

บทที่ 3

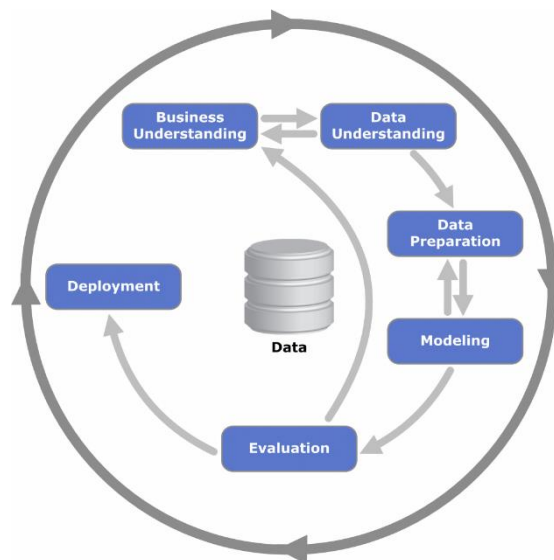
วิธีการดำเนินงานโครงการ

การพัฒนาโครงการจำเป็นต้องมีการจัดเก็บ รวบรวมข้อมูลจากแหล่งข้อมูลที่หลากหลาย เพื่อวิเคราะห์ข้อแตกต่างและสถิติการใช้บริการสายการบินต้นทุนต่ำระหว่างในประเทศและต่างประเทศ โดยมีการวิเคราะห์จากกรเก็บรวบรวมข้อมูลมาเปรียบเทียบการใช้บริการในแต่ละปี ซึ่งวิเคราะห์จากเที่ยวบินของแต่ละสนามบินระหว่างในประเทศและต่างประเทศ ผู้จัดทำจึงได้ทำการวิเคราะห์และออกแบบการแสดงผลสถิติ และการแสดงกราฟเปรียบเทียบ โดยใช้เครื่องมือในการออกแบบกราฟและสถิติ ผู้จัดทำจึงได้ทำการวิเคราะห์และเพื่อที่จะสามารถนำไปเป็นแนวทางในการออกแบบระบบงานใหม่โดยใช้โปรแกรม Tableau Public เพื่อนำมาแสดงข้อมูลในรูปแบบกราฟและสถิติ

3.1 กระบวนการวิเคราะห์ข้อมูลด้วย CRISP-DM

กระบวนการวิเคราะห์ข้อมูลมาตรฐานซึ่งเป็นเหมือน blueprint ที่ใช้กันอย่างกว้างขวาง เช่นเดียวกับกับกระบวนการ ISO ในโรงงานอุตสาหกรรม หรือกระบวนการ CMMI ซึ่งเป็นมาตรฐานในการพัฒนาซอฟต์แวร์ กระบวนการมาตรฐานในการวิเคราะห์ข้อมูลด้านดาต้าไมน์นิ่งนี้พัฒนาขึ้นในปี ค.ศ. 1996 โดยความร่วมมือกันของ 3 บริษัท คือ DaimlerChrysler SPSS และ NCR กระบวนการทำงานนี้เรียกว่า “Cross-Industry Standard Process for Data Mining” หรือเรียกย่อว่า “CRISP-DM”

โดยในกระบวนการ CRISP-DM นี้จะประกอบด้วย 6 ขั้นตอน แต่ละขั้นตอนในรูปแบบจะเป็นขั้นตอนที่ต่อเนื่องกันนั่นคือขั้นตอนถัดไปจะรอผลลัพธ์จากขั้นตอนก่อนหน้าซึ่งแสดงด้วยลูกศรที่เชื่อมระหว่างกล่องสี่เหลี่ยมแต่ละกล่องตัวอย่างเช่นเมื่อได้ผลลัพธ์จากขั้นตอนการเตรียมข้อมูล (Data Preparation) แล้วจะนำไปสร้างโมเดลจำแนกประเภทข้อมูลในขั้น Modeling และหลังจากนั้นอาจจะย้อนกลับมาเปลี่ยนแปลงข้อมูลให้ถูกต้องมากขึ้นเพื่อหวังว่าจะโมเดลที่ให้ความถูกต้องมากขึ้นก็ได้ เป็นต้น



ภาพที่ 3.1 กระบวนการ CRISP-DM 6 ขั้นตอน
ที่มา : dataminingtrend.com (2562)

(1) Business Understanding เป็นขั้นตอนแรกในกระบวนการ CRISP-DM ซึ่งเน้นไปที่การเข้าใจปัญหาและแปลงปัญหาที่ได้ให้อยู่ในรูปโจทย์ของการวิเคราะห์ข้อมูลทางดาต้าไมน์นิ่ง

(2) Data Understanding ขั้นตอนนี้เริ่มจากการเก็บรวบรวมข้อมูล หลังจากนั้นจะเป็นการตรวจสอบข้อมูลที่ได้ทำการรวบรวมมาได้เพื่อดูความถูกต้องของข้อมูล

(3) Data Preparation ขั้นตอนนี้เป็นการแปลงข้อมูลที่ได้ทำการเก็บรวบรวมมา (raw data) ให้กลายเป็นข้อมูลที่สามารถนำไปวิเคราะห์ในขั้นถัดไปได้ โดยขั้นตอนนี้จะเป็นขั้นตอนที่ใช้เวลามากที่สุดของกระบวนการ CRISP-DM

(4) Modeling ขั้นตอนนี้จะเป็นขั้นตอนการวิเคราะห์ข้อมูลด้วยเทคนิคทางดาต้า ไมน์นิ่ง ซึ่งในขั้นตอนนี้หลายเทคนิคจะถูกนำมาใช้เพื่อให้ได้คำตอบที่ดีที่สุด ดังนั้นในบางครั้งอาจจะต้องมีการย้อนกลับไปขั้นตอนที่ (3) Data Preparation เพื่อแปลงข้อมูลบางส่วนให้เหมาะสมกับแต่ละเทคนิค

(5) Evaluation ในขั้นตอนนี้เราจะได้ผลการวิเคราะห์ข้อมูลด้วยเทคนิคทางดาต้า ไมน์นิ่งแล้วแต่ก่อนที่จะนำผลลัพธ์ที่ได้ไปใช้งานต่อไปก็จะต้องมีการวัดประสิทธิภาพของผลลัพธ์ที่ได้ว่าตรงกับวัตถุประสงค์ที่ได้ตั้งไว้ในขั้นตอนแรก หรือมีความน่าเชื่อถือมากน้อยเพียงใด ซึ่งอาจจะย้อนกลับไปยังขั้นตอนก่อนหน้าเพื่อเปลี่ยนแปลงแก้ไขเพื่อให้ได้ผลลัพธ์ตามที่ต้องการได้

(6) Deployment ในกระบวนการทำงานของ CRISP-DM นั้นไม่ได้หยุดเพียงแค่ผลลัพธ์ที่ได้จากการวิเคราะห์ข้อมูลด้วยเทคนิคทางดาต้า ไม่นิ่งเท่านั้น แม้ว่าผลลัพธ์ที่ได้จะแสดงถึงองค์ความรู้ที่มีประโยชน์ แต่จะต้องนำองค์ความรู้ที่ได้เหล่านั้นไปใช้ได้จริงในองค์กร

3.1.1 Business Understanding ผู้วิเคราะห์ข้อมูลทำความเข้าใจปัญหาและ แปลงปัญหาที่ได้ให้อยู่ในรูปโจทย์ของการวิเคราะห์ข้อมูลทางดาต้าไม่ว่าหนึ่งจากนั้นวางแผนในการดำเนินการโดยการวิเคราะห์ข้อมูลรูปแบบความสัมพันธ์ที่ซับซ้อนอยู่ในข้อมูลของข้อมูลการใช้บริการสายการบินต้นทุนต่ำ

3.1.2 Data Understanding ผู้วิเคราะห์ข้อมูลเริ่มทำการเก็บรวบรวมข้อมูลและตรวจสอบข้อมูลที่ได้ดูความถูกต้องและความน่าเชื่อถือของข้อมูลและพิจารณาว่าจะใช้ข้อมูลทั้งหมดหรือจำเป็นต้องเลือกข้อมูลบางส่วนมาใช้ในการวิเคราะห์จากเว็บไซต์ <http://aot-th.listedcompany.com> ซึ่งเป็นเว็บไซต์ที่เปิดเผย (Open Data) ข้อมูลในเชิงศึกษาให้แก่ประชาชน ดังนั้นผู้วิเคราะห์ข้อมูลจึงได้ดำเนินการตามขั้นตอนดังต่อไปนี้

1) ผู้วิเคราะห์ได้ทำการรวบรวมเก็บข้อมูลโดยเลือกข้อมูลสายการบินต้นทุนต่ำของแต่ละท่าอากาศยาน จากเว็บไซต์ <http://aot-th.listedcompany.com> และได้เลือกข้อมูลปริมาณการจราจรทางอากาศของสายการบินต้นทุนต่ำ โดยผู้วิเคราะห์ได้เลือกข้อมูลมาจำนวน 6 ปี เพื่อนำมาวิเคราะห์ข้อมูล

ปีงบประมาณ	ปริมาณการจราจรทางอากาศรวม	ปริมาณการจราจรทางอากาศของสายการบินต้นทุนต่ำ
2563 (กรกฎาคม)	✖	✖
2562	✖	✖
2561	✖	✖
2560	✖	✖
2559	✖	✖
2558	✖	✖
2557	✖	✖
2556	✖	✖

ภาพที่ 3.2 ข้อมูลปริมาณการจราจรทางอากาศของสายการบินต้นทุนต่ำ ที่เลือกใช้ในการวิเคราะห์ จากเว็บไซต์ <http://aot-th.listedcompany.com>

2) ตรวจสอบข้อมูลว่าข้อมูลพร้อมที่จะทำการวิเคราะห์ข้อมูลและนำเชื่อถือหรือไม่

Table 1 LCC INTERNATIONAL AIRCRAFT MOVEMENT							
MONTH	FY 2013			FY 2014			Change
	Arrival	Departure	Total	Arrival	Departure	Total	
OCT.	2,269	2,268	4,537	3,223	3,227	6,450	36.2%
NOV.	2,404	2,401	4,805	3,251	3,245	6,497	35.2%
DEC.	2,567	2,570	5,137	3,473	3,463	6,942	35.1%
OCT - DEC.	7,239	7,239	14,478	9,947	9,942	19,889	36.5%
JAN.	2,551	2,709	5,260	3,856	3,970	7,826	35.9%
FEB.	2,596	2,594	5,190	3,226	3,222	6,448	24.2%
MAR.	2,906	2,909	5,815	3,334	3,337	6,671	18.8%
JAN - MAR.	8,053	8,100	16,153	10,416	10,536	20,952	29.2%
APR.	2,823	2,825	5,648	3,373	3,380	6,753	19.6%
MAY.	2,896	2,898	5,794	3,295	3,289	6,584	13.6%
JUN.	2,974	2,973	5,947	2,950	2,945	5,895	2.5%
APR - JUN.	8,693	8,696	17,389	9,224	9,614	18,838	10.9%
JUL.	3,111	3,103	6,214	3,016	3,016	6,034	-3.1%
AUG.	3,229	3,227	6,456	3,300	3,098	6,398	-4.2%
SEP.	3,065	3,063	6,128	3,017	3,016	6,033	-1.5%
JUL - SEP.	9,405	9,413	18,818	9,033	9,050	18,083	-5.0%
JAN - SEP.	28,103	28,117	56,220	32,382	32,970	65,352	19.3%
Total	33,445	33,454	66,899	38,530	38,916	77,446	16.3%

Source : Air Transport Information and Slot Coordination Division, ADT.

Table 2 LCC DOMESTIC AIRCRAFT MOVEMENT							
MONTH	FY 2013			FY 2014			Change
	Arrival	Departure	Total	Arrival	Departure	Total	
OCT.	4,837	4,835	9,672	5,640	5,645	11,285	21.0%
NOV.	4,503	4,497	9,000	5,828	5,835	11,664	29.6%
DEC.	5,056	5,081	10,137	6,636	6,638	13,274	30.9%

Table 4 LCC INTERNATIONAL PASSENGER											
MONTH	FY 2013					FY 2014					Change
	OutEmb.	Emb.	Disemb+Emb	Transit	Total	OutEmb.	Emb.	Disemb+Emb	Transit	Total	
OCT.	329,121	341,643	670,764	3	670,767	436,920	438,246	875,166	3	875,169	29.9%
NOV.	366,076	347,091	713,167	595	713,722	458,286	439,363	897,629	131	897,760	25.7%
DEC.	384,850	382,873	767,723	388	778,111	487,727	473,616	961,343	80	961,423	23.5%
OCT - DEC.	1,080,047	1,071,607	2,151,654	986	2,152,640	1,382,933	1,351,225	2,734,138	224	2,734,362	26.9%
JAN.	391,956	407,693	799,649	203	799,852	460,745	456,623	917,371	472	917,843	15.0%
FEB.	376,743	387,690	764,433	189	764,623	405,207	428,323	833,530	250	833,780	9.0%
MAR.	428,919	432,761	861,680	228	861,908	452,708	463,835	916,543	380	916,923	6.5%
JAN - MAR.	1,197,618	1,228,144	2,425,772	619	2,426,391	1,369,660	1,381,281	2,731,444	1,302	2,732,746	12.5%
APR.	412,956	436,323	849,279	60	849,339	466,599	461,691	928,290	3	928,293	12.0%
MAY.	407,001	404,363	811,364	-	811,364	428,927	430,283	859,210	2	859,212	5.8%
JUN.	427,586	419,797	847,383	-	847,383	391,985	383,870	775,855	116	776,071	-8.3%
APR - JUN.	1,247,143	1,239,283	2,486,426	20	2,486,446	1,279,511	1,275,844	2,553,355	121	2,553,476	11.0%
JUL.	444,176	446,303	890,479	1	890,480	447,294	423,343	870,637	157	870,794	-1.0%
AUG.	445,375	457,135	902,510	328	902,838	457,259	463,314	920,573	583	921,156	1.8%
SEP.	420,905	419,958	840,863	243	841,106	441,621	432,899	874,520	449	874,969	4.0%
JUL - SEP.	1,310,456	1,323,496	2,633,952	572	2,634,524	1,346,174	1,319,556	2,765,756	1,189	2,766,945	11.5%
JAN - SEP.	3,769,101	3,753,233	7,522,334	1,240	7,523,574	3,933,256	3,894,418	7,827,774	2,403	7,830,177	4.8%
Total	4,845,754	4,881,555	9,727,309	2,188	9,729,497	5,038,288	4,936,641	10,974,929	2,823	10,977,752	9.7%

Source : Air Transport Information and Slot Coordination Division, ADT.

