

โครงงาน

เรื่อง การวิเคราะห์ข้อมูลและพฤติกรรมการซื้อของลูกค้า ในบริษัทการค้า ผ่านการจัดกลุ่มโดยใช้ การแจกแจงผสมแบบเกาส์ (Gaussian mixture model)

จัดทำโดย

นายชนาธิป ศรีนวล เลขทะเบียนนักศึกษา 6309610076

นางสาวพัชรพร มณีฉาย เลขทะเบียนนักศึกษา 6309610084

นายธรณินทร์ แก้วฟู เลขทะเบียนนักศึกษา 6309610126

เสนอ

ผู้ช่วยศาสตราจารย์ ดร. วนิดา พฤทธิวิทยา

โครงงานนี้เป็นส่วนหนึ่งของการศึกษาในรายวิชา

หลักการวิทยาการข้อมูล Principles of Data Science (คพ.245)

คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยธรรมศาสตร์ ศูนย์รังสิต

ภาคเรียนที่ 1 ปีการศึกษา 2654

1. **บทนำ**

ในปัจจุบัน เราปฏิเสธไม่ได้เลยว่าการทำธุรกิจหรือการค้าขายเป็นสิ่งสำคัญและขับเคลื่อนเศรษฐกิจ กลุ่มของพวกเรามีความสนใจในข้อมูลของลูกค้าที่ถูกจัดเก็บเพื่อนำมาวิเคราะห์บุคลิกภาพและพฤติกรรมการใช้จ่ายของลูกค้า เพราะการวิเคราะห์บุคลิกภาพของลูกค้าเป็นการวิเคราะห์โดยละเอียดเกี่ยวกับลูกค้าในอุดมคติของบริษัท ช่วยให้บริษัทหรือธุรกิจเข้าใจลูกค้าได้ดีขึ้น และช่วยให้ปรับเปลี่ยนผลิตภัณฑ์ตามความต้องการเฉพาะ พฤติกรรม ของลูกค้าประเภทต่าง ๆ ได้ง่ายขึ้น และการวิเคราะห์บุคลิกภาพและพฤติกรรมการซื้อของลูกค้ายังช่วยให้ธุรกิจปรับเปลี่ยนผลิตภัณฑ์ตามลูกค้าเป้าหมายจากกลุ่มลูกค้าประเภทต่างๆ ตัวอย่างเช่น แทนที่จะใช้จ่ายเงินเพื่อทำการตลาดผลิตภัณฑ์ใหม่ให้กับลูกค้าทุกรายในฐานข้อมูลของบริษัท บริษัทสามารถวิเคราะห์ได้ว่าลูกค้ากลุ่มใดมีแนวโน้มที่จะซื้อผลิตภัณฑ์มากที่สุด และทำการตลาดผลิตภัณฑ์เฉพาะกลุ่มนั้นเท่านั้น

โดยโครงงานของกลุ่มเรานั้นมุ่งเน้นไปยังการหาคำตอบว่าลูกค้ากลุ่มหลักของบริษัทคือใคร มีพฤติกรรมการใช้จ่ายอย่างไร และข้อมูลที่กลุ่มของเราวิเคราะห์มาได้นั้นจะสามารถนำไปต่อยอดในการวางแผนการตลาดของบริษัทได้หรือไม่ ในการดำเนินงานโครงงานของกลุ่มจึงมีการนำโมเดล Gaussian Mixture Model มาช่วยในการวิเคราะห์กลุ่มลูกค้า

**1.1. วัตถุประสงค์ของโครงงาน**

1.1.1 ทราบขั้นตอนวิธีการนำชุดข้อมูลมาวิเคราะห์ เพื่อให้ได้ซึ่งข้อมูลสำหรับการวางแผนการตลาดเบื้องต้น

1.1.2 ทราบข้อมูลเบื้องต้นของลูกค้าผู้เป็นสมาชิกของบริษัทจากการวิเคราะห์พฤติกรรมการใช้จ่ายของลูกค้าจากชุดข้อมูล

1.1.3 ทราบข้อมูลพฤติกรรมการซื้อของลูกค้าผู้เป็นสมาชิกแต่ละกลุ่ม

1.1.4 ข้อมูลที่ต้องทำการจัดกลุ่มเพื่อนำไปประกอบการวิเคราะห์

1. กลุ่มลูกค้าที่มีรายได้สูง, รายจ่ายสูง และเป็นสมาชิกกับบริษัทนาน ข้อมูลที่ต้องการทราบเกี่ยวกับกลุ่มลูกค้า คือ พฤติกรรมการใช้จ่ายของกลุ่มลูกค้า สัดส่วนของสินค้าที่ลูกค้าซื้อ ช่องทางการซื้อสินค้าของลูกค้า และยอดขายสินค้าจากลูกค้ากลุ่มนี้
2. กลุ่มลูกค้าที่มีรายได้สูง, รายจ่ายสูง และระยะเวลาที่เป็นสมาชิกกับบริษัทสั้น ข้อมูลที่ต้องการทราบเกี่ยวกับกลุ่มลูกค้า คือ พฤติกรรมการใช้จ่ายของกลุ่มลูกค้า สัดส่วนของสินค้าที่ลูกค้าซื้อ ช่องทางการซื้อสินค้าของลูกค้า และยอดขายสินค้าจากลูกค้ากลุ่มนี้
3. กลุ่มลูกค้าที่มีรายได้น้อย, รายจ่ายน้อย และระยะเวลาที่เป็นสมาชิกกับบริษัทสั้น ข้อมูลที่ต้องการทราบเกี่ยวกับกลุ่มลูกค้า คือ พฤติกรรมการใช้จ่ายของกลุ่มลูกค้า สัดส่วนของสินค้าที่ลูกค้าซื้อ ช่องทางการซื้อสินค้าของลูกค้า และยอดขายสินค้าจากลูกค้ากลุ่มนี้
4. กลุ่มลูกค้าที่มีรายได้น้อย, รายจ่ายน้อย และเป็นสมาชิกกับบริษัทนาน ข้อมูลที่ต้องการทราบเกี่ยวกับกลุ่มลูกค้า คือ พฤติกรรมการใช้จ่ายของกลุ่มลูกค้า สัดส่วนของสินค้าที่ลูกค้าซื้อ ช่องทางการซื้อสินค้าของลูกค้า และยอดขายสินค้าจากลูกค้ากลุ่มนี้

**1.2. ประโยชน์ของโครงงาน**

1.2.1. ฝึกการวิเคราะห์ข้อมูลและนำข้อมูลที่วิเคราะห์แล้วไปวางแผนเพื่อประยุกต์ใช้เพื่อสร้างประโยชน์ต่อไป

1.2.2. ได้ศึกษาการทำData Scienceเพื่อวางแผนการตลาดเบื้องต้น

1.2.3. ได้ฝึกการจัดการเวลาและการทำงานกันเป็นทีม

**2. ชุดข้อมูล**

ชุดข้อมูลที่นำมาใช้ประกอบในการทำโครงงานเป็นข้อมูลเกี่ยวกับลูกค้าผู้เป็นสมาชิกของบริษัทหนึ่ง ประกอบด้วยข้อมูลส่วนตัวของลูกค้า ข้อมูลการใช้จ่ายของลูกค้าแยกตามชนิดของสินค้า ช่องทางการซื้อสินค้า และการยอมรับคอมเปญที่ห้างสรรพสินค้าจัดขึ้น โดยชุดข้อมูลที่นำมาทำการศึกษามีแหล่งที่มาจาก https://www.kaggle.com

| คำอธิบายข้อมูลที่จัดเก็บในแต่ละคอลัมน์ของชุดข้อมูล | | |
| --- | --- | --- |
| ลำดับ | ชื่อคอลัมน์ | ความหมาย |
| 1 | ID | รหัสประจำตัวลูกค้า |
| 2 | Year\_Birth | ปีเกิดของลูกค้า |
| 3 | Education | วุฒิการศึกษาของลูกค้า |
| 4 | Marital\_Status | สถานภาพสมรสของลูกค้า |
| 5 | Income | รายได้ในครัวเรือนต่อปีของลูกค้า |
| 6 | Kidhome | จำนวนสมาชิกในครอบครัวที่เป็นเด็กของลูกค้า |
| 7 | Teenhome | จำนวนสมาชิกในครอบครัวที่เป็นวัยรุ่นของลูกค้า |
| 8 | Dt\_Customer | วันที่ลูกค้าสมัครเป็นสมาชิก |
| 9 | Recency | จำนวนวันหลังจากการใช้จ่ายครั้งล่าสุด |
| 10 | Complain | แสดงค่า 1 สำหรับลูกค้าที่มีการcomplainในช่วงสองปีที่ผ่านมา  แสดงค่า 0 สำหรับลูกค้าที่ไม่มีการ complain |
| 11 | MntWines | จำนวนเงินที่ลูกค้าซื้อไวน์ในช่วง 2 ปีที่ผ่านมา |
| 12 | MntFruits | จำนวนเงินที่ลูกค้าซื้อผลไม้ในช่วง 2 ปีที่ผ่านมา |
| 13 | MntMeatProducts | จำนวนเงินที่ลูกค้าซื้อเนื้อในช่วง 2 ปีที่ผ่านมา |
| 14 | MntFishProducts | จำนวนเงินที่ลูกค้าซื้อปลาในช่วง 2 ปีที่ผ่านมา |
| 15 | MntSweetProducts | จำนวนเงินที่ลูกค้าซื้อขนมในช่วง 2 ปีที่ผ่านมา |
| 16 | MntGoldProds | จำนวนเงินที่ลูกค้าซื้อทองในช่วง 2 ปีที่ผ่านมา |
| 17 | NumDealsPurchases | จำนวนการซื้อสินค้าที่มีส่วนลด |
| 18 | AcceptedCmp1 | แสดงค่า 1 สำหรับลูกค้าที่เข้าร่วมแคมเปญครั้งที่1 และ  แสดงค่า 0 สำหรับลูกค้าที่ไม่ได้เข้าร่วม |
| 19 | AcceptedCmp2 | แสดงค่า 1 สำหรับลูกค้าที่เข้าร่วมแคมเปญครั้งที่2 และ  แสดงค่า 0 สำหรับลูกค้าที่ไม่ได้เข้าร่วม |
| 20 | AcceptedCmp3 | แสดงค่า 1 สำหรับลูกค้าที่เข้าร่วมแคมเปญครั้งที่3 และ  แสดงค่า 0 สำหรับลูกค้าที่ไม่ได้เข้าร่วม |
| 21 | AcceptedCmp4 | แสดงค่า 1 สำหรับลูกค้าที่เข้าร่วมแคมเปญครั้งที่4 และ  แสดงค่า 0 สำหรับลูกค้าที่ไม่ได้เข้าร่วม |
| 22 | AcceptedCmp5 | แสดงค่า 1 สำหรับลูกค้าที่เข้าร่วมแคมเปญครั้งที่5 และ  แสดงค่า 0 สำหรับลูกค้าที่ไม่ได้เข้าร่วม |
| 23 | Response | แสดงค่า 1 สำหรับลูกค้าที่รับข้อเสนอในแคมเปญครั้งล่าสุด |
| 24 | NumWebPurchases | จำนวนครั้งการซื้อสินค้าผ่านทางเว็บไซต์ |
| 25 | NumCatalogPurchases | จำนวนครั้งการซื้อสินค้าผ่านทางแคตตาล็อค |
| 26 | NumStorePurchases | จำนวนครั้งการซื้อสินค้าที่ร้าน |
| 27 | NumWebVisitsMonth | จำนวนครั้งการเข้าชมเว็บไซต์ในเดือนที่ผ่านมา |

**3. การวิเคราะห์ข้อมูล**

**ลิงค์สำหรับ Google Colab ในหัวข้อ 3.1, 3.2 และ 3.3 :**

https://colab.research.google.com/drive/15uw1G\_IvR1XdtZAnTI6YUPbGd1qlxT-e?usp=sharing

**ลิงค์สำหรับ Google Colab ในหัวข้อ 3.4 :**

https://colab.research.google.com/drive/1HOltALXIVBzMYnQfZH7B2jq59KCZikDz?usp=sharing

**3.1. ค่าสถิติ คุณลักษณะเบื้องต้น ของชุดข้อมูล**

แสดงคุณลักษณะเบื้องต้นต่างๆของชุดข้อมูลและมีการแสดงค่าสถิติต่างๆเช่น ค่าเฉลี่ย ค่าสูงสุด ค่าต่ำสุด ส่วนเบี่ยงเบนมาตรฐาน เป็นต้น

**3.2. Data Cleansing and Data Validity**

ในการทำcleansing & data validity ได้มีการกระทำดังนี้

1. ตรวจสอบ null ในแต่ละคอลัมน์ของชุดข้อมูล และทำการลบข้อมูลในส่วนของแถวที่มีข้อมูลที่เป็น null
2. เพิ่มคอลัมน์ Age ซึ่งมีข้อมูลอายุของลูกค้า โดยคำนวนจากปีเกิดของลูกค้าในคอลัมน์ Year\_Birth ทำให้พบว่าข้อมูลอายุของลูกค้าบางส่วนไม่สมเหตุสมผล
3. ทำการลบข้อมูลของลูกค้าที่มีอายุไม่สมเหตุสมผลทิ้ง และลบคอลัมน์Year\_Birth
4. ทำการลบคอลัมน์ Z\_CostContact และ Z\_Revenue เนื่องจากแหล่งข้อมูลไม่มีการอธิบายข้อมูลในคอลัมน์นี้
5. ตรวจสอบคอลัมน์ ID ว่ามีข้อมูลในแต่ละแถวที่ซ้ำกันหรือไม่เนื่องจากเป็นคอลัมน์ที่แสดงเลขสมาชิกของลูกค้า เมื่อพบว่าไม่มีข้อมูลที่ซ้ำจึงทำการลบคอลัมน์นี้

**3.3. Data Wrangling**

ในการทำdata wrangling ได้มีการกระทำดังนี้

1) เพิ่มคอลัมน์ Age ข้อมูลในคอลัมน์แสดงอายุของลูกค้า โดยข้อมูลในคอลัมน์นี้อ้างอิงจากคอลัมน์ Year\_Birth และอิมพอร์ตไลบรารี datetime เพื่อช่วยในการคำนวนอายุ

