Assignment_Score_Project

2024-09-27

```
# Load packages
library(ggplot2)
library(tinytex)
```

```
# Reading CSV files for each campaign
campaign_A <- read.csv("~/Desktop/Campaign IDs/19139.csv", header = TRUE)
campaign_B <- read.csv("~/Desktop/Campaign IDs/19140.csv", header = TRUE)
campaign_C <- read.csv("~/Desktop/Campaign IDs/19141.csv", header = TRUE)
campaign_D <- read.csv("~/Desktop/Campaign IDs/19142.csv", header = TRUE)

# Combining all campaign data into one dataframe
campaigns <- rbind(campaign_A, campaign_B, campaign_C, campaign_D)
```

```
# Cleaning data to make sure theres a dwid number and assignment scores from 1-100
campaigns_dwid <- campaigns[campaigns$assignment_score >= 1 &
                            campaigns$assignment_score <= 100 & !is.na(campaigns$dwid), ]
```

```
# Creating buckets from 1-100 and cutting by 20
campaigns_dwid$score_bucket <- cut(campaigns_dwid$assignment_score, breaks =
                                     seq(1, 101, by = 20))
```

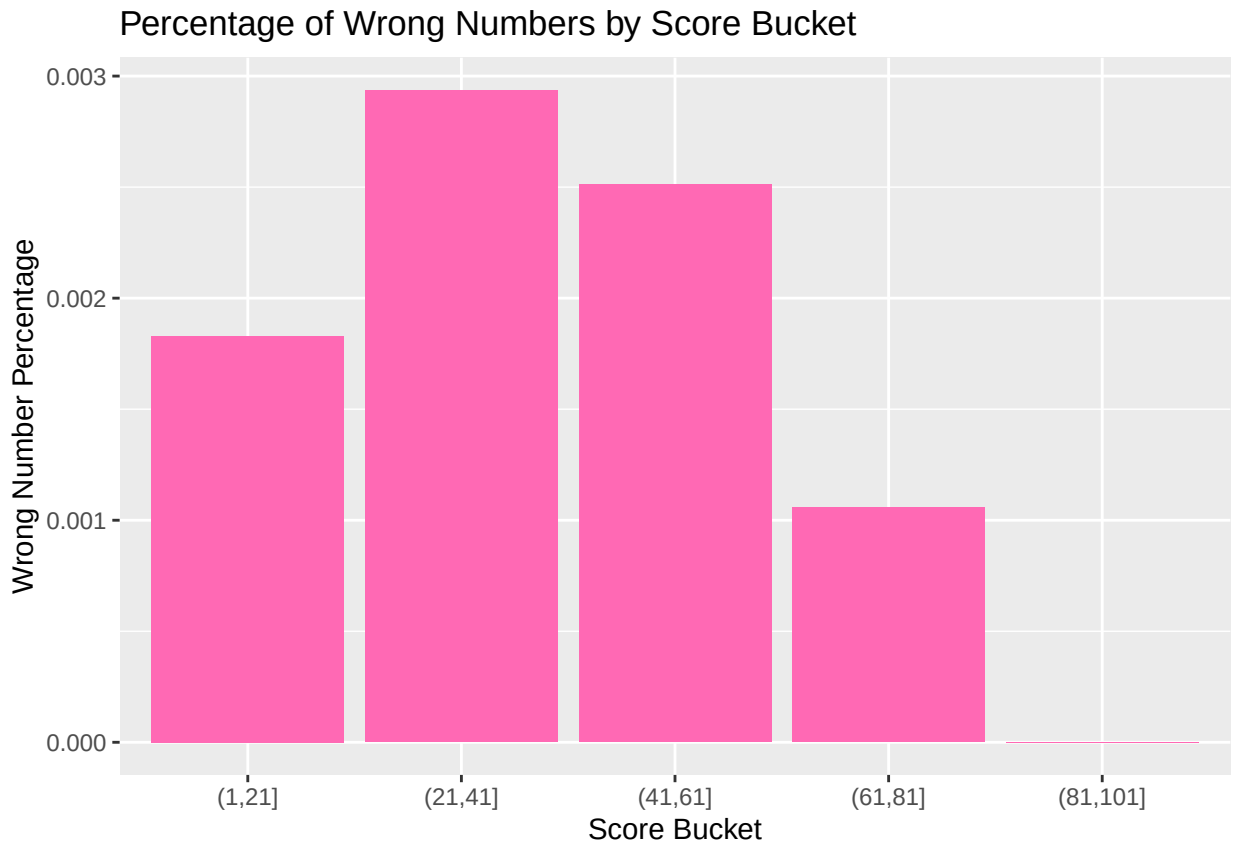
```
# Calculating percentage of wrong numbers
wrong_numbers <- aggregate(campaigns_dwid$tag_title == "Wrong Number",
                           by = list(campaigns_dwid$score_bucket), mean)
colnames(wrong_numbers) <- c("score_bucket", "wrong_number_pct")
```

```
# Calculating assignment score
avg_scores <- aggregate(campaigns_dwid$assignment_score,
                        by = list(campaigns_dwid$score_bucket), mean)
colnames(avg_scores) <- c("score_bucket", "avg_assignment_score")
```