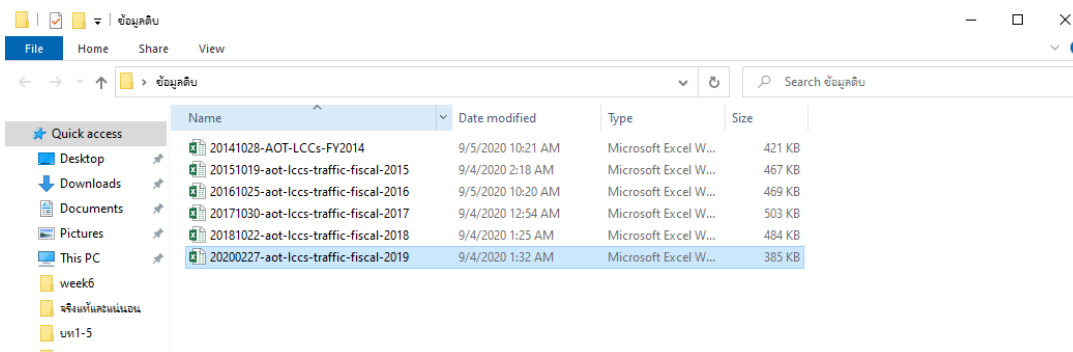
Table 5 LCC DOMESTIC PASSENGER											
MONTH	FY 2013					FY 2014					Change
	OutEmb.	Emb.	Disemb+Emb	Transit	Total	OutEmb.	Emb.	Disemb+Emb	Transit	Total	
OCT.	644,259	647,392	1,291,651	122	1,291,773	819,211	819,841	1,639,052	853	1,639,905	26.9%
NOV.	639,796	634,744	1,274,540	-	1,274,540	812,674	811,653	1,624,327	272	1,624,599	27.5%
DEC.	704,679	727,611	1,432,290	-	1,432,290	889,937	925,555	1,815,492	1,233	1,816,725	28.5%

รูปที่ 3.3 ข้อมูลปริมาณการจราจรทางอากาศของสายการบินต้นทุนต่ำ

ขั้นตอนที่ทำการแปลงข้อมูลที่ได้ทำการเก็บรวบรวมมา (raw data) ให้กลายเป็นข้อมูลที่
สามารถนำไปวิเคราะห์ในขั้นถัดไปได้ โดยการแปลงข้อมูลนี้อาจจะต้องมีการทำข้อมูลให้ถูกต้อง
(data cleaning) ผู้วิเคราะห์จะทำการเลือกข้อมูลที่เหมาะสมเพื่อให้ลดปริมาณข้อมูลที่ไม่จำเป็น
ออกไป เพื่อลดความผิดพลาดให้น้อยที่สุด โดยขั้นตอนนี้จะใช้เวลามากที่สุด

3.1.3 Data Preparation ผู้วิเคราะห์ข้อมูลทำการแปลงข้อมูลที่ได้ทำการเก็บรวบรวมมาให้
กลายเป็นข้อมูลที่นำไปวิเคราะห์ในขั้นถัดไปได้ โดยการทำให้ข้อมูลให้ถูกต้องโดยใช้
กระบวนการ data cleaning เพื่อให้ได้ข้อมูลอยู่ในรูปแบบและมาตรฐานเดียวกันและกำจัดข้อมูล
เสียออก โดยมีขั้นตอนดังต่อไปนี้

1) การปรับมาตรฐาน (Standardizing) ผู้วิเคราะห์ได้ทำการดาวน์โหลดข้อมูลปริมาณ
การจราจรทางอากาศของสายการบินต้นทุนต่ำเพื่อนำมาวิเคราะห์ข้อมูลและผู้วิเคราะห์ได้เลือก
ข้อมูลมาจำนวน 6 ปี โดยเลือกปี 2556-2561 และทำการแปลงไฟล์ ให้อยู่ในรูปแบบเดียวกัน
พร้อมสร้างแอตทริบิวต์เพื่อกำหนดคีย์หลักของของแต่ละไฟล์และระบุความสัมพันธ์ของข้อมูลที่
แสดง โดยในที่นี้ผู้จัดทำได้ยกตัวอย่างท่าอากาศยานเชียงใหม่



ภาพที่ 3.4 ไฟล์ข้อมูลดิบปริมาณการจราจรทางอากาศของสายการบินต้นทุนต่ำ

	A	B	C	D	E	F	G	H
1	เดือน	ชื่อสายการบิน	จำนวนเที่ยวบินขาเข้าในประเทศ	จำนวนเที่ยวบินขาออกในประเทศ	จำนวนเที่ยวบินขาเข้าระหว่างประเทศ	จำนวนเที่ยวบินขาออกระหว่างประเทศ	จำนวนเที่ยวบินขาเข้าทั้งหมด	จำนวนเที่ยวบินขาออกทั้งหมด
2	มกราคม	สายการบินเอมิเรตส์ เอ็มิเรตส์	96,160	100,545	5,077	5,792	1	1
3	กุมภาพันธ์	สายการบินเอมิเรตส์ เอ็มิเรตส์	88,895	89,375	9,645	8,824	1	1
4	มีนาคม	สายการบินเอมิเรตส์ เอ็มิเรตส์	88,506	88,794	9,662	8,416	-	-
5	เมษายน	สายการบินเอมิเรตส์ เอ็มิเรตส์	77,830	78,859	10,138	9,536	-	-
6	พฤษภาคม	สายการบินเอมิเรตส์ เอ็มิเรตส์	78,758	79,805	12,902	13,074	-	-
7	มิถุนายน	สายการบินเอมิเรตส์ เอ็มิเรตส์	74,484	76,801	10,817	10,448	-	-
8	กรกฎาคม	สายการบินเอมิเรตส์ เอ็มิเรตส์	79,272	79,088	9,752	8,660	-	-
9	สิงหาคม	สายการบินเอมิเรตส์ เอ็มิเรตส์	80,484	79,814	8,513	9,157	-	-
10	กันยายน	สายการบินเอมิเรตส์ เอ็มิเรตส์	73,647	73,053	6,781	6,608	-	-
11	ตุลาคม	สายการบินเอมิเรตส์ เอ็มิเรตส์	93,033	92,835	4,120	3,457	2	2
12	พฤศจิกายน	สายการบินเอมิเรตส์ เอ็มิเรตส์	94,556	92,602	4,225	3,853	1	1
13	ธันวาคม	สายการบินเอมิเรตส์ เอ็มิเรตส์	97,494	96,430	4,581	4,639	-	-
14	มกราคม	สายการบินยูไนเต็ด	20,233	204,935	226,472	225,969	-	-
15	กุมภาพันธ์	สายการบินยูไนเต็ด	201,395	194,336	206,544	216,345	-	-
16	มีนาคม	สายการบินยูไนเต็ด	206,055	205,251	203,776	206,293	-	-
17	เมษายน	สายการบินยูไนเต็ด	202,900	196,801	188,565	184,031	-	-
18	พฤษภาคม	สายการบินยูไนเต็ด	177,906	187,175	165,781	166,576	-	-
19	มิถุนายน	สายการบินยูไนเต็ด	177,015	173,887	183,939	177,761	-	-
20	กรกฎาคม	สายการบินยูไนเต็ด	178,676	173,526	209,465	202,778	-	-
21	สิงหาคม	สายการบินยูไนเต็ด	190,615	191,144	209,139	208,517	-	-
22	กันยายน	สายการบินยูไนเต็ด	155,164	150,879	179,796	174,381	-	-
23	ตุลาคม	สายการบินยูไนเต็ด	192,989	184,252	146,104	151,885	-	-
24	พฤศจิกายน	สายการบินยูไนเต็ด	194,407	179,363	161,957	147,871	-	-
25	ธันวาคม	สายการบินยูไนเต็ด	217,053	189,292	208,519	189,239	-	-
26	มกราคม	สายการบินเวียดเจ็ท	129,151	132,700	11,553	11,184	-	-
27	กุมภาพันธ์	สายการบินเวียดเจ็ท	120,072	120,901	11,627	11,240	-	-
28	มีนาคม	สายการบินเวียดเจ็ท	118,414	116,400	11,644	11,151	-	-
รวม			2560(2017)	2561(2018)	2562(2019)	2554		

ภาพที่ 3.5 ข้อมูลที่ปรับมาตรฐาน

2) การแจกแจงข้อมูล (Parsing) ข้อมูลที่ผู้วิเคราะห์ได้ทำการปรับมาตรฐาน

P10	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	
1	report name	Month	year	inbound population	in inbound populati	of international	inbound per of	inbound pac	outgoing package	international	inc	Number of interio	of domestic flight	domestic outb	of international	international outb
2	สายการบิน	มกราคม	2555	194,182	212,874	26,922	26,180	708	1,163	5	19	1,475	1,476	233	234	
3	สายการบิน	กุมภาพันธ์	2555	169,943	184,260	23,572	24,201	712	1,212	9	25	1,321	1,319	209	211	
4	สายการบิน	มีนาคม	2555	155,061	168,796	20,461	22,289	729	1,165	11	18	1,381	1,378	205	204	
5	สายการบิน	เมษายน	2555	153,213	155,097	18,486	17,915	587	912	14	2	1,363	1,365	181	180	
6	สายการบิน	พฤษภาคม	2555	134,367	139,161	16,650	16,257	623	901	14	4	1,237	1,236	186	185	
7	สายการบิน	มิถุนายน	2555	131,459	131,901	17,282	16,326	632	853	9	1	1,171	1,172	185	183	
8	สายการบิน	กรกฎาคม	2555	149,598	156,239	21,196	20,282	580	948	24	3	1,242	1,239	208	209	
9	สายการบิน	สิงหาคม	2555	152,299	166,821	22,611	22,662	595	998	11	3	1,167	1,165	213	212	
10	สายการบิน	กันยายน	2555	136,118	139,939	19,113	17,782	514	908	7	4	1,198	1,201	202	201	
11	สายการบิน	ตุลาคม	2555	163,206	143,522	15,680	15,153	1,020	1,236	8	4	1,192	1,191	170	164	
12	สายการบิน	พฤศจิกายน	2555	158,098	163,661	21,064	18,704	889	1,299	20	15	1,240	1,239	179	182	
13	สายการบิน	ธันวาคม	2555	195,761	188,810	23,656	23,508	803	1,239	6	20	1,429	1,428	205	205	
14	สายการบิน	มกราคม	2556	106,563	116,690	11,012	10,614	29	46	8	0	775	775	76	76	
15	สายการบิน	กุมภาพันธ์	2556	98,040	104,731	12,113	11,200	25	52	5	0	698	698	75	75	
16	สายการบิน	มีนาคม	2556	102,753	106,116	12,897	12,411	80	43	7	0	694	694	118	118	
17	สายการบิน	เมษายน	2556	91,126	91,821	12,812	12,225	22	47	1	0	635	635	114	114	
18	สายการบิน	พฤษภาคม	2556	78,625	78,250	12,984	12,467	36	69	2	0	545	545	121	121	
19	สายการบิน	มิถุนายน	2556	89,508	89,360	13,469	13,028	35	76	4	0	519	519	118	118	
20	สายการบิน	กรกฎาคม	2556	74,645	77,281	14,607	12,971	35	2	0	482	482	127	127		
21	สายการบิน	สิงหาคม	2556	80,825	86,333	15,559	15,682	85	50	6	0	546	546	142	142	
22	สายการบิน	กันยายน	2556	79,105	81,299	12,831	12,108	72	46	5	0	510	510	119	119	
23	สายการบิน	ตุลาคม	2556	80,272	80,312	10,312	10,379	29	4	0	574	573	77	76		
24	สายการบิน	พฤศจิกายน	2556	84,082	84,120	11,409	10,264	22	49	4	0	574	575	75	75	
25	สายการบิน	ธันวาคม	2556	107,751	105,957	11,238	11,031	32	43	3	0	683	712	75	75	
26	สายการบิน	มกราคม	2557	135,070	137,498	26,211	23,852	64	118	4	0	928	928	198	198	
27	สายการบิน	กุมภาพันธ์	2557	116,432	123,722	24,525	26,270	81	154	3	1	812	812	186	187	
28	สายการบิน	มีนาคม	2557	110,000	116,400	26,400	24,400	0	0	0	0	600	600	100	100	

ภาพที่ 3.6 ข้อมูลท่าอากาศยานเชียงใหม่ที่ใช้ในการวิเคราะห์ข้อมูล

3) การแก้ไขข้อมูลที่ผิดพลาด (Correcting) ผู้วิเคราะห์นำข้อมูล Dataset มาทำการคัดเลือกข้อมูลที่เป็นต่อการนำไปวิเคราะห์ข้อมูล โดยตรวจสอบข้อมูลที่ผิดพลาด ไม่ถูกต้อง หรือสูญหาย หรือข้อมูลที่ยังไม่สมบูรณ์และเพิ่มลบแก้ไขข้อมูลให้เหมาะสมแก่การนำไปวิเคราะห์ข้อมูลและตัดแอตทริบิวต์ที่ไม่สำคัญออกไปจากข้อมูลเพื่อให้เหลือแต่เพียงข้อมูลที่เป็นแก่การวิเคราะห์ข้อมูล โดยจากข้อมูลในที่นี้ผู้วิเคราะห์ได้ทำการตัด แอตทริบิวต์ การขนส่งสินค้าทั้งในประเทศและต่างประเทศออกไป โดยจากเดิมนั้นจะมีแอตทริบิวต์ จำนวนประชากรขาเข้าขาออกในประเทศและระหว่างประเทศ จำนวนขนส่งสินค้าขาเข้าขาออกในประเทศและระหว่างประเทศและจำนวนเที่ยวบินขาเข้าขาออกในประเทศและระหว่างประเทศ โดยผู้วิเคราะห์จะให้เหลือแค่ จำนวนประชากรขาเข้าขาออกในประเทศและระหว่างประเทศและจำนวนเที่ยวบินขาเข้าขาออกในประเทศและระหว่างประเทศ

