powerAnalysis

Research Question:

Does the gender or race of a potential customer affect response rates when requesting catering orders from U.S. states that historically supported slavery?

Data Structure Plan of Record:

- Factor 1 Geographic Block: union, confederacy
  - Factor 2 Race Treatment: white / black names
    - Factor 3 Gender Treatment: male / female names
      - Outcome Variable 1: response (binary)
      - Outcome Variable 2: budget acceptance (binary / NA)
- Total combinations: 2^3 = 2x2x2 = 8 experimental cells

Draft Response Rate Map of Scenario 1:

Rosen, J. (2010). *Legislative responsiveness to constituent ethnicity and grammar quality: A field experiment.*

Geography	Race	Gender	Experiment Mapping	Gender Adjustment	Final Rate
Union	White	Male	52	0	52
Union	White	Female	52	TBD	TBD
Union	Black	Male	37	0	37
Union	Black	Female	37	TBD	TBD
Confederate	White	Male	29	0	29
Confederate	White	Female	29	TBD	TBD
Confederate	Black	Male	34	0	34
Confederate	Black	Female	34	TBD	TBD

Drawn directly from the cell averages observed in the Jose versus Colin multi-factor experiment, which analyzed response rates in correspondence audits based on perceived ethnicity and grammar quality. Specifically, the 52% rate represents the response rate observed in the most favorable condition of that study (Colin with good grammar), which is used here to anchor the White Male response rate in the expected high-response Union region. The other source rates, 37% and 34%, correspond to less favorable experimental conditions in that study (e.g., Colin with bad grammar or Jose with bad grammar, respectively), reflecting lower response probability due to disadvantageous traits. These rates are applied across the geographies (Union/Confederate) and races (White/Black) to model expected discrimination, with lower rates generally assigned to the historically constrained Confederate region and to Black profiles. The table sets these initial rates as the male baseline (Gender Adjustment = 0), establishing the foundation for future analysis aimed at determining the necessary "Gender Adjustment" to calculate the final response rates for females in the study.

Draft Response Rate Map Scenario 2

Bertrand, M., & Mullainathan, S. (2004). Are Emily and Greg more employable than Lakish and Jamal? A field experiment on labor market discrimination. American Economic Review.

Geography	Race	Gender	Experiment Mapping	Gender Adjustment	Final Rate
Union	White	Male	11.88	0.00	11.88
Union	White	Female	11.88	+1.46	13.34
Union	Black	Male	7.83	0.00	7.83
Union	Black	Female	7.83	+0.87	8.70
Confederate	White	Male	8.61	0.00	8.61
Confederate	White	Female	8.61	+1.05	9.66
Confederate	Black	Male	5.81	0.00	5.81
Confederate	Black	Female	5.81	+0.64	6.45

Callback rates derived from the Bertrand and Mullainathan (B&M) field experiment, "Are Emily and Greg More Employable than Lakisha and Jamal?". The rates used here are the percentage callback rates from the B&M study. To proxy for the geographical split (Union/Confederate), this simulation uses the observed callback rates from B&M's Boston data (higher rates, proxy for Union) and Chicago data (lower rates, proxy for Confederate), aligning with the expectation that rates might be lower in historically slave-owning regions. The initial 'Jose Mapping' rates represent the anchored male rates observed by race in these proxy cities, with the 'Gender Adjustment' calculated by applying the difference in gender ratios (female rate relative to male rate) observed across the overall B&M sample to these anchors. The final rates illustrate the core finding of B&M that African American names received approximately 50 percent fewer callbacks for interviews compared to White names, with the gender adjustment reflecting the fact that females in that experiment often received slightly higher rates than males within the same racial group.

Draft Response Rate Map of Scenario 3

Block, R., Crabtree, C., Holbein, J. B., & Monson, J. Q. (2021). Are Americans less likely to reply to emails from Black people relative to White people. Proceedings of the National Academy of Sciences.

Geography	Race	Gender	Experiment Mapping	Gender Adjustment	Final Rate
Union	White	Male	4.20	0.00	4.20
Union	White	Female	4.20	TBD	TBD
Union	Black	Male	3.90	0.00	3.90
Union	Black	Female	3.90	TBD	TBD
Confederate	White	Male	1.60	0.00	1.60
Confederate	White	Female	1.60	TBD	TBD
Confederate	Black	Male	1.40	0.00	1.40
Confederate	Black	Female	1.40	TBD	TBD

Observed response rates reported in Block et al. (2021), focus on the differential treatment of putatively White and Black senders. The initial 'Jose Mapping' rates for the Union proxy are anchored to the higher response rates observed among elected officials (4.2% for White senders and 3.9% for Black senders). The Confederate proxy uses the lower response rates observed across the general public sample (1.6% for White senders and 1.4% for Black senders), reflecting the expected lower rates in historically slave-owning regions. This structure directly models the finding that Black senders received fewer responses than White senders, a difference that Block et al. found to be statistically significant. Since the Block et al. study held gender status constant in the initial design summary and did not provide gender-specific rate breakdowns in the excerpts, the 'Gender Adjustment' remains TBD for the female categories.

