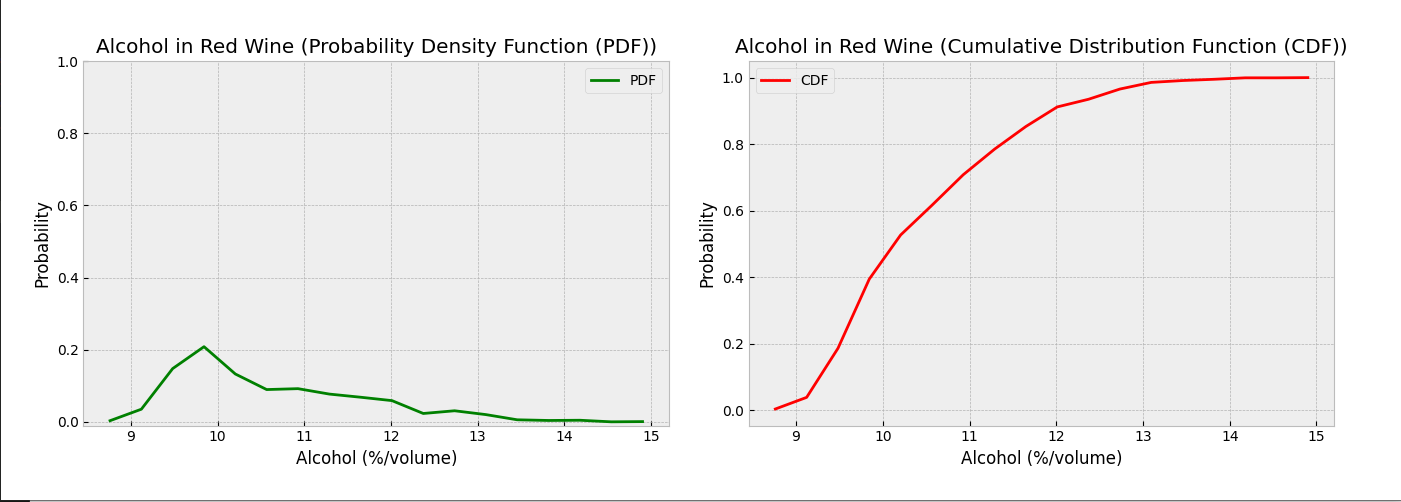
**Homework #3**

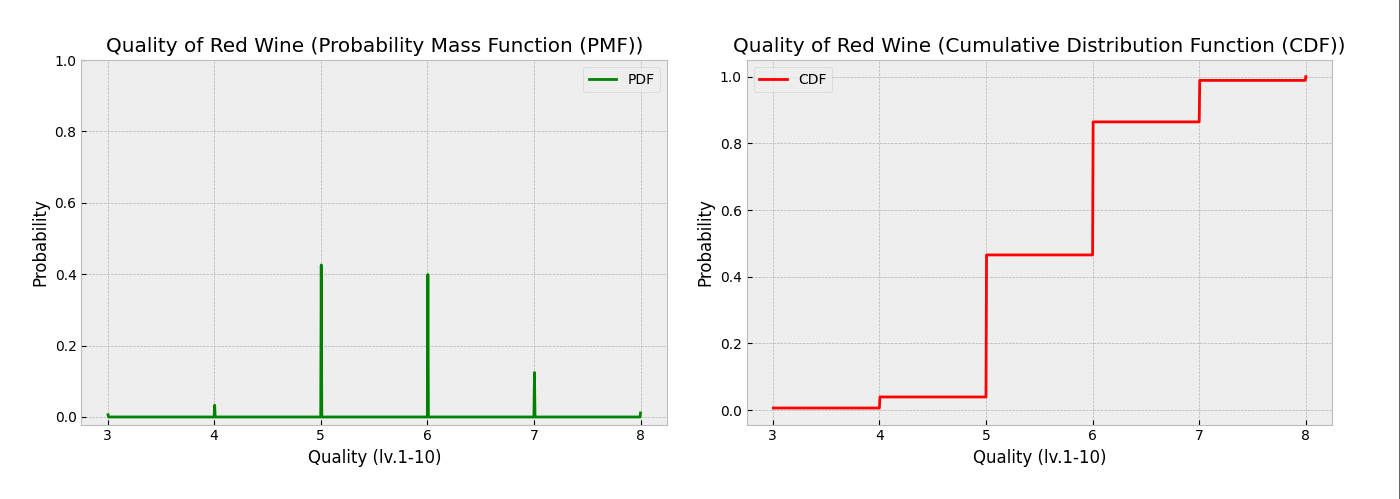
**กราฟ Probability Density Function (PDF)/ Cumulative Distribution Function (CDF)**

