````
---
title: "Project 2"
author: "Justin Williams"
date: "`r Sys.Date()`"
output: html_document
---

---
title: "Project 2"
author: "Justin Williams"
date: "`r Sys.Date()`"
output: html_document
---

Let's load some data and packages!
```{r}
knitr::opts_chunk$set(echo = TRUE)
library(tidyverse)
library(dplyr)
library(readxl)
Project2_JW <- read_excel("/Users/justinwilliams/Code/9050advresearch/Project
2/Project2_JW.xlsx")
if (interactive()) {
  View(Project2_JW)
}
```

# 1. Report the descriptive statistics along with the frequency distribution and provide a
detailed interpretation of how you would characterize the salary variable.
## Summary Statistics
```{r setup, include=TRUE}
summary(Project2_JW$Salary)
```

## Mean
```{r}
#mean
mean_salary <- mean(Project2_JW$Salary)
mean_salary
```

## Median
```{r}
#median
median_salary <- median(Project2_JW$Salary)
median_salary
```

## Mode
```{r}
#mode
mode_salary <- function(x) {
  ux <- unique(x)
  ux[which.max(tabulate(match(x, ux)))]
}
mode_salary_value <- mode_salary(Project2_JW$Salary)
mode_salary_value
```

## Frequency Table
```{r}
#frequency table
salary_table <- table(Project2_JW$Salary)
````
````

```
print(salary_table)
barplot(salary_table, main="Salary Distribution", xlab="Salary", ylab="Frequency",
col="blue")
```

Interpretation
Summary Interpretation
```{r, echo=FALSE}
cat("The mean salary is", mean_salary, "\n")
cat("The median salary is", median_salary, "\n")
cat("The mode salary is", mode_salary_value, "\n")
cat("The frequency distribution of salaries is shown in the bar plot above.\n")
```

Detailed Interpretation
```{r, echo=FALSE}
cat("The salary variable shows the following characteristics:\n")
cat("1. The mean salary is", mean_salary, "which indicates the average salary of the
employees.\n")
cat("2. The median salary is", median_salary, "which is the middle value when the salaries
are sorted in ascending order. This suggests that half of the employees earn less than",
median_salary, "and half earn more.\n")
cat("3. The mode salary is", mode_salary_value, "which is the most frequently occurring
salary in the dataset.\n")
cat("4. The frequency distribution bar plot shows the distribution of salaries across
different ranges. This helps in visualizing how salaries are spread out among the
employees.\n")
cat("Overall, the salary variable can be characterized by its central tendency measures
(mean, median, mode) and its distribution as shown in the frequency table and bar
plot.\n")
summary(Project2_JW$Salary)
```