Bacterial Meningitis Incidence and Linear Regression

201561866

2024-03-23

```
## Incidence and Linear Regression

install.packages("rmarkdown")
```

```
## Installing package into '/cloud/lib/x86_64-pc-linux-gnu-library/4.3'
## (as 'lib' is unspecified)
```

```
install.packages("knitr")
```

```
## Installing package into '/cloud/lib/x86_64-pc-linux-gnu-library/4.3'
## (as 'lib' is unspecified)
```

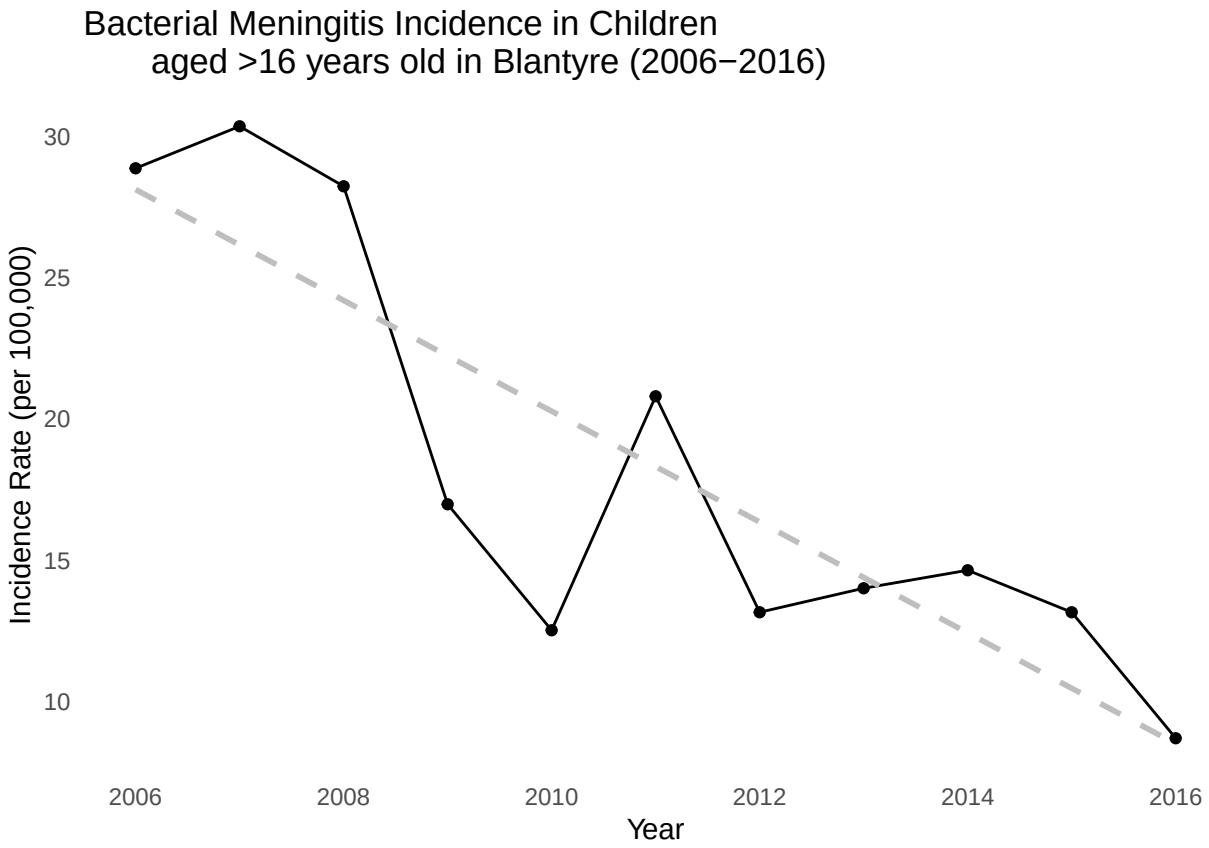
```
library(ggplot2)

data <- data.frame(
  Year = 2006:2016,
  Incidence = c(28.85898777, 30.34437685, 28.22239245, 16.97587516,
                12.51970793, 20.79544707, 13.15630325, 14.00509701,
                14.64169232, 13.15630325, 8.700136019)
)

ggplot(data, aes(x = Year, y = Incidence)) +
  geom_line(group=1) +  # Connect points with a line
  geom_point() +  # Add points for each year
  geom_smooth(method = "lm", se = FALSE, color = "grey", linetype = "dashed") +

  # Add linear trendline, without confidence interval
  labs(title = "Bacterial Meningitis Incidence in Children
       aged >16 years old in Blantyre (2006-2016)",
       x = "Year",
       y = "Incidence Rate (per 100,000)") +
  theme_minimal() +
  theme(panel.grid.major = element_blank(),   # Remove major grid lines
        panel.grid.minor = element_blank())   # Remove minor grid lines
```

```
## `geom_smooth()` using formula = 'y ~ x'
```




```
# Save the plot
ggsave("Incidence_Trend_with_Trendline.png", width = 8, height = 4)
```

```
## `geom_smooth()` using formula = 'y ~ x'
```

```
# statistical analysis to assess the trend
# fit a linear model to the data
model <- lm(Incidence ~ Year, data = data)

# Model Summary: check the p-value for the Year coefficient
summary_result <- summary(model)
print(summary_result)
```

```
##
## Call:
## lm(formula = Incidence ~ Year, data = data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -7.7472 -1.7863  0.7523  2.5888  4.1977
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 3959.7502   784.7517   5.046 0.000694 ***
## Year          -1.9599     0.3902  -5.023 0.000717 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Residual standard error: 4.093 on 9 degrees of freedom
## Multiple R-squared:  0.737,  Adjusted R-squared:  0.7078
## F-statistic: 25.23 on 1 and 9 DF,  p-value: 0.0007166
```

```
# Model Summary: trend and statistical significance
summary(model)
```

```
##
## Call:
## lm(formula = Incidence ~ Year, data = data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -7.7472 -1.7863  0.7523  2.5888  4.1977
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 3959.7502   784.7517   5.046 0.000694 ***
