

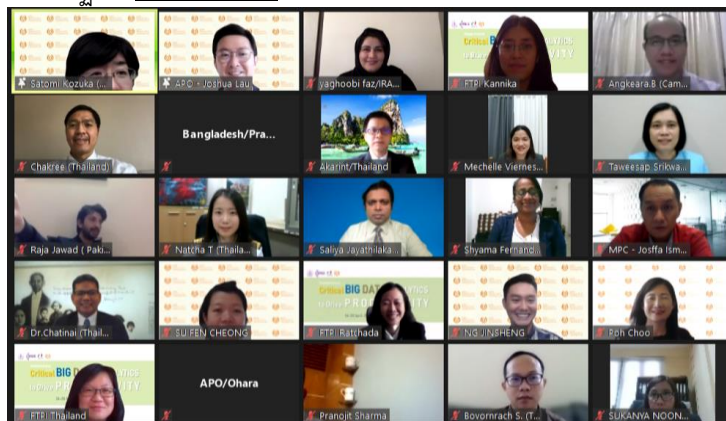
รายงานการเข้าร่วมโครงการเอพีไอ  
20-IN-97-GE-TRC-A: Training Course on Critical Big Data Analytics to Drive Productivity  
ระหว่างวันที่ 26-28 เมษายน 2564  
ผ่านระบบการประชุมออนไลน์

จัดทำโดย นางสาวณัฐชา เตชะชัยนรินทร์  
วิศวกรปฏิบัติการระดับสูง  
สำนักงานคณะกรรมการกิจการกระจายเสียง กิจการโทรทัศน์ และกิจการโทรคมนาคมแห่งชาติ  
วันที่ 29 เมษายน 2564

## ส่วนที่ 1 เนื้อหา/องค์ความรู้จากการเข้าร่วมโครงการ

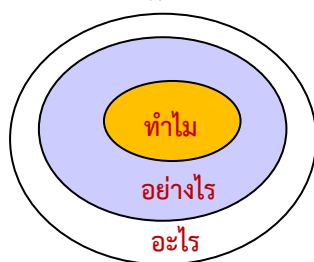
### 1.1 ที่มาหรือวัตถุประสงค์ของโครงการ

- Big data มีบทบาทสำคัญต่อการผลักดันให้เกิดการเปลี่ยนแปลงไปสู่ยุคดิจิทัล (Digital Transformation) โดยกระบวนการวิเคราะห์ข้อมูลขนาดใหญ่ หรือที่เรียกว่า Big Data Analytics เป็นการนำเทคโนโลยีมาประยุกต์ใช้เพื่อเพิ่มประสิทธิภาพในการตัดสินใจ และวางกลยุทธ์ขององค์กรให้บรรลุเป้าหมาย
- โครงการ 20-IN-97-GE-TRC-A: Training Course on Critical Big Data Analytics to Drive Productivity มีวัตถุประสงค์เพื่อให้ผู้เข้าร่วมโครงการเข้าใจการใช้เครื่องมือ และกระบวนการวิเคราะห์ข้อมูลขนาดใหญ่ ซึ่งประกอบด้วย การรวบรวมข้อมูล การวิเคราะห์ข้อมูล และการแสดงผลในรูปแบบที่เข้าใจง่าย
- โครงการนี้มีผู้เข้าร่วมจำนวน 48 คน จาก 13 ประเทศ รายละเอียดปรากฏตามเอกสารแนบ 1 และมีกำหนดการปรากฏตามเอกสารแนบ 2



### 1.2 เนื้อหา/องค์ความรู้

- Critical Big Data Analytics to Drive Productivity โดย Ms. TAN Poh Choo (เอกสารประกอบการบรรยายปรากฏตามเอกสารแนบ 3.1)



การเปลี่ยนแปลงไปสู่ยุคดิจิทัล โดยใช้ Golden Circle ของ Mr. Simon Sineks

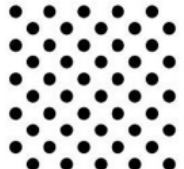
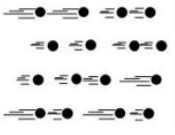
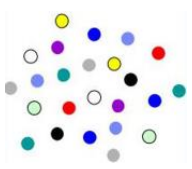

ทำไม – กระบวนการวิเคราะห์ข้อมูลขนาดใหญ่

อย่างไร – การบริหารข้อมูลสำหรับบริการสาธารณะ

อย่างไร – การเพิ่มผลผลิตโดยใช้เทคโนโลยี

อะไร – การจัดการให้มีบริการดิจิทัลที่ยั่งยืน

ข้อมูลขนาดใหญ่ หรือ Big Data มีลักษณะสำคัญ คือ

<b>Volume</b> มีจำนวนมาก	<b>Velocity</b> มีการเปลี่ยนแปลงที่รวดเร็ว	<b>Variety</b> มีความหลากหลาย	<b>Veracity</b> มีความไม่ชัดเจน
			

กระบวนการวิเคราะห์ข้อมูลขนาดใหญ่ และประเภทบริการสาธารณะที่เกี่ยวข้อง

ขั้นตอน	อธิบาย	วินิจฉัย	คาดการณ์	กำหนด	ดำเนินการล่วงหน้า
คำถาม	เกิดอะไรขึ้น	เกิดขึ้นได้อย่างไร	จะเกิดอะไรขึ้นต่อไป	จะต้องทำอะไร	ทำอะไรเพิ่มได้อีกหรือไม่
สิ่งที่ต้องทำ	ข้อมูลเบื้องต้น	ข้อมูลเชิงลึก		การตัดสินใจ	การปฏิบัติ
ประเภทบริการสาธารณะ	การบริหารส่วนกลาง	โครงสร้างพื้นฐานและสิ่งแวดล้อม	สังคม	ความปลอดภัย	เศรษฐกิจ