2) ตรวจสอบข้อมูลในคอลัมน์ Marital\_Status พบว่าในคอลัมน์มีการใช้คำที่ต่างกันแต่มีความหมายที่คล้ายคลึงกัน จึงจัดกลุ่มคำที่มีความหมายคล้ายกันไว้ด้วยกัน ดังนี้

1. Divorced, Widow, Alone, Absurd และ YOLO จัดไว้ในกลุ่ม Single
2. Together และ Married จัดไว้ในกลุ่ม Partner

3) จัดกลุ่มข้อมูลการศึกษาในคอลัมน์ Education เป็นสามกลุ่ม ดังนี้

1. Basic และ 2n Cycle จัดไว้ในกลุ่ม Undergraduate
2. Graduation จัดไว้ในกลุ่ม Graduate
3. Master และ PhD จัดไว้ในกลุ่ม Postgraduate

4) เพิ่มคอลัมน์ Spending ข้อมูลในคอลัมน์นี้จะแสดงจำนวนเงินที่ลูกค้าใช้จ่ายซื้อสินค้าจากห้างสรรพสินค้าในสองปีที่ผ่านมา ข้อมูลในคอลัมน์นี้มาจากการรวบรวมคอลัมน์: MntWines, MntFruits, MntMeatProducts, MntFishProducts, MntSweetProduct, และ MntGoldProducts

5) เพิ่มคอลัมน์Time\_With\_Company ข้อมูลในคอลัมน์นี้แสดงจำนวนเดือนที่ลูกค้าเป็นสมาชิกของห้างสรรพสินค้า โดยข้อมูลจากคอลัมน์นี้คำนวนจากคอลัมน์Dt\_Customer

6) เพิ่มคอลัมน์ Children ข้อมูลในคอลัมน์แสดงจำนวนบุตรของลูกค้าที่เป็นสมาชิกโดยข้อมูลจากคอลัมน์นี้คำนวนจากคอลัมน์ Kidhome และ Teenhome นำมารวมกัน

7) เพิ่มคอลัมน์ Has\_Child ข้อมูลในคอลัมน์แสดงว่าลูกค้ามีบุตรหรือไม่โดยแสดงข้อมูลHas Child สำหรับลูกค้าผู้มีบุตร และNo Child สำหรับลูกค้าผู้ไม่มีบุตร คอลัมน์นี้อ้างอิงจากคอลัมน์ Children

8) เพิ่มคอลัมน์ Family\_size ข้อมูลในคอลัมน์แสดงจำนวนสมาชิกทั้งหมดในครอบครัวของลูกค้า โดยข้อมูลในคอลัมน์นี้รวบรวมข้อมูลจากคอลัมน์Marital\_Statu และ Children

9) เพิ่มคอลัมน์ TotalAcceptedCmp ข้อมูลในคอลัมน์แสดงจำนวนการเข้าร่วมแคมเปญทั้งหมดของลูกค้าแต่ละคนภายในระยะเวลา 2 ปีที่ผ่านมา โดยข้อมูลในคอลัมน์นี้มาจากการรวบรวมข้อมูลในคอลัมน์ AcceptedCmp1, AcceptedCmp2, AcceptedCmp3, AcceptedCmp4, และ AcceptedCmp5

10) เปลี่ยนชื่อคอลัมน์เพื่อให้เข้าใจได้ง่ายเมื่อทำการแสดงผล ดังนี้

1. คอลัมน์ MntWines เปลี่ยนชื่อเป็น Wine
2. คอลัมน์ MntFruit เปลี่ยนชื่อเป็น Fruit
3. คอลัมน์ MntMeatProducts เปลี่ยนชื่อเป็น Meat
4. คอลัมน์ MntFishProducts เปลี่ยนชื่อเป็น Fish
5. คอลัมน์ MntSweetProducts เปลี่ยนชื่อเป็น Sweets
6. คอลัมน์ MntGoldProds เปลี่ยนชื่อเป็น Gold
7. คอลัมน์ NumWebPurchases เปลี่ยนชื่อเป็น Web
8. คอลัมน์ NumCatalogPurchases เปลี่ยนชื่อเป็น Catalog
9. คอลัมน์ NumStorePurchases เปลี่ยนชื่อเป็น Store
10. คอลัมน์ NumWebVisitsMonth เปลี่ยนชื่อเป็น WebVisits

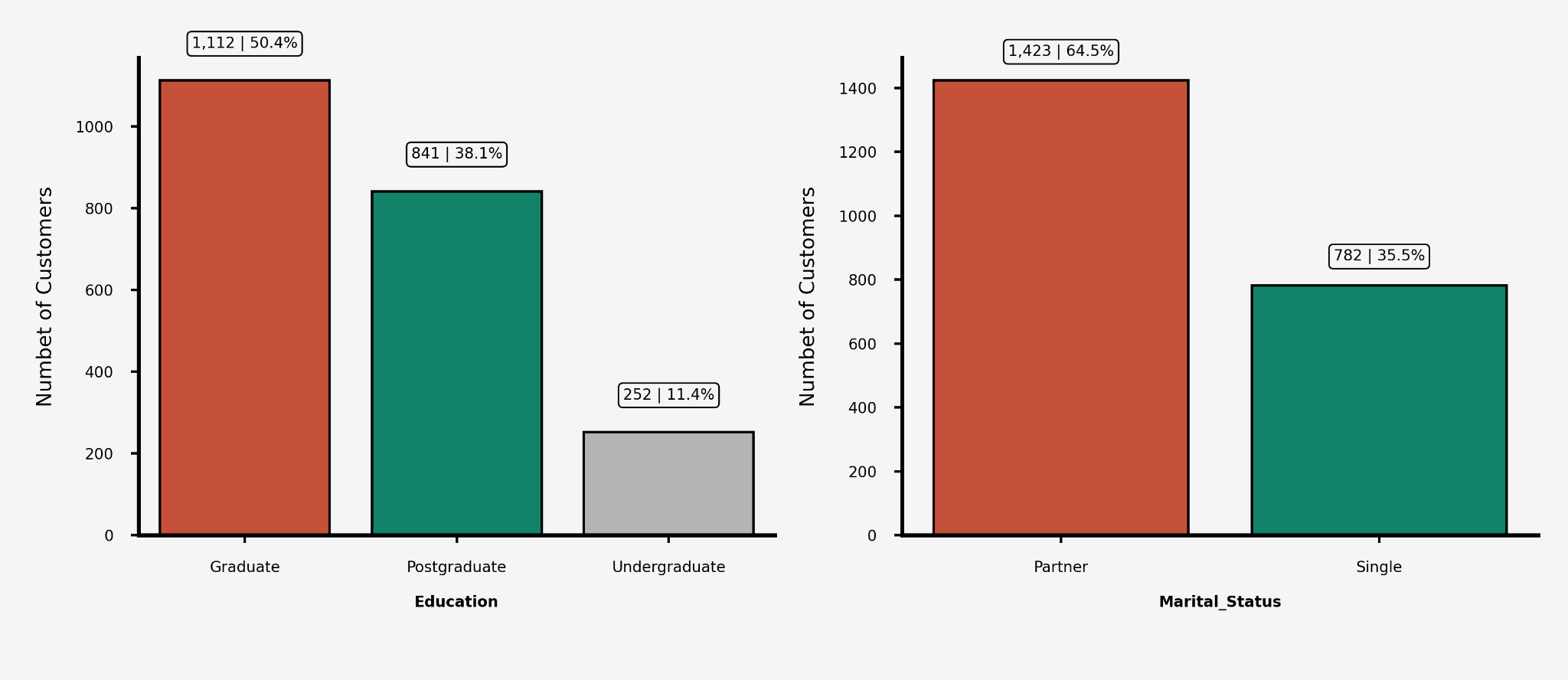
11) นำข้อมูลมาสร้างกราฟเพื่อดูการกระจายของข้อมูล พบว่ามีข้อมูลจำนวนหนึ่งกระจายตัวแยกออกจากกลุ่มข้อมูลส่วนใหญ่ จึงทำการลบข้อมูลในส่วนนั้นเพื่อลดความคลาดเคลื่อนที่อาจจะเกิดขึ้นเมื่อนำข้อมูลไปวิเคราะห์ต่อไป ข้อมูลส่วนที่ได้ทำการลบไปมีดังนี้

1. ข้อมูลในคอลัมน์ Web และ Catalog ที่มีค่าสูงกว่า 20
2. ข้อมูลในคอลัมน์ Income ที่มีค่าสูงกว่า 600,000