Assessing Feasibility

1. Responses have massively different scales
2. Gender is unspecified for Scenarios 1 and 3
3. We map Geography from 3 constructs: (Direct from grammar, Boston/Chicago proxy, and Officials/Public proxy)
4. Race effects have inconsistent magnitudes

Final Scenarios, Imputation Strategy, & Assumptions:

Scenario 1: Colin Good Grammar. Rosen, J. (2010). *Legislative responsiveness to constituent ethnicity and grammar quality: A field experiment.*

Geography	Race	Gender	Experiment Mapping	Gender Adjustment*	Final Rate
Union	White	Male	52	0	52
Union	White	Female	52	+1.46*	53.46
Union	Black	Male	37	0	37
Union	Black	Female	37	+0.87*	37.87
Confederate	White	Male	29	0	29
Confederate	White	Female	29	+1.05*	30.05
Confederate	Black	Male	34	0	34
Confederate	Black	Female	34	+0.64*	34.64

Note that this particular scenario shows that Black Males will receive higher responses in Confederate states. *This contradicts our assumptions regarding the distribution of responses*, because we are using the effect sizes observed in the Jose vs. Colin experimental data literally (adjusted only for gender). Since this is merely a counterfactual state of the world, and a science fiction table, we are leaving it at is without manipulation or pre-selection. The plots will directly show measurable comparisons across block cohorts, irrespective of these theoretical response rates.

Scenario 2: Bertrand, M., & Mullainathan, S. (2004). Are Emily and Greg more employable than Lakisha and Jamal?

Geography	Race	Gender	Experiment Mapping	Gender Adjustment	Final Rate
Union	White	Male	11.88	0.00	11.88
Union	White	Female	11.88	+1.46	13.34
Union	Black	Male	7.83	0.00	7.83
Union	Black	Female	7.83	+0.87	8.70
Confederate	White	Male	8.61	0.00	8.61
Confederate	White	Female	8.61	+1.05	9.66
Confederate	Black	Male	5.81	0.00	5.81
Confederate	Black	Female	5.81	+0.64	6.45

Scenario 3: Block, R., Crabtree, C., Holbein, J. B., & Monson, J. Q. (2021)

Geography	Race	Gender	Experiment Mapping	Gender Adjustment*	Final Rate
Union	White	Male	4.20	0	4.20
Union	White	Female	4.20	+1.46*	5.66
Union	Black	Male	3.90	0	3.90
Union	Black	Female	3.90	+0.87*	4.77
Confederate	White	Male	1.60	0	1.60
Confederate	White	Female	1.60	+1.05*	2.65
Confederate	Black	Male	1.40	0	1.40
Confederate	Black	Female	1.40	+0.64*	2.04

Assumptions

1. **Gender imputation method**: Scenarios 1 & 3 took each male rate and multiplied by weighted gender adjustments applying a closer approximation than is currently available (nothing) according to the female advantage in callback rates from scenario 2. The major problem is that is a employer context, and not a customer context, in which you could argue the incentives between hiring men or women is very different when it comes to responding to men or women's catering email requests. Employment discrimination may have different gender dynamics than service positions measured in 2004 - and this is likely confounded easily.
2. **Independence**: Each restaurant responds independently
3. **Homogeneity within cells**: All Union restaurants behave similarly to each other
4. **Fixed response probabilities**: Every restaurant in a cell has the same exact probability of responding
5. **No temporal effects**: Response rates don't change over the study period
6. **Undefined message content**: We haven't modeled the actual catering request content yet
7. **Percentages**: Table data are communicated in percentages

Experiment Proposal Task Decomposition:

- Design: JH/DS
  - Geographic Blocking, done
  - Race treatment, done
  - Gender treatment, done
  - One restaurant receives one treatment, done
- Outcomes: JH/DS
  - Response rate (binary), done
  - Budget acceptance (binary / NA), done
- Sample: JH/DS
  - Restaurants / caterers as subjects, done
  - Regional stratification & balance, done
- Analysis: JH/DS
  - Regression with robust SE, done
  - Discrimination effects, done
  - Regional differences, done
  - Plots, done

```r
# Total sample size
n <- 400

geography <- c("Union", "Confederate")
race <- c("White", "Black")
gender <- c("Male", "Female")
```

```r
# Create a balanced design: 100 observations per combination
design <- expand.grid(
    geography = geography,
    race = race,
    gender = gender
)

design$n <- n / nrow(design) # 50 per cell at n=400

# Assign reply rates for Scenario 1 (Jose/Colin based):
# order: union-white, union-black, confed-white, confed-black
multi_female1and3 <- c(1.46, 0.87, 1.05, 0.64)
add_female2 <- c(1.46, 0.87, 1.05, 0.64)

# Scenario 1
design$reply_rate1 <- c(
  52,
  37,
  29,
  34,
  52 + 1.46,
  37 + 0.87,
  29 + 1.05,
  34 + 0.64
)

# Scenario 2
design$reply_rate2 <- c(
  0.1188 * 100,   # white-union-male
  0.0783 * 100,     # black-union-male
  0.0861 * 100,  # white-confederate-male
  0.0581 * 100,    # black-confederate-male
  0.1188 * 100 + add_female2[1],    # white-union-female
  0.0783 * 100 + add_female2[2],     # black-union-female
  0.0861 * 100 + add_female2[3],     # white-confederate-female
  0.0581 * 100 + add_female2[4]  # black-confederate-female
)

# Scenario 3
design$reply_rate3 <- c(
  4.20,
  3.90,
  1.60,
  1.40,
  4.20 + 1.46,
  3.90 + 0.87,
  1.60 + 1.05,
  1.40 + 0.64
)

design
```

```
##     geography  race gender  n reply_rate1 reply_rate2 reply_rate3
## 1       Union White   Male 50       52.00       11.88        4.20
```

```
## 2 Confederate White   Male 50       37.00        7.83        3.90
## 3       Union Black   Male 50       29.00        8.61        1.60
## 4 Confederate Black   Male 50       34.00        5.81        1.40
## 5       Union White Female 50       53.46       13.34        5.66
## 6 Confederate White Female 50       37.87        8.70        4.77
## 7       Union Black Female 50       30.05        9.66        2.65
## 8 Confederate Black Female 50       34.64        6.45        2.04
```

```
# Generate the data
generate_data <- function(race_level, geo_level, gender_level, n, reply_rate1, reply_rate2, reply_rate3
    response1 <- rbinom(n, size = 1, prob = reply_rate1/100)
    response2 <- rbinom(n, size = 1, prob = reply_rate2/100)
    response3 <- rbinom(n, size = 1, prob = reply_rate3/100)

    # assume 70% baseline, reduced by 15% for Black names
    budget_base_prob <- ifelse(race_level == "White", 0.70, 0.55)

    # budget acceptance conditioned on response=1 (aka receiving a response at all)
    budget_accept1 <- ifelse(response1 == 1, rbinom(n, 1, budget_base_prob), NA)
    budget_accept2 <- ifelse(response2 == 1, rbinom(n, 1, budget_base_prob), NA)
    budget_accept3 <- ifelse(response3 == 1, rbinom(n, 1, budget_base_prob), NA)

    data.frame(
        race = rep(race_level, n),
        geography = rep(geo_level, n),
        gender = rep(gender_level, n),
        received_reply1 = response1,
        received_reply2 = response2,
        received_reply3 = response3,
        budget_accepted1 = budget_accept1,
        budget_accepted2 = budget_accept2,
        budget_accepted3 = budget_accept3
    )
}
```

```
experiment_data <- design %>%
    rowwise() %>%
    do(generate_data(
        .$race,
        .$geography,
        .$gender,
        .$n,
        .$reply_rate1,
        .$reply_rate2,
        .$reply_rate3
    )) %>%
    ungroup()

head(experiment_data)
```

```
## # A tibble: 6 x 9
##   race  geography gender received_reply1 received_reply2 received_reply3
##   <fct> <fct>     <fct>            <int>           <int>           <int>
## 1 White Union     Male                 1               0               0
## 2 White Union     Male                 0               0               0
```

```
## 3 White Union     Male                   1               0               0
## 4 White Union     Male                   0               0               0
## 5 White Union     Male                   0               0               0
## 6 White Union     Male                   1               0               0
## # i 3 more variables: budget_accepted1 <int>, budget_accepted2 <int>,
## #   budget_accepted3 <int>
```

```r
# Reply rates by cell
experiment_data %>%
    group_by(race, geography, gender) %>%
    summarise(
        n = n(),
        scenario1_actual = mean(received_reply1),
        scenario2_actual = mean(received_reply2),
        scenario3_actual = mean(received_reply3),
        .groups = "drop"
    )
```

```
## # A tibble: 8 x 7
##   race  geography   gender     n scenario1_actual scenario2_actual
##   <fct> <fct>       <fct>  <int>            <dbl>            <dbl>
## 1 White Union       Male      50             0.5              0.08
## 2 White Union       Female    50             0.6              0.08
## 3 White Confederate Male      50             0.38             0.12
## 4 White Confederate Female    50             0.32             0.04
## 5 Black Union       Male      50             0.4              0.06
## 6 Black Union       Female    50             0.32             0.14
## 7 Black Confederate Male      50             0.38             0.06
## 8 Black Confederate Female    50             0.38             0.1
## # i 1 more variable: scenario3_actual <dbl>
```

```r
# Regional Balance Check
balance_check <- experiment_data %>%
    group_by(geography, race, gender) %>%
    summarise(n = n(), .groups = "drop")
balance_check
```