**ปริมาณแอลกอฮอล์ในไวน์แดง**



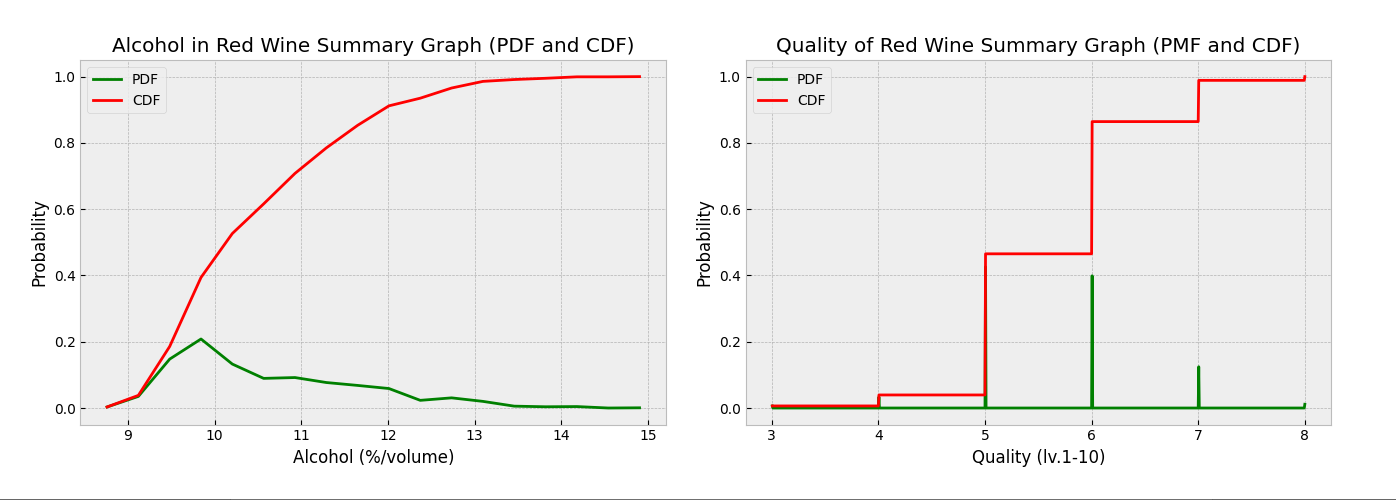
**กราฟ Probability Mass Function (PMF)/ Cumulative Distribution Function (CDF)**

**คุณภาพของไวน์แดง**



**บทวิเคราะห์ข้อมูลจากกราฟทั้งหมด**

จะเห็นได้ว่าเมื่อนำข้อมูลกราฟทั้งสองมารวมกัน จะเห็นความสัมพันธ์ระหว่างกราฟ PDF - CDF และ PMF – CDF ดังนี้



**ปริมาณแอลกอฮอล์ในไวน์แดง**

**ในกราฟปริมาณแอลกอฮอล์ เส้นสีเขียว (PDF) สังเกตได้ ดังนี้**

1. จะมีลักษณะเป็น **ภูเขาสูง** ระหว่างช่วงปริมาณแอลกอฮอล์**ตั้งแต่ 9.2 – 10.5 %/volume**
2. จะมีลักษณะเป็น **ที่ราบแบน** ระหว่างช่วงปริมาณแอลกอฮอล์**ตั้งแต่ 8.5 – 9.1 %/volume และ 10.6 – 15.0 %/volume**

