

# Quiz Regression 20%

ให้ส่งไฟล์เป็น .pdf เท่านั้น โดยตั้งชื่อไฟล์เป็น รหัสนิสิต.pdf

โดยในไฟล์ประกอบด้วย

- python code,
- กราฟรูปแบบต่างๆ ที่ใช้ในการวิเคราะห์
- อธิบายการวิเคราะห์ เหตุผลและการเลือกใช้โมเดลรูปแบบนั้น และ
- output ค่า Y ที่ได้จากการพยากรณ์ new\_X โดยใช้โมเดลที่เหมาะสมที่สุด

ไฟล์ data\_quiz.csv ประกอบด้วยข้อมูล 3 คอลัมน์คือ Y, X และ new\_X

ให้อ่านไฟล์ข้อมูล data\_quiz.csv วิเคราะห์ความสัมพันธ์ระหว่าง X และ Y โดยทำการวิเคราะห์ด้วยรูปแบบ

**Simple linear regression** และ **Polynomial regression** ด้วย degree ต่างๆ เเท่าที่คิดว่าเหมาะสม

แสดงกราฟแยกตามการวิเคราะห์แต่ละรูปแบบ โดยสามารถแสดงทีละกราฟ

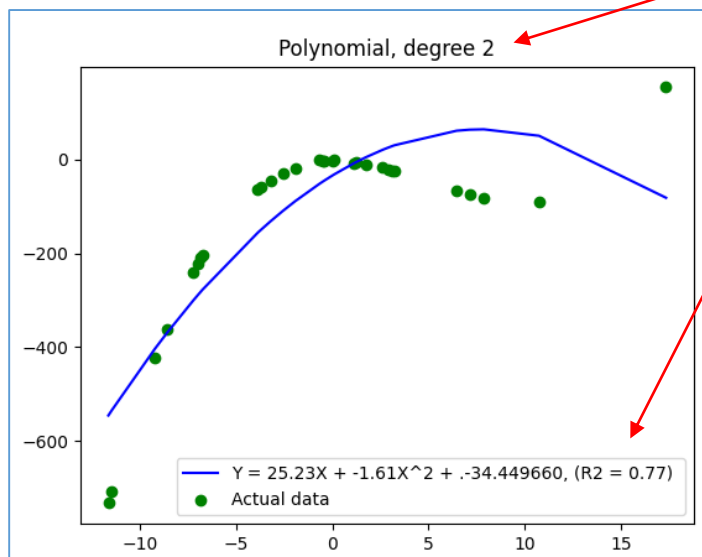
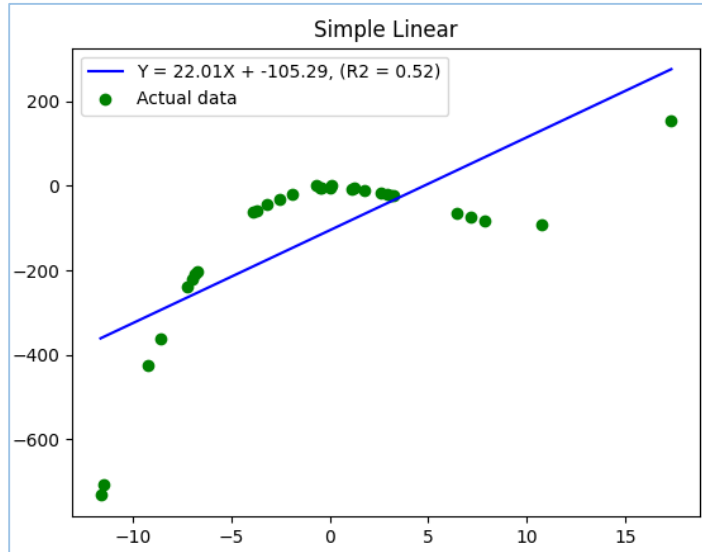
ในกราฟแต่ละรูปให้แสดง

- จุดสีเขียว แสดงข้อมูลจริง Actual data
- เส้นสีน้ำเงิน แสดงเส้น regression โดยแสดงสมการ regression และ  $R^2$  ของสมการนั้น
- title ของกราฟ แสดง รูปแบบการวิเคราะห์เช่น Simple linear หรือ Polynomial degree 2

วิเคราะห์ สรุปผลว่าโมเดลใดดีที่สุดที่จะใช้พยากรณ์ข้อมูลชุดนี้ เพราะเหตุใด

อ่านข้อมูลจากคอลัมน์ new\_X เพื่อใช้ในการพยากรณ์ค่า, output ค่า Y ที่ได้จากการพยากรณ์ค่าโดยใช้โมเดลที่ดีที่สุดเพียงโมเดลเดียวเท่านั้นในการพยากรณ์

## ตัวอย่างการแสดงผลกราฟ



title

label แสดงสมการและค่า  $R^2$