**ดิชั้นนางสาว.....รหัส 61051070 sec 02 วันนี้จะมานำเสนอโปรเจ็คของรายวิชา** Data Science**เรื่อง : โมเดลการทำนายอัตราการได้เหรียญของประเทศที่เข้าร่วมกีฬา Olympic ตั้งแต่ปี 1976 – 2008 ในโปรเจ็คนี้รับผิดชอบในส่วนเขียนโค้ดเกี่ยวกับกานจัดการข้อมูลเพื่อนำไปวิเคราะหาสถิติเพื่อทำนายผล และเขียนรายงาน**

**เราจะมาเริ่มพูดกันที่วัตถุประสงค์ของโปรเจ็คนี้นะคะ**

**ผู้จัดทำ**

นายภานุพันธ์ ตำปาน รหัสนักศึกษา 610510701 sec 2

นางสาวศศิวิมล วิทาทาน รหัสนักศึกษา 610510707 sec 2

**เสนอ**

อาจารย์ สุธาสินี โทวุฒิกุล

อาจารย์ เบญจมาศ ปัญญางาม

รายงานฉบับนี้เป็นส่วนหนึ่งของวิชา หลักการเขียนโปรแกรมสำหรับวิเคราะห์ข้อมูล (Programming for Data Science) 229223 ภาคเรียนที่ 2 ปีการศึกษา 2563

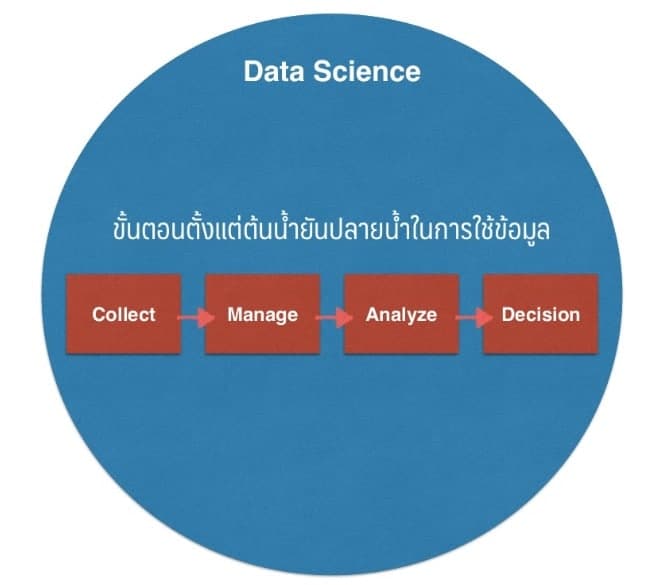
1. วัตถุประสงค์

- เพื่อสร้างโมเดลการทำนายผลอัตราการได้รับเหรียญในปีต่อๆไป

- เพื่อใช้ python ในการกรองข้อมูลและสร้างแบบจำลอง

2. หลักการและเหตุผล

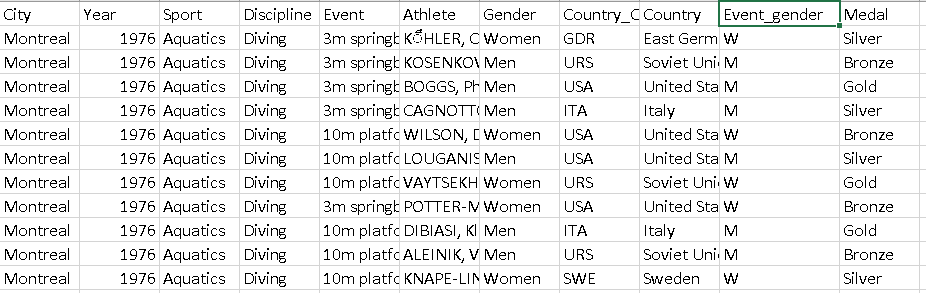
เนื่องจากทางเราได้ศึกษาข้อมูลเกี่ยวกับการทำนายผลศึกฟุตบอลชิงแชมป์สโมสรยุโรปซึ่งใช้หลักการของวิชา นี้ จึงทำให้ผู้จัดทำมีความสนใจในการนำข้อมูลผลการแข่งของโอลิมปิกมาทำนายผลผลการแข่งของโอลิมปิกในครั้งถัดไป และนำเสนอในรูปของกราฟเส้น



3. แหล่งที่มาของข้อมูล

[**https://www.kaggle.com/divyansh22/summer-olympics-medals**](https://www.kaggle.com/divyansh22/summer-olympics-medals)

