(Hypothesis Testing with 2 Samples)

ขั้นตอนที่ 1: สรุปข้อมูลจำนวนคนรับชมวิดีโอโดยเฉลี่ยได้ดังนี้

ตาราง 7 ตารางสรุปจำนวนคนรับชมวิดีโอเฉลี่ย
โดยแบ่งตามการมีพิธีกรรับเชิญ

	จำนวนคนรับชมวิดีโอ	
	มี (1)	ไม่มี (2)
ค่าเฉลี่ย (\bar{X})	89893.90	86190.23
ส่วนเบี่ยงเบนมาตรฐาน (s)	128250.73	141426.15
จำนวน (n)	31	1008

ขั้นตอนที่ 2: การทดสอบสมมติฐาน

วิธีทำ $H_0 : \mu_1 \leq \mu_2$

$H_a : \mu_1 > \mu_2$ (Upper tailed test)

กำหนดระดับนัยสำคัญ $\alpha = 0.05$

เนื่องจากจำนวนข้อมูลทั้งสองมีค่ามากกว่า 30 ข้อมูล

โดยทฤษฎีบทลิมิตเข้าสู่ศูนย์กลาง (Central Limit Theorem) จะได้ว่าทั้งสองกลุ่ม-

ข้อมูลมีแนวโน้มที่จะมีการแจกแจงเป็นการแจกแจงปกติ

ดังนั้น จึงใช้การทดสอบแบบ Z – Test

หาค่าสถิติทดสอบ (Test Statistics)

$$\begin{aligned} \text{จาก } Z_{ob} &= \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \\ &= \frac{(89893.90 - 86190.23) - 0}{\sqrt{\frac{128250.73^2}{31} + \frac{141426.15^2}{1008}}} \end{aligned} \quad (4.2.1)$$

$$\therefore Z_{ob} = 0.156$$

เพราะฉะนั้น

$$\begin{aligned} P - \text{value} &= P(Z > Z_{ob}) \\ &= P(Z > 0.156) \\ &= 0.438 \end{aligned} \quad (4.2.2)$$

เนื่องจาก $P - \text{value} > \alpha$

ดังนั้น ไม่สามารถที่จะปฏิเสธ H_0 ได้ที่ระดับนัยสำคัญ 0.05 หรือ

ระดับความเชื่อมั่นที่ 95% (Do not Reject H_0)

สรุปได้ว่า

“โดยเฉลี่ยจำนวนคนดูรูปแบบไม่มีพิธีกรรับเชิญมากกว่าหรือเท่ากับแบบมีพิธีกรรับเชิญ”

หมายเหตุ :

จากการทดสอบสมมติฐานดังกล่าว แสดงให้เห็นว่า ในรูปแบบรายการ Podcast ที่มีหลากหลายแตกต่างกันไป โดยหนึ่งสิ่งที่สร้างมุมมองการรับชมวิดีโอของผู้คนที่แตกต่างจากเนื้อหาที่นำเสนอแล้ว ยังมีในส่วนของผู้จัดรายการที่มีส่วนทำให้แตกต่าง ทั้งนี้เป็นเพราะในรูปแบบที่มีพิธีกรรับเชิญหรือมีผู้จัดรายการมากกว่าหนึ่งคน รูปแบบการนำเสนอ มักจะเป็นรูปแบบการสัมภาษณ์เป็นส่วนใหญ่ ซึ่งพฤติกรรมของคนไทยอาจมองว่าเป็นการจัดรูปแบบที่วุ่นวายและเข้าใจเนื้อหาได้ยาก หากพิธีกรรับเชิญไม่คุ้นชินกับการนำเสนอสื่อนอกจากนี้ยังรวมไปถึงความกังวลในความไม่ต่อเนื่องของการนำเสนอข้อมูล ปัญหาต่าง ๆ ที่อาจเกิดขึ้นระหว่างการจัดทำรายการ เช่น คุณภาพเสียงที่อาจไม่เท่ากัน เป็นต้น

สรุปขั้นตอนการทดสอบสมมติฐาน (ขั้นตอนที่ 2)

1. กำหนดให้

$$H_0 : \mu_1 \leq \mu_2$$

$$H_a : \mu_1 > \mu_2 \text{ (Upper tailed test)}$$

2. Test Statistics: $Z_{ob} = 0.156$

3. ตรวจสอบ Test Statistics ด้วยสองวิธี

3.1. P – value

$$\begin{aligned} P - \text{value} &= P(Z > Z_{ob}) \\ &= 0.438 \\ &> \alpha \end{aligned}$$

3.2. Critical Region Approach

$$\begin{aligned} Z_{\text{critical}} &= Z_{1-\alpha} = 0.95 = 0.171 \\ Z_{\text{critical}} &> Z_{ob} \end{aligned}$$

4. สรุปผล

เนื่องจาก $P - \text{value} > \alpha$ หรือ $Z_{\text{critical}} > Z_{ob}$

ดังนั้น ไม่สามารถที่จะปฏิเสธ H_0 ได้ที่ระดับนัยสำคัญ 0.05 หรือระดับความเชื่อมั่นที่ 95% (Do not Reject H_0)#

รูปภาพ 18 สรุปขั้นตอนการทำทดสอบสมมติฐานของจำนวนคน

รับชมวิดีโอระหว่างแบบมีพิธีกรรับเชิญ และไม่มีพิธีกรรับเชิญ

2.2. การทดสอบความเป็นอิสระต่อกันระหว่างหมวดหมู่ Podcast และความยาวคลิปด้วย

การทดสอบแบบไคสแควร์

(Test of Independent with Chi-Square Test)

ขั้นตอนที่ 1: สรุปข้อมูลทั้งหมดให้อยู่ในรูปแบบของตารางระหว่างหมวดหมู่
และความยาวคลิป

ตาราง 8 ตารางแสดงข้อมูลระหว่างหมวดหมู่ Podcast และความยาวคลิป

Duration Range	Category						Total _i
	เศรษฐกิจ	Life Style	ธุรกิจและการค้า	สินค้าและบริการ	การเมือง	อื่น ๆ	
0.5 - 1 hours	85	143	156	96	47	82	609
0 - 0.5 hours	27	82	55	19	15	52	250
1 - 1.5 hours	14	22	34	13	13	21	117
1.5 - 2 hours	12	10	12	8	10	11	63
Total _j	138	257	257	136	85	166	1039

ขั้นตอนที่ 2: การทดสอบสมมติฐานการเป็นอิสระต่อกัน

วิธีทำ H_0 : หมวดหมู่ Podcast ไม่มีความสัมพันธ์กับความยาวคลิป

H_a : H_0 เป็นเท็จ (ทั้งสองมีความสัมพันธ์กัน)

กำหนดระดับนัยสำคัญ $\alpha = 0.05$

จาก

$$\text{ค่าคาดหวัง } e_{i/j} = \frac{\text{Total}_i \times \text{Total}_j}{\text{Total}} \quad (4.2.3)$$

ตัวอย่างเช่น

$$e_{\text{เศรษฐกิจ} / 0.5 - 1 \text{ hours}} = \frac{615 \times 137}{1039} = 81.0924$$

$$e_{\text{การเมือง} / 1 - 1.5 \text{ hours}} = \frac{83 \times 133}{1039} = 10.6246$$

เมื่อทำเช่นนั้นจนครบทั้ง 24 ข้อมูลจะได้ค่าคาดหวังแต่ละตำแหน่ง ดังตาราง

ตาราง 9 ตารางแสดงค่าคาดหวังระหว่างหมวดหมู่ และความยาวคลิบ

Duration Range	Category						Total _j
	เศรษฐกิจ	Life Style	ธุรกิจและการค้า	สินค้าและบริการ	การเมือง	อื่น ๆ	
0.5 - 1 hours	80.8874	150.6381	150.6381	79.7151	49.8219	97.2993	609
0 - 0.5 hours	33.2050	61.8383	61.8383	32.7238	20.4524	39.9423	250
1 - 1.5 hours	15.5399	28.9403	28.9403	15.3147	9.5717	18.6930	117
1.5 - 2 hours	8.3677	15.5833	15.5833	8.2464	5.1540	10.0654	63
Total _j	138	257	257	136	85	166	1039

ทั้งนี้เพื่อเป็นไปตามข้อกำหนดของการทดสอบแบบไคสแควร์ที่ว่า “ค่าคาดหวัง (Expected value) $e_{i,j}$ จำเป็นต้องมากกว่าเท่ากับ 5 สำหรับทุกช่องในตาราง” ซึ่งจากการคำนวณในตารางดังกล่าว พบว่าทุกช่องมีค่ามากกว่าเท่ากับ 5 ซึ่งสอดคล้องกับข้อกำหนด

หาค่าสถิติทดสอบ (Test Statistics)

$$\begin{aligned}
 \text{จาก } \chi_{ob}^2 &= \sum_{\text{all cells}} \frac{(\text{observer}-\text{expected})^2}{\text{expected}} & (4.2.4) \\
 &= \sum_{j=1}^6 \sum_{i=1}^4 \frac{(o_{i,j}-e_{i,j})^2}{e_{i,j}} \\
 &= \frac{(85-80.8874)^2}{80.8874} + \frac{(143-150.6381)^2}{150.6381} + \\
 &\quad \frac{(156-150.6381)^2}{150.8381} + \dots + \frac{(10-5.1540)^2}{5.1540} + \\
 &\quad \frac{(11-10.0654)^2}{10.0654}
 \end{aligned}$$

$$\therefore \chi_{ob}^2 = 39.6336, df = (4 - 1)(6 - 1) = 15$$

หมายเหตุ: สำหรับตารางแสดงค่า $\frac{(\text{observer}-\text{expected})^2}{\text{expected}}$ จะแสดงบน

ตาราง 28 ตารางแสดงการหาค่าไคสแควร์ ในส่วนของภาคผนวก

เพราะฉะนั้น

$$\begin{aligned}
 P - value &= P(\chi^2 > \chi_{ob}^2, df) \\
 &= P(\chi^2 > 39.6336, df = 15) \\
 &= 0.000515
 \end{aligned}
 \tag{4.2.5}$$

เนื่องจาก $P - value < \alpha$

ดังนั้น สามารถที่จะปฏิเสธ H_0 ได้ที่ระดับนัยสำคัญ 0.05 หรือ

ระดับความเชื่อมั่นที่ 95% (Reject H_0)

สรุปได้ว่า

“ข้อมูลระหว่างหมวดหมู่ Podcast และความยาวคลิปมีความสัมพันธ์กัน หรือไม่เป็นอิสระต่อกัน”

สรุปขั้นตอนการทดสอบสมมติฐาน (ขั้นตอนที่ 2)

1. กำหนดให้

H_0 : หมวดยุค Podcast ไม่มีความสัมพันธ์กับความยาวคลิป

H_a : H_0 เป็นเท็จ (ทั้งสองมีความสัมพันธ์กัน)

2. Test Statistics: $\chi^2_{ob} = 39.6336$, $df = 15$

3. ตรวจสอบ Test Statistics ด้วยสองวิธี

3.1. P – value

$$\begin{aligned} P - \text{value} &= P(\chi^2 > \chi^2_{ob} = 39.6336, df = 15) \\ &= 0.000515 \\ &< \alpha \end{aligned}$$

3.2. Critical Region Approach

$$\begin{aligned} \chi^2_{critical} &= \chi^2_{\alpha = 0.05, df = 15} = 24.9958 \\ \chi^2_{critical} &< \chi^2_{ob} \end{aligned}$$

4. สรุปผล

เนื่องจาก $P - \text{value} < \alpha$ หรือ $\chi^2_{critical} < \chi^2_{ob}$

ดังนั้น เราสามารถที่จะปฏิเสธ H_0 ได้ที่ระดับความเชื่อมั่น 95%

หรือ เลือกที่จะยอมรับ H_a ที่ระดับความเชื่อมั่นดังกล่าว#

รูปภาพ 19 สรุปขั้นตอนการทำทดสอบสมมติฐานด้วยการทดสอบแบบไคสแควร์

2.3. การวิเคราะห์ความแปรปรวนของข้อมูล

(Analysis of Variance; ANOVA)

2.3.1. การวิเคราะห์ความแปรปรวนระหว่างหมวดหมู่ Podcast และปีที่อัปโหลด

ขั้นตอนที่ 1: สรุปข้อมูลจำนวนคนดูโดยเฉลี่ยในแต่ละปีในรูปแบบตารางดังนี้

ตาราง 10 ตารางแสดงจำนวนคนดูโดยเฉลี่ย
ในแต่ละประเภทและแต่ละหมวดหมู่

Year	Category					
	เศรษฐกิจ	Life Style	ธุรกิจ และการค้า	สินค้า และบริการ	การเมือง	อื่น ๆ
2018	44439.75	90310.89	35296.38	24706.20	0.00	121893.61
2019	44093.43	114938.83	54039.95	26648.28	12474.67	99039.98
2020	42410.36	52849.45	54842.54	44262.52	74284.73	75540.48
2021	60689.53	78599.79	122352.09	58778.70	168037.21	65434.67
2022	134593.49	94336.11	96957.07	140157.48	245826.42	86904.27
2023	106431.00	150040.54	95540.31	82697.50	19606.00	160613.67

ขั้นตอนที่ 2: การทดสอบสมมติฐานด้วย ANOVA

วิธีทำ สำหรับ Treatment (Category)

$$H_{0,t} : \mu_{1,t} = \mu_{2,t} = \mu_{3,t} = \mu_{4,t} = \mu_{5,t} = \mu_{6,t}$$

$$H_{a,t} : H_{0,t} \text{ เป็นเท็จ}$$

สำหรับ Block (Year)

$$H_{0,b} : \mu_{1,b} = \mu_{2,b} = \mu_{3,b} = \mu_{4,b} = \mu_{5,b} = \mu_{6,b}$$

$$H_{a,b} : H_{0,b} \text{ เป็นเท็จ}$$

กำหนดระดับนัยสำคัญ $\alpha = 0.05$

โดยในที่นี้จะใช้การวิเคราะห์ในรูปแบบ The Randomized Block Design (ANOVA: Two ways without Replication) ซึ่งจากการใช้คำสั่ง Data Analysis บนโปรแกรม Microsoft Excel จะให้ผลลัพธ์ดังตาราง

ตาราง 11 ตาราง ANOVA สำหรับค่าเฉลี่ยคนดูในแต่ละหมวดหมู่ และแต่ละปี

Anova: Two-Factor Without
Replication

<i>SUMMARY</i>	<i>Count</i>	<i>Sum</i>	<i>Average</i>	<i>Variance</i>
2018	6	316646.8	52774.47	2026872381
2019	6	351235.1	58539.19	1637825354
2020	6	344190.1	57365.01	207738922.6
2021	6	553892	92315.33	1934318558
2022	6	798774.8	133129.1	3540455076
2023	6	614929	102488.2	2592948906
เศรษฐกิจ	6	432657.6	72109.59	1529036797
Life Style	6	581075.6	96845.94	1094927381
ธุรกิจและการค้า	6	459028.3	76504.72	1110935361
สินค้าและบริการ	6	377250.7	62875.11	1899590156
การเมือง	6	520229	86704.84	9923798687
อื่น ๆ	6	609426.7	101571.1	1220797330

ANOVA

<i>Source of Variation</i>	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>P-value</i>	<i>F crit</i>
Year (B)	3.09E+10	5	6.18E+09	2.913768841	0.033095	2.602987
Category (T)	6.69E+09	5	1.34E+09	0.631514222	0.677454	2.602987
Error	5.3E+10	25	2.12E+09			
Total	9.06E+10	35				

จากตาราง 11 ตาราง ANOVA สำหรับค่าเฉลี่ยคนดูในแต่ละหมวดหมู่ และแต่ละปี

พบว่าได้ $P - \text{value}_{\text{block}} = 0.033095$

และ $P - \text{value}_{\text{treatment}} = 0.677454$

ซึ่ง $P - \text{value}_{\text{block}} < \alpha$ แต่ $P - \text{value}_{\text{treatment}} > \alpha$

ดังนั้น ไม่สามารถที่จะปฏิเสธ $H_{0,t}$ แต่สามารถจะปฏิเสธ $H_{0,b}$ ที่ระดับนัยสำคัญ 0.05

