R Notebook

```
#packages
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 4.2.3
```

```
library(knitr)
library(dplyr)
```

```
## Warning: package 'dplyr' was built under R version 4.2.3
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(gridExtra)
```

```
##
## Attaching package: 'gridExtra'
```

```
## The following object is masked from 'package:dplyr':
##
##     combine
```

```
#define function
set.seed(123)

simulate_t_tests <- function(n_sims, n1, n2, true_diff, sim_id) {
  p_values <- numeric(n_sims)
  test_stats <- numeric(n_sims)
  estimates <- numeric(n_sims)
  rejections <- logical(n_sims)

  for (i in 1:n_sims) {
    x1 <- rnorm(n1, mean = 0, sd = 1)
    x2 <- rnorm(n2, mean = 0.5, sd = 3)
```

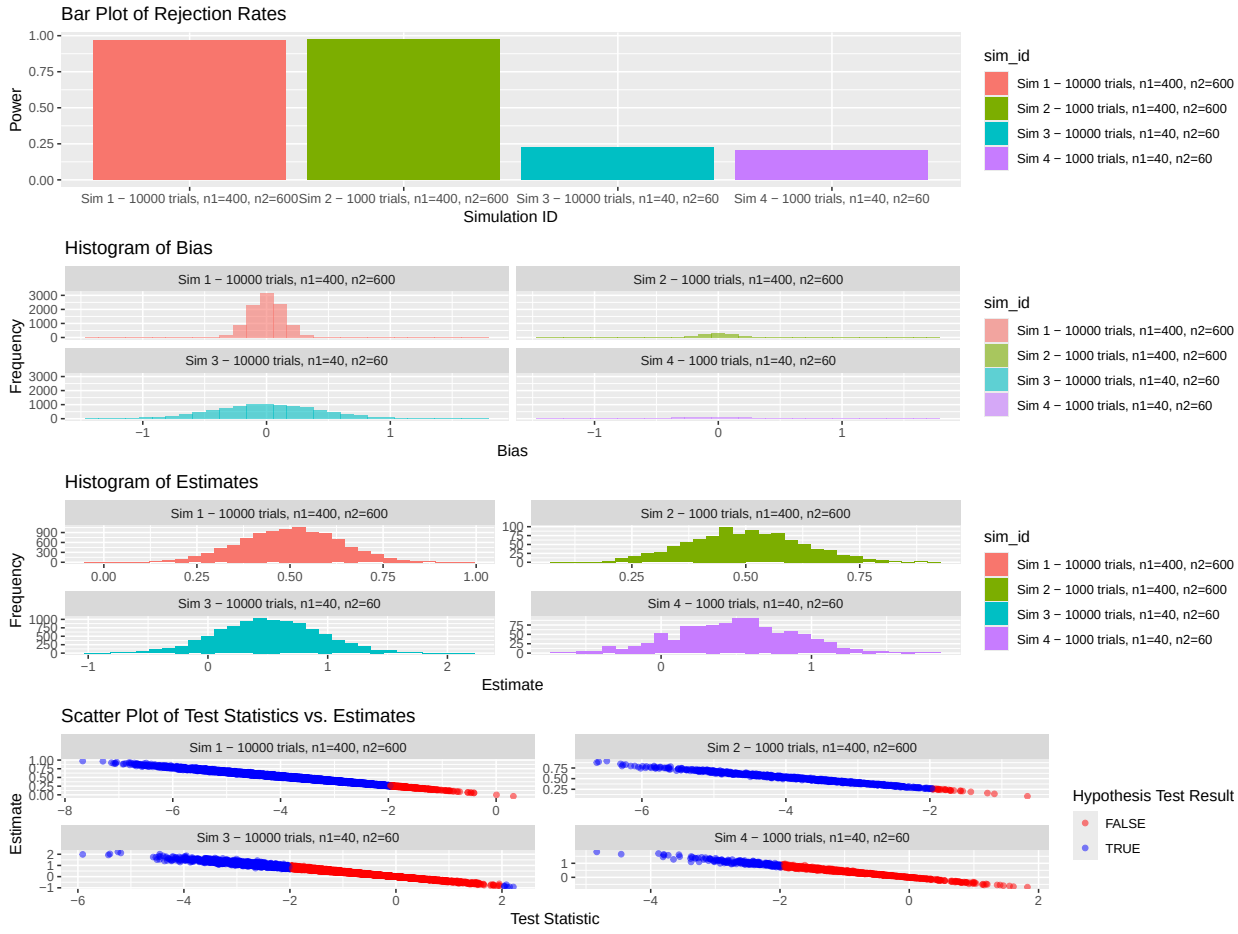
```
    test_result <- t.test(x1, x2)

    p_values[i] <- test_result$p.value
    test_stats[i] <- test_result$statistic
    estimates[i] <- test_result$estimate[2] - test_result$estimate[1]
    rejections[i] <- test_result$p.value < 0.05
  }

  simulation_data <- data.frame(
    sim_id = rep(sim_id, n_sims),
    p_values = p_values,
    test_stats = test_stats,
    estimates = estimates,
    rejections = rejections,
    bias = estimates - true_diff
  )

   return(simulation_data)
}
```

```
all_data <- bind_rows(
  simulate_t_tests(10000, 400, 600, 0.5, "Sim 1"),
  simulate_t_tests(10, 400, 600, 0.5, "Sim 2"),
  simulate_t_tests(10000, 40, 60, 0.5, "Sim 3"),
  simulate_t_tests(10, 40, 60, 0.5, "Sim 4")
)

# Calculate mean statistics for each simulation
mean_stats <- all_data %>%
  group_by(sim_id) %>%
  summarise(
    Mean_P_Value = mean(p_values),
    Mean_Test_Stat = mean(test_stats),
    Mean_Estimate = mean(estimates),
    Power = mean(rejections),
    Mean_Bias = mean(bias)
  )
```

```
all_data <- bind_rows(
  simulate_t_tests(10000, 400, 600, 0.5, "Sim 1 - 10000 trials, n1=400, n2=600"),
  simulate_t_tests(1000, 400, 600, 0.5, "Sim 2 - 1000 trials, n1=400, n2=600"),
  simulate_t_tests(10000, 40, 60, 0.5, "Sim 3 - 10000 trials, n1=40, n2=60"),
  simulate_t_tests(1000, 40, 60, 0.5, "Sim 4 - 1000 trials, n1=40, n2=60")
)
```

```
power_plot <- all_data %>%
  group_by(sim_id) %>%
  summarise(Power = mean(rejections)) %>%
  ggplot(aes(x = sim_id, y = Power, fill = sim_id)) +
  geom_bar(stat = "identity") +
  labs(title = "Bar Plot of Rejection Rates", x = "Simulation ID", y = "Power")

histogram_bias <- ggplot(all_data, aes(x = bias, fill = sim_id)) +