1	A	B	C	D	E	F	G	H	I	J
1	Airport	Date	Number of domestic departures	Number of international arrivals	Number of international	Number of domestic inbound flights	Number of domestic outbound flights	Number of international inbound flights	Number of international outbound flights	
2	cnk	Jan-13	106,563	116,690	11,012	10,614	775	76	76	
3	cnk	Feb-13	98,040	104,731	12,113	11,200	658	658	75	
4	cnk	Mar-13	102,753	105,116	12,897	12,411	694	694	116	
5	cnk	Apr-13	91,126	91,821	12,912	12,225	635	635	114	
6	cnk	May-13	78,625	78,250	12,984	12,467	545	545	121	
7	cnk	Jun-13	69,508	69,360	13,469	13,028	516	516	118	
8	cnk	Jul-13	74,545	77,281	14,067	12,971	482	482	127	
9	cnk	Aug-13	80,825	85,333	15,593	15,662	546	546	142	
10	cnk	Sep-13	79,105	81,299	12,831	12,106	510	510	119	
11	cnk	Oct-13	80,272	81,927	10,312	10,376	574	573	77	
12	cnk	Nov-13	84,082	84,120	11,409	10,284	574	575	75	
13	cnk	Dec-13	107,751	105,957	11,238	11,031	683	712	75	
14	cnk	Jan-14	155,070	137,498	26,211	23,852	928	928	186	
15	cnk	Feb-14	116,432	123,722	24,525	26,270	812	812	187	
16	cnk	Mar-14	130,000	135,450	26,382	26,598	1,013	1,013	206	
17	cnk	Apr-14	131,160	131,202	24,411	24,850	939	939	193	
18	cnk	May-14	121,070	120,633	24,120	23,074	885	885	190	
19	cnk	Jun-14	113,099	113,326	22,984	21,969	751	752	186	
20	cnk	Jul-14	126,436	127,137	27,795	25,070	792	791	197	
21	cnk	Aug-14	132,281	137,289	27,658	27,603	868	868	197	
22	cnk	Sep-14	121,573	122,821	24,836	22,336	798	799	184	
23	cnk	Oct-14	106,119	100,063	13,252	12,190	708	708	126	
24	cnk	Nov-14	110,311	109,278	20,059	18,151	672	672	138	
25	cnk	Dec-14	138,828	134,948	19,459	18,866	903	901	138	
26	cnk	Jan-15	184,819	198,524	43,787	40,851	1,235	1,235	303	
27	cnk	Feb-15	167,645	177,213	41,889	42,102	1,122	1,122	300	
28	cnk	Mar-15	186,993	176,829	37,997	36,840	1,206	1,206	290	
29	cnk	Apr-15	157,838	157,995	40,430	37,974	1,102	1,102	292	
30	cnk	May-15	155,208	154,942	35,965	34,562	1,074	1,074	288	
31	cnk	Jun-15	148,508	151,028	38,412	36,139	1,001	1,001	271	
32	cnk	Jul-15	166,993	166,995	42,123	38,600	1,040	1,040	285	
33	cnk	Aug-15	181,828	171,868	40,638	40,922	1,044	1,044	274	

ภาพที่ 3.7 ข้อมูลท่าอากาศยานใหม่ที่สมบูรณ์ผ่านกระบวนการ Data Cleaning

4) การจัดหมวดหมู่ (Transform) ผู้วิเคราะห์ข้อมูลนำ Dataset มาทำการคัดเลือกและดำเนินการ Data Cleaning มาทำการจัดหมวดหมู่ เพื่อง่ายแก่การนำข้อมูลมาวิเคราะห์ และทราบถึงความสัมพันธ์ข้อมูลในแต่ละแถว

4.1 โหลดข้อมูลแต่ละท่าอากาศยานที่ต้องการจัดกลุ่ม

A	B	C	D	E	F	G	H	I	J
1	Airport	Date	Number of domestic departures	Number of international arrivals	Number of international	Number of domestic inbound flights	Number of domestic outbound flights	Number of international inbound flights	Number of international outbound flights
2	cnx	Jan-13	106,563	116,630	11,012	10,614	775	775	76
3	cnx	Feb-13	98,040	104,731	12,113	11,200	858	858	75
4	cnx	Mar-13	102,753	105,116	12,837	12,411	894	894	118
5	cnx	Apr-13	91,126	91,821	12,812	12,225	635	635	114
6	cnx	May-13	78,625	78,250	12,984	12,467	545	545	121
7	cnx	Jun-13	63,508	63,360	13,469	13,028	516	516	118
8	cnx	Jul-13	74,545	77,281	14,067	12,911	482	482	127
9	cnx	Aug-13	80,825	85,333	15,533	15,862	546	546	142
10	cnx	Sep-13	79,105	81,239	12,831	12,106	510	510	115
11	cnx	Oct-13	80,272	81,327	10,312	10,376	574	573	76
12	cnx	Nov-13	84,082	84,120	11,409	10,264	574	575	77
13	cnx	Dec-13	107,751	105,357	11,238	11,031	683	712	75
14	cnx	Jan-14	135,070	137,496	28,211	23,852	329	328	158
15	cnx	Feb-14	116,432	123,722	24,525	26,270	612	612	186
16	cnx	Mar-14	130,000	135,450	26,182	26,598	1,013	1,013	206
17	cnx	Apr-14	131,160	131,202	24,411	24,850	939	939	193
18	cnx	May-14	121,070	120,633	24,120	23,014	865	865	196
19	cnx	Jun-14	113,059	113,336	22,584	21,569	751	752	186
20	cnx	Jul-14	126,436	127,137	27,795	25,070	792	791	197
21	cnx	Aug-14	132,281	137,289	27,658	27,603	868	868	197
22	cnx	Sep-14	121,573	122,821	24,836	22,936	798	799	184
23	cnx	Oct-14	136,113	136,083	13,252	12,730	708	708	126
24	cnx	Nov-14	110,311	109,278	20,059	18,151	672	672	138
25	cnx	Dec-14	138,828	134,948	19,459	18,866	903	901	138
26	cnx	Jan-15	184,819	198,524	43,767	40,851	1,235	1,235	303
27	cnx	Feb-15	167,845	177,213	41,869	42,102	1,122	1,122	300
28	cnx	Mar-15	166,559	176,829	37,387	36,840	1,206	1,206	290
29	cnx	Apr-15	157,838	157,935	40,430	37,974	1,102	1,102	292
30	cnx	May-15	155,208	154,942	35,365	34,562	1,074	1,074	287
31	cnx	Jun-15	148,508	151,028	38,412	36,139	1,001	1,001	271
32	cnx	Jul-15	166,993	166,965	42,123	38,600	1,040	1,040	285
33	cnx	Aug-15	161,828	171,068	40,838	40,522	1,044	1,044	274
34	cnx	Sep-15	154,540	154,404	36,404	35,341	999	998	243

ภาพที่ 3.8 แสดงการจัดกลุ่มท่าอากาศยานเชียงใหม่

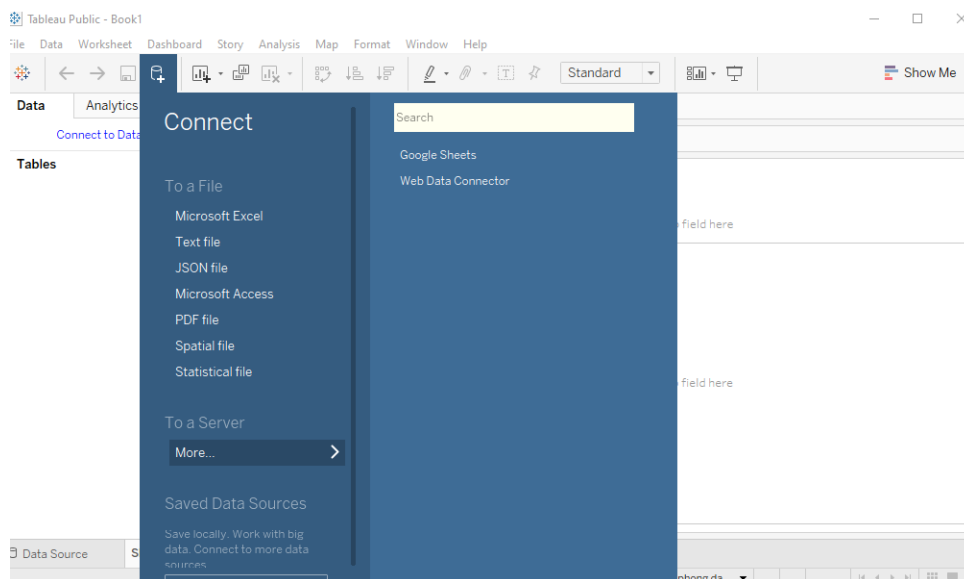
5) การโหลดข้อมูล (Loading) โดยใช้โปรแกรม Tableau Public ดำเนินกระบวนการโหลดข้อมูลเข้าสู่ระบบฐานข้อมูลโปรแกรม Tableau Public เพื่อเป็นการควบคุมคุณภาพของข้อมูล เพื่อที่จะนำข้อมูลมาวิเคราะห์หรืออย่างถูกต้องสมบูรณ์และสร้างออกมาเป็นรายงานภาพ Visualization โดยมีขั้นตอนดังต่อไปนี้

5.1 จัดเตรียมข้อมูลที่ต้องการนำมาแสดงผลลัพธ์

Name	Date modified	Type	Size
BKK_ท่าอากาศยานสุวรรณภูมิ_แก้	9/5/2020 10:44 AM	Microsoft Excel W...	15 KB
CEI_ท่าอากาศยานแม่ฟ้าหลวง เชียงราย_แก้	9/6/2020 11:15 PM	Microsoft Excel W...	15 KB
cnx_ท่าอากาศยานเชียงใหม่_แก้	9/4/2020 2:26 AM	Microsoft Excel W...	18 KB
DMK_ท่าอากาศยานดอนเมือง_แก้	9/5/2020 11:23 AM	Microsoft Excel W...	16 KB
HDY_ท่าอากาศยานหาดใหญ่_แก้	9/6/2020 10:05 PM	Microsoft Excel W...	15 KB
HKT_ท่าอากาศยานภูเก็ต_แก้	9/6/2020 11:32 PM	Microsoft Excel W...	16 KB
...

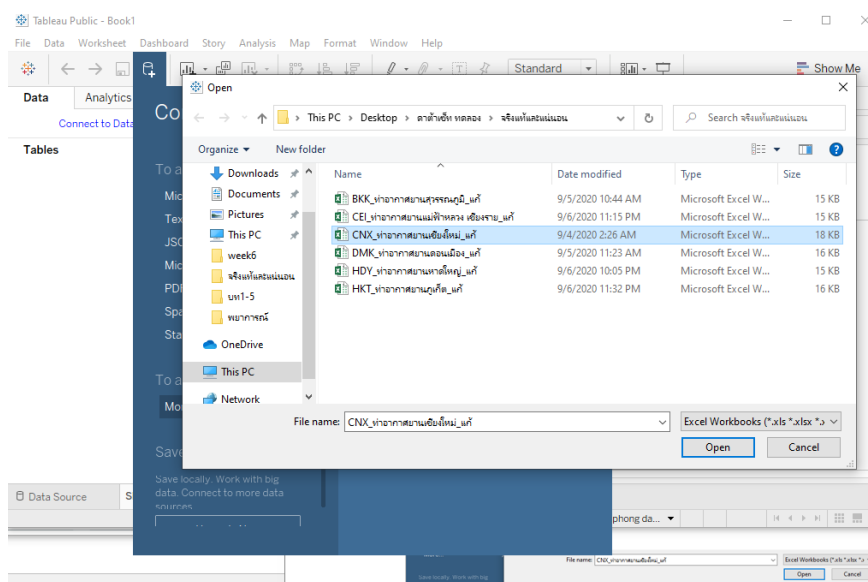
ภาพที่ 3.9 ไฟล์ข้อมูลที่นำมาแสดงผลในโปรแกรม Tableau Public

5.2 ใช้ฟังก์ชัน New Data Source ในการนำข้อมูลเข้าสู่โปรแกรม Tableau Public ในรูปแบบไฟล์ Microsoft Excel เพื่อนำมาแสดงผลยังโปรแกรม Tableau Public



ภาพที่ 3.10 การ Import ข้อมูล

5.3 เลือกไฟล์ข้อมูล Microsoft Excel ที่จะนำมาแสดงผลลัพธ์



ภาพที่ 3.11 การ Import ข้อมูลที่จะหาผลลัพธ์

3.1.4 Modeling ผู้วิเคราะห์ข้อมูลวิเคราะห์ข้อมูลด้วยเทคนิคทางดาต้าไมน์นิ่ง Classification ทำการเรียนรู้จากข้อมูลเดิมของการใช้บริการสายการบินต้นทุนต่ำเพื่อให้ได้คำตอบที่ดีที่สุด โดยการใช้โมเดล Time series หาค่าข้อมูลที่เปลี่ยนแปลงไปตามเวลา ในการทำนายผล โดยมีขั้นตอนการทำดังต่อไปนี้

ในการพยากรณ์โดยใช้โมเดล Time series เพื่อที่จะหาข้อมูลสถิติในอนาคตจำเป็นต้องใช้ข้อมูลย้อนหลังในอดีตเท่านั้น โดยตัวแปรอื่น ๆ จะไม่นำมาพิจารณา ซึ่งรูปแบบของข้อมูล (Data Pattern) มีเทคนิคการพยากรณ์หลายรูปแบบ แต่โดยปกติข้อมูลจะแบ่งได้หลายรูปแบบ ซึ่งทางผู้วิเคราะห์ข้อมูลได้เลือกข้อมูลในรูปแบบมีแนวโน้ม (Trend) ความต้องการเปลี่ยนไปอย่างมีทิศทาง