## Year          -1.9599     0.3902  -5.023 0.000717 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.093 on 9 degrees of freedom
## Multiple R-squared:  0.737,  Adjusted R-squared:  0.7078
## F-statistic: 25.23 on 1 and 9 DF,  p-value: 0.0007166
```

```
years <- 2006:2016
incidence <- c(28.85898777, 30.34437685, 28.22239245, 16.97587516,
               12.51970793, 20.79544707, 13.15630325, 14.00509701,
               14.64169232, 13.15630325, 8.700136019)

data <- data.frame(Year = years, Incidence = incidence)

# Linear model
model <- lm(Incidence ~ Year, data = data)

# Annual rate of decrease
rate_of_decrease <- coef(model)["Year"]

# Mean incidence over the period
mean_incidence <- mean(data$Incidence)

# Average percentage change per year based on the rate of decrease
average_percentage_change <- (rate_of_decrease / mean_incidence) * 100

#Print
print(paste("The incidence was decreasing at an average rate of",
            rate_of_decrease, "per year."))
```

```
## [1] "The incidence was decreasing at an average rate of -1.95994195022714 per year."
```

```
print(paste("This is an average percentage change of", average_percentage_change,
            "% per year."))
```

```
## [1] "This is an average percentage change of -10.7060063224419 % per year."
```

```r
## Incidence and Linear Regression

install.packages("rmarkdown")
```

```
## Installing package into '/cloud/lib/x86_64-pc-linux-gnu-library/4.3'
## (as 'lib' is unspecified)
```

```r
install.packages("knitr")
```

```
## Installing package into '/cloud/lib/x86_64-pc-linux-gnu-library/4.3'
## (as 'lib' is unspecified)
```

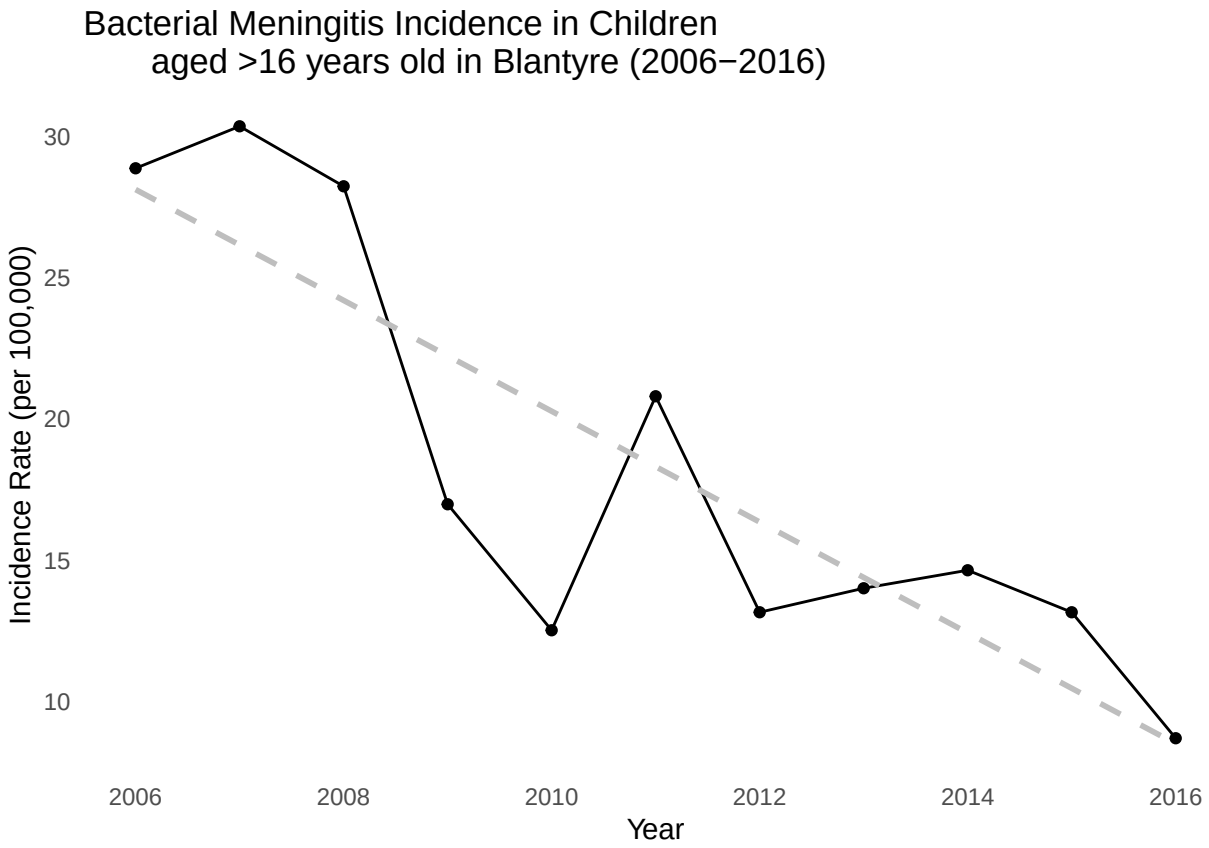
```r
library(ggplot2)

data <- data.frame(
  Year = 2006:2016,
  Incidence = c(28.85898777, 30.34437685, 28.22239245, 16.97587516,
                12.51970793, 20.79544707, 13.15630325, 14.00509701,
                14.64169232, 13.15630325, 8.700136019)
)

ggplot(data, aes(x = Year, y = Incidence)) +
  geom_line(group=1) +  # Connect points with a line
  geom_point() +  # Add points for each year
  geom_smooth(method = "lm", se = FALSE, color = "grey", linetype = "dashed") +

  # Add linear trendline, without confidence interval
  labs(title = "Bacterial Meningitis Incidence in Children
       aged >16 years old in Blantyre (2006-2016)",
       x = "Year",
       y = "Incidence Rate (per 100,000)") +
  theme_minimal() +
  theme(panel.grid.major = element_blank(),   # Remove major grid lines
        panel.grid.minor = element_blank())   # Remove minor grid lines
```

```