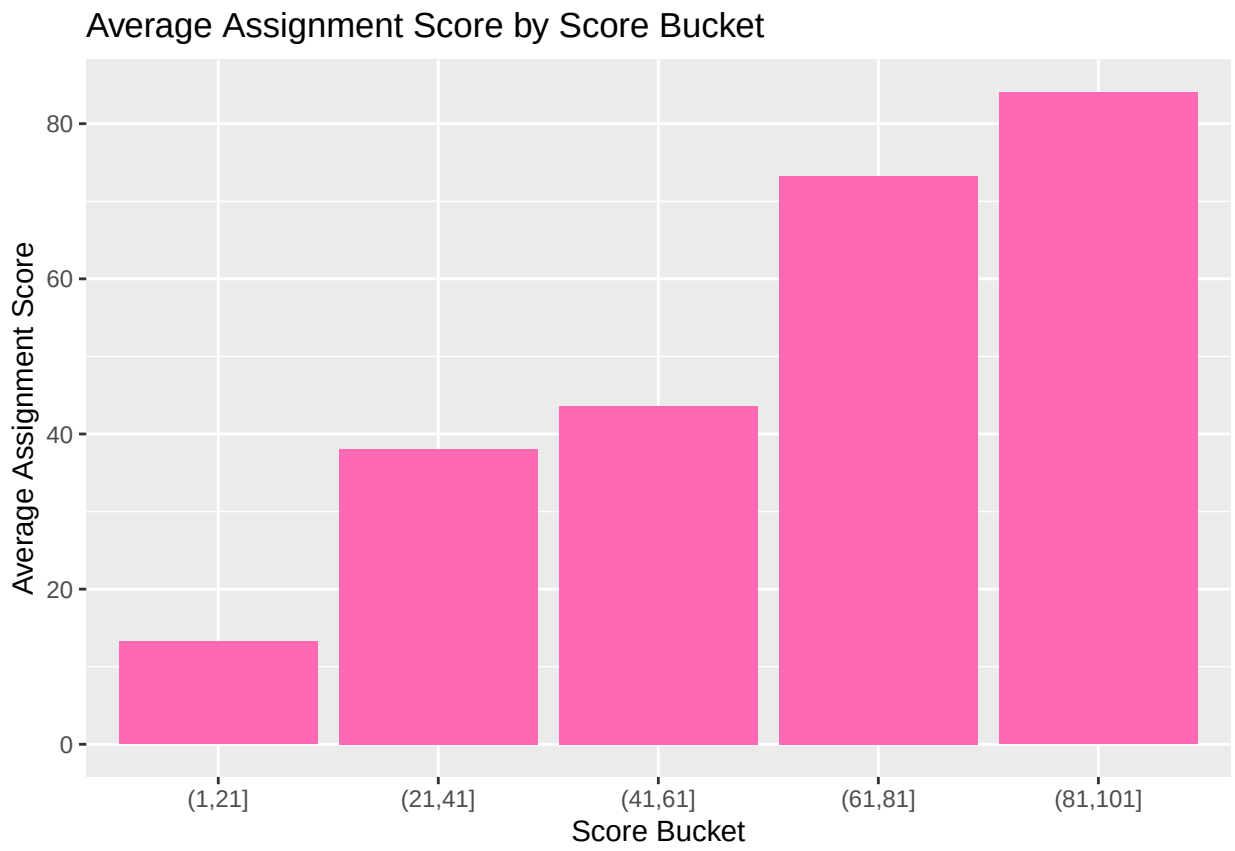
```
# Calculating percentage of wrong numbers by bucket and assignment score
results <- merge(wrong_numbers, avg_scores, by = "score_bucket")
```

