

Activity-asp

2025-05-28

R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.

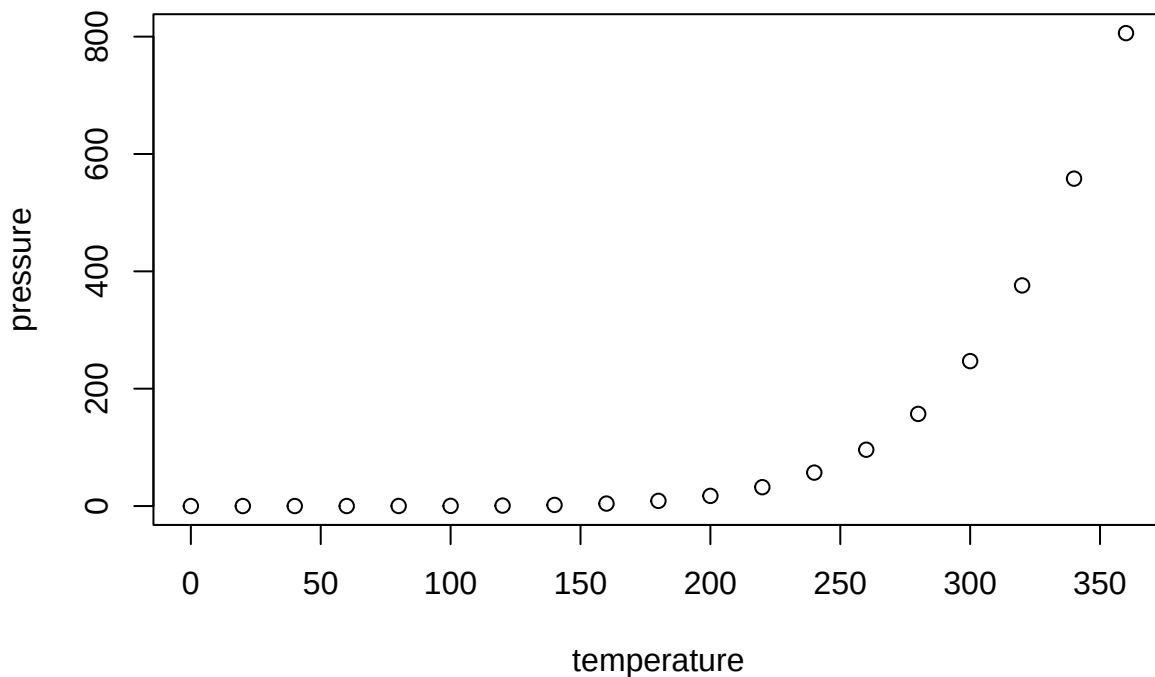
When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

```
summary(cars)
```

```
##      speed      dist
##  Min.   : 4.0    Min.   :  2.00
## 1st Qu.:12.0    1st Qu.: 26.00
##  Median:15.0    Median : 36.00
##  Mean   :15.4    Mean   : 42.98
## 3rd Qu.:19.0    3rd Qu.: 56.00
##  Max.   :25.0    Max.   :120.00
```

Including Plots

You can also embed plots, for example:



Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.

```
install.packages("gapminder") # Run only once library(gapminder) data("gapminder")
```

```
# Load library and dataset
library(gapminder)
data("gapminder")
```

1. Subset countries with GDP > threshold (GDP = gdpPercap * pop)

```
gapminder$gdp <- gapminder$gdpPercap * gapminder$pop
high_gdp <- subset(gapminder, gdp > 1e11) # GDP > 100 billion
print("Countries with high GDP:")
```

```
## [1] "Countries with high GDP:"
```

```
print(unique(high_gdp$country))
```

```
## [1] Algeria      Argentina    Australia    Austria
## [5] Bangladesh   Belgium      Brazil        Canada
## [9] Chile         China        Colombia      Cuba
## [13] Czech Republic Denmark      Egypt         Finland
## [17] France        Germany      Greece         Hong Kong, China
## [21] Hungary       India        Indonesia      Iran
## [25] Iraq          Ireland      Israel         Italy
## [29] Japan         Korea, Rep.  Kuwait         Malaysia
## [33] Mexico        Morocco     Netherlands    New Zealand
## [37] Nigeria       Norway       Pakistan       Peru
## [41] Philippines   Poland       Portugal       Romania
## [45] Saudi Arabia  Serbia       Singapore      Slovak Republic
## [49] South Africa  Spain        Sudan          Sweden
## [53] Switzerland  Taiwan       Thailand       Turkey
## [57] United Kingdom United States Venezuela      Vietnam
## 142 Levels: Afghanistan Albania Algeria Angola Argentina Australia ... Zimbabwe
```