```

```
  geom_histogram(bins = 30, alpha = 0.6) +
  facet_wrap(~sim_id) +
  labs(title = "Histogram of Bias", x = "Bias", y = "Frequency")

estimate_plot <- ggplot(all_data, aes(x = estimates, fill = sim_id)) +
  geom_histogram(position = "dodge", bins = 30) +
  facet_wrap(~sim_id, scales = "free") +
  labs(title = "Histogram of Estimates", x = "Estimate", y = "Frequency")

scatter_plot <- ggplot(all_data, aes(x = test_stats, y = estimates, color = rejections)) +
  geom_point(alpha = 0.5) +
  facet_wrap(~sim_id, scales = "free") +
  scale_color_manual(values = c("red", "blue")) +
  labs(title = "Scatter Plot of Test Statistics vs. Estimates", x = "Test Statistic", y = "Estimate", c

grid.arrange(power_plot,histogram_bias, estimate_plot, scatter_plot, ncol = 1)
```




```
print(mean_stats)
```

```
## # A tibble: 4 x 6
##   sim_id Mean_P_Value Mean_Test_Stat Mean_Estimate Power Mean_Bias
##   <chr>         <dbl>          <dbl>         <dbl> <dbl>     <dbl>
```

```
## 1 Sim 1       0.00802          -3.78         0.499 0.964 -0.000796
## 2 Sim 2       0.00371          -3.55         0.471 1     -0.0289
## 3 Sim 3       0.324            -1.20         0.496 0.209 -0.00381
## 4 Sim 4       0.388            -1.17         0.476 0.1   -0.0237
```

Decreasing Sample Size: Smaller sample sizes (Simulations 3 and 4 with n1=40, n2=60) led to higher variance in estimates and reduced power, making it less likely to correctly reject the null hypothesis. This also resulted in a higher mean bias, with a large variance as denoted by the bias histogram for simulation 3.

Increasing Sample Size: Larger sample sizes (Simulations 1 and 2 with n1=400, n2=600) improved the precision and power of the test. This is reflected in the broader distributions of estimates, suggesting more reliable detection of true effects.

Impact of Number of Simulations (n_sims): More simulations (Sim 1 and 3 with 10,000 trials) enhanced the stability and accuracy of the estimates, shown by increased power and reduces bias (note that power for sim 3 is 0.2, which is still more than double of sim 4).

Confidence Intervals and Histogram of Estimates: Narrow distributions in Simulations 1 and 2 indicate tight confidence intervals, which get us pretty close to the true value of our parameter (0.5). Distributions in Simulations 3 and 4 suggest wide confidence intervals, increasing the uncertainty in the estimate of our true_diff parameter.