Bar Charts

```
# Bar chart for Score Bucket and Wrong Numbers
ggplot(results, aes(x = score_bucket, y = wrong_number_pct)) +
  geom_bar(stat = "identity", fill = "hotpink") +
  xlab("Score Bucket") + ylab("Wrong Number Percentage") +
  ggtitle("Percentage of Wrong Numbers by Score Bucket")
```

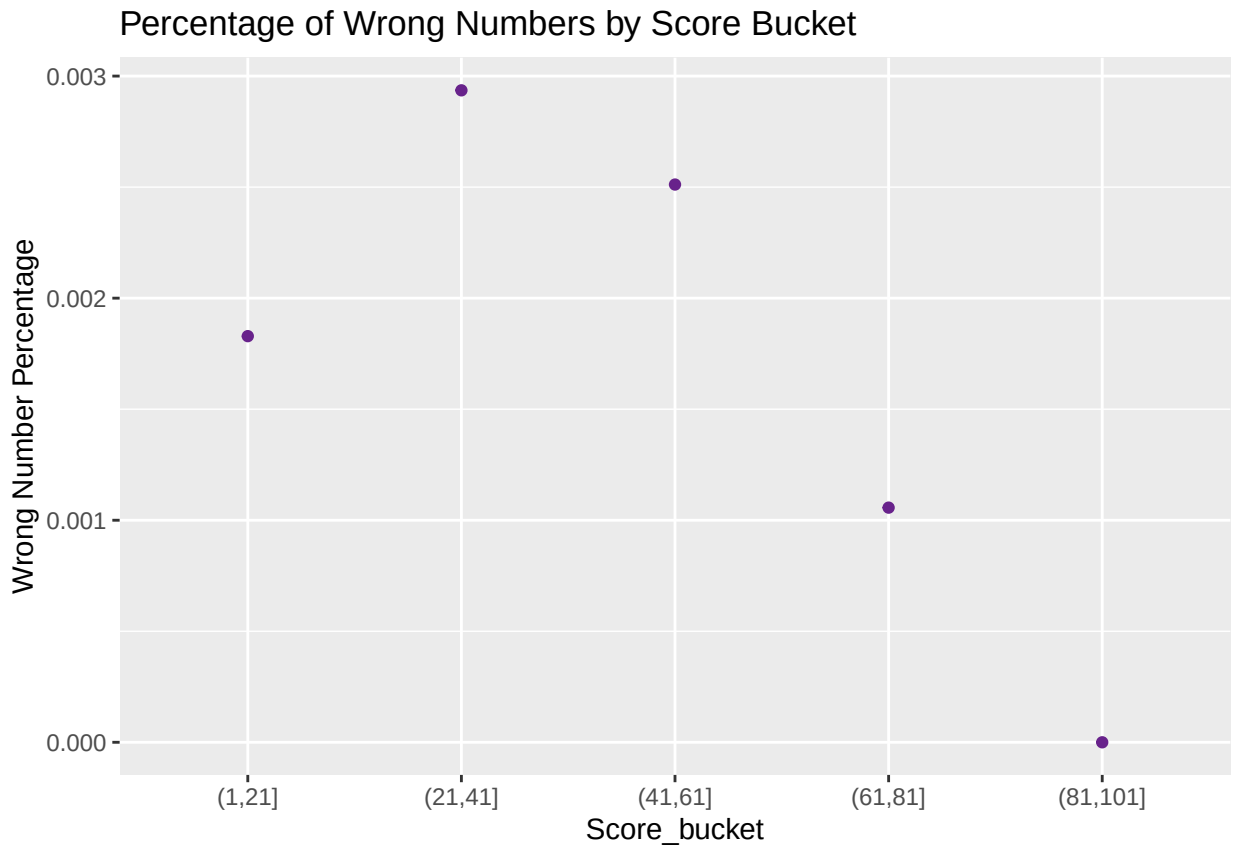



```
# Bar chart for Score Bucket and Average Assignment Score
ggplot(results, aes(x = score_bucket, y = avg_assignment_score)) +
  geom_bar(stat = "identity", fill = "hotpink") +
  xlab("Score Bucket") + ylab("Average Assignment Score") +
  ggtitle("Average Assignment Score by Score Bucket")
```