Airport name	Date	Number of domestic arrivals	Number of domestic departures	Number of international arrivals	Number of international departures
cnx	Jan-13	106,563	116,690	11,012	10,614
cnx	Feb-13	98,040	104,731	12,113	11,200
cnx	Mar-13	102,753	106,116	12,897	12,411
cnx	Apr-13	91,126	91,821	12,812	12,225
cnx	May-13	78,625	78,250	12,984	12,467
cnx	Jun-13	69,508	69,360	13,469	13,028
cnx	Jul-13	74,545	77,281	14,067	12,971
cnx	Aug-13	80,825	86,333	15,559	15,682
cnx	Sep-13	79,105	81,299	12,831	12,106
cnx	Oct-13	80,272	81,927	10,312	10,376
cnx	Nov-13	84,082	84,120	11,409	10,264
cnx	Dec-13	107,751	105,957	11,238	11,031
cnx	Jan-14	135,070	137,498	26,211	23,852
cnx	Feb-14	116,432	123,722	24,525	26,270
cnx	Mar-14	130,000	135,450	26,182	26,598
cnx	Apr-14	131,160	131,202	24,411	24,850
cnx	May-14	121,010	120,633	24,120	23,014
cnx	Jun-14	113,099	113,326	22,564	21,569
cnx	Jul-14	126,436	127,137	27,795	25,070
cnx	Aug-14	132,281	137,289	27,658	27,603
cnx	Sep-14	121,573	122,821	24,836	22,936
cnx	Oct-14	106,113	106,063	13,252	12,730
cnx	Nov-14	110,311	109,278	20,059	18,151
cnx	Dec-14	138,828	134,948	19,459	18,866
cnx	Jan-15	184,819	198,524	43,767	40,851
cnx	Feb-15	167,645	177,213	41,869	42,102

ภาพที่ 3.12 ตารางข้อมูลจำนวนผู้โดยสารขาเข้าขาออกในประเทศและระหว่างประเทศ

Number of domestic inbound flights	Number of domestic outbound flights	Number of international inbound flights	Number of international outbound flights
775	775	76	76
658	658	75	75
694	694	118	118
635	635	114	114
545	545	121	121
516	516	118	118
482	482	127	127
546	546	142	142
510	510	119	119
574	573	77	76
574	575	75	75
683	712	75	75
928	928	198	198
812	812	186	186
1,013	1,013	206	206
939	939	193	193
885	885	198	198
751	752	186	186
792	791	197	197
868	868	197	197
798	799	184	185
708	708	126	126
672	672	138	138
903	901	138	138
1,235	1,235	305	303
1,122	1,122	300	300
1,206	1,206	288	288

ภาพที่ 3.13 ตารางข้อมูลจำนวนเที่ยวบินขาเข้าขาออกในประเทศและระหว่างประเทศซึ่งเป็นไฟล์ Microsoft Excel เดียวกับ ภาพที่ 3.8

จากภาพที่ 3.8 และ 3.9 ประกอบด้วย 10 แอตทริบิวต์

- 1) Airport name เก็บข้อมูลชื่อตัวย่อท่าอากาศยาน
- 2) Date เก็บข้อมูล เดือนและปี ของข้อมูลท่าอากาศยาน
- 3) Number of domestic arrivals เก็บข้อมูลจำนวนประชากรขาเข้าภายในประเทศ
- 4) Number of domestic departures เก็บข้อมูลจำนวนประชากรขาออกภายในประเทศ
- 5) Number of international arrivals เก็บข้อมูลจำนวนประชากรขาเข้าระหว่างประเทศ
- 6) Number of international departures เก็บข้อมูลจำนวนประชากรขาออกระหว่างประเทศ
- 7) Number of domestic inbound flights เก็บข้อมูลจำนวนเที่ยวบินขาเข้าภายในประเทศ
- 8) Number of domestic outbound flights เก็บข้อมูลจำนวนเที่ยวบินขาออกภายในประเทศ
- 9) Number of international inbound flights เก็บข้อมูลจำนวนเที่ยวบินขาเข้าระหว่างประเทศ
- 10) Number of international inbound flights เก็บข้อมูลจำนวนเที่ยวบินขาออกระหว่าง

ประเทศ

3.1.5 Evaluation ผู้วิเคราะห์ข้อมูลได้ผลการวิเคราะห์ข้อมูลด้วยเทคนิคทางดาต้าไมนิงจากการสร้างโมเดล Time series สามารถวัดประสิทธิภาพของ ผลลัพธ์ที่ได้ว่าตรงกับวัตถุประสงค์ที่ตั้งไว้ โดยผู้วิเคราะห์ได้นำผลลัพธ์มาเทียบดังต่อไปนี้

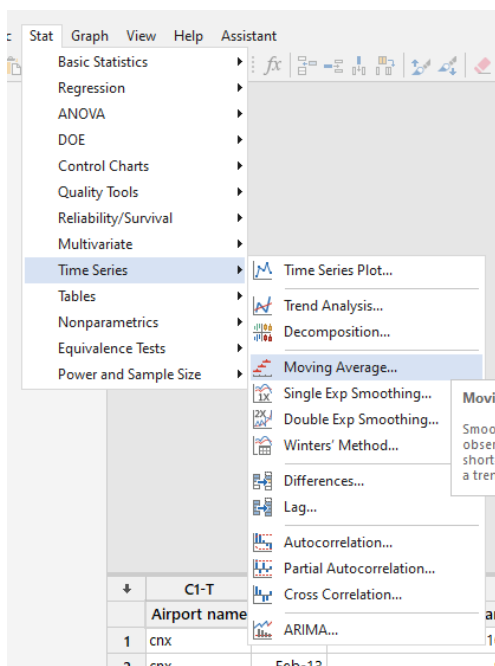
1) การพยากรณ์โดยการวิเคราะห์อนุกรมเวลาการทำเหมืองข้อมูล (Data mining time series analysis) โดยการใช้โปรแกรม Minitab 19 ด้วยวิธี Moving Average ซึ่งการพยากรณ์เป็นการพยากรณ์ในรูปแบบระยะสั้น โดยนำชุดข้อมูลแทนที่ชุดข้อมูลเก่าที่สุดแล้วทำการหาค่าเฉลี่ยใหม่ในช่วงเวลา 3 เดือน 5 เดือน โดยการพยากรณ์วิธี Moving Average เหมาะกับการพยากรณ์ยอดขาย ไม่เหมาะกับข้อมูลรูปแบบแนวโน้มและรูปแบบฤดูกาล โดยผลลัพธ์ที่ได้จากการพยากรณ์ได้ค่าความน่าเชื่อถือ MAPE เท่ากับ 9% โดยมีขั้นตอนการทำงานดังต่อไปนี้

ขั้นตอนที่ 1 ทำการจัดเตรียมข้อมูล Data set และทำการ insert ข้อมูลเข้าไปในโปรแกรม Minitab 19

+	C1-T	C2-D	C3	C4	C5	C6	C7
	Airport name	Date	Number of domestic arrivals	Number of domestic departures	Number of international arrival	Number of international departu	Number of domestic in
7	cnx	Jul-13	74545	77281	14067	12971	
8	cnx	Aug-13	80825	86333	15559	15682	
9	cnx	Sep-13	79105	81299	12831	12106	
10	cnx	Oct-13	80272	81927	10312	10376	
11	cnx	Nov-13	84082	84120	11409	10264	
12	cnx	Dec-13	107751	105957	11238	11031	
13	cnx	Jan-14	135070	137498	26211	23852	
14	cnx	Feb-14	116432	123722	24525	26270	
15	cnx	Mar-14	130000	135450	26182	26598	
16	cnx	Apr-14	131160	131202	24411	24850	
17	cnx	May-14	121010	120633	24120	23014	
18	cnx	Jun-14	113099	113326	22564	21569	
19	cnx	Jul-14	126436	127137	27795	25070	
20	cnx	Aug-14	132281	137289	27658	27603	
21	cnx	Sep-14	121573	122821	24836	22936	
22	cnx	Oct-14	106113	106063	13252	12730	
23	cnx	Nov-14	110311	109278	20059	18151	
24	cnx	Dec-14	138828	134948	19459	18866	
25	cnx	Jan-15	184819	198524	43767	40851	
26	cnx	Feb-15	167645	177213	41869	42102	
27	cnx	Mar-15	168559	176829	37397	38840	
28	cnx	Apr-15	157838	157935	40430	37974	
29	cnx	May-15	155208	154942	35965	34562	
30	cnx	Jun-15	148508	151028	38412	36139	
31	cnx	Jul-15	166993	166965	42123	38600	

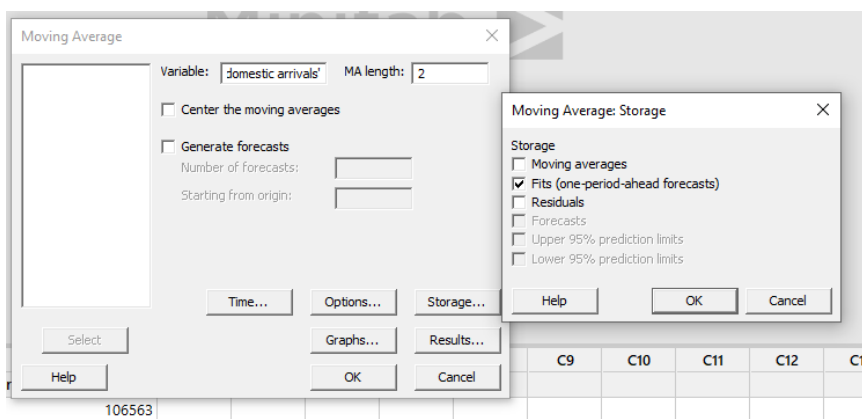
ภาพที่ 3.14 ข้อมูลที่ insert ข้อมูลเข้าไปในโปรแกรม Minitab 19

ขั้นตอนที่ 2 ทำการเลือกรูปแบบวิธีที่ต้องการจะนำข้อมูลไปทำการวิเคราะห์
(Moving Average)



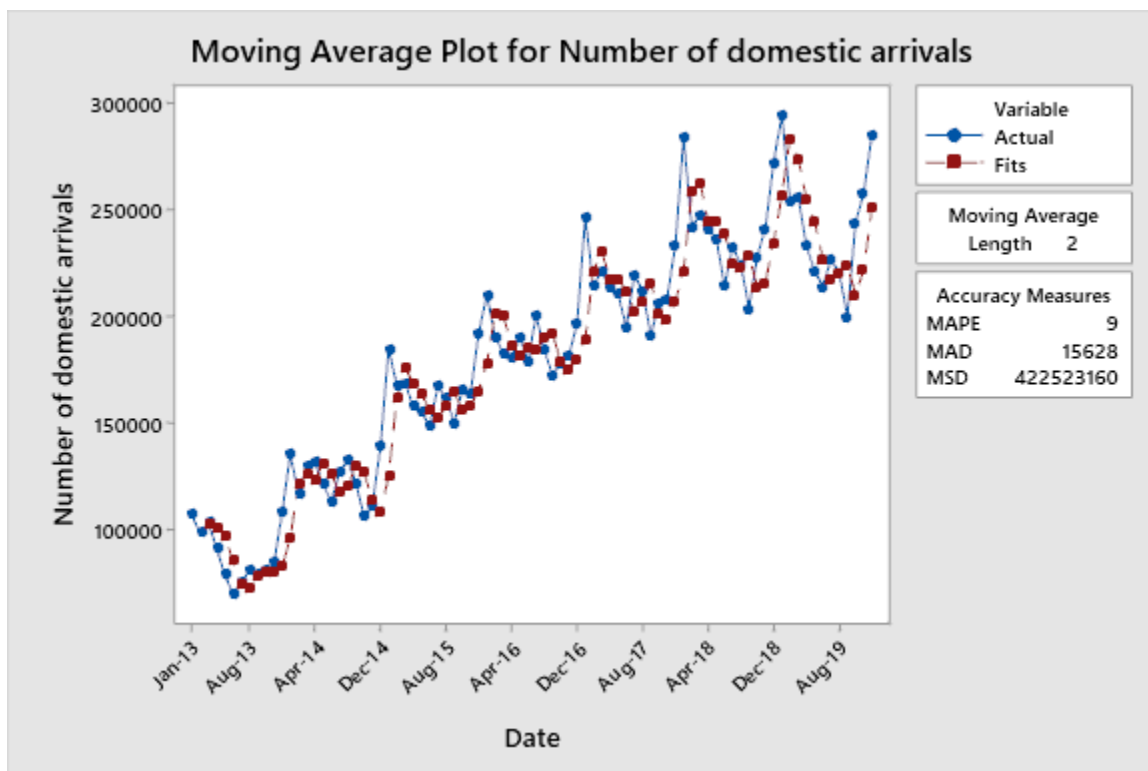
ภาพที่ 3.15 แสดงแถบเมนู Moving Average

ขั้นตอนที่ 3 ทำการพยากรณ์ล่วงหน้า โดยกำหนดช่วงเวลา 2 ปี



ภาพที่ 3.16 แสดงรูปแบบ ทำการพยากรณ์ล่วงหน้า 2 ปี

ขั้นตอนที่ 4 Minitab 19 จะแสดงผลลัพธ์ออกมาในรูปแบบกราฟและสามารถดูค่าที่อยู่ภายในได้



ภาพที่ 3.17 ผลลัพธ์ที่ได้ จาก โปรแกรม Minitab 19 ด้วยวิธี Moving Average

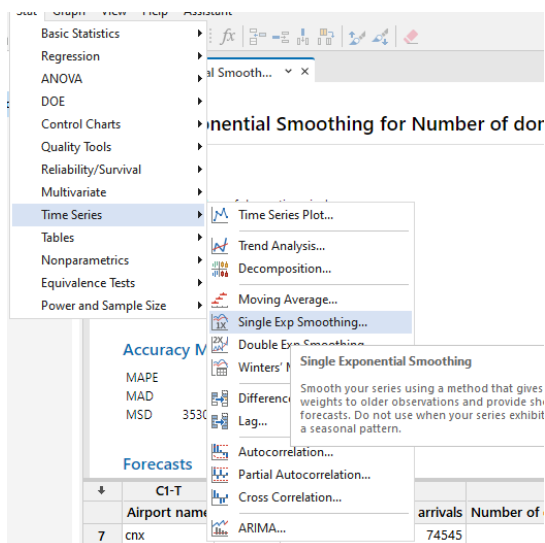
2) การพยากรณ์โดยการวิเคราะห์อนุกรมเวลาการทำเหมืองข้อมูล (Data mining time series analysis) โดยการใช้โปรแกรม Minitab 19 ด้วยวิธี Single Exponential smoothing ซึ่งการพยากรณ์วิธี Single Exponential smoothing เหมาะกับการพยากรณ์ระยะสั้น เทคนิคเชิงคุณภาพ และเทคนิคการวิเคราะห์การถดถอยเหมาะสำหรับการพยากรณ์ระยะยาว โดยผลลัพธ์ที่ได้จากการพยากรณ์ได้ค่าความน่าเชื่อถือ MAPE เท่ากับ 10% โดยมีขั้นตอนการทำงานดังต่อไปนี้

