

Final Project Report

1. Introduction

Dataset:

Disneyland Reviews Dataset บนเว็บไซต์ kaggle ซึ่งใน Dataset ประกอบด้วยคำวิจารณ์ 42,000 รายการ โพสต์บนแอปพลิเคชันแนะนำสถานที่ท่องเที่ยว Tripadvisor โดยผู้เข้าชม Disneyland สาขา California, Hong Kong, และ Paris ซึ่งข้อมูลภายใน Dataset ประกอบด้วย

1. Review_ID : รหัสเฉพาะของแต่ละคำวิจารณ์
2. Rating : ระดับความพึงพอใจ มีระดับตั้งแต่ 1 (ไม่พอใจ) ถึง 5 (พอใจมากที่สุด)
3. Year month : เดือนและปีที่ผู้วิจารณ์เข้าชม Disneyland
4. Review location : ประเทศต้นทางของผู้วิจารณ์
5. Review text : ความคิดเห็นของผู้วิจารณ์
6. Branch : สถานที่ตั้งของสาขานั้น ๆ

Review_ID	Rating	Year_Month	Reviewer_Location	Review_Text	Branch
670772142	4	2019-4	Australia	If you've ever been to Disneyland anywhere you'll find Disneyland Hong Kong very similar in the layout when you wal	Disneyland_HongKong
670682799	4	2019-5	Philippines	Its been a while since d last time we visit HK Disneyland .. Yet, this time we only stay in Tomorrowland ... AKA Marvel	Disneyland_HongKong
670623270	4	2019-4	United Arab Emirates	Thanks God it wasn't too hot or too humid when I was visiting the park otherwise it would be a big issue (there is n	Disneyland_HongKong
670607911	4	2019-4	Australia	HK Disneyland is a great compact park. Unfortunately there is quite a bit of maintenance work going on at present sc	Disneyland_HongKong
670607296	4	2019-4	United Kingdom	the location is not in the city, took around 1 hour from Kowlon, my kids like disneyland so much, everything is fine. I	Disneyland_HongKong
670591897	3	2019-4	Singapore	Have been to Disney World, Disneyland Anaheim and Tokyo Disneyland but I feel that Disneyland Hong Kong is really	Disneyland_HongKong
670585330	5	2019-4	India	Great place! Your day will go by and you won't even know it. Obviously went there for my daughter and she absolute	Disneyland_HongKong
670574142	3	2019-3	Malaysia	Think of it as an intro to Disney magic for the little ones. Almost all of the attractions can be completed in 1.5days.O	Disneyland_HongKong
670571027	2	2019-4	Australia	Feel so let down with this place,the Disneyland train was fantastic until you get past the station,bad signage,terrible	Disneyland_HongKong
670570869	5	2019-3	India	I can go on talking about Disneyland. Whatever I say about it, is less. Disneyland is all about your childhood dreams c	Disneyland_HongKong
670443403	5	2019-4	United States	Disneyland never cease to amaze me! I've been to Disneyland florida and I thought I have exhausted the kid in me bu	Disneyland_HongKong
670435886	5	2019-4	Canada	We spent the day here with our grown kids and I have to admit we had a great time. It seems the kids never tire of D	Disneyland_HongKong
670376905	4	2019-4	Australia	We spend two days, the second day went early then went straight to the back of the park, no lineups for so children j	Disneyland_HongKong
670324965	5	2019-4	Philippines	It was indeed the happiest place on earth! My family had really fun and it made my daughter's birthday a memorable	Disneyland_HongKong
670274554	5	2018-9	Australia	This place is HUGE! Definitely need more than one day. We had 3 children aged 11, 9 & 6 and they loved it. A great v	Disneyland_HongKong
670205135	3	2019-1	United Kingdom	We brought tickets before we left and got 2 days for less than the price of 1. After visiting would advise everyone to	Disneyland_HongKong
670199487	4	2019-4	Myanmar (Burma)	Its huge , not enough to visit in one day. We did with 2 days pass ... all the scenes were amazing ... you were like in m	Disneyland_HongKong
670129921	3	2019-4	United Kingdom	Around 60 per person and if you want to eat or drink at any point it was cost you a small fortune. If you go on a wei	Disneyland_HongKong
670099231	4	2019-4	Australia	It's Disneyland! It just does not need reviewing! Place speaks for itself. However smaller than I expected. Great ex	Disneyland_HongKong
670033848	5	2018-11	Hong Kong	There is nothing more to say except that you become a child once again as you step inside Disneyland in HK! Easy to	Disneyland_HongKong
670030877	5	2019-4	United Kingdom	We bought a 2 day tourist ticket via Klook which made it better value. It was good to have a day to get our bearings i	Disneyland_HongKong
670022335	3	2019-4	United States	First off I'm not a really big Disneyland fan. I live about 20 minutes from the California Disneyland, and haven't been i	Disneyland_HongKong
670007081	3	2019-4	United States	Mickey and the Wounderous book was great and the kids love it. The new Ant man attraction was good for all ages.	Disneyland_HongKong
669996858	3	2019-4	Singapore	I've been to Hong Kong Disneyland twice and even with new attractions such as the Ant Man and the Wasp, the exp	Disneyland_HongKong
669783708	5	2019-4	Australia	GO early if you want to do photos with the signboards and Mickey Donald Duck. Otherwise, you could spend an hour	Disneyland_HongKong
669548161	3	2019-4	Hong Kong	Disneylands the world over are generic apart from the language. This was no different than the others. It was a shari	Disneyland_HongKong
669528186	3	2019-4	Australia	This was my second trip here, the first been about 3 years ago. This is a smallish Disneyland park. The attractions are	Disneyland_HongKong
669403596	5	2018-12	Philippines	We have been to Hong Kong Disneyland a few times already but we still want to come back. Why? Because there are	Disneyland_HongKong