สรุปได้ว่า

“หมวดหมู่ของวิดีโอ Podcast ไม่ส่งผลต่อจำนวนคนดูโดยเฉลี่ย
ในขณะที่เวลา (ปี) ส่งผลต่อจำนวนคนดูโดยเฉลี่ย”

หมายเหตุ :

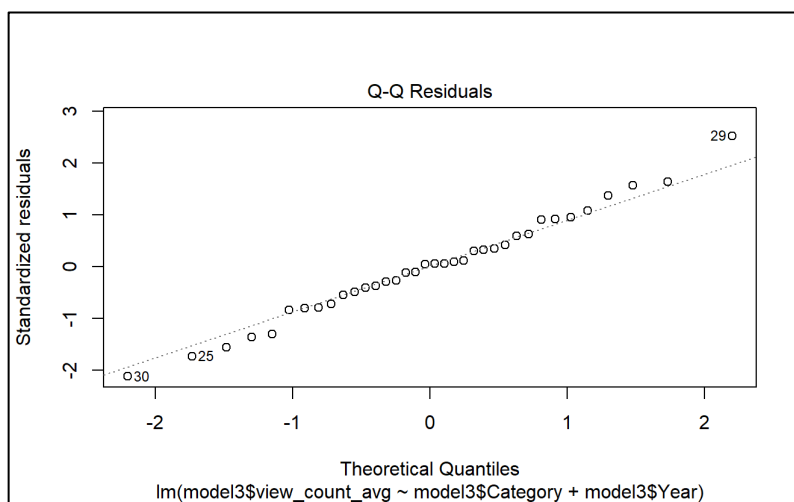
การวิเคราะห์ความแปรปรวนนี้สามารถอธิบายได้ว่าค่าเฉลี่ยจำนวนคนรับชมวิดีโอในแต่ละหมวดหมู่ตลอด 5 ปี 28 วันมีค่าเท่ากัน (เนื่องจากการทดสอบดังกล่าว ได้ยอมรับว่า $H_{0,t}: \mu_{1,t} = \mu_{2,t} = \mu_{3,t} = \mu_{4,t} = \mu_{5,t} = \mu_{6,t}$ เป็นจริง) ซึ่งหมายความว่าหมวดหมู่วิดีโอทั้ง 6 หมวดหมู่นี้ไม่ได้ส่งผลต่อจำนวนคนดูโดยเฉลี่ย แต่ในขณะเดียวกันในทุก ๆ ปีค่าเฉลี่ยของคนดูกลับไม่เท่ากันซึ่งแสดงว่าจำนวนคนรับชมวิดีโอในแต่ละปีมีการเปลี่ยนแปลงตามเวลาซึ่งสมเหตุสมผลกับความเป็นจริง โดยที่สามารถสะท้อนถึงผลของการพิจารณาส่วนความแปรปรวนของ Block ได้

และจากผลของการทดสอบสมมติฐานดังกล่าว และหากอ้างอิงจากแผนภูมิแท่งแสดงจำนวนคนรับชมวิดีโอโดยเฉลี่ยโดยแบ่งตามหมวดหมู่ จะเห็นได้ว่าในทุก ๆ หมวดหมู่นั้นมีลักษณะการเปลี่ยนแปลงไปในทิศทางเดียวกัน

ขั้นตอนที่ 3: การทดสอบข้อกำหนด หรือข้อตกลงสำหรับ ANOVA

(Check Assumption for ANOVA)

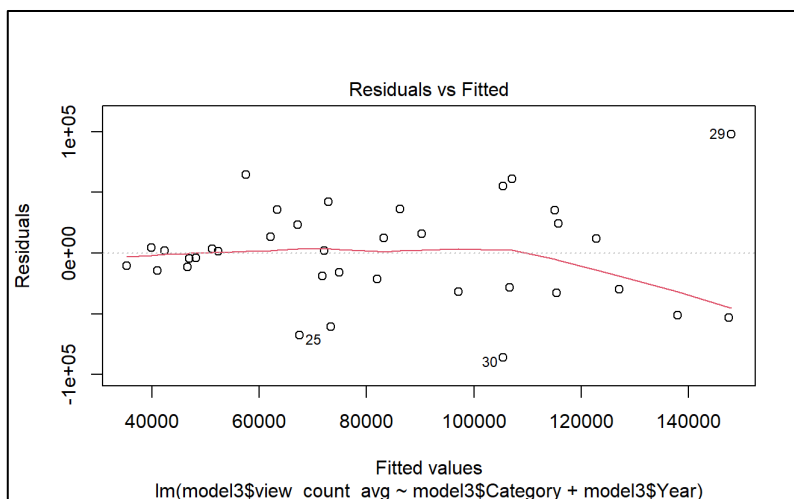
1. Normal Probability Plot of Residuals



รูปภาพ 20 Normal Probability Plot of Residuals

สำหรับ ANOVA (2.3.1)

2. Plot of Residual vs Fitted value



รูปภาพ 21 Plot of Residuals vs Fitted Value

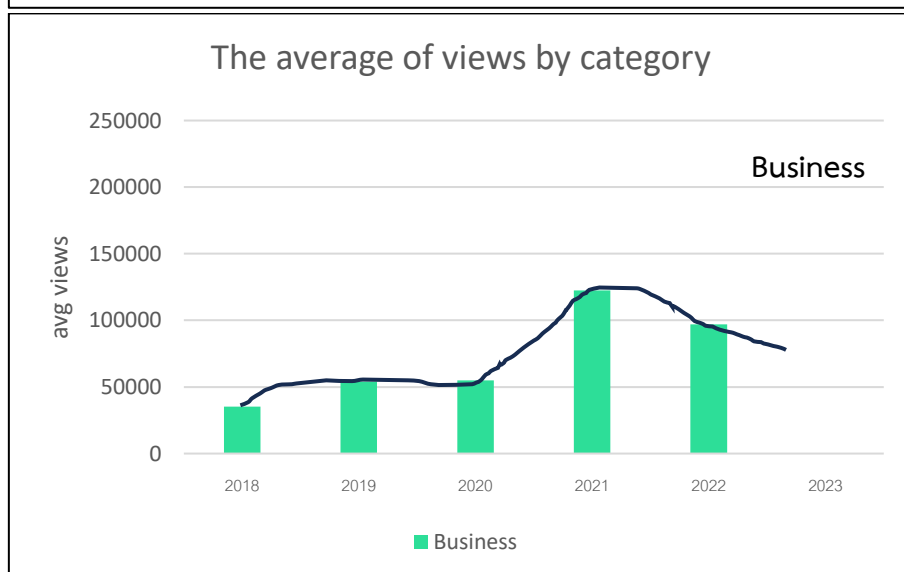
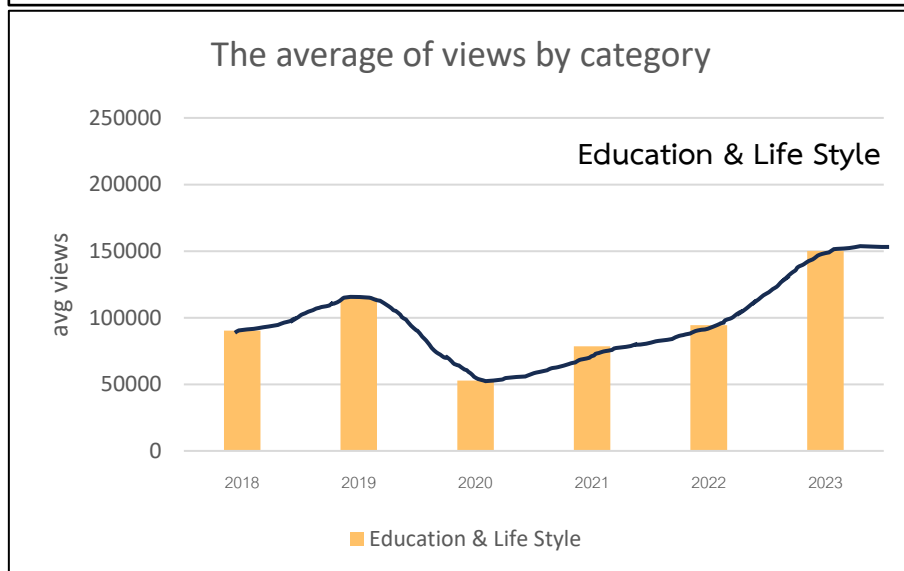
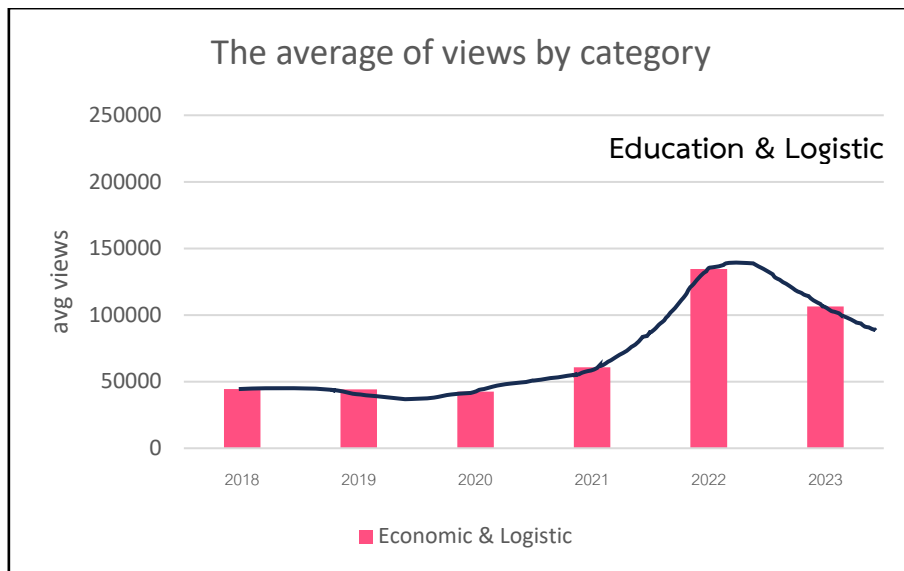
สำหรับ ANOVA (2.3.1)

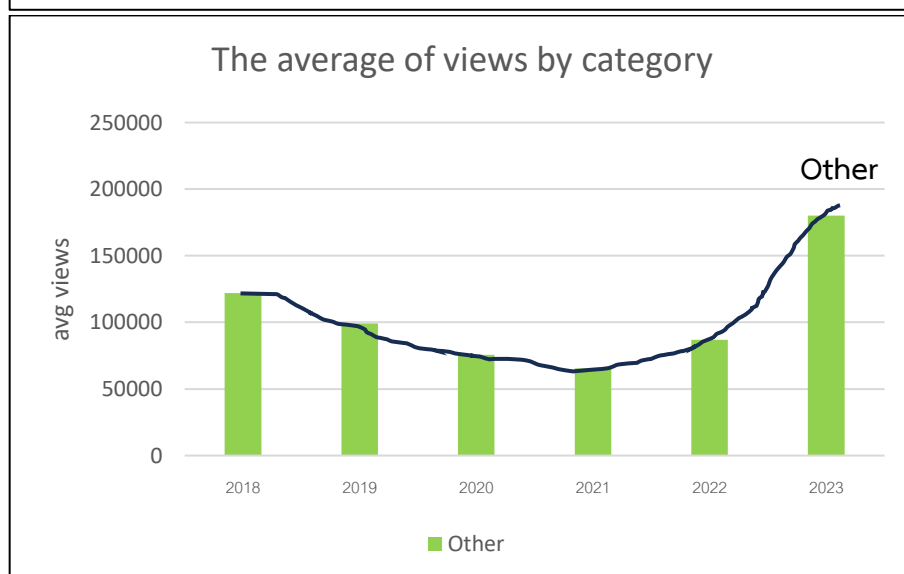
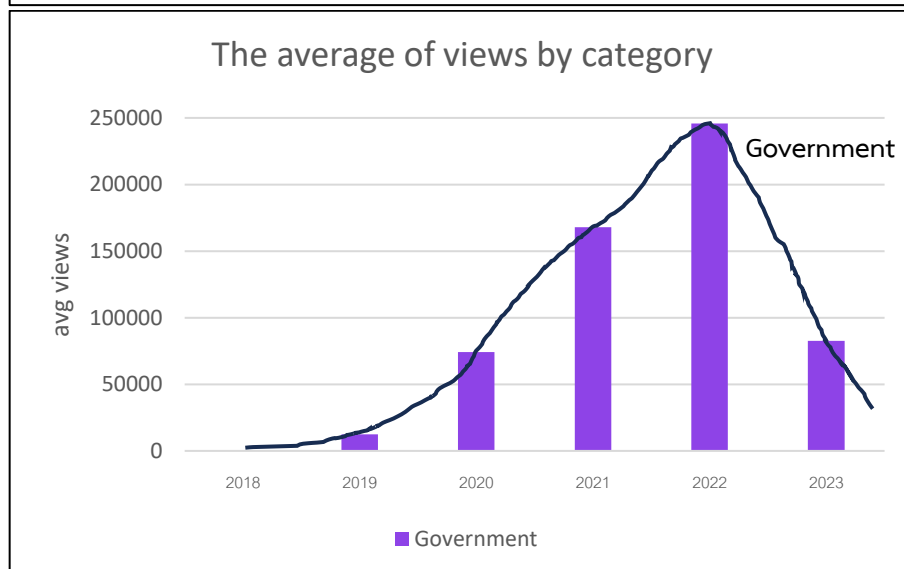
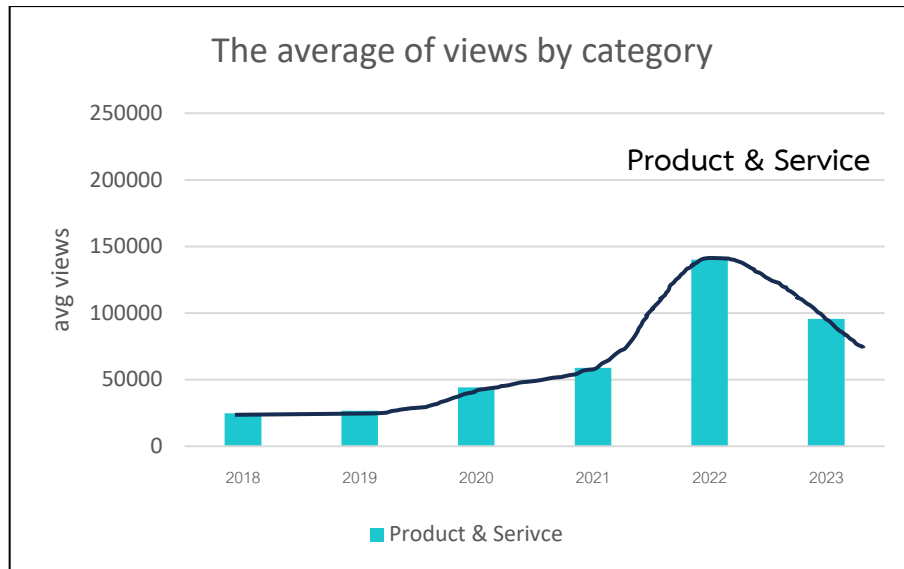
ตาราง 12 ตารางสรุปข้อมูลที่ได้จากกราฟ
สำหรับการตรวจสอบสมมติฐาน ANOVA (2.2.1)

กราฟ	คำอธิบาย (ข้อสรุป)
Normal Probability Plot of Residuals	ข้อมูลชุดนี้มีการแจกแจงที่สามารถอนุมานได้ว่าเป็นการแจกแจงปกติ เนื่องจากทุก ๆ จุดบนกราฟอยู่ใกล้เคียง หรือทับเส้นตรง
Residual VS Fitted Value	ข้อมูลมีค่า Error (Residuals) โดยเฉลี่ยแล้วไม่เป็นศูนย์ รวมถึงไม่มีการกระจายตัวระหว่าง Residuals ค่าบวก และลบที่ใกล้เคียงกัน สังเกตได้จากเส้นสีแดงที่มีการลดลงไปทางลบในช่วงหลัง