ตัวอย่างประเภทบริการสาธารณะที่สามารถนำการบริหารข้อมูลมาประยุกต์ใช้

<b>โครงสร้างพื้นฐานและสิ่งแวดล้อม</b>	การป้องกันโรคระบาดหรือความผิดปกติทางพันธุกรรม การคาดการณ์การเปลี่ยนแปลงสภาพภูมิอากาศ การลดมลภาวะ และการจัดการเพื่อลดผลกระทบจากภัยพิบัติทางธรรมชาติ การปรับปรุงประสิทธิภาพของระบบการประปา การแก้ไขปัญหาห่วงโซ่อุปทาน
<b>สังคม</b>	การสนับสนุน และตอบสนองความต้องการของประชาชน การเพิ่มประสิทธิภาพการตัดสินใจของบุคลากร
<b>ความปลอดภัย</b>	ความปลอดภัยทางไซเบอร์ การคาดการณ์ภัยคุกคาม หรือกิจกรรมที่ผิดกฎหมายที่อาจเกิดขึ้น การป้องกันและลดอาชญากรรม

การเพิ่มผลผลิตโดยใช้เทคโนโลยีเพื่อประชาชนมีหลักการ คือ 1) รับฟังสิ่งที่ประชาชนต้องการ 2) ทำความเข้าใจและวิเคราะห์เพื่อกำหนดแนวปฏิบัติ และ 3) ดำเนินการด้วยวิธีที่ถูกต้องภายในกรอบเวลาเหมาะสม ยกตัวอย่างเช่น Life SG Application ของประเทศสิงคโปร์ ซึ่งจัดทำขึ้นเพื่ออำนวยความสะดวกให้ประชาชนสามารถเข้าถึงข้อมูลและบริการที่จำเป็นของภาครัฐได้ผ่านช่องทางเดียว โดย Application ดังกล่าวจะปรับเปลี่ยนตามความต้องการของผู้ใช้งานแต่ละคน หรือ Application สำหรับเฝ้าระวังการแพร่ระบาดของโรคติดเชื้อไวรัสโคโรนา 2019 เป็นต้น



อย่างไรก็ตาม ความไว้วางใจและความโปร่งใสในการเข้าถึงข้อมูลเป็นสิ่งสำคัญในการนำเทคโนโลยีมาใช้เพิ่มผลผลิต โดยภาครัฐต้องสร้างความมั่นใจให้กับประชาชนที่อาจมีความกังวลในประเด็นเรื่องความเป็นส่วนตัวและข้อมูลส่วนบุคคล

### การจัดการให้มีบริการดิจิทัลที่ยั่งยืน

ประเทศออสเตรเลียสนับสนุนให้มีบริการดิจิทัลเพื่อเพิ่มผลผลิตในหลากหลายด้าน เช่น การรายงานภัยพิบัติ การบริการสุขภาพและสาธารณสุข การจัดการสวัสดิการ เป็นต้น ส่วนประเทศสิงคโปร์ในปัจจุบันมีการดำเนินการของภาครัฐผ่านระบบดิจิทัลคิดเป็น 95% และในปี พ.ศ. 2563 ประชาชนมีความพึงพอใจในการผลักดันให้เกิดการเปลี่ยนแปลงไปสู่ยุคดิจิทัลของภาครัฐมากถึง 85%

### ความท้าทายในการผลักดันให้เกิดการเปลี่ยนแปลงไปสู่ยุคดิจิทัล



People



Process



Technology

บุคลากรไม่มีความพร้อมหรือไม่มีความรู้ที่จะเปลี่ยนแปลงไปสู่ยุคดิจิทัล หรือต่อต้านไม่ให้เกิดการเปลี่ยนแปลง

ระบบการทำงานมีหลายขั้นตอนหรือหลายลำดับขั้น ทำให้การเปลี่ยนแปลงไปสู่ยุคดิจิทัล ดำเนินการได้ช้าและมีอุปสรรคมาก

มีเทคโนโลยีหลากหลาย และอาจทำให้เลือกเทคโนโลยีที่ไม่เหมาะสมกับงาน ซึ่งเป็นผลให้การทำงานไม่มีประสิทธิภาพและไม่คุ้มค่าต่อการลงทุน

### การสร้างความมั่นใจให้กับสาธารณะทำได้โดยใช้หลักการ

**Goal**

เป้าหมาย

**Reality**

ความเป็นจริง

**Opportunity**

โอกาส

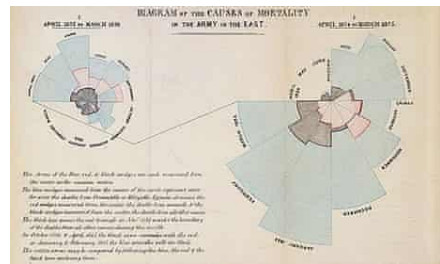
**Will**

ความตั้งใจ

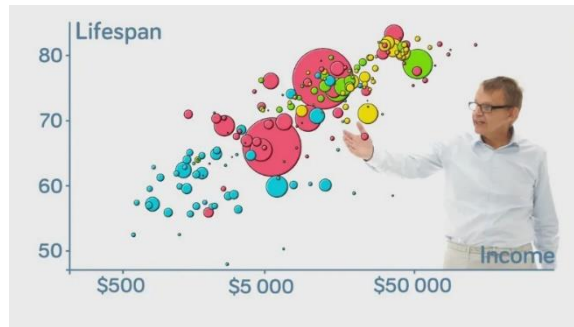
- Training of Trainers on Critical Big Data Analytics to Drive Productivity โดย Mr. NG Jinsheng (เอกสารประกอบการบรรยายปรากฏตามเอกสารแนบ 3.2)

