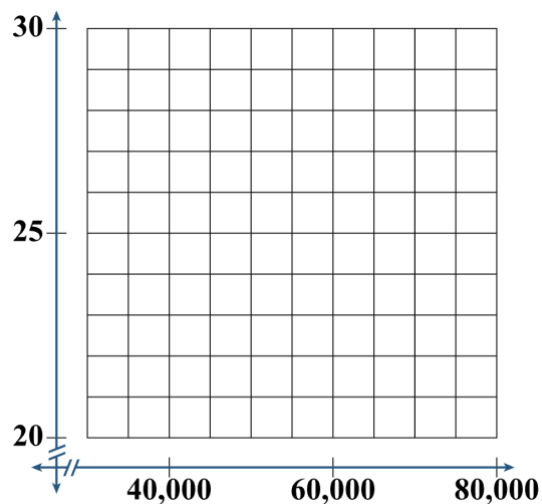


## DATA EXPLORATION

### Dataset 1: Income and Obesity Rates

- The **income rate** is the average amount of money earned per year by individuals or households in a given area.
- The **obesity rate** is the percentage of adults, aged 18 years or older, in a given area who have obesity.

| Income Rate (\$) | Obesity Rate (%) |
|------------------|------------------|
| 35,000           | 28.5             |
| 40,000           | 27.3             |
| 45,000           | 26.8             |
| 50,000           | 25.9             |
| 55,000           | 24.7             |
| 60,000           | 23.2             |
| 65,000           | 22.1             |
| 70,000           | 21.6             |
| 75,000           | 20.8             |



1) Plot the data to create a scatter plot. Let the income rate be the  $x$ -value and the obesity rate be the  $y$ -value. What trend do you notice in the data?

2) Input the data into your calculator and then use your calculator to find the line of best fit (linear regression model). Record the following values.

(slope)  $a$ :

( $y$ -intercept)  $b$ :

(correlation coefficient)  $r$ :

3) Use these values (from Step 2) to sketch the line of best fit.

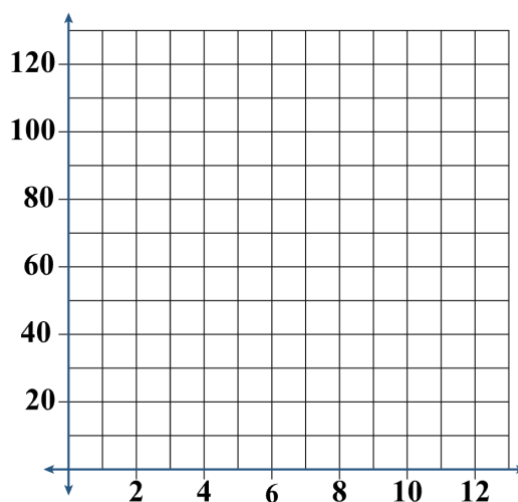
4) Do you think this line is an accurate representation of the data? Why or why not?



## Dataset 2: Days and Number of Zombies

- The **days** are the number of days since the infection began in town.
- The **number of zombies** is the quantity of zombies in town counted before the sun goes down.

| Days | Number of Zombies |
|------|-------------------|
| 0    | 2                 |
| 2    | 4                 |
| 4    | 8                 |
| 6    | 16                |
| 8    | 32                |
| 10   | 64                |
| 12   | 128               |



1) Plot the data to create a scatter plot. Let the days be the  $x$ -value and the number of zombies be the  $y$ -value. What trend do you notice in the data?

2) Input the data into your calculator and then use your calculator to find the line of best fit (linear regression model). Record the following values.

(slope)  $a$ :

( $y$ -intercept)  $b$ :

(correlation coefficient)  $r$ :

3) Use these values (from Step 2) to sketch the line of best fit.

4) Do you think this line is an accurate representation of the data? Why or why not?