<https://www.kaggle.com/arushchillar/disneyland-reviews>

โดยปัจจัยที่ใช้ในการเลือก Dataset มี 4 ข้อ ดังนี้

1. ชุดข้อมูลต้องมีความหลากหลาย โดยข้อมูลที่เลือก มีการเก็บรวบรวมคำวิจารณ์ของผู้วิจารณ์ที่เข้าชม Disneyland หลายสาขา และมีผู้เข้าชมจากหลากหลายประเทศ
2. ชุดข้อมูลต้องมีความน่าเชื่อถือ โดยข้อมูลที่นำมาใช้มาจากแหล่งที่มาที่มีความน่าเชื่อถือ คือ เว็บไซต์ Kaggle ซึ่งอ้างอิงข้อมูลมาจากแอปพลิเคชัน Tripadvisor ที่เป็นแอปพลิเคชันที่มีความนิยม เป็นที่รู้จัก และมีผู้ใช้งานจำนวนมาก
3. ชุดข้อมูลต้องน่าสนใจ เพราะเป็นข้อมูลเกี่ยวกับ Disneyland ซึ่งเป็นสถานที่ที่มีชื่อเสียงและเป็นที่รู้จักในวงกว้าง ซึ่งมีส่วนช่วยในการเพิ่มความน่าสนใจของข้อมูลต่อการนำเสนอ
4. ชุดข้อมูลต้องเพียงพอต่อการนำมาทำ Dashboard เพื่ออธิบายและตอบคำถามเชิง Business intelligence ได้ โดยข้อมูลที่เลือกมาสามารถนำไปทำ visualization เพื่อตอบคำถาม Descriptive Analytics ได้อย่างเหมาะสม

จากการศึกษาข้อมูลที่มีจึงนำมาตั้งเป็นคำถามเชิง Business intelligence
จะได้ 3 ข้อ ดังนี้

Predictive analytics question

1. Rating ในปีต่อไป ของ Disneyland จะเป็นอย่างไร?

Target audience:

- Disneyland

Objective:

- นำข้อมูลมาประกอบการตัดสินใจในการปรับปรุงและพัฒนา Disneyland แต่ละสาขาเพื่อตอบสนองความต้องการ รวมถึงเพิ่มระดับความพึงพอใจของผู้เข้าชม

2. การจัดกลุ่มปัจจัย ว่ามีอะไรบ้างที่มีผลต่อการดึงดูดนักท่องเที่ยว?

Target audience:

- Disneyland

Objective:

- แยกประเภทระหว่างปัจจัยที่ส่งผลให้ผู้เข้าชมรู้สึกดี (positive) และปัจจัยที่ทำให้ผู้เข้าชมรู้สึกแย่ (negative) เพื่อนำไปวางแผนปรับปรุงสถานที่ สิ่งอำนวยความสะดวก จัดสรรหรือสร้างโปรโมชั่น รวมถึงปัจจัยอื่น ๆ ให้สอดคล้องกับความต้องการของผู้เข้าชม

Descriptive analytics question**3. การเปลี่ยนแปลงของค่าเฉลี่ย Rating ในแต่ละปีของแต่ละสาขาเป็นอย่างไร?****Target audience:**

- Disneyland

Objective:

- ดูการเปลี่ยนแปลงของค่าเฉลี่ยระดับความพึงพอใจของผู้เข้าชม Disneyland ในแต่ละปี เพื่อเป็นการติดตามผลตอบรับจากการปรับปรุงและพัฒนาของ Disneyland แต่ละสาขา

2. Organizational memory

- **Data warehouse (DW) + Data lake (DL)**

ระบบจัดเก็บข้อมูลที่เกิดขึ้นภายใน Disneyland ใช้เทคโนโลยี Data Lake และ Data Warehouse ร่วมกัน ซึ่งเลือกใช้พื้นที่จัดเก็บตามคุณสมบัติ โดย Data Lake จะใช้กับการจัดเก็บข้อมูลที่มีการ generate เข้าสู่ระบบจำนวนมากๆ เพราะมีความยืดหยุ่น รองรับการขยายตัวของข้อมูลได้มาก เช่น ข้อมูล log การเข้าใช้งานหน้าเว็บไซต์ Disneyland ส่วน Data Warehouse จะใช้กับการเก็บข้อมูลสำหรับงานเฉพาะ ที่มีรูปแบบในการเก็บข้อมูลที่ชัดเจนเป็นแบบแผน เพื่อให้สามารถดึงข้อมูลไปใช้งานต่อได้สะดวก โดยใช้เทคโนโลยี ETL การเคลื่อนย้ายและตรวจสอบคุณภาพข้อมูล

- **Enterprise Resource Planning (ERP)**

เทคโนโลยี ERP เป็นซอฟต์แวร์สำหรับการผนวกระบบการทำงานของแผนกต่างๆภายใน Disneyland ให้ทั้งองค์กรทำงานระบบการทำงานบนระบบเดียวกัน สามารถนำมาปรับใช้ในการวางแผนธุรกิจ การดำเนินการแก้ไขปัญหภายในองค์กร ซึ่งเป็นการดำเนินการที่จำเป็นต้องทำงานร่วมกันหลายๆแผนก ERPจะเป็นตัวช่วยในการอำนวยความสะดวก เพิ่มความคล่องตัวให้กับการทำงานระหว่างหน่วยงาน และทำให้ข้อมูลมีความเป็นปัจจุบัน เช่น การดำเนินงานวางแผนการทำโปรโมชั่นหรือวางแผนจัดเทศกาลที่ต้องมีการปรับเปลี่ยนไปในทุก ๆ ไตรมาส เพื่อเพิ่มความสนใจให้กับ Disneyland การตอบปัญหาเรื่องการจัดกลุ่มปัจจัยที่ใช้ดึงดูดนักท่องเที่ยว

- **Knowledge repository (KR)**

ระบบเก็บรวบรวมข้อมูล ความรู้ เทคนิคและวิธีการจัดการแก้ไขปัญหาที่เกิดขึ้น การประชาสัมพันธ์ การจัดโปรโมชั่น หรือการดำเนินแผนทางธุรกิจภายใน Disneyland ที่เหมาะสมและได้ผลลัพธ์ที่ดี (Best Practice) ที่มีการนำมาใช้จริงในสาขาหนึ่ง ซึ่งอาจจะนำมาปรับใช้ในสาขาอื่นได้

3. Information integration

- **Environmental scanning (ES) + Future wheel (FW)**

นำเทคโนโลยี Environmental scanningมาใช้เป็นเครื่องมือในการรวบรวมข้อมูล ตรวจสอบเหตุการณ์ต่าง ๆ โดยมองหาปัจจัยที่เกิดขึ้นจากทั้งภายในและภายนอก Disneyland เช่น สภาพเศรษฐกิจ วัฒนธรรมของแต่ละพื้นที่ โดยนำข้อมูลที่ได้จากกระบวนการใช้ Environmental scanning นำไปวิเคราะห์หาแนวโน้มและความเชื่อมโยงของแต่ละปัจจัย ด้วยเทคโนโลยี Future wheel เพื่อจัดลำดับความสัมพันธ์ของแต่ละปัจจัยให้เกิดเป็นผลลัพธ์ที่มีประโยชน์ ซึ่งสามารถนำมาใช้ในการตอบปัญหาการจัดกลุ่มปัจจัยที่มีผลต่อการดึงดูดนักท่องเที่ยวของดิสนีย์แลนด์ เพื่อการวางแผนการดำเนินการทางธุรกิจ การจัดโปรโมชั่นให้นักท่องเที่ยวที่เหมาะสมต่อไป