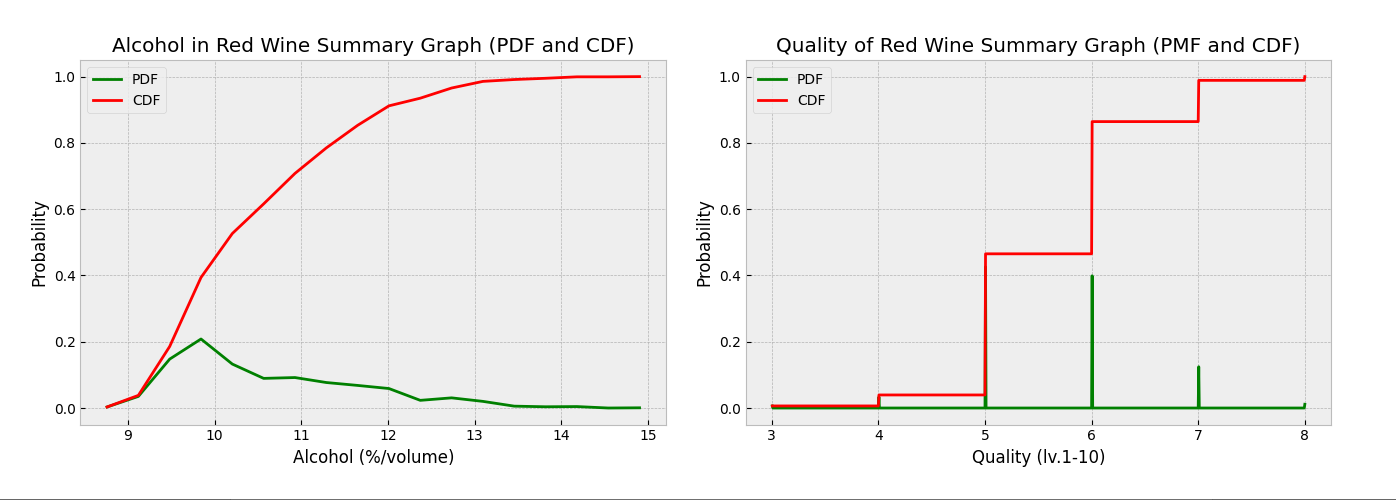
**ในกราฟปริมาณแอลกอฮอล์ เส้นสีแดง (CDF) สังเกตได้ ดังนี้**

1. เป็นการรวมค่า**ความน่าจะเป็น(ความถี่สัมพัทธ์)** จากเส้นกราฟสีเขียว เพิ่มขึ้นเรื่อย ๆ

สะสมจนครบ **1.0** เมื่อสิ้นสุดกราฟนั้น ๆ ซึ่งเป็น**ความถี่สะสมสัมพัทธ์**

1. จะมีค่า**ความน่าจะเป็น = 0.5 ที่ประมาณ 10.2 %/volume** ซึ่งเป็นจุดกึ่งกลางของกราฟพอดี (Median) ทำให้ทางซ้ายและทางขวาของพื้นที่ใต้กราฟสีเขียวแบ่งออกเป็น 50% เท่า ๆ กัน
2. **ความชันของกราฟช่วงประมาณ 9.2 - 9.7 %/volume** มีความชันโดยรวมมาก และความชันจะค่อย ๆ ลดหลั่นลงไปเรื่อย ๆ เพื่อมีปริมาณแอลกอฮอล์มากขึ้น

**กล่าวโดยสรุป จากข้อมูลที่ได้ ไวน์จะมีปริมาณแอลกอฮอล์ประมาณ 9.2 – 10.5 %/volume เป็นส่วนมาก เนื่องจากมีพื้นที่ใต้กราฟเส้นสีเขียวเยอะ (Density หนาแน่น) และไวน์จะมีปริมาณแอลกอฮอล์ประมาณ 8.5 – 9.1 %/volume และ 10.6 – 13.0 %/volume เป็นส่วนที่น้อย เนื่องจากพื้นที่ใต้กราฟเส้นสีเขียวน้อย (Density เบาบาง) และถ้ายิ่งมีปริมาณแอลกอฮอล์ที่มากขึ้นไปกว่า 13.0 %/volume ก็จะมีจำนวนไวน์ที่จะลดลงตามลำดับลงไป (Density เบาบางมาก ๆ) ซึ่งข้อมูลข้างต้นสามารถนำมาดูเปรียบเทียบได้กับเส้นสีแดงควบคู่กัน ถ้าเส้นสีแดงมีความชันมาก ก็จะมีการเพิ่มจำนวนไวน์ในปริมาณแอลกอฮอล์ช่วงนั้นที่มากขึ้น ส่วนถ้ามีความชันน้อย ก็จะมีการเพิ่มจำนวนไวน์ในปริมาณแอลกอฮอล์ช่วงนั้นที่น้อยลง**