### ประวัติบุคคลที่มีชื่อเสียงที่ใช้ประโยชน์จากการวิเคราะห์ข้อมูล

Florence Nightingale เป็นพยาบาลและนักสถิติจากสหราชอาณาจักร ในช่วงปี ค.ศ. 1820-1910 หรือถูกขนานนามว่า The Lady with the Lamp เธอได้ทำสถิติของทหารอังกฤษที่เข้ารับรักษาตัวในช่วงสงคราม Crimean โดยใช้แผนภาพ Pie Chart และค้นพบว่าสาเหตุการเสียชีวิตของทหารส่วนใหญ่เกิดจากการติดเชื้อและสุขอนามัยที่ไม่ดี



\*<https://www.theguardian.com/news/datablog/2010/aug/13/florence-nightingale-graphics>



**Hans Rosling** วิเคราะห์ข้อมูล อายุขัยและรายได้ของประชากรโลก ในช่วงเวลา 200 ปี โดยใช้แผนภาพ Scatterplot และพบว่าในช่วงปี ค.ศ. 1810-1948 ประชากรในทวีปยุโรป มีอายุขัยมากขึ้นและรายได้สูงขึ้น ในขณะที่ประชากรในทวีปเอเชียมี อายุขัยและรายได้น้อยกว่า ซึ่งในช่วง

ต่อมาจนถึงปี ค.ศ. 2009 ประชากรในทวีปเอเชียเริ่มมีอายุขัยและรายได้ที่มากขึ้น แต่ประชากรใน ทวีปแอฟริกายังคงมีอายุขัยและรายได้น้อย ทั้งนี้ ประชากรโลกในภาพรวมมีแนวโน้มที่จะมีอายุขัย และรายได้ที่มากขึ้นเรื่อยๆ

\* <https://www.youtube.com/watch?v=jbkSRLYSojo>

แพลตฟอร์มหลักที่ใช้สำหรับวิเคราะห์ข้อมูล ตามที่ได้รับการรับรองจากรายงาน Gartner Magic Quadrant คือ Microsoft BI (Power BI) Tableau และ Qlik Sense ซึ่งแต่ละแพลตฟอร์มมี คุณลักษณะและประโยชน์ที่แตกต่างกัน จึงควรเลือกใช้ให้เหมาะสม

แผนภาพแบ่งออกตามลักษณะการวิเคราะห์ข้อมูลได้ดังนี้

การวิเคราะห์เชิงเปรียบเทียบ	การวิเคราะห์แนวโน้ม
Bar Chart	Line Chart
Pie Chart	Histogram

ขั้นตอนการเลือกแผนภาพที่เหมาะสม คือ



ข้อมูล สามารถถูกแบ่งออกเป็น

<b>ข้อมูลเชิงปริมาณ</b> ตัวเลขที่สามารถคำนวณเชิงคณิตศาสตร์ได้	<b>ข้อมูลเชิงคุณภาพ</b> ชื่อ ประเภท สถานะ คุณสมบัติ ไม่สามารถ คำนวณเชิงคณิตศาสตร์ได้
<b>ข้อมูลที่มีโครงสร้าง</b> มีโครงสร้างชัดเจนในรูปแบบตาราง สามารถ นำไปประมวลผลได้ทันที	<b>ข้อมูลที่ไม่มีโครงสร้าง</b> ไม่มีโครงสร้างชัดเจน เช่น รูปภาพ เสียง วิดีโอ ข้อความ เป็นต้น ต้องทำการปรับเปลี่ยนข้อมูล ก่อนนำไปประมวลผล

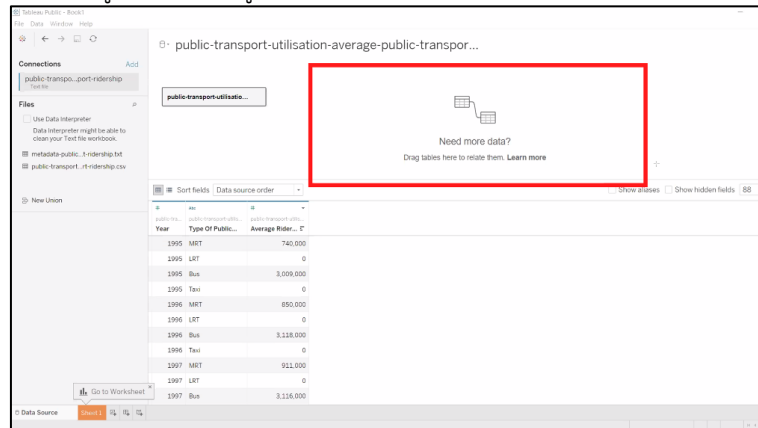
การใช้งานโปรแกรม Tableau

ตัวอย่างการใช้งานโปรแกรม Tableau โดยใช้ข้อมูลการใช้งานระบบขนส่งสาธารณะของประเทศ สิงคโปร์ ซึ่งข้อมูลดังกล่าวประกอบด้วย ปีที่เก็บข้อมูล ประเภทระบบขนส่งสาธารณะ และจำนวน ผู้ใช้งาน ระหว่างปี ค.ศ. 1995-2016