2. Factorise the continent variable

```
gapminder$continent <- factor(gapminder$continent)
print("Levels in continent factor:")
```

```
## [1] "Levels in continent factor:"
```

```
print(levels(gapminder$continent))
```

```
## [1] "Africa" "Americas" "Asia" "Europe" "Oceania"
```

```
# 3. Compute mean life expectancy per continent (vectorised)
```

```
mean_lifeExp <- tapply(gapminder$lifeExp, gapminder$continent, mean)
print("Mean Life Expectancy per Continent:")
```

```
## [1] "Mean Life Expectancy per Continent:"
```

```
print(mean_lifeExp)
```

```
## Africa Americas Asia Europe Oceania
## 48.86533 64.65874 60.06490 71.90369 74.32621
```

```

# 4. Extract matrix of population for selected countries
selected_countries <- c("India", "China", "United States", "Nigeria")
pop_matrix <- subset(gapminder, country %in% selected_countries)[, c("country", "year", "pop")]
# Convert to matrix with countries as rows, years as columns
pop_wide <- reshape(pop_matrix, timevar = "year", idvar = "country", direction = "wide")
rownames(pop_wide) <- pop_wide$country

## Warning: Setting row names on a tibble is deprecated.

pop_matrix_only <- as.matrix(pop_wide[, -1])
print("Population Matrix:")

## [1] "Population Matrix:"
print(pop_matrix_only)

##      pop.c(1952, 1957, 1962, 1967, 1972, 1977, 1982, 1987, 1992, 1997, 2002, 2007)
## [1,]                                                                                     NA
## [2,]                                                                                     NA
## [3,]                                                                                     NA
## [4,]                                                                                     NA

# 5. Convert country names to uppercase
gapminder$country_upper <- toupper(gapminder$country)
print("Sample country names in uppercase:")

## [1] "Sample country names in uppercase:"
print(unique(gapminder$country_upper)[1:10])

## [1] "AFGHANISTAN" "ALBANIA"      "ALGERIA"      "ANGOLA"      "ARGENTINA"
## [6] "AUSTRALIA"   "AUSTRIA"     "BAHRAIN"     "BANGLADESH"  "BELGIUM"

```

Activity 2

```

# Mean of life expectancy
# Calculate Mean of life expectancy
mean_lifeExp <- mean(gapminder$lifeExp)
mean_lifeExp

## [1] 59.47444

# Calculate Median of life expectancy
median_lifeExp <- median(gapminder$lifeExp)
median_lifeExp

## [1] 60.7125

# Define a function to calculate Mode
getmode <- function(v) {
  univq <- unique(v)
  univq[which.max(tabulate(match(v, univq)))]
}

# Calculate Mode of continent
mode_continent <- getmode(gapminder$continent)
mode_continent

```

```
## [1] Africa
## Levels: Africa Americas Asia Europe Oceania
```

Activity 3

```
# Variance of life expectancy
variance_lifeExp <- var(gapminder$lifeExp)
variance_lifeExp
```

```
## [1] 166.8517
```

```
# Standard deviation of life expectancy
sd_lifeExp <- sd(gapminder$lifeExp)
sd_lifeExp
```

```
## [1] 12.91711
```

```
# Range of life expectancy
range_lifeExp <- range(gapminder$lifeExp)
range_lifeExp
```

```
## [1] 23.599 82.603
```

```
# Summary statistics of life expectancy
summary_lifeExp <- summary(gapminder$lifeExp)
summary_lifeExp
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  23.60   48.20   60.71   59.47   70.85   82.60
```

```
# Interquartile range (IQR) of life expectancy
iqr_lifeExp <- IQR(gapminder$lifeExp)
iqr_lifeExp
```

```
## [1] 22.6475
```