**คุณภาพของไวน์แดง**

**ในกราฟคุณภาพไวน์แดง เส้นสีเขียว (PMF) สังเกตได้ ดังนี้**

1. **คุณภาพระดับ** **5 และ 6 จะมีปริมาณมาก** เมื่อเทียบกับคุณภาพไวน์อื่นๆ เนื่องจากมีแท่งสีเขียวที่ค่อนข้างสูงโดดเด่นเป็นพิเศษและมีค่า**ความน่าจะเป็น(ความถี่สัมพัทธ์) โดยประมาณ = 0.4**
2. **คุณภาพระดับ 7** จะมี **ความน่าจะเป็น(ความถี่สัมพัทธ์) โดยประมาณ = 0.1**
3. **คุณภาพระดับ 3 4 และ 8** จะมี **ความน่าจะเป็น(ความถี่สัมพัทธ์) โดยประมาณน้อยกว่า 0.05**ซึ่งเป็นค่าที่น้อยมากๆ

**ในกราฟคุณภาพไวน์แดง เส้นสีแดง (CDF) สังเกตได้ ดังนี้**

1. เป็นการรวมค่า**ความน่าจะเป็น(ความถี่สัมพัทธ์)** จากเส้นกราฟสีเขียว เพิ่มขึ้นเรื่อย ๆ

สะสมจนครบ **1.0** เมื่อสิ้นสุดกราฟนั้น ๆ ซึ่งเป็น**ความถี่สะสมสัมพัทธ์**

1. จะมีค่า**ความน่าจะเป็น = 0.5 ที่คุณภาพระดับ 6** ซึ่งเป็นจุดกึ่งกลางของกราฟพอดี (Median)
2. **ความชันของกราฟคุณภาพระดับที่ 5 และ 6** จะลากยาว และมีความสูงเป็นพิเศษกว่าระดับคุณภาพอื่นๆ เนื่องจากมีจำนวนคุณภาพระดับ 5 และ 6 ที่มีเยอะมากกว่าระดับอื่นๆ

**กล่าวโดยสรุป จากข้อมูลที่ได้จากเส้นสีเขียว ไวน์จะมีระดับคุณภาพอยู่ที่ประมาณ 5 และ 6 เป็นส่วนใหญ่ (คิดเป็นประมาณ 80% ของไวน์ทั้งหมด) รองลงมาไวน์จะมีคุณภาพอยู่ในระดับ 7 มีปริมาณปลานกลาง(คิดเป็นประมาณ 12% ของไวน์ทั้งหมด) และ คุณภาพระดับ 3 4 และ 8 มีปริมาณที่ค่อนข้างน้อย (คิดเป็นประมาณ 8% ของไวน์ทั้งหมด) ซึ่งสามารถนำข้อมูลข้างต้นมาดูเปรียบเทียบได้กับเส้นสีแดงควบคู่กัน ถ้าเส้นสีแดงมีความชันมาก ก็จะมีการจำนวนคุณภาพของไวน์ในระดับนั้นมาก ถ้ามีความชันน้อย ก็จะมีจำนวนคุณภาพของไวน์ในระดับนั้นน้อย**