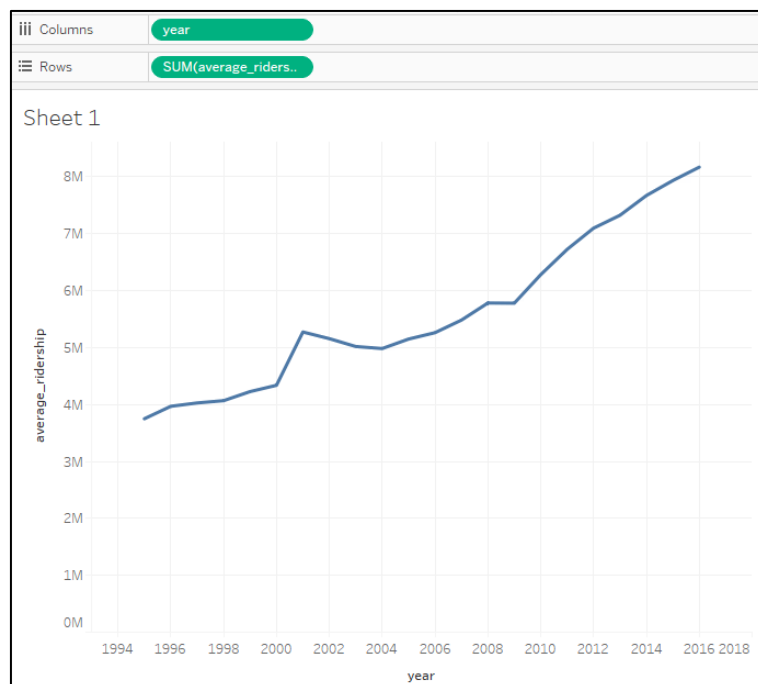
year	type_of_public_transport	average_ridership
1995	MRT	740000
1995	LRT	0
1995	Bus	3009000
1995	Taxi	0

\* <https://data.gov.sg/dataset/public-transport-utilisation-average-public-transport-ridership>

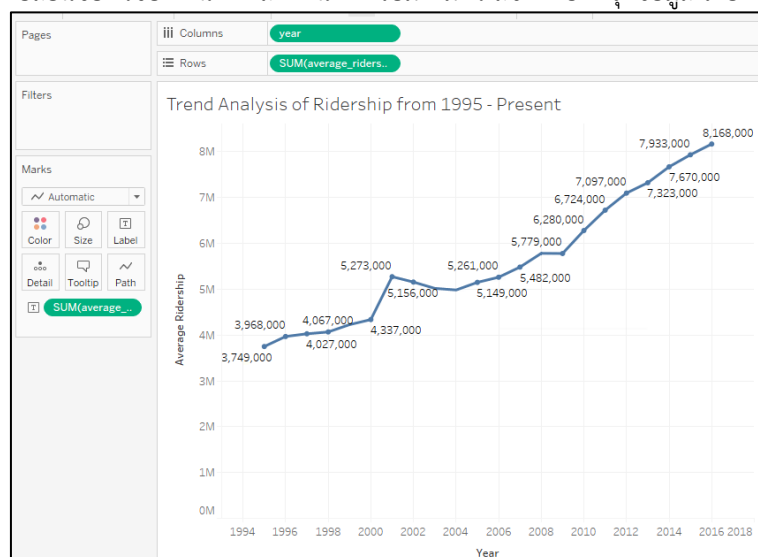
## นำเข้าข้อมูลดังกล่าวเข้าสู่โปรแกรม Tableau



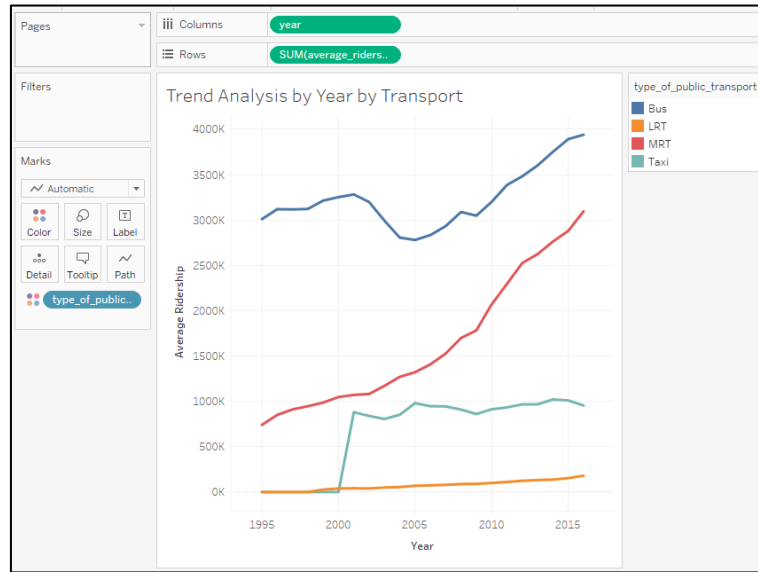
สร้าง Worksheet และนำตัวแปรเพิ่มลงใน Columns และ Rows โดยที่ Columns แสดงแกน X ของแผนภาพ และ Rows แสดงแกน Y ของแผนภาพ



เปลี่ยนชื่อหัวข้อ แกน X และแกน Y พร้อมเพิ่มตัวเลขกำกับแต่ละจุดข้อมูล โดยเพิ่มตัวแปรที่ Label



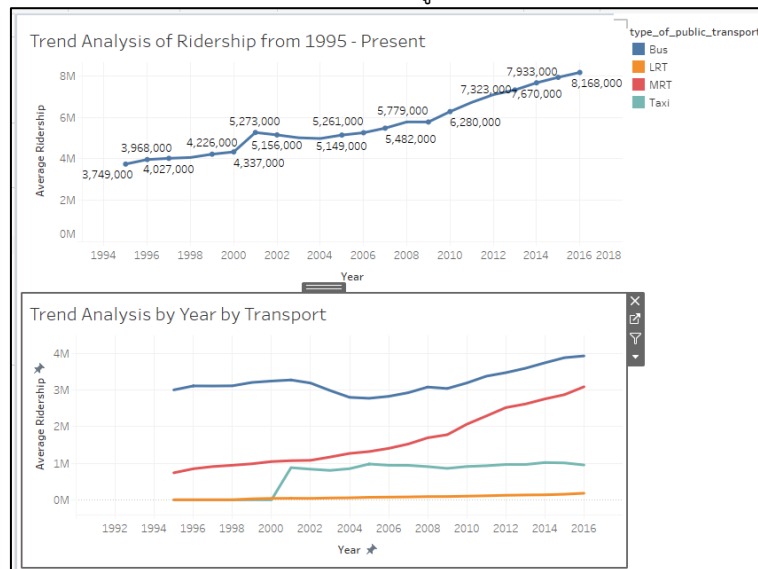
เพิ่มประเภทระบบขนส่งสาธารณะที่ Color เพื่อแยกจำนวนผู้ใช้งานตามประเภทระบบขนส่งสาธารณะ