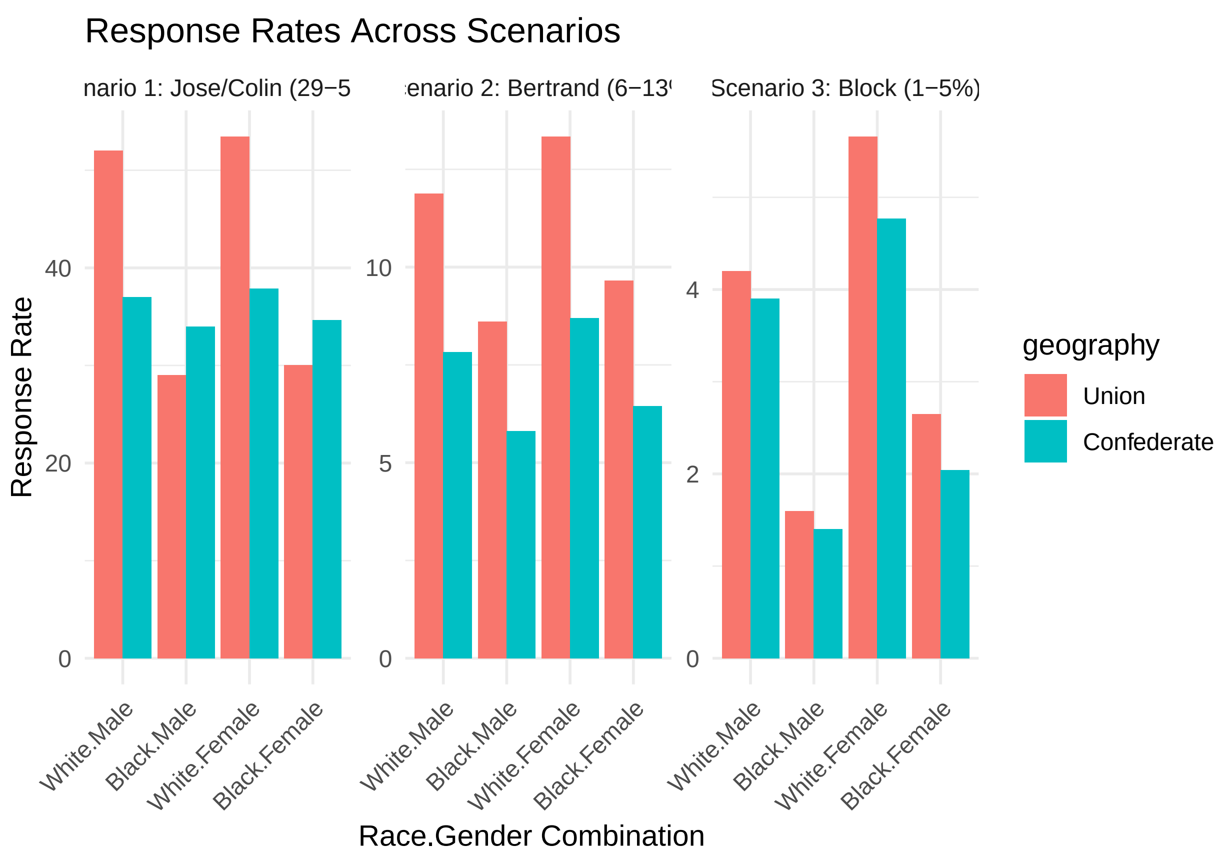
```
## # A tibble: 8 x 4
##   geography   race  gender     n
##   <fct>       <fct> <fct>  <int>
## 1 Union       White Male      50
## 2 Union       White Female    50
## 3 Union       Black Male      50
## 4 Union       Black Female    50
## 5 Confederate White Male      50
## 6 Confederate White Female    50
## 7 Confederate Black Male      50
## 8 Confederate Black Female    50
```

```r
# summary statistics by reply rates
summary_stats <- design %>%
  select(-n) %>%
  mutate(
    expected_rate1 = reply_rate1,
    expected_rate2 = reply_rate2,
    expected_rate3 = reply_rate3
```

```r
)
```

```r
plot_data <- design %>%
  select(race, geography, gender, reply_rate1, reply_rate2, reply_rate3) %>%
  pivot_longer(cols = starts_with("reply_rate"),
               names_to = "scenario",
               names_prefix = "reply_rate",
               values_to = "rate")

ggplot(plot_data, aes(x = interaction(race, gender), y = rate, fill = geography)) +
  geom_bar(stat = "identity", position = "dodge") +
  facet_wrap(~scenario, scales = "free_y",
             labeller = labeller(scenario = c("1" = "Scenario 1: Jose/Colin (29-58%)",
                                              "2" = "Scenario 2: Bertrand (6-13%)",
                                              "3" = "Scenario 3: Block (1-5%)"))) +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1)) +
  labs(title = "Response Rates Across Scenarios",
       x = "Race.Gender Combination",
       y = "Response Rate")
```




```r
# budget acceptance rates
experiment_data %>%
  group_by(race) %>%
  summarise(
    scenario1_budget_accept = mean(budget_accepted1, na.rm = TRUE),
```

```r
    scenario2_budget_accept = mean(budget_accepted2, na.rm = TRUE),
    scenario3_budget_accept = mean(budget_accepted3, na.rm = TRUE),
    .groups = "drop"
  ) %>%
  print(row.names = FALSE)
```

```
## Warning: `...` must be empty in `format.tbl()`
## Caused by error in `format_tbl()`:
## ! `...` must be empty.
## x Problematic argument:
## * row.names = FALSE

## # A tibble: 2 x 4
##   race  scenario1_budget_accept scenario2_budget_accept scenario3_budget_accept
##   <fct>                   <dbl>                   <dbl>                   <dbl>
## 1 White                   0.744                   0.812                   0.75
## 2 Black                   0.568                   0.667                   0.667
```

```r
# simplify conditioning with indicator variable option
experiment_data <- experiment_data %>%
  mutate(
    R = ifelse(race == "White", 0, 1), # race indicator
    G = ifelse(geography == "Union", 0, 1), # geography indicator
    S = ifelse(gender == "Male", 0, 1)# sex/gender indicator
  )
```

```r
head(experiment_data)
```

```
## # A tibble: 6 x 12
##   race  geography gender received_reply1 received_reply2 received_reply3
##   <fct> <fct>     <fct>            <int>           <int>           <int>
## 1 White Union     Male                 1               0               0
## 2 White Union     Male                 0               0               0
## 3 White Union     Male                 1               0               0
## 4 White Union     Male                 0               0               0
## 5 White Union     Male                 0               0               0
## 6 White Union     Male                 1               0               0
## # i 6 more variables: budget_accepted1 <int>, budget_accepted2 <int>,
## #   budget_accepted3 <int>, R <dbl>, G <dbl>, S <dbl>
```

```r
ftable(experiment_data$geography, experiment_data$race, experiment_data$gender)
```

```
##                    Male Female
##
## Union       White    50     50
##             Black    50     50
## Confederate White    50     50
##             Black    50     50
```

```r
# test regional differences in discrimination
test_regional_differences <- function(data, scenario_col) {
    # subset by geography
    union_data <- filter(data, geography == "Union")
    confed_data <- filter(data, geography == "Confederate")

    # calculate discrimination effect in each region
    union_model <- lm(as.formula(paste(scenario_col, "~ race")), data = union_data)