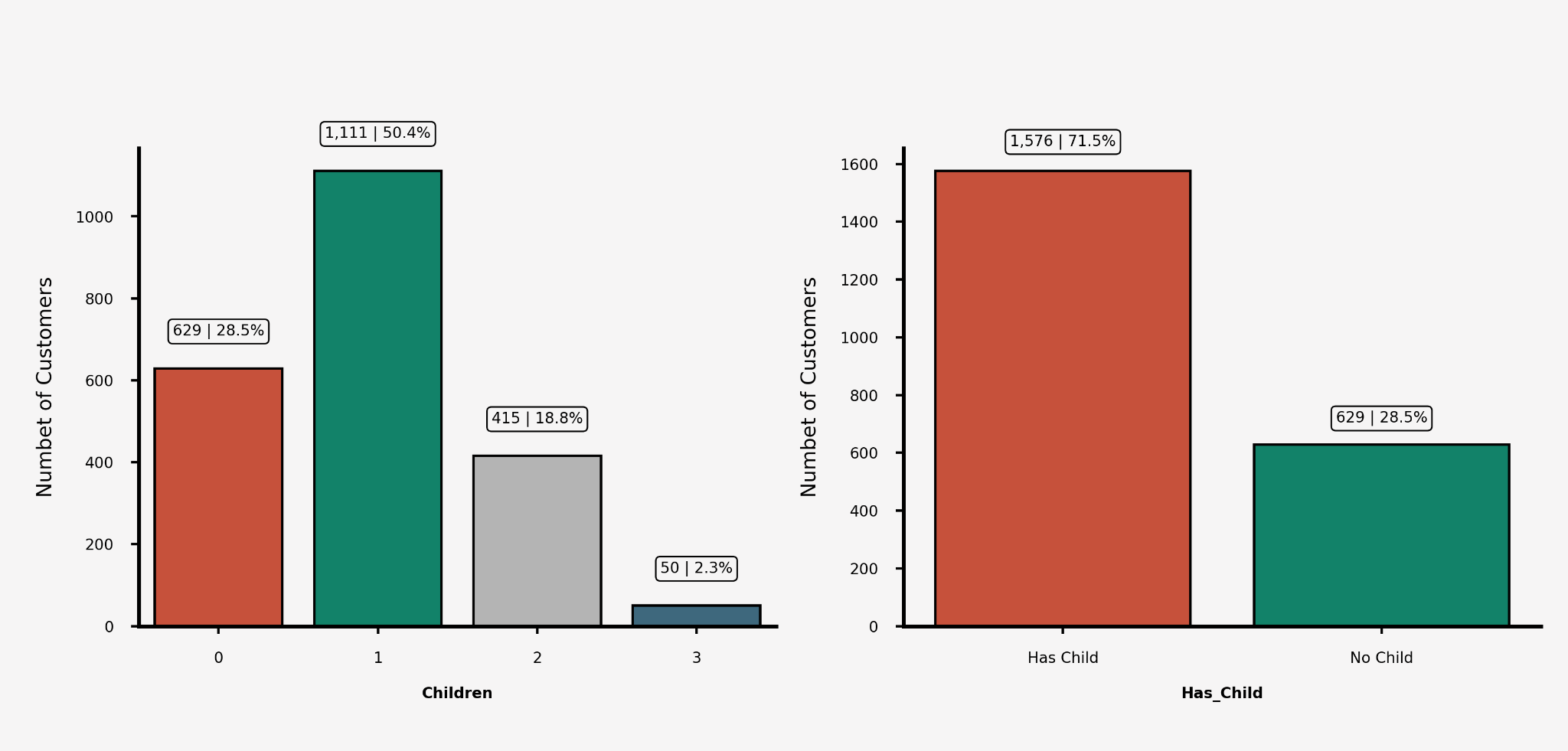
**3.4. Data Visualization**

3.4.1 วิเคราะห์ข้อมูลทั่วไปและพฤติกรรมการซื้อของลูกค้า

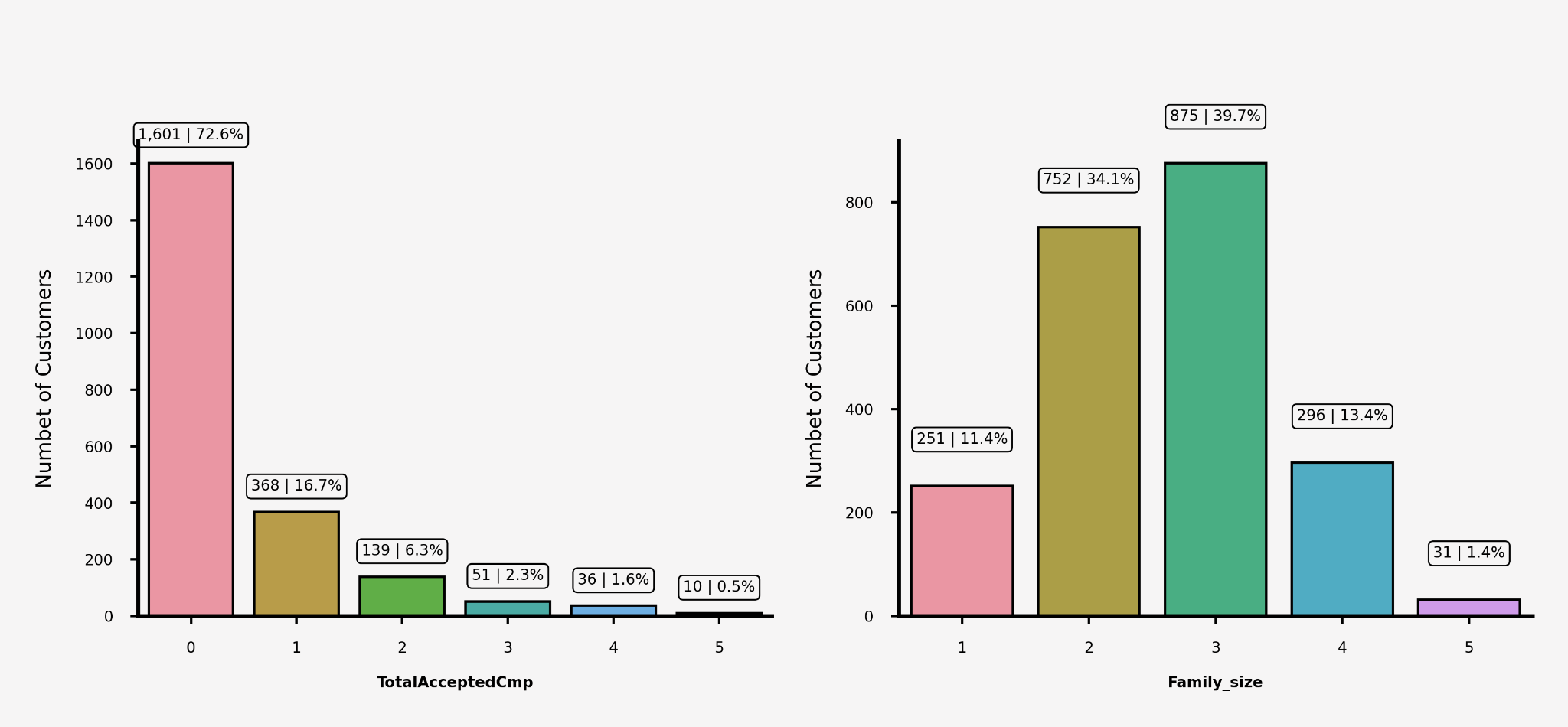
นำข้อมูลที่ผ่านการทำความสะอาด มาใช้ในการวิเคราะห์ข้อมูลทั่วไปของลูกค้า เพื่อที่จะได้ทราบว่าลูกค้าเป็นคนกลุ่มไหนและมีพฤติกรรมการใช้จ่ายอย่างไร วิธีการทำงานคือ แสดง DataFrame และ Plot Graph เพื่อดูข้อมูลทั่วไป ได้แก่ อายุ ระดับการศึกษา สถานภาพสมรส จำนวนสมาชิกภายในครอบครัวของลูกค้า รายได้ และข้อมูลที่เกี่ยวกับพฤติกรรมการซื้อของลูกค้าที่เกี่ยวข้องบริษัท ได้แก่ จำนวนเงินที่ใช้ซื้อของ ระยะเวลาที่เป็นสมาชิก นิยมซื้อสินค้าชนิดใดจำนวนครั้งในการเข้าร่วมแคมเปญ



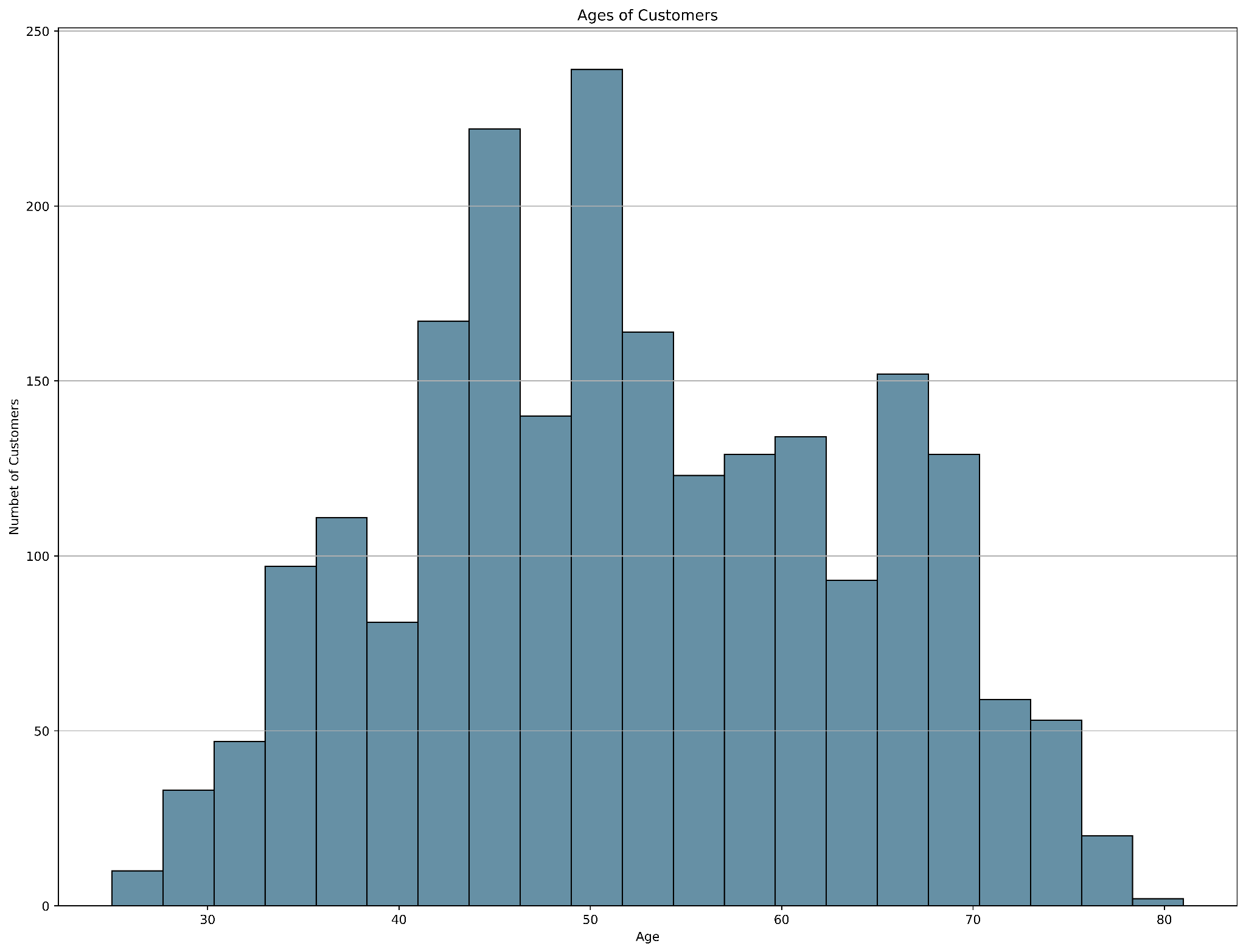
จากกราฟข้างต้นแสดงกลุ่มของลูกค้า โดยแบ่งกลุ่มตามการศึกษา สถานภาพสมรสและจำนวนบุตรของลูกค้า ซึ่งแสดงให้เห็นได้ว่าลูกค้ามากกว่า 50% จบการศึกษาแล้วและ 64.5% ของลูกค้าแต่งงานแล้วแต่ 35.5% ยังไม่แต่งงาน

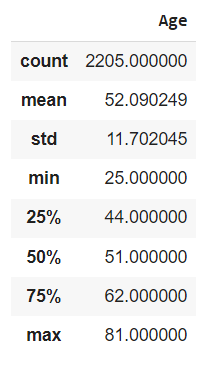


จากกราฟข้างต้นแสดงกลุ่มของลูกค้า โดยแบ่งกลุ่มจำนวนบุตร ซึ่งจากกราฟแสดงให้เห็นว่า 71.5% ของลูกค้ามีบุตรแล้ว แต่อีก 28.5 % ยังไม่ได้มีบุตร และจากจำนวนลูกค้า 1,576 คน ที่มีบุตรแล้ว ลูกค้า 1,111 คน มีบุตรเพียงคนเดียว ลูกค้า 415 คนมีบุตร 2 คน และมีลูกค้าเพียง 50 คน ที่มีบุตร 3 คน

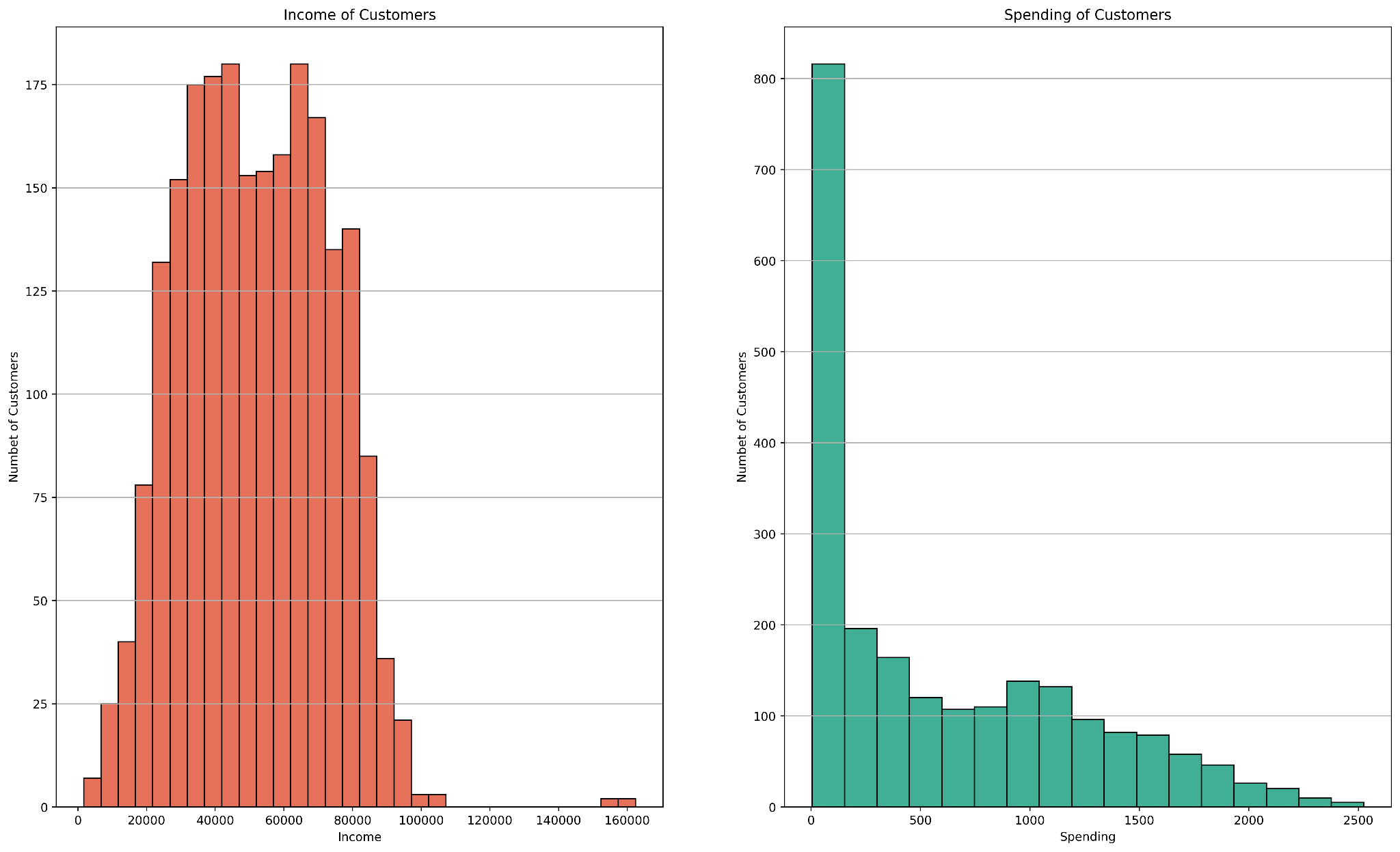


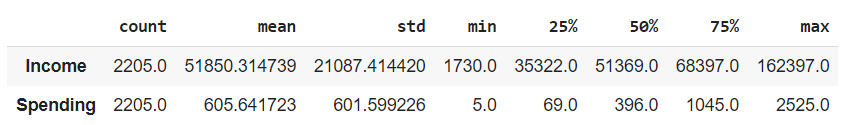
จากกราฟข้างต้นแสดงให้เห็นถึงจำนวนของลูกค้า โดยแบ่งตามจำนวนครั้งที่ลูกค้าเข้าร่วมเคมเปญและแบ่งตามจำนวนสมาชิกภายในครอบครัวของลูกค้า ทำให้เห็นได้ว่าลูกค้ามากกว่าครึ่งไม่เคยเข้าร่วมเคมเปญใดเลย และจำนวนสมาชิกภายในครอบครัวของลูกค้าส่วนใหญ่มีจำนวนสมาชิก 2-3 คน



