# Activity Guide - Zombie Prediction

## Overview

A zombie outbreak is happening across the world! The zombies are spreading in your town, but you don't know where they are congregating! Luckily, you have data from a neighboring town about where the zombies tended to gather. Using their data, can we predict how many zombies will be at locations in our town and pick the safest location to hide?
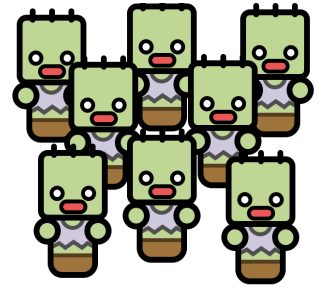
### Our Data

The table below shows the data we've received from a neighboring town

| Location | Building | Type | Noise Level | Sidewalk? | ZOMBIES 🧟 |
|---|---|---|---|---|---|
| 1 | indoor | school | medium | yes | 51 |
| 2 | outdoor | zoo | medium | yes | 66 |
| 3 | indoor | library | low | no | 9 |
| 4 | indoor | restaurant | medium | yes | 45 |
| 5 | outdoor | park | low | no | 21 |
| 6 | outdoor | school | medium | no | 54 |
| 7 | outdoor | restaurant | high | no | 60 |
| 8 | outdoor | park | high | yes | 96 |

**What are some of the things you notice in the data above?** Fill in the blanks below

- The greatest number of zombies are __96__ at location __8__
- The smallest number of zombies are __9__ at location __3__
- There tend to be more zombies in __outdoor, high/Medium noise__ locations
- There tend to be less zombies in __quiet, indoor__ locations
- What is something else that you notice? __Low noise correlates with lower counts (locations 3 & 5 have 9 and 21), While higher noise correlates with higher counts.__

_____

## Creating a Model

In order to turn this data into a model that can make a prediction, we'll take unknown data from our town and find similarities with the data you already have. Here is information about one location that is close to us:

| Location | Building | Type | Noise Level | Sidewalk? | ZOMBIES 🧟 |
|----------|----------|------|-------------|-----------|------------|
| A | indoor | restaurant | low | no | 25 |

**Do This:** Compare Location A with the features from the previous data:
- Anytime a column from Location A matches a column for another location, that's one point of similarity.
- After comparing each column, you can add up all the points of similarity.
- Use the table below to keep track of how similar Location A is to the other locations.
- The first two rows have already been completed.

| Location | Building | Type | Noise Level | Sidewalk? | ZOMBIES 🧟 | A Similarities |
|----------|----------|------|-------------|-----------|------------|----------------|
| 1 | indoor | school | medium | yes | 51 | 1 |
| 2 | outdoor | zoo | medium | yes | 66 | 0 |
| 3 | indoor | library | low | no | 9 | 3 |
| 4 | indoor | restaurant | medium | yes | 45 | 2 |
| 5 | outdoor | park | low | no | 21 | 2 |
| 6 | outdoor | school | medium | no | 54 | 1 |
| 7 | outdoor | restaurant | high | no | 60 | 2 |
| 8 | outdoor | park | high | yes | 96 | 0 |

**What are three locations that are most similar to Location A?** If some locations have the same level of similarity, choose the ones highest in the list.

The three locations most similar to Location A are: __3__ , __4__ , and __5__

To predict how many zombies could be at this location, we need to take the average of the three most similar locations.
1. Calculate the <u>sum</u> by adding together all the zombie values from these locations: __75__
2. Divide the sum by the number of values: __3__
3. Write this value in the Zombies column at the top of this page. That's the prediction!

## Predicting Multiple Locations

There are two other locations that are nearby. We need to decide whether to go hide at location A, B, or C. Use the same strategy to predict how many zombies will be at location B and at location C.

| Location | Indoor/Outdoor | Type | Noise Level | Sidewalk? | ZOMBIES 🧟 |
|---|---|---|---|---|---|
| B | outdoor | restaurant | high | yes | 74 |
| C | indoor | school | low | yes | 35 |

| Location | Building | Type | Noise Level | Sidewalk | ZOMBIES | B Similarities | C Similarities |
|---|---|---|---|---|---|---|---|
| 1 | indoor | school | medium | yes | 51 | 1 | 3 |
| 2 | outdoor | zoo | medium | yes | 66 | 2 | 1 |
| 3 | indoor | library | low | no | 9 | 0 | 2 |
| 4 | indoor | restaurant | medium | yes | 45 | 2 | 2 |
| 5 | outdoor | park | low | no | 21 | 1 | 1 |
| 6 | outdoor | school | medium | no | 54 | 1 | 1 |
| 7 | outdoor | restaurant | high | no | 60 | 3 | 0 |
| 8 | outdoor | park | high | yes | 96 | 3 | 1 |

**What are three locations that are most similar to Location B?** If some locations have the same level of similarity, choose the ones highest in the list.

The three locations most similar to Location B are: __7__, __8__, and __2__

To predict how many zombies could be at this location, we need to take the **average** of the three most similar locations.
1. Calculate the <u>sum</u> by adding together all the zombie values from these locations: __222__
2. Divide the sum by the number of values: __74__
3. Write this value in the Zombies column at the top of this page. That's the prediction!

**What are three locations that are most similar to Location C?** If some locations have the same level of similarity, choose the ones highest in the list.

The three locations most similar to Location C are: __1__, __3__, and __4__

To predict how many zombies could be at this location, we need to take the **average** of the three most similar locations.
1. Calculate the <u>sum</u> by adding together all the zombie values from these locations: __105__
2. Divide the sum by the number of values: __35__
3. Write this value in the Zombies column at the top of this page. That's the prediction!

**Which location is predicted to have the least amount of zombies - location A, location B, or location C?**

Location A