```

```r
    union_race_effect <- coef(union_model)["raceBlack"]

    confed_model <- lm(as.formula(paste(scenario_col, "~ race")), data = confed_data)
    confed_race_effect <- coef(confed_model)["raceBlack"]

    # difference in discrimination between regions
    regional_diff <- confed_race_effect - union_race_effect

    return(list(
        union_effect = union_race_effect,
        confederate_effect = confed_race_effect,
        difference = regional_diff
    ))
}

# test regional differences for each scenario
regional1 <- test_regional_differences(experiment_data, "received_reply1")
regional2 <- test_regional_differences(experiment_data, "received_reply2")
regional3 <- test_regional_differences(experiment_data, "received_reply3")

regional_results <- data.frame(
  Scenario = c("Scenario 1", "Scenario 2", "Scenario 3"),
  Union_Effect = c(regional1$union_effect, regional2$union_effect, regional3$union_effect),
  Confederate_Effect = c(regional1$confederate_effect, regional2$confederate_effect, regional3$confedera
  Difference = c(regional1$difference, regional2$difference, regional3$difference)
)

regional_results
```

```
##     Scenario  Union_Effect Confederate_Effect Difference
## 1 Scenario 1 -1.900000e-01       3.000000e-02       0.22
## 2 Scenario 2  2.000000e-02       8.635509e-16      -0.02
## 3 Scenario 3 -1.962616e-17      -2.000000e-02      -0.02
```

```r
test_discrimination_effects <- function(data, scenario_col) {
    formula <- as.formula(paste(scenario_col, "~ race * geography * gender"))
    model <- lm(formula, data = data)
    robust_se <- vcovHC(model, type = "HC0")
    coef_test <- coeftest(model, vcov = robust_se)

    # extract all discrimination-relevant effects
    effects <- list(
        race_main = coef_test["raceBlack", ],
        geography_main = coef_test["geographyConfederate", ],
        gender_main = coef_test["genderFemale", ],
        race_x_geography = coef_test["raceBlack:geographyConfederate", ],
        race_x_gender = coef_test["raceBlack:genderFemale", ],
        geography_x_gender = coef_test["geographyConfederate:genderFemale", ],
        three_way = coef_test["raceBlack:geographyConfederate:genderFemale", ]
    )

    return(effects)
}
```

```r
# test all three scenarios
disc_effects_s1 <- test_discrimination_effects(experiment_data, "received_reply1")
disc_effects_s2 <- test_discrimination_effects(experiment_data, "received_reply2")
disc_effects_s3 <- test_discrimination_effects(experiment_data, "received_reply3")

# display results
# Scenario 1 Discrimination Effect
disc_effects_s1$race_main
```

```
##    Estimate  Std. Error     t value    Pr(>|t|)
## -0.10000000  0.09899495 -1.01015254  0.31304532
```

```r
disc_effects_s1$race_x_geography
```

```
##   Estimate Std. Error    t value   Pr(>|t|)
##  0.1000000  0.1386506  0.7212372  0.4711935
```

```r
# function to run single simulation and test race effect
run_single_simulation <- function(n, scenario_num) {
  # calculate observations per cell
  n_per_cell <- n %/% 8
  remainder <- n %% 8
  cell_sizes <- c(rep(n_per_cell + 1, remainder), rep(n_per_cell, 8 - remainder))
  expanded_design <- design[rep(1:8, times = cell_sizes), ]

  if (scenario_num == 1) {
    response_probs <- expanded_design$reply_rate1
  } else if (scenario_num == 2) {
    response_probs <- expanded_design$reply_rate2
  } else if (scenario_num == 3) {
    response_probs <- expanded_design$reply_rate3
  }

  sim_data <- data.frame(
    race = expanded_design$race,
    geography = expanded_design$geography,
    gender = expanded_design$gender,
    response = rbinom(nrow(expanded_design), size = 1, prob = response_probs/100)
  )

  # fit linear model: response ~ race + geography + gender
  model <- lm(response ~ race + gender + geography + race:geography, data = sim_data)

  # compute robust se
  robust_se <- vcovHC(model, type = "HC0")

  # extract p-value for race coefficient
  coef_test <- coeftest(model, vcov = robust_se)
  p_value <- coef_test["raceBlack:geographyConfederate", "Pr(>|t|)"]

  return(p_value)
}
```

```r
# calculate power through repeated simulations
calculate_power <- function(n, scenario_num, n_sims = 1000) {
```