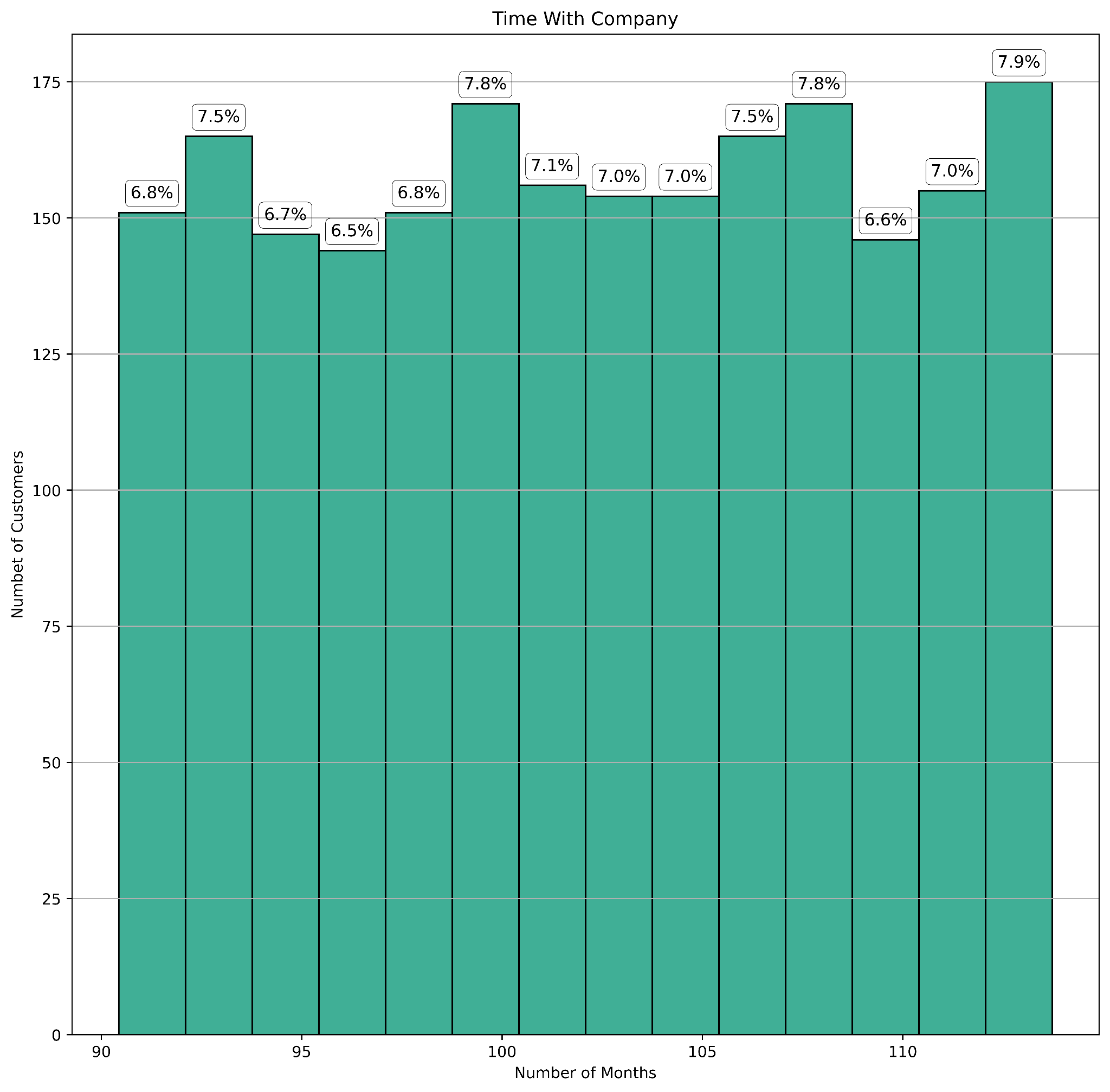


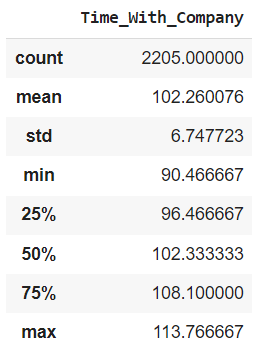
จากกราฟและ DataFrame ข้างต้นแสดงจำนวนของลูกค้าโดยแบ่งตามอายุ ซึ่งจากกราฟและdataframeแสดงให้เห็นได้ว่าลูกค้าที่มีอายุ 50 ปีมีจำนวนมากที่สุด ลูกค้าที่มีอายุน้อยที่สุดมีอายุ 25 ปี ลูกค้าที่มีอายุมากที่สุดมีอายุ 81 ปี และโดยเฉลี่ยแล้วอายุของลูกค้าทั้งหมดอยู่ที่ 52 ปี



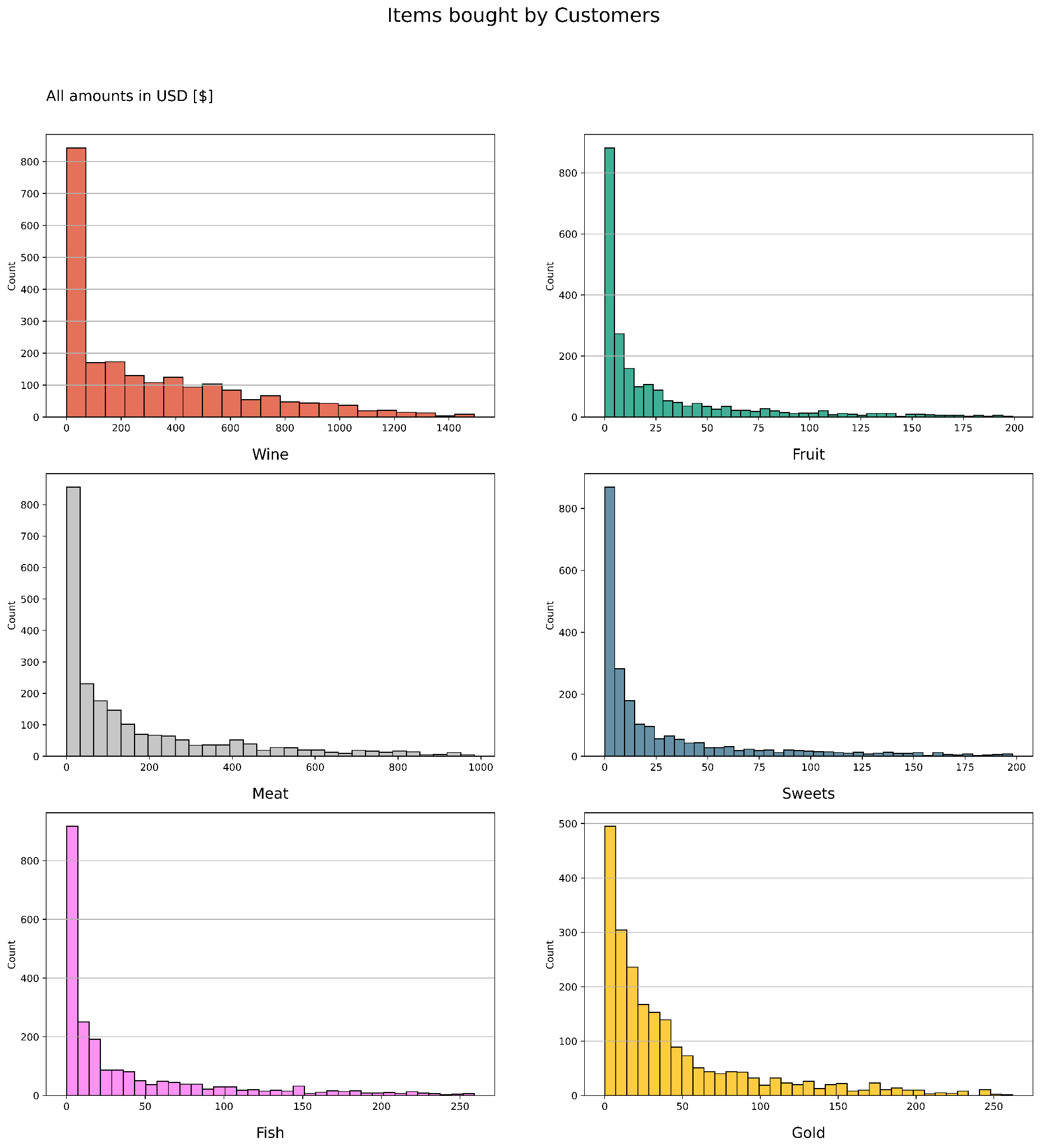


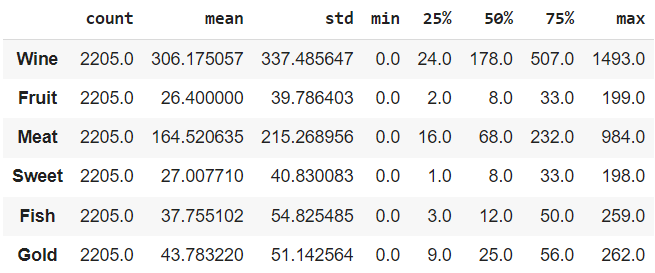
จากกราฟและDataFrameแสดงจำนวนของลูกค้าโดยแบ่งจากรายได้และยอดรายการใช้จ่ายทั้งหมด ซึ่งจากกราฟและdataframe แสดงให้เห็นว่ารายได้สูงสุดของกลุ่มลูกค้าคือ 162,397 USD รายได้เฉลี่ยอยู่ที่ประมาณ 52,000 USD และยอดรายการใช้จ่ายสูงสุดอยู่ที่ 2,525 USD มีค่าเฉลี่ยอยู่ที่ 605 USD





จากกราฟและDataFrameข้างต้นแสดงให้เห็นถึงจำนวนลูกค้าเมื่อแบ่งตามระยะเวลาที่เป็นสมาชิกกับบริษัท โดยระยะเวลาต่ำที่สุดที่ลูกค้าเป็นสมาชิกคือ 90 เดือน และระยะเวลาที่สูงที่สุดคือ 114 เดือน โดยระยะเวลาเฉลี่ยที่ลูกค้าเป็นสมาชิกคือ 102 เดือน



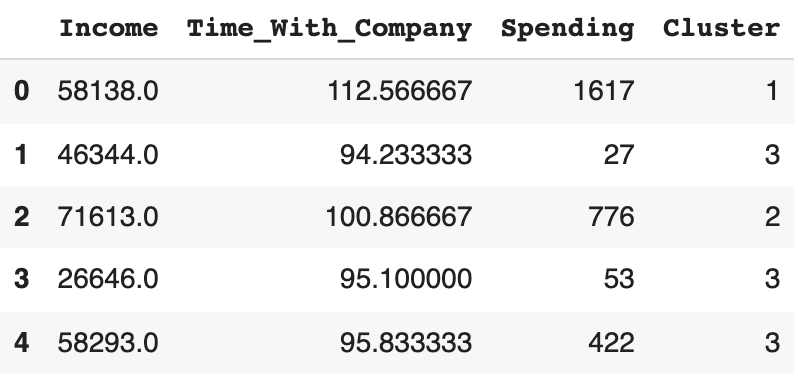


จากกราฟและ DataFrame นี้แสดงถึงจำนวนของสินค้าและจำนวนเงินที่ใช้ซื้อไปของสินค้าโดยแบ่งตามประเภทของสินค้าแต่ละชนิด ทำให้เห็นได้ว่าไวน์เป็นสินค้าที่ลูกค้าใช้เงินเพื่อซื้อมากที่สุด คิดเป็น 50% ของจำนวนเงินที่ลูกค้าใช้ซื้อสินค้าทุกชนิด โดยมีค่าเฉลี่ยอยู่ที่ 306 USD รองลงมาคือเนื้อคิดเป็น 27% ของจำนวนเงินที่ลูกค้าใช้ซื้อสินค้าทุกชนิด มีค่าเฉลี่ยอยู่ที่ 164 USD และสินค้าที่ลูกค้าใช้เงินเพื่อซื้อน้อยที่สุดคือผลไม้ คิดเป็น 4.36% ของจำนวนเงินที่ลูกค้าใช้ซื้อสินค้าทุกชนิด

จากข้อมูลที่ได้กราฟและ DataFrame ทั้งหมดสามารถสรุปได้ว่า ลูกค้าส่วนใหญ่เป็นคนวัยกลางคนที่มีอายุเฉลี่ยประมาณ 52 ปี จบการศึกษาและแต่งงานแล้วเป็นส่วนใหญ่ มีบุตรอย่างน้อยหนึ่งคน มีรายได้เฉลี่ยอยู่ที่ 52,000 USD และทำให้สรุปข้อมูลที่เกี่ยวข้องกับพฤติกรรมการซื้อของลูกค้าได้ว่า ลูกค้าส่วนใหญ่มีค่าใช้จ่ายโดยเฉลี่ยอยู่ที่ 600 USD และเป็นสมาชิกมาแล้วเฉลี่ย 102 เดือน และพบว่าลูกค้าจำนวนมากไม่เคยเข้าร่วมแคมเปญใดเลยสักครั้ง ส่วนสินค้าที่ลูกค้าซื้อส่วนมากจะเป็นไวน์และเนื้อสัตว์

3.4.2. การแบ่งกลุ่มลูกค้า (Clustering)

ทำการแบ่งกลุ่มข้อมูล (Cluster Analysis) เพื่อแบ่งลูกค้าออกเป็นกลุ่ม ๆ ให้ง่ายต่อการวิเคราะห์ โดยใช้โมเดล Gaussian Mixture Model เนื่องจากโมเดลนี้เหมาะกับการแบ่งข้อมูลเป็นคลัสเตอร์ที่มีความแตกต่างกันซึ่งตอบสนองความต้องการของที่มีความต้องการแบ่งกลุ่มลูกค้าเป็น 4 กลุ่มโดยให้น้ำหนักความสำคัญกับข้อมูลจากน้อยไปมาก ได้แก่ รายได้ของลูกค้า ระยะเวลาการที่ลูกค้าเป็นสมาชิกของบริษัท และการใช้จ่าย



จากการใช้งานโมเดล Gaussian Mixture Model ทำให้ได้ข้อมูลของลูกค้าแบ่งเป็นกลุ่ม 4 กลุ่ม ผลลัพธ์ที่แสดงออกมาในขั้นแรกข้อมูลจำแนกแต่ละคลัสเตอร์จะเป็นตัวเลข 0, 1, 2 และ 3 โดยตัวเลขแต่ละจำนวนแบ่งกลุ่มของลูกค้า ดังนี้

กลุ่มที่ 1 คือ กลุ่มลูกค้าที่ลูกค้าผู้ที่มีรายได้สูง, มีการใช้จ่ายสูง และเป็นสมาชิกในระยะเวลานาน

กลุ่มที่ 2 คือ กลุ่มลูกค้าที่ลูกค้าผู้ที่มีรายได้สูง, มีการใช้จ่ายสูง และเป็นสมาชิกในระยะเวลาสั้น

กลุ่มที่ 3 คือ กลุ่มลูกค้าที่ลูกค้าผู้ที่มีรายได้น้อย, มีการใช้จ่ายต่ำ และเป็นสมาชิกในระยะเวลาสั้น

กลุ่มที่ 0 คือ กลุ่มลูกค้าที่ลูกค้าผู้ที่มีรายได้น้อย, มีการใช้จ่ายต่ำ และเป็นสมาชิกในระยะเวลานาน

