**เรื่อง Google Scholar**

**จัดทำโดย**

นาย ธนกฤษต ฤทธิ์เดช

เลขนักศึกษา 6209650404

**เสนอ**

อาจารย์ วสิศ ลิ้มประเสริฐ

รายงานนี้เป็นส่วนหนึ่งของวิชา DSI200

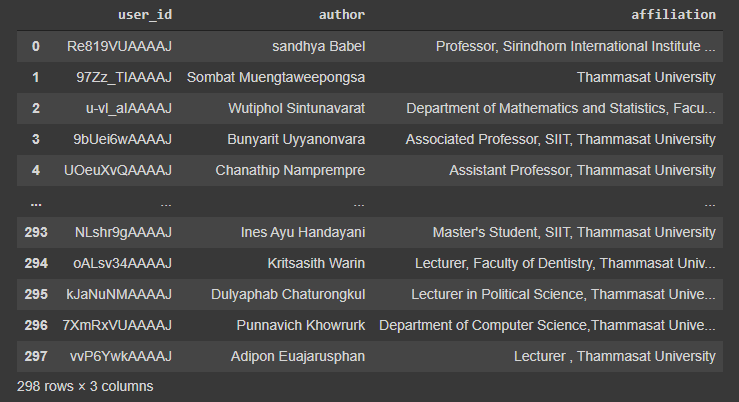
ภาคเรียนที่ 2 ปีการศึกษา 2563

มหาวิทยาลัยธรรมศาสตร์(ศูนย์รังสิต)

**Project: Google Scholar**

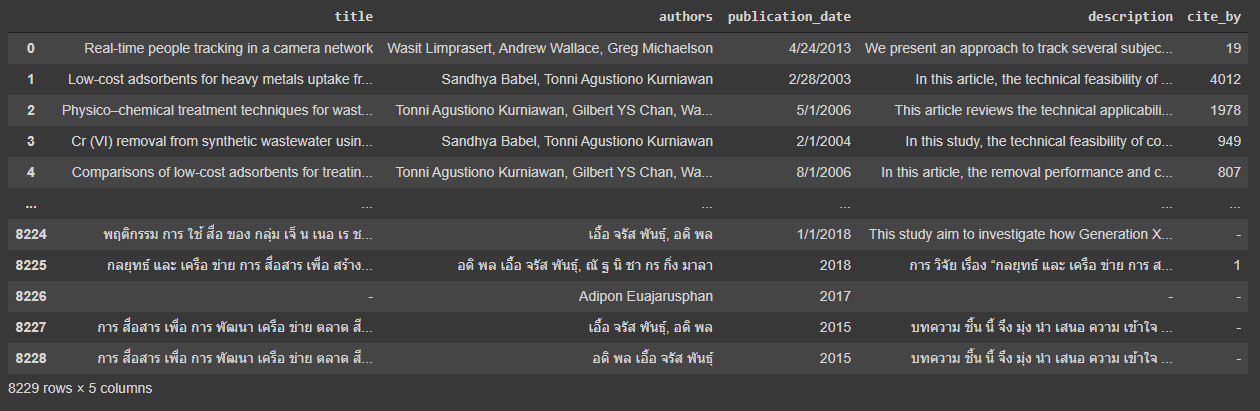
จากการศึกษาเกี่ยวกับการทำ web scraping เพื่อเก็บรวบรวมข้อมูล โดยใช้ HTTP request เก็บข้อมูลต่างๆ ชื่ออาจารย์ และบทวิจัยต่างๆ ของมหาวิทยาลัยธรรมศาสตร์ จากเว็บ Google Scholar ได้ข้อมูลมาดังนี้

1. ตารางเก็บข้อมูล User\_ID ชื่อของผู้เขียน และสังกัด/คณะ



จากข้อมูลที่ได้มาทั้งหมดนั้น สามารถสรุปได้ว่า ในวันที่ทำการเก็บรวบรวมข้อมูลล่าสุดนั้น และข้อมูลสรุปรวมแถวที่ท้ายตาราง ทำให้ทราบว่ามีผู้เขียนบทวิจัย ที่อยู่ในมหาวิทยาธรรมศาสตร์อยู่ที่ประมาณ 298 คน