3.1ข้อมูลประกอบด้วย field ต่างๆดังนี้



- City

เป็นชื่อเมืองที่แข่งของประเทศเจ้าภาพ

- Year

เป็นปีที่แข่งขัน

- Country Code

เป็นรหัสย่อ 3 ตัวเพื่อแทนชื่อเต็มประเทศ

- Country

ชื่อประเทศที่เข้าแข่งขัน

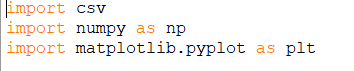
- Medal

ชนิดเหรียญที่ได้รับ

4. การทำงานในโปรแกรม

4.1 ความสามารถของโปรแกรม

1. โปรแกรมนี้สามารถอ่านและเขียนไฟล์สกุล Excel ได้ เนื่องมาจากการนำเข้าไลบารี่ csv
2. โปรแกรมนี้สามารถแสดงกราฟเส้นแสดงค่าโอกาสเหรียญที่ได้รับของประเทศที่เข้าร่วมแข่งได้ เนื่องมีการนำเข้าไลบารี่ numpy และ ไลบารี่ matplotlib



4.2 หลักการการทำงานภายในโปรแกรม

แผนพังภาพรวมของโปรแกรม

INPUT

summer-olympics-medals.csv

ไฟล์ excel แสดงข้อมูลในโอลิมปิก

PROGRAM

* อ่านข้อมูลแล้วคัดข้อมูลที่จำเป็นต่อการหาสถิติแล้วเก็บลงใน dict
* นำข้อมูลมาหาค่าสถิติแล้ววาดกราฟ

OUTPUT

1. ไฟล์ excel แสดงข้อมูลตัวเลขโอกาสเหรียญโอลิมปิกที่ได้ของแต่ละประเทศในแต่ละปี
2. กราฟแสดงข้อมูลสถิติที่ได้รับเหรียญของแต่ละประเทศพร้อมทำนายในปีถัดไป

5. เหตุผลในการเลือกโครงสร้างและลักษณะข้อมูลออกของโปรแกรม

5.1. เหตุผลในการเลือกใช้โครงสร้างข้อมูล

- เก็บข้อมูลเป็นชนิด String เนื่องจากง่ายต่อการเก็บข้อมูลเนื่องจากข้อมูลที่ได้อ่านจากไฟล์ excel ทำให้สามารถเช็คข้อมูลจากไฟล์ที่อ่านมาได้สะดวกและถูกต้อง

- เลือกใช้การเก็บข้อมูลด้วยโครงสร้างข้อมูลชนิด Dictionary เนื่องจากง่ายต่อข้อมูลที่มากและมีชื่อเฉพาะทำให้ง่ายต่อการเรียกใช้ข้อมูล

- เลือกใช้แสดงข้อมูลจากข้อมูลชนิด List เนื่องจากง่ายต่อการนำมาวาดกราฟด้วยไลบารี่ matplotlib