ต่อมาจึงทำการใช้ชื่อในการกำกับแต่ละกลุ่มแทนจำนวนตัวเลข ดังนี้

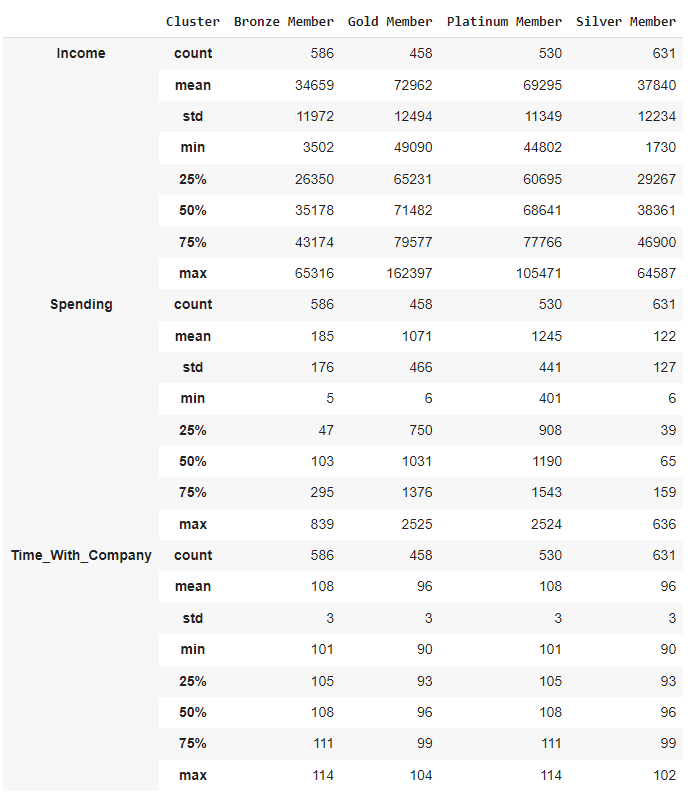
Platinum สำหรับกลุ่มลูกค้าที่ลูกค้าผู้ที่มีรายได้สูง, มีการใช้จ่ายสูง และเป็นสมาชิกในระยะเวลานาน

Gold สำหรับกลุ่มลูกค้าที่ลูกค้าผู้ที่มีรายได้สูง, มีการใช้จ่ายสูง และเป็นสมาชิกในระยะเวลาสั้น

Silver สำหรับกลุ่มลูกค้าที่ลูกค้าผู้ที่มีรายได้น้อย, มีการใช้จ่ายต่ำ และเป็นสมาชิกในระยะเวลาสั้น

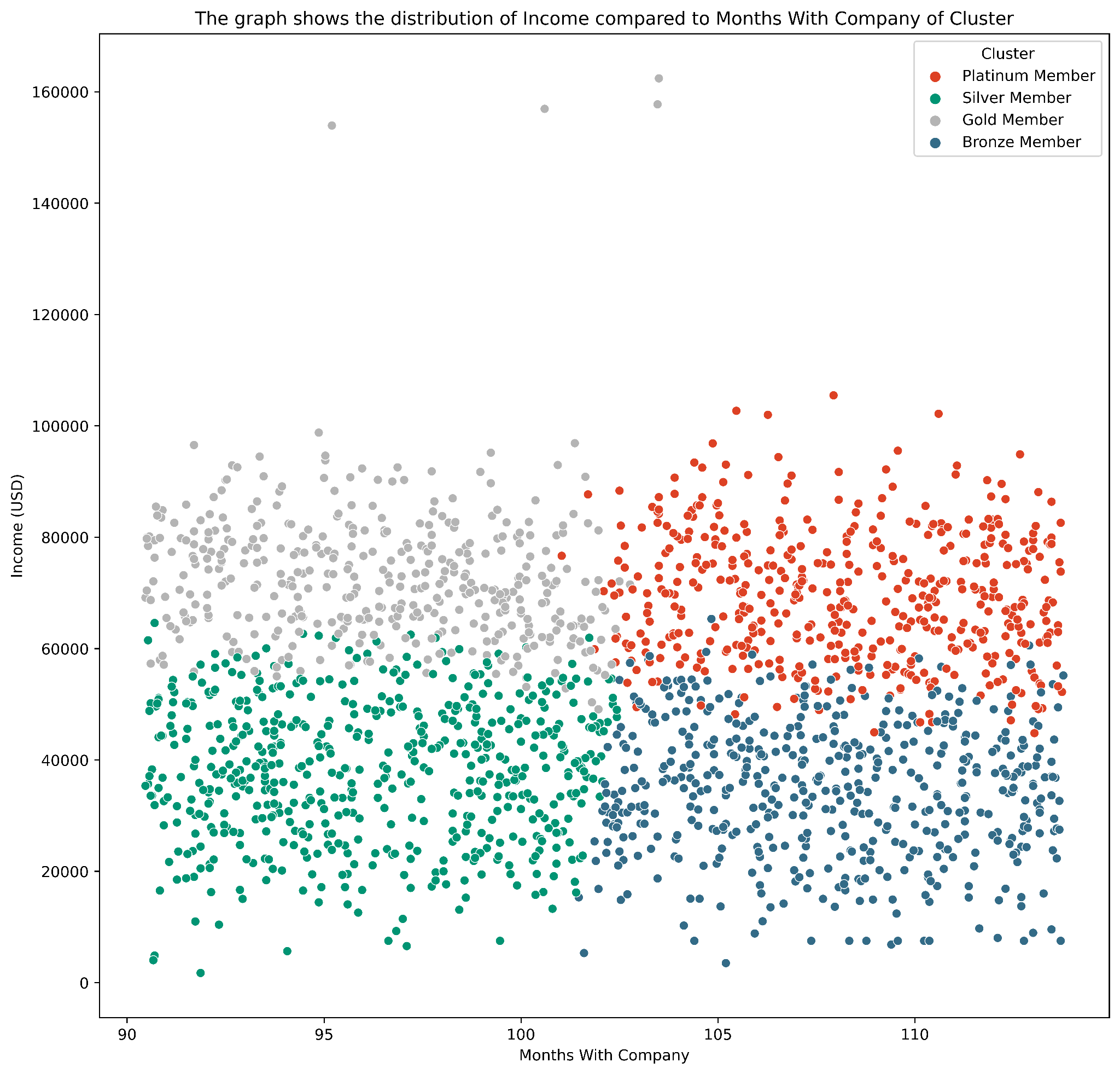
Bronze สำหรับกลุ่มลูกค้าที่ลูกค้าผู้ที่มีรายได้น้อย, มีการใช้จ่ายต่ำ และเป็นสมาชิกในระยะเวลานาน

หลังจากการการจัดกลุ่มการแจกแจงผสมแบบเกาส์ (Gaussian mixture model) และทำการแบ่งคลัสเตอร์ตามที่ได้ข้อมูลที่ได้ให้เงื่อนไขไว้ เพื่อพิสูจน์ว่าเงื่อนไขที่ตั้งไว้สำหรับแต่ละคลัสเตอร์ตรงตามกลุ่มที่ได้แบ่งออกมา จึงจำเป็นต้องทำการแสดงข้อมูลและวิเคราะห์ในส่วนที่ได้ให้เงื่อนไหนไว้ในแต่ละกลุ่ม



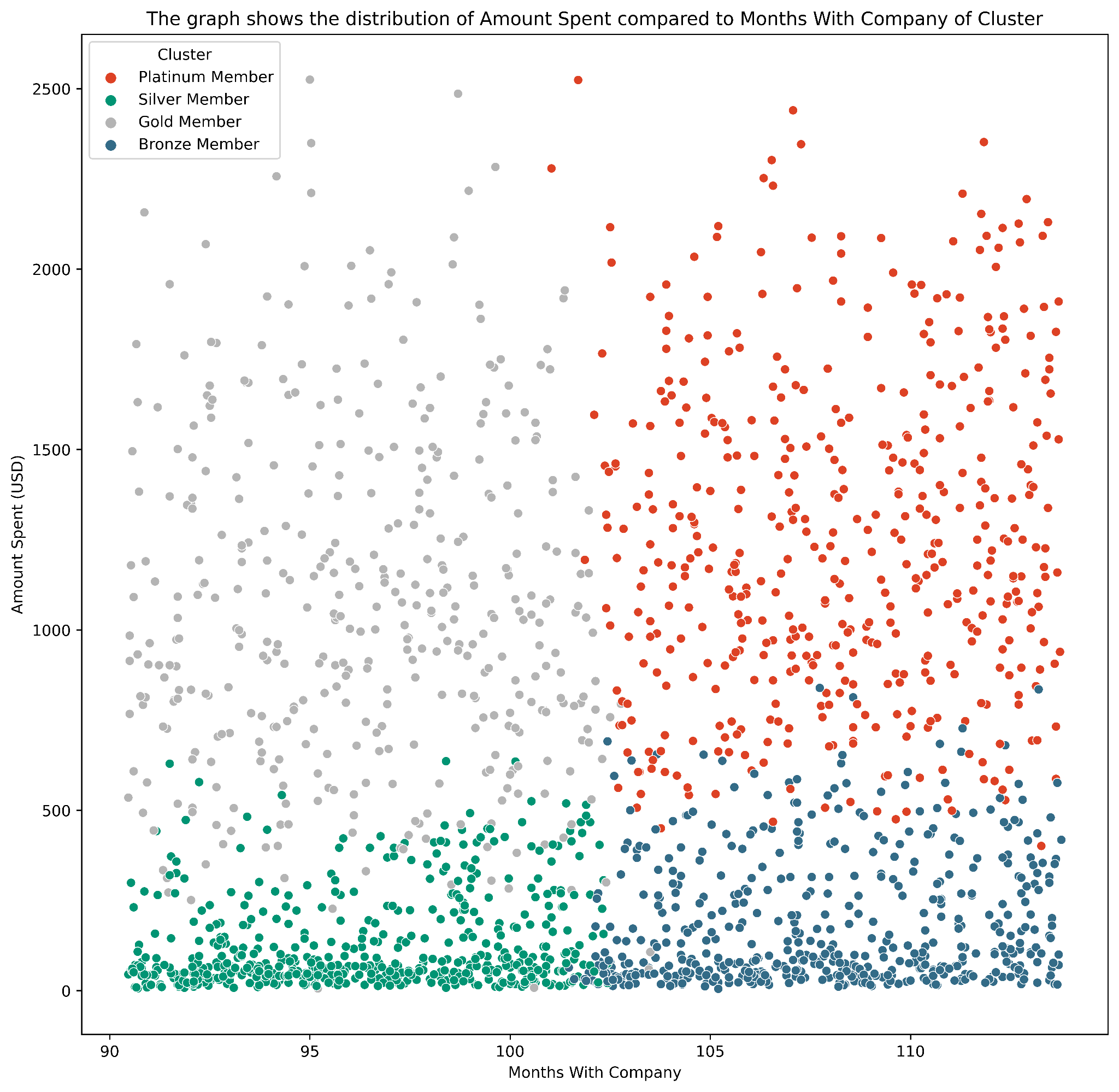


จากกราฟและ DataFrame นี้แสดงให้เห็นถึงรายได้และจำนวนเงินที่ใช้ โดยแบ่งลูกค้าเป็นกลุ่มตามระดับการเป็นสมาชิกทำให้เห็นได้ว่ากลุ่มลูกค้าที่มีรายได้สูงและมีจำนวนเงินที่ใช้สูงจะเป็นกลุ่มลูกค้าที่เป็นสมาชิกระดับ Platinum และ Gold โดยรายได้เฉลี่ยอยู่ที่ 69,295 USD และจำนวนเงินที่ใช้เฉลี่ยอยู่ที่ 1,245 USD สำหรับลูกค้าระดับ Platinum รายได้เฉลี่ยอยู่ที่ 72,962 USDและจำนวนเงินที่ใช้เฉลี่ยอยู่ที่ 1,071 USD สำหรับลูกค้าระดับ Gold และลูกค้าที่เป็นสมาชิกระดับ Silver และ Bronze จะมีรายได้และจำนวนเงินที่ใช้ลดลงมาสำหรับลูกค้าระดับ Silver จะมีรายได้โดยเฉลี่ยที่ 37,840 USDและมีจำนวนเงินที่ใช้เฉลี่ยอยู่ที่ 122 USD สำหรับลูกค้าระดับ Bronze จะมีรายได้โดยเฉลี่ยที่ 34,659 USD และมีจำนวนเงินที่ใช้เฉลี่ยอยู่ที่ 185 USD

