

Fathom: Least Square Regression

โจทย์ จงหา Least Square Regression line ของข้อมูลในตารางต่อไปนี้

ข้อมูลเกี่ยวกับค่าตั๋วโดยสารและระยะทางจากเมืองหลวงแห่งหนึ่งไปยังเมืองต่างๆ

City	Distance (ไมล์)	Airfare (\$)	City	Distance (ไมล์)	Airfare (\$)
A	576	178	G	946	198
B	370	138	H	998	188
C	612	194	I	189	98
D	1216	278	J	787	179
E	409	158	K	210	138
F	1502	258	L	737	98

กิจกรรม

1. เปิดโปรแกรม Fathom เลือกเพิ่มใหม่
2. **สร้างตาราง** การสร้างตารางทำได้โดยการ *Drag* รูปตาราง (case table) จาก *object shelf* มาปล่อยวางที่หน้าต่างจอของ Fathom

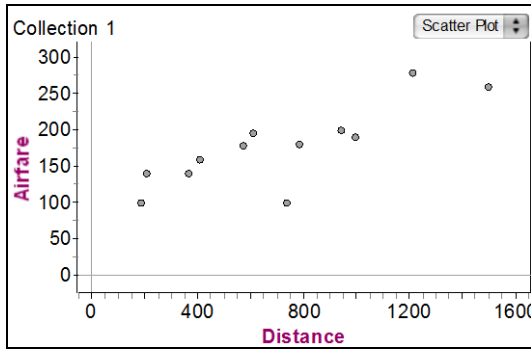
3. Click <new>

- พิมพ์ชื่อ attribute แรกคือ City
- กดแป้น Enter แล้วบันทึกข้อมูล
- กด Tab เพื่อตั้งชื่อ attribute ใน column ถัดไปแล้วบันทึกข้อมูลจนครบ

	City	Distance	Airfare
1	A	576	178
2	B	370	138
3	C	612	194
4	D	1216	278

4. **Scatter plot** การสร้างscatter plot ทำได้ดังต่อไปนี้

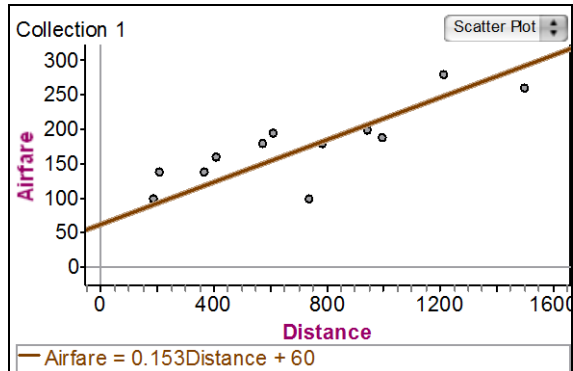
- Drag รูปกราฟ จาก *object shelf* มาปล่อยวางที่หน้าต่างจอของ Fathom
 - นำ cursor ไปที่ตาราง วางบน attribute ที่ชื่อ Distance แล้ว Drag มาปล่อยบนแกน x
 - นำ cursor ไปที่ตาราง วางบน attribute ที่ชื่อ Airfare แล้ว Drag มาปล่อยบนแกน y
- จะมีรูป Scatter plot ปรากฏที่หน้าจอตั้งรูป



- คลิกที่รูปกราฟ
- คลิกที่เมนู Collection และ
- เลือกคำสั่ง Prevent Changing Values in Graph

5. การสร้างกราฟเส้นตรงที่เลื่อนไปมาได้

- คลิกที่รูปกราฟ
- คลิกที่เมนู Graph เลือกคำสั่ง Movable Line จะมีรูปกราฟเส้นตรง และสมการปรากฏที่หน้าจอ (ดังรูป)
- นำเคอร์เซอร์วางบนกราฟและปรับรูปกราฟเส้นตรงตามต้องการ
- กรณีที่ไม่ต้องการแสดงรูปกราฟ และสมการที่หน้าจอ ให้คลิกที่รูปกราฟ และคลิกที่เมนู Graph เลือกคำสั่ง Remove Movable Line

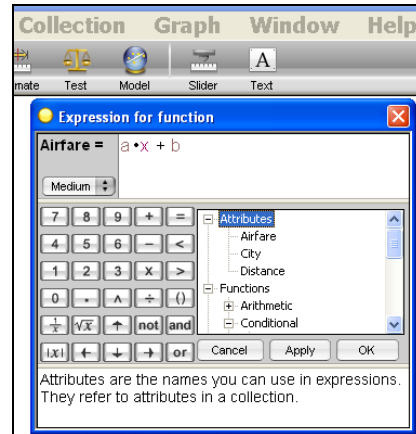


6. การสร้าง Sliders และสมการ Regression line

- Drag รูปSlider จาก *object shelf* มาปล่อยวางที่หน้าต่างจอของ Fathom
- พิมพ์ชื่อ slider เป็น a
- นำเคอร์เซอร์วางบนเส้นจำนวน ดับเบิ้ลคลิกจะมีกล่องคำสั่งของ Inspect Slider เกิดขึ้น กำหนดให้ค่าต่ำสุดของ a เป็น -1 และให้ค่าสูงสุดเป็น 1

Property	Value	Formula
a	0.36	
Max_up...		
Lower_	-1	
Upper_	1	
Restrict...		
Revers...	false	