ขั้นตอนที่ 4: สรุปขั้นตอน

จากการวิเคราะห์ความแปรปรวนระหว่างหมวดหมู่ Podcast และค่าเฉลี่ยของคนดูในแต่ละปี ดังตารางที่ 9 โดยใช้การทดสอบแบบ F – Test พบว่าประเภทของวิดีโอไม่ได้ส่งผลต่อจำนวนคนดูคลิปโดยเฉลี่ยในแต่ละปี หรือก็คือหมวดหมู่ไม่ได้เป็นปัจจัยหลักสำคัญสำหรับค่าเฉลี่ยของจำนวนคนดูในแต่ละปี ทั้งนี้สามารถอธิบายได้ว่าคนไทยมีความสนใจในการฟัง Podcast ในแต่ละหมวดหมู่เท่า ๆ กันในแต่ละปี ซึ่งเป็นเหตุผลเพียงพอในการสนับสนุนผลที่ได้จากแผนภูมิแท่งแสดงค่าเฉลี่ยของจำนวนคนดูโดยแบ่งตามหมวดหมู่ในหัวข้อที่ 4.1.1.3 ที่เมื่อได้พิจารณาแนวโน้มของจำนวนคนดูโดยเฉลี่ยต่อปีแล้วในทุก ๆ หมวดหมู่จะมีลักษณะไปในทิศทางเดียวกัน เว้นแต่หมวดหมู่ Other ที่มีลักษณะแนวโน้มผิดไปจากหมวดหมู่อื่น ๆ





2.3.2. การวิเคราะห์ความแปรปรวนระหว่าง 3 ปัจจัยต่าง ๆ ที่เกี่ยวข้อง

โดยในการวิเคราะห์นี้จะใช้หลักการ ANOVA: Three Ways with Replication ซึ่งประกอบด้วย

1. หมวดหมู่ของ Podcast ได้แก่ เศรษฐกิจ, Life Style, ธุรกิจและการค้า, สินค้าและบริการ, การเมือง, และอื่น ๆ (6 จำพวก)
2. ไตรมาสของการอัปโหลดวิดีโอ (4 จำพวก)
3. ช่วงความยาววิดีโอ ได้แก่ 0 – 30 นาที, 30 – 60 นาที, 60 – 90 นาที และ 90 – 120 นาที (4 จำพวก)
4. ปีที่อัปโหลดวิดีโอ ตั้งแต่ พ.ศ. 2561 จนถึง 2565 (5 จำพวก)

ซึ่งจะมีจำนวนข้อมูลทั้งหมด 480 ข้อมูล

ขั้นตอนที่ 1: สรุปข้อมูลจำนวนคนดูโดยเฉลี่ยได้ดังนี้

ตาราง 13 ตารางสรุปข้อมูล (เบื้องต้น) ของปัจจัยต่าง ๆ

quarter	category	duration	year	response_avg
1	เศรษฐกิจ	0 - 0.5 hours	2018	4718
1	เศรษฐกิจ	0.5 - 1 hours	2018	0
1	เศรษฐกิจ	1 - 1.5 hours	2018	67036
1	เศรษฐกิจ	1.5 - 2 hours	2018	0
.
.
1	เศรษฐกิจ	0.5 - 1 hours	2019	0
.
.
1	Life Style	0 - 0.5 hours	2018	0
.
.
2	เศรษฐกิจ	0 - 0.5 hours	2018	0
.
.
4	อื่น ๆ	1 - 1.5 hours	2022	24818
4	อื่น ๆ	1.5 - 2 hours	2022	463810

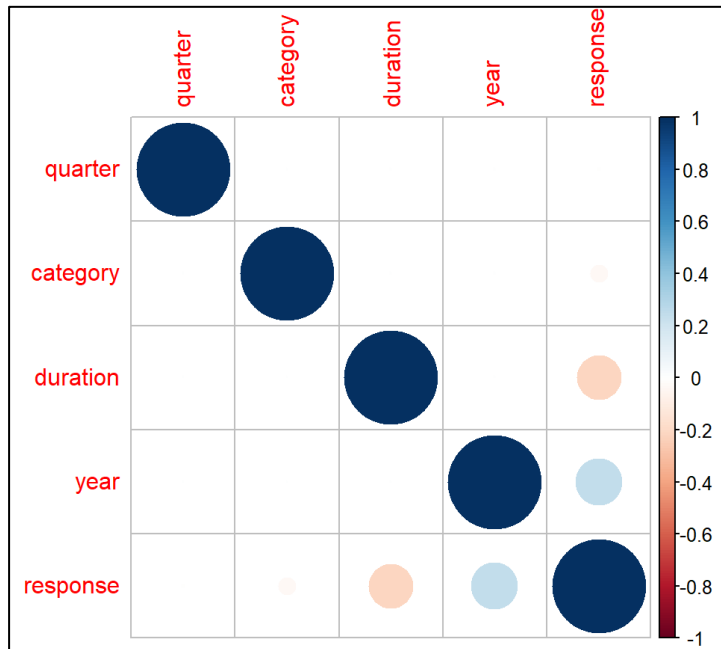
ขั้นตอนที่ 2: พิจารณาความสัมพันธ์ระหว่างตัวแปรต่าง ๆ

เมื่อพิจารณาความสัมพันธ์ระหว่างตัวแปรเป็นคู่ ๆ โดยใช้ค่าสหสัมพันธ์ (Co-relation; r) ซึ่งสามารถใช้อธิบายการกระจายระหว่างชุดข้อมูล 2 ข้อมูลได้ โดยสามารถสรุปข้อมูลได้ดังตาราง

ตาราง 14 ตารางแสดงค่าสหสัมพันธ์ระหว่างปัจจัยต่าง ๆ

	quarter	category	duration	year	response_avg
quarter	1.00	0.00	0.05	0.00	0.02
category	0.00	1.00	0.00	0.00	0.00
duration	0.00	0.05	1.00	0.04	-0.20
year	0.00	0.00	0.04	1.00	0.30
response_avg	0.02	0.00	-0.20	0.30	1.00

โดยสามารถนำมาสร้างแผนภูมิได้ดังรูป



รูปภาพ 22 แผนภาพแสดงความสัมพันธ์
ระหว่างปัจจัยต่าง ๆ ที่เกี่ยวข้อง

ขั้นตอนที่ 3: การทดสอบสมมติฐานด้วย ANOVA

วิธีทำ $H_{0,i} : \mu_{1,i} = \mu_{2,i} = \mu_{3,i} = \mu_{4,i} = \mu_{5,i} = \mu_{6,i}$

$H_{a,i} : H_{0,i}$ เป็นเท็จ, สำหรับทุก ๆ $i \in [\text{Factors} \times \text{Factors}]$

กำหนดระดับนัยสำคัญ $\alpha = 0.05$

โดยจากการใช้คำสั่งบนโปรแกรม R – Studio จะได้ผลลัพธ์ดังตาราง

ตาราง 15 ตาราง ANOVA สำหรับแสดงปัจจัยต่าง ๆ

กับค่าเฉลี่ยจำนวนคนดู

ANOVA: Three-factors

with Replication

Source of Variation	SS	df	MS	F	P-value
quarter	1.345E+12	1	1.345E+12	12.6723	0.000
category	6.814E+10	1	6.814E+10	0.6423	0.423
duration	2.727E+12	1	2.727E+12	25.7071	5.731E-07
year	3.477E+12	1	3.477E+12	32.7736	1.850E-08
quarter:category	2.739E+10	1	2.739E+10	0.2582	0.612
quarter:duration	2.980E+10	1	2.980E+10	0.2809	0.596
quarter:year	3.343E+11	1	3.343E+11	3.1514	0.077
category:duration	8.920E+09	1	8.920E+09	0.0841	0.772
category:year	7.678E+11	1	7.678E+11	7.2365	0.007
duration:year	1.164E+11	1	1.164E+11	1.0975	0.295
Error	4.976E+13	469	1.061E+11		
Total	5.733E+13	479			

จากตาราง 15 ตาราง ANOVA สำหรับแสดงปัจจัยต่าง ๆ
กับค่าเฉลี่ยจำนวนคนดู
พบว่า

$H_{0,i}$	P - value	ข้อสรุป
quarter	0.000	Reject H_0
category	0.423	-
duration	5.731E-07	Reject H_0
year	1.850E-08	Reject H_0
quarter:category	0.612	-
quarter:duration	0.596	-
quarter:year	0.077	-
category:duration	0.772	-
category:year	0.007	Reject H_0
duration:year	0.295	-

สรุปได้ว่า

“ปัจจัยที่ส่งผลต่อจำนวนคนรับชมวิดีโอ Podcast ที่พิจารณานั้น ได้แก่ ช่วงไตรมาสที่อัปโหลด, ความยาววิดีโอ และหมวดหมู่ของ Podcast นั้นมีเพียง**ช่วงไตรมาสและความยาววิดีโอ**เท่านั้นที่ส่งผลต่อจำนวนคนรับชมวิดีโอ”

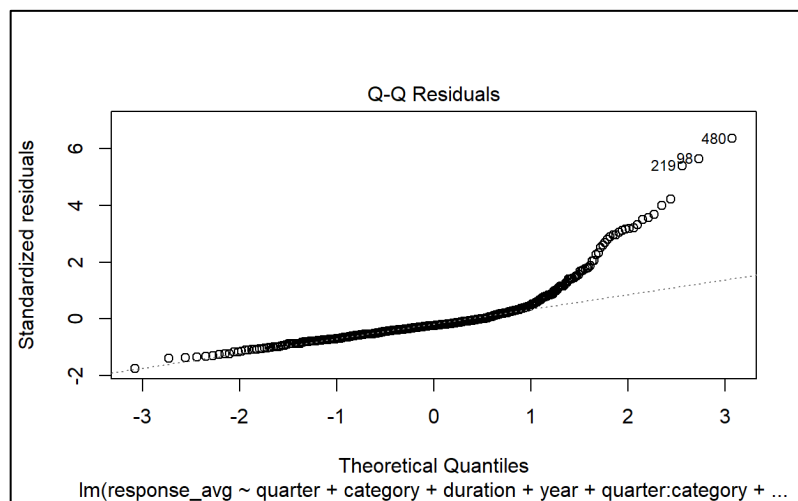
หมายเหตุ :

จากการวิเคราะห์ความแปรปรวนนี้ สามารถอธิบายได้ว่าช่วงไตรมาสที่อัปโหลด วิดีโอส่งผลต่อจำนวนคนรับชมวิดีโอ เนื่องจากในแต่ละช่วงไตรมาสของแต่ละปีนั้นจะมีการนำเสนอข้อมูลหรือสื่อต่าง ๆ ที่ไม่เท่ากัน เช่น ในช่วงไตรมาสที่ 2 จะมีการนำเสนอเกี่ยวกับธุรกิจมากเป็นพิเศษ รวมถึงจำนวนกลุ่มคนดูที่เปลี่ยนแปลงในแต่ละไตรมาส เนื่องจากเวลาและความสะดวกของแต่ละกลุ่ม อาทิ ช่วงไตรมาสที่ 2 จะมีกลุ่มนักเรียนและนักศึกษาที่เพิ่มขึ้นจากช่วงปิดเทอม เป็นต้น ในขณะที่ความยาววิดีโอซึ่งเป็นอีกหนึ่งปัจจัยที่ส่งผลนั้น พบว่าในช่วงที่ความยาววิดีโอมากเกินกว่า 60 นาทีจะมีคนดูที่น้อย เนื่องจากความไม่สะดวกและการใช้ชีวิตที่เร่งรีบของผู้คนในปัจจุบัน

ขั้นตอนที่ 4: การทดสอบข้อกำหนด หรือข้อตกลงสำหรับ ANOVA

(Check Assumption for ANOVA)

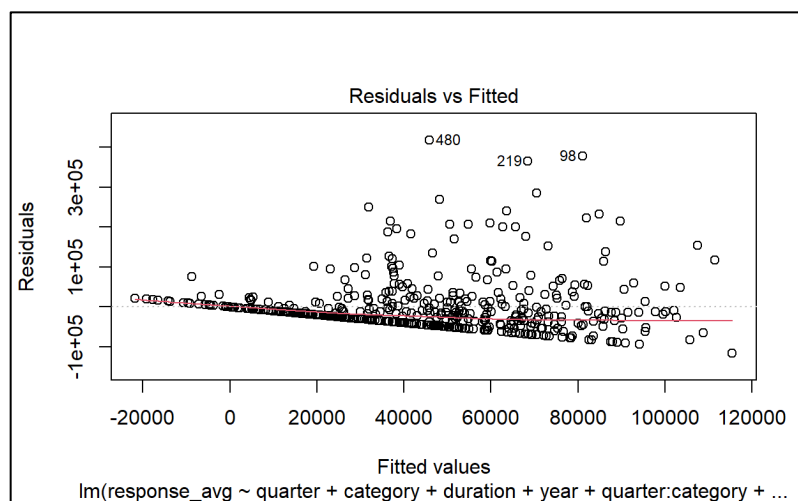
1. Normal Probability Plot of Residuals



รูปภาพ 23 Normal Probability Plot of Residuals

สำหรับ ANOVA (2.3.2)

2. Plot of Residual vs Fitted value



รูปภาพ 24 Plot of Residual vs Fitted value

สำหรับ ANOVA (2.3.2)

ตาราง 16 ตารางสรุปข้อมูลที่ได้จากกราฟ
จากการตรวจสอบสมมติฐานของ ANOVA (2.2.2)

กราฟ	คำอธิบาย (ข้อสรุป)
Normal Probability Plot of Residuals	ข้อมูลชุดนี้มีการแจกแจงที่สามารถอนุมานได้ว่าเป็นการแจกแจงปกติได้เพียงช่วงต้น ๆ เท่านั้น โดยที่ภายหลังจะมีการกระจายอยู่ห่างจากเส้นตรง $y = x$ เป็นอย่างมาก
Residual VS Fitted Value	ข้อมูลมีค่า Error (Residuals) โดยเฉลี่ยแล้วไม่เป็นศูนย์ รวมถึงไม่มีการกระจายตัวระหว่าง Residuals ค่าบวก และลบที่ใกล้เคียงกัน สังเกตได้จากเส้นสีแดงที่มีการลดลงไปทางลบในช่วงหลัง

ขั้นตอนที่ 5: สรุปขั้นตอน

จากการวิเคราะห์ความแปรปรวน 3 ปัจจัยได้แก่หมวดหมู่ Podcast, ความยาววิดีโอ และช่วงไตรมาสที่อัปโหลดวิดีโอ พบว่ามีเพียงความยาววิดีโอและช่วงไตรมาสที่อัปโหลดเท่านั้นที่ส่งผลต่อจำนวนคนรับชมวิดีโอโดยเฉลี่ย ซึ่งการจากการวิเคราะห์แล้วมีความเป็นไปได้ เนื่องจากความยาวของวิดีอนั้น โดยส่วนใหญ่หากนานเกินไปผู้คนมักจะไม่ค่อยสนใจกันมากนัก เว้นเสียแต่หัวข้อหรือเนื้อหาของวิดีอนั้นน่าสนใจจริง ๆ และจากการใช้ชีวิตแบบเร่งรีบของผู้คนก็มีส่วนเช่นกัน ในขณะที่ช่วงไตรมาสที่อัปโหลดนั้นมีเป็นอยู่บางส่วน เนื่องจากบางไตรมาสส่งผลต่อจำนวนคนดู และความน่าสนใจของสื่อเช่นกัน อาทิ ในช่วงไตรมาสที่สองของแต่ละปีเป็นช่วงปิดเทอมของนักเรียนและนักศึกษาในหลาย ๆ แห่ง ทำให้มีกลุ่มคนในช่วงวัยดังกล่าวมีเวลาในการรับฟัง Podcast มากกว่าไตรมาสอื่น ๆ ซึ่งได้สรุปการเปลี่ยนแปลงของจำนวนคนดูโดยเฉลี่ยไว้ในภาคผนวก⁹ เป็นต้น