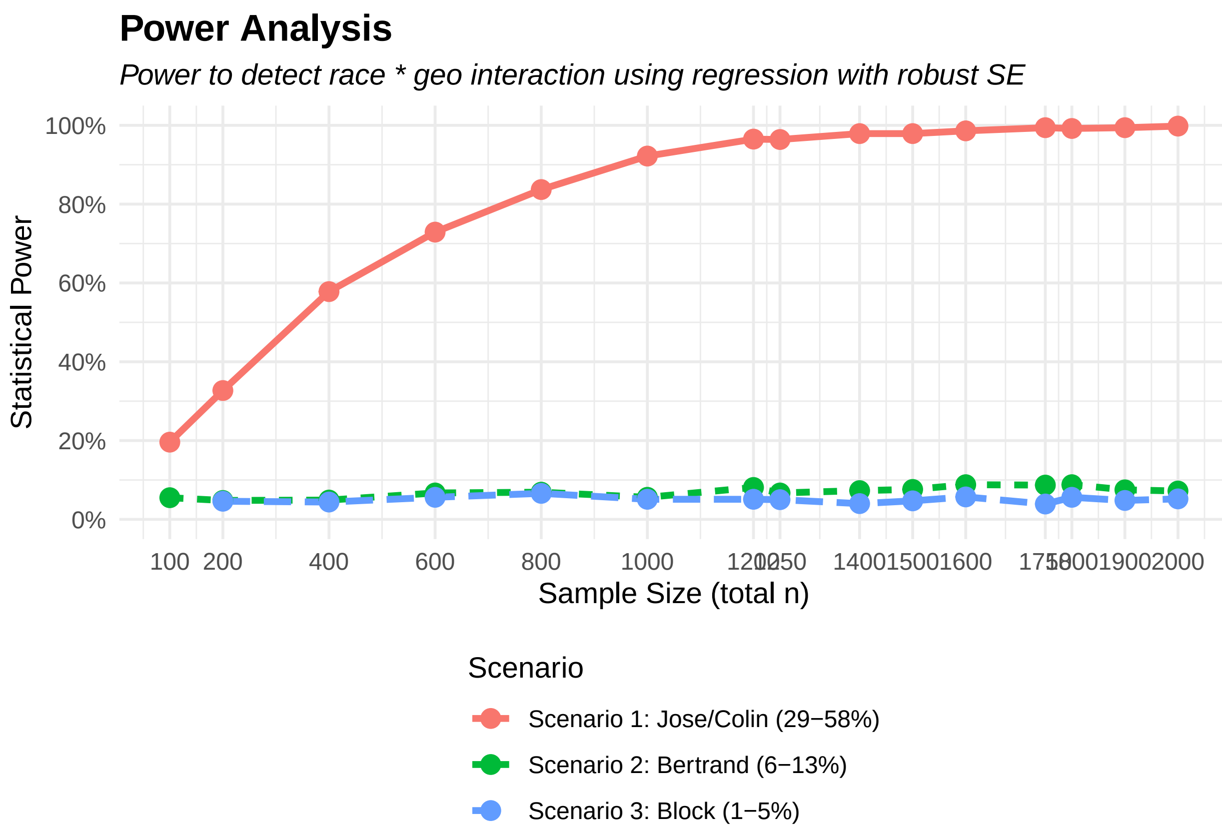
```r
  # run n_sims simulations and collect p-values
  p_values <- replicate(n_sims, run_single_simulation(n, scenario_num))
  # calculate proportion of simulations with p < 0.05
  power <- mean(p_values < 0.05)
  return(power)
}
```

```r
generate_power_curves <- function() {
  sample_sizes <- c(100, 200, 400, 600, 800, 1000, 1200, 1250, 1400, 1500, 1600, 1750, 1800, 1900, 2000
  power_results <- data.frame(
    scenario = integer(),
    sample_size = integer(),
    power = numeric())
  for (scenario in 1:3) {
    for (n in sample_sizes) {
      # calculate power for this combination
      power <- calculate_power(n, scenario, n_sims = 1000)
      power_results <- rbind(power_results,
                             data.frame(scenario = scenario,
                                        sample_size = n,
                                        power = power))}}
  power_results$scenario_label <- factor(power_results$scenario,
                                        levels = 1:3,
                                        labels = c("Scenario 1: Jose/Colin (29-58%)",
                                                   "Scenario 2: Bertrand (6-13%)",
                                                   "Scenario 3: Block (1-5%)"))
  return(power_results)}
```

```r
plot_power_curves <- function(power_results) {
  p <- ggplot(power_results, aes(x = sample_size, y = power,
                                 color = scenario_label,
                                 linetype = scenario_label)) +
    geom_line(linewidth = 1.2) +
    geom_point(size = 3) +
    labs(title = "Power Analysis",
         subtitle = "Power to detect race * geo interaction using regression with robust SE",
         x = "Sample Size (total n)",
         y = "Statistical Power",
         color = "Scenario",
         linetype = "Scenario") +
    scale_y_continuous(limits = c(0, 1),
                       breaks = seq(0, 1, by = 0.2),
                       labels = scales::percent) +
    scale_x_continuous(breaks = c(100, 200, 400, 600, 800, 1000, 1200, 1250, 1400, 1500, 1600, 1750, 18
    theme_minimal() +
    theme(legend.position = "bottom",
          legend.direction = "vertical",
          plot.title = element_text(size = 14, face = "bold"),
          plot.subtitle = element_text(size = 11, face = "italic"))
  p <- p + annotate("text", x = 900, y = 78,
                    label = "80% power",
                    size = 3,
                    color = "red")
  return(p)
```

```
}
```

```
power_results <- generate_power_curves()
power_plot <- plot_power_curves(power_results)
print(power_plot)
```




```
write.csv(power_results, "power_analysis_results.csv", row.names = FALSE)
```

```
power_results
```

```
##    scenario sample_size power                   scenario_label
## 1         1         100 0.196 Scenario 1: Jose/Colin (29-58%)
## 2         1         200 0.327 Scenario 1: Jose/Colin (29-58%)
## 3         1         400 0.578 Scenario 1: Jose/Colin (29-58%)
## 4         1         600 0.729 Scenario 1: Jose/Colin (29-58%)
## 5         1         800 0.837 Scenario 1: Jose/Colin (29-58%)
## 6         1        1000 0.922 Scenario 1: Jose/Colin (29-58%)
## 7         1        1200 0.965 Scenario 1: Jose/Colin (29-58%)
## 8         1        1250 0.964 Scenario 1: Jose/Colin (29-58%)
## 9         1        1400 0.979 Scenario 1: Jose/Colin (29-58%)
## 10        1        1500 0.979 Scenario 1: Jose/Colin (29-58%)
## 11        1        1600 0.986 Scenario 1: Jose/Colin (29-58%)
## 12        1        1750 0.994 Scenario 1: Jose/Colin (29-58%)
## 13        1        1800 0.992 Scenario 1: Jose/Colin (29-58%)
## 14        1        1900 0.994 Scenario 1: Jose/Colin (29-58%)
## 15        1        2000 0.998 Scenario 1: Jose/Colin (29-58%)
## 16        2         100 0.055    Scenario 2: Bertrand (6-13%)
```