- **Radio Frequency Identification (RFID)**

การใช้อุปกรณ์ RFID tag กับผู้เข้าชม Disneyland เพื่อเก็บลักษณะการใช้งาน การเข้าชมบูธต่าง ๆ ภายใน Disneyland เวลาที่ใช้ในแต่ละจุด ในสถานที่ ช่วงเวลาและตำแหน่งที่มีผู้เข้าชมจำนวนมาก ซึ่งอาจนำมาใช้ปรับปรุงการจัดวางตำแหน่งของบูธต่าง ๆ ให้เหมาะสมกับการใช้งานของผู้เข้าชม

- **Text mining**

เทคโนโลยีที่ใช้กับการดึงข้อมูลความคิดเห็นของผู้คนที่มีการพูดถึงเกี่ยวกับ Disneyland บน Social Media ซึ่งเป็น platform ที่มีการใช้งานอย่างแพร่หลายในปัจจุบัน Social Media ได้เข้ามามีส่วนเกี่ยวข้องกับชีวิตประจำวัน ในการใช้เป็นพื้นที่ในการเล่าเรื่องราวประสบการณ์ของตนเอง หรือพบเจอจากสื่ออื่น ๆ มา แบ่งปันให้กับผู้อื่น ซึ่งถือเป็นข้อมูลสำคัญที่ใช้ร่วมกับข้อมูลจากแหล่งอื่นข้างต้น เพื่อวิเคราะห์หาปัจจัยที่ส่งผลต่อการดึงดูดนักท่องเที่ยว

4. Insight creation (includes the ML canvas)

For predictive analytics question

1. Rating ในปีต่อไป ของ disneyland จะเป็นอย่างไร?

Rating ในปีต่อไป ของ disneyland จะเป็นอย่างไร?					Designed by: อ้อชื่อกลุ่ม	Date:	Iteration:
Decisions How are predictions used to make decisions that provide the proposed value to the end-user? ช่วยในการตัดสินใจเกี่ยวกับทิศทางในการปรับปรุงและพัฒนา Disneyland ให้สามารถตอบสนองต่อความต้องการของผู้เข้าชมได้มากขึ้น	ML task Input, output to predict, type of problem. Q: Rating ในปีต่อไป ของ disneyland จะเป็นอย่างไร I/P : review data O/P : rating Problem : Regression	Value Propositions What are we trying to do for the end-user(s) of the predictive system? What objectives are we serving? Target audience: disneyland Rating ในปีต่อไป ของ Disneyland จะเป็นอย่างไร? Objectives: - นำข้อมูลมาประกอบการตัดสินใจในการปรับปรุงและพัฒนา Disneyland ให้ได้สาขาที่ตอบสนองความต้องการ รวมถึงเพิ่มระดับความพึงพอใจของผู้เข้าชม	Data Sources Which raw data sources can we use (internal and external)? - Disneyland review database from Tripadvisor - แบบประเมินความพึงพอใจต่อ Disneyland	Collecting Data How do we get new data to learn from (inputs and outputs)? เชื่อมต่อกับ database โดยตรง และเก็บ data ทุก ๆ วัน ดึงข้อมูล review จากแอปพลิเคชัน Tripadvisor			
Making Predictions When do we make predictions on new inputs? How long do we have to featureize a new input and make a prediction? ทำการปรับปรุงและพัฒนา Disneyland แต่ละสาขาในทุก ๆ ปีตามแผนธุรกิจที่อ้างอิงจากข้อมูล ที่ได้จากการทำ model ที่เหมาะสมกับความพึงพอใจ	Offline Evaluation Methods and metrics to evaluate the system before deployment. Regression problem; 1. R Square/Adjusted R Square 2. Mean Square Error(MSE)/Root Mean Square Error(RMSE) 3. Mean Absolute Error(MAE)	Building Models When do we create/update models with new training data? How long do we have to featureize training inputs and create a model? เก็บ data ทุก ๆ วันและนำมาวิเคราะห์เพื่อไปปรับปรุงแผนในการพัฒนาเป็นรายปี					
Live Evaluation and Monitoring Methods and metrics to evaluate the system after deployment, and to quantify value creation. ติดตามค่าเฉลี่ยของ Rating รวมถึงยอดเข้าชม ยอดนักท่องเที่ยวที่เข้ามา					Features Input representations extracted from raw data sources. - Rating ที่ลูกค้าประเมินโดยมีเกณฑ์ประเมินดังนี้ 1-(ไม่พึงพอใจ) ถึง 5-(พึงพอใจมาก) -Year and Month when the reviewer visited the theme park		

machinelearningcanvas.com by Louis Dorand, Ph.D. Licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

- Machine Learning

สำหรับการทำ Model จะเก็บ Data ทุก ๆ วันและนำมาวิเคราะห์เพื่อทำแผนในการพัฒนาเป็นรายปี ซึ่ง Model ที่ได้นำไปใช้ในการทำนายระดับความพึงพอใจของผู้เข้าชม Disneyland ในปีต่อไป ซึ่งทำเพื่อให้ได้ข้อมูลไปปรับปรุง Disneyland ให้ตอบสนองความต้องการของผู้เข้าชมและเพิ่มระดับความพึงพอใจ

สำหรับ Model ที่ใช้จะเป็นแบบ Regression เนื่องจากเป็นการวิเคราะห์ค่าความพึงพอใจและได้เป็นข้อมูลเชิงปริมาณ

2. การเปลี่ยนแปลงของค่าเฉลี่ย Rating ในแต่ละปีของแต่ละสาขาเป็นอย่างไร?