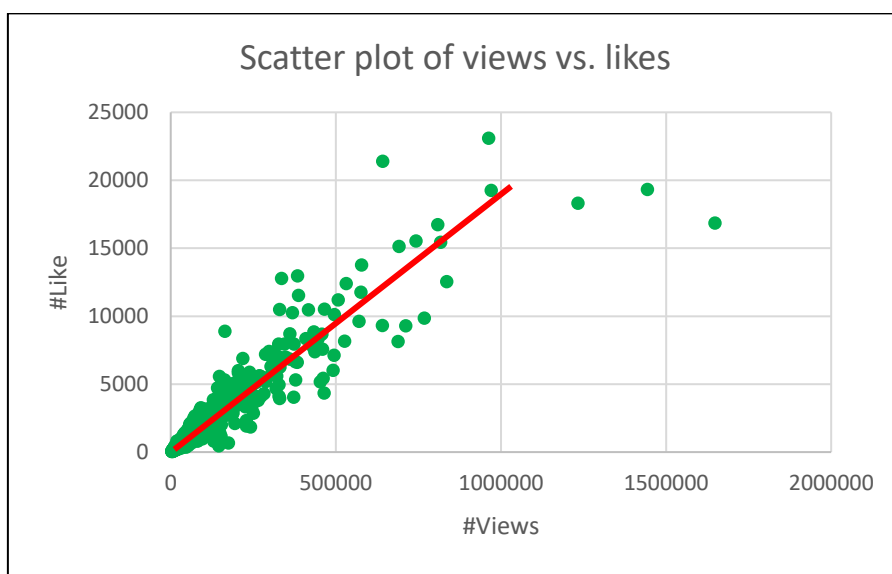
⁹ รูปภาพ 35 แผนภูมิเส้นแสดงการเปลี่ยนแปลง
ของค่าเฉลี่ยคนดูในแต่ละไตรมาส

2.4. การสร้างแบบจำลองสมการถดถอยเชิงเส้นอย่างง่าย

(Simple Linear Regression)

สมการเพื่อแสดงความสัมพันธ์ระหว่างจำนวนคนดู และคนกดถูกใจ

ขั้นตอนที่ 1 : เมื่อพิจารณาจากแผนภูมิการกระจายระหว่างจำนวนของคนดูและจำนวนคนกดถูกใจในหัวข้อที่ 4.1.8 พบว่ามีแนวโน้มของข้อมูลมีลักษณะเป็นเส้นตรงในช่วง ๆ หนึ่ง ดังรูป



รูปภาพ 25 ภาพแสดงแนวโน้มของข้อมูลระหว่าง
จำนวนคนดู และจำนวนคนกดถูกใจ

ขั้นตอนที่ 2 : กำหนดสมการเชิงเส้นเพื่อแสดงความสัมพันธ์ระหว่างจำนวนคนดู
และคนกดถูกใจของกลุ่มตัวอย่าง ดังสมการ

$$\hat{y} = a + bx \quad (4.2.6)$$

โดยกำหนดให้ \hat{y} แทน จำนวนคนกดถูกใจ (ต่อคลิป)

x แทน จำนวนคนรับชมวิดีโอ (ต่อคลิป)

a, b เป็นค่าคงที่

ขั้นตอนที่ 3 : คำนวณหาของ a, b จาก Least Square Method (LSM) ด้วย
โปรแกรม R ได้ผลดังตาราง

ตาราง 17 ตารางแสดง Regression ของสมการที่ 1

<i>Regression Statistics</i>				
Multiple R	0.931253442			
R Square	0.867232973			
Adjusted R Square	0.867104943			
Standard Error	954.7909764			
Observations	1039			

ANOVA Table				
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>
Regression	1	6175056261	6.18E+09	6773.674
Residual	1037	945355963.5	911625.8	
Total	1038	7120412224		

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>
Intercept	207.1990643	34.70861515	5.969672	3.26E-09
#view	0.017267719	0.000209808	82.30233	0

เพราะฉะนั้น จากตารางดังกล่าวจะได้ค่า $a = 207.199$ และ $b = 0.017268$

ขั้นตอนที่ 4 : ทดสอบสมมติฐานรายพจน์ (Individual Test) a และ b
พจน์ค่าคงที่ a โดยมีพารามิเตอร์เป็น α

วิธีทำ $H_0 : \alpha = 0$

$H_a : \alpha \neq 0$ (Two tailed Test)

กำหนดระดับนัยสำคัญ $\alpha^+ = 0.05$

จากตารางที่ 5 จะได้ว่า

$$t_{ob} = 5.9697, df = 1037$$

เพราะฉะนั้น

$$\begin{aligned} P - value &= 2 \times P(T > |t_{ob}|, df) \\ &= 2 \times P(T > 5.9697, \\ &\quad df = 1037) \\ &= 3.26 \times 10^{-9} \end{aligned}$$

เนื่องจาก $P - value < \alpha^+$

ดังนั้น เราสามารถที่จะปฏิเสธ H_0 ได้ที่ระดับนัยสำคัญ 0.05 หรือ

ระดับความเชื่อมั่นที่ 95% (Reject H_0)

สรุปได้ว่า

$\alpha \neq 0$

พจน์ค่าคงที่ b โดยมีพารามิเตอร์เป็น β

วิธีทำ $H_0 : \beta = 0$

$H_a : \beta \neq 0$ (Two tailed Test)

กำหนดระดับนัยสำคัญ $\alpha = 0.05$

จากตารางที่ 5 จะได้ว่า

$$t_{ob} = 82.3023, df = 1037$$

เพราะฉะนั้น

$$\begin{aligned} P - value &= 2 \times P(T > |t_{ob}|, df) \\ &= 2 \times P(T > 82.3023, \\ &\quad df = 1037) \\ &= 0 \end{aligned}$$

เนื่องจาก $P - value < \alpha$

ดังนั้น เราสามารถที่จะปฏิเสธ H_0 ได้ที่ระดับนัยสำคัญ 0.05 หรือ

ระดับความเชื่อมั่นที่ 95% (Reject H_0)

สรุปได้ว่า

$\beta \neq 0$

สรุปว่าทั้ง α และ β ต่างไม่เท่ากับศูนย์ ซึ่งสรุปได้ว่าแบบจำลองมีความเหมาะสม

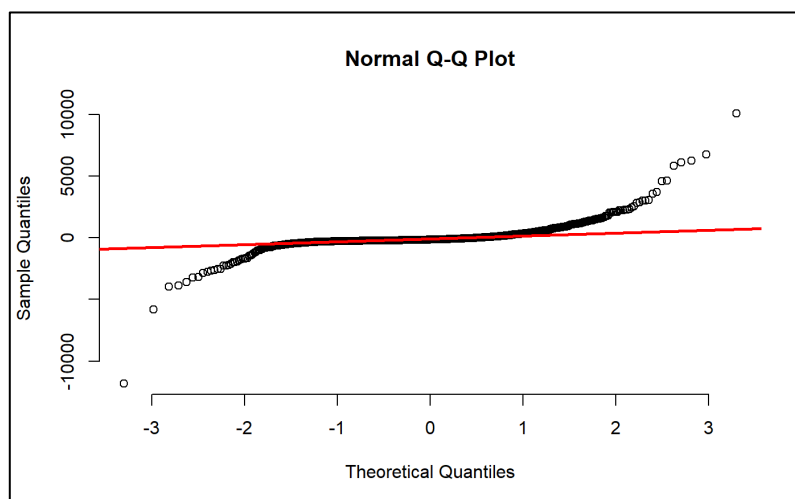
(Model is significant / useful.)

หมายเหตุ : สามารถใช้ตาราง Anova เพื่อทำการทดสอบแบบ F – Test สำหรับ

Overall Test ซึ่งจะได้ข้อสรุปว่า $\beta \neq 0$ เช่นเดียวกัน

ขั้นตอนที่ 5 : การทดสอบข้อกำหนด หรือข้อตกลงสำหรับสมการถดถอยเชิงเส้น
(Check Assumption for Simple Linear Regression)

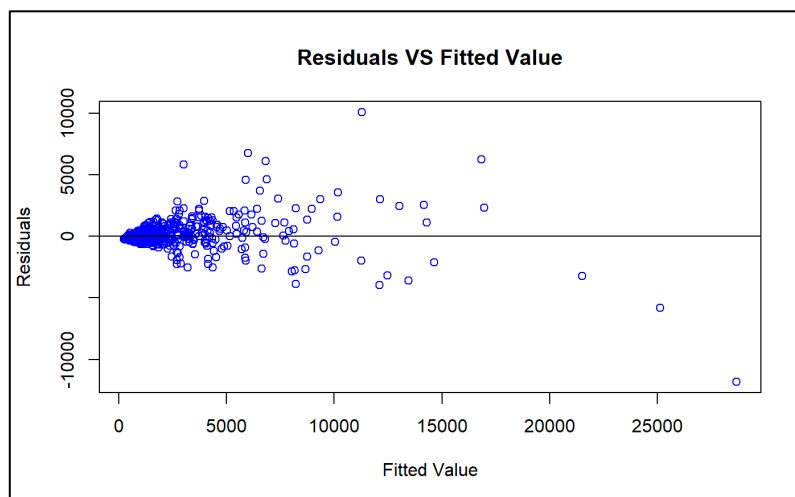
1. Normal Probability Plot of Residuals



รูปภาพ 26 Normal Probability Plot of Residuals

สำหรับสมการที่ 1

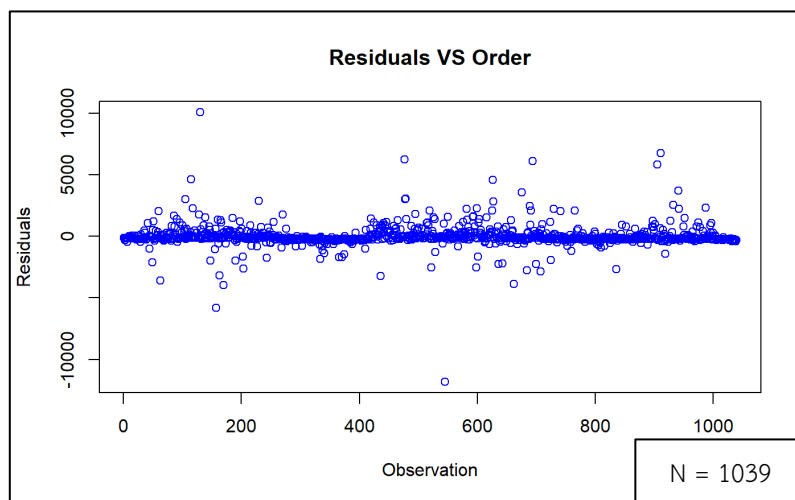
2. Plot of Residual vs Fitted value



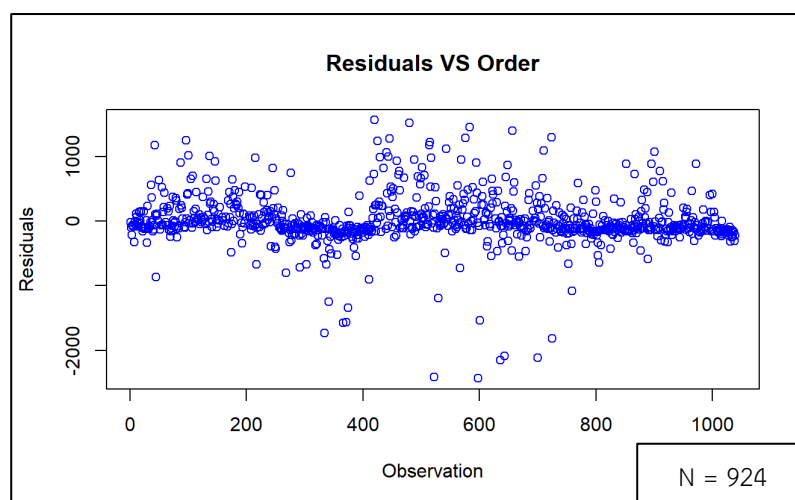
รูปภาพ 27 Residuals VS Fitted Value plot

สำหรับสมการที่ 1

3. Plot of residual vs Order



รูปภาพ 28 Residuals VS Order สำหรับสมการที่ 1 (a)



รูปภาพ 29 Residuals VS Order สำหรับสมการที่ 1 (b)

โดยรูปแบบ (b) คือข้อมูลที่ตัดส่วนของค่านอกเกณฑ์ (Outlier) จำนวน 115 ข้อมูล คิดเป็น 11.07% ออกเพื่อให้วิเคราะห์ข้อมูลได้โดยง่าย

ตาราง 18 ตารางสรุปข้อมูลที่ได้จากกราฟสำหรับ

การทดสอบสมมติฐานของสมการการถดถอย

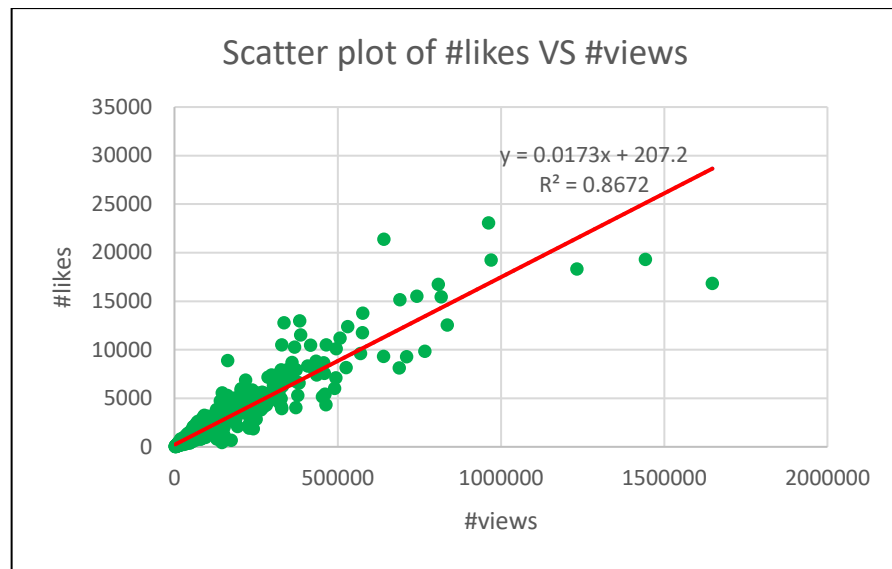
กราฟ	คำอธิบาย (ข้อสรุป)
Normal Probability Plot of Residuals	สำหรับข้อมูลที่เก็บมานี้ มีช่วงบริเวณหนึ่ง ๆ ของข้อมูลที่อยู่บริเวณเส้นตรงสีแดง สามารถอนุมานได้ว่าข้อมูลมีการแจกแจงแบบปกติ ซึ่งสามารถนำแบบจำลองนี้ไปใช้คาดคะเนได้ค่อนข้างแม่นยำ
Residual VS Fitted Value	ข้อมูลมีค่าเฉลี่ยของ Error (Residual) เข้าใกล้ ศูนย์ และมีค่าความแปรปรวนเท่ากัน
Residual VS Order	ข้อมูลทั้งหมดเป็นอิสระต่อกัน

ขั้นตอนที่ 6 : สรุปขั้นตอน

จากการสร้างแบบจำลองสมการถดถอยเชิงเส้นเพื่อคาดคะเนความสัมพันธ์ระหว่างจำนวนคนดูและจำนวนคนกดถูกใจในแต่ละคลิป พบว่ามีความสัมพันธ์เป็นเส้นตรง ซึ่งมีสมการสำหรับการคาดคะเนเป็น

$$\hat{y} = 207.199 + 0.017268 x \quad (4.2.7)$$

ซึ่งเมื่อทำการทดสอบสมมติฐานรายพจน์ด้วยการทดสอบ T (T-Test) แล้วพบว่าค่า α และ β ซึ่งเป็นค่าของพารามิเตอร์สำหรับ a และ b ต่างไม่เท่ากับศูนย์ จึงทำให้แบบจำลองนี้เหมาะสม โดยสอดคล้องกับข้อสมมติฐานสำหรับแบบจำลองทั้ง 3 สมมติฐานด้วยเช่นกัน นอกจากนี้ในตาราง 17 ตารางแสดง Regression ของสมการ ข้อมูลชุดนี้มีค่า $R\text{-squared} = 0.8672 = 86.72\%$ ซึ่งมีความแม่นยำสำหรับการคาดคะเนที่สูงมาก