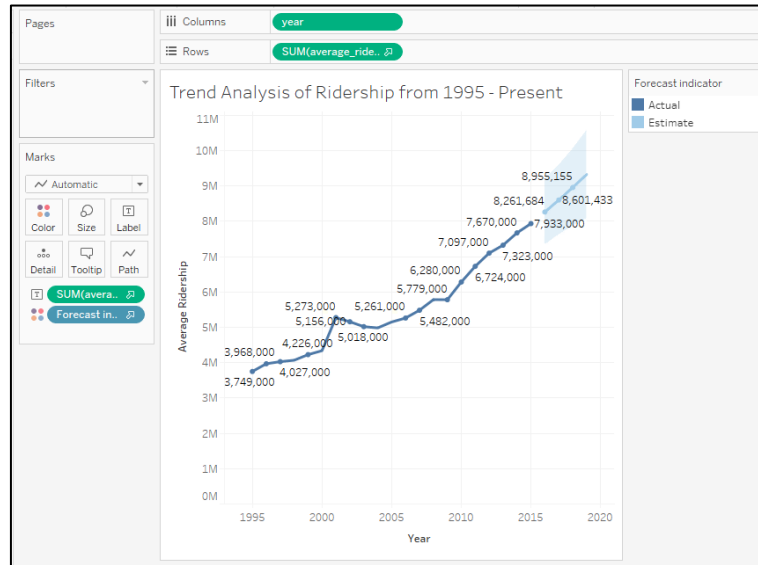
จากแผนภาพข้างต้นสามารถอธิบายการเปลี่ยนแปลงจำนวนผู้ใช้งานของระบบขนส่งสาธารณะแต่ละประเภทได้ดังนี้

1. ในช่วงปี ค.ศ. 2005-2011 มีการใช้งาน MRT เพิ่มขึ้นอย่างรวดเร็ว (Exponentially Increase)
2. ในช่วงปี ค.ศ. 2004-2015 มีการใช้งาน LRT เพิ่มขึ้นเล็กน้อย (Linearly increase)
3. ในช่วงปี ค.ศ. 2014-2016 มีการใช้งาน BUS เพิ่มขึ้นในปริมาณที่ลดลง (Decrease of increase)

จัดทำ Dashboard สำหรับการนำเสนอข้อมูล โดยนำแผนภาพที่เกี่ยวข้องแสดงในหน้าเดียวกัน

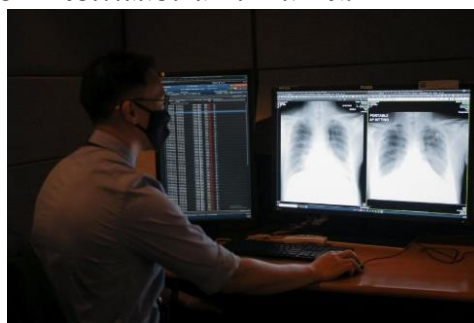


คาดการณ์จำนวนผู้ใช้งานระบบขนส่งสาธารณะ โดยการใช้การ Forecast จากจุดข้อมูลสุดท้าย ทั้งนี้ การ Forecast มีแบบ Additive (โมเดล Linear) และแบบ Multiplicative (โมเดล Compound)



จักรกลเรียนรู้ หรือ Machine Learning มีตัวอย่างกรณีการนำจักรกลเรียนรู้มาสร้างประโยชน์ต่อสังคม เช่น

1. ตำรวจในประเทศสิงคโปร์ใช้อากาศยานไร้คนขับเพื่อช่วยสำรวจและเฝ้าระวังการก่ออาชญากรรม โดยส่งอากาศยานไร้คนขับไปถ่ายภาพหรือวิดีโอในพื้นที่ต่างๆ โดยตำรวจไม่จำเป็นต้องลงพื้นที่เอง
2. เนื่องจากประเทศสิงคโปร์เริ่มเข้าสู่สังคมสูงอายุ และหนึ่งในสาเหตุหลักของการเสียชีวิตของผู้สูงอายุคือการหกล้ม จึงมีการนำกล้องดักจับความร้อนของร่างกายมาช่วยเฝ้าดูผู้สูงอายุ หากมีความผิดปกติของอุณหภูมิจะแจ้งเตือนให้สถานพยาบาลทราบเพื่อให้ผู้สูงอายุได้รับความช่วยเหลืออย่างเร่งด่วน
3. ประเทศญี่ปุ่นได้นำหุ่นยนต์สัตว์เลี้ยงอัจฉริยะมาให้ความช่วยเหลือและดูแลผู้สูงอายุ
4. ในสถานการณ์การแพร่ระบาดของโรคติดเชื้อไวรัสโคโรนา 2019 (COVID-19) ประเทศสิงคโปร์ได้นำเทคโนโลยี RadiLogic มาช่วยวิเคราะห์ภาพ X-ray ของปอดผู้มีความเสี่ยงติดเชื้อไวรัส โดยเทคโนโลยีดังกล่าวสามารถประมวลผลได้ภายในเวลาเพียง 3 วินาที ซึ่งทำให้กระบวนการคัดกรองผู้มีความเสี่ยงติดเชื้อไวรัสมีประสิทธิภาพมากขึ้น