## `geom_smooth()` using formula = 'y ~ x'
```




```
# Save the plot
ggsave("Incidence_Trend_with_Trendline.png", width = 8, height = 4)
```

```
## `geom_smooth()` using formula = 'y ~ x'
```

```
# statistical analysis to assess the trend
# fit a linear model to the data
model <- lm(Incidence ~ Year, data = data)

# Model Summary: check the p-value for the Year coefficient
summary_result <- summary(model)
print(summary_result)
```

```
##
## Call:
## lm(formula = Incidence ~ Year, data = data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -7.7472 -1.7863  0.7523  2.5888  4.1977
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 3959.7502   784.7517   5.046 0.000694 ***
## Year          -1.9599     0.3902  -5.023 0.000717 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Residual standard error: 4.093 on 9 degrees of freedom
## Multiple R-squared:  0.737,  Adjusted R-squared:  0.7078
## F-statistic: 25.23 on 1 and 9 DF,  p-value: 0.0007166
```

```
# Model Summary: trend and statistical significance
summary(model)
```

```
##
## Call:
## lm(formula = Incidence ~ Year, data = data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -7.7472 -1.7863  0.7523  2.5888  4.1977
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 3959.7502   784.7517   5.046 0.000694 ***
## Year          -1.9599     0.3902  -5.023 0.000717 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.093 on 9 degrees of freedom
## Multiple R-squared:  0.737,  Adjusted R-squared:  0.7078
## F-statistic: 25.23 on 1 and 9 DF,  p-value: 0.0007166
```

```
years <- 2006:2016
incidence <- c(28.85898777, 30.34437685, 28.22239245, 16.97587516,
               12.51970793, 20.79544707, 13.15630325, 14.00509701,
               14.64169232, 13.15630325, 8.700136019)

data <- data.frame(Year = years, Incidence = incidence)

# Linear model
model <- lm(Incidence ~ Year, data = data)

# Annual rate of decrease
rate_of_decrease <- coef(model)["Year"]

# Mean incidence over the period
mean_incidence <- mean(data$Incidence)

# Average percentage change per year based on the rate of decrease
average_percentage_change <- (rate_of_decrease / mean_incidence) * 100

#Print
print(paste("The incidence was decreasing at an average rate of",
            rate_of_decrease, "per year."))
```

```
## [1] "The incidence was decreasing at an average rate of -1.95994195022714 per year."
```

```
print(paste("This is an average percentage change of", average_percentage_change,
            "% per year."))
```

```
## [1] "This is an average percentage change of -10.7060063224419 % per year."
```