```
## 17        2         200 0.048    Scenario 2: Bertrand (6-13%)
## 18        2         400 0.049    Scenario 2: Bertrand (6-13%)
## 19        2         600 0.067    Scenario 2: Bertrand (6-13%)
## 20        2         800 0.069    Scenario 2: Bertrand (6-13%)
## 21        2        1000 0.056    Scenario 2: Bertrand (6-13%)
## 22        2        1200 0.081    Scenario 2: Bertrand (6-13%)
## 23        2        1250 0.067    Scenario 2: Bertrand (6-13%)
## 24        2        1400 0.073    Scenario 2: Bertrand (6-13%)
## 25        2        1500 0.076    Scenario 2: Bertrand (6-13%)
## 26        2        1600 0.088    Scenario 2: Bertrand (6-13%)
## 27        2        1750 0.087    Scenario 2: Bertrand (6-13%)
## 28        2        1800 0.088    Scenario 2: Bertrand (6-13%)
## 29        2        1900 0.075    Scenario 2: Bertrand (6-13%)
## 30        2        2000 0.072    Scenario 2: Bertrand (6-13%)
## 31        3         100    NA        Scenario 3: Block (1-5%)
## 32        3         200 0.046        Scenario 3: Block (1-5%)
## 33        3         400 0.044        Scenario 3: Block (1-5%)
## 34        3         600 0.056        Scenario 3: Block (1-5%)
## 35        3         800 0.066        Scenario 3: Block (1-5%)
## 36        3        1000 0.051        Scenario 3: Block (1-5%)
## 37        3        1200 0.051        Scenario 3: Block (1-5%)
## 38        3        1250 0.050        Scenario 3: Block (1-5%)
## 39        3        1400 0.040        Scenario 3: Block (1-5%)
## 40        3        1500 0.047        Scenario 3: Block (1-5%)
## 41        3        1600 0.057        Scenario 3: Block (1-5%)
## 42        3        1750 0.039        Scenario 3: Block (1-5%)
## 43        3        1800 0.056        Scenario 3: Block (1-5%)
## 44        3        1900 0.048        Scenario 3: Block (1-5%)
## 45        3        2000 0.052        Scenario 3: Block (1-5%)
```

```r
create_discrimination_plot <- function() {
  # Calculate effect sizes for each scenario
  effect_data <- data.frame(
    scenario = rep(c("Scenario 1", "Scenario 2", "Scenario 3"), each = 4),
    geography = rep(rep(c("Union", "Confederate"), each = 2), 3),
    gender = rep(c("Male", "Female", "Male", "Female"), 3),
    white_rate = c(
      52, 53.46, 29, 30.05,   # Scenario 1
      11.88, 13.34, 8.61, 9.66,   # Scenario 2
      4.20, 5.66, 1.60, 2.65   # Scenario 3
    ),
    black_rate = c(
      37, 37.87, 34, 34.64,   # Scenario 1
      7.83, 8.70, 5.81, 6.45,   # Scenario 2
      3.90, 4.77, 1.40, 2.04   # Scenario 3
    ))

  effect_data$discrimination <- effect_data$white_rate - effect_data$black_rate
  ggplot(effect_data, aes(x = interaction(geography, gender), y = discrimination, fill = geography)) +
    geom_bar(stat = "identity", position = "dodge") +
    geom_hline(yintercept = 0, linetype = "dashed", alpha = 0.5) +
    facet_wrap(~scenario, scales = "free_y", nrow = 1) +
    labs(title = "Discrimination effect sizes across experimental cells",
         subtitle = "Observable tau percentage point difference : White - Black response rates",
```

```r
         x = "Geography × Gender",
         y = "Discrimination Effect (pp)") +
    scale_fill_manual(values = c("Union" = "#2166ac", "Confederate" = "#b2182b")) +
    theme_minimal() +
    theme(axis.text.x = element_text(angle = 45, hjust = 1),
          legend.position = "top",
          plot.title = element_text(size = 16, face = "bold"))}
```

```r
generate_summary_report <- function(power_results) {
  # Sample sizes for 80% power
  for(s in unique(power_results$scenario_label)) {
    scenario_data <- filter(power_results, scenario_label == s)
    power_80_n <- scenario_data$sample_size[which(scenario_data$power >= 0.80)[1]]
    if(is.na(power_80_n)) power_80_n <- ">2000"
  }
}
```