\* <https://www.straitstimes.com/singapore/ai-powered-tool-detects-lung-infection-in-chest-x-rays-quickly-during-covid-19-screening>

5. สหรัฐอเมริกาใช้อัลกอริทึม Cluster Analysis เพื่อค้นหาเกษตรกรที่ไม่ได้นำเงินงบประมาณไปใช้อย่างถูกต้องตามวัตถุประสงค์ โดยใช้วิธีการหาเกษตรกรที่มีลักษณะผิดปกติจากกลุ่มที่ควรมีลักษณะคล้ายคลึงกัน พิจารณาจากจำนวนงบประมาณที่ขอ ปริมาณที่ดินที่ถือครอง และวัตถุประสงค์ในการของบประมาณ เป็นต้น

- บริษัท Walmart ในสหรัฐอเมริกาคาดการณ์สินค้าที่จะขายดีในสถานการณ์พายุเฮอริเคน และพบว่าสินค้าที่ขายดีที่สุดคือ Poptarts รสสตรอเบอร์รี่
- บริษัทขอรับบริจาคส่งเอกสารเชิญชวนการบริจาคไปที่กลุ่มคนที่คาดการณ์ว่าจะมีโอกาสการบริจาคสูงและบริจาคปริมาณมาก โดยวิเคราะห์จากลักษณะของผู้ที่อาจจะร่วมบริจาค เช่น พื้นที่อยู่อาศัย รายได้ การบริจาคในอดีต เป็นต้น

วิธีการเรียนรู้ของจักรกลทำได้โดยการกำหนดชุดเงื่อนไขในการตัดสินใจ ซึ่งประกอบด้วย ตัวแปร (Predictor) และเป้าหมายหรือผลลัพธ์ (Target/Outcome) ยกตัวอย่างในกรณีที่ได้ก็นำมือไปสัมผัสกาน้ำร้อนหรือกาน้ำเย็น และผลลัพธ์คือเด็กจะร้องไห้หรือไม่ สามารถกำหนดชุดเงื่อนไขการตัดสินใจได้ดังนี้

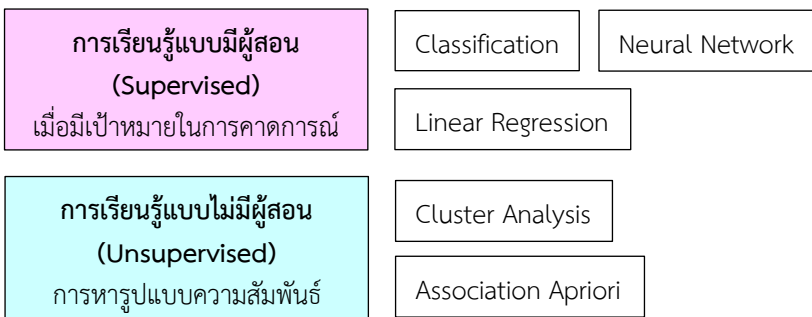
ตัวแปร		ผลลัพธ์
กาน้ำ	สัมผัสกาน้ำหรือไม่	ร้องไห้หรือไม่
ร้อน	สัมผัส	ร้องไห้
เย็น	ไม่สัมผัส	ไม่ร้องไห้
ร้อน	ไม่สัมผัส	ไม่ร้องไห้
เย็น	สัมผัส	ไม่ร้องไห้

โปรแกรมที่ใช้งานสำหรับการเรียนรู้เชิงลึก หรือ Deep learning ได้แก่ Python (neural network) R/R libraries และ SAS Enterprise Miner (commercial platform)/IBM SPSS ซึ่ง Python และ R ไม่มีค่าใช้จ่ายในการใช้งาน

คำถามที่สำคัญสำหรับจักรกลเรียนรู้ คือ 1. จะเกิดอะไรขึ้น และ 2. เครื่องมือที่ใช้ในการวิเคราะห์ข้อมูลเหมาะสมหรือไม่ ซึ่งการเลือกเครื่องมือต้องพิจารณาจากประเภทของตัวแปรและเป้าหมายที่ต้องการกล่าวคือ ตัวแปรเป็นข้อมูลที่มีโครงสร้างหรือไม่มีโครงสร้าง ผลลัพธ์ที่ต้องการเป็นเชิงคุณภาพหรือเชิงปริมาณ โดยสามารถเลือกโมเดลในเบื้องต้นได้ดังนี้

ผลลัพธ์ที่ต้องการเป็นข้อมูลเชิงคุณภาพ	ผลลัพธ์ที่ต้องการเป็นข้อมูลเชิงปริมาณ
Decision Tree	Simple Linear Regression
Neural Network	Multiple Linear Regression

อัลกอริทึมของจักรกลเรียนรู้ แบ่งออกเป็นการเรียนรู้แบบมีผู้สอน และการเรียนรู้แบบไม่มีผู้สอน โดยมีตัวอย่างอัลกอริทึมแต่ละประเภทดังนี้





## การสร้างโมเดล Simple Linear Regression โดยใช้โปรแกรม Excel

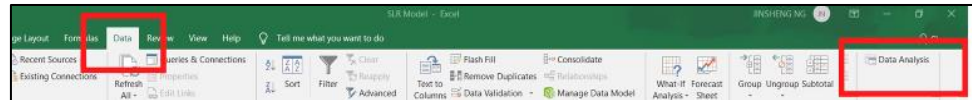
ตัวอย่างการสร้างโมเดลจากข้อมูลส่วนสูงและน้ำหนัก โดยต้องตั้งคำถามและตอบคำถามต่อไปนี้เพื่อเลือกโมเดลที่เหมาะสม