**จากการวิเคราะห์ทั้งหมด จะมีข้อสรุปได้ว่า ไวน์ส่วนใหญ่ที่ได้จากการผลิตในความสัมพันธ์ระหว่างปริมาณแอลกอฮอล์ในไวน์และคุณภาพของไวน์ ปริมาณแอลกอฮอล์ในไวน์ช่วงตั้งแต่ประมาณ 9.2 – 10.5 %/volume จะมีไวน์ปริมาณมากที่อยู่ในช่วงนี้ ซึ่งจะมีคุณภาพของไวน์เฉลี่ยอยู่ในระดับ 5 และ 6 เป็นส่วนใหญ่ด้วยเช่นกัน ถือว่าไวน์ส่วนใหญ่ที่มีระดับแอลกอฮอล์ 9.2 – 10.5 %/volume จะมีระดับคุณภาพที่พอใช้-ดี ในขณะที่ปริมาณแอลกอฮอล์ในไวน์ในช่วงตั้งแต่ประมาณ 8.5-9.1 %/volume และ 10.6 – 13.0 %/volume เป็นส่วนที่มีจำนวนไวน์ที่หาได้ยากหรือมีจำนวนที่ลดลง จะมีคุณภาพของไวน์เฉลี่ยอยู่ในระดับ 7 จึงอาจตีความได้ว่า เมื่อปริมาณแอลกอฮอล์ในไวน์ยิ่งสูงขึ้น จำนวนไวน์จะยิ่งหาได้ยากขึ้นหรือมีจำนวนที่ลดลง คุณภาพของไวน์อาจจะเพิ่มขึ้นไปอยู่ที่ระดับ 7 ซึ่งเป็นคุณภาพในระดับที่ดี-ดีมากส่วนคุณภาพในระดับ 3 4 และ 8 จะไม่นำมาคิด เพราะมีค่าความน่าจะเป็นที่น้อยมากๆ จากค่า outlier ของข้อมูล รวมถึงไม่คิดในส่วนปริมาณแอลกอฮอล์ที่มากเกินกว่า 13.0 %/volume ด้วย**

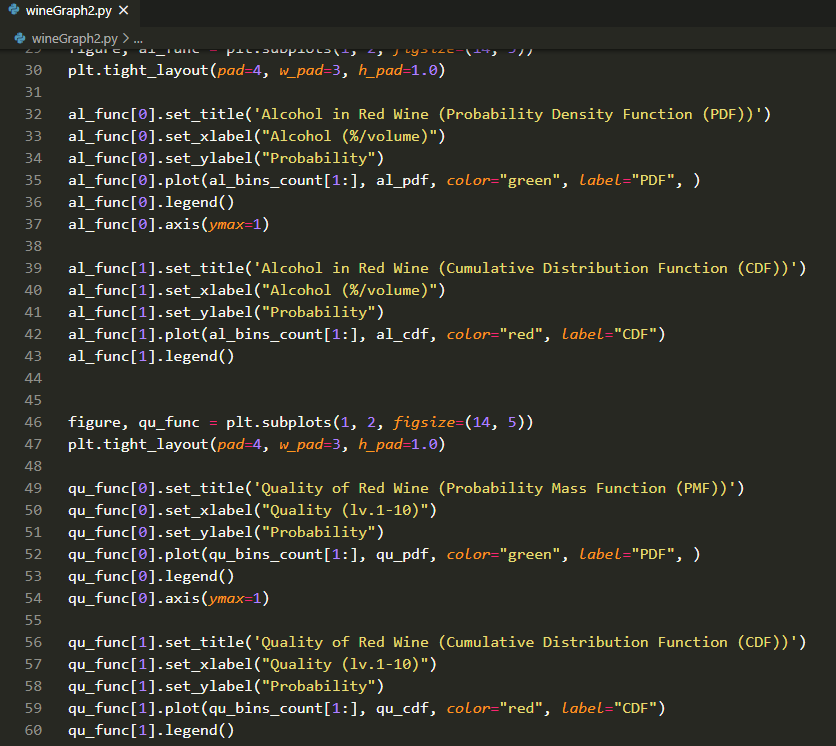
**ดังนั้น ปริมาณแอลกอฮอล์ในไวน์แดงที่มากขึ้น อาจมีแนวโน้มที่จะทำให้คุณภาพของไวน์แดงเพิ่มขึ้นตามไปด้วย โดยความสัมพันธ์ระหว่างปริมาณแอลกอฮอล์ในไวน์แดง และคุณภาพของไวน์แดงเป็นความสัมพันธ์แบบแปรผันตรงหรือคล้อยตามกัน**