การวัดกลุ่มปัจจัย ว่ามีอะไรบ้างที่มีผลต่อการดึงดูดนักท่องเที่ยว ? signed by: อ้อชื่อกลุ่ม Date: Iteration:

Decisions How are predictions used to make decisions that provide the proposed value to the end-user? ช่วยทำให้สามารถตัดสินใจอะไรได้บ้าง เช่น การตอบรับของนักท่องเที่ยว และการช่วยในการหาแนวทางเพิ่มจำนวนนักท่องเที่ยวให้มากขึ้น	ML task Input, output to predict, type of problem. Q: ปัจจัยอะไรที่มีผลต่อการดึงดูดนักท่องเที่ยว? I/P: review data, social media, RFID O/P: ปัจจัยที่ส่งผลต่อการดึงดูดนักท่องเที่ยว Problem: -Clustering : Topic modeling -Classification : Sentiment Analysis	Value Propositions What are we trying to do for the end-user(s) of the predictive system? What objectives are we serving? Target audience: disneyland การวัดกลุ่มปัจจัย ว่ามีอะไรบ้างที่มีผลต่อการดึงดูดนักท่องเที่ยว?	Data Sources Which raw data sources can we use (internal and external)? - แบบประเมินความพึงพอใจต่อ Disneyland - Disneyland review database from tripadvisor - ข้อมูลการพูดถึงบน social media (Text mining) - ข้อมูลจาก ES - RFID tag data	Collecting Data How do we get new data to learn from (inputs and outputs)? -เชื่อมต่อกับ database โดยตรง และเก็บ data ทุก ๆ วัน -ดึงข้อมูล review จากแอปพลิเคชัน Tripadvisor -รวบรวมข้อมูลจาก RFID, ES, Text mining
Making Predictions When do we make predictions on new inputs? How long do we have to featureize a new input and make a prediction? ทำการปรับปรุงและพัฒนา Disneyland แต่ละสาขาในทุก ๆ ปีตามแผนธุรกิจ กำหนดให้โมเดลใหม่เข้ามาของ Disneyland ที่ดึงดูดความสนใจของลูกค้านักท่องเที่ยว ทั่วโลก ที่ได้อ้างอิงจากข้อมูล ที่ได้จากการทำ model การวัดกลุ่มปัจจัย	Offline Evaluation Methods and metrics to evaluate the system before deployment. Clustering problem: Silhouette Coefficient, Dunn's Index Classification problem: Accuracy, Recall, Precision, F1 Score, Confusion Matrix, Area Under ROC Curve (AUC)	Objectives: -แยกประเภทระหว่างปัจจัยที่ส่งผลให้ผู้ใช้บริการรู้สึกดี (positive) และปัจจัยที่ทำให้ผู้ใช้บริการรู้สึกแย่ (negative) เพื่อนำไปวางแผนปรับปรุงสถานที่ สิ่งอำนวยความสะดวก จัดสรรหรือสร้างโปรแกรม รวมถึงปัจจัยอื่น ๆ ให้สอดคล้องกับความต้องการของผู้เข้าชม	Features Input representations extracted from raw data sources. - Review Text ของลูกค้า - Rating ที่ลูกค้าประเมินโดยนักเดินทางประเมิน (1-5 (ไม่พึงพอใจ) ถึง 5 (พึงพอใจมาก)) - ข้อมูลของลูกค้า disneyland	Building Models When do we create/update models with new training data? How long do we have to featureize training inputs and create a model? เก็บ data ทุก ๆ วัน และนำมาวิเคราะห์เป็นรายปี โดยใช้ data จากปีก่อนหน้าเพื่อเอามาพัฒนา Model ไปด้วยทุก ๆ ปี
Live Evaluation and Monitoring Methods and metrics to evaluate the system after deployment, and to quantify value creation.		ติดตามการเปลี่ยนแปลงของจำนวนนักท่องเที่ยว ติดตาม Review และ Rating ของผู้เข้าชม		

machinelearningcanvas.com by Louis Dorard, Ph.D. Licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

- Machine Learning

สำหรับการทำ Model จะเก็บ Data ทุก ๆ วันและนำมาวิเคราะห์ทุกไตรมาสเพื่อทำแผนในการพัฒนา Model เป็นรายปี การวิเคราะห์ทุกไตรมาสเป็นการทำเพื่อนำไปวางแผนปรับปรุงสถานที่ สิ่งอำนวยความสะดวก จัดสรรหรือสร้างโปรโมชั่น รวมถึงปัจจัยอื่น ๆ ให้สอดคล้องกับความต้องการของผู้เข้าชม

สำหรับ Model ที่ใช้จะเป็นการใช้ร่วมกันระหว่าง โดยใช้ Topic model Clustering มาสกัดเพื่อแยกว่าประโยคหรือความคิดเห็นที่นักท่องเที่ยวพูดถึงเกี่ยวข้องกับเรื่องอะไรเป็นส่วนไหนก่อน จากนั้นจะใช้ Sentiment analysis Classification เพื่อแยกประเภทความคิดเห็นนั้น ๆ เป็น แง่บวก(Positive) แง่ลบ(Negative) หรือ กลางๆ(Nature)

5. Presentation

Descriptive analytics question

1. การเปลี่ยนแปลงของค่าเฉลี่ย Rating ในแต่ละปีของแต่ละสาขาเป็นอย่างไร?

Target audience:

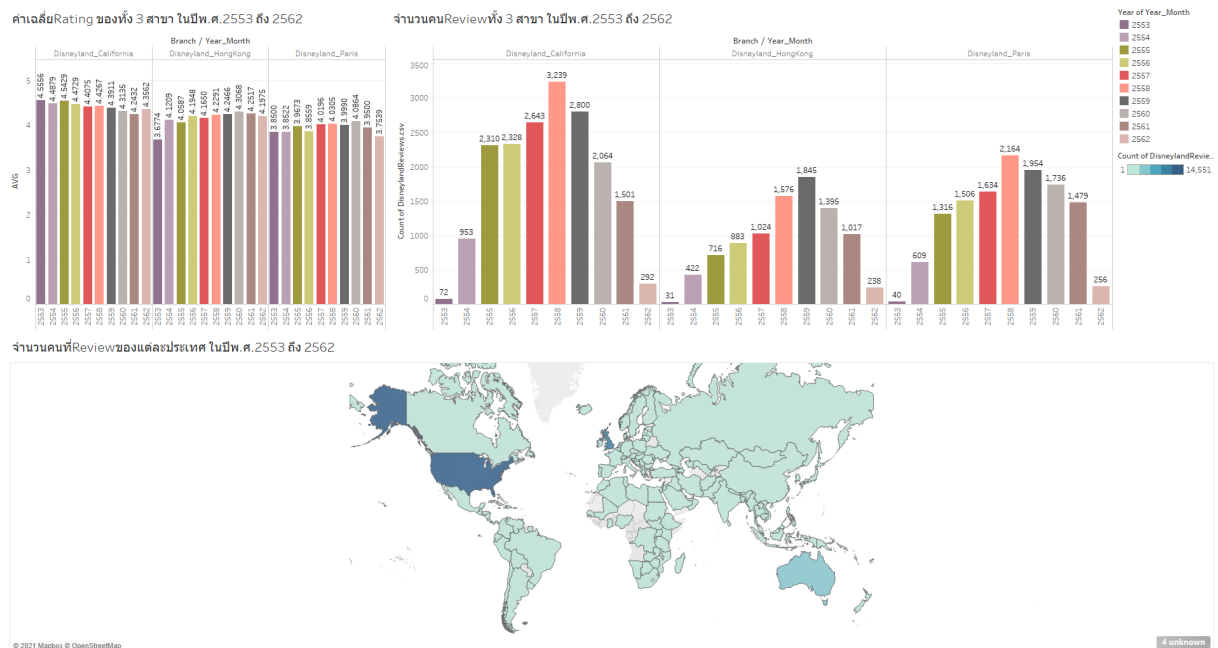
- Disneyland

Objective:

- ดูการเปลี่ยนแปลงของค่าเฉลี่ยระดับความพึงพอใจของผู้เข้าชม Disneyland ในแต่ละปี เพื่อเป็นการติดตามผลลัพธ์จากการปรับปรุงและพัฒนาของ Disneyland แต่ละสาขา

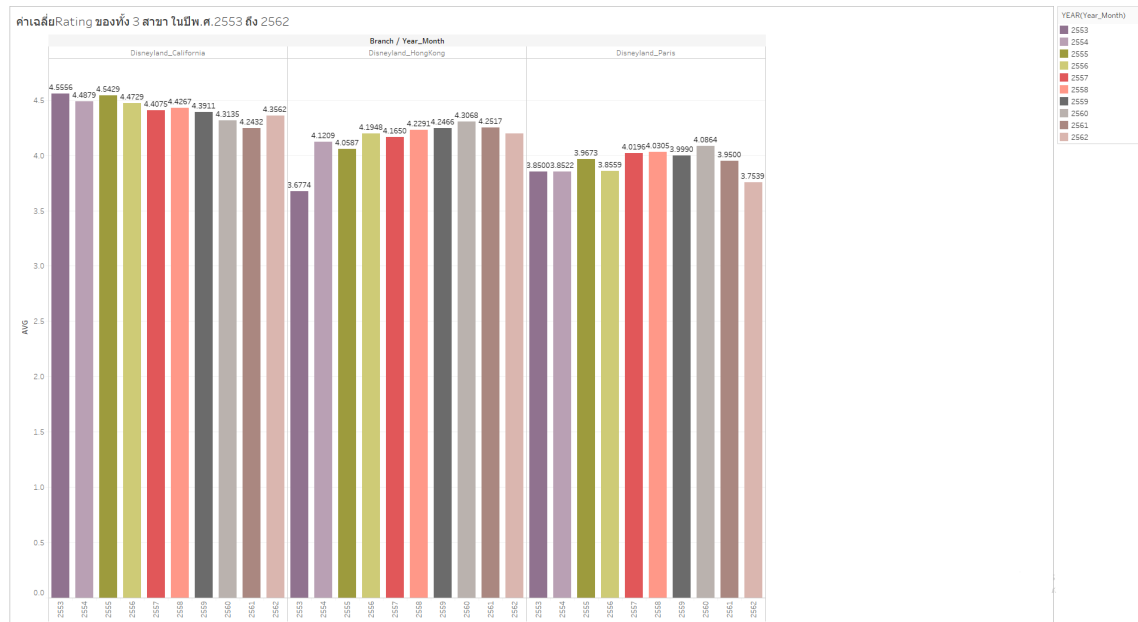
- Visualization

จากการนำชุดข้อมูลคำวิจารณ์ข้างต้นมาทำ visualization ด้วยโปรแกรม Tableau ได้ผลลัพธ์ออกมาเป็น Dashboard ดังนี้