ขั้นตอนที่ 1 ทำการจัดเตรียมข้อมูล Data set และทำการ insert ข้อมูลเข้าไปในโปรแกรม Minitab 19

#	C1-T	C2-D	C3	C4	C5	C6	C7
	Airport name	Date	Number of domestic arrivals	Number of domestic departures	Number of international arrival	Number of international departu	Number of domestic in
7	cnx	Jul-13	74545	77281	14067	12971	
8	cnx	Aug-13	80825	86333	15559	15682	
9	cnx	Sep-13	79105	81299	12831	12106	
10	cnx	Oct-13	80272	81927	10312	10376	
11	cnx	Nov-13	84082	84120	11409	10264	
12	cnx	Dec-13	107751	105957	11238	11031	
13	cnx	Jan-14	135070	137498	26211	23852	
14	cnx	Feb-14	116432	123722	24525	26270	
15	cnx	Mar-14	130000	135450	26182	26598	
16	cnx	Apr-14	131160	131202	24411	24850	
17	cnx	May-14	121010	120633	24120	23014	
18	cnx	Jun-14	113099	113326	22564	21569	
19	cnx	Jul-14	126436	127137	27795	25070	
20	cnx	Aug-14	132281	137289	27658	27603	
21	cnx	Sep-14	121573	122821	24836	22936	
22	cnx	Oct-14	106113	106063	13252	12730	
23	cnx	Nov-14	110311	109278	20059	18151	
24	cnx	Dec-14	138828	134948	19459	18866	
25	cnx	Jan-15	184819	198524	43767	40851	
26	cnx	Feb-15	167645	177213	41869	42102	
27	cnx	Mar-15	168559	176829	37397	38840	
28	cnx	Apr-15	157838	157935	40430	37974	
29	cnx	May-15	155208	154942	35965	34562	
30	cnx	Jun-15	148508	151028	38412	36139	
31	cnx	Jul-15	166993	166965	42123	38600	

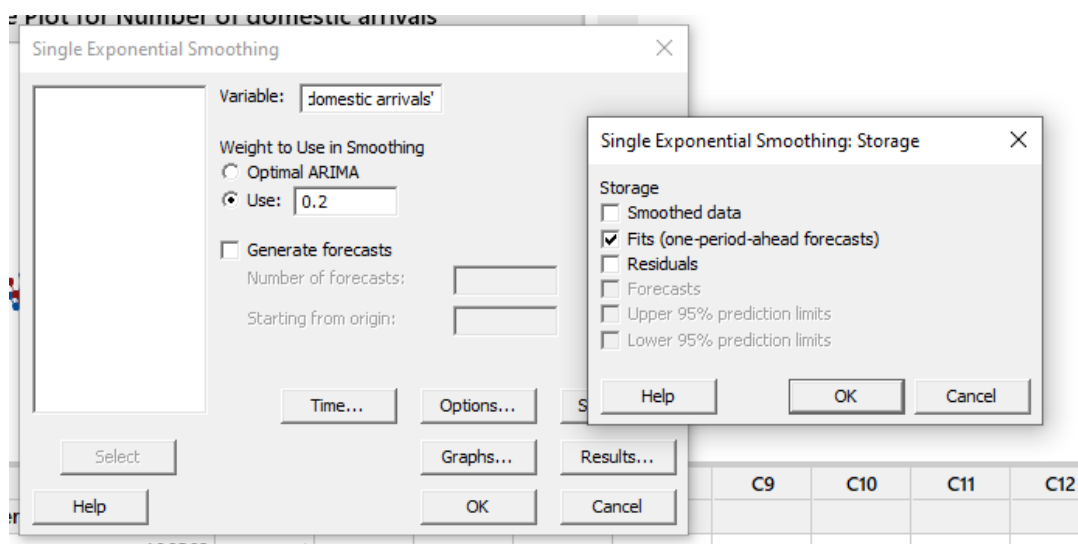
ภาพที่ 3.18 ข้อมูลที่ insert ข้อมูลเข้าไปในโปรแกรม Minitab 19

ขั้นตอนที่ 2 ทำการเลือกรูปแบบวิธีที่ต้องการจะนำข้อมูลไปทำการวิเคราะห์ (Single Exponential smoothing)



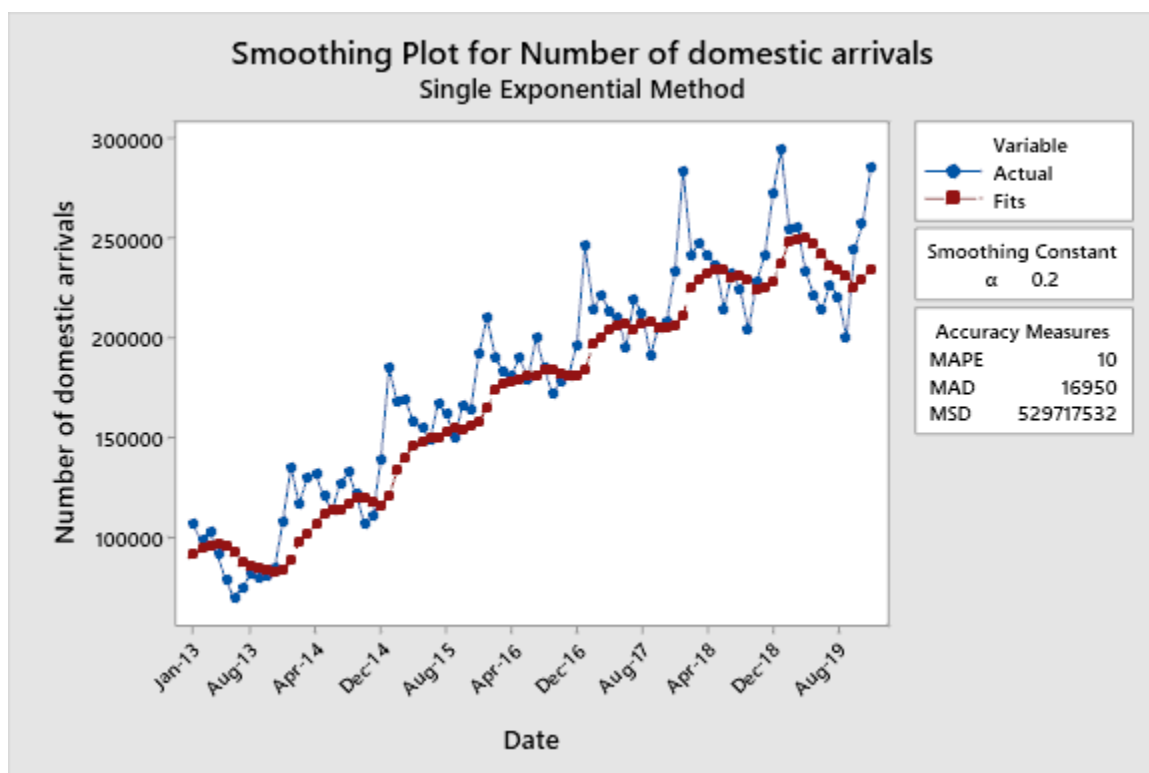
ภาพที่ 3.19 แสดงแถบเมนู Single Exponential smoothing

ขั้นตอนที่ 3 ทำการพยากรณ์ ล่วงหน้า โดยกำหนดช่วงเวลา 2 ปี



ภาพที่ 3.20 แสดงรูปแบบ ทำการพยากรณ์ ล่วงหน้า 2 ปี

ขั้นตอนที่ 4 Minitab 19 จะแสดงผลลัพธ์ออกมาในรูปแบบกราฟและสามารถดูค่าที่อยู่ภายในได้



ภาพที่ 3.21 ผลลัพธ์ที่ได้ จาก โปรแกรม Minitab 19 ด้วยวิธี Single Exponential smoothing

3) การพยากรณ์โดยการวิเคราะห์อนุกรมเวลาการทำเหมืองข้อมูล (Data mining time series analysis) โดยการใช้โปรแกรม Minitab 19 ด้วยวิธี Double Exponential smoothing ซึ่งวิธี Double Exponential smoothing เป็นวิธีการปรับเรียบโดยการนำค่าการพยากรณ์มาเปรียบเทียบกับอีกครั้ง เพื่อพยายามลดปัจจัยอันเกิดจากการเปลี่ยนแปลงที่อธิบายไม่ได้ เหมาะสำหรับรูปแบบความต้องการที่มีการเปลี่ยนแปลงแบบมีทิศทาง โดยผลลัพธ์ที่ได้จากการพยากรณ์ได้ค่าความน่าเชื่อถือ MAPE เท่ากับ 10% โดยมีขั้นตอนการทำงานดังต่อไปนี้

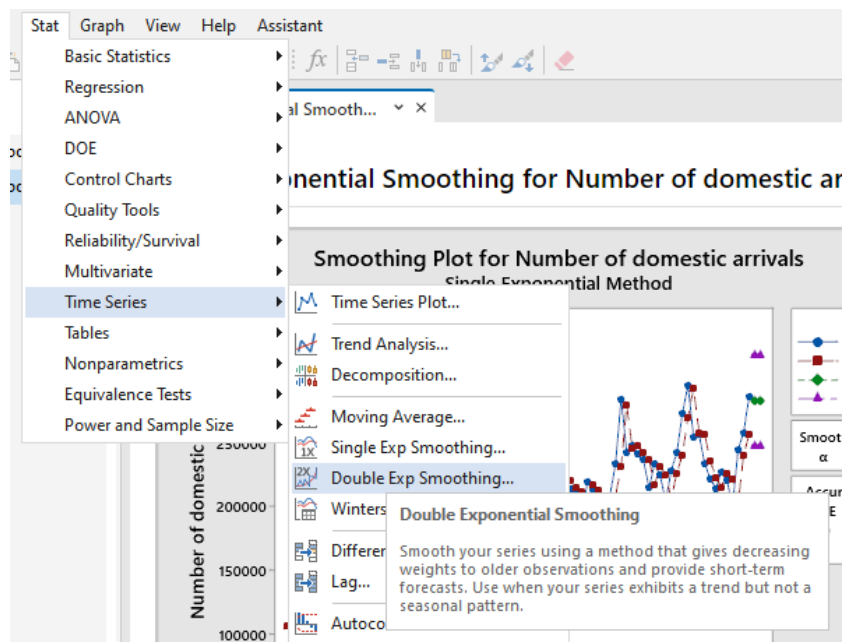
ขั้นตอนที่ 1 ทำการจัดเตรียมข้อมูล Data set และทำการ insert ข้อมูลเข้าไปในโปรแกรม Minitab 19

#	C1-T	C2-D	C3	C4	C5	C6	C7
	Airport name	Date	Number of domestic arrivals	Number of domestic departures	Number of international arrival	Number of international departu	Number of domestic in
7	cnx	Jul-13	74545	77281	14067	12971	
8	cnx	Aug-13	80825	86333	15559	15682	
9	cnx	Sep-13	79105	81299	12831	12106	
10	cnx	Oct-13	80272	81927	10312	10376	
11	cnx	Nov-13	84082	84120	11409	10264	
12	cnx	Dec-13	107751	105957	11238	11031	
13	cnx	Jan-14	135070	137498	26211	23852	
14	cnx	Feb-14	116432	123722	24525	26270	
15	cnx	Mar-14	130000	135450	26182	26598	
16	cnx	Apr-14	131160	131202	24411	24850	
17	cnx	May-14	121010	120633	24120	23014	
18	cnx	Jun-14	113099	113326	22564	21569	
19	cnx	Jul-14	126436	127137	27795	25070	
20	cnx	Aug-14	132281	137289	27658	27603	
21	cnx	Sep-14	121573	122821	24836	22936	
22	cnx	Oct-14	106113	106063	13252	12730	
23	cnx	Nov-14	110311	109278	20059	18151	
24	cnx	Dec-14	138828	134948	19459	18866	
25	cnx	Jan-15	184819	198524	43767	40851	
26	cnx	Feb-15	167645	177213	41869	42102	
27	cnx	Mar-15	168559	176829	37397	38840	
28	cnx	Apr-15	157838	157935	40430	37974	
29	cnx	May-15	155208	154942	35965	34562	
30	cnx	Jun-15	148508	151028	38412	36139	
31	cnx	Jul-15	166993	166965	42123	38600	

ภาพที่ 3.22 แสดงแสดงข้อมูลที่ insert ข้อมูลเข้าไปในโปรแกรม Minitab 19

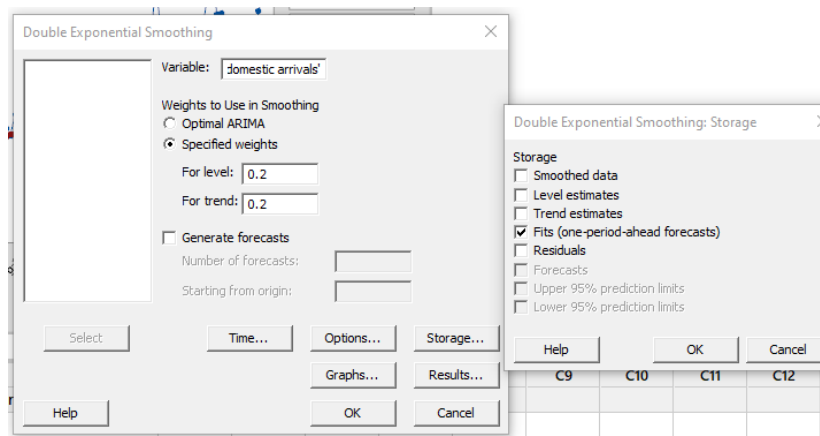
ขั้นตอนที่ 2 ทำการเลือกรูปแบบวิธีที่ต้องการจะนำข้อมูลไปทำการวิเคราะห์