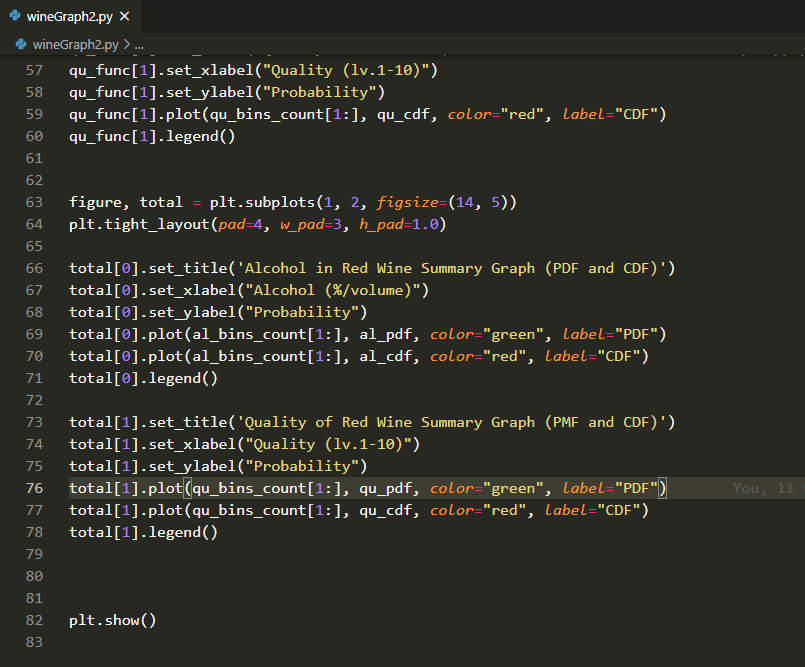
**อย่างไรก็ตาม ทั้งหมดนี้ยังไม่สามารถกล่าวได้อย่างชัดเจน 100% เป็นเพียงแค่แนวโน้มเท่านั้น เนื่องจากในการผลิตไวน์จริง จะมีส่วนผสมอื่นๆ และมีอีกหลายปัจจัยในการกำหนดคุณภาพของไวน์แดง เช่น ค่าความเป็นกรด, น้ำตาลคงค้างที่เหลือในไวน์แดง, ระยะเวลาการผลิตไวน์แดง, คุณภาพขององุ่นที่นำมาใช้ในการผลิต และ เกณฑ์การวัดคุณภาพของไวน์แดง เป็นต้น ซึ่งเกณฑ์การวัดคุณภาพของไวน์แดง(ระดับ 1-10)ในครั้งนี้ อ้างอิงมาจากโรงงานผลิตไวน์ ในจังหวัด Minho ทางตอนเหนือของประเทศโปรตุเกสเท่านั้น**



**รายละเอียด Source Code ของโปรแกรม WineGraph2.py**







**แหล่งที่มาของชุดข้อมูล (Reference/URL) :**

**- ที่มาของชุดข้อมูล** Winequality-red.csv

<https://www.kaggle.com/uciml/red-wine-quality-cortez-et-al-2009>

**- ที่มาคำอธิบายแต่ละส่วนประกอบของไวน์**

<https://archive.ics.uci.edu/ml/datasets/wine+quality>

<https://rstudio-pubs-static.s3.amazonaws.com/57835_c4ace81da9dc45438ad0c286bcbb4224.html>

<https://waterlibrary.com/th-รู้ไหมว่า-ระดับปริมาณแ/#:~:text=ปัจจุบันมีแอลกอฮอล์อยู่ใน,สูงขึ้นด้วยเช่นกัน>

**- วิธีการทำไวน์**

<https://www.youtube.com/watch?v=7gquYRxLMFI&ab_channel=Insider>

**- ประเภทของไวน์**

<https://www.unlockmen.com/terrazas-unlock-wine-101-1/>

<https://thewinelist.shop/blog/news/wine-101>

**- รายละเอียดอื่นๆ เกี่ยวกับคุณภาพและวิธีรับรสที่ดีของไวน์**

<https://www.blockdit.com/posts/5e5f68d77b00780ed6462939>

<https://www.dummies.com/food-drink/drinks/wine/the-special-technique-for-tasting-wine/>

<https://www.quickanddirtytips.com/house-home/entertaining/wine/4-ways-to-know-if-your-wine-is-good>