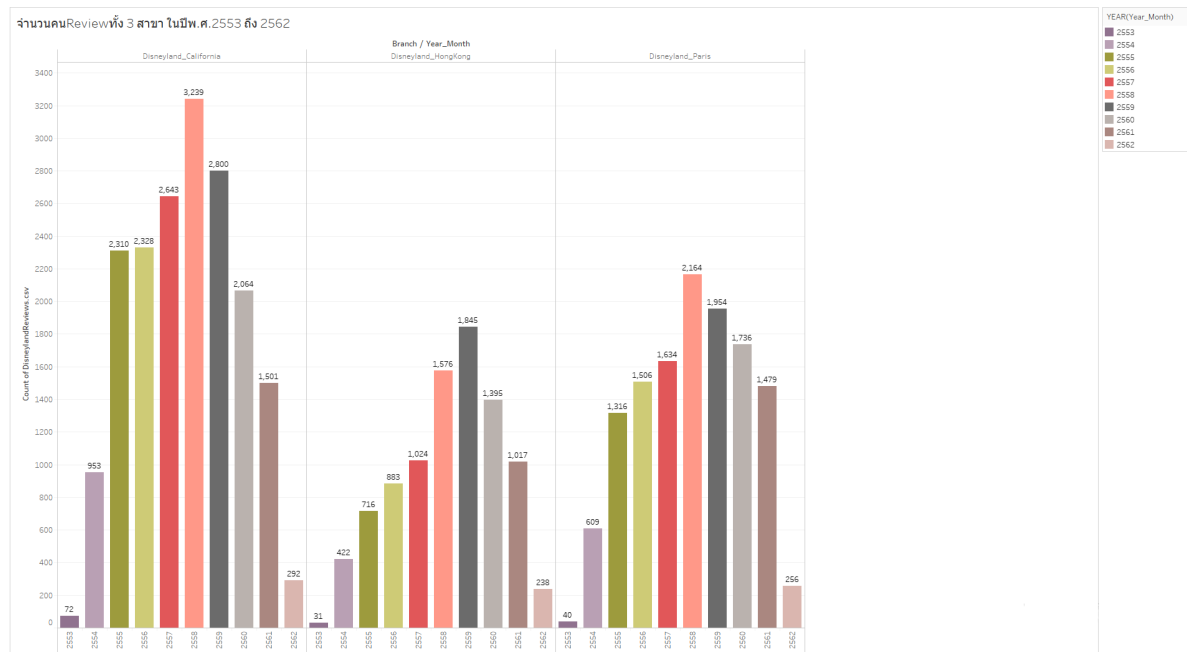


ซึ่งDashboard จะประกอบไปด้วย 3 ส่วน ได้แก่

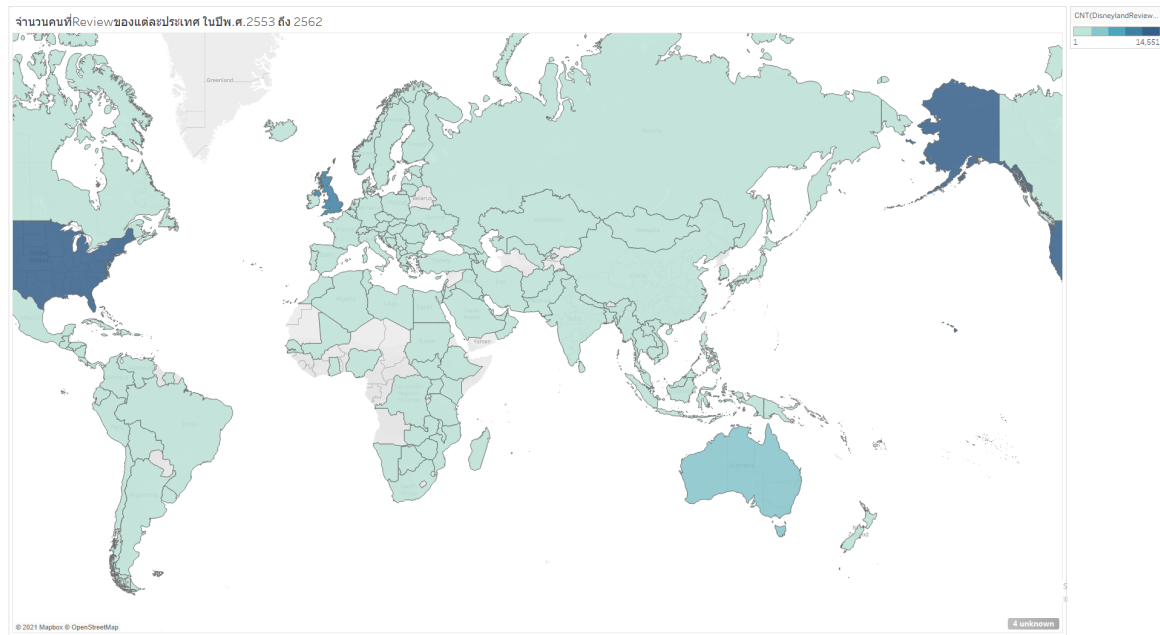
ส่วนที่ 1: กราฟแท่งแสดงค่าเฉลี่ย Rating ของทั้ง 3 สาขา ในปีพ.ศ.2553 ถึง 2562



ส่วนที่ 2: กราฟแท่งแสดงจำนวนคน Review ทั้ง 3 สาขา ในปีพ.ศ.2553 ถึง 2562

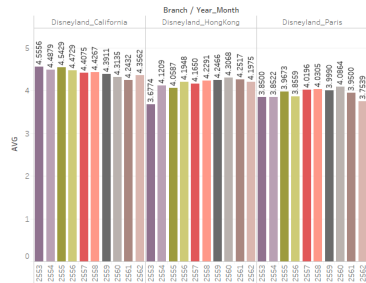


ส่วนที่ 3: แผนที่จำนวนคนที่ Review ของแต่ละประเทศ ในปีพ.ศ.2553 ถึง 2562

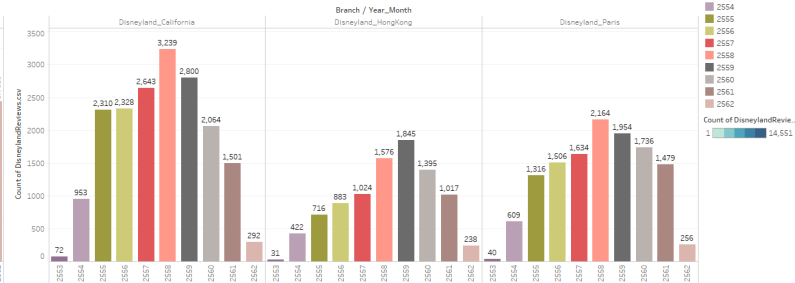


Dashboard:

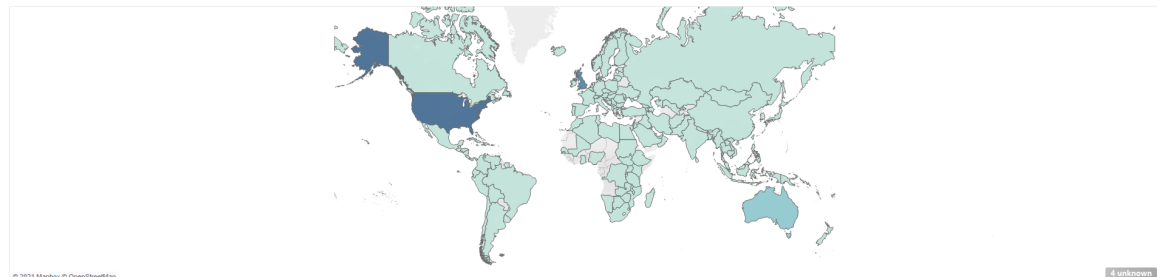
ค่าเฉลี่ยRating ของทั้ง 3 สาขา ในปีพ.ศ.2553 ถึง 2562