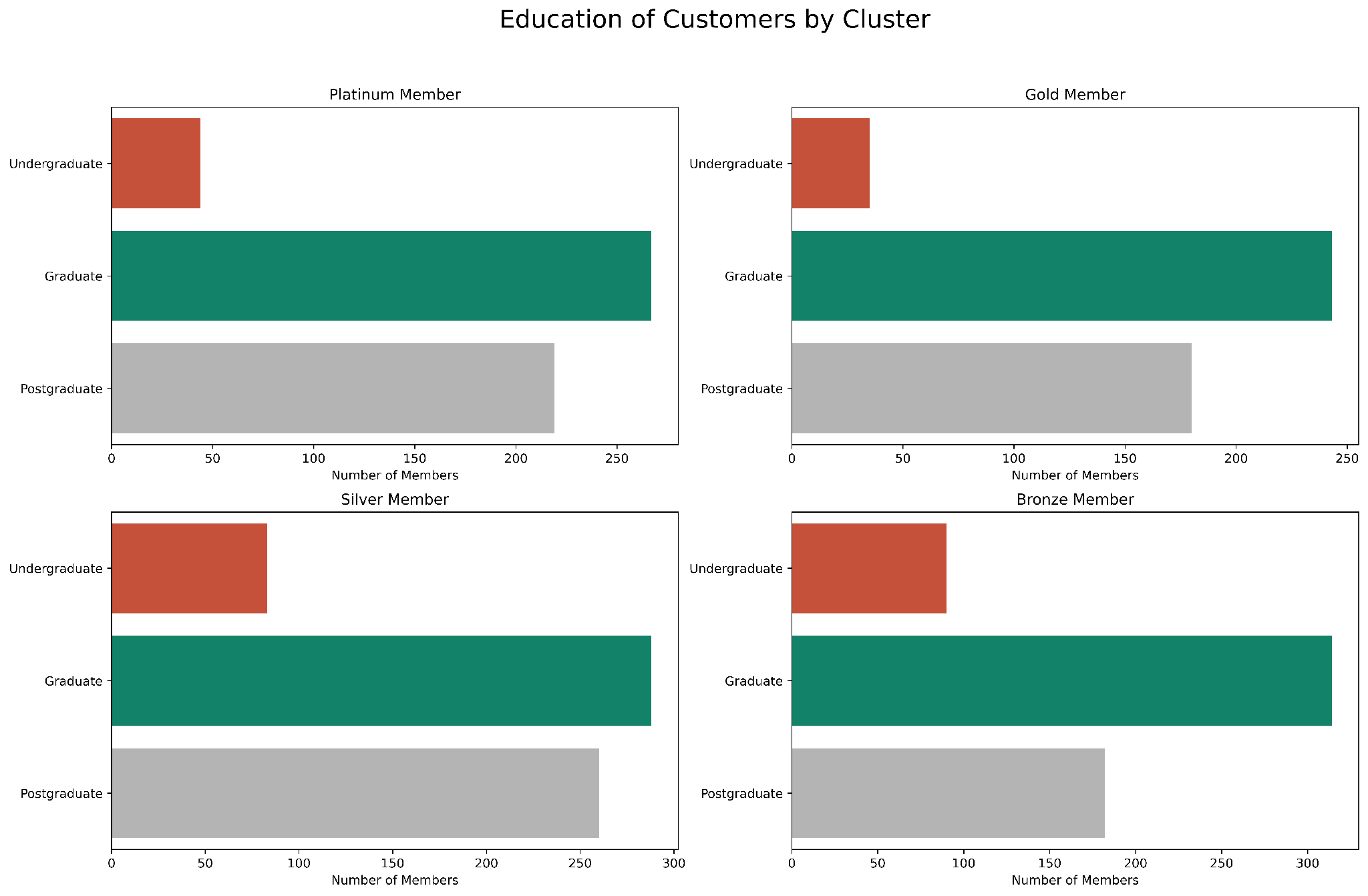


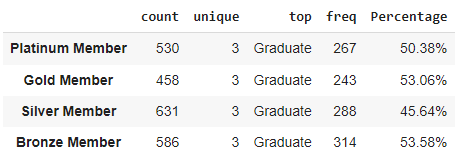
จากกราฟและ DataFrame นี้แสดงให้เห็นถึงระยะเวลาการเป็นสมาชิกกับบริษัทและรายได้ โดยแบ่งลูกค้าเป็นกลุ่มตามระดับการเป็นสมาชิก ทำให้เห็นได้ว่าลูกค้าที่เป็นสมาชิกระดับ Platinum และ Gold เป็นกลุ่มลูกค้าที่มีรายได้สูงแต่ลูกค้าระดับ Platinum จะมีระยะการเป็นสมาชิกกับบริษัทนานกว่า โดยระยะเวลาที่เป็นสมาชิกเฉลี่ยคือ 108 เดือน ในขณะที่ลูกค้าระดับบ Gold จะมีระยะเวลาการเป็นสมาชิกกับบริษัทน้อยกว่า โดยระยะเวลาเฉลี่ยคือ 96 เดือน

ส่วนลูกค้าระดับ Silver และ Bronze จะเป็นกลุ่มลูกค้าที่มีรายได้ต่ำกว่า แต่ลูกค้าระดับ Bronze จะมีระยะการเป็นสมาชิกกับบริษัทนานกว่า โดยระยะเวลาที่เป็นสมาชิกเฉลี่ยคือ 108 เดือน ในขณะที่ลูกค้าระดับบ Silver จะมีระยะเวลาการเป็นสมาชิกกับบริษัทน้อยกว่า โดยระยะเวลาเฉลี่ยคือ 96 เดือน



จากกราฟและ DataFrame นี้แสดงให้เห็นถึงระยะเวลาการเป็นสมาชิกกับบริษัทและการใช้จ่าย โดยแบ่งลูกค้าเป็นกลุ่มตามระดับการเป็นสมาชิก ทำให้เห็นได้ว่าลูกค้าที่เป็นสมาชิกระดับ Platinum และ Gold เป็นกลุ่มลูกค้าที่มีการใช้จ่ายสูงแต่ลูกค้าระดับ Platinum จะมีระยะการเป็นสมาชิกกับบริษัทนานกว่า โดยระยะเวลาที่เป็นสมาชิกเฉลี่ยคือ 108 เดือน ในขณะที่ลูกค้าระดับบ Gold จะมีระยะเวลาการเป็นสมาชิกกับบริษัทน้อยกว่า โดยระยะเวลาเฉลี่ยคือ 96 เดือนส่วนลูกค้าระดับ Silver และ Bronze จะเป็นกลุ่มลูกค้าที่มีการใช้จ่ายต่ำกว่า แต่ลูกค้าระดับ Bronze จะมีระยะการเป็นสมาชิกกับบริษัทนานกว่า โดยระยะเวลาที่เป็นสมาชิกเฉลี่ยคือ 108 เดือน ในขณะที่ลูกค้าระดับบ Silver จะมีระยะเวลาการเป็นสมาชิกกับบริษัทน้อยกว่า โดยระยะเวลาเฉลี่ยคือ 96 เดือน



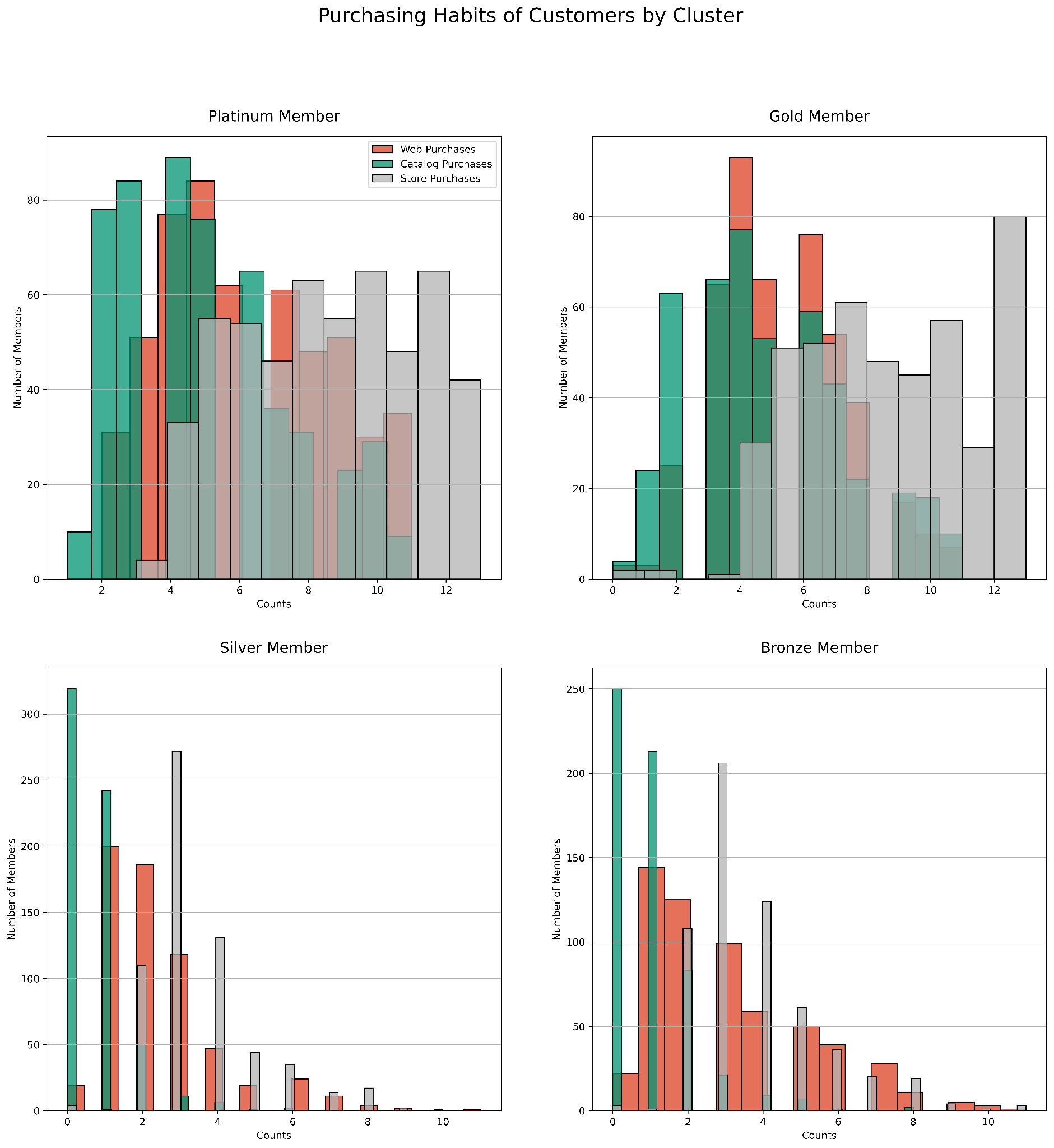


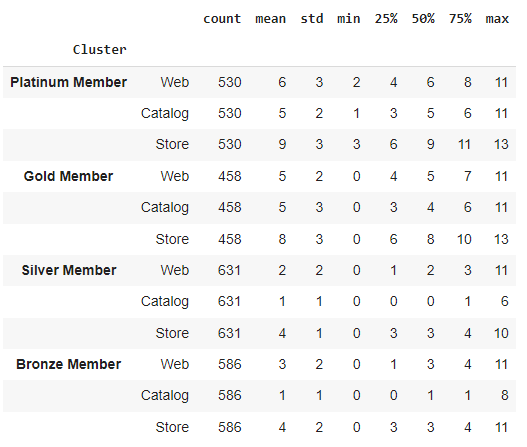
จากกราฟและDataFrame แสดงให้เห็นถึงระดับการศึกษาของลูกค้าโดยแบ่งตามระดับการเป็นสมาชิก ซึ่งทำให้เห็นได้ว่าลูกค้าที่เป็นสมาชิกในทุกระดับนั้นส่วนมากจบการศึกษาแล้ว โดยลูกค้าระดับ Platinum มีจำนวนลูกค้าที่จบการศึกษา 50% ลูกค้าระดับ Gold มีจำนวนลูกค้าที่จบการศึกษา 53% ลูกค้าระดับ Silver มีจำนวนลูกค้าที่จบการศึกษา 46% และลูกค้าระดับ Bronze มีจำนวนลูกค้าที่จบการศึกษา 56%

จากข้อมูลที่ได้จากกราฟและ DataFrame ทั้งหมดที่กล่าวสามารถสรุปได้ว่า ภายในกลุ่มลูกค้าทั้ง 4 กลุ่มนั้น ลูกค้าระดับ Platinum และ Gold เป็นกลุ่มลูกค้าที่มีรายได้และการใช้จ่ายที่สูง แต่ลูกค้าระดับ Platinum มีระยะเวลาการเป็นสมาชิกมานานกว่าลูกค้าระดับ Gold ส่วนลูกค้าระดับ Silver และ Bronze เป็นกลุ่มลูกค้าที่มีรายได้และการใช้จ่ายที่น้อยกว่า แต่ลูกค้าระดับ Bronze มีระยะเวลาการเป็นสมาชิกมานานกว่าลูกค้าระดับ Silver และในภาพรวมของกลุ่มลูกค้าทั้ง 4 กลุ่ม ลูกค้าส่วนใหญ่นั้นจบการศึกษาแล้ว ซึ่งตรงตามเงื่อนไขของการแบ่งคลัสเตอร์ที่ได้นิยามไว้

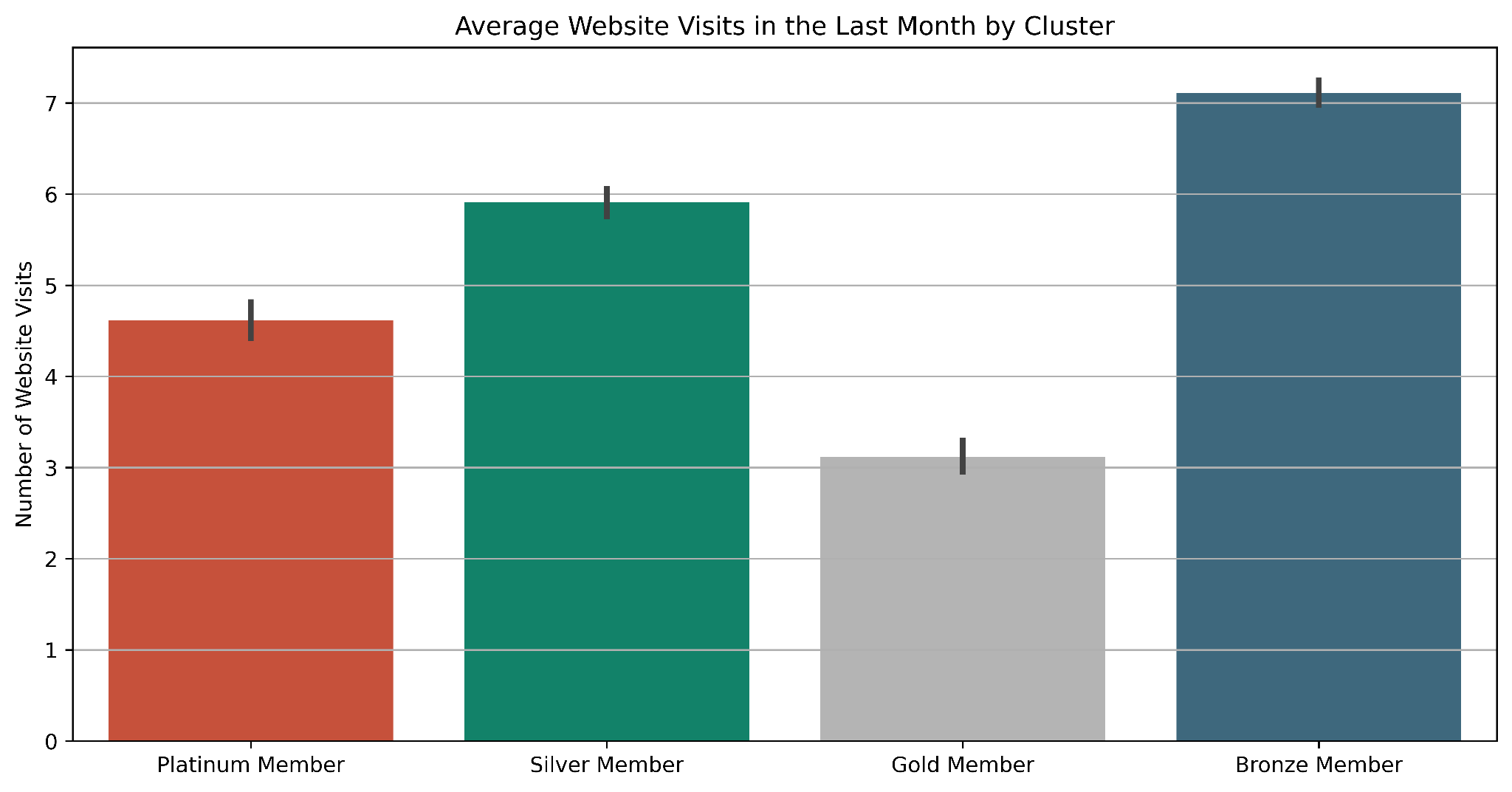
3.4.3 พฤติกรรมการซื้อของลูกค้าแต่ละกลุ่ม