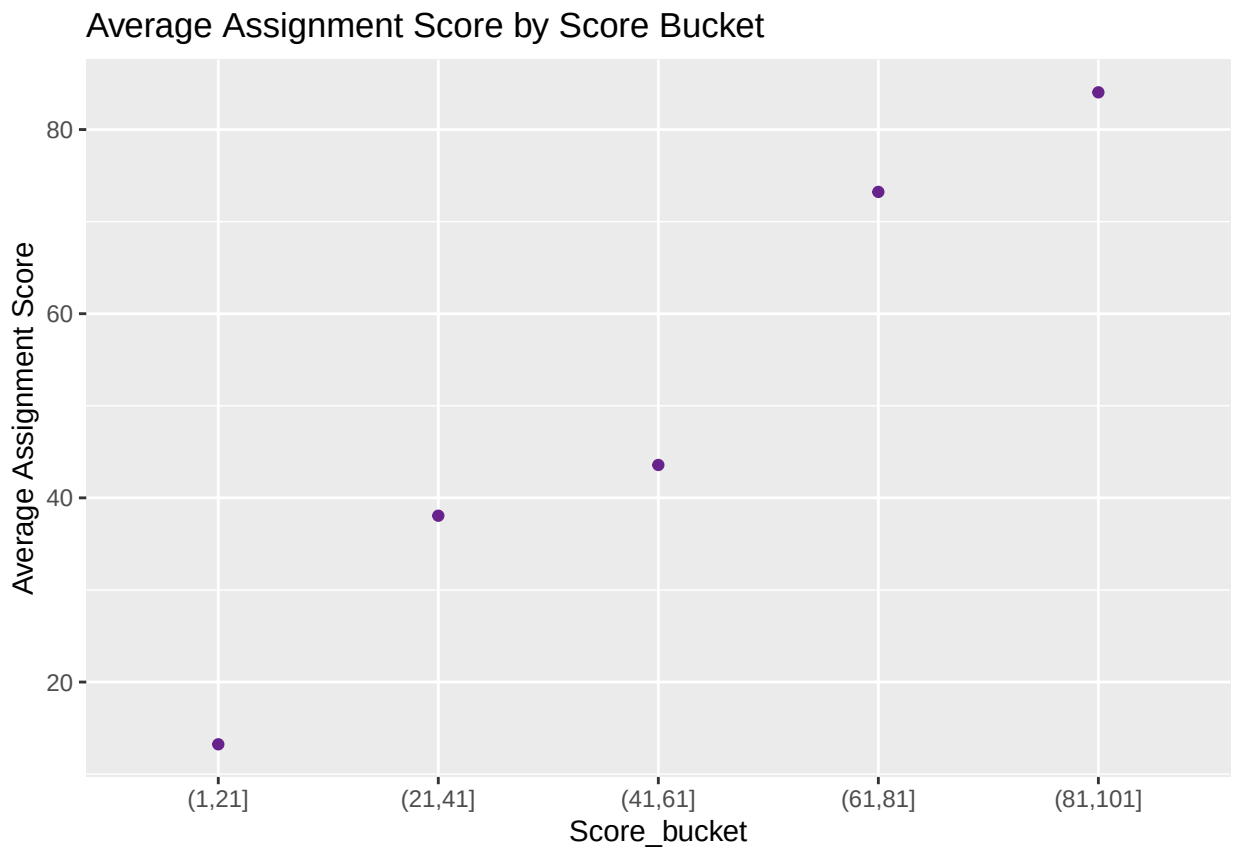


Scatter Plots

```
# Scatter plot for Score Bucket and Wrong Numbers
ggplot(results, aes(x = score_bucket, y = wrong_number_pct)) +
  geom_point(color = "darkorchid4") +
  xlab("Score_bucket") + ylab("Wrong Number Percentage") +
  ggtitle("Percentage of Wrong Numbers by Score Bucket")
```