จำนวนคนReviewทั้ง 3 สาขา ในปีพ.ศ.2553 ถึง 2562



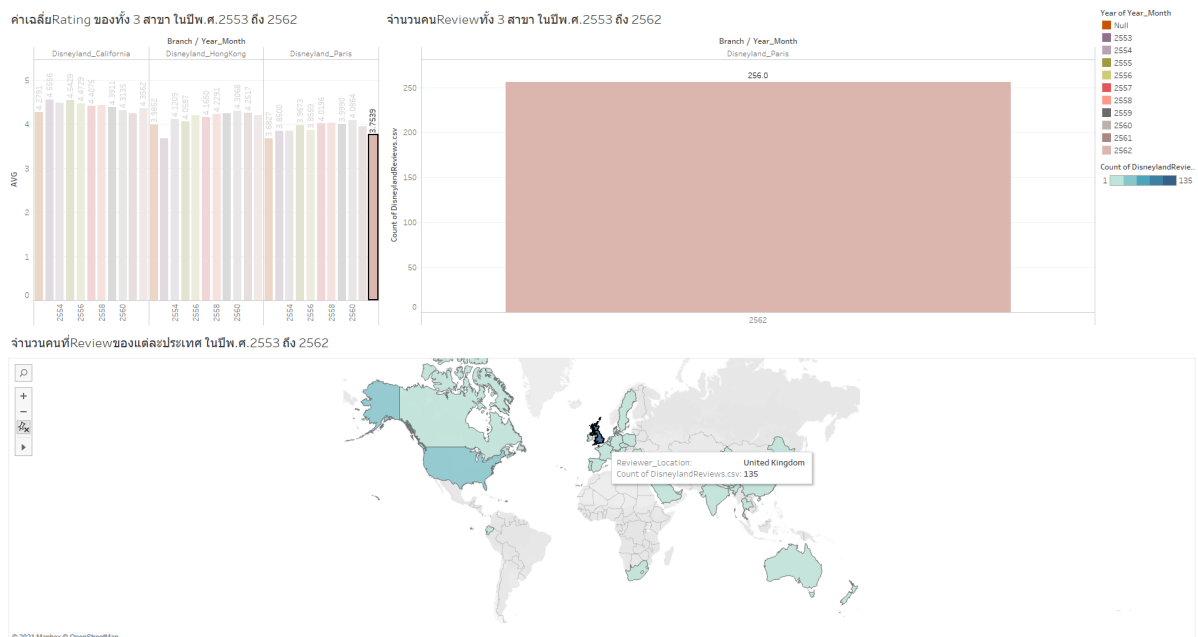
จำนวนคนที่Reviewของแต่ละประเทศ ในปีพ.ศ.2553 ถึง 2562



- ถ้าต้องการดูรายละเอียดค่าเฉลี่ย Rating ของ Disneyland สาขา Paris ใน พ.ศ. 2562 และอยากรู้ว่ามีประเทศไหนที่มี Review มากที่สุด

ให้ไปกดที่ ส่วนที่1 Disneyland สาขา Paris ปีพ.ศ. 2562

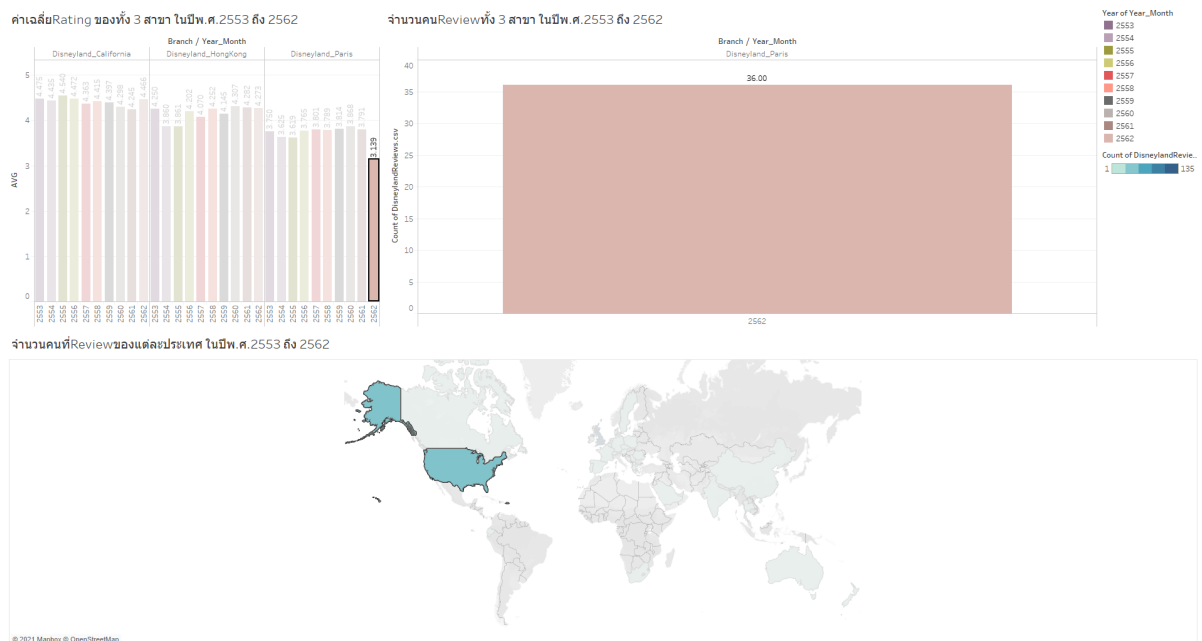
แล้วจะเห็นว่า ส่วนที่1 ใน Disneyland สาขา Paris ปีพ.ศ. 2562 โดยมีค่าเฉลี่ย Rating คือ 3.7539 ส่วนที่2 จะแสดงจำนวนของคนที่มี Review Disneyland สาขา Paris ปีพ.ศ. 2562 มาเป็นจำนวน 256 คน และส่วนที่3 ก็จะแสดงจำนวนคนที่ Review ของแต่ละประเทศ ของ Disneyland สาขา Paris ปีพ.ศ. 2562 โดยจะเห็นว่าที่ United Kingdom มีสีเข้มที่สุด นั่นแปลว่ามีจำนวนคนที่ Review ของแต่ละประเทศ ของ Disneyland สาขา Paris ปีพ.ศ. 2562 มากที่สุด



- ต่อมาถ้าต้องการทราบว่า ค่าเฉลี่ย Rating ของ Disneyland สาขา Paris ในปีพ.ศ. 2562 ที่ถูก Review โดยคนใน United States เท่านั้น มีค่าเท่าไร และมีจำนวนกี่คน?