เมื่อทำการแบ่งกลุ่มของลูกค้าได้เป็น 4 กลุ่ม และแสดง DataFrame และ Plot Graph เพื่อทำการวิเคราะห์พฤติกรรมการซื้อของลูกค้าแต่ละกลุ่ม โดยวิเคราะห์แหล่งที่ลูกค้าเลือกซื้อสินค้า ได้แก่ ซื้อจากเว็บไซต์ ซื้อจากแคตตาล็อก หรือซื้อจากร้านค้า และวิเคราะห์จากชนิดของสินค้าที่ลูกค้าแต่ละกลุ่มเลือกซื้อ ได้แก่ ไวน์ เนื้อ ผลไม้ และขนม

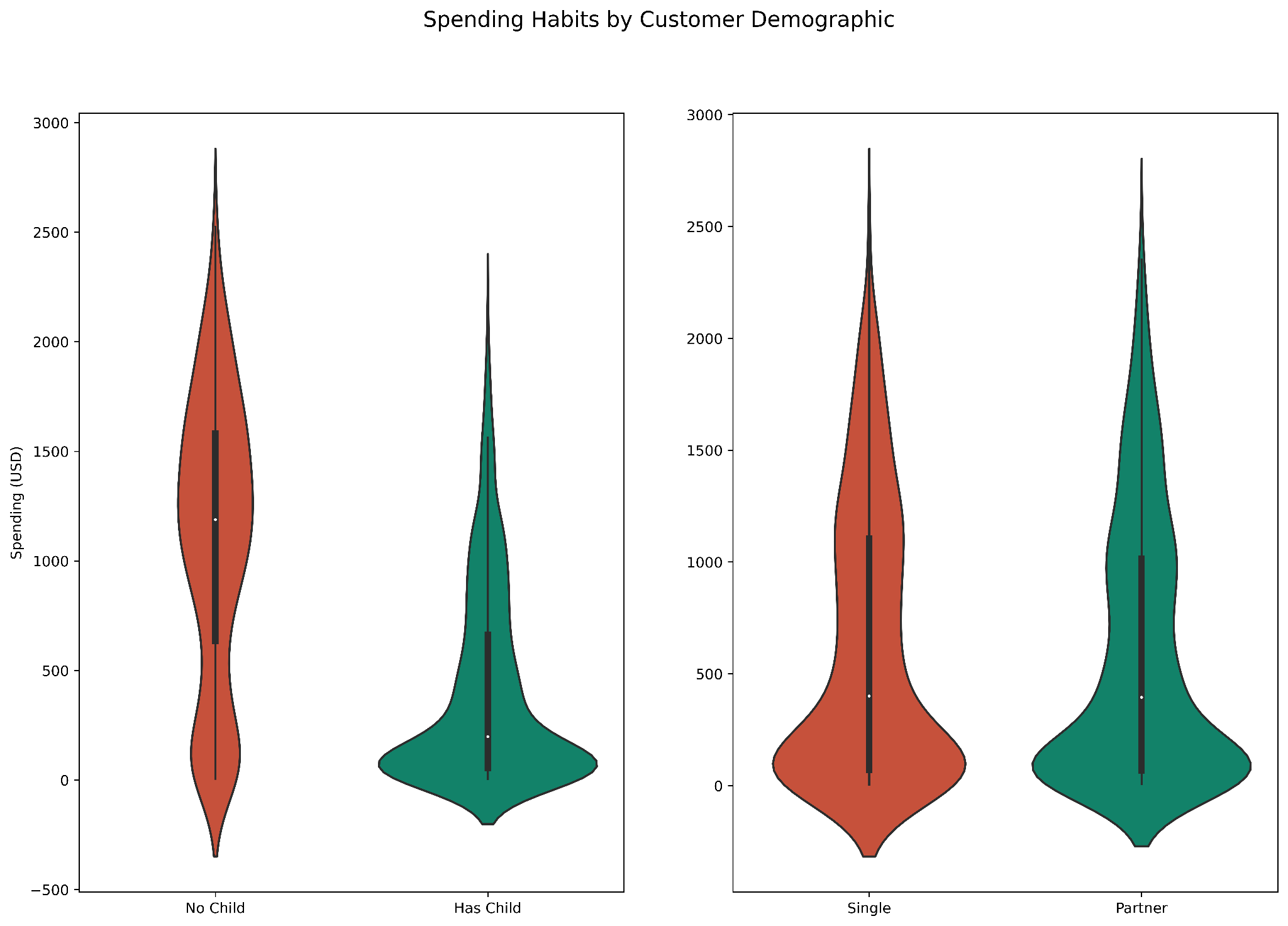




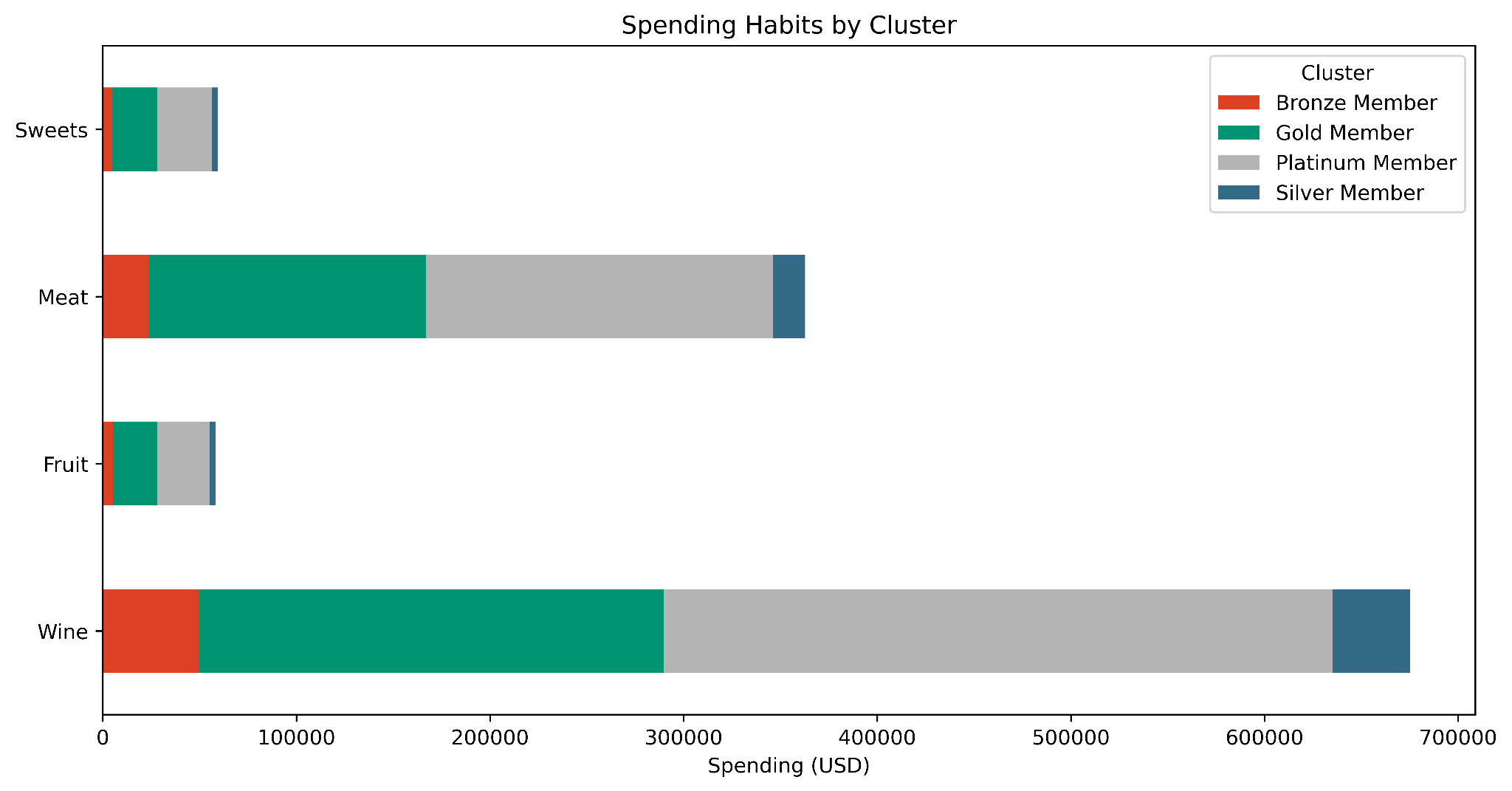
จากกราฟและ DataFrameนี้แสดงให้เห็นว่า ลูกค้าระดับ Platinum มีการซื้อของจากร้านค้ามากที่สุด โดยมีค่าเฉลี่ยอยู่ที่ 9 ครั้ง ลูกค้าระดับ Gold มีการซื้อของจากร้านค้ามากที่สุดเช่นกันโดยมีค่าเฉลี่ยอยู่ที่ 8 ครั้ง ลูกค้าระดับ Silver มีการซื้อของจากเว็บไซต์มากที่สุดจำนวน 11 ครั้ง แต่ซื้อของร้านค้ามีค่าเฉลี่ยมากกว่าอยู่ที่ 4 ครั้ง ลูกค้าระดับ Bronze จำนวนครั้งในการซื้อของจากเว็บไซต์และจากร้านค้ามีจำนวนเท่ากันคือ 11 ครั้ง แต่ซื้อของร้านค้ามีค่าเฉลี่ยมากกว่าอยู่ที่ 4 ครั้ง

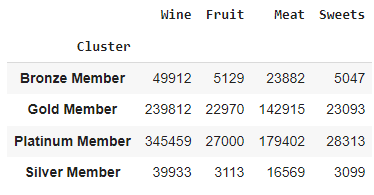


จากกราฟนี้แสดงให้เห็นถึงค่าเฉลี่ยของการเข้าชมเว็บไซต์โดยแบ่งตามระดับสมาชิกของลูกค้า ซึ่งแสดงให้เห็นว่าค่าเฉลี่ยของการเข้าชมเว็บไซต์ของลูกค้าระดับ Bronze มีค่ามากที่สุดประมาณ 7 ครั้ง รองลงมาคือลูกค้าระดับ Silver ที่มีค่าเฉลี่ยของการเข้าชมเว็บไซต์ประมาณ 6 ครั้ง ลูกค้าระดับ Platinum มีค่าเฉลี่ยของการเข้าชมเว็บไซต์ประมาณ 4-5 ครั้ง และลูกค้าระดับ Gold มีค่าเฉลี่ยของการเข้าชมเว็บไซต์ของน้อยที่สุดประมาณ 3 ครั้ง และจากกราฟนี้ทำให้เราเห็นได้ว่าลูกค้าระดับ Bronze ที่มีการใช้จ่ายน้อยที่สุดมีการเข้าชมเว็บไซต์มากที่สุด

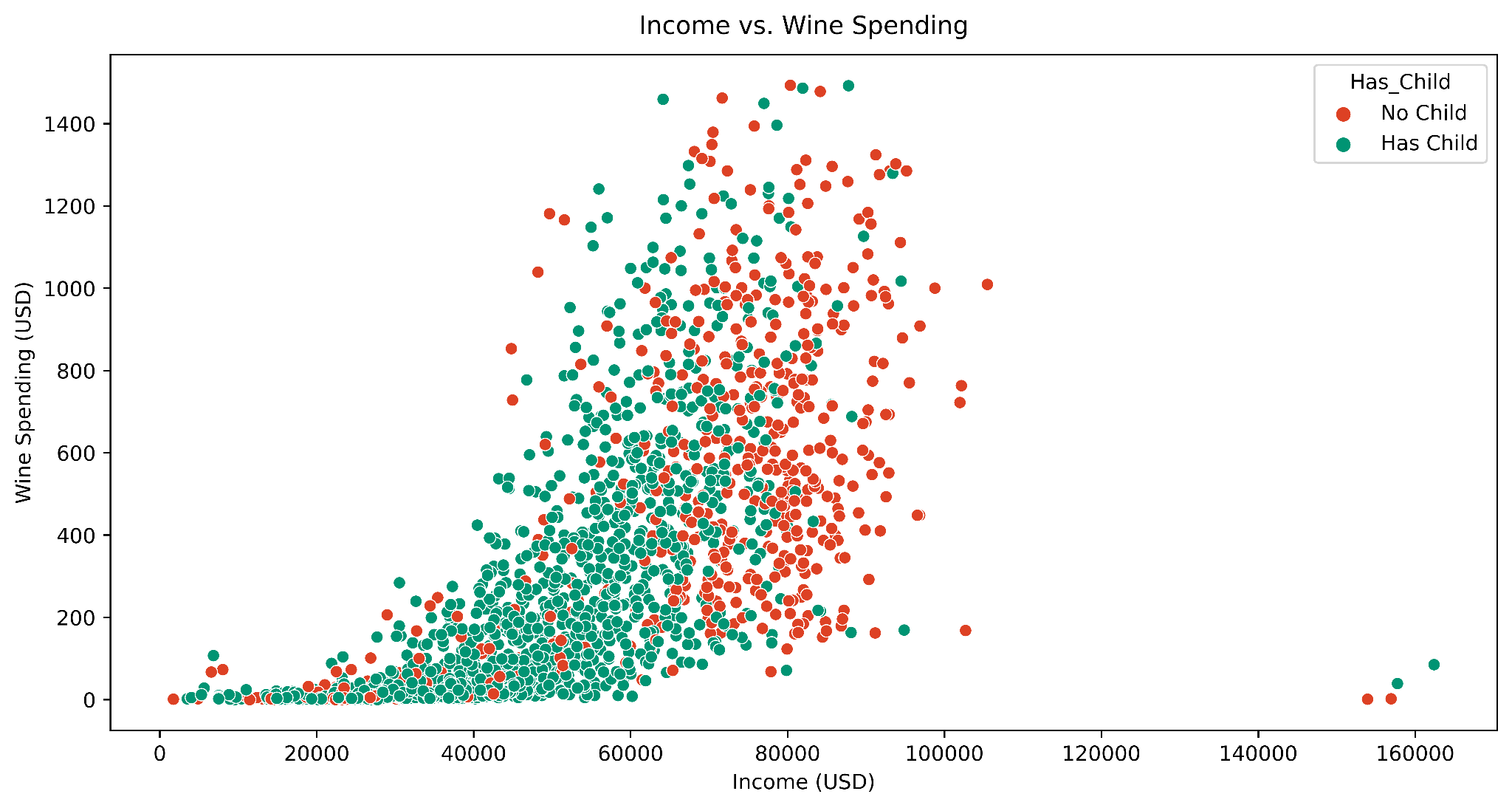


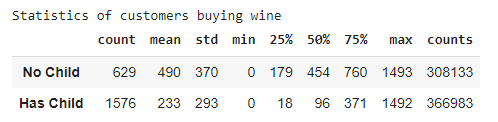
จากกราฟนี้แสดงให้เห็นถึงพฤติกรรมการซื้อของของลูกค้าโดยแบ่งตามสถานภาพสมรส และครอบครัวของลูกค้ามีบุตรหรือไม่ ซึ่งจะได้เห็นได้ว่าในกลุ่มลูกค้าที่ไม่มีบุตรส่วนใหญ่จะมีจำนวนเงินที่ใช้จ่ายอยู่ที่ 1,000-1,500 USD แต่ในกลุ่มลูกค้าที่มีบุตรมีจำนวนเงินที่ใช้จ่ายไม่ถึง 500 USD เป็นจำนวนมาก ในส่วนลูกค้าที่ยังไม่ได้แต่งงานและแต่งงานแล้วนั้นส่วนใหญ่มีจำนวนเงินที่ใช่จ่ายไม่ถึง 500 USD เท่า ๆ กัน