5.2. เหตุผลในการเลือกลักษณะข้อมูลออกของโปรแกรม

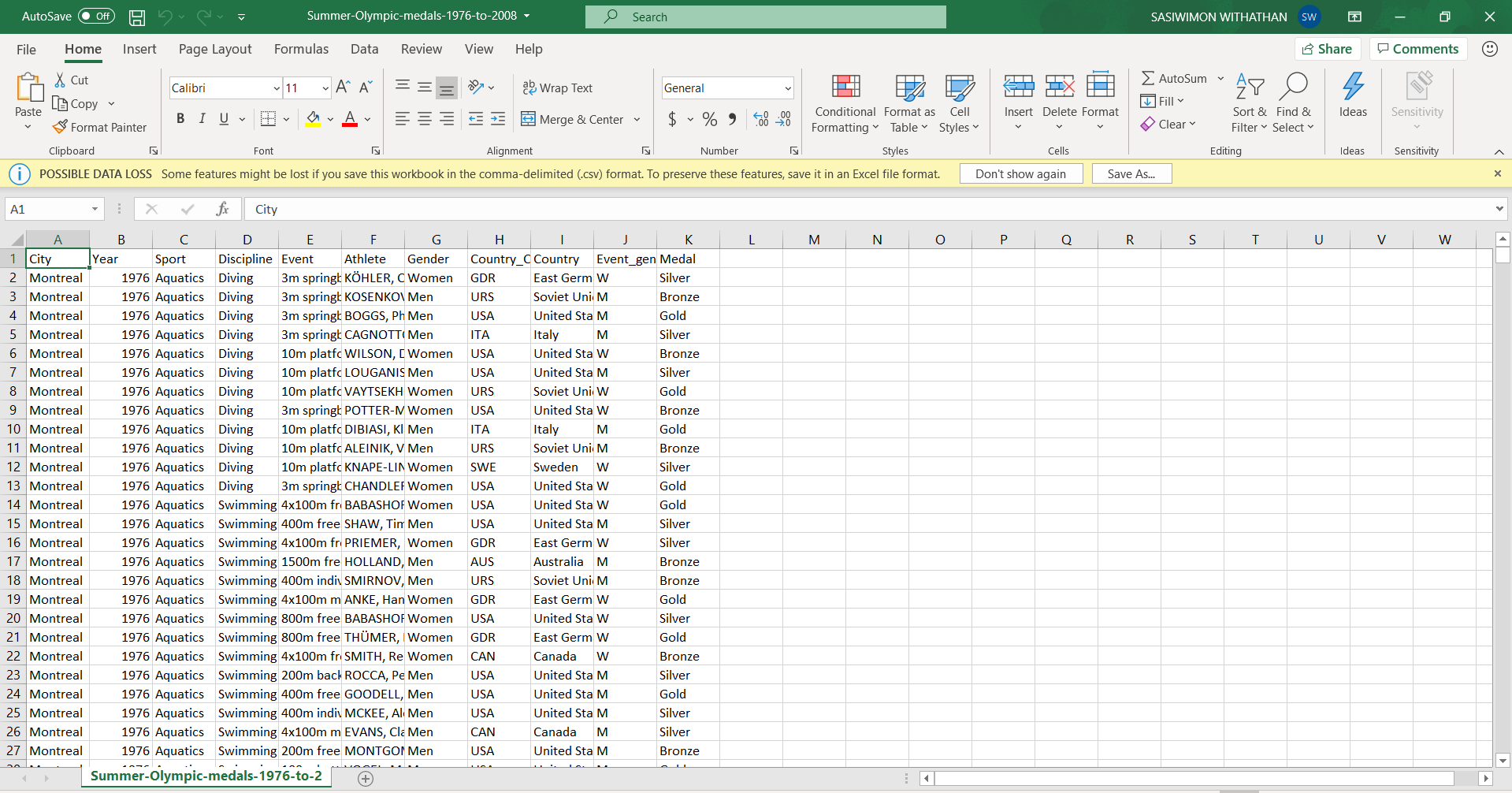
- สร้างไฟล์excelใหม่เพื่อแสดงข้อมูลที่ใช้งานในโปรแกรมที่เลือกสร้างไฟล์ใหม่เพราะเวลาที่รันโปรแกรมแต่ละครั้งจะได้ผลเหมือนเดิมทุกครั้ง

- เลือกการใช้ไลบารี่ matplotlib เพื่อวาดกราฟ เนื่องจากทำให้ผู้ที่ใช้โปรแกรมสามารถดูสถิติข้อมูลได้สะดวก และเข้าง่าย