รูปภาพ 30 ภาพแสดงระหว่างข้อมูลจริง

และข้อมูลที่ได้จากแบบจำลองที่ 1

ทั้งนี้เนื่องจากสมการดังกล่าวอาจไม่สามารถคาดคะเนข้อมูลแบบจุด (Point Estimation) ได้แม่นยำมากนักเนื่องมาจากการมีข้อผิดพลาด และค่านอกเกณฑ์ ดังนั้นจึงสร้างช่วงที่ครอบคลุมค่าที่เป็นไปได้ (Interval Estimation) โดยใช้ช่วงแห่งความเชื่อมั่น (Confident Interval) สำหรับค่า y ที่ได้จากสมการ (4.2.5) ดังสมการ

$$\hat{y}_0 \pm t_{\alpha/2} \cdot \sqrt{MSE \left(1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}} \right)} \quad (4.2.8)$$

โดย \hat{y}_0 แทน ค่าที่ได้จากสมการ (4.2.7) ด้วย $x = x_0$

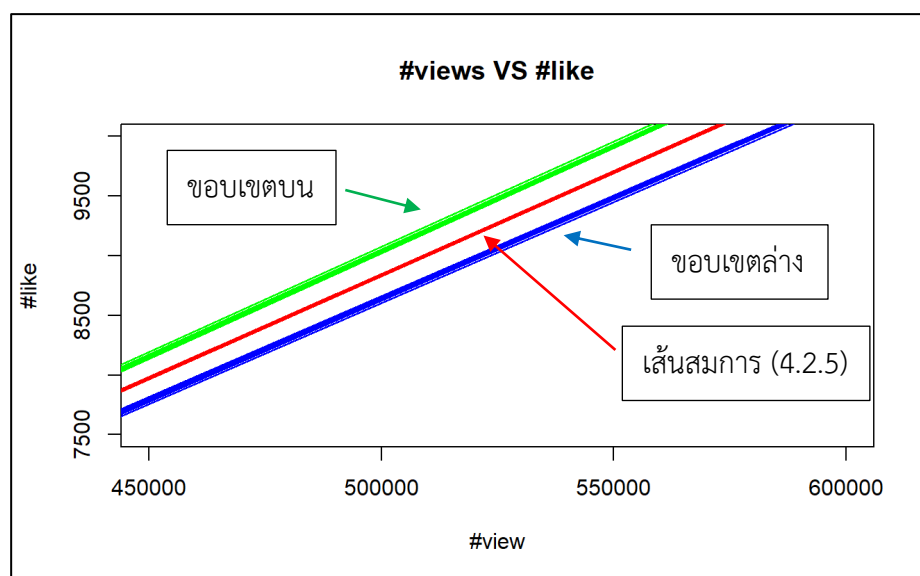
$$S_{xx} = \sum x^2 - \frac{(\sum x)^2}{n}$$

$$MSE = 911,625.8 \text{ ซึ่งได้จากตาราง Regression}$$

และกำหนดระดับนัยสำคัญ $\alpha = 0.05$

จึงได้ผลลัพธ์ตามภาพ โดยพิจารณาเพียงช่วง

$$x \in [450,000, 600,000] \text{ และ } y \in [7,500, 10,000]$$



รูปภาพ 31 ช่วงแห่งความเชื่อมั่นของแบบจำลองที่ 1

บทที่ 5

สรุป อภิปรายและข้อเสนอแนะ

การศึกษาค้นคว้าในครั้งนี้ เป็นการศึกษาปัจจัยต่าง ๆ ที่เกี่ยวข้องกับการเข้าดูคลิปพอดแคสต์ของคนไทยโดยอาศัยหลักการทางสถิติ ซึ่งมีขั้นตอนสรุปผล อภิปรายผลและข้อเสนอแนะดังนี้

1. วัตถุประสงค์ของการศึกษา
2. สรุปผลของการศึกษา
3. อภิปรายผล
4. ข้อเสนอแนะ

1. วัตถุประสงค์ของการศึกษา

1. เพื่อศึกษาพฤติกรรมและรูปแบบการเข้าชม Podcast ของคนไทยที่ส่งผลต่อจำนวนคนรับชมวิดีโอ Podcast บนแพลตฟอร์ม The Secret Sauce
2. เพื่อวิเคราะห์ปัจจัย 3 ปัจจัยที่ส่งผลต่อจำนวนคนดูในแต่ละวิดีโอได้แก่ หมวดหมู่วิดีโอ ความยาวของวิดีโอ และเวลาที่อัปโหลดใน YouTube บนแพลตฟอร์ม The Secret Sauce

2. สรุป และอภิปรายผลของการศึกษา

จากการศึกษาปัจจัยต่าง ๆ ที่เกี่ยวข้องกับการเข้าดูคลิปพอดแคสต์ของคนไทยโดยอาศัยหลักการทางสถิติ สามารถสรุป และอภิปรายผลได้ดังต่อไปนี้

1. จำนวนคนรับชมวิดีโอและจำนวนคนกดถูกใจมีความสัมพันธ์กัน

แซนเนล The Secret Sauce มีการกระจายตัวของจำนวนคนรับชมวิดีโอและจำนวนคนกดถูกใจที่ใกล้เคียงกันในแต่ละหมวดหมู่ นอกจากนี้ยังพบว่าโดยทั่วไปแล้วจำนวนคนรับชมวิดีโอและจำนวนคนกดไลค์มีลักษณะไปในทิศทางเดียวกัน ทั้งนี้เนื่องจากแผนภาพแสดงแนวโน้มของข้อมูลระหว่างจำนวนคนดูและจำนวนการกดถูกใจ ซึ่งอยู่ในส่วนของสมการถดถอยเชิงเส้นอย่างง่ายในหัวข้อ 4.2.2.3 ที่แสดงให้เห็นว่าทั้งสองนั้นมีแนวโน้มที่จะมีความสัมพันธ์เป็นเส้นตรงในทิศทางเดียวกัน โดยจากการสร้างสมการถดถอยเชิงเส้นอย่างง่ายแล้ว พบว่าความสัมพันธ์ดังกล่าวมีโอกาเป็นเส้นตรงที่สูง เนื่องจากมี R^2 เท่ากับ 86.72% ส่งผลให้การจัดทำวิดีโอเพื่อนำเสนอข้อมูลต่าง ๆ มีลักษณะที่จะเป็นได้ด้วยดี เพราะจำนวนของคนดูที่ยิ่งมากขึ้น แสดงให้เห็นว่าจำนวนการกดถูกใจก็จะมีโอกาสที่จะมากขึ้นด้วยซึ่งเป็นการทำให้ผู้จัดทำวิดีโอมีกำลังใจสำหรับการนำเสนอวิดีโอในอนาคตได้อีกด้วย แต่ในขณะเดียวกันกลับพบว่าที่สัดส่วนของจำนวนคนกดถูกใจต่อจำนวนคนรับชมวิดีโอในแต่ละคลิปอยู่ในช่วง 2% ถึง 3% เป็นส่วนใหญ่ซึ่งถือสัดส่วนที่ค่อนข้างต่ำมาก

2. Podcast ที่อัปโหลดในแต่ละปีมีจำนวนใกล้เคียงกันสำหรับทุก ๆ หมวดหมู่

การอัปโหลดวิดีโอ Podcast ในแต่ละปีของแซนเนล The Secret Sauce มีการกระจายหมวดหมู่เท่า ๆ กัน โดยจากการทดสอบด้วยไคสแควร์ระหว่างทั้งหมวดหมู่วิดีโอเหล่านั้นและความยาววิดีโอพบว่าทั้งสองต่างมีความสัมพันธ์กัน (ไม่ได้เป็นอิสระต่อกัน) ซึ่งการทดสอบดังกล่าวค่อนข้างสมเหตุสมผล เนื่องจากในวิดีโอที่เกี่ยวข้องกับประวัติศาสตร์ การเมือง หรือการศึกษาและจิตวิทยา ความยาวของวิดีโอจะมีความยาวค่อนข้างมาก เพราะเนื้อหาภายในหมวดหมู่ที่เกี่ยวข้องนี้มักมีเนื้อหาที่ละเอียดอ่อนและมีความซับซ้อนเป็นอย่างสูง ทำให้ไม่สามารถที่จะเล่าได้อย่างรวบรัดได้ เพราะทำให้ขาดสาระบางส่วนไปได้โดยง่าย จึงได้มีการแบ่งออกเป็นตอนย่อย ๆ ในหลายวิดีโอ ตัวอย่างเช่น “ดร. สุรเกียรติ์ เสถียรไทย ตอน 1 โลกาภิวัตน์แตกเป็นเสี่ยงเสื้อ EP.396 (18 พ.ย. 2565) และดร.สุรเกียรติ์ เสถียรไทย ตอน 2 โลกแบ่งขั้ว ไทยควรเลือกข้างไหน? EP.397 (19 พ.ย. 2565)” เป็นต้น

3. Podcast ที่มีระยะเวลาที่สั้น มักจะถูกเลือกรับฟังมากที่สุด

ความยาววิดีโอมีลักษณะการกระจายแบบเบ้ขวา โดยมีความยาววิดีโออยู่ในช่วง 30 ถึง 60 นาทีเป็นส่วนใหญ่ ถัดลงมาเป็นความยาว 0 ถึง 30 นาที โดยมีค่าเฉลี่ยของความยาววิดีโอ คิดเป็น 42.45 นาที เนื่องจากในช่วงวิดีโอ 0 ถึง 60 นาทีนั้นมักเป็นช่วงเวลาที่ผู้คนส่วนมากมักมีสมาธิที่ยังสามารถจดจ่อและรับฟังได้มากที่สุดสำหรับการเข้ารับชมวิดีโอในแต่ละคลิป รวมไปถึงเป็นช่วงเวลาที่ผู้คนมักฟัง Podcast ร่วมกับการทำกิจกรรมอื่น ๆ ร่วมกันอีกด้วย อย่างเช่น การทำอาหารร่วมกับการฟัง Podcast การฟัง Podcast ก่อนนอน เป็นต้น และนอกจากนี้ในช่วงเวลา 30 ถึง 60 นาทียังเป็นเวลาที่สามารถเล่าหรือนำเสนอข้อมูลที่มีความยาวหรือความสลับซับซ้อน ได้อยู่มากเช่นกันโดยที่ไม่เร่งรีบจนเกินไป

4. แนวโน้มของการรับชมวิดีโอมีแนวโน้มไปในทางเดียวกัน

จากการวิเคราะห์ความแปรปรวนที่ระดับนัยสำคัญ 0.05 พบว่าหมวดหมู่ของวิดีโอมีค่าเฉลี่ยใกล้เคียงกัน ซึ่งอธิบายได้ว่าหมวดหมู่ในแต่ละหมวดหมู่ไม่ได้ส่งผลต่อจำนวนคนดูโดยเฉลี่ย และพบว่าในแต่ละหมวดหมู่จะมีทิศทางหรือแนวโน้มของจำนวนคนดูโดยเฉลี่ยไปในทิศทางเดียวกัน ทั้งนี้ปัจจัยที่ส่งผลต่อจำนวนคนดูโดยเฉลี่ยของแขนงดังกล่าว ได้แก่ ความยาววิดีโอที่อัปโหลด และช่วงไตรมาสของแขนงที่ได้อัปโหลดวิดีโอเพียงเท่านั้น

5. คนไทยยุคปัจจุบันยังคงให้ความสำคัญกับการศึกษา

โดยภาพรวมของคนไทยในการเลือกตัดสินใจกับดูหรือฟัง Podcast พบว่าคนไทยยังคงสนใจกับเนื้อหาประเภท “การศึกษา” เป็นหลัก โดยมีประเภทอื่นที่รอง 2 อันดับแรก ได้แก่ “Life Style”, “Logistics” ทั้งนี้เนื่องจากหมวดหมู่ของการศึกษานั้นเป็นสิ่งที่คนไทยในยุคปัจจุบันให้ความสำคัญกับเรื่องนี้มากเป็นพิเศษเพื่อที่จะพัฒนาตนเองให้ทันและพร้อมเข้าสู่ยุคของเทคโนโลยีและดิจิทัล เพราะวิดีโอภายในประเภทนี้จะเน้นเนื้อหาที่เกี่ยวข้องกับการพัฒนาตนเอง จิตวิทยา และกรณีศึกษาจากบุคคลสำคัญต่าง ๆ มานำเสนอ ซึ่งตอบโจทย์ต่อผู้คนเป็นอย่างมาก ในขณะเดียวกัน Life Style และ Logistics เป็นเรื่องที่คนไทยให้ความสำคัญเป็นอันดับรองลงมา เนื่องจากเป็นเรื่องที่อยู่ใกล้ตัวของตัวเอง และสามารถนำเนื้อหาบางส่วนไปพัฒนากับตัวเองต่อไปได้แม้ว่าจะได้ไม่เท่ากับประเภทการศึกษาก็ตาม

6. จำนวนผู้จัดรายการยิ่งน้อย ยิ่งกระชับ ยิ่งน่าสนใจ

ในปัจจุบันมีรูปแบบรายการ Podcast ที่หลากหลายแตกต่างกันออกไป และหนึ่งในปัจจัยที่สร้างมุมมองการฟังที่แตกต่างนอกเหนือจากเนื้อหาและหมวดหมู่ที่น่าสนใจแล้ว ยังมีในส่วนของผู้จัดรายการที่แตกต่างกันออกไป บางรายการมีการเชิญพิธีกรหรือวิทยากรอื่น ๆ มาร่วมกันจัดการรายการ ทำให้บางรายการนำเสนอผ่านพิธีกรเหล่านั้นไปในเชิงการสัมภาษณ์ ซึ่งพบว่าโดยเฉลี่ยแล้วคนไทยนิยมในการรับชมวิดีโอโดยมีผู้จัดรายการเพียงคนเดียวเป็นส่วนใหญ่ เนื่องจากมองว่าสามารถเข้าใจเนื้อหาได้โดยง่ายและไม่วุ่นวายมากนัก

3. ข้อเสนอแนะ

1. เนื่องจากข้อมูลที่มีเป็นข้อมูลเพียงแขนงเดียว The Secret Sauce เพียงแขนงเดียว ซึ่งไม่ครอบคลุมทุก ๆ แขนงในประเทศไทยที่ทำรายการ Podcast ข้อมูลที่สรุปมาได้นั้นจึงใช้ได้แม่นยำเพียงแค่แขนงเดียวเท่านั้น ทั้งนี้เนื่องจากเป็นแขนงที่มีผู้ติดตามเป็นจำนวนมาก และยังมีจำนวนวิดีโอที่อัปโหลดจำนวนมาก จึงสามารถนำมาเป็นข้อมูลกลุ่มตัวอย่างแทนจำนวนแขนงอื่น ๆ ได้
2. เนื่องด้วยระยะเวลาที่มีจำกัดจากกิจกรรมต่าง ๆ ร่วมกันในขณะดำเนินโครงการ ทำให้ไม่สามารถทำงานได้เต็มประสิทธิภาพ