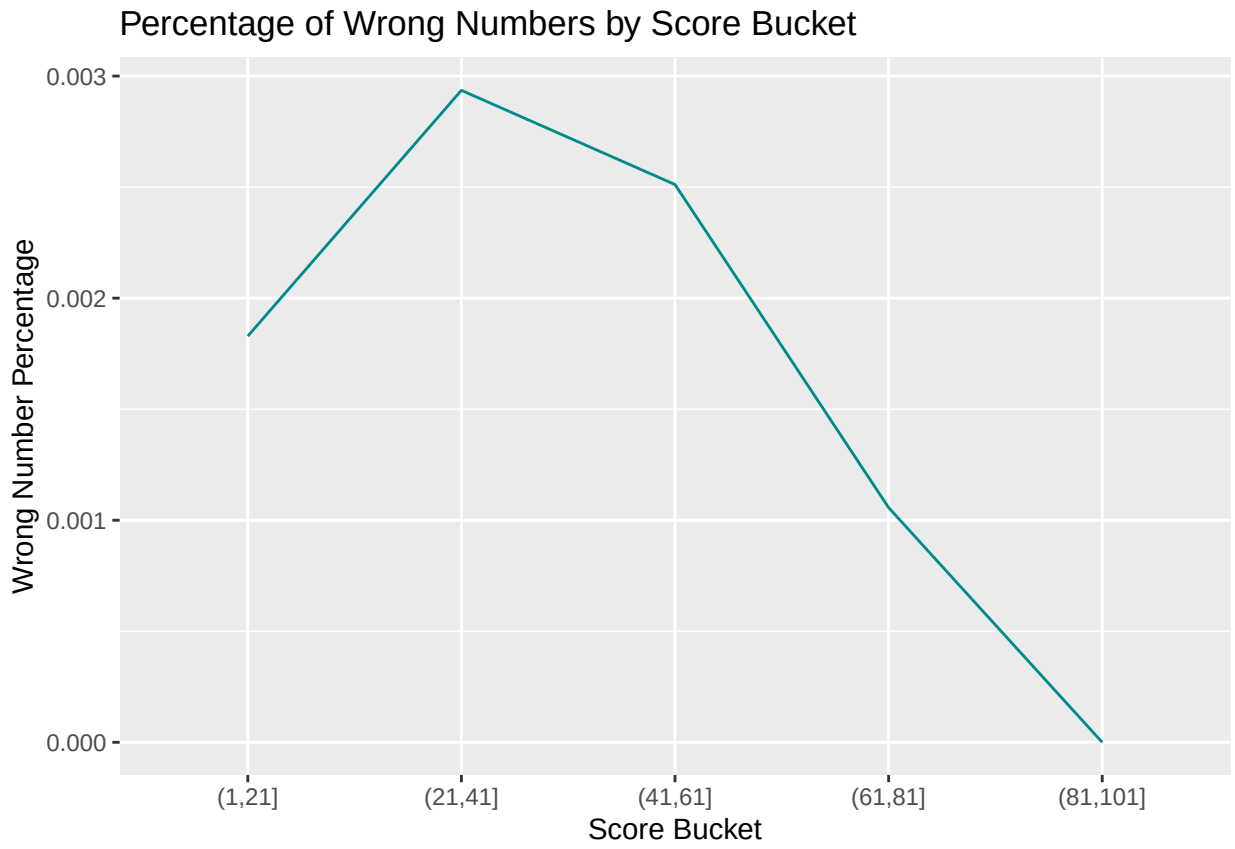



```
# Scatter plot for Score Bucket and Average Assignment Score
ggplot(results, aes(x = score_bucket, y = avg_assignment_score)) +
  geom_point(color = "darkorchid4") +
  xlab("Score_bucket") + ylab("Average Assignment Score") +
  ggtitle("Average Assignment Score by Score Bucket")
```