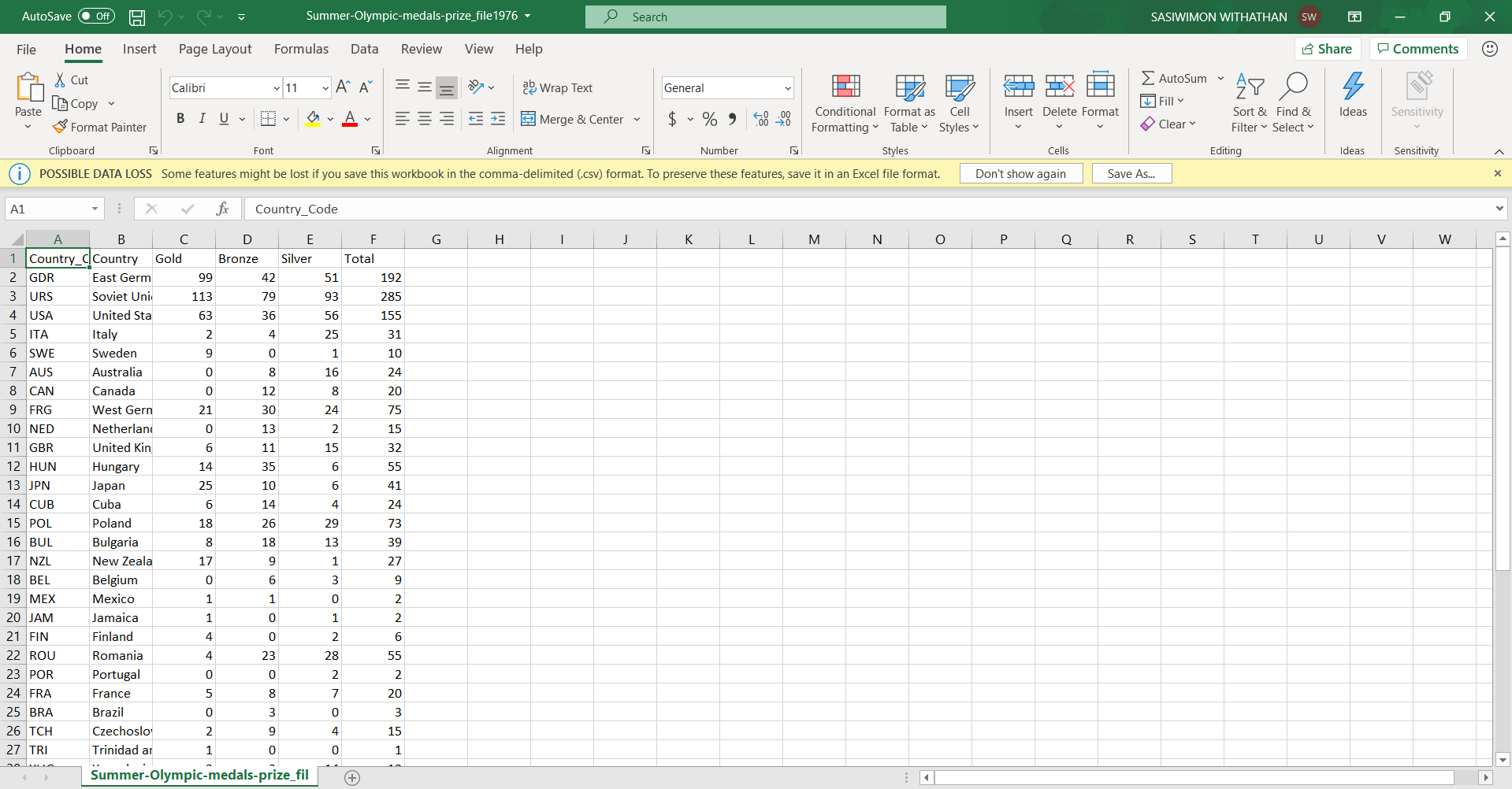
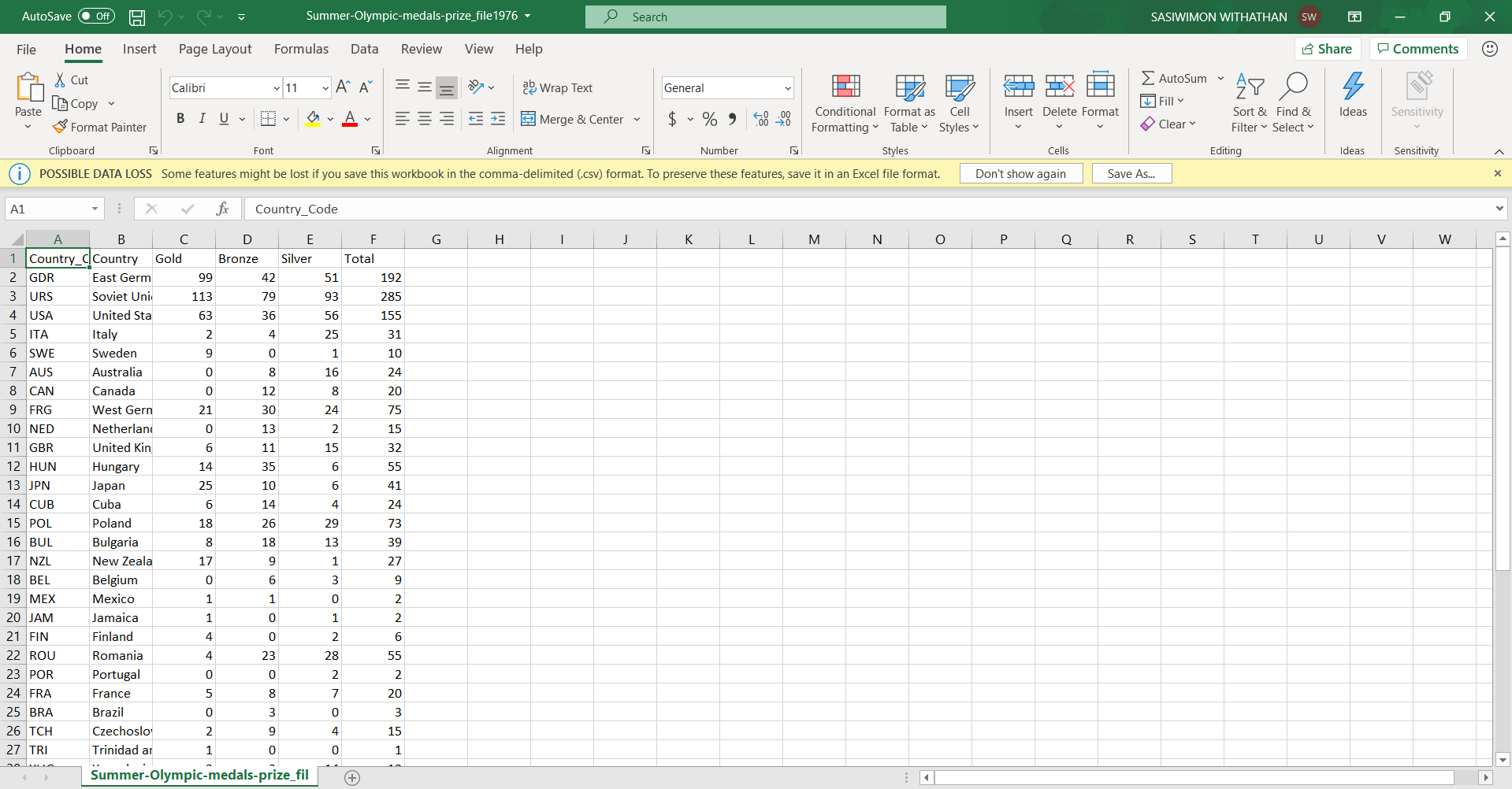
6. ตัวอย่างหน้าจอพร้อมคำอธบายการทำงานโปรแกรม