Height	Weight
146	72
191	86
196	107
156	91
194	92
197	75
193	88
189	91
189	100

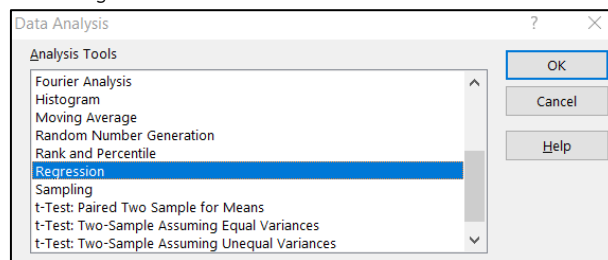
1. ตัวแปรคืออะไร (ตอบ: ส่วนสูง)
2. ผลลัพธ์คืออะไร (ตอบ: น้ำหนัก)
3. ผลลัพธ์เป็นข้อมูลประเภทใด (ตอบ: ข้อมูลเชิงปริมาณ)
4. มีกี่ตัวแปร (ตอบ: 1)

ดังนั้น เลือกใช้ Simple Linear Regression

เลือก Data และ Data Analysis



เลือก Regression



เพิ่มข้อมูลน้ำหนักลงใน Input Y และข้อมูลส่วนสูงลงใน Input X และเลือก Label เพื่อระบุว่าแถวแรกเป็นการระบุชื่อข้อมูล

Input

Input **Y** Range:

Input **X** Range:

Labels  Constant is Zero

Confidence Level:  %

Output options

Output Range:

New Worksheet Ply:

New Workbook

จะได้โมเดล Simple Linear Regression ดังนี้

SUMMARY OUTPUT								
<i>Regression Statistics</i>								
Multiple R	0.71187772							
R Square	0.506769888							
Adjusted R Square	0.502659637							
Standard Error	9.616764637							
Observations	122							
<i>ANOVA</i>								
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>			
Regression	1	11402.5094	11402.5094	123.2941482	3.87806E-20			
Residual	120	11097.85945	92.48216209					
Total	121	22500.36885						
	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>	<i>Lower 95.0%</i>	<i>Upper 95.0%</i>
Intercept	9.121829433	7.198639337	1.267160224	0.207551672	-5.130975466	23.37463433	-5.130975466	23.37463433
Height	0.453296034	0.040823542	11.10378981	3.87806E-20	0.372468261	0.534123807	0.372468261	0.534123807

## วิธีการอ่านโมเดล

- ค่า R-square อยู่ระหว่าง 0 ถึง 1 บอกความแม่นยำของโมเดล โมเดลมีความแม่นยำสูงเมื่อ R-square มีค่าเข้าใกล้ 1 ในกรณีนี้โมเดลมีความแม่นยำ 50.67%

Regression Statistics	
Multiple R	0.71187772
R Square	0.506769888
Adjusted R Square	0.502659637
Standard Error	9.616764637
Observations	122

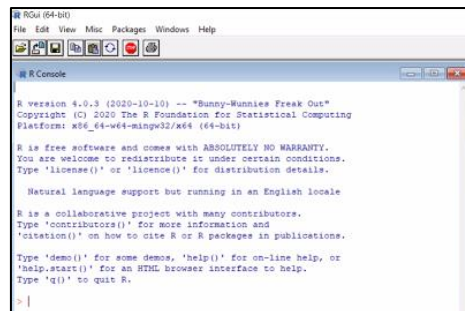
- ค่า Coefficients หรือสัมประสิทธิ์ของตัวแปร โดยพิจารณาจากค่า P-value ว่าจะปฏิเสธ Null Hypothesis หรือไม่ โดย Null Hypothesis กำหนดว่าตัวแปรมีค่าเป็นศูนย์ ซึ่งถ้า P-value มีค่าน้อยกว่า 5% แสดงว่า ตัวแปรดังกล่าวมีความสำคัญเชิงสถิติและจะต้องปฏิเสธ Null Hypothesis ในกรณีนี้ค่า P-value ของตัวแปรส่วนสูงมีค่าน้อยมาก ดังนั้น ตัวแปรส่วนสูงจึงมีความสำคัญเชิงสถิติและส่วนสูงที่เพิ่มขึ้นทุก 1 เซนติเมตร น้ำหนักควรจะเพิ่มขึ้น 0.453 กิโลกรัม

	Coefficients	Standard Error	t Stat	P-value
Intercept	9.121829433	7.198639	1.26716	0.207551672
Height	0.453296034	0.040824	11.10379	3.87806E-20

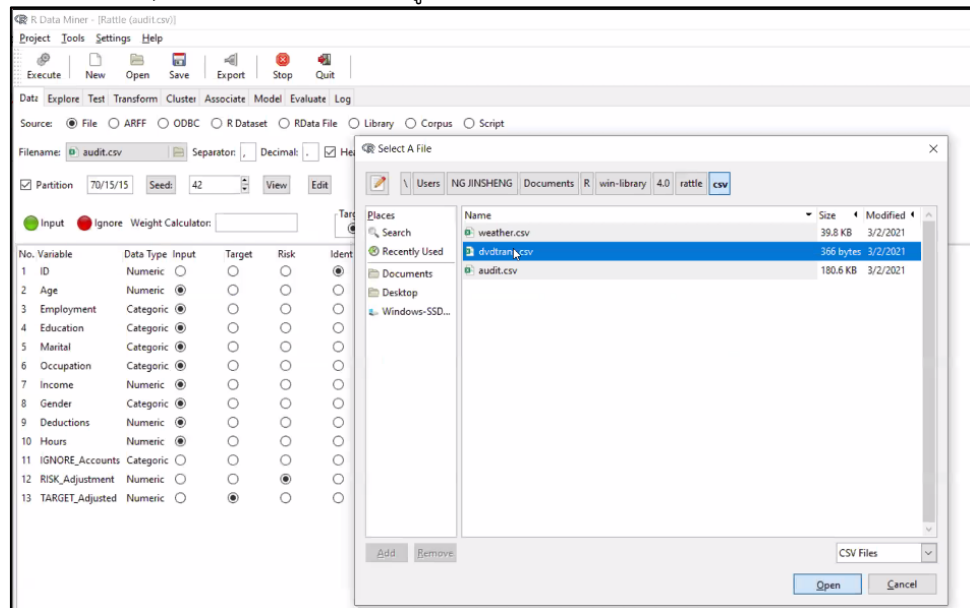
ทั้งนี้ ไม่จำเป็นต้องพิจารณาค่า P-value ของ Intercept