บรรณานุกรม

- ชุดิธารัฐ อุตมะสิริเสนี และสุธาสินี พวงพลับ. (10 กรกฎาคม 2566). *คู่มือการเขียนรายงานเชิงวิชาการ*. เข้าถึงได้จาก
https://reg.pim.ac.th/registrar/download/pdf/Work_writing_guide.pdf
- ดวิษ ปรภายนต์. (15). Customer Insights and Media Exposure of Podcast Listeners to Present and Promote Podcast Channels Effectively on Facebook in Thailand . *วารสารวิทยาลัยดุสิตธานี*, 251. เข้าถึงได้จาก <https://dtt.ac.th/wp-content/>
- นครินทร์ วนกิจไพบูลย์. (6 สิงหาคม 2560). *The Standard*. เข้าถึงได้จาก แนะนำ ‘The Secret Sauce’: <https://thestandard.co/podcast/thesecondsauce00/>
- นันทชัย กานตานันทะ และคณะ. (7 สิงหาคม 2566). *เอกสารประกอบการสอน Engineering Statistic I*. เข้าถึงได้จาก MyCourseville:
<https://www.mycourseville.com/?q=courseville/course/35196>
- วิราพร หงษ์เวียงจันทร์. (ม.ป.ป.). *การอ้างอิงแบบแทรกในเนื้อหาและการเขียนรายการอ้างอิง*. เข้าถึงได้จาก <https://arts.tu.ac.th/>:
<https://arts.tu.ac.th/uploads/arts/TU104/171061.pdf>
- วิโรจน์ อรุณมานะกุล. (11 เมษายน 2560). *สถิติและการใช้โปรแกรม R*. เข้าถึงได้จาก Pioneer Chula: <http://pioneer.chula.ac.th/~awirote/courses/res-tech-ling/statistics-and-r.pdf>
- อังศุมาลิน เสนจันทร์พิไชย. (9 เมษายน 2567). *ข้อควรระวังในการเขียนปริญญานิพนธ์และวิทยานิพนธ์*. กรุงเทพฯ, ประเทศไทย.
- อังศุมาลิน เสนจันทร์พิไชย และนระเกณท์ พุ่มชูศรี. (8 มกราคม 2567). *เอกสารประกอบการสอน Engineering Statistics II*. เข้าถึงได้จาก MyCourseville:
<https://www.mycourseville.com/?q=courseville/course/45922>

ภาคผนวก

1. รูปภาพเพิ่มเติม

1.1. Output ที่ได้จากสมการที่ 1 โดยใช้คำสั่งในโปรแกรม R – Studio

```
Call:
lm(formula = X.like ~ X.view, data = df)

Residuals:
    Min       1Q   Median       3Q      Max
-11808.3   -231.2   -142.4    88.9   10100.0

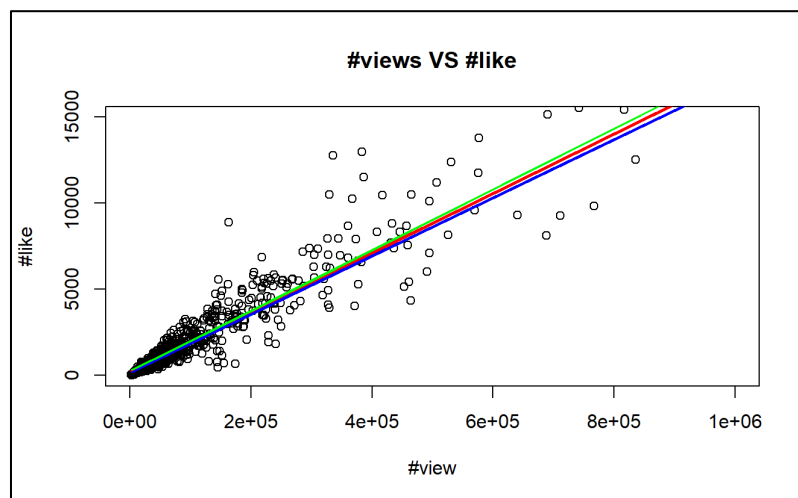
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.072e+02  3.471e+01   5.97 3.26e-09 ***
X.view       1.727e-02  2.098e-04  82.30 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 954.8 on 1037 degrees of freedom
Multiple R-squared:  0.8672,    Adjusted R-squared:  0.8671
F-statistic: 6774 on 1 and 1037 DF, p-value: < 2.2e-16
```

รูปภาพ 32 Output ที่ได้ของสมการที่ 1 (ตาราง Regression)

จากโปรแกรม R-Studio

1.2. Output ช่วงแห่งความเชื่อมั่นโดยใช้คำสั่งในโปรแกรม R – Studio



รูปภาพ 33 Output ช่วงแห่งความเชื่อมั่น

โดยใช้คำสั่งในโปรแกรม R - Studio

จะเห็นได้ว่าข้อมูลโดยรวม (ดังภาพ) ช่วงแห่งความเชื่อมั่นสำหรับค่า X ใด ๆ ค่อนข้างแคบ เพราะฉะนั้นเพื่อการตรวจสอบโดยสะดวก จึงสนใจในบริเวณบางช่วง X และ Y ดังที่อธิบายรูปภาพ 31 ช่วงแห่งความเชื่อมั่นของแบบจำลองที่ 1

1.3. หน้าปกสไลด์สำหรับการนำเสนอวันที่ 18 เมษายน 2567



รูปภาพ 34 หน้าปกสไลด์สำหรับการนำเสนอวันที่ 18 เมษายน 2567

2. ตารางเพิ่มเติม

2.1. ตารางแจกแจงหน้าที่ในการรับผิดชอบภาระงานแต่ละส่วน

ตาราง 19 ตารางแจกแจงภาระหน้าที่ของคณะผู้จัดทำ

ลำดับที่	ภาระงาน	บุคคลผู้รับผิดชอบ
1	วิเคราะห์สถิติเชิงพรรณนา	วันพิชิต, ศิวพร
2	วิเคราะห์สถิติเชิงอนุมาน	ธนานพ
3	ออกแบบแผนภูมิ กราฟ และแผนภาพ	คณะผู้จัดทำ
4	ปรึกษา และสอบถามอาจารย์	ธนานพ, วันพิชิต
5	ทำสไลด์ประกอบการนำเสนอ	วันพิชิต, ศิวพร
6	ทำเอกสารรายงาน	ธนานพ
7	ตรวจสอบเนื้อหาและพิสูจน์อักษร	ศิวพร, วันพิชิต

2.2. ค่าสถิติพื้นฐานของจำนวนคนดูและจำนวนคนกดถูกใจ

ตาราง 20 ตารางแสดงค่าสถิติพื้นฐานของจำนวนคนดู และจำนวนคนกดถูกใจ

<i>view</i>		<i>like</i>	
Mean	86225.87	Mean	1696.12
Standard Error	4382.07	Standard Error	81.25
Median	39039	Median	769
Mode	12670	Mode	708
Standard Deviation	141249.47	Standard Deviation	2619.11
Sample Variance	19951411579	Sample Variance	6859742.03
Skewness	4.75	Skewness	3.73
Range	1645318	Range	23051
Minimum	2273	Minimum	28
Maximum	1647591	Maximum	23079
1st Quartile	19374.5	1st Quartile	352.5
3rd Quartile	87624	3rd Quartile	1758
Interquartile Range	68249.50	Interquartile Range	1405.5
Sum	89588682	Sum	1762272
Count	1039	Count	1039
Confidence Level(95.0%)	8598.72	Confidence Level(95.0%)	159.44

2.3. ค่าสถิติพื้นฐานของความยาววิดีโอ

ตาราง 21 ตารางแสดงค่าสถิติพื้นฐานของความยาววิดีโอ

<i>duration_min</i>	
Mean	42.74
Standard Error	0.51
Median	40.9
Mode	40.6
Standard Deviation	16.58
Sample Variance	274.96
Skewness	0.58
Range	101.60
Minimum	5.8
Maximum	107.40
1st Quartile	29.85
3rd Quartile	53.40
Interquartile Range	23.55
Sum	44408.20
Count	1039
Confidence Level(95.0%)	1.01

2.4. ตารางแจกแจงจำนวนคลิปในแต่ละช่วงเวลา สำหรับแต่ละไตรมาส

ตาราง 22 ตารางแจกแจงความยาววิดีโอในแต่ละช่วงเวลา สำหรับแต่ละไตรมาส

<i>yearQtr</i>	0 - 0.5 hours	0.5 - 1 hours	1 - 1.5 hours	1.5 - 2 hours	Total
2018-Q1	2	22	3	0	27
2018-Q2	0	5	1	0	6
2018-Q3	7	7	0	0	14
2018-Q4	12	14	0	0	26
2019-Q1	9	16	0	0	25
2019-Q2	12	16	1	0	29
2019-Q3	17	12	3	0	32
2019-Q4	9	16	11	0	36
2020-Q1	18	37	6	0	61
2020-Q2	12	71	18	0	101
2020-Q3	17	47	7	0	71
2020-Q4	19	32	15	1	67
2021-Q1	13	42	13	0	68
2021-Q2	19	48	5	1	73
2021-Q3	21	40	13	2	76
2021-Q4	10	50	13	0	73
2022-Q1	13	33	12	1	59
2022-Q2	12	23	8	0	43
2022-Q3	12	34	7	1	54
2022-Q4	12	36	7	1	56
2023-Q1	15	22	5	0	42
Total	261	623	148	7	1039

2.5. ตารางแจกแจงจำนวนคลิปที่พบ Keyword 9 อันดับแรก

ตาราง 23 ตารางแจกแจงจำนวนคลิปที่พบ Keyword 9 อันดับแรก

Keyword	Count	Avg of views
สำเร็จ	22	108,422
โลก	143	101,598
วิธี	48	93,274
ล้าน	89	90,960
อย่างไร	139	65,424
ธุรกิจ	56	65,021
สร้าง	63	63,352
เคล็ดลับ	21	58,601
แบรนด์	28	35,558

2.6. ตารางสถิติพื้นฐานแจกแจงจำนวนคนรับชมในแต่ละปี

ตาราง 24 ตารางแจกแจงจำนวนคนรับชมในแต่ละปี

	2018	2019	2020	2021	2022	2023
Mean	73745.247	79848.943	56964.137	91318.631	122173.33	118840.88
Max	969692	808380	526044	1647591	1442899	742414
Min	3213	5109	2273	3803	7125	9748
n	74	123	301	291	213	43
Quartile						
1	13147	17450.25	13707	19713.25	32510	30669.25
2	24893	36789.5	30183.5	38583.5	60295	48926
3	73605	103865	61416.5	76799.5	155063.25	151225.75
IQR	60458	86414.75	47709.5	57086.25	122553.25	120556.5

2.7. ตารางแสดงจำนวนคนดูทั้งหมดในแต่ละประเภท

ตาราง 25 ตารางแสดงค่าเฉลี่ยของจำนวนคนดูในแต่ละประเภทวิดีโอ

Industry	2018	2019	2020	2021	2022	2023	total
การศึกษา	408474	174548	45608	103784	98938	153781	985133
Life Style	190411	111563	117556	72729	130909	164759	787927
Logistics	0	29440	33108	525555	59363	0	647466
Unspecified	121894	99040	75540	65435	86904	160614	609427
SMEs	0	104471	38589	177893	68019	171677	560649
ผู้นำภาครัฐ	0	12475	74285	168037	245826	19606	520229
พลังงาน	0	21320	38270	25030	196929	184446	465995
Design	32659	19482	78541	184516	141850	0	457048
ธุรกิจค้าปลีก	55561	12694	122710	67204	66773	116551	441491
อสังหาริมทรัพย์	17779	55186	55367	125404	99833	76513	430081
ธนาคาร	44440	44093	42410	54653	134593	106431	426621
Innovation	31207	62991	52240	76696	89027	65999	378159
E-Commerce	0	9004	98037	31516	171625	0	310181
Entertainment	15173	24064	43620	129902	78373	0	291132
รถยนต์	0	20824	6379	23601	225097	0	275901
สินค้าอุปโภค และบริโภค	24229	28122	71264	74050	24497	0	222161
การท่องเที่ยว	0	18578	95983	20162	30260	50268	215252
Platform	21329	37649	35197	61166	22311	20981	198633
ประกัน	0	0	9123	152622	32664	0	194409
Food Delivery	28642	0	50723	74433	16353	0	170151
โทรคมนาคม	7438	0	51035	13170	80082	0	151725
อาหาร	33784	26936	26534	47242	0	0	134495
Health & Wellness	21034	0	25217	45256	33288	0	124795
สินเชื่อ	0	0	0	123066	0	0	123066

ตาราง 26 ตารางแสดงค่าเฉลี่ยของจำนวนคนดูในแต่ละประเภทวิดีโอ (ต่อ)

Industry	2018	2019	2020	2021	2022	2023	total
Fashion	32990	38695	0	11602	0	0	83287
เอเจนซีโฆษณา	9550	0	7394	12943	0	42666	72552
ธุรกิจเครื่องดื่ม	0	0	39619	0	0	0	39619

2.8. ตารางแจกแจงจำนวนวิดีโอที่อัปโหลดในแต่ละปี

ตาราง 27 ตารางแจกแจงจำนวนวิดีโอที่อัปโหลดในแต่ละปี

Year	Quarter				Total
	1	2	3	4	
2018	27	6	14	26	73
2019	25	29	32	36	122
2020	61	101	71	67	300
2021	68	73	76	73	290
2022	59	43	54	56	212
2023	42	0	0	0	42
Total	282	252	247	258	1039

2.9. ตารางแสดงการหาค่าโคสแควร์

ตาราง 28 ตารางแสดงการหาค่าโคสแควร์

Duration (i)	Category (j)						Total _i
	1	2	3	4	5	6	
1	0.2091	0.3873	0.1909	3.3268	0.1598	2.4057	6.6796
2	1.1595	6.5735	0.7562	5.7555	1.4535	3.6400	19.3383
3	0.1526	1.6644	0.8846	0.3499	1.2279	0.2847	4.5641
4	1.5768	2.0004	0.8239	0.0074	4.5564	0.0868	9.0517
Total _j	3.0980	10.6256	2.6556	9.4395	7.3977	6.4171	39.63357

3. เนื้อหา หรือข้อมูลเพิ่มเติม

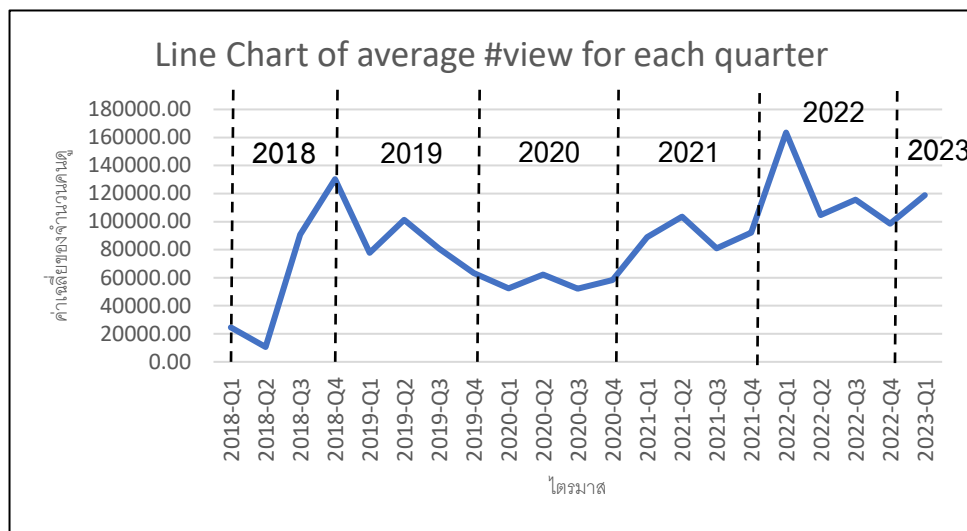
3.1. ตัวแบบพยากรณ์สำหรับค่าเฉลี่ยของจำนวนคนดูคลิปแต่ละไตรมาส

จุดประสงค์: เพื่อพยากรณ์ค่าเฉลี่ยของคนรับชมวิดีโอในไตรมาสที่ 2 ปี 2023

ขั้นตอนที่ 1: สรุปข้อมูลระหว่างจำนวนคนดูในแต่ละไตรมาส ได้ดังตาราง

ตาราง 19 ตารางแสดงค่าเฉลี่ยของจำนวนคนรับชมวิดีโอในแต่ละไตรมาส

No	yearQtr	Year	Quarter	Avg_View
1	2018-Q1	2018	1	24665.26
2	2018-Q2	2018	2	10598.83
3	2018-Q3	2018	3	90704.43
4	2018-Q4	2018	4	130153.31
5	2019-Q1	2019	1	77718.36
6	2019-Q2	2019	2	101252.66
7	2019-Q3	2019	3	80752.03
8	2019-Q4	2019	4	63283.89
9	2020-Q1	2020	1	52338.48
10	2020-Q2	2020	2	62234.53
11	2020-Q3	2020	3	52178.25
12	2020-Q4	2020	4	58302.24
13	2021-Q1	2021	1	88881.03
14	2021-Q2	2021	2	103524.15
15	2021-Q3	2021	3	81029.49
16	2021-Q4	2021	4	92095.74
17	2022-Q1	2022	1	163504.73
18	2022-Q2	2022	2	104626.63
19	2022-Q3	2022	3	115630.20
20	2022-Q4	2022	4	98410.57
21	2023-Q1	2023	1	118840.88