Double Exponential smoothing



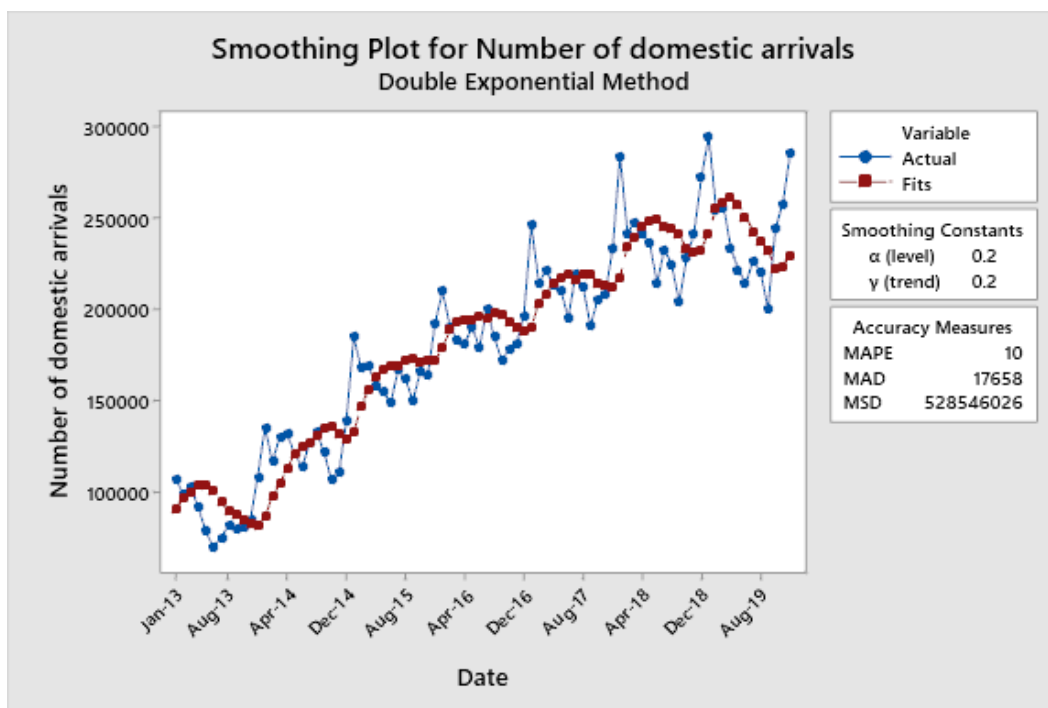
ภาพที่ 3.23 แสดงแถบเมนู Double Exponential smoothing

ขั้นตอนที่ 3 ทำการพยากรณ์ ล่วงหน้า โดยกำหนดช่วงเวลา 2 ปี



ภาพที่ 3.24 แสดงรูปแบบ ทำการพยากรณ์ ล่วงหน้า 2 ปี

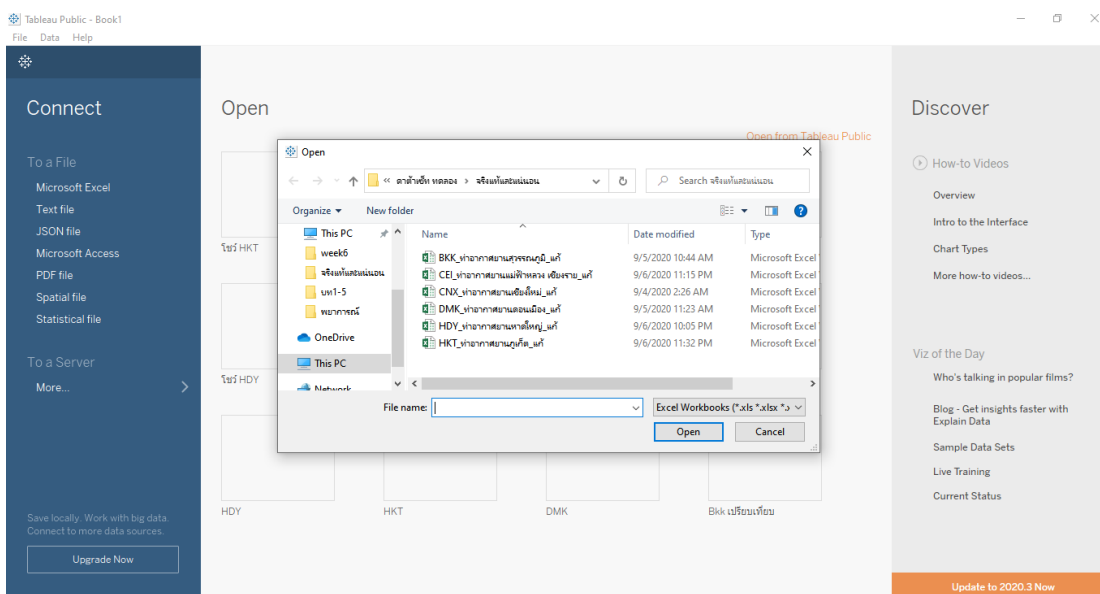
ขั้นตอนที่ 4 Minitab 19 จะแสดงผลลัพธ์ออกมาในรูปแบบกราฟและสามารถดูค่าที่อยู่ภายในได้



ภาพที่ 3.25 แสดงผลลัพธ์ Minitab 19 โดยใช้วิธี Double Exponential smoothing

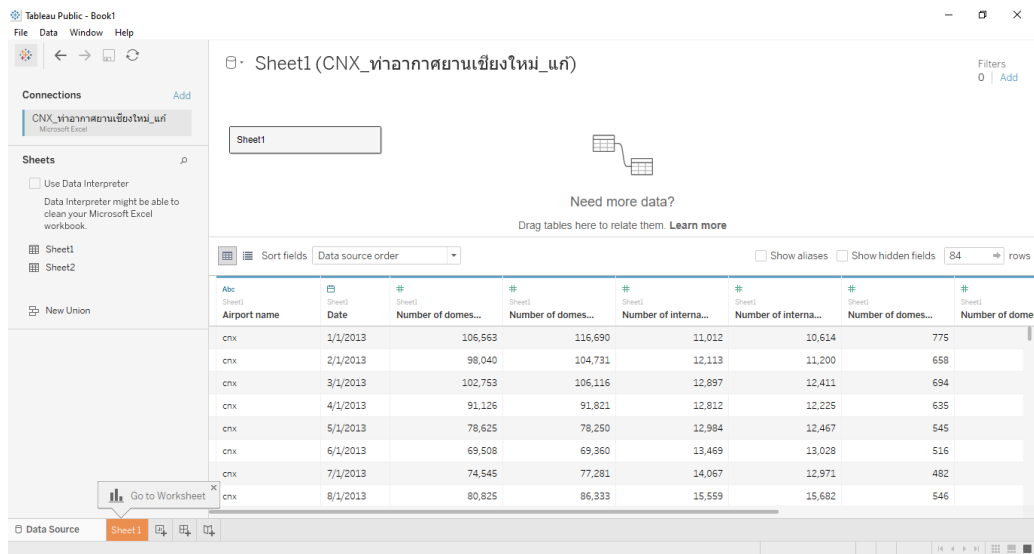
3) การพยากรณ์โดยการวิเคราะห์อนุกรมเวลาการทำเหมืองข้อมูล (Data mining time series analysis) โดยการใช้โปรแกรม Tableau Public โดยการพยากรณ์ล่วงหน้าจำนวน 2 ปี ซึ่งใช้ฟังก์ชัน Forecast ซึ่งเป็นวิธีคาดคะเนแนวโน้ม (Trend Projections) เป็นการใช้ข้อมูลในอดีตเพื่อพยากรณ์แนวโน้มในอนาคตโดยอาศัยวิธีการกำลังสองน้อยที่สุดมาช่วย โดยมีเงื่อนไขว่า ค่าข้อมูลต่าง ๆ เมื่อนำมาเขียนในเส้นกราฟจะต้องมีความสัมพันธ์เป็นเส้นตรง และมีค่าเบี่ยงเบนของข้อมูลห่างจากเส้นพยากรณ์ไม่มาก เหมาะกับข้อมูลเชิงสถิติที่ต้องการทราบจำนวนหรือข้อมูลสถิติในอนาคต ซึ่งได้ผลลัพธ์ออกมาจากการพยากรณ์ได้ค่าความน่าเชื่อถือ MAPE เท่ากับ 9% โดยมีขั้นตอนการทำงานดังต่อไปนี้

ขั้นตอนที่ 1 เลือกข้อมูลจากไฟล์ Excel ที่ได้ทำการเตรียมข้อมูลไว้



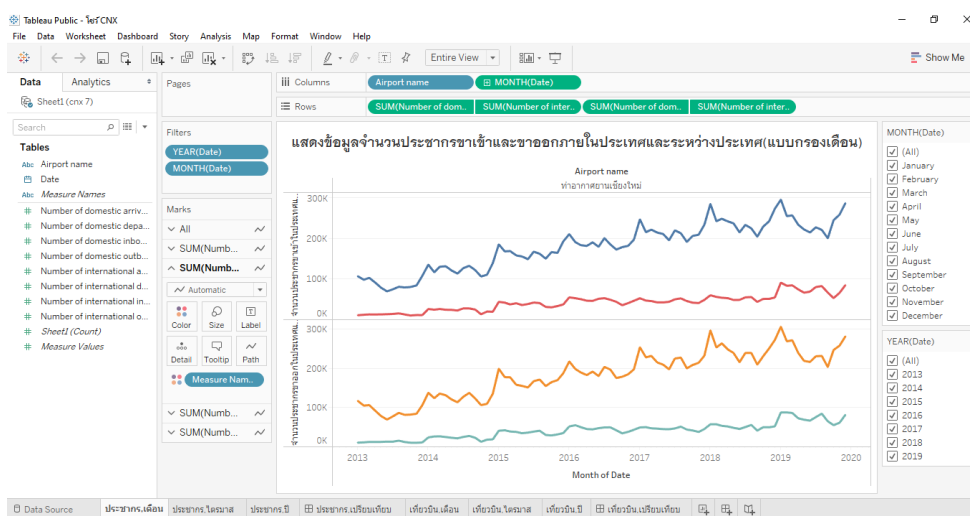
ภาพที่ 3.26 เปิดไฟล์ Data Set จากไฟล์ Excel ด้วยโปรแกรม Tableau

ขั้นตอนที่ 2 insert ข้อมูล Data Set



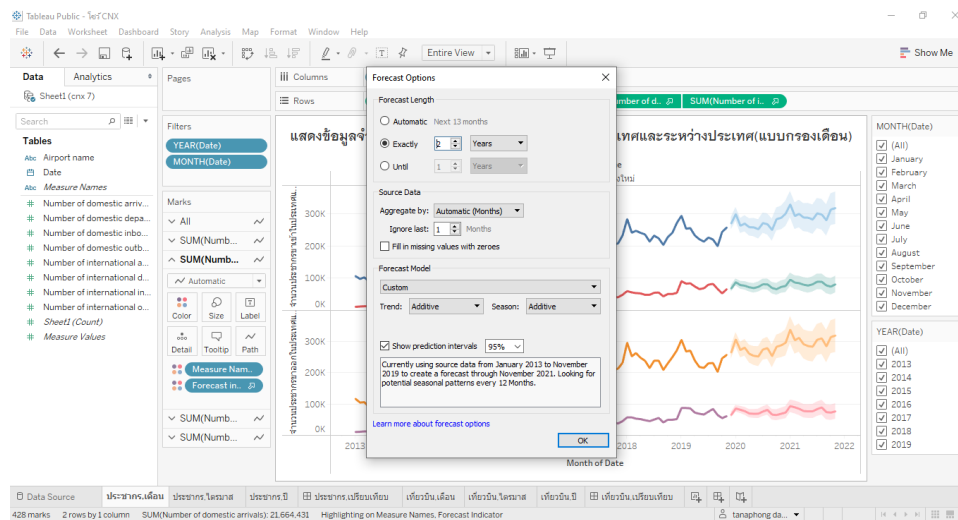
ภาพที่ 3.27 insert ข้อมูล Data Set ของสายการบินต้นทุนต่ำ เข้าไปในโปรแกรม Tableau

ขั้นตอนที่ 3 ทำการเลือกข้อมูลที่เหมาะสมเพื่อสร้างกราฟเปรียบเทียบแสดงจำนวนเที่ยวบินภายในประเทศของแต่ละท่าอากาศยาน



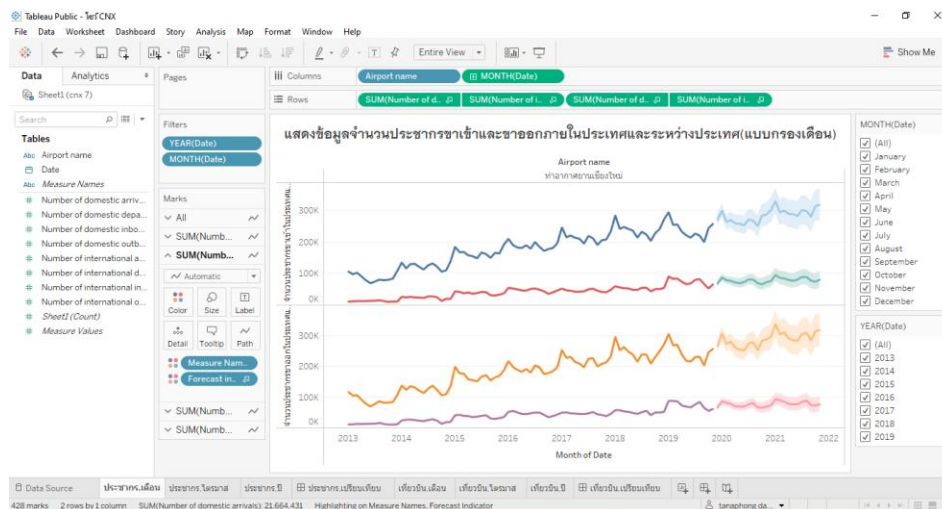
ภาพที่ 3.28 ข้อมูลเที่ยวบินสายการบินภายในประเทศของสายการบินต้นทุนต่ำ ที่นำมาแสดงในโปรแกรม Tableau

ขั้นตอนที่ 4 ทำการทำพยากรณ์ข้อมูลโดยใช้โปรแกรม Tableau โดยใช้ฟังก์ชัน Forecast โดยจะทำการพยากรณ์ล่วงหน้าไปอีก 2 ข้างหน้า



ภาพที่ 3.29 พยากรณ์ข้อมูลโดยใช้โปรแกรม Tableau โดยใช้ฟังก์ชัน Forecast

ขั้นตอนที่ 5 แสดงกราฟเปรียบเทียบแสดงจำนวนประชากรและเที่ยวบินที่ใช้บริการภายในประเทศและระหว่างประเทศ โดยทำการพยากรณ์สถิติการให้บริการล่วงหน้า 2 ปี



ภาพที่ 3.30 แสดงกราฟเปรียบเทียบแสดงจำนวนประชากรและเที่ยวบินที่ทำกรพยากรณ์สถิติการให้บริการล่วงหน้า 2 ปี