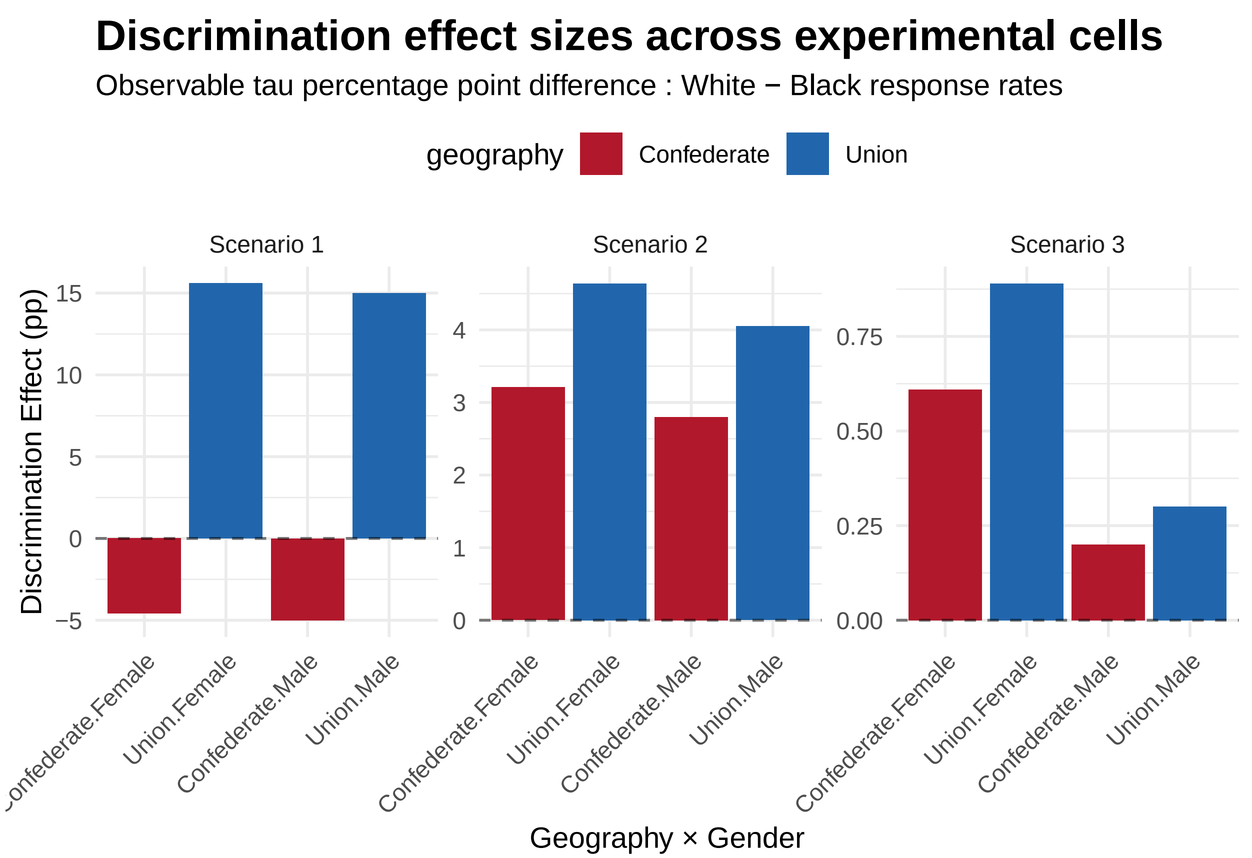
```r
print_scenario_tables <- function() {
  # Combine all rates and their labels into one data frame
  all_rates <- data.frame(
    Scenario = rep(c("Scenario 1", "Scenario 2", "Scenario 3"), each = 8),
    Geography = rep(c("Union", "Union", "Confederate", "Confederate"), times = 6),
    Race = rep(c("White", "Black"), each = 4, times = 3),
    Gender = rep(c("Male", "Female"), times = 12),
    Rate = c(
      # Scenario 1
      design$reply_rate1,
      # Scenario 2
      design$reply_rate2,
      # Scenario 3
      design$reply_rate3
    )
  )
  # Print a simple, readable table (sorted for clarity)
  print(
    all_rates[order(all_rates$Scenario, all_rates$Geography, all_rates$Race, all_rates$Gender), ],
    row.names = FALSE
  )
}

print_scenario_tables()
```

```
##    Scenario   Geography  Race Gender  Rate
##  Scenario 1 Confederate Black Female 34.64
##  Scenario 1 Confederate Black   Male 30.05
##  Scenario 1 Confederate White Female 34.00
##  Scenario 1 Confederate White   Male 29.00
##  Scenario 1       Union Black Female 37.87
##  Scenario 1       Union Black   Male 53.46
##  Scenario 1       Union White Female 37.00
##  Scenario 1       Union White   Male 52.00
##  Scenario 2 Confederate Black Female  6.45
##  Scenario 2 Confederate Black   Male  9.66
##  Scenario 2 Confederate White Female  5.81
```

```
## Scenario 2 Confederate White   Male  8.61
## Scenario 2       Union Black Female  8.70
## Scenario 2       Union Black   Male 13.34
## Scenario 2       Union White Female  7.83
## Scenario 2       Union White   Male 11.88
## Scenario 3 Confederate Black Female  2.04
## Scenario 3 Confederate Black   Male  2.65
## Scenario 3 Confederate White Female  1.40
## Scenario 3 Confederate White   Male  1.60
## Scenario 3       Union Black Female  4.77
## Scenario 3       Union Black   Male  5.66
## Scenario 3       Union White Female  3.90
## Scenario 3       Union White   Male  4.20
```

```
create_discrimination_plot()
```




```
generate_summary_report(power_results)
```

```
experiment_data %>%
  group_by(geography, race, gender) %>%
  summarise(
    BudgetAccepted_S1 = mean(budget_accepted1, na.rm = TRUE),
    BudgetAccepted_S2 = mean(budget_accepted2, na.rm = TRUE),
    BudgetAccepted_S3 = mean(budget_accepted3, na.rm = TRUE),
    n_S1 = sum(!is.na(budget_accepted1)),
    n_S2 = sum(!is.na(budget_accepted2)),
    n_S3 = sum(!is.na(budget_accepted3)),
```

```r
    .groups = "drop"
  ) %>%
  print(row.names = FALSE)
```

```
## Warning: `...` must be empty in `format.tbl()`
## Caused by error in `format_tbl()`:
## ! `...` must be empty.
## x Problematic argument:
## * row.names = FALSE

## # A tibble: 8 x 9
##   geography   race  gender BudgetAccepted_S1 BudgetAccepted_S2 BudgetAccepted_S3
##   <fct>       <fct> <fct>              <dbl>             <dbl>             <dbl>
## 1 Union       White Male               0.72              0.75              0.5
## 2 Union       White Female             0.733             1                 0.5
## 3 Union       Black Male               0.65              0.667             0.667
## 4 Union       Black Female             0.625             0.857             0
## 5 Confederate White Male               0.789             0.833             1
## 6 Confederate White Female             0.75              0.5               1
## 7 Confederate Black Male               0.368             0.667             1
## 8 Confederate Black Female             0.632             0.4               1
## # i 3 more variables: n_S1 <int>, n_S2 <int>, n_S3 <int>
```

Final Notes

- Race and geography are measurably able to affect reply probability, and the magnitude varies by the scenario
- Gender effects are present but secondary to race
- We were able to balance region throughout our designs and blocks
- Budget was modeled at 70% for white and 55% for black names, an arbitrary baseline
- Power analysis shows ample statistical power to detect discrimination in high-rate scenarios, but it's marginal when the mean differences are subtle and reply rates are low.
- Our key takeaway is that under scenario 1, which is most similar to the Colin and Jose experiment where we see large differences in assumed response rates, we reach 80% power at ~800 participants. So this will be the number we will strive for and search for a programmatic way to reach out to these restaurants for catering quotes.