รูปภาพ 35 แผนภูมิเส้นแสดงการเปลี่ยนแปลง
ของค่าเฉลี่ยคนดูในแต่ละไตรมาส

ขั้นตอนที่ 2: สร้างสมการสำหรับการพยากรณ์

จากกราฟดังกล่าวจะเห็นได้ว่าข้อมูลมีลักษณะไม่เป็นรูปแบบ (No Pattern) แต่พบว่าสำหรับไตรมาสที่ 2 ของปี 2019 จนถึง 2022 ค่าเฉลี่ยของคนดูในช่วงเวลาดังกล่าวจะมีค่าสูงที่สุดสำหรับในแต่ละปี ในขณะที่ไตรมาสที่ 3 ในปีเดียวกันจะมีค่าลดลงเมื่อเทียบกับไตรมาสที่ 2 เสมอ

กรณีที่ 2.1: ใช้สมการ Holt's Exponential Smoothing

(Adjusted Exponential Smoothing) สำหรับการพยากรณ์

วิธีทำ โดยพิจารณาจากสมการ

$$F_{t+1} = \alpha A_t + (1 - \alpha)F_t$$

$$T_{t+1} = \beta(F_{t+1} - F_t) + (1 - \beta)T_t$$

$$AF_{t+1} = F_{t+1} + T_{t+1}$$

โดยที่ A_t แทน ค่าจริง (Actual) ที่เกิดขึ้นในเวลา t

F_t แทน ค่าพยากรณ์ (Forecast) ที่เกิดขึ้นในเวลา t โดยที่ $F_1 = 1$

T_t แทน ค่าแนวโน้ม (Trend) ที่เกิดขึ้นในเวลา t โดยที่ $T_1 = 0$

α แทน ค่าคงที่ที่ทำให้เรียบระหว่างข้อมูลกับค่าพยากรณ์ โดย $0 \leq \alpha \leq 1$

β แทน ค่าคงที่ที่ทำให้เรียบระหว่างแนวโน้มจริงกับค่าประมาณของแนวโน้ม โดย $0 \leq \beta \leq 1$

AF_t แทน ค่าพยากรณ์ที่มีการปรับค่าแนวโน้มที่เกิดขึ้นเวลา t

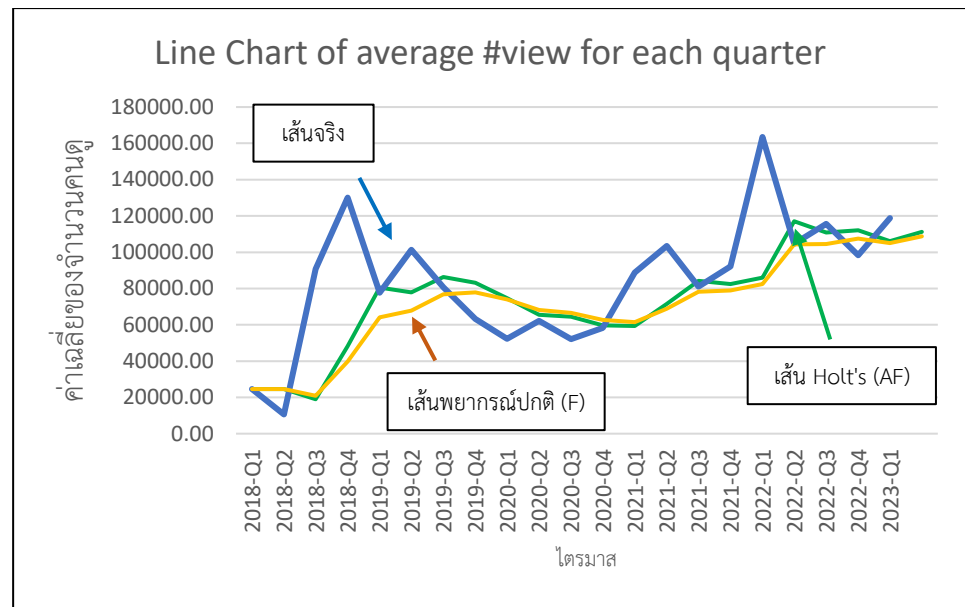
และ $t = 1, 2, 3, \dots, n$ เมื่อ n แทนจำนวนข้อมูล

ในที่นี้กำหนดให้ $\alpha = 0.27$ และ $\beta = 0.50$ สำหรับการพยากรณ์ โดยที่จะได้

ผลลัพธ์การคำนวณดังตาราง

ตาราง 29 ตารางแสดงการคำนวณการพยากรณ์แบบ Holt's

No	yearQtr	Avg_View	F_t	T_t	Adjust
1	2018-Q1	24665.26	24665.26	0.00	24665.26
2	2018-Q2	10598.83	24665.26	0.00	24665.26
3	2018-Q3	90704.43	20867.32	-1898.97	18968.36
4	2018-Q4	130153.31	39723.34	8478.53	48201.87
5	2019-Q1	77718.36	64139.43	16447.31	80586.74
6	2019-Q2	101252.66	67805.74	10056.81	77862.55
7	2019-Q3	80752.03	76836.41	9543.74	86380.15
8	2019-Q4	63283.89	77893.63	5300.48	83194.11
9	2020-Q1	52338.48	73949.00	677.92	74626.92
10	2020-Q2	62234.53	68114.16	-2578.46	65535.70
11	2020-Q3	52178.25	66526.66	-2082.98	64443.68
12	2020-Q4	58302.24	62652.59	-2978.52	59674.07
13	2021-Q1	88881.03	61477.99	-2076.56	59401.44
14	2021-Q2	103524.15	68876.81	2661.13	71537.94
15	2021-Q3	81029.49	78231.59	6007.96	84239.55
16	2021-Q4	92095.74	78987.03	3381.69	82368.72
17	2022-Q1	163504.73	82526.38	3460.52	85986.90
18	2022-Q2	104626.63	104390.53	12662.34	117052.87
19	2022-Q3	115630.20	104454.28	6363.04	110817.32
20	2022-Q4	98410.57	107471.78	4690.27	112162.05
21	2023-Q1	118840.88	105025.25	1121.87	106147.13
	2023-Q2		108755.47	2426.05	111181.52



รูปภาพ 36 แผนภูมิเส้นแสดงค่าการพยากรณ์แบบ

Holt's Exponential Smoothing

เพราะฉะนั้นด้วยวิธีการพยากรณ์แบบ Holt's Exponential Smoothing ด้วย $\alpha = 0.27$ และ $\beta = 0.50$ นั้น จะได้ผลลัพธ์จากการพยากรณ์ด้วยวิธีดังกล่าว สำหรับในปี 2566 ไตรมาสที่ 2 ว่ามีค่าเฉลี่ยของคนรับชมวิดีโอทั้งสิ้นเป็น 111,181.52 คนดูภายในช่วงเวลานี้

กรณีที่ 2.2: ใช้สมการ Weight Moving Average สำหรับการพยากรณ์

วิธีทำ โดยพิจารณาจากสมการ

$$F_{t+1} = \sum_{i=0}^{n-1} w_{t-i} A_{t-i}$$

โดยที่ A_t แทน ค่าจริง (Actual) ที่เกิดขึ้นในเวลา t

F_t แทน ค่าพยากรณ์ (Forecast) ที่เกิดขึ้นในเวลา t โดย $t = n+1, n+2, \dots$

w_t แทน ค่าน้ำหนัก (Weight) สำหรับข้อมูลจริง A_t โดยที่

$$w_1 + w_2 + w_3 + \dots + w_{n-1} + w_n = 1$$

n แทน จำนวนข้อมูลที่จะนำมาเฉลี่ย

หมายเหตุ: สำหรับการพยากรณ์นี้จะใช้ $n = 3$

โดยที่ $w_1 = 0.32, w_2 = 0.37$ และ $w_3 = 0.31$

ตัวอย่างเช่น

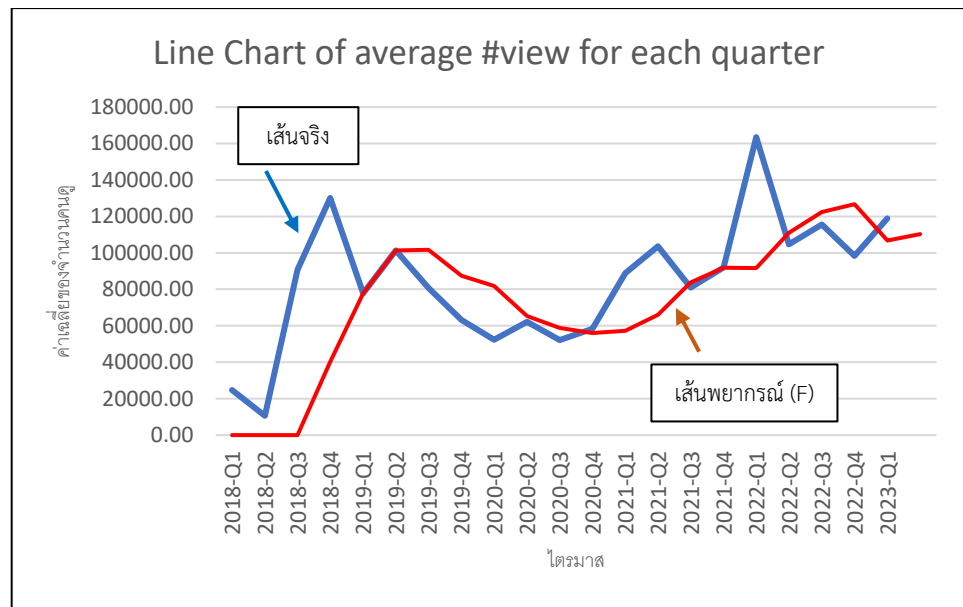
แทน $t = 3$:

$$\begin{aligned} F_4 &= \sum_{i=0}^{3-1} w_{3-i} A_{3-i} \\ &= w_3 A_3 + w_2 A_2 + w_1 A_1 \\ &= (0.31)(90704.43) + \\ &\quad (0.37)(10598.83) + \\ &\quad (0.32)(24665.26) \\ &= 40135.95 \end{aligned}$$

ซึ่งจะได้ข้อมูลดังตาราง

ตาราง 30 ตารางแสดงการคำนวณการพยากรณ์แบบ Moving Average

No	yearQtr	Avg_View	Forecast /w WMA
1	2018-Q1	24665.26	0
2	2018-Q2	10598.83	0
3	2018-Q3	90704.43	0
4	2018-Q4	130153.31	40135.95
5	2019-Q1	77718.36	77718.64
6	2019-Q2	101252.66	101252.65
7	2019-Q3	80752.03	101675.91
8	2019-Q4	63283.89	87387.53
9	2020-Q1	52338.48	81766.48
10	2020-Q2	62234.53	65381.83
11	2020-Q3	52178.25	58900.15
12	2020-Q4	58302.24	55954.63
13	2021-Q1	88881.03	57277.33
14	2021-Q2	103524.15	65941.63
15	2021-Q3	81029.49	83793.75
16	2021-Q4	92095.74	91847.14
17	2022-Q1	163504.73	91611.10
18	2022-Q2	104626.63	110958.87
19	2022-Q3	115630.20	122476.06
20	2022-Q4	98410.57	126698.03
21	2023-Q1	118840.88	106756.54
22	2023-Q2		110256.12



รูปภาพ 37 แผนภูมิเส้นแสดงค่าการพยากรณ์

แบบ Weight Moving Average

เพราะฉะนั้นด้วยวิธีการพยากรณ์แบบ Weight Moving Average (WMA) ด้วยการกำหนดค่า $n = 3$ โดยกำหนด $w_1 = 0.32$, $w_2 = 0.37$ และ $w_3 = 0.31$ นั้น จะได้ผลลัพธ์จากการพยากรณ์ด้วยวิธีดังกล่าวสำหรับในปี 2566 ไตรมาสที่ 2 ซึ่งเป็นช่วงเวลาที่สนใจ พบว่ามีค่าเฉลี่ยของคนรับชมวิดีโอทั้งสิ้นเป็น 110,256.12 คนดู ภายในช่วงเวลานี้

กรณีที่ 2.3: การใช้ Linear Trend Line สำหรับการพยากรณ์

วิธีทำ พิจารณาผลรวมค่าเฉลี่ยจำนวนคนดูในแต่ละปีตั้งแต่ 2561 ถึง 2565

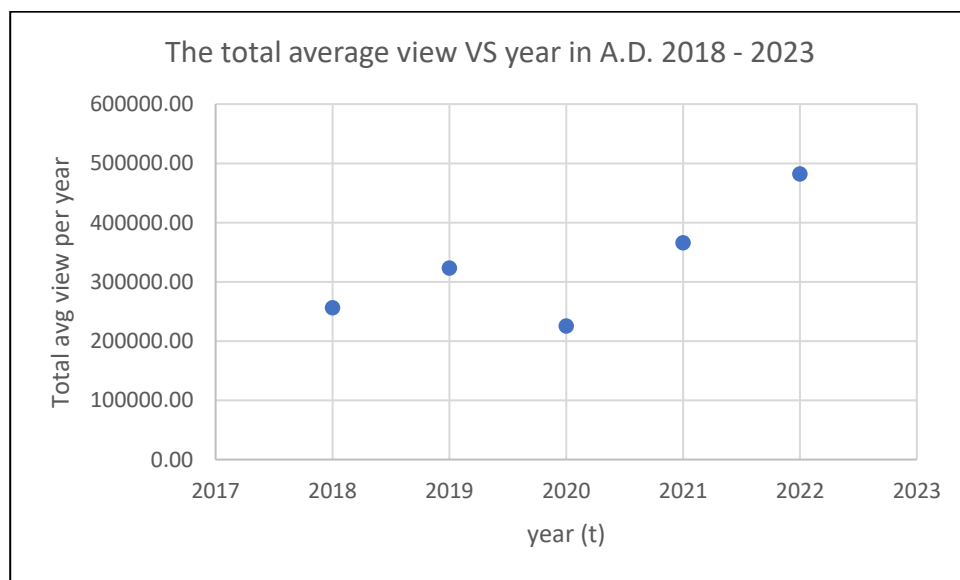
ดังตาราง

ตาราง 31 ตารางสรุปจำนวนคนดูโดยเฉลี่ย

ในแต่ละปีตั้งแต่ พ.ศ. 2561 ถึง 2565

Year (AD), t	Quarter (i)				Total, A(t)
	1	2	3	4	
2018	24665.26	10598.83	90704.43	130153.3	256,121.83
2019	77718.36	101252.7	80752.03	63283.89	323,006.94
2020	52338.48	62234.53	52178.25	58302.24	225,053.50
2021	88881.03	103524.2	81029.49	92095.74	365,530.41
2022	163504.7	104626.6	115630.2	98410.57	482,172.13
Average (D_i)	81,421.57	76,447.36	84,058.88	88,449.15	330,376.96

ซึ่งจะได้แผนภูมิแท่งระหว่างเวลา (ปี) กับจำนวนคนดูโดยเฉลี่ยรวมดังภาพ



รูปภาพ 38 แผนภูมิกระจายระหว่างปีและค่าเฉลี่ยจำนวนคนดู

ในช่วง พ.ศ. 2561 ถึง 2565

พิจารณาจากสมการ

$$y(t) = a + bt$$

โดย y แทน จำนวนค่าเฉลี่ยรวมของจำนวนคนดูที่พยากรณ์ได้ในปีที่ ค.ศ. ที่ t

โดย $t = 2018, 2019, \dots$

$$b = \frac{\sum t.A - \frac{(\sum t)(\sum A)}{n}}{\sum t^2 - \frac{(\sum t)^2}{n}} \quad \text{และ} \quad a = \bar{A} - b.\bar{t}$$