* Input / Output

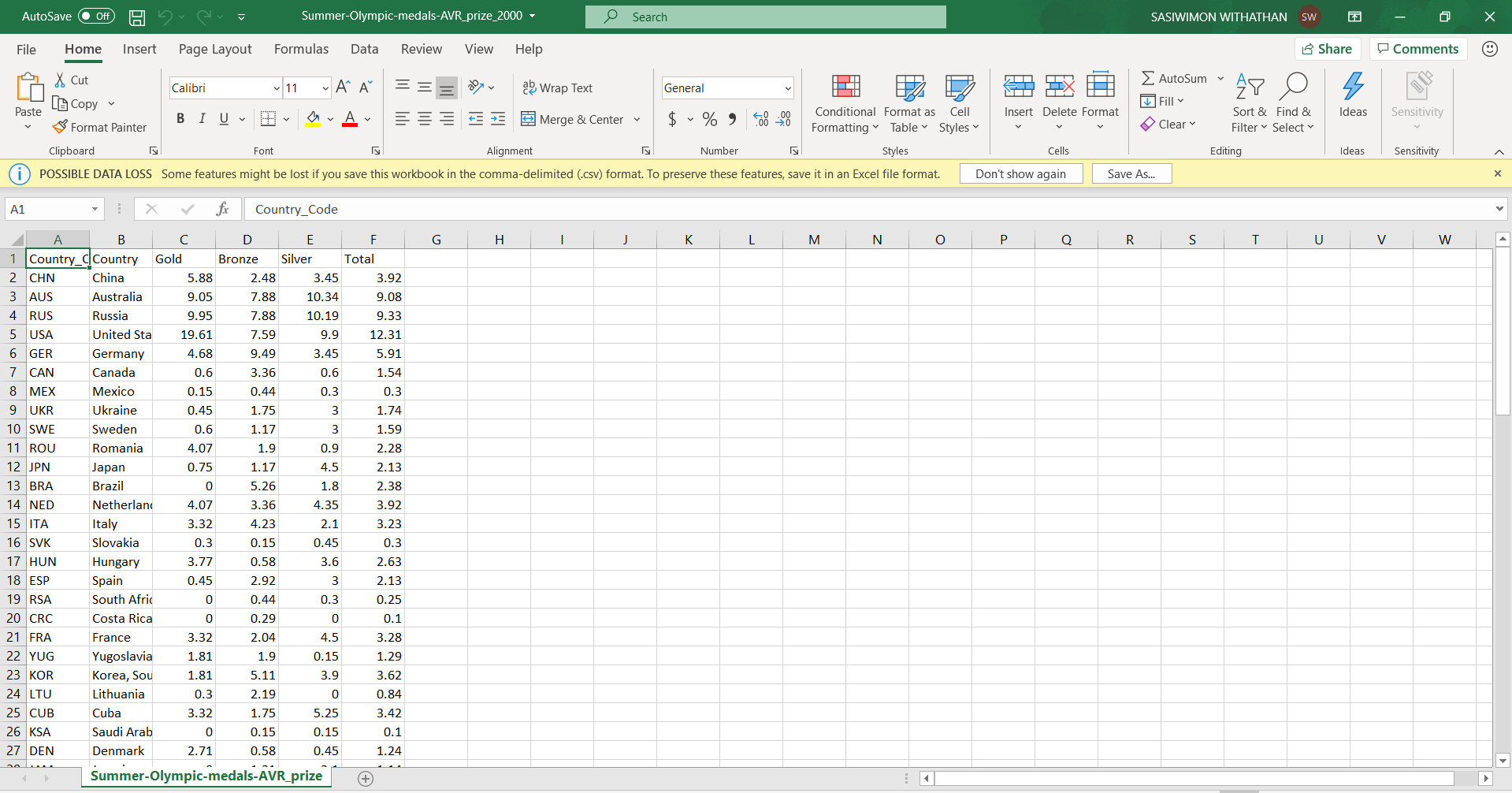
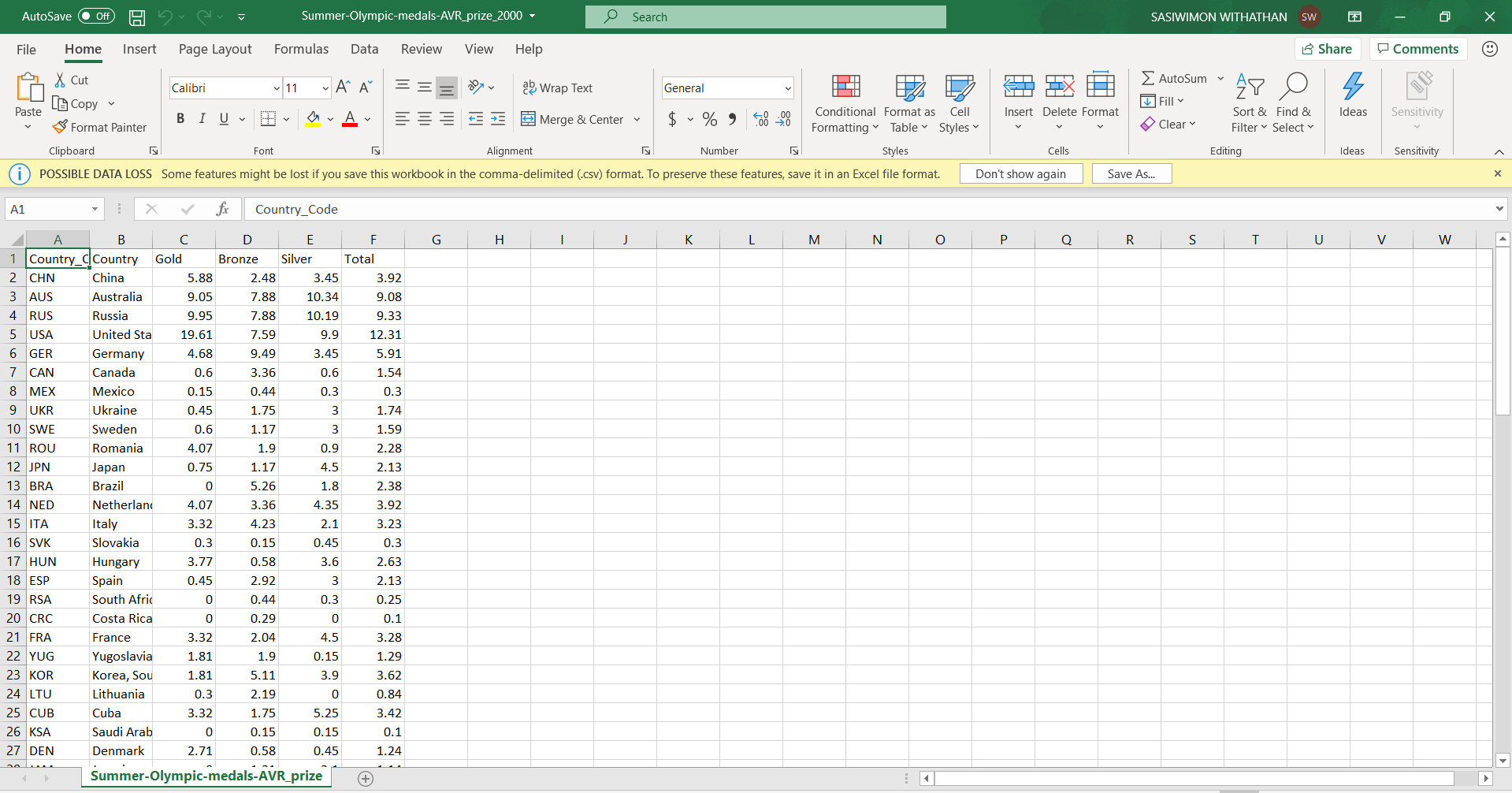
Input : Summer-Olympic-medals-1976-to-2008.csv