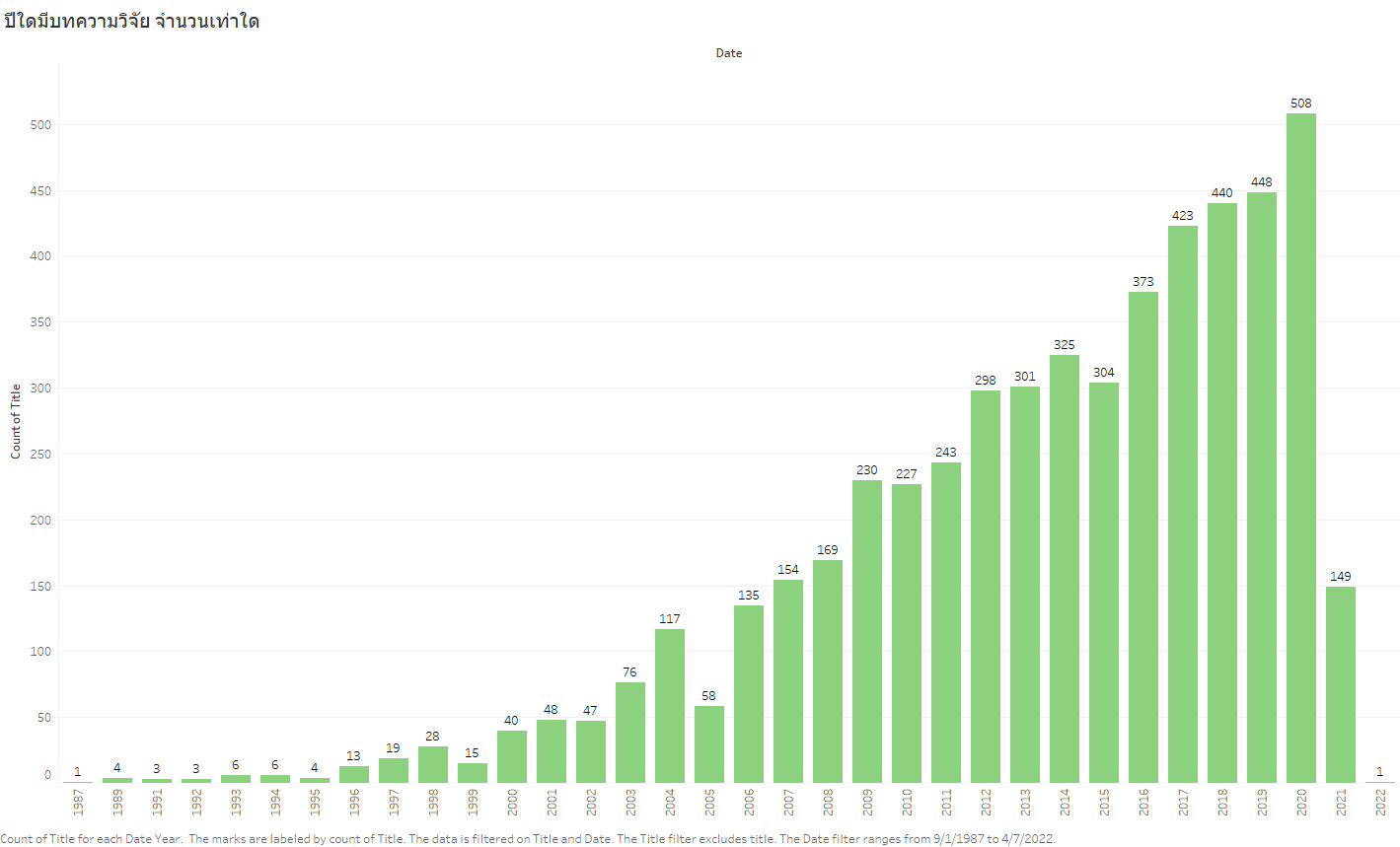
2. ตารางเก็บชื่อบทความ ผู้เขียน วันที่ และคำอธิบาย



จากข้อมูลที่ได้มาทั้งหมดนั้น สามารถสรุปได้ว่า ในวันที่ทำการเก็บรวบรวมข้อมูลล่าสุดนั้น และข้อมูลสรุปรวมแถวที่ท้ายตาราง ทำให้ทราบว่ามีบทวิจัยที่ถูกเขียนบน Google Scholar จากอาจารย์ในมหาวิทยาลัยธรรมศาสตร์เป็น เป็นจำนวน 8229 เรื่อง

และจากข้อมูลที่ได้มานั้น มีข้อมูลที่เป็นประเภทวันที่ จึงนำออกมาวิเคราะห์ และทำเป็นกราฟได้ จากการวิเคราะห์ทำให้เห็นว่า มีบทวิจัยกว่า 3000 เรื่อง ที่ไม่ถูกใส่วันที่เขียน หรือไม่สามารถดึงข้อมูลวันที่มาได้ และนำข้อมูลที่เหลือนั้น มาทำเป็นกราฟ เพื่อเปรียบเทียบว่า ในแต่ละปีนั้น มีผู้เขียนบทวิจัยเป็นจำนวนเท่าใด

ดังนี้



และจากข้อมูลที่เก็บมานั้น มีข้อมูลที่บอกถึงจำนวนผู้เข้าชมบทวิจัย (Cite by) จึงนำมาเปรียบเทียบ และเรียงลำดับ จากจำนวนผู้เข้าชม ตั้งแต่จำนวนผู้เข้าชมมากที่สุด จนไปถึงจำนวนผู้เข้าชมน้อยที่สุด และนำมาทำเป็นกราฟแท่งได้ ดังนี้