ซึ่งสามารถคำนวณแต่ละพจน์ได้ดังตารางต่อไปนี้

Data	t	A(t)	t.A	t ²
1	2018	256121.83	516853850.6	4072324
2	2019	323006.94	652151002.4	4076361
3	2020	225053.50	454608074.8	4080400
4	2021	365530.41	738736951.9	4084441
5	2022	482172.13	974952050.6	4088484
Total	10100	1651885	3337301930	20402010

จะได้ว่า

$$b = \frac{3337301930 - \frac{(10100)(1651885)}{5}}{20402010 - \frac{(10100)^2}{5}} = 49462$$

$$a = \frac{1651885}{5} - (49462) \frac{10100}{5} = -9.95837 \times 10^7$$

จากสมการ Linear Trend Line จะได้สมการเป็น

$$y(t) = 49462t - 9.95837 \times 10^7$$

และพิจารณาสมการ

$$\text{Seasonal Index } i, \quad S_i = \frac{D_i}{\bar{D}}$$

ได้ผลลัพธ์ดังนี้

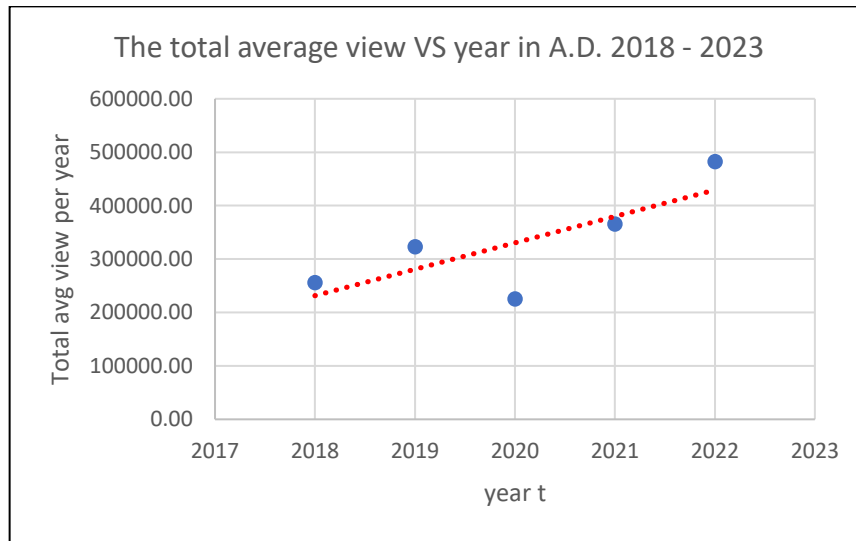
Quarter (i)	1	2	3	4
S_i	0.2465	0.2314	0.2544	0.2677

เพราะฉะนั้น

สำหรับ Quarter 2 ในปี 2023 จะได้จำนวนคนดูโดยเฉลี่ย

$$\begin{aligned} \text{มีค่าเป็น } y(t = 2023) \cdot S_2 &= [49462(2023) - 9.95837 \times 10^7](0.2314) \\ &= 110,592.08 \end{aligned}$$

หมายเหตุ



รูปภาพ 39 แผนภูมิเส้นแสดงค่าพยากรณ์

แบบ Linear Trend Line

หากพิจารณา Quarter 1 ในปี 2023 จากสมการเส้นตรงด้วยวิธีเดียวกัน
จะได้ค่าเฉลี่ยที่ไตรมาสดังกล่าวเป็น $y(t = 2023) \cdot S_1 = 117,808.76$ ซึ่งจากข้อมูล
ที่เก็บมา พบว่าค่าเฉลี่ยจำนวนคนรับชม ณ ที่แท้จริงเวลาดังกล่าวคือ 118,840.88
โดยถือว่าการพยากรณ์แบบดังกล่าวมีความผิดพลาดคิดเป็น 0.87%

ขั้นตอนที่ 3: พิจารณา MAPE (Mean Absolute percent Error)

จากสมการ

$$MAPE = \frac{1}{n} \sum_{t=1}^n \frac{|A_t - F_t|}{A_t}$$

โดย A_t แทน ค่าจริง (Actual) ที่เกิดขึ้นในเวลา t

F_t แทน ค่าพยากรณ์ (Forecast) ที่เกิดขึ้นในเวลา t

n แทน จำนวนข้อมูลที่ได้จากการพยากรณ์ในขณะที่เดิมมีค่าจริงอยู่แล้ว

ซึ่งจากการคำนวณ (แสดงในหน้าถัดไป) จะได้ดังนี้

ตาราง 32 ตารางสรุปค่า MAPE จากการพยากรณ์ทั้ง 3 รูปแบบ

	Holt's Exponential	Weight Moving Average	Linear Trend
MAPE	0.249	0.212	0.479

ขั้นตอนที่ 4: สรุปผล

จากการสร้างสมการสำหรับการพยากรณ์ทั้ง 3 รูปแบบทั้ง Holt's Exponential Smoothing, Weight Moving Average และ Linear Trend Line เพื่อพยากรณ์จำนวนค่าเฉลี่ยคนรับชมวิดีโอในไตรมาสที่ 2 ปีพ.ศ. 2566 พบว่าได้ผลลัพธ์เป็น 111,181.52 110,256.12 และ 110,592.08 และมี MAPE เป็น 0.249 0.213 และ 0.479 ตามลำดับ

นอกเหนือจากนี้ หากพิจารณาค่าของ MAPE สำหรับการพยากรณ์แบบ Weight Moving Average มีค่าที่ต่ำที่สุดจากทั้ง 3 แบบ แสดงว่าการพยากรณ์แบบดังกล่าวมีความเหมาะสมมากที่สุด

ทั้งนี้ผลจากการพยากรณ์นี้อาจไม่สามารถพยากรณ์ได้อย่างแม่นยำเนื่องจากในแต่ละไตรมาสนั้นมีจำนวนวิดีโอที่อัปโหลด หมวดหมู่ที่อัปโหลด หรือความน่าสนใจของเนื้อหาที่นำเสนอไม่เหมือนกัน

ตาราง 33 ตารางแสดงการหา MAPE ของการพยากรณ์ 3 รูปแบบ

No	yearQtr	Holt Error /Actual	MWA Error /Actual	Linear Tread Error /Actual
1	2018-Q1	-	-	1.313
2	2018-Q2	-	-	4.053
3	2018-Q3	0.770	-	0.351
4	2018-Q4	0.695	0.692	0.524
5	2019-Q1	0.175	0.000	0.109
6	2019-Q2	0.330	0.000	0.358
7	2019-Q3	0.048	0.259	0.115
8	2019-Q4	0.231	0.381	0.188
9	2020-Q1	0.413	0.562	0.556
10	2020-Q2	0.094	0.051	0.228
11	2020-Q3	0.275	0.129	0.611
12	2020-Q4	0.075	0.040	0.517
13	2021-Q1	0.308	0.356	0.053
14	2021-Q2	0.335	0.363	0.151
15	2021-Q3	0.035	0.034	0.193
16	2021-Q4	0.142	0.003	0.104
17	2022-Q1	0.495	0.440	0.353
18	2022-Q2	0.002	0.061	0.051
19	2022-Q3	0.097	0.059	0.055
20	2022-Q4	0.092	0.287	0.168
21	2023-Q1	0.116	0.102	0.009
Total		4.73	3.82	10.06

เพราะฉะนั้น MAPE จึงคำนวณได้เป็น

$$MAPE_{Holt} = 4.73 / 19 = 0.249$$

$$MAPE_{WMA} = 3.82 / 18 = 0.212$$

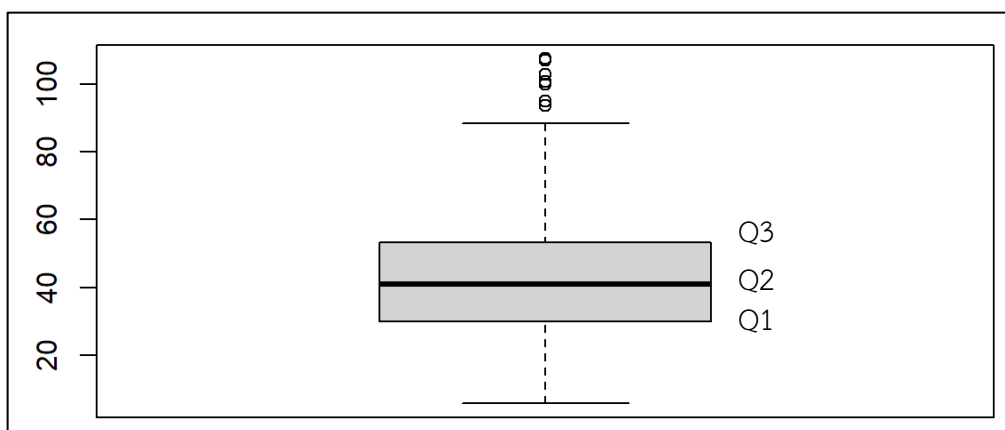
$$MAPE_{Linear} = 10.06 / 21 = 0.479$$

3.2. การทำ Normality Test กับความยาววิดีโอด้วยการทดสอบแบบ Shapiro - Wilk

จุดประสงค์: เพื่อทดสอบการแจกแจงภายใต้เส้นโค้งปกติของความยาววิดีโอ

ขั้นตอนที่ 1: พิจารณาแผนภาพ

- รูปภาพ 13 แผนภูมิฮิสโทแกรมแสดงการกระจายตัวของความยาววิดีโอ
- แผนภาพกล่องแสดงการกระจายของความยาววิดีโอ ดังรูป



รูปภาพ 40 แผนภาพกล่องแสดงการกระจายตัวของความยาววิดีโอ

จากแผนภาพกล่องแสดงให้เห็นว่าข้อมูลมีค่านอกเกณฑ์ (Outlier) จำนวนหนึ่ง ๆ และขนาดหรือความสูงกล่องในแต่ละส่วนไม่เท่ากัน

ขั้นตอนที่ 2: ทำการทดสอบสมมติฐาน

วิธีทำ H_0 : ข้อมูลมีการแจกแจงเป็นโค้งปกติ

H_a : ข้อมูลมีการแจกแจงไม่เป็นโค้งปกติ

กำหนดระดับนัยสำคัญ $\alpha = 0.05$

ซึ่งจากการใช้โปรแกรม R ได้ผลลัพธ์ ดังรูป

```
Shapiro-wilk normality test
data:  df$duration_min
w = 0.97801, p-value = 1.945e-11
```

รูปภาพ 41 Output ที่ได้จากการทำ Shapiro - Wilk Test จากโปรแกรม R

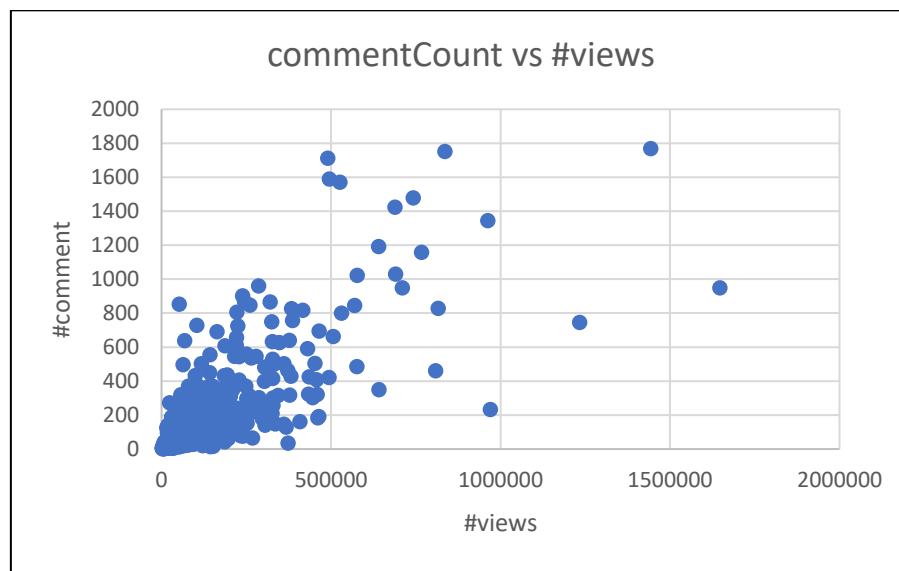
เนื่องจาก $P - \text{value} < \alpha$

ดังนั้น การแจกแจงความยาววิดีโอนี้ไม่เป็นการแจกแจงโค้งปกติ#

หมายเหตุ : สามารถใช้การทดสอบดังกล่าวกับการทดสอบข้อกำหนดในหัวข้อ 4.2. ได้เช่นกัน

3.3. การวิเคราะห์ด้วยสถิติเชิงพรรณนา (เพิ่มเติม)

3.3.1. การกระจายระหว่างจำนวนคนรับชมวิดีโอและข้อความในแสดงความคิดเห็น

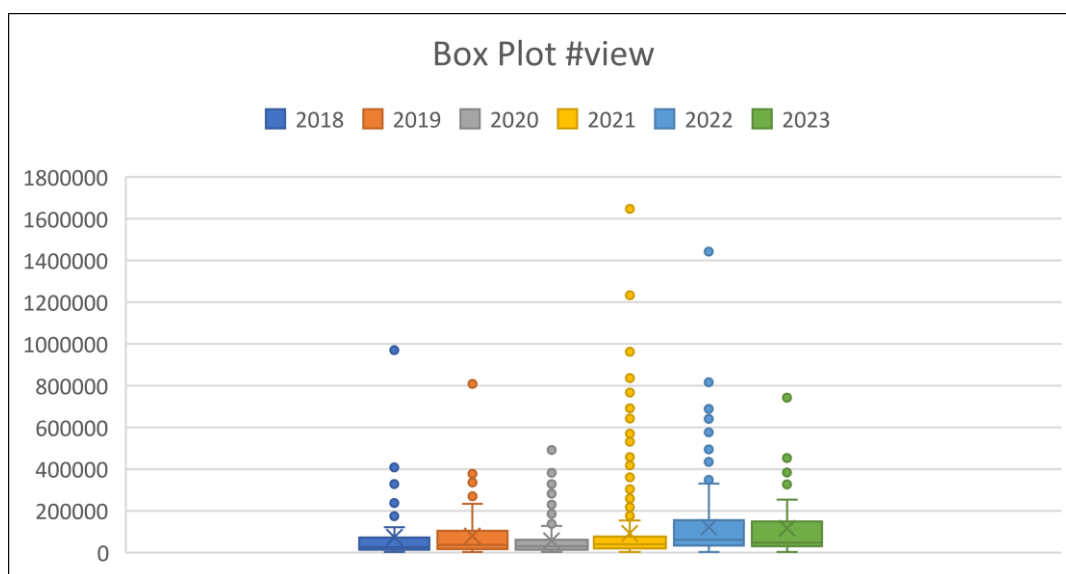


รูปภาพ 42 แผนภาพกระจายระหว่างจำนวนคนรับชม

และจำนวนข้อความแสดงความคิดเห็น

จากแผนภูมิดังกล่าวแสดงให้เห็นว่าจำนวนคนรับชมวิดีโอและจำนวนข้อความที่แสดงความคิดเห็นในแต่ละวิดีโอมีความสัมพันธ์กันในทิศทางบวก ซึ่งยิ่งจำนวนคนรับชมวิดีโอมากขึ้นจะทำให้มีการแสดงความคิดเห็นมากขึ้น

3.3.2. แผนภาพกล่องแสดงการแจกแจงจำนวนคนรับชมวิดีโอ



รูปภาพ 43 แผนภาพกล่องแสดงจำนวนคนรับชมวิดีโอในแต่ละปี

จากแผนภาพกล่องแสดงให้เห็นว่าจำนวนคนรับชมวิดีโอในแต่ละคลิปนั้น มีความไม่แน่นอน เนื่องจากมีจำนวนค่านอกเกณฑ์ (Outlier) ที่เปลี่ยนแปลงไปอย่างไม่แน่นอน โดยมีบางข้อมูลที่มีค่ามากกว่าค่าเฉลี่ยในแต่ละปีเป็นอย่างมาก ซึ่งจากตาราง 24 ตาราง แจกแจงจำนวนคนรับชมในแต่ละปี จะเห็นได้ชัดเจนอย่างยิ่ง