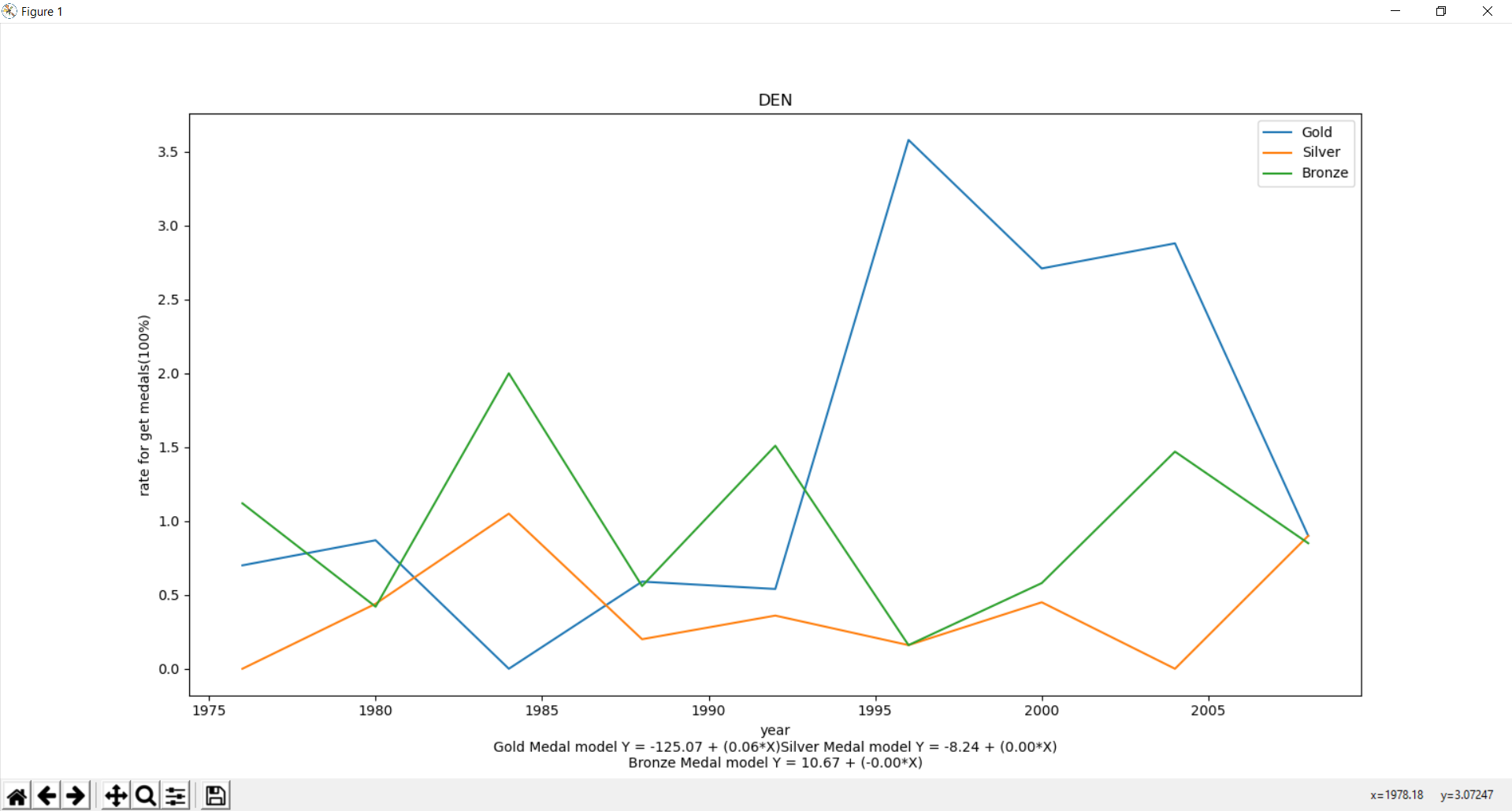
Output : 1.ไฟล์ excel แสดงข้อมูลจำนวนเหรียญโอลิมปิกที่ได้ของแต่ละประเทศในแต่ละปี



