Problem Setup

Suppose there are 8000 people in a town. Out of the 8000 people:

- 800 are females
- 7200 are males

In that town, 1600 people are employed, of which:

- 120 are females
- 1480 are males

We'll use this information to calculate the following metrics:

- **Relative Risk**
- **Odds of Success for Each Gender**
- **Odds Ratio**
- **Verification of**

$$\pi_i = \frac{\Omega_i}{1 + \Omega_i} \quad \text{for } i = 1, 2$$

- **Comment on the Association Based on the Value of** θ (Odds Ratio)
- **Verification that**

$$\text{Odds Ratio} = \text{Relative Risk} \times \frac{1 - \pi_2}{1 - \pi_1}$$

Step 1: Create the Data Frame and Contingency Table

First, we set up the data frame with the given information and create a contingency table to work from.

```
# Define the data frame with gender, employment status, and count
employee_data <- data.frame(
  Sex = rep(c("Female", "Male"), each = 2),
  Status = rep(c("Employed", "Unemployed"), times = 2),
  Count = c(120, 680, 1480, 5720)
)

# Display the data frame
print(employee_data)
```

```
##      Sex     Status Count
## 1 Female   Employed   120
## 2 Female Unemployed   680
## 3   Male   Employed  1480
## 4   Male Unemployed  5720
```

```r
# Create a contingency table from the data frame
contingency_table <- xtabs(Count ~ Sex + Status, data = employee_data)
print(contingency_table)
```

```
##         Status
## Sex      Employed Unemployed
##   Female      120        680
##   Male       1480       5720
```

Part (a): Calculate Relative Risk

Relative risk compares the probability of employment for females to the probability of employment for males.

Steps to Calculate Relative Risk:

1. **Calculate the employment probability (π) for each gender.**
2. **Compute the relative risk as the ratio of these probabilities.**

```r
# Calculate employment probabilities for each gender
pi <- c()
for (i in 1:2) {
  pi[i] <- contingency_table[i, "Employed"] / sum(contingency_table[i, ])
}

# Calculate relative risk
relative_risk <- pi[1] / pi[2]
relative_risk
```

```
## [1] 0.7297297
```

Part (b): Calculate Odds of Success (Employment) for Each Gender

The odds of success, or the odds of employment for each gender, are calculated by dividing the probability of employment by the probability of non-employment for each gender.

Steps to Calculate Odds of Success:

1. **Use the contingency table to find the counts of employed and unemployed individuals for each gender.**
2. **Calculate the odds as the ratio of employed to unemployed individuals for each gender.**

```r
# Calculate odds of success (employment) for each gender
odds_success <- c()
for (i in 1:2) {
  odds_success[i] <- contingency_table[i, "Employed"] / contingency_table[i, "Unemployed"]
}
odds_success
```

```
## [1] 0.1764706 0.2587413
```

Part (c): Calculate Odds Ratio

The odds ratio compares the odds of employment for females to the odds of employment for males.

Steps to Calculate Odds Ratio:

1. **Use the odds of success for each gender calculated in Part (b).**
2. **Calculate the odds ratio as the ratio of the odds of success for females to the odds of success for males.**

```
# Calculate odds ratio
odds_ratio <- odds_success[1] / odds_success[2]
odds_ratio
```

```
## [1] 0.682035
```

Part (d): Verify $\pi_i = \frac{\Omega_i}{1+\Omega_i}$ for $i = 1, 2$

The formula $\pi_i = \frac{\Omega_i}{1+\Omega_i}$ shows that the probability π_i for each gender can be derived from the odds Ω_i.

Verification Steps:

1. **Calculate π_i using the odds Ω_i for each gender.**
2. **Compare these values to the employment probabilities calculated in Part (a) to verify they match.**

```
# Verify the relationship pi_i = omega_i / (1 + omega_i)
pi_verified <- odds_success / (1 + odds_success)
all.equal(pi, pi_verified) # Checks if the calculated values are the same
```

```
## [1] TRUE
```

Part (e): Comment on Association Based on Odds Ratio

The odds ratio (θ) helps to understand the association between being female and the likelihood of employment:

- **If** $\theta = 1$: No association between gender and employment.
- **If** $\theta > 1$: Positive association (females are more likely to be employed than males).
- **If** $\theta < 1$: Negative association (females are less likely to be employed than males).

Interpretation:

Based on the value of θ calculated in Part (c), we will conclude if there is a positive or negative association between gender and employment likelihood.

```r
# Interpretation of association based on the odds ratio calculated
if (odds_ratio > 1) {
  association <- "Positive association: females are more likely to be employed than males."
} else if (odds_ratio < 1) {
  association <- "Negative association: females are less likely to be employed than males."
} else {
  association <- "No association: gender does not affect employment likelihood."
}
association
```

```
## [1] "Negative association: females are less likely to be employed than males."
```

Part (f): Verify that Odds Ratio = Relative Risk $\times \frac{1-\pi_2}{1-\pi_1}$

This verification step ensures that the odds ratio can be expressed as the product of relative risk and the adjustment factor involving probabilities π_1 and π_2.

Verification Steps:

1. **Calculate the adjustment factor $\frac{1-\pi_2}{1-\pi_1}$.**
2. **Multiply the relative risk by this factor and check if it equals the odds ratio.**

```r
# Verify the relationship
verification_value <- relative_risk * ((1 - pi[2]) / (1 - pi[1]))
all.equal(odds_ratio, verification_value) # Checks if both values are approximately equal
```

```
## [1] TRUE
```