จากการทดสอบโมเดล Time series สามารถวัดประสิทธิภาพของผลลัพธ์ในแต่ละวิธีที่แตกต่างกันซึ่งเป็นการพยากรณ์ล่วงหน้า 2 ปี เช่นเดียวกันทั้ง 4 ดังต่อไปนี้ โดยเริ่มจาก วิธีที่ 1 การพยากรณ์โดยการวิเคราะห์หอนุกรมเวลาการทำเหมืองข้อมูล (Data mining time series analysis) โดยการใช้โปรแกรม Minitab 19 ด้วยวิธี Moving Average โดยผลลัพธ์ที่ได้จากการพยากรณ์ได้ค่าความน่าเชื่อถือ MAPE เท่ากับ 9% วิธีที่ 2 การพยากรณ์โดยการวิเคราะห์หอนุกรมเวลาการทำเหมืองข้อมูล (Data mining time series analysis) โดยการใช้โปรแกรม Minitab 19 ด้วยวิธี Single Exponential โดยผลลัพธ์ที่ได้จากการพยากรณ์ได้ค่าความน่าเชื่อถือ MAPE เท่ากับ 10% วิธีที่ 3 การพยากรณ์โดยการวิเคราะห์หอนุกรมเวลาการทำเหมืองข้อมูล (Data mining time series analysis) โดยการใช้โปรแกรม Minitab 19 ด้วยวิธี Double Exponential โดยผลลัพธ์ที่ได้จากการพยากรณ์ได้ค่าความน่าเชื่อถือ MAPE เท่ากับ 10% และวิธีที่ 4 การใช้โปรแกรม Tableau Public โดยการพยากรณ์ล่วงหน้าจำนวน 2 ปี ซึ่งใช้ฟังก์ชัน Forecast ซึ่งได้ผลลัพธ์ออกมาจากการพยากรณ์ได้ค่าความน่าเชื่อถือ MAPE เท่ากับ 4.5%

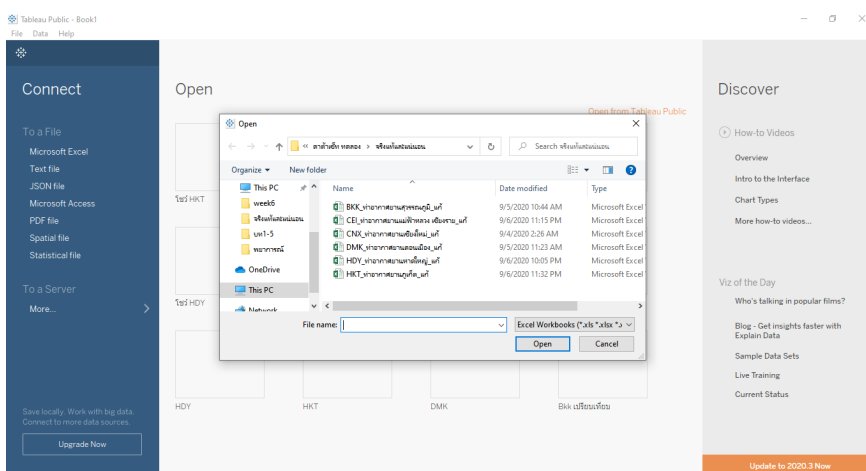
จากผลการทดสอบประสิทธิภาพโมเดลและประเมินค่า MAPE พบว่าหากวิธีไหนได้ค่าความคลาดเคลื่อน MAPE ที่น้อยที่สุดคือวิธีนั้นมีประสิทธิภาพที่ดีที่สุดเหมาะสมแก่การนำไปพยากรณ์ข้อมูล ดังนั้นผู้วิเคราะห์จึงได้ใช้โปรแกรม Tableau Public มาการพยากรณ์ล่วงหน้าจำนวน 2 ปี ตารางแสดงการเปรียบเทียบค่าความคลาดเคลื่อน MAPE

ประเภทของโปรแกรม / วิธี	MAPE
1. โปรแกรม Minitab 19	
- วิธี Moving Average	9%
- วิธี Single Exponential smoothing	10%
- วิธี Double Exponential smoothing	10%
2. โปรแกรม Tableau Public	
- ฟังก์ชัน Forecast	4.5%

ตารางที่ 3.1 การเปรียบเทียบค่าความคลาดเคลื่อน MAPE โดยใช้ 3 วิธี

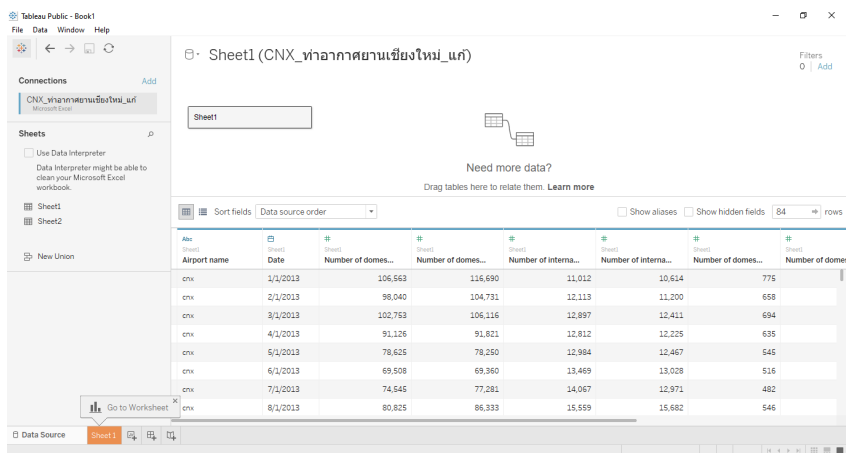
3.1.6 Deployment ผู้วิเคราะห์ข้อมูลนำผลองค์ความรู้ที่ได้เหล่านี้ไปเผยแพร่ผ่าน ทางแสดงผลข้อมูลบน Web browser โดยใช้ชุดคำสั่ง HTML และ CSS ร่วมกับการนำเสนอข้อมูลแบบ visualization ด้วยการแสดงผลข้อมูลในรูปแบบ ของภาพโดยใช้โปรแกรม Tableau โดยมีขั้นตอนดังต่อไปนี้

ขั้นตอนที่ 1 เลือกข้อมูลจากไฟล์ Excel ที่ได้ทำการเตรียมข้อมูลไว้



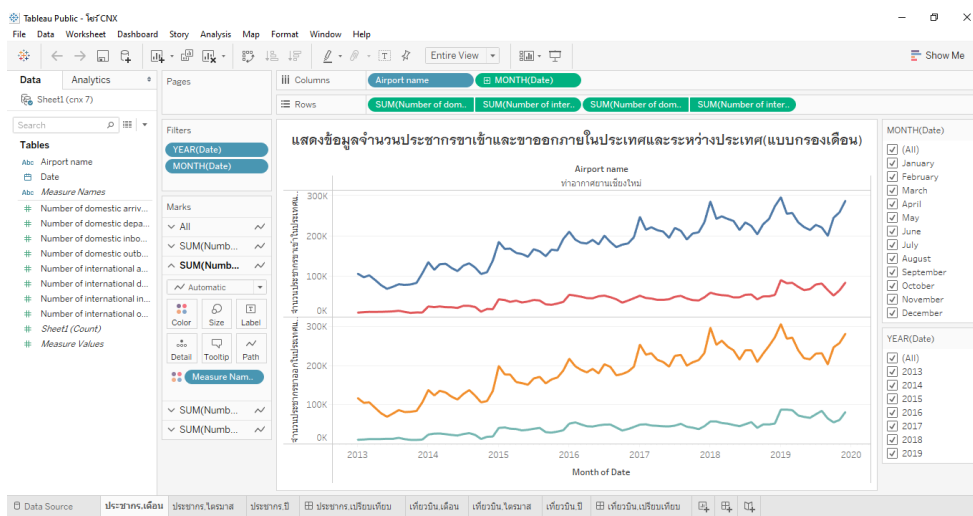
ภาพที่ 3.31 เปิดไฟล์ Data Set จากไฟล์ Excel ด้วยโปรแกรม Tableau

ขั้นตอนที่ 2 insert ข้อมูล Data Set



ภาพที่ 3.32 insert ข้อมูล Data Set ของสายการบินต้นทุนต่ำ เข้าไปในโปรแกรม Tableau

ขั้นตอนที่ 3 ทำการเลือกข้อมูลที่เหมาะสมเพื่อสร้างกราฟเปรียบเทียบแสดงจำนวนเที่ยวบินภายในประเทศของแต่ละท่าอากาศยาน



ภาพที่ 3.33 ข้อมูลเที่ยวบินสายการบินภายในประเทศของสายการบินต้นทุนต่ำ ที่นำมาแสดงในโปรแกรม Tableau

3.2 การออกแบบเว็บไซต์ด้วยโปรแกรม Adobe XD

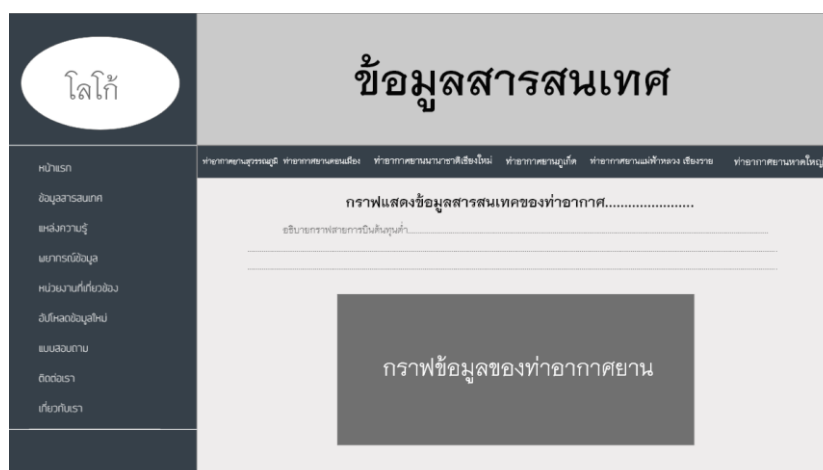
การวางแผนนำเนื้อหาต่าง ๆ มานำเสนอในรูปแบบการจัดทำเว็บไซต์ ซึ่งจะแบ่งเนื้อหาออกเป็นการตกแต่งออกมาเป็นในรูปแบบแถบเมนูซึ่งผู้ใช้งานสามารถใช้งานได้อย่างเข้าใจ และเว็บไซต์ก็มีการอัปเดตข้อมูลใหม่ให้ผู้ใช้งานสามารถนำไปต่อยอดได้

3.2.1 หน้าแรกของเว็บไซต์



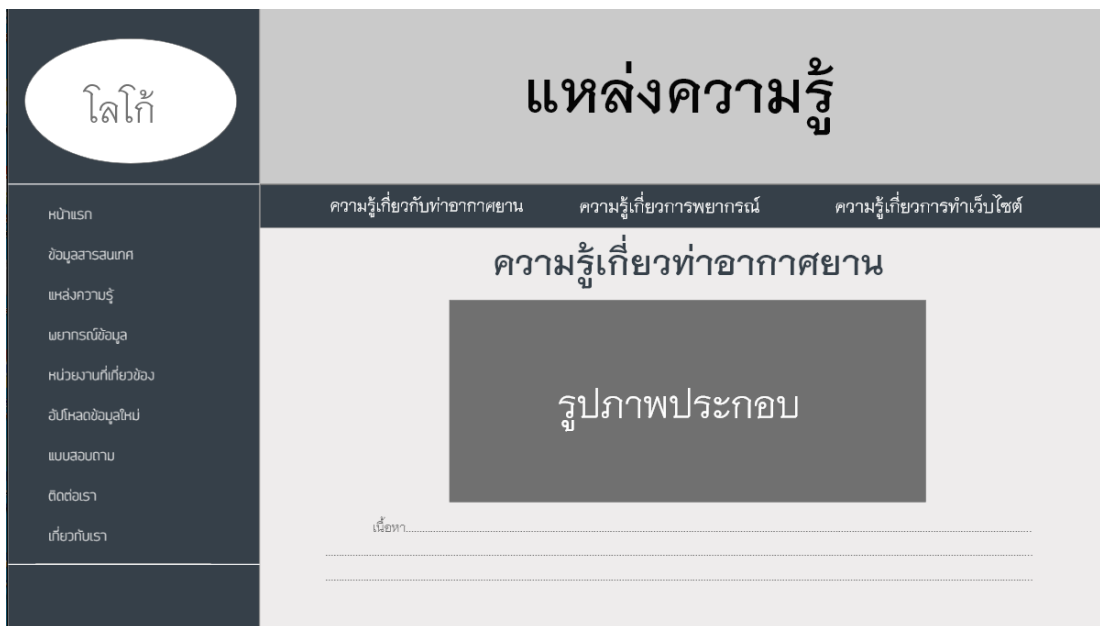
ภาพที่ 3.34 หน้าแรกของเว็บไซต์ที่สามารถลิ้งค์ไปยังแถบเมนูอื่นได้