2.ไฟล์ excel แสดงข้อมูลตัวเลขโอกาสเหรียญโอลิมปิกที่ได้ของแต่ละประเทศในแต่ละปี



3.กราฟแสดงข้อมูลสถิติที่ได้รับเหรียญของแต่ละประเทศพร้อมทำนายในปีถัดไป



* การทำงานของโปรแกรม

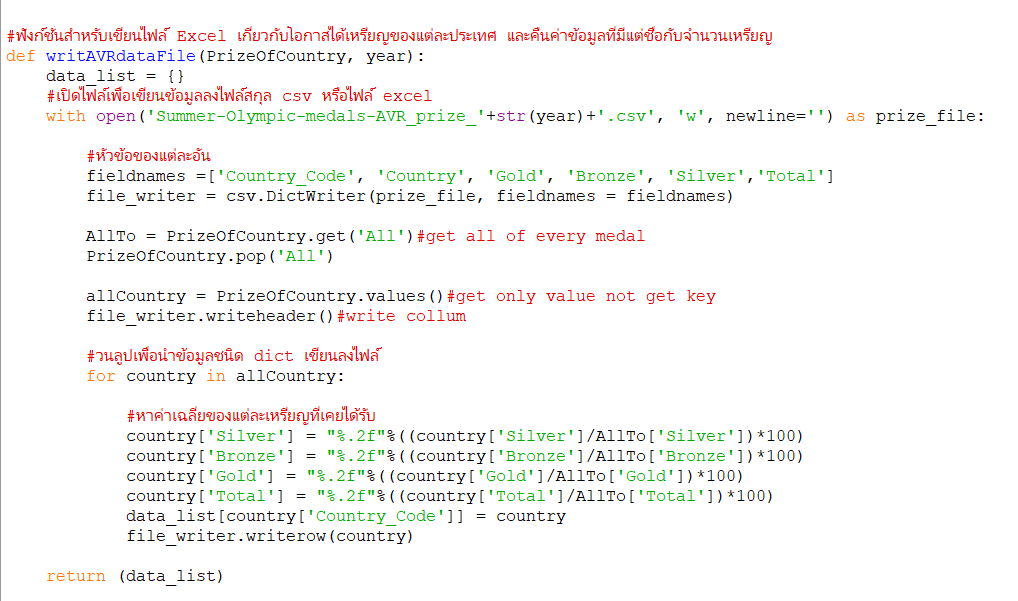
1. สร้างฟังก์ชันเพื่ออ่านข้อมูลจากไฟล์ excel และคืนค่าข้อมูลที่จะใช้ในการหาค่าสถิติ