จากกราฟและ DataFrame นี้แสดงถึงจำนวนเงินทีใช้สินค้าชนิดต่างๆของลูกค้าโดยแบ่งตามระดับการเป็นสมาชิกของลูกค้า ซึ่งสินค้าที่ลูกค้าทุกระดับใช้เงินซื้อมากที่สุดก็คือไวน์ โดยลูกค้าระดับ Platinum ใช้เงินซื้อไวน์เป็นจำนวนเงินประมาณ 345,000 USD ลูกค้าระดับ Gold ใช้เงินซื้อไวน์เป็นจำนวนเงินประมาณ 240,000 USD　ลูกค้าระดับ Silver ใช้เงินซื้อไวน์เป็นจำนวนเงินประมาณ 40,000 USD และลูกค้าระดับ Bronze ใช้เงินซื้อไวน์เป็นจำนวนเงินประมาณ 50,000 USD และสินค้าที่ลูกค้าทุกระดับใช้เงินซื้อรองลงมาก็คือเนื้อ โดยคิดเป็นเงินประมาณ 360,000 USD ส่วนสินค้าจำพวกขนมและผลไม้เป็นสินค้าที่ลูกค้าทุกระดับใช้เงินซื้อในราคาใกล้เคียงกันคือประมาณ 60,000 USD สำหรับขนมและประมาณ 58,000 USD สำหรับผลไม้





จากกราฟและ DataFrame ข้างต้นแสดงให้เห็นถึงจำนวนเงินที่ใช้ในการซื้อไวน์เมื่อเปรียบเทียบกับรายได้โดยแบ่งตามกลุ่มลูกค้าที่มีบุตรและไม่มีบุตร จะเห็นได้ว่าลูกค้าที่ไม่มีบุตรใช้จำนวนเงินในการซื้อไวน์ โดยมีค่าเฉลี่ยอยู่ที่ 490 USD ในขณะที่ลูกค้าที่มีบุตรใช้เงินในการซื้อไวน์เฉลี่ย 233 USD ทำให้เห็นว่าลูกค้าที่ไม่มีบุตรใช้เงินซื้อไวน์โดยเฉลี่ยมากกว่าลูกค้าที่มีบุตร ถึงแม้ว่าลูกค้าที่มีบุตรจะใช้เงินในการซื้อไวน์มากกว่าเนื่องจากจำนวนลูกค้าที่มีบุตรมีจำนวนมากกว่าจำนวนลูกค้าที่ไม่มีบุตร

จากข้อมูลที่ได้จากกราฟและ DataFrame ทั้งหมดที่กล่าวมาข้างต้น สามารถสรุปได้ว่าภายในกลุ่มลูกค้าทั้ง 4 กลุ่มนั้นโดยเฉลี่ยแล้วลูกค้าทุกระดับซื้อของจากร้านค้ามากทีุ่สด รองลงมาคือซื้อของจากเว็บไซต์ โดยกลุ่มของลูกค้าระดับ Bronze มีการเข้าชมเว็บไซต์มากที่สุด และจากพฤติกรรมการซื้อสินค้าของลูกค้าทำให้สรุปได้ว่า ไวน์ คือสินค้าที่ลูกค้าใช้เงินซื้อมากที่สุด โดยกลุ่มลูกค้าที่ไม่มีบุตรจากลูกค้าทุกระดับจะเป็นกลุ่มลูกค้าที่ใช้เงินในการซื้อไวน์โดยเฉลี่ยมากที่สุด

**4. สรุปผลการดำเนินงาน**

จากการดำเนินงานโครงงานคณะผู้จัดทำได้ทราบข้อมูลทั่วไปของลูกค้าผู้เป็นสมาชิกของบริษัทซึ่งเป็นข้อมูลเกี่ยวกับอายุ การมีบุตร สถานภาพการสมรส รายได้ในครัวเรือนต่อปี และการศึกษา โดยลูกค้าผู้เป็นสมาชิกของบริษัทส่วนใหญ่เป็นวัยกลางคน มีอายุเฉลี่ยอยู่ที่ 52 ปี จบการศึกษาแล้ว อาศัยอยู่ด้วยกันกับคู่สมรสและมีบุตรอย่างน้อย 1 คน โดยลูกค้ากลุ่มคนโสดคิดเป็นเกือบครึ่งหนึ่งของกลุ่มลูกค้าผู้อยู่กับคู่สมรส และลูกค้ากลุ่มผู้ไม่มีบุตรคิดเป็นเกือบครึ่งหนึ่งของกลุ่มลูกค้าผู้มีบุตร ซึ่งจากการวิเคราะห์ข้อมูลผ่านกราฟและ Data Frame พบว่าในสองปีที่ผ่านมากลุ่มลูกค้าผู้ไม่มีบุตรใช้จ่ายสูงกว่ากลุ่มลูกค้าผู้มีบุตร แสดงให้เห็นว่าสินค้าของบริษัทอาจยังไม่ตอบสนองความต้องการของเด็กและวัยรุ่นส่วนใหญ่ สังเกตได้จากการที่กลุ่มลูกค้าผู้มีบุตรใช้จ่ายน้อยกว่ากลุ่มลูกค้าผู้ไม่มีบุตรในขณะที่กลุ่มลูกค้าผู้มีบุตรมีจำนวนสมาชิกในครอบครัวมากกว่า ในส่วนของผลิตภัณฑ์ ลูกค้าผู้เป็นสมาชิกของบริษัทใช้จ่ายเงินกับไวน์และเนื้อมากที่สุด ในขณะที่ทองคำเป็นสินค้าที่มีราคาสูงแต่กลับมียอดใช้จ่ายซื้อสินค้าชนิดนี้ต่ำ แสดงให้เห็นว่าลูกค้าผู้เป็นสมาชิกของบริษัทมีความต้องการซื้อทองคำจากบริษัทน้อยแต่ความต้องการซื้อไวน์และเนื้อสัตว์จากบริษัทสูง สำหรับช่องทางการสั่งซื้อสินค้าจากบริษัทมีสามช่องทาง คือ ร้านค้า แคตตาล็อค และเว็บไซต์ ช่องทางที่ลูกค้าผู้เป็นสมาชิกใช้บริการมากที่สุดคือ ร้านค้า รองลงมาเป็นเว็บไซต์ และแคตตาล็อคมีผู้ใช้บริการสั่งซื้อสินค้าน้อยที่สุด

ในการจัดกลุ่มข้อมูลลูกค้าผู้เป็นสมาชิกของบริษัทด้วยโมเดล Gaussian Mixture Models ทำให้ได้กลุ่มลูกค้าผู้เป็นสมาชิกทั้งหมดสี่กลุ่ม พบว่าลูกค้าทุกกลุ่มมีการเลือกซื้อสินค้าที่คล้ายคลึงกัน สินค้าที่ได้รับการซื้อมากที่สุดจากลูกค้าทุกกลุ่มคือไวน์ โดยลูกค้ากลุ่มที่มีรายได้สูงจะมีการซื้อที่สูงตามไปด้วย ถึงแม้ว่าจะมีอายุการเป็นสมาชิกของบริษัทที่น้อยกว่าแต่ปริมาณการใช้จ่ายของกลุ่มลูกค้าที่มีรายได้ในระดับเดียวกันนั้นมีปริมาณการใช้จ่ายที่ใกล้เคียงกัน แสดงให้เห็นว่าลูกค้าที่เป็นสมาชิกของบริษัทมานานอาจจะมีแผนการซื้อสินค้าจากบริษัทที่ค่อนข้างตายตัวและไม่มีแผนการซื้อเพิ่มขึ้นมากนักเมื่อเปรียบเทียบกับกลุ่มลูกค้าใหม่ การซื้อสินค้าผ่านทางช่องทางต่างๆจากกราฟได้แสดงให้เห็นว่ากลุ่มลูกค้าผู้มีรายได้สูงและเป็นสมาชิกมานานสมาชิกส่วนใหญ่จะซื้อสินค้าผ่านทางเว็บไซต์ แสดงให้เห็นถึงความไว้วางใจของลูกค้าผู้เป็นสมาชิกกลุ่มนี้ที่เชื่อใจในผลิตภัณฑ์ของบริษัทจนไว้วางใจที่จะสั่งสินค้าผ่านทางเว็บไซต์ แต่ในขณะเดียวกันพบว่ากลุ่มลูกค้าผู้เป็นสมาชิกของบริษัทที่เข้าชมสินค้าทางเว็บไซต์มากที่สุด คือ กลุ่มลูกค้าผู้เป็นสมาชิกที่มีรายจ่ายกับบริษัทน้อยแต่เป็นสมาชิกนาน การสร้างเคมเปญการตลาดในส่วนของเว็บไซต์จึงควรเน้นไปที่ลูกค้ากลุ่มนี้เพื่อกระตุ้นให้ลูกค้ากลุ่มนี้ทำการซื้อสินค้าผ่านทางเว็บไซต์มากยิ่งขึ้นจากที่ปกติมักเป็นการเข้ามาเยี่ยมชมเพียงอย่างเดียว

จากข้อมูลที่ผ่านการวิเคราะห์เมื่อนำมาพิจารณาดูแล้ว ทำให้ทราบถึงพฤติกรรมการใช้จ่ายของลูกค้าแต่ละกลุ่ม การแบ่งกลุ่มลูกค้าผู้เป็นสมาชิกของบริษัทโดยการใช้ข้อมูลรายได้ รายจ่าย และระยะเวลาในการเป็นสมาชิกของบริษัทนั้นถึงแม้จะไม่สามารถแสดงให้เห็นถึงความสนใจต่อสินค้าที่แตกต่างกันของแต่ละกลุ่มได้ เนื่องจากเมื่อแบ่งกลุ่มลูกค้าแล้วลูกค้าทุกกลุ่มมีการใช้จ่ายในสินค้าที่คล้ายคลึงกัน แต่ช่วยแสดงให้เห็นถึงความแตกต่างในปริมาณการใช้จ่ายของแต่ละกลุ่มลูกค้า ความแตกต่างในการซื้อสินค้าผ่านทางช่องทางต่างๆ ความแตกต่างในการเข้าชมสินค้าในเว็บไซต์ของบริษัท และ ความนิยมของสินค้า ผลจากการวิเคราะห์นี้ทำให้ทราบว่าบริษัทมีสินค้าที่เป็นจุดเด่น คือ ไวน์ ซึ่งได้รับการตอบรับเป็นอย่างดีจากทุกกลุ่มลูกค้า การที่บริษัทต้องการจะเพิ่มยอดขายของสินค้าอื่นๆอาจจะมีการนำไวน์มาช่วยในการวางแผนการตลาด เช่น เมื่อซื้อไวน์แล้วจะได้รับส่วนลดในการซื้อสินค้าชนิดอื่น หรือการจัดโปรโมชั่นลดราคาไวน์เพื่อเพิ่มยอดขายนั้นไม่จำเป็นที่จะต้องจัดบ่อย ๆ เนื่องจากเป็นสินค้าที่ขายได้อยู่แล้ว แต่อาจจะมีการลดราคาไวน์หรือเพิ่มของแถมในช่วงเทศกาลเพื่อเป็นการสร้างความประทับใจแก่ลูกค้า

**5. เอกสารประกอบโครงงาน**

5.1 .ไฟล์ชุดข้อมูล

5.1.1 ชุดข้อมูลตั้งต้น :

https://drive.google.com/file/d/1CMZ55MEuH4YM22ldS2ZGa8whFkCf8bc7/view?usp=sharing

5.1.2 ชุดข้อมูลที่ผ่านการทำความสะอาดแล้ว :

https://drive.google.com/file/d/1iq1Wbh1OLijGknBjYbrVe19JbglB5f6l/view?usp=sharing

5.2. ซอร์สโคดภาษาไพทอน

5.2.1 ซอร์สโคดภาษาไพทอนในส่วนของการทำความสะอาดข้อมูล :

https://colab.research.google.com/drive/15uw1G\_IvR1XdtZAnTI6YUPbGd1qlxT-e?usp=sharing

5.2.2 ซอร์สโคดภาษาไพทอนในส่วนของการวิเคราะห์ข้อมูล :

https://colab.research.google.com/drive/1HOltALXIVBzMYnQfZH7B2jq59KCZikDz?usp=sharing

5.3 กูเกิ้ลไดรฟ์ที่รวมไฟล์งานทั้งหมด : https://drive.google.com/drive/folders/1DNeBrbPFIRdb6l8Z962G4N2Shbq296iL?usp=sharing