## การสร้างโมเดล Decision Tree โดยใช้โปรแกรม R

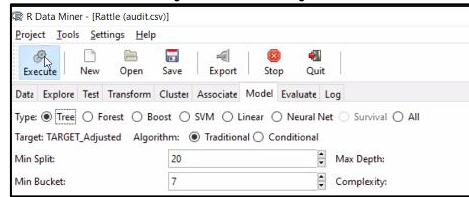
ตัวอย่างการสร้างโมเดลจากข้อมูลลักษณะของบุคคลเพื่อพิจารณาความเป็นไปได้ในการเลี้ยงภาษี โดยใช้โปรแกรม R4.0



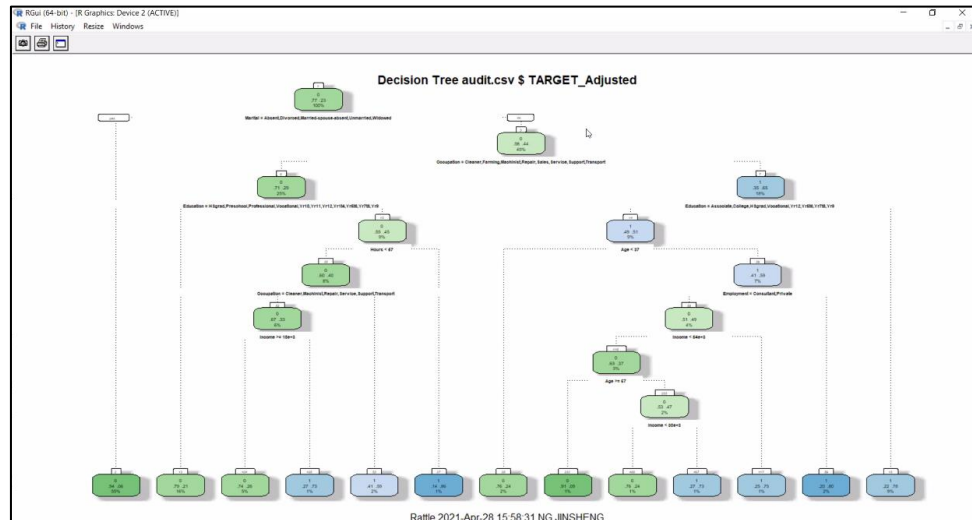
เลือกใช้ Library(rattle) เลือกประเภทข้อมูลของตัวแปร



## เลือกการแบ่งข้อมูลและความสูงของ Decision Tree



## จะได้โมเดล Decision Tree ดังนี้



## ส่วนที่ 2 ประโยชน์ที่ได้รับและการขยายผลจากการเข้าร่วมโครงการ

- ประโยชน์ต่อตนเอง/หน่วยงานต้นสังกัด/สายงานหรือวงการวิชาชีพ

การนำความรู้ในการวิเคราะห์ข้อมูลขนาดใหญ่โดยใช้ Analytics tool มาประยุกต์ใช้กับข้อมูล การใช้งานคลื่นความถี่ในปัจจุบัน ประกอบกับการคาดการณ์ความต้องการใช้งานคลื่นความถี่ ในอนาคต จะสามารถสนับสนุนการกำหนดนโยบายและการวางแผนคลื่นความถี่ให้เหมาะสมกับ การใช้งานจริง มีประสิทธิภาพและเกิดประโยชน์สูงสุดต่อประชาชน อุตสาหกรรม และประเทศ และ รองรับความก้าวหน้าของเทคโนโลยีใหม่ๆ
- กิจกรรมการขยายผลที่ได้ดำเนินการ

เสนอรายงานสรุปผลการเข้าร่วมโครงการ 20-IN-97-GE-TRC-A: Training Course on Critical Big Data Analytics to Drive Productivity ขององค์การเพิ่มผลผลิตแห่งเอเชีย ต่อเลขาธิการสำนักงาน คณะกรรมการกิจการกระจายเสียง กิจการโทรทัศน์ และกิจการโทรคมนาคมแห่งชาติ เพื่อทราบ และ นำรายงานดังกล่าวเผยแพร่ทางเว็บไซต์ของสำนักบริหารคลื่นความถี่ สำนักงานคณะกรรมการกิจการ กระจายเสียง กิจการโทรทัศน์ และกิจการโทรคมนาคมแห่งชาติ (spectrum.nbtc.go.th)

## ส่วนที่ 3 เอกสารแนบ

- เอกสารแนบ 1 - รายชื่อผู้เข้าร่วมโครงการและประเทศที่เข้าร่วมโครงการ
- เอกสารแนบ 2 - กำหนดการฉบับล่าสุด (Program)
- เอกสารแนบ 3 - เอกสารประกอบการประชุม/สัมมนา (Training Materials)

  - เอกสารแนบ 3.1 – เอกสารประกอบการบรรยาย โดย Ms. TAN Poh Choo
  - เอกสารแนบ 3.2 – เอกสารประกอบการบรรยาย โดย Mr. NG Jinsheng