ให้ไปที่ ส่วนที่1 แล้วก็กดไปที่ Disneyland สาขา Paris ปีพ.ศ. 2562 และ กด ส่วนที่2 ให้เลือกกดไปที่ United States

แล้วจะเห็นว่า ส่วนที่1 จะแสดงค่าเฉลี่ย Rating ของ Disneyland สาขา Paris ในปีพ.ศ. 2562 ที่ถูก Review โดยคนใน United States เท่านั้น โดยมีค่าเฉลี่ย Rating คือ 3.139 ส่วนที่2 จะแสดงจำนวนคนที่ Review ใน United States เป็นจำนวน 36 คน



- จากการทำ Visualization โดยใช้ Dashboard จะเห็นว่าสามารถช่วยให้เห็นภาพรวมของข้อมูล ทำให้เข้าใจข้อมูลได้ง่าย การเลือกใช้สีที่เหมาะสมจะช่วยลดความสับสนและสามารถสื่อสารข้อมูลให้เข้าใจได้ตรงกัน อีกทั้งยังนำไปใช้ประกอบการตัดสินใจเพื่อวางแผนในการพัฒนา Disneyland

6. Summary

การนำ 4 BI Capabilities มาใช้งานในDisneyland จะได้ดังนี้

- **Organizational memory** เลือกใช้ Data lake (DL) เป็นฐานข้อมูลขององค์กร สำหรับจัดเก็บข้อมูลของนักท่องเที่ยวก่อน แล้วนำข้อมูลไปใส่ใน Data warehouses เพื่อให้สะดวกในการทำตามวัตถุประสงค์ และใช้ Enterprise Resource Planning (ERP) ในการวางแผนธุรกิจเพื่อให้ Data flow ในองค์กรมีความเป็นระบบ นอกจากนั้นยังเลือกใช้ Knowledge repository เป็นระบบในการรวมข้อมูลวิธีการจัดการปัญหาที่เกิดขึ้นภายใน Disneyland
- **Information integration** เลือกใช้ Environmental scanning (ES) รวบรวมข้อมูล ตรวจสอบปัจจัยจากทั้งภายในและภายนอก Disneyland และใช้ Future wheel (FW) เพื่อจัดลำดับความสัมพันธ์ของแต่ละปัจจัย และใช้ Radio Frequency Identification (RFID) เป็น tag เพื่อเก็บลักษณะการเข้าชมบุรเพื่อนำมาปรับปรุงการจัดวางตำแหน่งของบุร นอกจากนั้นยังเลือก Text mining มาวิเคราะห์ความคิดเห็นที่นักท่องเที่ยวพูดถึง Disneyland ใน Social media
- **Insight creation** เลือกใช้ Machine Learning โดยทำ Model ออกมาเป็นแบบ Regression เพื่อใช้ในการทำนาย ทำ Model ออกมาเป็นแบบ Clustering เพื่อจัดกลุ่มความคิดเห็นที่คล้ายๆกัน และใช้ Model แบบ Classification แยกว่าแต่ละกลุ่มเป็นความคิดเห็นประเภทใด
- **Presentation** จากการนำชุดข้อมูลคำวิจารณ์ข้างต้นมาทำ visualization ด้วยโปรแกรม Tableau ได้ผลลัพธ์ออกมาเป็น Dashboard ซึ่งประกอบด้วย
ส่วนที่ 1: กราฟแท่งแสดงค่าเฉลี่ย Rating ของทั้ง 3 สาขา ในปีพ.ศ.2553 ถึง 2562
ส่วนที่ 2: กราฟแท่งแสดงจำนวนคน Review ทั้ง 3 สาขา ในปีพ.ศ.2553 ถึง 2562
ส่วนที่ 3: แผนที่จำนวนคนที่ Review ของแต่ละประเทศ ในปีพ.ศ.2553 ถึง 2562

ได้ข้อสรุปว่า ค่าเฉลี่ยของระดับความพึงพอใจในภาพรวมอยู่ในช่วง พอใจระดับปานกลาง (3) ถึงพอใจมาก(4)

จากการนำ 4 BI Capabilities มาใช้งานใน Disneyland จะทำให้ทราบเป้าหมายที่ชัดเจนของ Disneyland ซึ่งช่วยในการวางแผนการต่าง ๆ ให้มีประสิทธิภาพมากขึ้น และช่วยในการประกอบการตัดสินใจให้ได้ผลลัพธ์ที่มีประโยชน์สูงสุดต่อ Disneyland

7. Reference

1. Arush Chillar. Disneyland Reviews [อินเทอร์เน็ต]. 2564 [เข้าถึงเมื่อ 4 พ.ค. 2564]. เข้าถึงได้จาก:
<https://www.kaggle.com/arushchillar/disneyland-reviews>
2. bhimantoro suryo admodjo. Disneyland SA - Random Forest [อินเทอร์เน็ต]. 2564 [เข้าถึงเมื่อ 4 พ.ค. 2564]. เข้าถึงได้จาก:
<https://www.kaggle.com/bhimantoros/disneyland-sa-random-forest>
3. Ahmed Ashour. Disneyland Reviews NLP +Sentiment Analysis [อินเทอร์เน็ต]. 2564 [เข้าถึงเมื่อ 4 พ.ค. 2564]. เข้าถึงได้จาก:
<https://www.kaggle.com/ahmedterry/disneyland-reviews-nlp-sentiment-analysis>
4. Siritwat Limwattana. สร้าง Topic Modeling ง่ายๆ ด้วย LSA [อินเทอร์เน็ต]. 2562 [เข้าถึงเมื่อ 4 พ.ค. 2564]. เข้าถึงได้จาก:
<http://bigdataexperience.org/topic-modeling-with-lsa/>
5. Songhao Wu. 3 Best metrics to evaluate Regression Model? [อินเทอร์เน็ต]. 2562 [เข้าถึงเมื่อ 4 พ.ค. 2564]. เข้าถึงได้จาก:
[3 Best metrics to evaluate Regression Model? | by Songhao Wu | Towards Data Science](#)
6. Kasidis Satangmongkol. อธิบาย 10 Metrics พื้นฐานสำหรับวัดผลโมเดล Machine Learning [อินเทอร์เน็ต]. 2562 [เข้าถึงเมื่อ 4 พ.ค. 2564]. เข้าถึงได้จาก:
[อธิบาย 10 Metrics พื้นฐานสำหรับวัดผลโมเดล Machine Learning – DataRockie](#)
7. MANN. Quick Guide to Evaluation Metrics for Supervised and Unsupervised Machine Learning [อินเทอร์เน็ต]. 2562 [เข้าถึงเมื่อ 4 พ.ค. 2564]. เข้าถึงได้จาก: [Evaluation Metrics For Machine Learning For Data Scientists \(analyticsvidhya.com\)](#)