3.2.2 หน้าเว็บไซต์ที่แสดงถึงข้อมูลสารสนเทศในรูปแบบกราฟ

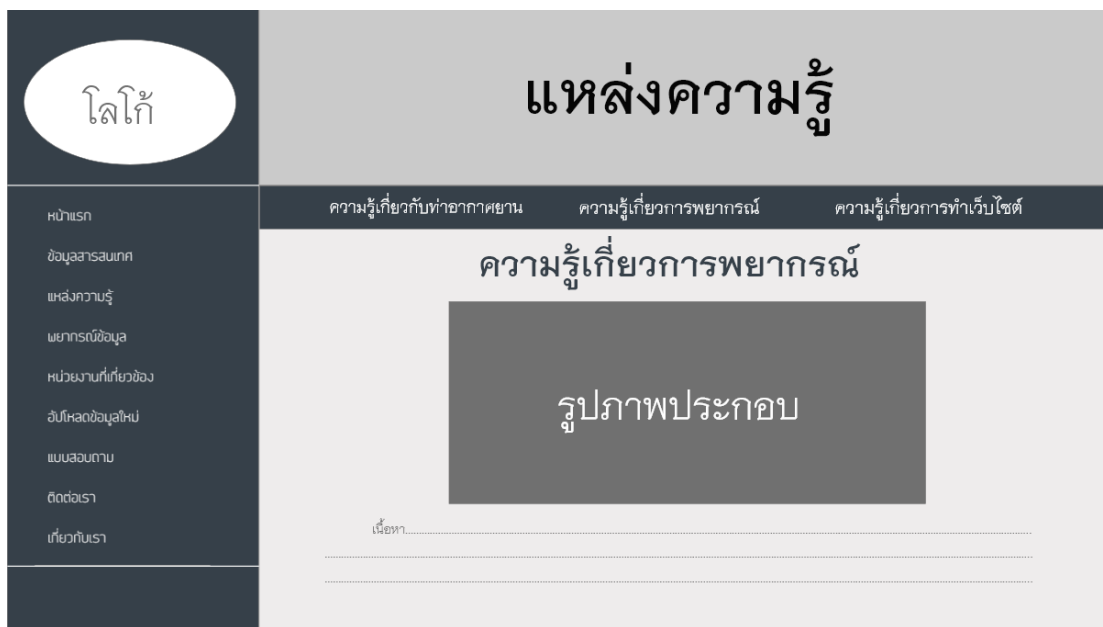


ภาพที่ 3.35 หน้าเว็บไซต์แสดงข้อมูลสารสนเทศในรูปแบบกราฟของท่าอากาศยานสุวรรณภูมิ

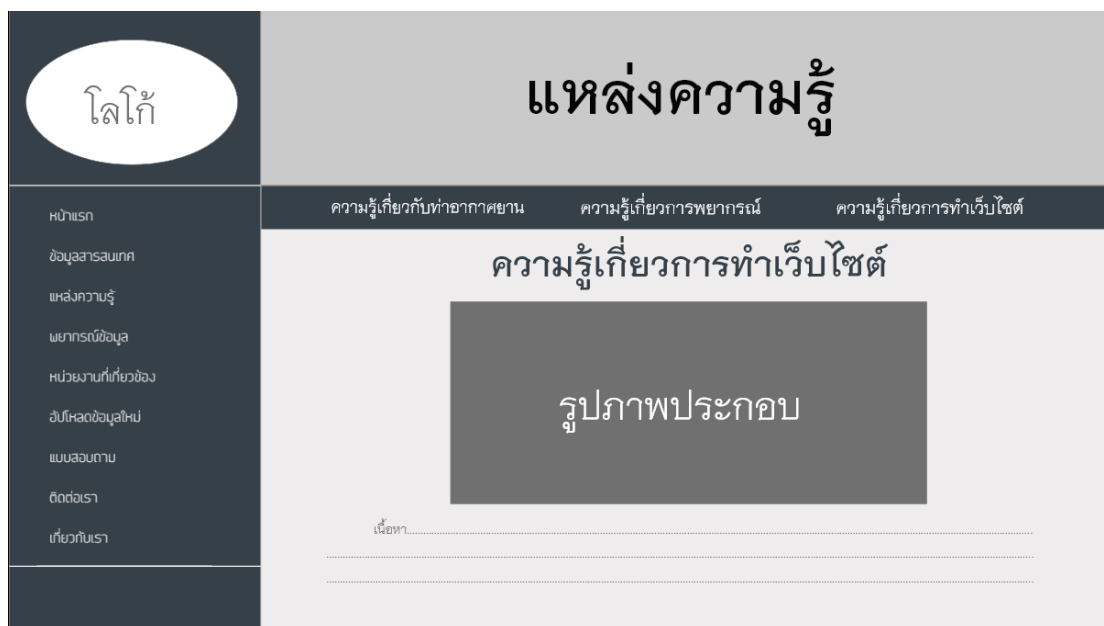
3.2.3 หน้าเว็บไซต์ที่แสดงถึงแหล่งความรู้ที่เกี่ยวกับการทำโครงการ



ภาพที่ 3.36 หน้าเว็บไซต์ที่ประกอบด้วยเนื้อหาของหน่วยงานทำอากาศยานไทย

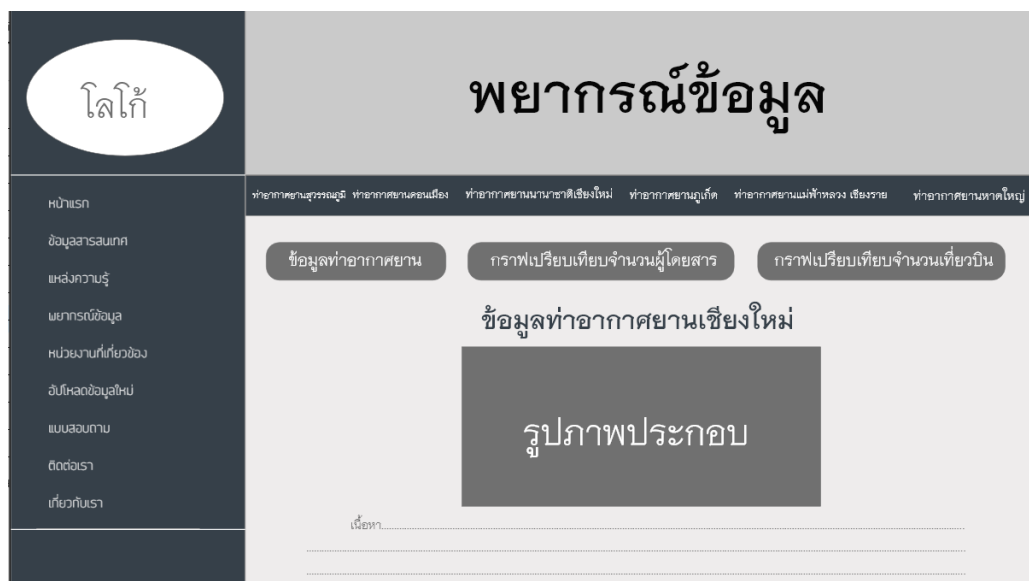


ภาพที่ 3.37 หน้าเว็บไซต์ที่ประกอบด้วยเนื้อหาเกี่ยวกับการพยากรณ์ข้อมูลล่วงหน้า



ภาพที่ 3.38 หน้าเว็บไซต์ที่ประกอบด้วยเนื้อหาการทำเว็บไซต์เบื้องต้น

3.2.4 หน้าเว็บไซต์ที่แสดงข้อมูลกราฟที่ทำการพยากรณ์ข้อมูลล่วงหน้า



ภาพที่ 3.39 หน้าเว็บไซต์ที่ประกอบด้วยเนื้อหาของท่าอากาศยาน



ภาพที่ 3.40 หน้าเว็บไซต์ที่แสดงกราฟพยากรณ์ประชากรของท่าอากาศยานเชียงใหม่



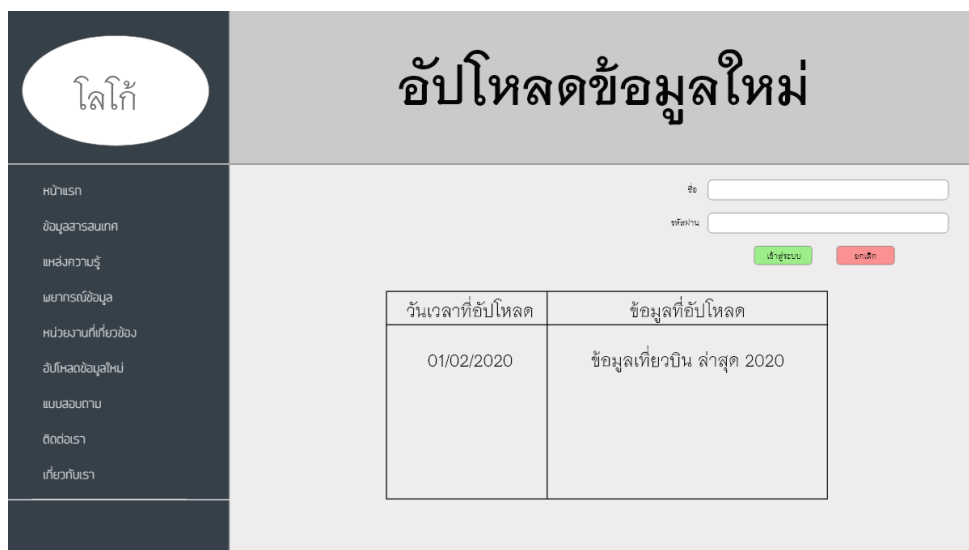
ภาพที่ 3.41 หน้าเว็บไซต์ที่แสดงกราฟพยากรณ์เที่ยวบินของท่าอากาศยานเชียงใหม่

3.2.5 หน้าเว็บไซต์แสดงข้อมูลสารสนเทศเกี่ยวกับหน่วยงานที่เกี่ยวข้องในการศึกษา



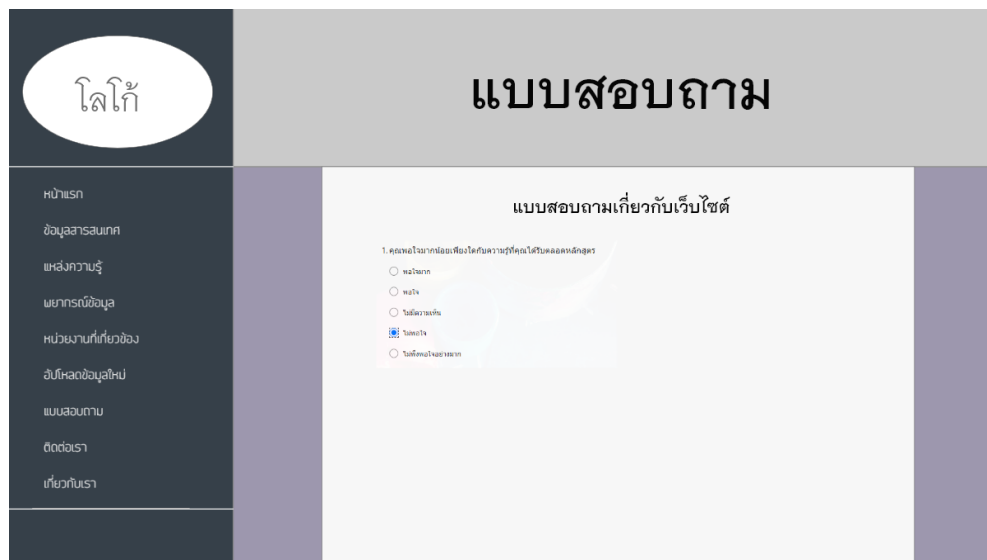
ภาพที่ 3.42 หน้าเว็บไซต์ที่แสดงข้อมูลสารสนเทศเกี่ยวกับหน่วยงานที่เกี่ยวข้อง

3.2.6 หน้าเว็บไซต์แสดงเมนูการอัปเดตข้อมูลใหม่



ภาพที่ 3.43 หน้าเว็บไซต์ที่แสดงแถบเมนูการอัปเดตข้อมูลใหม่

3.2.7 หน้าเว็บไซต์ที่ประกอบด้วยแบบสอบถามเกี่ยวกับเว็บไซต์



ภาพที่ 3.44 หน้าเว็บไซต์แบบสอบถามเกี่ยวกับเว็บไซต์

3.2.8 หน้าเว็บไซต์ที่ประกอบด้วยข้อมูลของเว็บไซต์ ข้อมูลอาจารย์ที่ปรึกษาและประวัติของผู้จัดทำ



ภาพที่ 3.45 หน้าเว็บไซต์ที่ประกอบด้วยข้อมูลของเว็บไซต์

3.3 บทสรุป

จากการดำเนินงานกระบวนการวิเคราะห์ข้อมูลด้วย CRISP-DM พบว่ากระบวนการ CRISP-DM นี้จะประกอบด้วย 6 ขั้นตอน แต่ละขั้นตอนในรูปแบบจะเป็นขั้นตอนที่ต่อเนื่องกันนั่นคือขั้นตอนถัดไปจะรอผลลัพธ์จากขั้นตอนก่อนหน้า ซึ่งกระบวนการ CRISP-DM จะมีด้วยกัน 6 ขั้นตอน คือ (Business Understanding) ผู้วิเคราะห์ข้อมูลทำความเข้าใจปัญหาและ แปลงปัญหาที่ได้ให้อยู่ในรูปแบบโจทย์ของการวิเคราะห์ข้อมูลทางดาต้าไมน์นิ่ง (Data Understanding) ผู้วิเคราะห์ข้อมูลเริ่มทำการเก็บรวบรวมข้อมูลและ ตรวจสอบข้อมูลที่ได้ (Data Preparation) ผู้วิเคราะห์ข้อมูลทำการแปลงข้อมูลที่ได้ทำการเก็บ รวบรวมมาให้กลายเป็นข้อมูลที่สามารถนำไปวิเคราะห์ในขั้นถัดไปได้ (Modeling) ผู้วิเคราะห์ข้อมูลวิเคราะห์ข้อมูลด้วยเทคนิคทางดาต้าไมน์นิ่ง (Classification) ทำการเรียนรู้จากข้อมูลเดิมของการใช้บริการสายการบินต้นทุนต่ำเพื่อให้ได้คำตอบที่ดีที่สุด โดยการใช้โมเดล Time series หาค่าข้อมูลที่เปลี่ยนแปลงไปตามเวลา ในการทำนายผล (Evaluation) ผู้วิเคราะห์ข้อมูลได้ผลการวิเคราะห์ข้อมูลด้วยเทคนิคทางดาต้าไมน์นิ่งจากการสร้างโมเดล Time series สามารถวัดประสิทธิภาพของ ผลลัพธ์ที่ได้ว่าตรงกับวัตถุประสงค์ที่ได้ตั้งไว้ (Deployment) ผู้วิเคราะห์ข้อมูลนำผลองค์ความรู้ที่ได้เหล่านี้ไปเผยแพร่ผ่าน ทางแสดงผลข้อมูลบน Web browser โดยใช้ชุดคำสั่ง HTML และ CSS ร่วมกับการนำเสนอข้อมูลแบบ visualization ด้วยการแสดงผลข้อมูลในรูปแบบ ของภาพโดยใช้โปรแกรม Tableau

บทถัดไปจะกล่าวถึงผลการดำเนินงานทั้งหมดที่ผู้วิเคราะห์ได้จัดทำโครงการขึ้นมา และ อธิบายผลการดำเนินงานต่าง ๆ ซึ่งจะแสดงผลการดำเนินงานในรูปแบบเว็บไซต์เชิงศึกษาข้อมูล พร้อมอธิบายว่าผลการศึกษาผลการทำโครงการนั้นจะเหมือนหรือแตกต่างจากการศึกษาของผู้อื่นหรือไม่ ซึ่งจะกล่าวในบทถัดไป