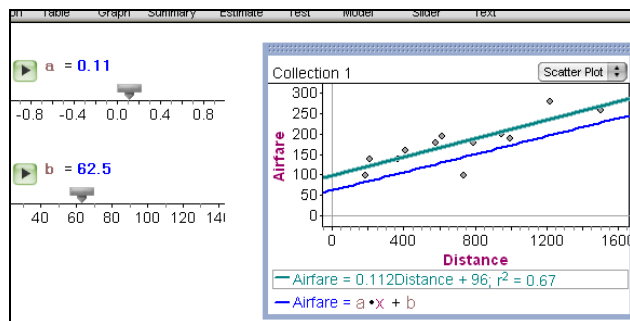
- Drag รูปSlider จาก *object shelf* มาปล่อยวางที่หน้าต่างจอ พิมพ์ชื่อ slider เป็น b
 - คลิกรูปกราฟ คลิกเมนู Graph และเลือกคำสั่ง Plot Function
 - พิมพ์ a คลิกเครื่องหมาย คูณ (x) ที่อยู่ในกล่องคำสั่ง
 - พิมพ์ x คลิกเครื่องหมายบวก และพิมพ์ b คลิก OK
- ที่รูปกราฟจะมีรูปกราฟเส้นตรงสีน้ำเงิน และสมการเกิดขึ้น



- เลื่อน sliders a และ b ไปมา เพื่อปรับค่า และให้มีรูปกราฟเส้นตรงตามต้องการ

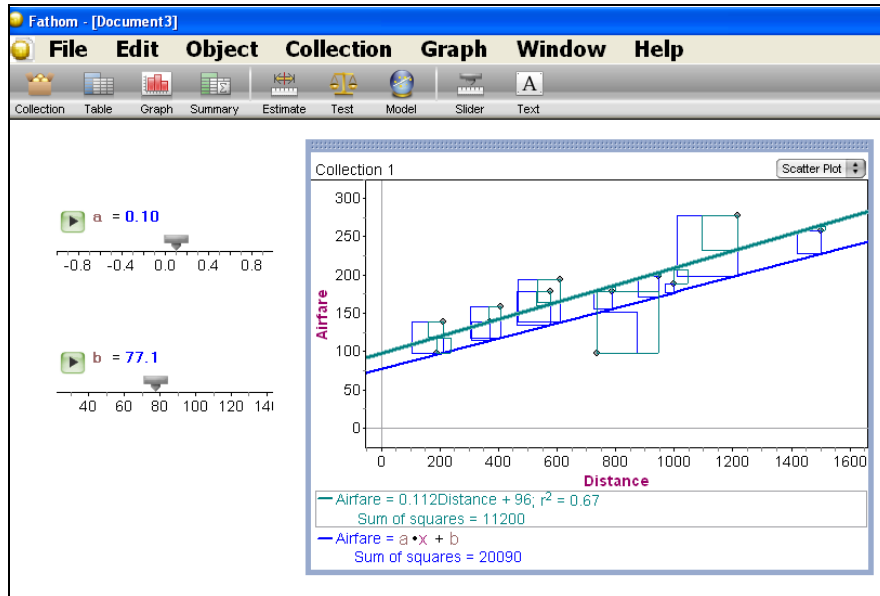
7. การสร้างรูปกราฟ Least Square Line

- คลิกที่รูปกราฟ
- คลิกที่เมนู Graph เลือกคำสั่ง Least-Squares Line จะมีรูปกราฟเส้นตรงสีเขียว สมการ และค่า r^2 ปรากฏที่หน้าจอ



8. การสร้าง Sum Squares

- คลิกที่รูปกราฟ
- คลิกที่เมนู Graph เลือกคำสั่ง Sum Squares จะมีรูปสี่เหลี่ยมจัตุรัส และ ค่า Sum of Squares เป็นสีเขียวและสีน้ำเงินปรากฏที่หน้าจอ
- เลื่อน sliders a และ b ไปมา เพื่อปรับค่า Sum of Squares ของเส้นกราฟสีน้ำเงินให้มีค่าที่ใกล้เคียงกับค่า Sum of Squares ของเส้นกราฟสีเขียวที่เป็น Least-Squares Line และมีค่า error น้อยที่สุด



9. หาLeast Squares line (Regression Line) โดยการคำนวณจากสูตร $\hat{y} = ax + b$

เมื่อ a คือ slope: $a = \frac{\sum(x-\bar{x})(y-\bar{y})}{\sum(x-\bar{x})^2}$ และ b คือ y- intercept $b = \frac{\sum y - (a\sum x)}{n}$

10. ใช้สมการ Regression Line จากข้อ 6 ข้อ 7 และข้อ 9 ทำนาย (predict) ค่าโดยสาร เมื่อต้องการเดินทางไปเป็นระยะทาง 100 ไมล์ และทำนาย (predict) ค่าโดยสาร เมื่อต้องการเดินทางไปเป็นระยะทาง 900 ไมล์, 901 ไมล์, 902 ไมล์, และ 903 ไมล์

11. การหาค่า Sum of Squared Residuals

The **residual** is the difference between the actual y-value and the fitted value \hat{y} ($residual = actual - fitted$), so that residual measures the vertical distance from the observed y-value to the regression line.

- คลิกที่รูปกราฟ
- คลิกที่เมนู Graph เลือกคำสั่ง **Make Residual Plot** จะมีรูปกราฟของ Residual เกิดขึ้นที่หน้าจอ (ดังรูป)
- เลื่อน sliders a และ b ไปมา เมื่อ Sum of Squares ของเส้นกราฟสีน้ำเงินมีค่าใกล้เคียงกับค่า Sum of Squares ของเส้นกราฟสีเขียวที่เป็น Least-Squares Line สังเกตรูปกราฟ Residual จุดที่ได้จะอยู่ใกล้เส้นตรงมากที่สุด เนื่องจาก residual มีค่าerror น้อยที่สุด