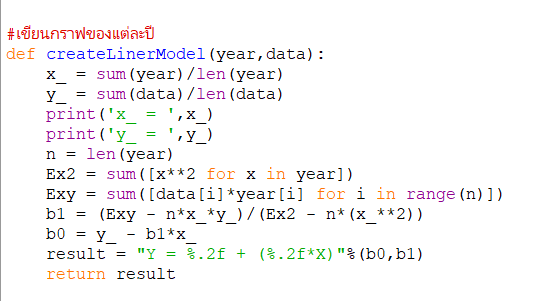
1. สร้างฟังก์ชันสำหรับแสดงผลของข้อมูลจำนวนเหรียญที่ได้รับของแต่ละประเทศในแต่ละปีในรูปแบบของไฟล์ excel



1. สร้างฟังก์ชันสำหรับแสดงผลของค่าเฉลี่ยจำนวนเหรียญที่ได้รับของแต่ละประเทศในแต่ละปีในรูปแบบของไฟล์ excel และคืนค่าข้อมูลนี้เป็นชนิดdict

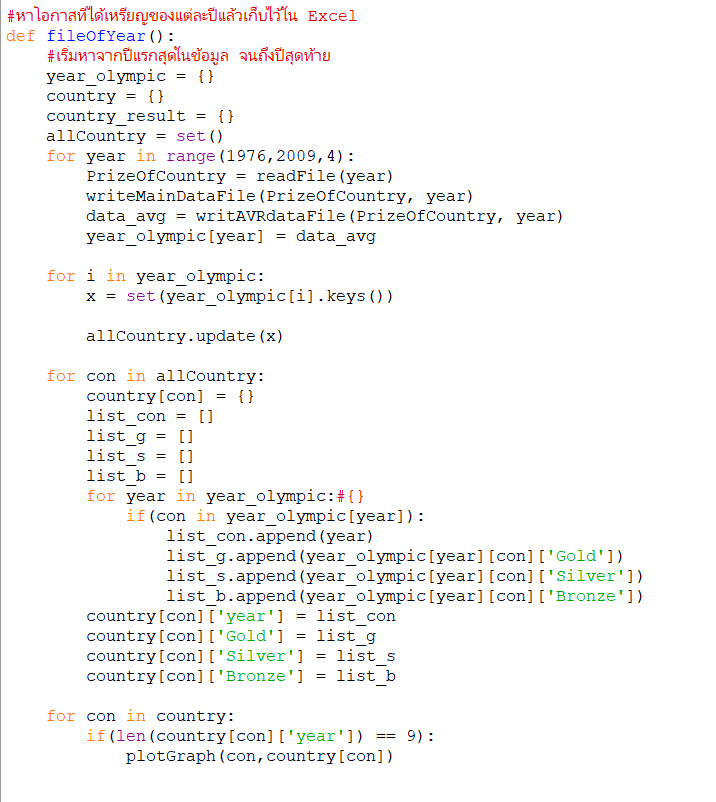


1. ฟังก์ชันการสร้างโมเดลเพื่อทำนายเหรียญที่จะได้ในโอลิมปิกครั้งถัดไปโดยใช้วิธี ความสัมพันธ์เชิงเส้นเชิงเดียวจากวิธีกำลังสองน้อยที่สุด (Least Square Method)โดยคำตอบจะได้สมการเชิงเส้นออกมาเป็น Y = b0 + b1X



1. ฟังก์ชันแสดงผลกราฟแสดงข้อมูลสถิติที่ได้รับเหรียญของแต่ละประเทศพร้อมทำนายในปีถัดไป



1. ฟังก์ชันหลักในการดำเนินการฟังก์ชันทั้งหมดของโปรแกรม

ตัวอย่างโปรแกรม