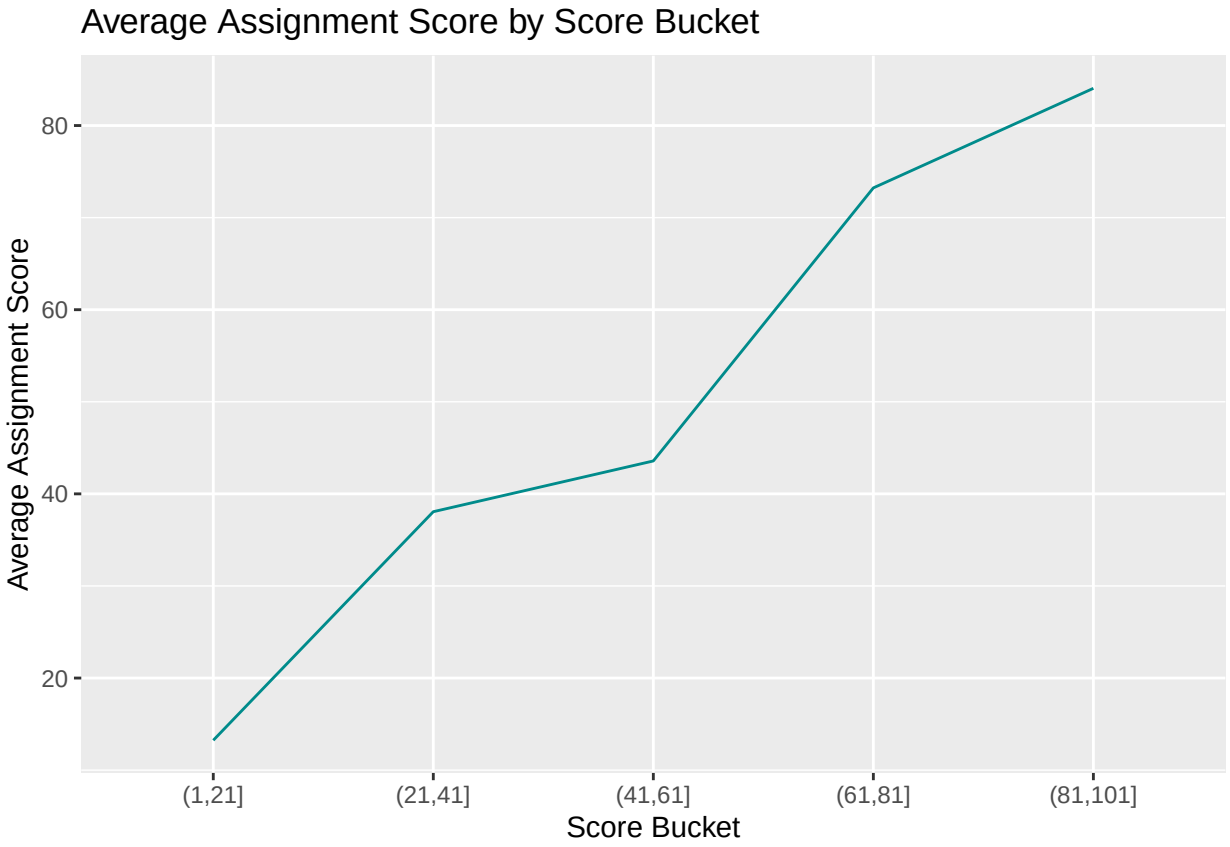


Line Graphs

```
# Line graph for Score Bucket and Wrong Numbers
ggplot(results, aes(x = score_bucket, y = wrong_number_pct, group = 1)) +
  geom_line(color = "darkcyan") +
  xlab("Score Bucket") + ylab("Wrong Number Percentage") +
  ggtitle("Percentage of Wrong Numbers by Score Bucket")
```




```
# Line graph for Score Bucket and Average Assignment Score
ggplot(results, aes(x = score_bucket, y = avg_assignment_score, group = 1)) +
  geom_line(color = "darkcyan") +
  xlab("Score Bucket") + ylab("Average Assignment Score") +
  ggtitle("Average Assignment Score by Score Bucket")
```



Analysis

- Wrong numbers go up and are highest in score buckets 21-40 then taper down.
- As the average assignment score increases after 40, the number of wrong percentages decreases.
- Higher score buckets & assignment scores may mean that these contacts are more accurate and up to date.
- Lower score buckets correspond to higher percentage of wrong numbers.
- These contacts in lower score buckets may not as accurate and not up to date.